Research Article

Daniel Li*, Hervé Queffélec, and Luis Rodríguez-Piazza

Approximation and entropy numbers of composition operators

https://doi.org/10.1515/conop-2020-0106
Received October 29, 2019; accepted September 15, 2020

Abstract: We give a survey on approximation numbers of composition operators on the Hardy space, on the disk and on the polydisk, and add corresponding new results on their entropy numbers, revealing how they are different.

Keywords: approximation numbers; composition operator; entropy numbers; polydisk; pluricapacity

MSC: primary: 47B33; secondary: 30H10; 32A35; 46B06

1 Introduction

This paper surveys results on approximation numbers of composition operators on the Hardy space, and gives new results, on their entropy numbers, in one or several dimensions. In various papers (see [3, 19–22]), pretty sharp estimates are obtained for them, either for classes of examples like the lens maps or the cusp maps, or in the general case. In particular, a few properties are investigated, related with the so-called “spectral radius type formula”, obtained, in dimension one through a result of Widom in [21], and, partially in dimension $N \geq 2$ [22, 23], through a result of Nivoche [26] and Zakharyuta [33]. One of our main results (quoted in dimension one) was the implication (using Green capacity considerations):

$$\|\varphi\|_\infty = 1 \implies \lim_{n \to \infty} [a_n(C \varphi)]^{1/n} = 1,$$

where $a_n(C \varphi)$ is the $n$-th approximation number of $C \varphi$. Note that it is straightforward that if $\|\varphi\|_\infty = r < 1$, then $\lim_{n \to \infty} [a_n(C \varphi)]^{1/n} \leq r < 1$.

Another way of measuring the compactness of operators is using the entropy numbers instead of the approximation numbers. Those numbers stand a little apart in the jungle of “$s$-numbers”, even though they seem to be the most natural for the study of compactness, since their membership in $c_0$ characterizes compactness, even in the general framework of arbitrary Banach spaces.

Given a compact operator $T: H_1 \to H_2$ between Hilbert spaces, the relation between its entropy numbers $e_n(T)$ and its approximation (if one prefers singular) numbers $a_n(T)$ is theoretically known, through a general result on diagonal operators on $l^2$, recalled in Theorem 2.2 to follow, and through the Schmidt decomposition of $T$. This comparison can be thought useless, since in principle we do not know better the numbers $a_n(T)$ than the numbers $e_n(T)$. But in our case, with $T = C \varphi$, namely $T(f) = f \circ \varphi$ where $\varphi$ is an analytic self-map of the polydisk $\mathbb{D}^N$, the situation is slightly different. Answering a question of J. Wengenroth [30] about the
behavior of the entropy numbers of composition operators, we give in this paper estimates for these numbers, analog to that on the approximation numbers. In particular, we have:

\[ \| \varphi \|_\infty = 1 \implies \lim_{n \to \infty} \left[ e_n(C_\varphi) \right]^{1/\sqrt{n}} = 1 . \] (1.2)

The proofs are not difficult, but as indicated for example by the comparison between (1.1) and (1.2), the statements feature a very different behavior of those entropy numbers, which deserves attention. This difference is actually more transparent in the polydisk \( \mathbb{D}^N \), and the main interest of this paper is to point out how the dependence of the entropy numbers with respect to the dimension \( N \) differs from that of the approximation numbers.

The paper is organized as follows. Section 1 is this introduction. In Section 2 we recall the necessary background. In Section 3, we survey results on approximation numbers and give the corresponding new results on entropy numbers. We first begin with general facts, and then give specific results, with a particular interest to the examples of the lens maps and the cusp map. For any non-constant analytic map \( \varphi : \mathbb{D} \to \mathbb{D} \), we have:

\[ \lim_{n \to \infty} \left[ a_n(C_\varphi) \right]^{1/n} = \exp \left( - \frac{1}{\text{Cap} [\varphi(\mathbb{D})]} \right) , \] (1.3)

where \( \text{Cap} [\varphi(\mathbb{D})] \) is the Green capacity of \( \varphi(\mathbb{D}) \), from which it follows that \( \lim_{n \to \infty} \left[ a_n(C_\varphi) \right]^{1/n} = 1 \) if and only if \( \| \varphi \|_\infty = 1 \) (Theorem 3.1). We moreover prove that:

\[ \lim_{n \to \infty} \left[ e_n(C_\varphi) \right]^{1/\sqrt{n}} = \exp \left( - \sqrt{\log 2 / \text{Cap} [\varphi(\mathbb{D})]} \right) , \] (1.4)

from which it follows that \( \lim_{n \to \infty} \left[ e_n(C_\varphi) \right]^{1/\sqrt{n}} = 1 \) if and only if \( \| \varphi \|_\infty = 1 \) (Theorem 3.2).

For the lens map \( \lambda_\theta \) with parameter \( \theta \), we have (the constants depending only on \( \theta \)):

\[ a e^{-C n^{1/2}} \leq a_n(C_{\lambda_\theta}) \leq b e^{-c n^{1/2}} , \] (1.5)

whereas:

\[ a' e^{-C' n^{1/3}} \leq e_n(C_{\lambda_\theta}) \leq b' e^{-c' n^{1/3}} . \] (1.6)

For the cusp map \( \chi \), we have, for absolute constants:

\[ a e^{-C n / \log n} \leq a_n(C_\chi) \leq b e^{-c n / \log n} , \] (1.7)

and

\[ a' e^{-C' \sqrt{n / \log n}} \leq e_n(C_\chi) \leq b' e^{-c' \sqrt{n / \log n}} . \] (1.8)

Section 4 is concerned by the multivariate case, again first in the general case, then on specific cases. We are in particular interested by the multi-lens map \( \Lambda_\theta \), defined by:

\[ \Lambda_\theta(z_1, \ldots, z_N) = (\lambda_\theta(z_1), \lambda_\theta(z_2), \ldots, \lambda_\theta(z_N)) , \]

and the multi-cusp map \( \Xi \) defined by:

\[ \Xi(z_1, \ldots, z_N) = (\chi(z_1), \chi(z_2), \ldots, \chi(z_N)) . \]

We prove that:

\[ \exp (-C n^{1/2N}) \leq a_n(C_{\Lambda_\theta}) \leq b \exp (-c n^{1/2N}) \] (1.9)

\[ a' \exp (-C' n^{1/(2N+1)}) \leq e_n(C_{\Lambda_\theta}) \leq b' \exp (-c' n^{1/(2N+1)}) \] (1.10)

and

\[ \exp [-C n^{1/N} / \log n] \leq a_n(C_\Xi) \leq b \exp [-c n^{1/N} / \log n] \] (1.11)

\[ a' \exp [-C' n^{1/(N+1)} (\log n)^{-N/(N+1)}] \leq e_n(C_\Xi) \leq b' \exp [-c' n^{1/(N+1)} (\log n)^{-N/(N+1)}] . \] (1.12)

Section 5 is more specifically devoted to the multidimensional case, in connection with the notion of Monge-Ampère (or Bedford-Taylor) pluricapacity, which recently turned out to play an important role in connection with composition operators [23].
2 Background and notation

We denote by $\mathbb{D}$ the open unit disk and by $T = \partial \mathbb{D}$ the unit circle; $m$ is the normalized Lebesgue measure on $T$: $dm(t) = dt/2\pi$, and $H^2$ is the usual Hardy space on $\mathbb{D}$. By Littlewood's subordination principle, every analytic self-map $\varphi : \mathbb{D} \to \mathbb{D}$ (also called Schur function) defines a bounded operator $C_{\varphi} : H^2 \to H^2$ by $C_{\varphi}(f) = f \circ \varphi$, called the composition operator of symbol $\varphi$.

2.1 Two types of Schur functions

Lens maps

For $0 < \theta < 1$, the lens map $\lambda_\theta$ with parameter $\theta$ is obtained by sending conformally the unit disk $\mathbb{D}$ onto the right half-plane $\Pi = \{z \in \mathbb{C}; \Re z > 0\}$; then making $u \mapsto u^\theta$, and coming back to $\mathbb{D}$ (see [29, page 27]). Namely:

$$\lambda_\theta(z) = \frac{(1 + z)^\theta - (1 - z)^\theta}{(1 + z)^\theta + (1 - z)^\theta}.$$  \hfill (2.1)

It is a conformal map from $\mathbb{D}$ onto the domain represented on Figure 1.

Cusp map

The cusp map is a conformal mapping $\chi$ sending the unit disk $\mathbb{D}$ onto the domain represented on Figure 2. This map was first introduced in [14].

To obtain it, we first map $\mathbb{D}$ onto the half-disk $\mathbb{D}^+ = \{z \in \mathbb{D}; \Re z > 0\}$. To do that, map $\mathbb{D}$ onto itself by $z \mapsto iz$; then map $\mathbb{D}$ onto the upper half-plane $\mathbb{H} = \{z \in \mathbb{C}; \Im z > 0\}$ by

$$T(u) = \frac{1 + u}{1 - u}.$$  \hfill (2.2)

Take the square root to map $\mathbb{H}$ in the first quadrant $Q_1 = \{z \in \mathbb{H}; \Re z > 0\}$, and go back to the half-disk $\{z \in \mathbb{D}; \Im z < 0\}$ by $T^{-1}; T^{-1}(s) = \frac{1 + is}{1 - is}$; finally, make a rotation by $i$ to go onto $\mathbb{D}^+$. We get

$$\chi_0(z) = \frac{(z - i)^{1/2} - i}{(z - i)^{1/2} + 1}.$$  \hfill (2.3)

One has $\chi_0(1) = 0, \chi_0(-1) = 1, \chi_0(i) = -i$ and $\chi_0(-i) = i$. The half-circle $\{z \in \mathbb{T}; \Re z \geq 0\}$ is mapped onto the segment $[-i, i]$ and the segment $[-1, 1]$ onto the segment $[0, 1]$.

Set now, successively,

$$\chi_1(z) = \log \chi_0(z), \quad \chi_2(z) = -\frac{2}{\pi} \chi_1(z) + 1, \quad \chi_3(z) = \frac{1}{\chi_2(z)}.$$  \hfill (2.3)
and finally:

\[ \chi(z) = 1 - \chi(z). \]  

(2.4)

Hence:

\[ 1 - \chi(z) = \frac{1}{1 + \frac{2}{\pi} \log \left( \frac{1}{|\chi_0(z)|} \right) - i \frac{2}{\pi} \arg \chi_0(z)}. \]

(2.5)

Note that \( \chi_2 \) maps \( \mathbb{D} \) onto the half-strip \( \{ z \in \mathbb{C}; \Re z > 1 \text{ and } |\Im z| < 1 \} \). One has \( \varphi(1) = 1, \varphi(-1) = 0, \varphi(i) = (1 + i)/2 \) and \( \varphi(-i) = (1 - i)/2 \).

The domain \( \varphi(\mathbb{D}) \) is edged by three circular arcs of radii \( 1/2 \) and of respective centers \( 1/2, 1 + i/2 \) and \( 1 - i/2 \). The real interval \( ]-1, 1[ \) is mapped onto the real interval \( ]0, 1[ \) and the half-circle \( \{ e^{i\theta}; |\theta| \leq \pi/2 \} \) is sent onto the two circular arcs tangent at 1 to the real axis.

### 2.2 Approximation and entropy numbers

Given an operator \( T : X \rightarrow Y \) between Banach spaces, recall (see [6]) that we can attach to this operator five non-increasing sequences \( (a_n), (b_n), (c_n), (d_n), (e_n) \) of non-negative numbers (depending on \( T \)), respectively the sequences of approximation, Bernstein, Gelfand, Kolmogorov, and entropy numbers of \( T \). We only define here the first one and the last one. The approximation numbers are defined as:

\[ a_n(T) = \inf \{ \| T - R \| ; \text{ rank } (R) < n \}, \quad n \geq 1. \]

(2.6)

The entropy numbers are defined for \( n \geq 1 \) as:

\[ e_n(T) = \inf \{ \varepsilon > 0 ; N(T(B_X), \varepsilon B_Y) \leq 2^{n-1} \}, \quad n \geq 1, \]

(2.7)

where \( B_X \) and \( B_Y \) are the respective closed unit balls of \( X \) and \( Y \), and where, for \( A, B \subseteq Y, N(A, B) \) denotes the smallest number of translates of \( B \) needed to cover \( A \) (see [6, Chapter 1], or [28, Chapter 5]).

All those sequences \( (a_n), (b_n), (c_n), (d_n), (e_n) \), say \( (u_n) \), share the ideal property:

\[ u_n(ATB) \leq \| A \| u_n(T) \| B \|. \]

(2.8)

For Hilbert spaces, it turns out that

\[ a_n = b_n = c_n = d_n = s_n, \]

(2.9)

where \( (s_n) \) designates the sequence of singular numbers; but entropy numbers stay a little apart.

For Banach spaces \( X \) and \( Y \) and \( T : X \rightarrow Y \), we have, in general, for \( \alpha > 0 \):

\[ \sup_{1 \leq k \leq n} k^\alpha e_k(T) \leq C_n \sup_{1 \leq k \leq n} k^\alpha a_k(T) \]

([5, Theorem 1], see also [28, Theorem 5.2]), and, if \( X \) and \( Y^* \) are of type 2:

\[ a_n(T) \leq K e_n(T), \quad \text{ for all } n \geq 1 \]

([12, Corollary 1.6]), where \( K = \kappa \left( T_2(X)T_2(Y^*) \right)^2 \); in particular, if \( T \) acts between Hilbert spaces:

\[ a_n(T) \leq 4 e_n(T), \quad \text{ for all } n \geq 1 \]

(see [28, Theorem 5.3]).
A useful corollary of Theorem 2.1 is the following:

\[
\left( \frac{1}{2^n} \sum_{\epsilon_1, \ldots, \epsilon_n=1} \left\| \sum_{k=1}^{n} \epsilon_k x_k \right\| \right)^{1/2} \leq C \left( \sum_{k=1}^{n} \|x_k\|^2 \right)^{1/2},
\]

the smallest such constant \(C\) is denoted \(T_2(X)\). Every Hilbert space has type 2, thanks to the parallelogram identity.

Those inequalities indicate that entropy numbers are always bigger than singular numbers, up to a constant, and that, as far as the scale of powers \(n^\alpha\) is implied, they are dominated by approximation numbers in a weak sense. But it turns out that, individually, they can be much bigger than the latter for composition operators, as we shall see.

We will rely on the following estimate ([6, Proposition 1.3.2, p. 17]), in which \(\ell^2\) denotes the space of square-summable sequences \(x = (x_k)_{k=1}^\infty\) of complex numbers. This estimate is given for the sequence \((\epsilon_n)\), but \(e_n = \epsilon_{2^{n-1}}\), by definition.

**Theorem 2.1.** (see [6, p. 17]) There exist absolute constants \(0 < a < b\) such that, for any diagonal compact operator \(\Delta: \ell^2 \to \ell^2\) with positive and non-increasing eigenvalues \((\sigma_k)_{k=1}^\infty\), namely \(\Delta(x_k) = (\sigma_k x_k)_k\), we have, for all \(n \geq 1\):

\[
a \sup_{k=1} 2^{-n/2k} \left( \prod_{j=1}^k \sigma_j \right)^{1/k} \leq e_n(\Delta) \leq b \sup_{k=1} 2^{-n/2k} \left( \prod_{j=1}^k \sigma_j \right)^{1/k},
\]

(2.10)

A useful corollary of Theorem 2.1 is the following.

**Theorem 2.2.** Let \(T: H_1 \to H_2\) be a compact operator between the complex Hilbert spaces \(H_1\) and \(H_2\), and let \((a_n)_{n=1}^\infty\) be its sequence of approximation numbers. Then, for all \(n \geq 1\):

\[
a \sup_{k=1} 2^{-n/2k} \left( \prod_{j=1}^k a_j \right)^{1/k} \leq e_n(T) \leq b \sup_{k=1} 2^{-n/2k} \left( \prod_{j=1}^k a_j \right)^{1/k},
\]

(2.11)

where \(a\) and \(b\) are positive numerical constants.

**Proof.** Let \(Tx = \sum_{n=1}^\infty s_n(x \mid u_n) v_n\) the Schmidt decomposition of \(T\), where \((u_n)_n\) and \((v_n)_n\) are orthonormal sequences of \(H_1\) and \(H_2\), respectively, and \((s_n)_n\) is the sequence of singular numbers of \(T\). Let \(\Delta: \ell^2 \to \ell^2\) the diagonal operator with diagonal values \(s_n\), \(n \geq 1\). Then \(T = V_1 \Delta U_1\) and \(\Delta = V_2 T U_2\), with \(U_1 x = (x \mid u_n)_n\), \(V_1((t_n)_n) = \sum_n t_n v_n\), \(U_2((t_n)_n) = \sum_n t_n u_n\) and \(V_2 x = (x \mid v_n)_n\). We have \(\|U_1\|, \|V_1\|, \|U_2\|, \|V_2\| \leq 1\); hence the result follows from Theorem 2.1 and the ideal property (2.8).

This theorem might be thought useless, because we do not know better the numbers \(a_n\) than the numbers \(e_n\). In our situation, this is not the case, since we made a more or less systematic study of the approximation numbers \(a_n\) for composition operators in [3, 19–21] for example.

## 3 The 1-dimensional case

### 3.1 General results

In [21], we coined the parameter

\[
\beta_1(T) = \lim_{n \to \infty} \left[ a_n(T) \right]^{1/n}
\]

(3.1)

and its versions \(\beta_1(T), \beta_1(T)\) with an upper limit and a lower limit respectively. The following result, proved in [21, Theorem 3.1] for \(\|\phi\|_\infty < 1\) and in [21, Theorem 3.14] for \(\|\phi\|_\infty = 1\), shows in particular that no lower or
The first item says in particular that we always have:

$$\beta_1(C_\varphi) = \exp[-1/\text{Cap}(\Omega)] > 0.$$  \hfill (3.2)

2) In particular, one has the equivalence

$$\beta_1(C_\varphi) = 1 \iff \|\varphi\|_\infty = 1.$$  \hfill (3.3)

The first item says in particular that we always have:

$$a_n(C_\varphi) \gtrsim r^n$$  \hfill (3.4)

for some positive constant $r < 1$ [19, Theorem 3.1]. This was actually first pointed out by Parfenov [27].

The second item says that the behavior $a_n(C_\varphi) = r^n$ is only obtained when $\|\varphi\|_\infty < 1$ ([19, Theorem 3.4] or [21, Theorem 3.14]).

For entropy numbers, another parameter emerges:

$$\gamma_1(T) = \lim_{n \to \infty} \left[ e_n(T) \right]^{1/\sqrt{n}}$$  \hfill (3.5)

and its $\gamma_1^-(T)$ and $\gamma_1^+(T)$ versions.

**Theorem 3.2.** Let $\varphi : \mathbb{D} \to \mathbb{D}$ be a non-constant symbol and $\Omega = \varphi(\mathbb{D})$. Then

1) $\gamma_1^-(C_\varphi) = \gamma_1^+(C_\varphi) =: \gamma_1(C_\varphi)$ and

$$\gamma_1(C_\varphi) = \exp[-\sqrt{\log 2/\text{Cap}(\Omega)}] > 0.$$  \hfill (3.6)

2) In particular, one has the equivalence

$$\gamma_1(C_\varphi) = 1 \iff \|\varphi\|_\infty = 1.$$  \hfill (3.7)

**Proof.** Set $\rho = 1/\text{Cap}(\Omega)$ for simplicity of notations. Let $\varepsilon > 0$, and $C_\varepsilon$ a positive constant which depends only on $\varepsilon$ and can vary from a formula to another. Theorem 3.1 implies $a_k \leq C_\varepsilon e^{\varepsilon k^2} e^{-k \rho}$, whence

$$(a_1 \cdots a_k)^{1/k} \leq C_\varepsilon e^{\varepsilon k/2} e^{-k \rho/2}.$$  

Theorem 2.2 gives

$$e_n(C_\varphi) \leq C_\varepsilon \sup_{k \geq 1} \left\{ e^{\varepsilon k/2} e^{-n(k)(\log 2/\rho)-(pk/2)} \right\}.$$  

This supremum is essentially attained for $k = \lceil \sqrt{(n \log 2)/\rho} \rceil$ where $\lceil \cdot \rceil$ stands for the integer part, and gives

$$e_n(C_\varphi) \leq C_\varepsilon e^{\frac{\varepsilon k}{2}} \sqrt{n \log 2/\rho} e^{-\sqrt{n \rho \log 2}}.$$  

This implies $\gamma_1^-(C_\varphi) \leq e^{\frac{\varepsilon k}{2}} \sqrt{\log 2/\rho} e^{-\sqrt{\rho \log 2}}$, and finally

$$\gamma_1^-(C_\varphi) \leq e^{-\sqrt{\rho \log 2}}.$$  

The lower bound $\gamma_1^-(C_\varphi) \geq e^{-\sqrt{\rho \log 2}}$ is proved similarly.

This clearly ends the proof, since we know that $\text{Cap}(\Omega) = \infty$ if and only if $\|\varphi\|_\infty = 1$ (see [21, Theorem 3.13]).
3.2 Estimates for approximation numbers

Estimates of approximation numbers of composition operators can be obtained by using the boundary behavior of the symbol and Blaschke products for the upper estimates, and the radial behavior of this symbol, reproducing kernels, and interpolating sequences, for the lower estimates.

In order to treat simultaneously the two cases of the lens maps and of the cusp map, we will put ourselves in a more general situation.

We say that a continuous function \( \omega : [0, 1] \to \mathbb{R}_+ \) is a modulus of continuity if it is increasing, subadditive, i.e. \( \omega(s + t) \leq \omega(s) + \omega(t) \), and vanishes at 0.

**Upper estimates**

Let \( \varphi \) be a symbol in the disk algebra, i.e. \( \varphi : \mathbb{D}_+ \to \mathbb{D} \) is continuous on \( \partial \mathbb{D} \) and analytic in \( \mathbb{D} \), and such that \( \varphi(\partial \mathbb{D}) \cap \partial \mathbb{D} = \{\xi_1, \ldots, \xi_p\} \). We say that this symbol is boundary regular if, writing \( \xi_j = e^{i\theta_j} \), we have:

1) for some positive constant \( C \), we have \( |\varphi(e^{i\theta}) - \varphi(e^{i\theta_j})| \leq C (1 - |\varphi(e^{i\theta_j})|) \), for \( t \) in a neighborhood of \( t_j \) and for \( j = 1, \ldots, p \);

2) for some modulus of continuity \( \omega \) and for some positive constant \( c \), we have \( c \omega(|t - t_j|) \leq |\varphi(e^{i\theta}) - \varphi(e^{i\theta_j})| \), for \( t \) in a neighborhood of \( t_j \) and for \( j = 1, \ldots, p \).

The following result is proved in [20, Theorem 2.3].

**Theorem 3.3.** Let \( \varphi \) be a symbol in the disk algebra whose image touches \( \partial \mathbb{D} \) at the points \( \xi_1, \ldots, \xi_p \), and nowhere else, and such that \( \varphi \) is boundary regular. Then, there are constants \( k, K, L > 0 \), depending only on \( \varphi \), such that, for every \( q \geq 1 \):

\[
a_q(C_\varphi) \leq K \sqrt{\frac{\omega^{-1}(k^2 N_q)}{2^{-N_q}}},
\]

where \( N_q \) is the largest integer such that \( p N_d N < q \), with \( d_N \) the integer part of \( \sigma \log \frac{k^2}{(k^2)^n} + 1 \).

This theorem allows to give an upper estimate for all approximation numbers \( a_n(C_\varphi) \), \( n \geq 1 \) when we can interpolate between the integers \( N_d N \) and \( (N + 1) d_N + 1 \), but this is not the case in general. Nevertheless, this is the case in the examples below.

**Theorem 3.4.**

1) For the lens map \( \lambda_\theta \) with parameter \( \theta \), we have, for some positive constants \( \alpha \) and \( \beta \), depending on \( \theta \):

\[
a_n(C_\lambda_\theta) \leq \alpha e^{-\beta n^{1/\log n}}.
\]

2) For the cusp map \( \chi \), we have, for some positive constants \( \alpha \) and \( \beta \):

\[
a_n(C_\chi) \leq \alpha e^{-\beta n^{1/\log n}}
\]

**Proof.**

1) The map \( \lambda_\theta \) satisfies the conditions of Theorem 3.3 with \( \omega(h) = h^\theta \), with \( p = 2 \). We have \( \omega^{-1}(h) = h^{1/\theta} \). Hence \( d_N = N, N_q = \sqrt{q} \), and we then get from (3.8) that \( a_q(C_\lambda_\theta) \leq \alpha 2^{-N} \) for \( q \geq N^2 \), with \( \delta > 0 \), which gives (3.9).

2) The map \( \chi \) satisfies the conditions of Theorem 3.3 with \( p = 1 \) and \( \omega(h) = 1/(\log 1/h) \). Then, \( \omega^{-1}(h) = e^{-1/h} \) and \( d_N = 2^N \), so that \( N_q = \log q \) and \( 2^{N_q} = q/\log q \), and (3.10) follows.

The proof of Theorem 3.3 is based on the following Lemma 3.5, with a suitable choice of the Blaschke product of length \( p N d \) (\( d \) a positive integer to be specified):

\[
B(z) = \prod_{j=1}^{p} \prod_{k=1}^{N} \left[ \frac{z - p_{j,k}}{1 - p_{j,k} z} \right]^d,
\]

where \( p_{j,k} = (1 - 2^{-k}) \xi_j \), for \( j = 1, \ldots, p \) and \( k = 1, 2, \ldots \). We do not give the details here and refer to [20, Proof of Theorem 2.3].
Lemma 3.5 ([15, Lemma 2.4]). For every Blaschke product \( B \) with less than \( N \) zeros (each of them being counted with its multiplicity), one has:

\[
[ a_N(C_\varphi) ]^2 \lesssim \sup_{0<h<1,|\xi|=1} \frac{1}{h} \int_{S(\xi,h)} |B(z)|^2 \, dm_\varphi(z),
\]

where \( S(\xi,h) = \{ z \in \mathbb{T} : |z-\xi| \leq h \} \) and \( m_\varphi \) is the pull-back measure by \( \varphi \) of the normalized Lebesgue measure \( m \) on \( \mathbb{T} \).

The proof of Lemma 3.5 comes from the estimate of the Carleson norm in the Carleson’s theorem and from the fact that the subspace \( BH^2 \) is of codimension \( \leq N - 1 \), leading to a majorization of the Gelfand number \( c_N(C_\varphi) \), and then using (2.9). For the convenience of the reader, we reproduce the proof.

Proof of Lemma 3.5. The subspace \( BH^2 \) is of codimension \( \leq N - 1 \). Therefore, \( a_N = c_N(C_\varphi) \leq \| C_\varphi|_{BH^2} \| \), where \( c_N \) is the \( n \)-th Gelfand number and where we used the equality \( a_N = c_N \) occurring in the Hilbertian case, as recalled in the introduction. Now, since \( \| Bf \|_{H^2} = \| f \|_{H^2} \) for any \( f \in H^2 \), we have:

\[
\| C_\varphi|_{BH^2} \| = \sup_{\| f \|_{H^2} = 1} \int_{\mathbb{T}} |B \circ \varphi|^2 |f \circ \varphi|^2 \, dm = \sup_{\| f \|_{H^2} = 1} \int_{\mathbb{T}} |B|^2 |f|^2 \, dm_\varphi = \| R_\mu \|^2,
\]

where \( \mu = |B|^2 m_\varphi \) and where \( R_\mu : H^2 \to L^2(\mu) \) is the restriction map. Of course, \( \mu \) is a Carleson measure for \( H^2 \) since \( \mu \leq m_\varphi \). Now, Carleson’s embedding theorem says us that \( \| R_\mu \|^2 \leq \kappa^2 \sup_{0<h<1,|\xi|=1} \frac{\int_{S(\xi,h)} |B|^2 \, dm_\varphi}{h} \) (see [9], Remark after the proof of Theorem 9.3, at the top of page 163; actually, in that book, Carleson’s windows \( W(\xi, h) \) are used instead of pseudo-Carleson’s windows \( S(\xi, h) \), but that does not matter, since \( W(\xi, h) \subseteq S(\xi, 2h) \); if \( r \gtrsim 1 - h \) and \( |t-t_0| \leq h \), then \( |re^{it} - e^{it_0}| \leq |re^{it} - e^{it}| + |e^{it} - e^{it_0}| \leq 2h \).

Lower estimates

We consider symbols \( \varphi \) taking real values in the real axis (i.e. its Taylor series has real coefficients) and such that \( \lim_{r \to 1} \varphi(r) = 1 \), with a given speed.

We say that \( \varphi \) is radially regular if it takes real values on \( ]-1,1[ \) and there exists a modulus of continuity \( \omega : [0, 1] \to [0, 2] \) such that \( 1 - \varphi(r) \leq \omega(1-r) \) for \( 0 \leq r < 1 \). Then we have the following result ([20, Theorem 3.2]).

Theorem 3.6. Let \( \varphi \) be a radially regular symbol. Then, for the approximation numbers \( a_n(C_\varphi) \) of the composition operator \( C_\varphi \) of symbol \( \varphi \), one has the following lower bound.

\[
a_n(C_\varphi) \geq c \sup_{0<\sigma<1} \sqrt{\frac{\omega^{-1}(a \sigma^m)}{a \sigma^n}} \exp \left[ -\frac{20}{1-\sigma} \right],
\]

where \( a = 1 - \varphi(0) > 0 \) and \( c \) is another constant depending only on \( \varphi \).

For our examples, we get:

Theorem 3.7.

1) For the lens map \( \lambda_\theta \) of parameter \( \theta \), we have, for some positive constants \( c \) and \( C \), depending on \( \theta \):

\[
a_n(C_{\lambda_\theta}) \geq c \exp \left( -\frac{C}{\sqrt{n}} \right).
\]

2) For the cusp map \( \chi \), we have, for some positive constants \( c \) and \( C \):

\[
a_n(C_\varphi) \geq c \exp(-C n/ \log n).
\]
Proof. 1) The lens map $\lambda_0$ satisfies the conditions of Theorem 3.6 with $\omega^{-1}(h) = h^{1/\beta}$. We get the result by adjusting $\sigma = 1 - 1/\sqrt{n}$.

2) The cusp map $\chi$ satisfies the conditions of Theorem 3.6 with $\omega^{-1}(h) = e^{-C'/h}$, and by taking $\sigma = \exp(-\log n/2n)$, we get the result. \hfill $\Box$

The tools for proving Theorem 3.6 are the following.

Recall (see [13, pages 194–195], or [25, pages 302–303]) that if $(z_j)$ is a Blaschke sequence, its Carleson constant $\delta$ is defined as

$$\delta = \inf_{j \geq 1} (1 - |z_j|^2) |B'(z_j)|,$$

where $B$ is the Blaschke product whose zeros are the points $z_j$. Recall also (see [7], [13, pages 194–195], or [25, pages 302–303]) that an interpolation sequence $(z_n)$ with (best) interpolation constant $C$ is a sequence $(z_n)$ (necessarily Blaschke, i.e. $\sum_{n=1}^{\infty} (1 - |z_n|) < \infty$) in the unit disk such that, for any bounded sequence $(w_n)$ of scalars, there exists a bounded analytic function $f$ (i.e. $f \in H^\infty$) such that:

$$f(z_n) = w_n, \quad \forall n \geq 1, \quad \text{and} \quad \|f\|_\infty \leq C \sup_{n \geq 1} |w_n|.$$ 

Now (see [11, Chapter VII, Theorem 1.1]), every $H^\infty$-interpolation sequence $(z_j)$ is a Blaschke sequence and its Carleson constant $\delta$ is connected to its interpolation constant $C$ by the inequalities

$$1/\delta \leq C \leq \kappa/\delta^2 \tag{3.15}$$

where $\kappa$ is an absolute constant (actually $C \leq \kappa_1 (1/\delta)(1 + \log 1/\delta)$). Now, if $(z_j)$ is a $H^\infty$-interpolation sequence with constant $C$, the sequence of the normalized reproducing kernels $f_j = K_{z_j}/\|K_{z_j}\|$, viz $f_j(z) = (1-|z|^2)/(1-\overline{z}z_j)$, satisfies

$$C^{-1} \left( \sum_j |\lambda_j|^2 \right)^{1/2} \leq \left\| \sum_j \lambda_j f_j \right\|_{H^2} \leq C \left( \sum_j |\lambda_j|^2 \right)^{1/2} \tag{3.16}$$

(see [19, Lemma 2.2]).

We then use the following lemma [20, Lemma 3.3], with $u_j \in [0, 1)$ defined inductively by $u_0 = 0$ and the relation:

$$1 - \varphi(u_{j+1}) = a(1 - \varphi(u_j)) \quad \text{with} \quad 1 > u_{j+1} > u_j,$$

using the intermediate value theorem, where $0 < a < 1$ is a fixed positive number. Note that, when setting $\nu_j = \varphi(u_j)$, we have $-1 < \nu_j < 1$,

$$\frac{1 - \nu_{j+1}}{1 - \nu_j} = \sigma, \tag{3.17}$$

and

$$1 - \nu_n = a \sigma^n, \tag{3.18}$$

with $a = 1 - \varphi(0)$.

Lemma 3.8. Let $\varphi : \mathbb{D} \to \mathbb{D}$ be an analytic self-map. Let $u = (u_1, \ldots, u_n)$ be a finite sequence in $\mathbb{D}$ and set $\nu_j = \varphi(u_j)$, $\nu = (\nu_1, \ldots, \nu_n)$. Denote by $\delta_n$ the Carleson constant of the finite sequence $\nu$ and set

$$\mu_n^2 = \inf_{1 \leq j \leq n} \frac{1 - |u_j|^2}{1 - |\varphi(u_j)|^2}.$$ 

Then, for some constant $c' > 0$, we have the lower bound:

$$a_n(C_{\varphi}) \geq c' \delta_n^2 \mu_n. \tag{3.19}$$

For the details of the proof, we refer to [20, Proof of Theorem 3.2, pages 556-557].
3.3 Entropy numbers

In this section, we give the behavior of the entropy numbers of the lens maps and the cusp map, seeing that they significantly differ from that of the approximation numbers.

**Theorem 3.9.** Let $\lambda_0$ be the lens map with parameter $\theta$. Then, with positive constants $a$, $b$, $a'$, $b'$ depending only on $\theta$:

$$a' e^{-b'n^{1/3}} \leq e_n(C_{\lambda_0}) \leq a e^{-bn^{1/3}}.$$  

**Proof.** We use Theorem 3.9. It follows, using Theorem 2.2, writing $a_k = a_k(C_{\lambda_0})$, that $(a_1 \cdots a_k)^{1/k} \leq a e^{-b\sqrt{\chi}}$ and that, for some positive constant $C$:

$$e_n(C_{\lambda_0}) \leq C \exp \left[ - (n/2k) \log 2 + b k^{1/2} \right].$$

Taking $k = \lceil n^{2/3} \rceil$ gives the claimed upper bound, with a different value of $b$.

The lower bound is proved similarly, using Theorem 3.13 and the left inequality in Theorem 2.2. \qed

**Theorem 3.10.** Let $\chi$ be the cusp map. Then, with positive constants $a$, $b$, $a'$, $b'$:

$$a' e^{-b'\sqrt{n/\log n}} \leq e_n(C_{\chi}) \leq a e^{-b'\sqrt{n/\log n}}.$$  

**Proof.** The proof follows the same lines as in Theorem 3.9. Using Theorem 3.10, we get, with another $b$:

$$\prod_{j=1}^{k} [a_j(C_{\chi})] \leq e^{-b k^2/\log k},$$

and Theorem 2.2, with the choice $k = \lceil \sqrt{n \log n} \rceil$, gives the upper bound, with a different value of $b$.

The lower bound is obtained similarly, using now Theorem 3.14 and Theorem 2.2. \qed

4 The multidimensional case

4.1 General results

Let $\varphi : \mathbb{D}^N \to \mathbb{D}^N$ be an analytic map. We will say that $\varphi$ is non-degenerate if $\varphi(\mathbb{D}^N)$ has non-empty interior, equivalently if $\det \varphi'(z) \neq 0$ for at least one point $z \in \mathbb{D}^N$.

Let now $\varphi : \mathbb{D}^N \to \mathbb{D}^N$ be a non-degenerate analytic map inducing a bounded composition operator $C_\varphi : H^2(\mathbb{D}^N) \to H^2(\mathbb{D}^N)$ (this is not always the case as soon as $N > 1$, even if $\varphi$ is injective and hence non-degenerate, see for example [8, p. 246], when the polydisk is replaced by the ball; but similar examples exist for the polydisk). Assume moreover that $C_\varphi$ is a compact operator.

**Theorem 4.1.** Let $C_\varphi : H^2(\mathbb{D}^N) \to H^2(\mathbb{D}^N)$ be a compact composition operator, with $\varphi$ non-degenerate. We have:

1) $e_n(C_{\varphi}) \geq c \exp \left( - C n \frac{\pi}{\lambda} \right)$, for some constants $C > c > 0$, depending on $\varphi$;

2) if $\|\varphi\|_{\infty} < 1$, then $e_n(C_{\varphi}) \leq C \exp \left( - c n \frac{\pi}{\lambda} \right)$, with $C > c > 0$ depending on $\varphi$.

**Proof.** 1) It is proved in [3, Theorem 3.1] that, for a non-degenerate map $\varphi$, it holds:

$$a_k(C_{\varphi}) \geq a' e^{-b' \lambda^{3/N}}.$$  

As in the previous section, it follows from Theorem 2.2, that $(a_1 \cdots a_k)^{1/k} \geq e^{-b' \lambda^{3/N}}$, and then, taking $k = \lceil n^{N/(N+1)} \rceil$, that:

$$e_n(C_{\varphi}) \geq c \exp^{-c \lambda^{3/(N-1)}}.$$


2) Similarly, for \( \| \varphi \|_\infty < 1 \), it is proved in [3, Theorem 5.2] that
\[
a_k(C_\varphi) \leq C e^{-k^{1/N}},
\]
and we get the result from Theorem 2.2.

Those estimates motivate the introduction of the parameter
\[
\frac{1}{2} \Phi = \text{two.tf.}
\]
\[
\phi = \text{four.tf.}
\]
map \([3, \text{Theorem 6.1 and Theorem 6.2}]\). First, we define the multi-lens map \([3]\). However, we recall the result that can be obtained from that for the multi-lens maps and the multi-cusp was obtained in \([3, \text{Theorem 5.5}]\) when the polydisk \(\text{compact composition operators on} \mathbb{D} \rightarrow \mathbb{D}\).

Theorem 4.3.

Theorem 4.2.

With the same proof as in Theorem 3.9 and Theorem 3.10, we obtain estimates for the entropy numbers.

4.2 Specific results

When \( \Phi : \mathbb{D}^N \rightarrow \mathbb{D}^N \) is a non degenerate analytic map inducing a compact composition operator \( C_\varphi : H^2(\mathbb{D}^N) \rightarrow H^2(\mathbb{D}^N) \), a version of Theorem 3.6 is proved in \([3, \text{Theorem 4.2}]\). For upper bounds, a result was obtained in \([3, \text{Theorem 5.5}]\) when \( \Phi = \varphi_1 \circ \cdots \circ \varphi_N \), where \( \varphi_1, \ldots, \varphi_N : \mathbb{D} \rightarrow \mathbb{D} \) are symbols inducing compact composition operators on \( H^2(\mathbb{D}) \). We do not give here the precise statements and refer to the paper \([3]\). However, we recall the result that can be obtained from that for the multi-lens maps and the multi-cusp map \([3, \text{Theorem 6.1 and Theorem 6.2}]\). First, we define the multi-lens map \( \Lambda_\theta \) and the multi-cusp map.

Let \( \lambda_\theta \) be lens maps with parameter \( \theta \). The multi-lens map \( \Lambda_\theta \) with parameter \( \theta \) the multi-cusp map on the polydisk \( \mathbb{D}^N \) are defined respectively as:
\[
\Lambda_\theta(z_1, \ldots, z_N) = (\lambda_\theta(z_1), \lambda_\theta(z_2), \ldots, \lambda_\theta(z_N)), \tag{4.2}
\]
and:
\[
\Xi(z_1, \ldots, z_N) = (\chi(z_1), \chi(z_2), \ldots, \chi(z_N)), \tag{4.3}
\]
for \((z_1, \ldots, z_N) \in \mathbb{D}^N\).

Theorem 4.2.

1) For positive constants \( a, b, a', b' \) depending only on \( \theta \) and \( N \), one has
\[
a' e^{-b' n^{1/(2N)}} \leq a_n(C_{\Lambda_\theta}) \leq a e^{-b n^{1/(2N)}}. \tag{4.4}
\]

2) For positive constants \( a, b, a', b' \) depending only on \( N \), one has
\[
a' e^{-b' n^{1/N} / \log n} \leq a_n(C_{\Xi}) \leq a e^{-b n^{1/N} / \log n}, \tag{4.5}
\]

With the same proof as in Theorem 3.9 and Theorem 3.10, we obtain estimates for the entropy numbers.

Theorem 4.3.

1) Let \( \Lambda_\theta \) be the multi-lens map with parameter \( \theta \). Then:
\[
a' \exp (-b' n^{1/(2N+1)}) \leq a_n(C_{\Lambda_\theta}) \leq a \exp (-b n^{1/(2N+1)}). \tag{4.6}
\]

2) Let \( \Xi : \mathbb{D}^N \rightarrow \mathbb{D}^N \) be the multi-cusp map. Then:
\[
a' \exp \left[ -b' \left( \log n \right)^{N/(N+1)} \right] \leq a_n(C_{\Xi}) \leq a \exp \left[ -b \left( \log n \right)^{N/(N+1)} \right]. \tag{4.7}
\]
5 Links with pluricapacity and Zakharyuta’s results

Here, in dimension \( N \geq 2 \), the situation is satisfactory for upper bounds (see [22]); for lower bounds, see [23]. The notion involved is now that of pluricapacity, or Monge-Ampère capacity, coined by Bedford and Taylor in [4]. More precisely, if \( A \) is a Borel subset of \( \mathbb{D}^N \), we refer to [22] or [23] for the definition of its pluricapacity \( \text{Cap}_N(A) \), belonging to \([0, +\infty)\), and set:

\[
\begin{align*}
\tau_N(A) &= \frac{1}{(2\pi)^N} \text{Cap}_N(A) \\
\Gamma_N(A) &= \exp \left[ - \left( \frac{N!}{\tau_N(A)} \right)^{1/N} \right] \\
\beta_N(T) &= \limsup_{n \to \infty} \left[ a_n(T) \right]^{1/n^{1/N}}.
\end{align*}
\]

We temporarily assume that \( \|\varphi\|_\infty < 1 \) so that \( K = \varphi(\mathbb{D}^N) \) is a compact subset of \( \mathbb{D}^N \). We proved in [22, Theorem 6.4], relying on positive results of Nivoche ([26]) and Zaharyuta ([33, Proposition 6.1]) on the so-called Kolmogorov conjecture, that:

**Theorem 5.1.** It holds:

\[
\beta_N(C\varphi) \leq \Gamma_N(K).
\]

We have the following result, which extends the previous result in dimension 1.

**Theorem 5.2.** The following upper bound holds:

\[
\gamma_N(C\varphi) \leq \exp \left( - \beta_N \rho^{N/(N+1)} \right),
\]

where

\[
\rho = \left( \frac{N!}{\tau_N(K)} \right)^{1/N} = 2\pi \left( \frac{N!}{\text{Cap}_N(K)} \right)^{1/N},
\]

and

\[
\beta_N = (\log 2)^{1/(N+1)} \left( \frac{2N}{N+1} \right)^{N/(N+1)}.
\]

**Proof.** Abbreviate \( a_n(C\varphi) \) and \( e_n(C\varphi) \) to \( a_n \) and \( e_n \), and set \( \alpha = N/(N + 1) \). Let \( \varepsilon > 0 \). Theorem 5.1 implies:

\[
a_k \leq C \varepsilon e^{\varepsilon k^{1/N}} e^{-\rho k^{1/N}},
\]

so:

\[
(a_1 \cdots a_k)^{1/k} \leq C \varepsilon e^{\varepsilon k^{1/N}} e^{-\rho k^{1/N}}.
\]

Apply once more Theorem 2.2 to obtain:

\[
e_n \leq C \varepsilon \sup_{k \geq 1} e^{\varepsilon k^{1/N}} \exp \left[ - \left( \frac{n}{2k} \log 2 + \rho \alpha k^{1/N} \right) \right].
\]

The supremum is essentially attained for \( k \) the integral part of \( (\log 2/2\rho \alpha)^n n^\alpha \) and then, in view of (5.7) and \( a/N = 1 - \alpha \), as \( n \) goes to infinity:

\[
\frac{n}{k} \log 2 + \rho \alpha k^{1/N} \sim (\log 2)^{1-a} (2\rho \alpha)^n n^{1-a}.
\]
Finally, with $c_\alpha = (\log 2/2\rho \alpha)^{1-\alpha}$, we have:

$$e_n \leq C e^c n^{1-\alpha} \exp(-\beta_n \rho^a n^{1-\alpha}) = C e^c n^{1/(N+1)} \exp(-\beta_n \rho^N/(N+1)n^{1/(N+1)}) .$$

This clearly ends the proof of Theorem 5.2.

**Remark.** We have so far no sharp lower bound for entropy numbers, at least when $\|\varphi\|_\infty = 1$, since we already fail to have one in general for approximation numbers (see however [23]). In dimension 1, we used the fact that every connected Borel subset $E$ of $\mathbb{D}$ such that $\overline{E} \subseteq \mathbb{D}$, the Green capacity of $E$ is equal to that of $\overline{E}$ ([21, Theorem 2.3]).

Besides, let $J : H^\infty(\mathbb{D}^N) \to C(K)$ be the canonical embedding, when $K \subseteq \mathbb{D}^N$ is a “condenser”, namely a compact subset of $\mathbb{D}^N$ such that any bounded analytic function on $\mathbb{D}^N$ which vanishes on $K$ vanishes identically, which is moreover “regular”. The positive solution to the Kolmogorov conjecture can be expressed in terms of the Kolmogorov numbers $d_n(J)$ of $J$ or equivalently, in terms of the entropy numbers $e_n(J)$ of $J$ ([32, Theorem 5], generalizing Erokhin’s result in dimension 1 appearing in his posthumous paper [10] and methods due to Mityagin [26] and Levin and Tikhomirov [16]; see also [33, Lemma 2.2]). The result is that, taking $K = \overline{\varphi(\mathbb{D}^N)}$, one has, with sharp constants $c_K$, $c'_K$ depending on the pluricapacity of $K$ in $\mathbb{D}^N$:

$$d_n(J) = e^{-c_K n^{1/N}} \quad \text{and} \quad e_n(J) = e^{-c'_K n^{1/(N+1)}} .$$

This jump from the exponent $1/N$ to the exponent $1/(N + 1)$ is reflected in our Theorem 5.2, through the new parameter $\gamma^+_N$.

**Acknowledgement:** L. Rodríguez-Piazza is partially supported by the project PGC2018-094215-B-I00 (Spanish Ministerio de Ciencia, Innovación y Universidades, and FEDER funds).

We thank the referee for a careful reading of this paper.

**References**

[1] F. Albiac and N. J. Kalton, Topics in Banach space theory, Graduate Texts in Mathematics 233, Springer, New York (2006).

[2] O. F. Bandtlow and S. Nivoche, New solution of a problem of Kolmogorov on width asymptotics in holomorphic function spaces, *preprint*.

[3] F. Bayart, D. Li, H. Queffélec, and L. Rodríguez-Piazza, Approximation numbers of composition operators on the Hardy and Bergman spaces of the ball and of the polydisk, Math. Proc. Cambridge Philos. Soc. 165 (2018), no. 1, 69–91.

[4] E. Bedford and B. A. Taylor, A new capacity for plurisubharmonic functions, Acta. Mathematica 149 (1982), 1–40.

[5] B. Carl, Entropy numbers, Cambridge Tracts in Mathematics 98, Cambridge University Press, Cambridge (1990).

[6] B. Carl and I. Stephani, Entropy, Compactness and the Approximation of Operators, Cambridge Tracts in Mathematics 98, Cambridge University Press, Cambridge (1990).

[7] L. Carleson, An interpolation problem for bounded analytic functions, Amer. J. Math. 80 (1958), 921–930.

[8] C. Cowen and B. MacCluer, Composition Operators on Spaces of Analytic Functions, Studies in Advanced Mathematics, CRC Press (1994).

[9] P. Duren, Theory of $H^p$-spaces, Second edition, Dover Publications (2000).

[10] V. D. Erokhin, Best linear approximations of functions analytically continuable from a given continuum into a given region, Russ. Math. Surv. 23 (1968), no. 1, 93–135.

[11] J. B. Garnett, Bounded Analytic Functions, Revised first version, Graduate Texts in Math. 236, Springer (2007).

[12] Y. Gordon, H. König, and C. Schütt, Geometric and probabilistic estimates for entropy and approximation numbers of operators, J. Approx. Theory 49 (1987), no. 3, 219–239.

[13] K. Hoffman, Banach Spaces of Analytic Functions, Prentice-Hall Series in Modern Analysis, Prentice-Hall, Inc., Englewood Cliffs, N. J. (1962).

[14] P. Lefèvre, D. Li, H. Queffélec, and L. Rodríguez-Piazza Compact composition operators on Bergman-Orlicz spaces, Trans. Amer. Math. Soc. 365 (2013), no. 8, 3943–3970.

[15] P. Lefèvre, D. Li, H. Queffélec, and L. Rodríguez-Piazza, Some new properties of composition operators associated with lens maps, Israel J. Math. 195 (2013), no. 2, 801–824.
Approximation and entropy numbers of composition operators

16. A. L. Levin and V. M. Tikhomirov, On theorem of V. D. Erokhin, appendix of Erokhin’s paper, Russian Math. Surveys 23 (1968), no. 1, 119–135.
17. D. Li and H. Queffélec, Introduction à l’étude des espaces de Banach. Analyse et probabilités, Cours Spécialisés 12, Société Mathématique de France, Paris (2004).
18. D. Li and H. Queffélec, Introduction to Banach spaces: analysis and probability, Vol. 1, Cambridge Studies in Advanced Mathematics 166, Cambridge University Press, Cambridge (2018).
19. D. Li, H. Queffélec, and L. Rodríguez-Piazza, On approximation numbers of composition operators, J. Approx. Theory 164 (2012), no. 4, 431–459.
20. D. Li, H. Queffélec, and L. Rodríguez-Piazza, Estimates for approximation numbers of some classes of composition operators on the Hardy space, Ann. Acad. Scient. Fennicae 38 (2013), 547–564.
21. D. Li, H. Queffélec, and L. Rodríguez-Piazza, A spectral radius type formula for approximation numbers of composition operators, J. Funct. Anal. 160 (2015), no. 12, 430–454.
22. D. Li, H. Queffélec, and L. Rodríguez-Piazza, Some examples of composition operators and their approximation numbers on the Hardy space of the bidisk, Trans. Amer. Math. Soc. 372 (2019), no. 4, 2631–2658.
23. D. Li, H. Queffélec, and L. Rodríguez-Piazza, Pluricapacity and approximation numbers of composition operators, J. Math. Anal. Appl. 474 (2019), no. 2, 1576–1600.
24. B. S. Mityagin, Approximative dimension and bases in nuclear spaces, Russian Math. Survey 16 (1963), 59–127.
25. N. K. Nikolski, Operators, Functions and Systems: An Easy Reading, Volume 1, Hardy, Hankel, and Toeplitz, Math. Surveys and Monographs 92, Amer. Math. Soc., Providence, RI (2002).
26. S. Nivoche, Proof of a conjecture of Zaharyuta concerning a problem of Kolmogorov on the ε entropy, Invent. Math. 158 (2004), no. 2, 413–450.
27. O. G. Parfenov, Estimates of the singular numbers of the Carleson embedding operator, Math. USSR Sbornik 59 (2) (1988), 497–511.
28. G. Pisier, The volume of convex bodies and Banach space geometry, Cambridge Tracts in Mathematics 94, Cambridge University Press, Cambridge (1989).
29. J. H. Shapiro, Composition operators and classical function theory, Universitext, Tracts in Mathematics, Springer-Verlag, New York (1993).
30. J. Wengenroth, Private discussion, Liège (2011).
31. H. Widom, Rational approximation and n-dimensional diameter, J. Approx. Theory 5 (1972), 342–361.
32. V. Zaharyuta, On asymptotics of entropy of a class of analytic functions, Funct. Approx. Comment. Math. 44, part 2 (2011), 307–315.
33. V. Zaharyuta, Extendible bases and Kolmogorov problem on asymptotics of entropy and widths of some classes of analytic functions, Annales de la Faculté des Sciences de Toulouse Vol. XX, numéro spécial (2011), 211–239.