Embeddings of model subspaces of the Hardy space: compactness and Schatten–von Neumann ideals

A. D. Baranov

Abstract. We study properties of the embedding operators of model subspaces \( K^p_\Theta \) (defined by inner functions) in the Hardy space \( H^p \) (coinvariant subspaces of the shift operator). We find a criterion for the embedding of \( K^p_\Theta \) in \( L^p(\mu) \) to be compact similar to the Volberg–Treil theorem on bounded embeddings, and give a positive answer to a question of Cima and Matheson. The proof is based on Bernstein-type inequalities for functions in \( K^p_\Theta \). We investigate measures \( \mu \) such that the embedding operator belongs to some Schatten–von Neumann ideal.

Keywords: Hardy space, inner function, embedding theorem, Carleson measure.

§ 1. Introduction and main results

Let \( \mathbb{D} = \{ z : |z| < 1 \} \) be the unit disc and \( \mathbb{T} = \{ z : |z| = 1 \} \) the unit circle. We denote the normalized Lebesgue measure on \( \mathbb{T} \) by \( m \). A bounded analytic function \( \Theta \) on \( \mathbb{D} \) is said to be inner if its non-tangential boundary values satisfy \( |\Theta| = 1 \) \( m \)-a.e. on \( \mathbb{T} \). We recall that every inner function \( \Theta \) admits a factorization

\[
\Theta(z) = B(z)I_\psi(z)
\]

(up to a constant factor of modulus 1). Here

\[
B(z) = \prod_n \frac{|z_n|}{z_n} \frac{z_n - z}{1 - z_n \bar{z}}, \quad z \in \mathbb{D},
\]

is the Blaschke product with zeros \( z_n \in \mathbb{D} \). The sequence \( \{ z_n \} \) satisfies the Blaschke condition \( \sum_n (1 - |z_n|) < \infty \) (we put \( |z_n|/z_n = 1 \) if \( z_n = 0 \)). The singular inner function \( I_\psi \) is defined by

\[
I_\psi(z) = \exp \left( - \int \frac{\zeta + z}{\zeta - z} d\psi(\zeta) \right), \quad z \in \mathbb{D},
\]

where \( \psi \) is a finite Borel measure on \( \mathbb{T} \) that is singular with respect to \( m \).

We denote the Hardy spaces on \( \mathbb{D} \) by \( H^p \), \( 1 \leq p \leq \infty \). Every inner function \( \Theta \) determines a subspace

\[
K^p_\Theta = H^p \cap \Theta \overline{H^p_0}
\]

This paper was written with financial support of the RFBR (grant no. 06-01-00313).

AMS 2000 Mathematics Subject Classification. 30D55, 47A45, 47B37.
of $H^p$, where $H^p_0 = \{ f \in H^p : f(0) = 0 \} = zH^p$. One can also define $K^p_{\Theta}$ as the set of all functions $f$ in $H^p$ such that $\langle f, \Theta g \rangle = \int_T f\Theta g \, dm = 0$ for all $g \in H^q$, where $1/p + 1/q = 1$.

We note that $K^2_{\Theta} = H^2 \ominus \Theta H^2$. It is known that if $1 \leq p < \infty$, then every closed subspace of $H^p$ which is invariant under the backward shift operator $(S^* f)(z) = \frac{f(z) - f(0)}{z}$ is given by $K^p_{\Theta}$ for some inner function $\Theta$ (see [1], Ch. II and [2]). The subspaces $K^p_{\Theta}$ are often referred to as *-invariant. They are important in operator theory and function theory (see [3]–[6]). In particular, they arise in the Sz.-Nagy–Foias model for contraction operators in a Hilbert space (this is why they are also called model subspaces). We note that if $\Theta$ is a Blaschke product, then $K^p_{\Theta}$ coincides with the closure in $H^p$ of the linear span of the partial fractions with poles of corresponding multiplicities at the points $1/z_n$.

We write $\sigma(\Theta)$ for the so-called spectrum of $\Theta$. This is the set of all $\zeta \in \overline{D}$ such that $\liminf_{z \to \zeta, z \in D} |\Theta(z)| = 0$. In other words, $\sigma(\Theta)$ is the smallest closed subset of $\overline{D}$ that contains the zeros $z_n$ and the support of $\psi$. It is well known that $\Theta$ (as well as any element of $K^p_{\Theta}$) admits analytic continuation across any arc lying in $\mathbb{T} \setminus \sigma(\Theta)$.

In this paper we consider the following problem. Given an inner function $\Theta$ and an exponent $p \geq 1$, describe the set of all Borel measures $\mu$ on the closed disc $\overline{D}$ such that the subspace $K^p_{\Theta}$ is embedded in $L^p(\mu)$ (or such that the embedding is compact). This problem was posed by Cohn [7] in 1982. It is still open despite a number of partial results. The embedding $K^p_{\Theta} \subset L^p(\mu)$ is equivalent to the estimate

$$\|f\|_{L^p(\mu)} \leq C\|f\|_p, \quad f \in K^p_{\Theta}. \quad (1.1)$$

We denote the set of measures $\mu$ with this property by $C_p(\Theta)$.

We recall that a finite Borel measure $\mu$ on the closed disc $\overline{D}$ is called a Carleson measure if there is a constant $M > 0$ such that

$$\mu(S(I)) \leq M|I| \quad (1.2)$$

for every arc $I \subset \mathbb{T}$. Here and in what follows $|I|$ is the length of $I$ and $S(I)$ is the Carleson square:

$$S(I) = \{ z = \rho e^{i\varphi} \in \overline{D} : e^{i\varphi} \in I, \ 1 - (2\pi)^{-1}|I| \leq \rho \leq 1 \}. \quad (1.3)$$

The set of Carleson measures is denoted by $\mathcal{C}$. Given a measure $\mu \in \mathcal{C}$, we write $M_\mu$ for the smallest constant $M$ in inequality (1.2). Carleson’s classical theorem says that $H^p \subset L^p(\mu)$ for some (any) $p > 0$ if and only if $\mu \in \mathcal{C}$ (see [1], [3]). The embedding $H^p \subset L^p(\mu)$ is compact if and only if $\mu$ is a vanishing Carleson measure, that is,

$$\lim_{|I| \to 0} \frac{\mu(S(I))}{|I|} = 0 \quad (1.4)$$

(see [8] and a generalization in [9]).

Clearly, $\mathcal{C} \subset C_p(\Theta)$. One might expect that the class $C_p(\Theta)$ depends essentially on the geometric properties of the function $\Theta$. A complete description of $C_p(\Theta)$ is
Embeddings of model subspaces 1079

currently known only for some special classes of inner functions. We say that $\Theta$ is a one-component inner function if the level set
\[
\Omega(\Theta, \varepsilon) = \{ z \in \mathbb{D} : |\Theta(z)| < \varepsilon \}
\]
is connected for some $\varepsilon \in (0,1)$. Cohn [7] proved that if $\Theta$ has this property, then it suffices to verify inequality (1.1) for the reproducing kernels of $K^2_{\Theta}$ (see the definition in §2). Nazarov and Volberg [10] have recently shown that this does not hold in the general case.

The following geometric condition on $\mu$ ensures the embeddability of $K^p_{\Theta}$ (Volberg and Treil [11]): we have $K^p_{\Theta} \subset L^p(\mu)$ if there is an $\varepsilon \in (0,1)$ such that $\mu(S(I)) \leq C|I|$ for all squares $S(I)$ satisfying $S(I) \cap \Omega(\Theta, \varepsilon) \neq \emptyset$.

Thus it suffices to verify Carleson’s condition (1.2) only for squares of a special form. Let $\mathcal{C}(\Theta)$ be the class of measures that satisfy the hypothesis of the Volberg–Treil theorem for some $\varepsilon \in (0,1)$. Aleksandrov [12] proved that the condition $\mu \in \mathcal{C}(\Theta)$ is necessary (that is, $\mathcal{C}_p(\Theta) = \mathcal{C}(\Theta)$) if and only if $\Theta$ is one-component. Moreover, if $\Theta$ is not one-component, then the class $\mathcal{C}_p(\Theta)$ depends essentially on the exponent $p > 0$ (in contrast to Carleson’s classical theorem). Some other embedding theorems have been obtained in [12]–[14]. The compactness of the embeddings was studied in [13], [15], [16].

Especially interesting is the case when $\mu = \sum_n a_n \delta_{\lambda_n}$ is a discrete measure. Then the embedding is equivalent to the Bessel property for the system $\{k_{\lambda_n}\}$ of reproducing kernels. Even the particular case when $\mu$ is a measure on the unit circle is of great interest. In contrast to embeddings of the whole Hardy space $H^p$ (we note that Carleson measures on $\mathbb{T}$ are measures of bounded density with respect to the Lebesgue measure $m$), the class $\mathcal{C}_p(\Theta)$ with $p \geq 1$ always contains non-trivial examples of singular measures on $\mathbb{T}$. In particular, if $p = 2$, then it contains Clark’s measures [17], for which the embedding $K^2_{\Theta} \subset L^2(\mu)$ is isometric. On the other hand, if $\mu = w m$, $w \in L^2(\mathbb{T})$, then the embedding problem turns out to be related to properties of the Toeplitz operator $T_w$ (see [13]).

A new approach to embedding theorems was suggested by the author [18], [19]. It is based on Bernstein inequalities for the spaces $K^p_{\Theta}$. By a Bernstein inequality we mean an estimate of the weighted norm of the derivative $f'$ in terms of the standard $L^p$-norm of $f \in K^p_{\Theta}$ in the space $L^p(\mathbb{T}, \mu)$. In other words, it is an estimate of the form
\[
\|f'\|_{L^p(\mu)} \leq C\|f\|_p, \quad f \in K^p_{\Theta},
\]
where $\mu$ is a measure on the closed disc $\overline{\mathbb{D}}$. This approach enables one to obtain essentially new embedding theorems that generalize the Volberg–Treil theorem and some results of Cohn. Another application of Bernstein inequalities concerns the stability of Bessel sequences and Riesz bases of reproducing kernels [20]. The results in [19], [20] are obtained for Hardy spaces in the upper half-plane. Their analogues for the disc are discussed in §3.
In this paper we use Bernstein inequalities for $K_p^\Theta$ to study whether the embedding operator is compact or belongs to Schatten–von Neumann ideals. One of the main results is a geometric condition (analogous to the Volberg–Treil theorem) guaranteeing that the embedding operator is compact. It now suffices to verify the ‘vanishing condition’ (1.4) only for squares that intersect the level set. For one-component functions, this condition also appears to be necessary.

**Theorem 1.1.** Suppose that $1 < p < \infty$, $\mu$ is a Borel measure on $\overline{D}$, and $\varepsilon \in (0, 1)$. Consider the following conditions.

(i) For every $\eta > 0$ there is a $\delta > 0$ such that $\mu(S(I))/|I| < \eta$ whenever $|I| < \delta$ and $S(I) \cap \Omega(\Theta, \varepsilon) \neq \emptyset$.

(ii) The embedding of $K_p^\Theta$ in $L^p(\mu)$ is compact.

Then condition (i) implies condition (ii). If the inner function $\Theta$ is one-component, then the converse holds: (ii) implies (i).

The implication (ii) $\implies$ (i) for one-component inner functions was proved in [16], where the question of the sufficiency of (i) was also raised. Theorem 1.1 gives a positive answer to this question. The paper [16] also suggests another ‘vanishing condition’ on $\mu$ which is sufficient for the embedding $K_p^\Theta \subset L^p(\mu)$ to be compact for all $p > 0$. We show in Proposition 4.1 that this condition implies condition (i) of Theorem 1.1 (this answers another question posed in [16]). We deduce Theorem 1.1 from a more general embedding theorem (Theorem 3.1), which generalizes the Volberg–Treil theorem.

In §§5, 6 we study whether the embedding operator $J_\mu: K_2^\Theta \to L^2(\mu)$, $J_\mu f = f$, belongs to the Schatten–von Neumann ideals $S_r$. We completely describe such measures for one-component inner functions $\Theta$ when $r \geq 1$. Given $\varepsilon \in (0, 1)$, we consider a Whitney-type decomposition of $T \setminus \sigma(\Theta)$ into a union of arcs $I_k$ with the property

$$\text{dist}(I_k, \Omega(\Theta, \varepsilon)) \asymp |I_k|$$

(see the detailed description in §3, Lemma 3.3). Then the following theorem holds.

**Theorem 1.2.** Let $\mu$ be a Borel measure such that $\text{supp} \mu \subset \bigcup_k S(I_k)$. Suppose that for some $r > 0$ we have

$$\mathcal{M}_r(\mu) = \sum_k \left( \frac{\mu(S(I_k))}{|I_k|} \right)^{r/2} < \infty. \quad (1.6)$$

Then $J_\mu \in S_r$ and $\|J_\mu\|_{S_r}^r \leq \mathcal{M}_r(\mu)$.

Let $R_{n,m}$ be the elements of the standard dyadic decomposition of the disc (see the definition in §5). We have the following necessary condition.

**Theorem 1.3.** Suppose that $\varepsilon \in (0, 1)$. If $J_\mu \in S_r$, $r \geq 1$, then

$$\sum_{R_{n,m} \cap \Omega(\Theta, \varepsilon) \neq \emptyset} (2^n \mu(R_{n,m}))^{r/2} < \infty. \quad (1.7)$$

The converses of Theorems 1.2 and 1.3 also hold for one-component inner functions. In this case we get a complete description of the embeddings of class $S_r$, $r \geq 1$. 

Theorem 1.4. Let $\Theta$ be a one-component inner function and $\mu$ a Borel measure on $D \cup (T \setminus \sigma(\Theta))$. The embedding operator $J_\mu$ belongs to $S_r$, $r \geq 1$, if and only if $\mu$ satisfies conditions (1.6) and (1.7) for every $\varepsilon \in (0,1)$.

Our conditions, which are stated in terms of the dyadic decomposition of the disc, are analogous to Luecking’s theorem [21] (which describes embeddings of class $S_r$ for the whole Hardy space) and to some results of Parfenov [22], [23]. We discuss these results further in §§ 5, 6.

As in [19], [20], essential use is made of the Bernstein inequalities for $K_\Theta^p$. For the sake of completeness we include a discussion of these results, which are of independent interest. In contrast to [19], we also give estimates for the higher derivatives. Let us state two such results. The first says that the growth of the derivatives at the boundary is controlled by the distance to the level set. For $\zeta \in T$ we put $d_\varepsilon(\zeta) = \text{dist}(\zeta, \Omega(\Theta, \varepsilon))$.

Theorem 1.5. Suppose that $\varepsilon \in (0,1)$, $1 < p < \infty$, $n \in \mathbb{N}$. Then

$$\| f(n) d_\varepsilon^n \|_p \leq C(p, n, \varepsilon) \| f \|_p, \quad f \in K_\Theta^p. \quad (1.8)$$

Let $\Theta$ be a one-component inner function. Then the boundary spectrum $\sigma(\Theta) \cap T$ has Lebesgue measure zero [24]. Thus the $n$th derivative $f^{(n)}(\zeta)$ is well defined for almost all $\zeta \in T$.

Theorem 1.6. Suppose that $\Theta$ is a one-component inner function, $1 < p < \infty$, $n \in \mathbb{N}$, $\mu \in C$. Then

$$\int_D |f^{(n)}(z)|^p \left( \frac{1 - |z|}{1 - |\Theta(z)|} \right)^{pn} d\mu(z) \leq C(\Theta, p, n, \mu) \| f \|_p^p, \quad f \in K_\Theta^p, \quad (1.8)$$

and, in particular,

$$\| f^{(n)}|\Theta'|^{-n} \|_p \leq C(\Theta, p, n) \| f \|_p, \quad f \in K_\Theta^p. \quad (1.9)$$

Theorems 1.5, 1.6 follow from a more general and somewhat more complicated Bernstein inequality, which is proved in § 2 (Theorem 2.1).

We shall use the following notation. Given non-negative functions $g$ and $h$, we write $g \lesssim h$ if $g \leq C h$ for some positive constant $C$ and all admissible values of the variables. We write $g \asymp h$ if $g \lesssim h \lesssim g$. The letters $C, C_1, \ldots$ stand for various positive constants that may be different in different formulas.

The author is grateful to A. B. Aleksandrov for some useful discussions and remarks.

§ 2. Bernstein inequalities for higher derivatives

We assume throughout that $p \in [1, \infty)$ and $q$ is the conjugate exponent: $1/p + 1/q = 1$. The symbol $L^p$ stands for the standard space $L^p(T, m)$.

We begin with discussing the local behaviour of the elements of a model subspace and their derivatives near the boundary. An important role is played by reproducing
kernels. The reproducing kernel (of the space $K^2_\Theta$) corresponding to a point $z \in \mathbb{D}$ is given by

$$k_z(\zeta) = \frac{1 - \overline{\Theta(z)}\Theta(\zeta)}{1 - z\overline{\zeta}}.$$  

Since $k_z \in K^\infty_\Theta$, we have

$$f(z) = \int_{\mathbb{T}} f(\tau) k_z(\tau) \, dm(\tau), \quad f \in K^p_\Theta \quad (2.1)$$

for all $1 \leq p \leq \infty$. There is a similar representation for the $n$th derivative:

$$f^{(n)}(z) = n! \int_{\mathbb{T}} \tau^n f(\tau) (k_z(\tau))^{n+1} \, dm(\tau), \quad f \in K^p_\Theta. \quad (2.2)$$

This follows from the inclusion $(1 - \tau z)^{n+1} - (k_z(\tau))^{n+1} \in \Theta H^\infty$.

The integral representations (2.1), (2.2) can be extended to some points $z = \zeta$ of $\mathbb{T}$. It is well known that every function in $K^p_\Theta$ admits analytic continuation across any arc lying in $\mathbb{T} \setminus \sigma(\Theta)$ and, therefore, equations (2.1), (2.2) hold for all $z = \zeta \in \mathbb{T} \setminus \sigma(\Theta)$. The boundary behaviour at a point of the spectrum depends on the ‘density’ of the spectrum near that point, and the answer depends on the ‘density’ of the spectrum near that point. For $z = \zeta \in \mathbb{T}$ we put

$$S_q(\zeta) = \sum_n \frac{1 - |z_n|^2}{|\zeta - z_n|^q} + \int_{\mathbb{T}} \frac{d\psi(\tau)}{|\zeta - \tau|^q}.$$  

Then the results of Ahern and Clark [25] and Cohn [26] yield that the derivative $f^{(n)}(\zeta)$ (understood as a non-tangential limit) exists for every $f \in K^p_\Theta$ if and only if $S_{(n+1)q}(\zeta) < \infty$. In this case we have $k^{n+1}_\zeta \in K^q_\Theta$ and equation (2.2) holds for $z = \zeta$. We are especially interested in the number $S_2$, which describes the inclusion $k_\zeta \in K^2_\Theta$. If $\Theta$ has a non-tangential limit at $\zeta$ and $\Theta(\zeta) \in \mathbb{T}$, then $S_2$ equals the modulus of the angular derivative of $\Theta$ at $\zeta$ (the angular derivative is defined as $\lim_{z \to \zeta} \frac{\Theta(z) - \Theta(\zeta)}{z - \zeta}$, where $z$ tends to $\zeta$ non-tangentially):

$$|\Theta'(\zeta)| = \sum_n \frac{1 - |z_n|^2}{|\zeta - z_n|^2} + \int_{\mathbb{T}} \frac{d\psi(\tau)}{|\zeta - \tau|^2}.$$  

The main result of this section is the following Bernstein inequality of the form (1.5), which holds for any inner function $\Theta$ and for measures of the form $w\mu$, where $\mu$ is a Carleson measure and the weight $w(z)$ depends on the norm of $k^{n+1}_z$ in $L^q$ (and thus, in fact, on the norm of the functional $f \mapsto f^{(n)}(z)$, $f \in K^p_\Theta$).

We put

$$w_{p,n}(z) = \|k^{n+1}_z\|_q^{-\frac{p}{q}+\frac{n}{q}}.$$  

We put $\|k^{n+1}_\zeta\|_q = \infty$ and $w_{p,n}(\zeta) = 0$ whenever $\zeta \in \mathbb{T}$ and $S_{(n+1)q}(\zeta) = \infty$. Then the product $f^{(n)}(z)w_{p,n}(z)$ is well defined for all $f \in K^p_\Theta$, $z \in \mathbb{D}$. 
Theorem 2.1. Suppose that $\mu \in \mathcal{C}$, $1 \leq p < \infty$, $n \in \mathbb{N}$. Then the operator 

$$(T_{p,n}f)(z) = f^{(n)}(z)w_{p,n}(z)$$

has weak type $(p, p)$ as an operator from $K^p_{\Theta}$ to $L^p(\mu)$ and is bounded as an operator from $K^p_{\Theta}$ to $L^r(\mu)$ for every $r > p$. Moreover, there is a constant $C = C(M_{\mu}, p, r, n)$ such that 

$$\|f^{(n)}w_{p,n}\|_{L^r(\mu)} \leq C\|f\|_r, \quad f \in K^r_{\Theta}. \quad (2.3)$$

To apply Theorem 2.1, we need effective estimates for the corresponding weights (that is, for the norms of reproducing kernels). When $p = 2$, there is an explicit formula 

$$\|k_z\|^2 = \frac{1 - |\Theta(z)|^2}{1 - |z|^2}, \quad z \in \mathbb{D},$$

and $\|k_{\zeta}\|^2 = |\Theta'(\zeta)|$ for $\zeta \in \mathbb{T}$. Sharp inequalities for the norms of reproducing kernels are known in the case when $\Theta$ is one-component (see [12], inequality (15)). Moreover, one can relate the weight $w_{p,n}$ to geometric properties of the level sets of $\Theta$: 

$$d^n_{\zeta}(\zeta) \lesssim w_{p,n}(\zeta) \lesssim |\Theta'(\zeta)|^{-n}, \quad \zeta \in \mathbb{T}. \quad (2.4)$$

Proofs of these estimates are contained in Lemma 4.5 of [19], where the corresponding inequality is obtained for $n = 1$ (the argument obviously works for $n \geq 2$). Similar results can be found in [12].

The proof of Theorem 2.1 is based on the integral representation (2.2), which reduces the study of the differentiation operator to the study of certain singular integral operators. We shall deduce inequality (2.3) from the boundedness of the following integral operators on the $L^p$-spaces related to Carleson measures ([19], Theorems 3.1, 3.2).

Theorem 2.2. Suppose that $\mu \in \mathcal{C}$, $h$ is a non-negative function on $\overline{\mathbb{D}}$, $h$ is measurable with respect to $\mu$ and $m$, and there is a constant $A > 0$ such that $h(z) \geq A(1 - |z|)$, $z \in \mathbb{D}$. Put

$$Tf(z) = h(z)\int_{|\zeta - z| > h(z)} \frac{f(\zeta)}{|\zeta - z|^2} dm(\zeta), \quad z \in \overline{\mathbb{D}}.$$ 

Then $T$ has weak type $(1, 1)$ as an operator from $L^1$ to $L^1(\mu)$ and is bounded as an operator from $L^r$ to $L^r(\mu)$ for all $r > 1$. The norm of $T$ does not exceed a constant $C$ depending only on $r$, $A$ and the Carleson constant $M_{\mu}$ of $\mu$. 

We also consider a class of integral operators with ‘diagonal’ kernels. Suppose that $\mu, \nu \in \mathcal{C}$, and let $K(z, u)$ be a $(\mu \times \nu)$-measurable function. For $z \in \overline{\mathbb{D}}$ we put 

$$\Delta_z(p) = \left\{ u \in \overline{\mathbb{D}} : \|u - z\| < \|K(z, \cdot)\|^{-p}_{L^p(\nu)} \right\}$$

(there is no loss of generality in assuming that $K(z, \cdot)$ is $\nu$-measurable for all $z$, and we put $\|K(z, \cdot)\|^{-p}_{L^p(\nu)} = 0$ if $K(z, \cdot) \notin L^p(\nu)$). Consider the following ‘truncation’ of the integral operator with kernel $K$: 

$$T_pf(z) = \int_{\Delta_z(p)} K(z, u)f(u) d\nu(u).$$
Theorem 2.3. If \( \|K(z, \cdot)\|^{-p}_{L^q(\nu)} \geq A(1 - |z|) \), then \( T_p \) has weak type \((p, p)\) as an operator from \( L^p(\nu) \) to \( L^p(\mu) \) and is bounded as an operator from \( L^r(\nu) \) to \( L^r(\mu) \) for all \( r > p \).

Detailed proofs of Theorems 2.2 and 2.3 are given in [19] for the half-plane case. The proofs for the disc are exactly the same, and so we omit them.

The idea behind the proof of Theorem 2.1 is to divide the integral (that represents the derivative) into ‘diagonal’ and ‘off-diagonal’ parts and estimate them separately using Theorems 2.2 and 2.3 respectively.

Proof of Theorem 2.1. Put \( h(z) = (w_{p, n}(z))^{1/n} \). The obvious inequality \( |k_z(w)| \leq 2(1 - |z|)^{-1}, \) \( z, w \in \mathbb{D} \), yields that \( h(z) \geq A(1 - |z|) \). We multiply the integral in (2.2) by \( w_{p, n}(z) \) and divide it into two parts:

\[
\frac{1}{n!} w_{p, n}(z) f^{(n)}(z) = I_1 f(z) + I_2 f(z),
\]

where

\[
I_1 f(z) = w_{p, n}(z) \int_{|\zeta - z| \geq h(z)} \overline{\zeta}^n f(\zeta) \overline{k_z^2(z)}^{n+1}(\zeta) \, dm(\zeta),
\]

\[
I_2 f(z) = w_{p, n}(z) \int_{|\zeta - z| < h(z)} \overline{\zeta}^n f(\zeta) \overline{k_z^2(z)}^{n+1}(\zeta) \, dm(\zeta).
\]

Since \( |1 - \zeta| = |\zeta - z| \) for \( \zeta \in \mathbb{T} \), we have

\[
|I_1 f(z)| \leq Ch^n(z) \int_{|\zeta - z| \geq h(z)} \frac{|f(\zeta)|}{|\zeta - z|^{n+1}} \, dm(\zeta) \leq Ch(z) \int_{|\zeta - z| \geq h(z)} \frac{|f(\zeta)|}{|\zeta - z|^2} \, dm(\zeta).
\]

Theorem 2.2 now yields that \( I_1 \) is bounded as an operator from \( L^r \) to \( L^r(\mu) \) for every Carleson measure \( \mu \) and \( r > 1 \).

To estimate \( I_2 f \), we put \( K(z, \zeta) = (h(z))^n k_z^2(z)^{n+1}(\zeta) \). Then \( \|K(z, \cdot)\|^{-p}_q = (h(z))^{-pn} \|k_z^{n+1}\|^{-p}_q = h(z) \). Thus,

\[
I_2 f(z) = \int_{|\zeta - z| \leq \|K(z, \cdot)\|_q^{-1}} \overline{\zeta}^n f(\zeta) K(z, \zeta) \, dm(\zeta).
\]

Applying Theorem 2.3, we conclude that \( I_2 \) has weak type \((p, p)\) as an operator from \( L^p \) to \( L^p(\mu) \) and is bounded as an operator from \( L^r \) to \( L^r(\mu) \) for every measure \( \mu \in \mathcal{C} \) and \( r > p \).

Proof of Theorem 1.5. By inequality (2.4), the desired assertion follows from Theorem 2.1 with \( \mu = m \).

Proof of Theorem 1.6. It is shown in [12] that for one-component \( \Theta \) we have

\[
\|k_z\|_s^p \asymp \|k_z\|_2^{2(s-1)}, \quad z \in \mathbb{D},
\]

where the constants depend on \( \Theta \) and \( s \in (1, \infty) \), but are independent of \( z \). Therefore,

\[
w_{p, n}(z) \asymp \left( \|k_z\|_2^{2(q(n+1)-1)} \right)^{-\frac{pn}{(p+1)(n+1)}} = \|k_z\|_2^{-2n} = \left( \frac{1 - |z|^2}{1 - |\Theta(z)|^2} \right)^n, \quad z \in \mathbb{D},
\]

and \( w_{p, n}(\zeta) \asymp |\Theta(\zeta)|^{-n}, \) \( \zeta \in \mathbb{T} \). The theorem is proved.
Remark 2.1. Inequality (1.9) in Theorem 1.6 should be compared to the Bernstein inequality for \( L^\infty \)-norms: if \( \Theta \) has a non-tangential limit at a point \( \zeta \in \mathbb{T} \), \( \Theta(\zeta) \in \mathbb{T} \) and \( |\Theta'(\zeta)| < \infty \), then the derivative \( f'(\zeta) \) of any function \( f \in K_\Theta^\infty \) exists as a non-tangential limit and
\[
\frac{|f'(\zeta)|}{|\Theta'(\zeta)|} \leq \|f\|_\infty. \tag{2.7}
\]
Indeed, we have
\[
f'(\zeta) = \int_{\mathbb{T}} \tau f(\tau) k_\Theta^2(\tau) d\tau(\tau),
\]
whence
\[
|f'(\zeta)| \leq \|f\|_\infty \|k_\Theta\|_2^2 = \|f\|_\infty |\Theta'(\zeta)|. \tag{2.7}
\]
Note that inequality (2.7) holds for any (not necessarily one-component) inner function and the constant 1 is sharp. This inequality is due to Levin [27]. It was later rediscovered by several authors (see, for example, [28]) for finite Blaschke products.

Remark 2.2. Example 5.2 in [19] shows that the exponent \( \frac{p}{p+1} \) in the definition of \( w_{p,1} \) is precise in some sense.

Remark 2.3. Bernstein inequalities for model subspaces \((K^p_\theta)_+\) in the upper half-plane \( \mathbb{C}^+ \) were studied by Dyakonov [29], [30], who showed that differentiation is a bounded operator from \((K^p_\theta)_+\) to \( L^p(\mathbb{R}) \) for \( 1 < p \leq \infty \), that is,
\[
\|f'\|_p \leq C(p,\theta) \|f\|_p, \quad f \in (K^p_\theta)_+ \tag{2.8}
\]
if and only if \( \theta' \in H^\infty(\mathbb{C}^+) \). In this case \( \theta \) is meromorphic on the whole complex plane and the subspaces \((K^p_\theta)_+\) are closely related to certain spaces of entire functions (in particular, to de Branges spaces if \( p = 2 \); see Proposition 1.1 of [31]). Weighted Bernstein inequalities of the form (1.5) (which were obtained in [18], [19]) essentially generalize (2.8). Weighted estimates have an advantage: the weight may compensate the possible growth of elements of \((K^p_\theta)_+\) and their derivatives near the boundary. We note that the Bernstein inequality for the standard \( L^p \)-norms (that is, \( \|f'\|_p \leq C \|f\|_p \)) holds for the model subspace \( K^p_\theta \) in the disc if and only if it is finite-dimensional (that is, \( \Theta \) is a finite Blaschke product). Hence the idea of using weighted Bernstein inequalities (with a ‘correcting’ weight) becomes even more natural when dealing with spaces in the disc.

§ 3. Embedding theorems

Bernstein inequalities for model subspaces turn out to be a convenient tool for proving new embedding theorems. This enables us to give new proofs and essential generalizations of nearly all known embedding results. Thus we shall see that Theorem 2.1 yields an embedding theorem that generalizes the Volberg–Treil theorem. This result also yields a sufficient condition for the embedding to be compact. We assume throughout that \( \mu \) is a finite Borel measure on the closed disc \( \overline{D} \).

By a square with side of length \( h \) in the unit disc we mean a set of the form
\[
S(h_0, \phi_0, h) = \left\{ \rho e^{i\phi} : h_0 - \frac{h}{2\pi} \leq \rho \leq h_0, \ 0 \leq \phi \leq \phi_0 + h \right\},
\]
where \( h_0 \in (0, 1], \phi_0 \in \mathbb{R} \) and \( 0 < h < 2\pi h_0 \). We write \( J(S) \) for the outer side of the square \( S \), that is, \( J(S) = \{h_0 e^{i\phi} : \phi_0 \leq \phi \leq \phi_0 + h \} \).
We note that particular cases of this definition are Carleson’s squares (1.3) (they arise when \( h_0 = 1 \)) and the dyadic squares (5.1) introduced below.

Let \( \{S_k\}_{k \in \mathbb{N}} \) be a sequence of squares in \( \mathbb{D} \), \( J_k \) the outer side of \( S_k \), and \( \delta J_k \) the Lebesgue measure on the arc \( J_k \). Suppose that \( 1 < r < p \) and the squares \( S_k \) satisfy two conditions,

\[
\sum_k \delta J_k \in C, \tag{3.1}
\]

\[
\sup_k |J_k| \|w_r^{-1}\|_{L^q(J_k)}^p < \infty, \tag{3.2}
\]

where \( w_r(z) = w_{r,1}(z) = \|k_z^2\|_{r \left( r' \right)} \) is the weight in the Bernstein inequality (see §2), \( 1/r + 1/r' = 1 \). Condition (3.1) means that the sequence \( \{S_k\} \) of squares is sufficiently rarefied, and inequality (3.2) controls their size.

**Theorem 3.1.** Let \( \{S_k\}_{k \in \mathbb{N}} \) be a sequence of squares satisfying (3.1) and (3.2), and let \( \mu \) be a Borel measure on \( \bigcup_k S_k \).

(i) If \( \mu(S_k) \leq C|J_k| \), then \( \mu \in C_p(\Theta) \).

(ii) If, moreover, \( \mu(S_k) = o(|J_k|) \) as \( k \to \infty \), then the embedding \( K^p_\Theta \subset L^p(\mu) \) is compact.

Note that here, as in the Volberg–Treil theorem, we are considering measures with the Carleson property for a special class of rather large squares. It will be seen that the squares in Theorem 3.1 may be considerably larger (see Proposition 3.1 below).

To prove Theorem 3.1 we shall use the following lemmas. The first says that the norms of reproducing kernels are in some sense monotone along the radii.

**Lemma 3.1.** Suppose that \( q > 1 \). Then there is a constant \( C = C(q) \) such that the following estimate holds for all \( z = \rho e^{i\phi} \) and \( \tilde{z} = \tilde{\rho} e^{i\phi} \), \( 0 \leq \tilde{\rho} \leq \rho \):

\[
\|k_{\tilde{z}}\|_q \leq C(q)\|k_z\|_q. \tag{3.3}
\]

**Proof.** This property is established for the half-plane in [19], Corollary 4.7. The same argument proves the assertion for the disc.

If a sequence \( \{S_k\} \) satisfies (3.2), then the estimate (3.3) yields that

\[
\sup_k |J_k| \left( \int_{S_k \cap \{|z| = \rho\}} w_{r,1}^{-q}(z) |dz| \right)^{p/q} \leq C \tag{3.4}
\]

for all \( \rho \in (0, 1] \) (where \(|dz|\) is the Lebesgue measure on the corresponding arc).

**Lemma 3.2.** If \( J_k \subset \mathbb{T} \), then it follows from condition (3.2) that the integral \( \int_{J_k} |(\Theta')| \|dm(\tau)\| \) is finite. In particular, \( \text{Int} J_k \cap \sigma(\Theta) = \emptyset \) (here \( \text{Int} J_k \) is the relative interior of the arc \( J_k \) in \( \mathbb{T} \)) and the function \( \Theta \) is continuous on every (closed) square \( S_k \).

**Proof.** By inequalities (2.4) and (3.2) we have

\[
\int_{J_k} |(\Theta')|^q \|dm(\tau)\| < \infty.
\]
Hence \( \int_{J_k} |\Theta'(\tau)| \, dm(\tau) < \infty \), and we conclude that \( \Theta \) is continuous on \( J_k \). It is easily seen that \( |\Theta'(r\zeta)| \leq C|\Theta'(\zeta)| \), \( \zeta \in \mathbb{T}, r \in (0, 1) \) and, therefore, \( \Theta \) is continuous on \( S_k \). The lemma is proved.

**Proof of Theorem 3.1.** (i) Clearly, the embedding \( K^p_{\Theta} \subset L^p(\mu_{\{|z| < 1/2\}}) \) is compact. Hence there is no loss of generality in assuming that \( \text{supp} \mu \subset \{1/2 \leq |z| \leq 1\} \).

It follows from Lemma 3.2 that those functions \( f \in K^p_{\Theta} \) which are continuous on every \( S_k \) form a dense subset of \( K^p_{\Theta} \), \( 1 < p < \infty \) (it suffices to consider the reproducing kernels). Thus it suffices to prove the inequality

\[
\|f\|_{L^p(\mu)} \leq C\|f\|_p, \quad f \in K^p_{\Theta},
\]

only for functions continuous on \( \bigcup_k S_k \).

Suppose that \( f \in K^p_{\Theta} \) is continuous on every \( S_k \). Then there are \( w_k \in S_k \) such that

\[
\|f\|_{L^p(\mu)}^p \leq \sum_k |f(w_k)|^p \mu(S_k) \leq \sup_k \frac{\mu(S_k)}{|J_k|} \sum_k |f(w_k)|^p |J_k|. \tag{3.5}
\]

Part (i) will be proved if we can show that

\[
\sum_k |f(w_k)|^p |J_k| \leq C\|f\|_p^p, \tag{3.6}
\]

where \( C \) is independent of \( f \) and of the choice of the points \( w_k \in S_k \).

Consider the arcs \( \tilde{J}_k = S_k \cap \{|z| = |w_k|\} \). Since \( \mu(\{|z| < 1/2\}) = 0 \), we have \( \tilde{J}_k \supset |J_k|/2 \). Put \( \nu = \sum_k \delta_{\tilde{J}_k} \). Then condition (3.1) yields that \( \nu \in \mathcal{C} \) (and the Carleson constants \( M_\nu \) of such measures \( \nu \) are uniformly bounded). We have

\[
\left( \sum_k |f(w_k)|^p |\tilde{J}_k| \right)^{1/p} \leq \|f\|_{L^p(\nu)} + \left( \sum_k \int_{J_k} |f(z) - f(w_k)|^p |dz| \right)^{1/p}, \tag{3.7}
\]

and \( \|f\|_{L^p(\nu)} \leq C_1 \|f\|_p \).

Let us estimate the last term in (3.7). Given \( z \in \tilde{J}_k \), we write \( \gamma(z, w_k) \) for the arc with endpoints \( z \) and \( w_k \) that lies inside the arc \( \tilde{J}_k \). Then \( f(z) - f(w_k) = \int_{\gamma(z, w_k)} f'(u) \, du \) (if \( J_k \subset \mathbb{T} \), it follows from Lemma 3.2 that every function \( f \in K^p_{\Theta} \) is analytic on \( J_k \), except possibly at the endpoints) and, therefore,

\[
\sum_k \int_{\tilde{J}_k} |f(z) - f(w_k)|^p |dz| = \sum_k \int_{\tilde{J}_k} \left| \int_{\gamma(z, w_k)} f'(u) \, du \right|^p |dz|
\]

\[
\leq \sum_k \int_{\tilde{J}_k} \left( \int_{\gamma(z, w_k)} w_r^{-q}(u) |du| \right)^{p/q} \left( \int_{\gamma(z, w_k)} |f'(u)|^p w_p^p(u) |du| \right)^{1/q} |dz|
\]

\[
\leq \sum_k |\tilde{J}_k| \left( \int_{J_k} w_r^{-q}(u) |du| \right)^{p/q} \left( \int_{J_k} |f'(u)|^p w_p^p(u) |du| \right).\]

Using inequality (3.4), we get

\[
\sum_k \int_{\tilde{J}_k} |f(z) - f(w_k)|^p |dz| \leq C_2 \sum_k \int_{\tilde{J}_k} |f'(u)|^p w_p^p(u) |du|
\]

\[
= C_2 \|f' w_r^p\|_{L^p(\nu)} \leq C_3 \|f\|_p^p,
\]

where the last inequality follows from Theorem 2.1.
(ii) Given a Borel set \( E \subset \overline{D} \), we define an operator \( \mathcal{I}_E : K^p_\Theta \to L^p(\mu) \) by putting \( \mathcal{I}_E f = \chi_E f \), where \( \chi_E \) is the characteristic function of \( E \). For \( N \in \mathbb{N} \) we put \( F_N = \bigcup_{k=1}^N S_k \). As above, we assume that the function \( f \in K^p_\Theta \) is continuous on \( \bigcup_k S_k \). Then the estimates (3.5) and (3.6) yield that
\[
\int_{\mathbb{D}\setminus F_N} |f|^p \, d\mu \leq C \sup_{k>N} \frac{\mu(S_k)}{|J_k|} \|f\|_p^p.
\]
Hence \( \|\mathcal{I}_{\overline{D}\setminus F_N}\| \to 0 \) as \( N \to \infty \). Part (ii) will be proved if we can show that the operator \( \mathcal{I}_{F_N} \) is compact for every \( N \) (thus the embedding operator \( J = \mathcal{I}_{\overline{D}} = \mathcal{I}_{F_N} + \mathcal{I}_{\overline{D}\setminus F_N} \) can be approximated in the operator norm by the compact operators \( \mathcal{I}_{F_N} \)). Clearly, it suffices to prove that the operator \( \mathcal{I}_{S_k} \) is compact for every fixed \( k \).

Let us approximate \( \mathcal{I}_{S_k} \) by finite-rank operators. To do this, we represent the square \( S_k \) as a finite union \( \{S_l\}_{l=1}^L \) of squares with disjoint interiors and fix an arbitrary point \( \zeta_l \) in each \( S_l \). By Lemma 3.1, for any \( \epsilon > 0 \) one can choose the squares \( S_l \) to be so small that
\[
\left( \int_{[\zeta_l, z_l]} w_r^{-q}(z) \, |dz| \right)^{p/q} < \epsilon
\]
for all \( l, 1 \leq l \leq L \), and all \( z_l \in S_l \). Here \([z, w]\) stands for the straight line interval with endpoints \( z \) and \( w \).

Now consider the finite-rank operator
\[
T : K^p_\Theta \to L^p(\mu), \quad (Tf)(z) = \sum_{l=1}^L f(\zeta_l) \chi_{S_l}(z).
\]
We claim that \( \|\mathcal{I}_{S_k} - T\|^p \leq C\epsilon \). As in the proof of part (i), we have
\[
\|(\mathcal{I}_{S_k} - T)f\|_{L^p(\mu)}^p = \sum_{l=1}^L \int_{S_l} |f(z) - f(\zeta_l)|^p \, d\mu(z)
\]
\[
\leq \sum_{l=1}^L \int_{S_l} \left( \int_{[\zeta_l, z]} |f'(u)|^p w_r^p(u) \, |du| \right) \left( \int_{[\zeta_l, z]} w_r^{-q}(u) \, |du| \right)^{p/q} \, d\mu(z).
\]
By Theorem 2.1 we get
\[
\int_{[\zeta_l, z]} |f'(u)|^p w_r^p(u) \, |du| \leq C_1 \|f\|_p^p,
\]
where \( C_1 \) is independent of \( f \in K^p_\Theta \), \( 1 \leq l \leq L \) and \( z \in S_l \). Hence, by inequality (3.8), we have
\[
\|(\mathcal{I}_{S_k} - T)f\|_{L^p(\mu)}^p \leq C_1 \epsilon \|f\|_p^p \sum_{l=1}^L \mu(S_l) = C_1 \epsilon \mu(S_k) \|f\|_p^p.
\]
We conclude that the operator \( \mathcal{I}_{S_k} \) is approximated by finite-rank operators and, therefore, is compact.
Remark 3.1. The proof of Theorem 3.1 uses the density in $K^p_	heta$ of the set of functions that are continuous on every square $S_k$ (Lemma 3.2). The results of [2] actually imply that even functions continuous on the closed disc $\mathbb{D}$ are dense in $K^p_	heta$, $p \geq 1$.

Theorem 3.1 describes a wider class of measures than the class $\mathcal{C}(\Theta)$ in the Volberg–Treil theorem. Namely, the following proposition holds.

**Proposition 3.1.** If $\mu \in \mathcal{C}(\Theta)$, then $\mu = \mu_1 + \mu_2$, where $\mu_1$ satisfies the hypotheses of Theorem 3.1, (i) for all $p > 1$ and $r \in (1, p)$ while $\mu_2 \in \mathcal{C}$.

We shall use the following special family of arcs on $T$ (easily seen to be analogous to the Whitney decomposition of $\mathbb{D} \setminus \Omega(\Theta, \varepsilon)$).

**Lemma 3.3.** Take $\varepsilon \in (0, 1)$. Suppose that $T \setminus \sigma(\Theta) \neq \emptyset$. Then there is a sequence $I_k \subset T$, $k \in \mathbb{N}$, of arcs with disjoint interiors such that $\bigcup_k I_k = T \setminus \sigma(\Theta)$ and

\[ |I_k| \leq \text{dist}(I_k, \Omega(\Theta, \varepsilon)) \leq 2|I_k|. \tag{3.9} \]

Moreover, if we put $F = \bigcup_k S(I_k)$ and $G = \mathbb{D} \setminus F$, then the following estimate holds for all $z \in G$, $z \neq 0$:

\[ \text{dist}\left(\frac{z}{|z|}, \Omega(\Theta, \varepsilon)\right) \leq 6\pi(1 - |z|). \tag{3.10} \]

**Proof.** Note that $\int_{T \setminus \sigma(\Theta)} d^{-1}_\varepsilon(\zeta) \, dm(\zeta) = \infty$. Therefore we can choose a sequence of arcs $I_k$ with pairwise-disjoint interiors such that $\bigcup_k I_k = T \setminus \sigma(\Theta)$ and

\[ \int_{I_k} d^{-1}_\varepsilon(\zeta) \, dm(\zeta) = \frac{1}{2}. \]

Hence there is a $\zeta_k \in I_k$ such that $d_\varepsilon(\zeta_k) = 2|I_k|$. Then $d_\varepsilon(\zeta) \geq d_\varepsilon(\zeta_k) - |I_k| \geq |I_k|$ for all $\zeta \in I_k$ and we obtain inequality (3.9).

Now suppose that $z = re^{i\phi} \in G$. Then either $e^{i\phi} \in \sigma(\Theta)$ (and hence $\text{dist}(e^{i\phi}, \Omega(\Theta, \varepsilon)) \leq 1 - |z|$), or $e^{i\phi} \in I_k$ for some $k$. Since $z \notin S(I_k)$, we have $1 - r \geq |I_k|/(2\pi)$, and inequality (3.9) yields that $\text{dist}(e^{i\phi}, \Omega(\Theta, \varepsilon)) \leq 3|I_k| \leq 6\pi(1 - r)$. The lemma is proved.

**Proof of Proposition 3.1.** As usual, given an arc $I \subset T$ and $a > 0$, we write $aI$ for the arc of length $a|I|$ with the same centre. We put $\mu_1 = \mu|_F$ and $\mu_2 = \mu|_G$, where the sets $F$, $G$ are defined in the statement of Lemma 3.3. It follows from (3.9) that

\[ |I_k| \left( \int_{I_k} d^{-q}_\varepsilon(\zeta) \, dm(\zeta) \right)^{p/q} \leq C. \]

Take $r \in (1, p)$. By inequality (2.4) we have $d_\varepsilon \lesssim w_r$ and, therefore,

\[ |I_k| \left( \int_{I_k} w^{-q}_r(\zeta) \, dm(\zeta) \right)^{p/q} \leq C_1. \tag{3.11} \]

Thus the family of squares $S(I_k)$ satisfies conditions (3.1) and (3.2). We claim that $\mu_1(S(I_k)) \lesssim C_2|I_k|$. Indeed, (3.9) yields that $S(AI_k) \cap \Omega(\Theta, \varepsilon) \neq \emptyset$ for a sufficiently
large absolute constant $A$. Since $\mu \in \mathcal{C}(\Theta)$, we see that $\mu_1(S(I_k)) \leq \mu(S(AI_k)) \leq C_2|I_k|$.

Let us show that $\mu_2$ is a standard Carleson measure. If $S(I) \cap G \neq \emptyset$ for some arc $I \subset \mathbb{T}$, then inequality (3.10) yields a constant $A_1 > 1$ such that

$$S(A_1I) \cap \Omega(\Theta, \varepsilon) \neq \emptyset.$$  \hspace{1cm} (3.12)

Hence $\mu(S(I)) \leq C_3|I|$ for some positive constant $C_3$. The proposition is proved.

The following example shows that Theorem 3.1 describes a much wider class of measures than that in the Volberg–Treil theorem.

**Example 3.1.** By Proposition 3.1, every measure $\mu \in \mathcal{C}(\Theta)$ takes the form $\mu = \mu_1 + \mu_2$, where $\mu_1$ satisfies hypothesis (i) of Theorem 3.1 and $\mu_2$ is a standard Carleson measure. Thus the Volberg–Treil theorem follows from Theorem 3.1. On the other hand, one can easily construct a measure $\mu$ that satisfies the hypotheses of Theorem 3.1, (i) but does not belong to $\mathcal{C}(\Theta)$.

For example, if $\mu \in \mathcal{C}(\Theta)$, then $\mu$ has no point weights at the boundary spectral points: $\mu(\{\zeta\}) = 0$ for all $\zeta \in \sigma(\Theta) \cap \mathbb{T}$ (note that then we have $S(I) \cap \Omega(\Theta, \varepsilon) \neq \emptyset$ for every $\varepsilon \in (0, 1)$ and any arc $I$ having $\zeta$ as an interior point). At the same time, the measures in Theorem 3.1 may admit non-trivial point weights on $\sigma(\Theta) \cap \mathbb{T}$. Let $B$ be the Blaschke product with zeros $z_n = (1 - 2^{-n})e^{i/n}$, $n \in \mathbb{N}$. Then for every $p \in (1, \infty)$ we have

$$\|k_2^1\|_q \leq C_1, \quad -\pi < \arg \zeta \leq 0,$$

whence $w_p^{-1}(\zeta) = \|k_2^1\|_q^{p/(p+1)} \leq C_2$ for $-\pi < \arg \zeta \leq 0$. Thus $\delta_1 \in \mathcal{C}_p(\Theta)$. One can similarly construct an infinite sum of point weights (that is, $\mu = \sum_n a_n \delta_{\zeta_n}$ with $a_n > 0$ and $\zeta_n \in \sigma(\Theta) \cap \mathbb{T}$) such that the embedding $K^p_\Theta \subset L^p(\mu)$ is bounded or even compact (see Example 6.3 of [19] for details).

---

**§ 4. Compact embeddings. Proof of Theorem 1.1**

In this section we prove Theorem 1.1 and discuss the relationship between the two ‘vanishing conditions’ introduced in [16].

*Proof of Theorem 1.1.* As already noted, the implication (ii) $\implies$ (i) for one-component inner functions was proved in [16]. Consider the implication (i) $\implies$ (ii), whose proof is similar to that of Proposition 3.1. Take an $\varepsilon \in (0, 1)$ and let $I_k$, $F$, $G$, $\mu_1$, $\mu_2$ have the same meaning as in Proposition 3.1 and Lemma 3.3. We claim that $\mu_1$ satisfies hypothesis (ii) of Theorem 3.1 while $\mu_2$ is a vanishing Carleson measure (see the definition in § 2) and, therefore, the embedding $H^p \subset L^p(\mu_2)$ is compact.

Indeed, we have the estimate (3.11) for all $p \in (1, \infty)$ and $r \in (1, p)$. Since $|I_k| \to 0$ as $k \to \infty$ and we have $S(AI_k) \cap \Omega(\Theta, \varepsilon) \neq \emptyset$ for a sufficiently large absolute constant $A > 1$, it follows from the hypotheses of the theorem that

$$\lim_{k \to \infty} \frac{\mu(S(I_k))}{|I_k|} = 0.$$
Hence the embedding $K^p_\Theta \subset L^p(\mu)$ is compact by Theorem 3.1, (ii). As shown in the proof of Proposition 3.1, any arc $I$ with $\mu_2(S(I)) \neq 0$ (that is, $S(I) \cap G \neq \emptyset$) satisfies (3.12) for a sufficiently large absolute constant $A_1 > 1$. By hypothesis (i) of Theorem 1.1 we have $\mu_2(S(I))/|I| \to 0$ as $|I| \to 0$ with $S(A_1 I) \cap \Omega(\Theta, \varepsilon) \neq \emptyset$. Hence $\mu_2$ is a vanishing Carleson measure. The theorem is proved.

Another vanishing condition (which guarantees that the embedding is compact) was suggested in [16]: given any $\delta > 0$, we put

$$H_\delta = \{ z \in \overline{D} : \text{dist}(z, \sigma(\Theta) \cap T) < \delta \}$$

and say that a measure $\mu$ satisfies the first vanishing condition (briefly, V1) if

$$M_{\mu_\delta} \to 0, \quad \delta \to 0, \quad (4.1)$$

where $\mu_\delta = \mu|_{H_\delta}$. We recall that $M_\nu$ stands for the Carleson constant of $\nu$. If a measure $\mu$ satisfies hypothesis (i) of Theorem 1.1, then we say that $\mu$ satisfies the second vanishing condition (V2).

As shown in [16], if $\mu$ satisfies V1, then the embedding $K^p_\Theta \subset L^p(\mu)$ is compact for all $p$, $0 < p < \infty$. The authors of [16] asked what is the relationship between the two vanishing conditions? Here we answer this question by showing that V1 always implies V2, but not conversely. (Thus the sufficient condition for compactness proved in [16] follows from Theorem 1.1 for any $p > 1$.)

**Proposition 4.1.** Condition V1 implies condition V2.

**Proof.** Assume that $\mu$ satisfies V1 but not V2. Then there is a sequence $\{J_n\}_{n \in \mathbb{N}}$ of arcs such that for some fixed $\varepsilon \in (0, 1)$ we have $S(J_n) \cap \Omega(\Theta, \varepsilon) \neq \emptyset$, $|J_n| \to 0$, $n \to \infty$, but

$$\mu(S(J_n)) \geq C|J_n| \quad (4.2)$$

for some constant $C > 0$.

We fix $\delta > 0$ and put $G_\delta = \overline{D} \setminus H_\delta$. It follows from the definition of the spectrum $\sigma(\Theta)$ that $\Theta$ is continuous on $G_\delta$ and $|\Theta(z)| \to 1$ uniformly as $|z| \to 1$, $z \in G_\delta$. Hence there is a constant $\delta_1 \in (0, \delta)$ such that

$$|\Theta(z)| > \varepsilon, \quad z \in G_\delta, \quad 1 - \delta_1 \leq |z| \leq 1. \quad (4.3)$$

Choose $N$ in such a way that $|J_n| < \delta_1$, $n \geq N$. Clearly, $S(J_n) \subset \{ z \in \overline{D} : 1-\delta_1 \leq |z| \leq 1 \}$. Since $S(J_n) \cap \Omega(\Theta, \varepsilon) \neq \emptyset$, it follows from condition (4.3) that

$$S(J_n) \notin G_\delta \cap \{ 1 - \delta_1 \leq |z| \leq 1 \}, \quad n \geq N.$$

We conclude that $S(J_n) \cap H_\delta \neq \emptyset$, $n \geq N$. Hence $S(J_n) \subset H_{2\delta}$ and, therefore, $\mu(S(J_n)) \leq M_{\mu_{2\delta}}|J_n|$, $n \geq N$. This contradicts the estimate (4.2) since we have $M_{\mu_{2\delta}} \to 0$ as $\delta \to 0$ (by condition (4.1)).

**Example 4.1.** Let us show that condition V1 is not necessary for the embedding even in the case of one-component functions (and so it does not follow from V2). Let $I_n$ be the sequence of arcs in Lemma 3.3 and let $\zeta_n$ be the midpoint of $I_n$. We put
\[ \mu = \sum_n a_n |I_n| \delta_{\zeta_n}, \text{ where } a_n \to 0 \text{ as } n \to \infty. \]

By Theorem 3.1, (ii) the embedding \( K^p_\Theta \subset L^p(\mu) \) is compact for every \( p \in (1, \infty) \). Suppose that the function \( \Theta \) is one-component. Then \( \mu \) satisfies condition \( V_2 \) by Theorem 1.1, (ii) (this may be verified as in the proof of Proposition 3.1). However, the measure \( \mu_\delta \) has non-trivial point weights on \( \mathbb{T} \) for any \( \delta > 0 \) and, therefore, \( \mu_\delta \) is not a Carleson measure. Thus \( \mu \) does not satisfy condition \( V_1 \).

§ 5. The classes \( S_r \). Sufficient conditions

The definition and basic properties of the Schatten–von Neumann operator ideals \( S_r \) can be found in [32].

Let \( R_{n,m} \) be the elements of the standard dyadic decomposition of the disc \( \mathbb{D} \):

\[ R_{n,m} = \left\{ z = \rho e^{i\phi} : 1 - \frac{1}{2^n - 1} \leq \rho \leq 1 - \frac{1}{2^n}, \frac{\pi m}{2n-1} \leq \phi < \frac{\pi(m+1)}{2n-1} \right\}, \quad (5.1) \]

where \( n \in \mathbb{N}, m = 0, 1, \ldots, 2^n - 1 \).

We recall a theorem of Luecking [21] that describes the \( S_r \)-properties of embeddings of the whole Hardy space \( H^2 \). Let \( \mu \) be a Carleson measure on \( \mathbb{D} \). The embedding of \( H^2 \) to \( L^2(\mu) \) belongs to \( S_r, r > 0 \), if and only if

\[ \sum_{n,m} \left( 2^n \mu(R_{n,m}) \right)^r < \infty, \quad (5.2) \]

where the sum is taken over all dyadic squares \( R_{n,m} \). An interesting general approach to embeddings of Hilbert spaces with reproducing kernels was suggested in [22]. We use the ideas of [22], especially when proving necessary conditions.

There is an obvious criterion for the embedding operator \( J_\mu : K^2_\Theta \to L^2(\mu), J_\mu f = f \), to belong to \( S_2 \).

**Proposition 5.1.** We have \( J_\mu \in S_2 \) if and only if \( \|k_z\|_2 \in L^2(\mu) \). In this case, \( \|J_\mu\|_{S_2}^2 = \int \|k_z\|_2^2 \, d\mu(z) \).

**Proof.** We have

\[ (J_\mu f)(z) = \int_{\mathbb{T}} f(w) \overline{k_z(w)} \, dm(w), \quad f \in K^2_\Theta. \quad (5.3) \]

Let \( \tilde{J}_\mu \) be the operator defined by formula (5.3) on the whole space \( L^2(\mathbb{T}) \). Then \( \tilde{J}_\mu \) is the orthogonal projection of \( L^2(\mathbb{T}) \) onto \( K^2_\Theta \). We have \( J_\mu \in S_2 \) if and only if \( \tilde{J}_\mu \in S_2 \). This is equivalent to the condition

\[ \int_{\mathbb{T}} \int_{\mathbb{T}} |k_z(w)|^2 \, dm(w) \, d\mu(z) = \int \|k_z\|_2^2 \, d\mu(z) < \infty. \]

The proposition is proved.

The following theorem yields a sufficient condition for the inclusion \( J_\mu \in S_r \), \( r > 0 \). This condition uses the arcs \( I_n \) in Lemma 3.3 and a special family of dyadic squares. The result contains Theorem 1.2. Given \( \varepsilon \in (0,1) \) and \( A > 0 \), we put

\[ \mathcal{R}(\varepsilon, A) = \left\{ R_{n,m} : \text{dist}(R_{n,m}, \Omega(\Theta, \varepsilon)) \leq 2^{-n} A \right\}. \]
Theorem 5.1. Suppose that $r > 0$, $\mu$ is a Borel measure on $\mathbb{D} \cup (\mathbb{T} \setminus \sigma(\Theta))$, and $\varepsilon \in (0, 1)$. Then there is an absolute constant $A > 0$ such that $\mathcal{J}_\mu \in S_r$ whenever

$$\sum_{n} \left( \frac{\mu(S(I_n))}{|I_n|} \right)^{r/2} < \infty, \quad (5.4)$$

$$\sum_{R_{n,m} \in \mathcal{R} (\varepsilon, A)} (2^n \mu(R_{n,m}))^{r/2} < \infty. \quad (5.5)$$

To prove Theorem 5.1, we use the following property of the arcs $I_n$ constructed in Lemma 3.3.

Lemma 5.1. Suppose that $\varepsilon \in (0, 1)$ and $\{I_n\}$ is the system of arcs in Lemma 3.3. Then there is a constant $C = C(\varepsilon) > 0$ such that

$$|k_z(w)| = \left| \frac{1 - \overline{\Theta(z)}\Theta(w)}{1 - \overline{z}w} \right| \leq C|I_n|^{-1}$$

for all $n$ and $z, w \in S(I_n)$. In particular, $|\Theta'(\zeta)| \leq C|I_n|^{-1}$ for $\zeta \in I_n$.

Proof. By the construction of $I_n$ we have $d_\varepsilon(\zeta) = \text{dist}(\zeta, \Omega(\Theta, \varepsilon)) \geq |I_n|$, $\zeta \in I_n$. We write $w = r\zeta$, where $r \in (0, 1)$, $\zeta \in I_n$. Clearly, $|k_z(w)| \leq \|k_z\|_2 \|k_w\|_2$, and Lemma 3.1 yields that

$$\|k_w\|_2^2 \leq C\|k_\zeta\|_2^2 = C|\Theta'(\zeta)|.$$ 

It follows from inequality (2.4) (see also [19], Theorem 4.9) that $|\Theta'(\zeta)| \leq C_1(d_\varepsilon(\zeta))^{-1} \leq C_1|I_n|^{-1}$. Hence $\|k_w\|_2^2 \leq C_2|I_n|^{-1}$ for $w \in S(I_n)$. The lemma is proved.

Turning to Theorem 5.1, we first treat the case $0 < r \leq 1$ using an idea in [22]. Then we use the method of complex interpolation between various ideals $S_r$ to give a proof in the case $r > 1$.

Proof of Theorem 5.1. Given any $\varepsilon \in (0, 1)$, consider the family of arcs $\{I_n\}$ constructed in Lemma 3.3. As in the proof of Proposition 3.1, put $F = \bigcup_n S(I_n)$ and $G = \mathbb{D} \setminus F$. Note that $G \cap \mathbb{T} = \sigma(\Theta) \cap \mathbb{T}$ and, therefore, $\mu(G \cap \mathbb{T}) = 0$.

We start by proving that, for every $r > 0$, condition (5.5) with an appropriate $A$ guarantees that the embedding operator $\mathcal{J}_{\mu|G} : H^2 \to L^2(\mu|_G)$ belongs to $S_r$. Let $R_{n,m}$ be a dyadic square such that $R_{n,m} \cap G \neq \emptyset$. We claim that $R_{n,m} \in \mathcal{R} (\varepsilon, A)$ for some $A > 0$. Indeed, take $z \in R_{n,m} \setminus F$, $z = (1 - \rho)\zeta$, $\rho \in (0, 1)$, $\zeta \in \mathbb{T}$. If $\zeta \in \sigma(\Theta)$, then $\text{dist}(z, \Omega(\Theta, \varepsilon)) \leq \rho \leq 2^{-(n-1)}$. Otherwise we have $\zeta \in I_k$ for some $k$, and the estimate (3.10) yields that $\text{dist}(\zeta, \Omega(\Theta, \varepsilon)) \leq 6 \pi \rho$. Hence $\text{dist}(z, \Omega(\Theta, \varepsilon)) \leq (6 \pi + 1)\rho \leq 2^{-(n-1)}(6 \pi + 1)$. We conclude that $R_{n,m} \in \mathcal{R} (\varepsilon, A)$ for $A = 12 \pi + 2$. Thus,

$$\sum_{R_{n,m} \cap G \neq \emptyset} (2^n \mu(R_{n,m}))^{r/2} < \infty,$$

and $\mathcal{J}_{\mu|G} \in S_r$ by Luecking’s theorem.
We now consider the embedding operator $J : K_{\Omega}^2 \to L^2(\mu)$, where $\mu$ is a measure on the set $F = \bigcup_n S(I_n)$.

Let $0 < r \leq 1$. We use an idea in [22]. Let $D_n$ be the smallest disc containing the Carleson square $S_n = S(I_n)$. By Lemma 3.3 we have $\text{dist}(I_n, \Omega(\Theta, \varepsilon)) \geq |I_n|$, whence the radius $d_n$ of $D_n$ does not exceed $2|I_n|/3$. Let $\widetilde{D}_n$ be the disc of radius $\widetilde{d}_n = 3|I_n|/4$ with the same centre. Then

$$\text{dist}(\widetilde{D}_n, \Omega(\Theta, \varepsilon)) \asymp |I_n|.$$ 

In this case the function $\Theta$ is analytic in $\widetilde{D}_n$, $\Theta(z) = 1/\Theta(1/z)$, $z \in \widetilde{D}_n \setminus \mathbb{D}$, and $|\Theta(z)| \leq 1/\varepsilon$, $z \in \widetilde{D}_n$. Moreover, every function $f \in K_{\Omega}^2$ is analytic in $\widetilde{D}_n$, and $k_z$ is a reproducing kernel for $K_{\Omega}^2$ at the point $z \in \widetilde{D}_n$. As in Lemma 5.1, we conclude that

$$\|k_z\|^2_2 \lesssim |I_n|^{-1}, \quad z \in \widetilde{D}_n. \tag{5.6}$$ 

By the well-known inequality of Rotfeld for ideals $S_r$, $r \leq 1$, we have $\|A + B\|_{S_r} \leq \|A\|_{S_r} + \|B\|_{S_r}$. Therefore we can represent the embedding operator $J : K_{\Omega}^2 \to L^2(\mu)$ as the sum of the embedding operators $J_n : K_{\Omega}^2 \to L^2(\mu|S_n)$ and estimate their $S_r$-norms separately. We now factorize $J_n$ as $J_n = J_n^{(2)} J_n^{(1)}$, where $J_n^{(1)}$ is the embedding of $K_{\Omega}^2$ in $H^2(\widetilde{D}_n)$ and $J_n^{(2)}$ is the embedding of $H^2(\widetilde{D}_n)$ in $L^2(\mu|S_n)$. Here $H^2(\widetilde{D}_n)$ stands for the Hardy space in the disc $\widetilde{D}_n$.

It follows from standard properties of the ideals $S_r$ that

$$\|J_n\|_{S_r} \leq \|J_n^{(2)}\|_{S_r} \|J_n^{(1)}\|_{S_2}. \tag{5.7}$$

Inequality (5.6) implies that

$$\|J_n^{(1)}\|^2_{S_2} = \frac{1}{2\pi d_n} \int_{\partial \widetilde{D}_n} \int_{T} |k_z(\zeta)|^2 dm(\zeta) |dz| = \frac{1}{2\pi d_n} \int_{\partial \widetilde{D}_n} \|k_z\|^2_2 |dz| \leq C|I_n|^{-1}.$$

Note that $d_n \leq \delta \widetilde{d}_n$, where $\delta < 1$ is an absolute constant. Let $s_l$ be the lth singular number of the operator $J_n^{(2)}$. We claim that

$$s_l \leq C(\delta)^{1/2}(\mu(D_n))^{1/2}. \tag{5.8}$$

Indeed, by applying a shift and a linear change of variable, we reduce the problem to the case when $D_n = \mathbb{D}$, $D_n = \delta \mathbb{D}$, and $\nu$ is a measure on $\delta \mathbb{D}$. Then $s_l$ does not exceed the norm of the restriction to the subspace $z^l H^2(\mathbb{D}) \subset H^2(\mathbb{D})$ of the embedding in $L^2(\nu)$. Note that $|f(z)| \leq C(\delta)\|f\|_2$, $z \in \delta \mathbb{D}$. Then

$$\|z^l f\|^2_{L^2(\nu)} = \int_{\delta \mathbb{D}} |z^l f(z)|^2 d\nu(z) \leq C^2(\delta)^{2l} \delta^{2l} \nu(\delta \mathbb{D})\|f\|^2_2; \quad f \in H^2(\mathbb{D}),$$

and the estimate (5.8) follows. Summing the $s_l^r$, we now obtain that $\|J_n^{(2)}\|^r_r \leq C(\mu(S_n))^{r/2}$. Hence, by inequality (5.7), we have

$$\|J_n\|^r_r \leq C\left(\frac{\mu(S_n)}{|I_n|}\right)^{r/2}.$$
We conclude that
\[ \|J\|_r^r \leq C \sum_n \left( \frac{\mu(S_n)}{|I_n|} \right)^{r/2}. \]

Now take \( r > 1 \). We put \( a_n = \mu(S_n) \) and consider the normalized measures \( \mu_n = a_n^{-1}\mu|_{S_n} \) on the Carleson squares \( S_n \). Then \( \mu = \sum_n a_n\mu_n \). We introduce the measure \( \nu = \sum_n \mu_n \). Clearly, the map \( f \mapsto \sum_n a_n^{1/2} f|_{S_n} \) is a unitary operator acting from \( L^2(\mu) \) onto \( L^2(\nu) \) (we recall that \( \chi_G \) stands for the characteristic function of the set \( G \)). Hence the embedding operator \( J \) belongs to \( S_r \) if and only if the operator
\[ T: K^2_{\Theta} \to L^2(\nu), \quad T f = \sum_n a_n^{1/2} f|_{S_n}, \]
belongs to \( S_r \), and \( \|J\|_{S_r} = \|T\|_{S_r} \). We claim that \( T \in S_r \) and \( \|T\|_{S_r} \leq C\mathfrak{M}_r(\mu) \), where \( \mathfrak{M}_r(\mu) = \sum_n (a_n/|I_n|)^{r/2} < \infty \).

We use the complex interpolation method (see [32]). Consider the following analytic family of operators \( T(\zeta): K^2_{\Theta} \to L^2(\nu) \) parametrized by \( \zeta \in \mathbb{C}, 0 \leq \Re \zeta \leq 1: \)
\[ T(\zeta)f = \sum_n a_n^{1/2} \left( \frac{a_n}{|I_n|^{r/(r-1)}} \right)^{\zeta r-1} f|_{S_n}. \]

Clearly, \( T(1/r) = T \). We shall see that \( T(\zeta) \) is bounded as an operator from \( K^2_{\Theta} \) to \( L^2(\nu) \) and
\[ \|T(\zeta)\| \leq A_0, \quad \Re \zeta \in [0, 1]. \quad \text{(5.9)} \]
We shall also see that \( T(\zeta) \in S_1 \) for \( \Re \zeta = 1 \) and
\[ \|T(\zeta)\|_{S_1} \leq A_1, \quad \Re \zeta = 1. \quad \text{(5.10)} \]

Then Theorem 13.1 of [32] yields that \( T = T(1/r) \in S_r \) and \( \|T\|_{S_r} \leq A_0^{1-1/r} A_1^{1/r}. \)

We note that \( |t^{r-1}| \leq t^{-1} \) for \( t > 0 \), whence
\[ |T(\zeta)f| \leq \sum_n |I_n|^{r/(r-1)} |f| |\chi_{S_n}| = Mf. \]
The non-linear operator \( M \) will be bounded as an operator from \( K^2_{\Theta} \) to \( L^2(\nu) \) provided that the embedding of \( K^2_{\Theta} \) in \( L^2(\nu_0) \) with measure \( \nu_0 = \sum_n |I_n|^{r/(r-1)} \mu_n \) is bounded. But this assumption holds by Theorem 3.1 because \( \nu_0(S_n) = |I_n|^{r/(r-1)} \leq C_0|I_n| \). Thus the estimate (5.9) holds.

Now suppose that \( \Re \zeta = 1 \). Then
\[ T(\zeta)f = \sum_n \left( \frac{a_n}{|I_n|} \right)^{r/2} e^{i\alpha_n} f|_{S_n}, \]
where \( \alpha_n = \frac{r}{2} \Im \zeta \log \frac{a_n}{|I_n|^{r/(r-1)}} \). Clearly, \( T(\zeta) \) belongs to \( S_1 \) (as an operator from \( K^2_{\Theta} \) to \( L^2(\nu) \)) if and only if the embedding of \( K^2_{\Theta} \) in \( L^2(\nu_1) \) belongs to \( S_1 \), where
\[ \nu_1 = \sum_n \left( \frac{a_n}{|I_n|} \right)^r \mu_n. \]
We have $\nu_1(S_n) = \left(\frac{a_n}{|T_n|}\right)^r$ and
\[
\mathcal{M}_1(\nu_1) = \sum_n \left(\frac{a_n}{|T_n|}\right)^{r/2} < \infty.
\]
Applying the result obtained above in the case of $\mathcal{S}_1$, we conclude that $T(\zeta) \in \mathcal{S}_1$ and the estimate (5.10) holds with $A_1 = C_1 \mathcal{M}_1(\nu_1) = C_1 \mathcal{M}_r(\mu)$. Therefore $T = T(1/r) \in \mathcal{S}_r$ and $\|T\|_{\mathcal{S}_r}^r \leq C_2 \mathcal{M}_r(\mu)$. The theorem is proved.

§ 6. Necessary conditions for the inclusion $\mathcal{J}_\mu \in \mathcal{S}_r$.

Proof of Theorem 1.4

In this section we consider conditions necessary for the inclusion $\mathcal{J}_\mu \in \mathcal{S}_r$, $r \geq 1$. We use general results obtained in [22]. Let $X$ be a Hilbert space consisting of functions analytic in a domain $D$, with reproducing kernel $K$. Suppose that $\{D_n\}$ is a decomposition of $D$ and for every $n$ there is a $w_n \in D_n$ such that the following estimate holds for all $z \in D_n$:
\[
|K(z, w_n)|^2 \geq cK(z, z)K(w_n, w_n), \tag{6.1}
\]
where $c$ is a positive constant (independent of $n$). We consider the discrete measure $\nu = \sum_n (K(w_n, w_n))^{-1} \delta_{w_n}$. Given a measure $\mu$ on $G$, we put
\[
j_n = \left(\int_{D_n} K(z, z) d\mu(z)\right)^{1/2}.
\]
If the embedding of $X$ in $L^2(\nu)$ is bounded and the embedding of $X$ in $L^2(\mu)$ belongs to $\mathcal{S}_r$, $r \geq 1$, then $\{j_n\} \in \ell^r$ ([22], Theorem 3).

Proof of Theorem 1.3. We fix an enumeration $R_n$, $n \in \mathbb{N}$, of the set of all squares $R_{l,m}$ such that $R_{l,m} \cap \Omega(\Theta, \varepsilon) \neq \emptyset$. Choose a point $w_n$ in each square $R_n$ in such a way that $|\Theta(w_n)| < \varepsilon$. If $R_n = R_{l,m}$, we put $d_n = 2^{-l}$. We easily see that there is a number $\delta < 1$ (depending only on $\varepsilon$) such that $|\Theta(z)| < \delta$, $z \in R_n$. In other words, $R_n \subset \Omega(\Theta, \delta)$ for all $n$.

We have $\|k_z\|_2^2 \sim d_n^{-1}$ for $z \in R_n$, with the constants depending only on $\delta$. Therefore the sets $D_n = R_n$ and points $w_n$ satisfy condition (6.1). Putting $\nu = \sum_n d_n \delta_{w_n}$, we see from the construction of the Carleson contours (see [1], Ch. VIII, § 5) that $\nu$ is a Carleson measure. Now let $\mu$ be a measure on $\bigcup_n R_n$, and let $\mathcal{J}_\mu : K^2_\Theta \to L^2(\mu)$ be the embedding operator. If $\mathcal{J}_\mu \in \mathcal{S}_r$, $r \geq 1$, then Theorem 3 of [22] yields that $\{j_n\} \in \ell^r$, where $j_n = \left(\int_{R_n} \|k_z\|_2^2 d\mu(z)\right)^{1/2} \simeq (\mu(R_n)/d_n)^{1/2}$. The theorem is proved.

We conclude this section by proving Theorem 1.4. First we obtain an elementary estimate for inner functions.

Lemma 6.1. Suppose that $\zeta \in \mathbb{T} \setminus \sigma(\Theta)$, $z \in \mathbb{D}$, and $|z - \zeta| < A \operatorname{dist}(\zeta, \sigma(\Theta))$ for some constant $A \in (0, 1)$. Then there is a constant $C = C(A) > 0$ such that
\[
\log |\Theta(z)| \leq -C(1 - |z|)|\Theta'(\zeta)|.
\]
Proof. By Frostman’s theorem, $\Theta_\alpha = \Theta - \frac{\alpha}{1 - \alpha}$ is a Blaschke product for almost all $\alpha$, $|\alpha| < 1$, and $\|\Theta_\alpha - \Theta\|_\infty \to 0$ as $\alpha \to 0$. We also have $|\Theta'_\alpha(\zeta)| \to |\Theta'(\zeta)|$ as $\alpha \to 0$ if $\zeta \in \mathbb{T} \setminus \sigma(\Theta)$. Thus it suffices to prove the estimate in the case when $\Theta$ is a Blaschke product.

Let $B$ be the Blaschke product with zeros $z_n$. Take $z \in \mathbb{D}$. Then
\[
\log |B(z)|^2 = \sum_n \log \left(1 - \frac{(1 - |z_n|^2)(1 - |z|^2)}{|1 - z_n z|^2}\right).
\]
We also recall that $|B'(\zeta)| = \sum_n \frac{1 - |z_n|^2}{|\zeta - z_n|^2}$ for $\zeta \in \mathbb{T}$. Since $|z - \zeta| < A \text{dist}(\zeta, \sigma(\Theta))$, we have $|z - \zeta| < A|z_n - \zeta|$ for all $n$. Therefore,
\[
(1 - A)|\zeta - z_n| < |1 - z_n z| < (1 + A)|\zeta - z_n|.
\]
Since $\log(1-t) < -t$ for $t \in (0, 1)$, we get
\[
\log |B(z)|^2 < -\sum_n \frac{(1 - |z_n|^2)(1 - |z|^2)}{|1 - z_n z|^2}
< -C(A)(1 - |z|) \sum_n \frac{1 - |z_n|^2}{|\zeta - z_n|^2} = -C(A)(1 - |z|)|B'(\zeta)|.
\]
The lemma is proved.

In the following lemma $\{I_n\}$ stands for the family of arcs constructed in Lemma 3.3 for a given $\varepsilon \in (0, 1)$.

**Lemma 6.2.** Let $\Theta$ be a one-component inner function. Then there is a $\delta \in (0, 1)$ (depending on $\Theta$ and $\varepsilon$ but not on $n$) such that $|\Theta(z)| \leq \delta$ for $z = (1 - |I_n|/(2\pi))\zeta$, $\zeta \in I_n$ (that is, for $z$ lying on the inner side of the square $S(I_n)$). We also have $k_z(z) = \|k_z\|^2 \simeq |I_n|^{-1}$, $z \in S(I_n)$.

**Proof.** First we prove that there are constants $C_j = C_j(\Theta, \varepsilon) > 0$, $j = 1, 2$, such that
\[
C_1 |I_n|^{-1} \leq |\Theta'(\zeta)| \leq C_2 |I_n|^{-1}, \quad \zeta \in I_n.
\] (6.2)
Only the first inequality must be proved (the second follows from Lemma 5.1).

By Lemma 3.3 there is a point $w \in \Omega(\Theta, \varepsilon)$ such that $|\zeta - w| \leq C_3 |I_n|$ for $\zeta \in I_n$, where $C_3 > 0$ is an absolute constant. Hence,
\[
|k_w(\zeta)| \geq \frac{1 - |\Theta(w)|}{|\zeta - w|} \geq C_3^{-1}(1 - \varepsilon)|I_n|^{-1}.
\]
On the other hand, an inequality obtained in [12] yields that
\[
|k_w(\zeta)| \leq C_4 |\Theta'(\zeta)|, \quad w \in \mathbb{D}, \quad \zeta \in \mathbb{T},
\] (6.3)
for the one-component function $\Theta$, whence we get the estimate (6.2).

We now fix $\zeta \in I_n$ and put $z = (1 - |I_n|/(2\pi))\zeta$. Since $\text{dist}(I_n, \sigma(\Theta)) \not\geq |I_n|$, we have $|\zeta - z| < A \text{dist}(\zeta, \sigma(\Theta))$ for some $A < 1$. It follows from inequality (6.2) and Lemma 6.1 that $|\Theta(z)| \leq \delta = \exp(-C(A)C_1/(2\pi))$. We conclude that $k_z(z) \simeq |I_n|^{-1}$ when $z = (1 - |I_n|/(2\pi))\zeta$, $\zeta \in I_n$. Lemma 3.1 now yields the estimate $k_z(z) = \|k_z\|^2 \simeq |I_n|^{-1}$ for all $z \in S(I_n)$. The lemma is proved.
Corollary 6.1. Let $\Theta$ be a one-component inner function, $\varepsilon \in (0, 1)$. Then there is a number $\delta \in (0, 1)$ such that $|\Theta(z)| \leq \delta$ for $z \in G \cap \mathbb{D}$.

Proof. Recall that $G = \overline{\mathbb{D}} \setminus \bigcup_n S(I_n)$. We claim that there is a $\delta \in (0, 1)$ such that $|\Theta(z)| \leq \delta$ for $z \in \partial G \cap \mathbb{D}$. Indeed, $\partial G$ is a rectifiable Jordan curve, we have $|\Theta(z)| \leq 1$ in $G$, and $\partial G \cap \mathbb{T} = \sigma(\Theta) \cap \mathbb{T}$ has Lebesgue measure zero (see [24]). Therefore we conclude $|\Theta(z)| \leq \delta$ for $z \in G \cap \mathbb{D}$.

Take $z \in \partial G \cap \mathbb{D}$. Then there are two possibilities: either $z = (1 - |I_n|/(2\pi))\zeta$, $\zeta \in I_n$, for some $n$ ($z$ lies on the inner side of some square), or there are adjacent squares $S(I_n)$ and $S(I_m)$ such that $|I_n| \leq |I_m|$ and $z = r\zeta$, where $\zeta$ is the common endpoint of the arcs $I_n$ and $I_m$, and we have $1 - |I_m|/(2\pi) \leq r \leq 1 - |I_n|/(2\pi)$. In the first case we have $|\Theta(z)| \leq \delta_1 < 1$ by Lemma 6.2. Note that $|I_n| \simeq |I_m| \simeq |\Theta'(\zeta)|^{-1}$ by inequality (6.2). In the second case we have $|\Theta(z)| \leq \delta_2 < 1$ by Lemma 6.1. Here $\delta_1$ and $\delta_2$ depend only on $\varepsilon$. The corollary is proved.

Proof of Theorem 1.4. We start by proving the sufficiency of conditions (1.6) and (1.7). As above, we put $F = \bigcup_n S(I_n)$, $G = \overline{\mathbb{D}} \setminus F$. It follows from Theorem 1.2 that the embedding of $K_\Theta^2$ in $L^2(\mu_F)$ belongs to $\mathcal{S}_r$. Now let $R_{n,m}$ be a dyadic square such that $R_{n,m} \cap G \neq \emptyset$. Then Corollary 6.1 yields a constant $\delta < 1$ such that $R_{n,m} \cap \Omega(\Theta, \delta) \neq \emptyset$. By condition (1.7) we have

$$\sum_{R_{n,m} \cap G \neq \emptyset} (2^n \mu(R_{n,m}))^{r/2} \leq \sum_{R_{n,m} \cap \Omega(\Theta, \delta) \neq \emptyset} (2^n \mu(R_{n,m}))^{r/2} < \infty,$$

and the inclusion $\mathcal{J}_{\mu|G} \in \mathcal{S}_r$ follows from Luecking’s theorem.

By Theorem 1.3, condition (1.7) is necessary for the inclusion $\mathcal{J}_\mu \in \mathcal{S}_r$ even in the case of an arbitrary inner function. To prove the necessity of (1.6), we verify that the hypotheses of Parfenov’s theorem hold for $D_n = S(I_n)$. By Lemma 6.2 one can choose points $w_n \in S(I_n)$ in such a way that $|\Theta(w_n)| \leq \delta$. Hence we have

$$|k_{w_n}(z)|^2 = \left| \frac{1 - \Theta(z)\Theta(w_n)}{1 - z\overline{w_n}} \right|^2 \geq C_1 |I_n|^{-2} \geq C_2 k_z(z) k_{w_n}(w_n).$$

Here we have used the estimates $|1 - z\overline{w_n}| \lesssim |I_n|$ and $k_z(z) \lesssim |I_n|^{-1}$ (see Lemma 6.2), which hold for all $z \in S(I_n)$. We also have $k_{w_n}(w_n) \asymp |I_n|^{-1}$, and the measure $\nu = \sum_n |I_n| \delta_{w_n}$ belongs to $C_2(\Theta)$ by Theorem 3.1, (i).

If $\mathcal{J}_{\mu|F} \in \mathcal{S}_r$, $r \geq 1$, then Theorem 3 of [22] yields that $\{j_n\} \in \ell^r$,

$$j_n = \left( \int_{S(I_n)} k_z(z) \, d\mu(z) \right)^{1/2}.$$

It remains to note that $j_n \asymp \left( \mu(S(I_n))/|I_n| \right)^{1/2}$ since $k_z(z) \asymp |I_n|^{-1}$, $z \in S(I_n)$.

The theorem is proved.

Remark 6.1. The condition $R_{n,m} \in \mathcal{R}(\varepsilon, A)$ in Theorem 5.1 means that the distance from the dyadic square $R_{n,m}$ to the level set $\Omega(\Theta, \varepsilon)$ is not much larger than the size of $R_{n,m}$. The constant $A = 12\pi + 2$ (which appears in the proof of Theorem 5.1) is far from being exact. There is a gap between the sufficient condition (5.5) and
the necessary condition (1.7). We note that the inclusion $R_{n,m} \in R(\varepsilon,A)$ does not generally imply that $R_{n,m} \cap \Omega(\Theta,\varepsilon_1) \neq \emptyset$ for some $\varepsilon_1 \in (0,1)$ independent of $n$ and $m$.

**Remark 6.2.** Theorem 1.4 (or rather its analogue for the upper half-plane) generalizes Parfenov’s theorem [23] on embeddings of the Paley–Wiener space $PW^2_a$ (the space of entire functions of exponential type at most $a$ that are square-summable on $\mathbb{R}$): if $\mu$ is a measure on $\mathbb{R}$ and $J_\mu : PW^2_a \to L^2(\mu)$, $J_\mu f = f$, then the inclusion $J_\mu \in S_r$, $r > 0$, is equivalent to

$$\sum_{n \in \mathbb{Z}} \left( \mu([n,n+1]) \right)^{r/2} < \infty.$$ 

The space $PW^2_a$ basically coincides with the model space generated by the one-component inner function $\theta(z) = \exp(2iaz)$ on $\mathbb{C}^+$. More precisely, $PW^2_a = e^{-iaz}(K^2_{\theta})_+$. We note that the intervals $J_n = [n,n+1]$ possess the same property (with respect to $\theta$) as the arcs $I_n$ in Lemma 3.3: for every $\varepsilon \in (0,1)$ the distance $\text{dist}(J_n, \Omega(\theta,\varepsilon))$ is comparable with the length of $J_n$.

**Bibliography**

[1] J. B. Garnett, *Bounded analytic functions*, Pure Appl. Math., vol. 96, Academic Press, New York–London 1981; Russian transl., Mir, Moscow 1984.

[2] A. B. Aleksandrov, “Invariant subspaces of the shift operator. Axiomatic approach”, *Studies on linear operator and function theory*. XI, Zap. Nauchn. Sem. LOMI, vol. 113, Nauka, Leningrad 1981, pp. 7–26; English transl., *J. Soviet Math.* 22:6 (1983), 1695–1708.

[3] N. K. Nikol’skii, *Treatise on the shift operator*, Nauka, Moscow 1980; English transl., Grundlehren Math. Wiss., vol. 273, Springer-Verlag, Berlin–Heidelberg 1986.

[4] N. K. Nikol’ski, *Operators, functions, and systems: an easy reading. Vol. 2: Model operators and systems*, Math. Surveys Monogr., vol. 93, Amer. Math. Soc., Providence, RI 2002.

[5] J. A. Cima and W. T. Ross, *The backward shift on the Hardy space*, Math. Surveys Monogr., vol. 79, Amer. Math. Soc., Providence, RI 2000.

[6] S. V. Hruščev, N. K. Nikol’skii, and B. S. Pavlov, “Unconditional bases of exponentials and of reproducing kernels”, *Complex analysis and spectral theory*, Lecture Notes in Math., vol. 864, Springer-Verlag, Berlin–New York 1981, pp. 214–335.

[7] B. Cohn, “Carleson measures for functions orthogonal to invariant subspaces”, *Pacific J. Math.* 103:2 (1982), 347–364.

[8] S. C. Power, “Vanishing Carleson measures”, *Bull. Lond. Math. Soc.* 12:3 (1980), 207–210.

[9] O. Blasco and H. Jarchow, “A note on Carleson measures for Hardy spaces”, *Acta Sci. Math. (Szeged)* 71:1–2 (2005), 371–389.

[10] F. Nazarov and A. Volberg, “The Bellman function, the two-weight Hilbert transform, and embeddings of the model spaces $K^2_\theta$”, *J. Anal. Math.* 87:1 (2002), 385–414.

[11] A. L. Vol’berg and S. R. Treil’, “Imbedding theorems for the invariant subspaces of the backward shift operator”, *Studies on linear operators and function theory*. XV, Zap. Nauchn. Sem. LOMI, vol. 149, Nauka, Leningrad 1986, pp. 38–51; English transl., *J. Soviet Math.* 42:2 (1988), 1562–1572.

[12] A. B. Aleksandrov, “On embedding theorems for coinvariant subspaces of the shift operator. II”, *Studies on linear operators and function theory*. 27, Zap. Nauchn. Sem. POMI, vol. 262, POMI, St. Petersburg 1999, pp. 5–48; English transl., *J. Math. Sci. (New York)* 110:5 (2002), 2907–2929.
1100 A. D. Baranov

[13] W. S. Cohn, “Carleson measures and operators on star-invariant subspaces”, J. Operator Theory 15:1 (1986), 181–202.

[14] K. M. Dyakonov, “Embedding theorems for star-invariant subspaces generated by smooth inner functions”, J. Funct. Anal. 157:2 (1998), 588–598.

[15] A. L. Volberg, “Thin and thick families of rational fractions”, Complex analysis and spectral theory, Lecture Notes in Math., vol. 864, Springer-Verlag, Berlin–New York 1981, pp. 440–480.

[16] J. A. Cima and A. L. Matheson, “On Carleson embeddings of star-invariant supspaces”, Quaest. Math. 26:3 (2003), 279–288.

[17] D. N. Clark, “One dimensional perturbations of restricted shifts”, J. Analyse Math. 25:1 (1972), 169–191.

[18] A. D. Baranov, “Weighted Bernstein-type inequalities, and embedding theorems for the model subspaces”, Algebra i Analiz 15:5 (2003), 138–168; English transl., St. Petersburg Math. J. 15:5 (2004), 733–752.

[19] A. D. Baranov, “Bernstein-type inequalities for shift-coinvariant subspaces and their applications to Carleson embeddings”, J. Funct. Anal. 223:1 (2005), 116–146.

[20] O. G. Parfenov, “On properties of imbedding operators of certain classes of analytic functions”, Algebra i Analiz 3:2 (1991), 199–222; English transl., St. Petersburg Math. J. 3:2 (1992), 425–446.

[21] O. G. Parfenov, “Weighted estimates of the Fourier transformation”, Studies on linear operators and function theory. 23, Zap. Nauchn. Sem. POMI, vol. 303, POMI, St. Petersburg 2003, pp. 5–33; English transl., J. Math. Sci. (N. Y.) 129:4 (2005), 3927–3943.