Abstract

In this paper we introduce a class of Bernstein–Durrmeyer operators with respect to an arbitrary measure \( \rho \) on the \( d \)-dimensional simplex, and a class of more general polynomial integral operators with a kernel function involving the Bernstein basis polynomials. These operators generalize the well-known Bernstein–Durrmeyer operators with respect to Jacobi weights. We investigate properties of the new operators. In particular, we study the associated reproducing kernel Hilbert space and show that the Bernstein basis functions are orthogonal in the corresponding inner product. We discuss spectral properties of the operators. We make first steps in understanding convergence of the operators.

© 2009 Elsevier Inc. All rights reserved.

Keywords: Bernstein basis polynomials; Bernstein–Durrmeyer operator; Jacobi weight; Reproducing kernel Hilbert space; Korovkin type theorem

1. Introduction

The motivation of this paper comes from paper [11], in which D.-X. Zhou and the second author have applied the univariate Bernstein–Durrmeyer operator to bias-variance estimates as they are common in learning theory. Bernstein–Durrmeyer operators are well known in approximation theory, and their properties have been studied in great detail for special weight functions, namely, for Jacobi weights. In this paper we study these operators under more general assumptions and from a somewhat different point of view.
Let us start with notation. The standard simplex in $\mathbb{R}^d$ is the set
$$\mathbb{S}^d := \left\{ x = (x_1, \ldots, x_d) \in \mathbb{R}^d : 0 \leq x_1, \ldots, x_d \leq 1, 0 \leq x_1 + \cdots + x_d \leq 1 \right\}.$$ Instead of Cartesian coordinates, we will often refer to barycentric coordinates, which we denote by a boldface symbol $x = (x_0, x_1, \ldots, x_d)$, $x_0 := 1 - x_1 - \cdots - x_d$.

Throughout the paper we will use more or less standard multi-index notation such as $x^\alpha = x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_d^{\alpha_d}$ or $|\alpha| = \alpha_0 + \cdots + \alpha_d$ for $x, \alpha \in \mathbb{R}^{d+1}$. Note that in $|\alpha|$ the sum of the components (and not the sum of their absolute values) is taken. In particular, for $x \in \mathbb{S}^d$, the components of $x$ are all non-negative, and they add up to 1.

Let $P_n^d$ denote the space of $d$-variate algebraic polynomials of (total) degree at most $n$. The $d$-variate Bernstein basis polynomials of degree $n$ are defined by
$$B_\alpha(x) := \binom{n}{\alpha} x^\alpha := \frac{n!}{\alpha_0! \alpha_1! \cdots \alpha_d!} x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_d^{\alpha_d}, \quad |\alpha| = n,$$
where the components of $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_d)$ are non-negative integers. All together there are $\binom{n+d}{d} = \dim P_n^d$ linear independent polynomials of degree $n$ that constitute a basis of $P_n^d$. This basis is often preferred to the monomial basis in numerical calculations, since it is known to be better conditioned. It is also used in computer aided geometric design for the design of curves and surfaces, since a polynomial represented in this basis can be fast and stably evaluated in a recursive way by the famous de Casteljau algorithm.

In approximation theory, the Bernstein polynomial basis is used in various types of positive polynomial approximation schemes. One of them is the famous Bernstein polynomial operator of degree $n$,
$$B_n : C(\mathbb{S}^d) \to P_n^d, \quad f \mapsto \sum_{|\alpha|=n} f \left( \frac{\alpha}{n} \right) B_\alpha,$$
which maps continuous functions on the simplex onto $P_n^d$. This positive linear operator reproduces constant functions and linear polynomials, i.e., we have
$$\sum_{|\alpha|=n} B_\alpha(x) = 1,$$
and
$$\sum_{|\alpha|=n} \frac{\alpha_i}{n} B_\alpha(x) = x_i, \quad i = 0, \ldots, d.$$
Assumption N. The measure $\rho$ satisfies $\int_{\mathbb{S}^d} d\rho = 1$.

The spaces $L^q_{\rho}(\mathbb{S}^d), 1 \leq q < \infty$, are defined in a usual way as spaces of (equivalence classes) of real-valued functions $f$, for which $|f|^q$ is integrable with respect to $\rho$, with the norm

$$\|f\|_{L^q_{\rho}(\mathbb{S}^d)} := \left(\int_{\mathbb{S}^d} |f(x)|^q d\rho(x)\right)^{1/q}, \quad 1 \leq q < \infty.$$ 

The space $L^\infty_{\rho}(\mathbb{S}^d)$ is the space of $\rho$-essentially bounded functions with the norm

$$\|f\|_{L^\infty_{\rho}(\mathbb{S}^d)} := \text{ess sup}_{x \in \mathbb{S}^d} |f(x)|.$$ 

We shall also consider the space $C(\mathbb{S}^d)$ of continuous, real-valued functions on the simplex, with the norm

$$\|f\|_{C(\mathbb{S}^d)} := \max_{x \in \mathbb{S}^d} |f(x)|.$$ 

Let $1$ denote the constant function equal to 1, i.e.,

$$1 : x \mapsto 1(x) \equiv 1, \quad x \in \mathbb{S}^d.$$ 

Due to Assumption N, $\|1\|_{L^q_{\rho}(\mathbb{S}^d)} = 1$ for all $q$. Note that $C(\mathbb{S}^d) \subset L^1_{\rho}(\mathbb{S}^d)$ and $L^q_{\rho}(\mathbb{S}^d) \subset L^1_{\rho}(\mathbb{S}^d)$ for all $1 \leq q < \infty$. On the space $L^2_{\rho}(\mathbb{S}^d)$ we consider the (semi-)inner product

$$(f|g)_{\rho} := \int_{\mathbb{S}^d} f(x) g(x) d\rho(x).$$

Assumption P guarantees that $(p|p)_{\rho} > 0$ for any polynomial $p$ which is not identically zero.

Definition 1.1. The $\rho$-weighted Bernstein–Durrmeyer operator of degree $n$ on the simplex $\mathbb{S}^d$ is the operator

$$M_{n,\rho} : L^1_{\rho}(\mathbb{S}^d) \to \mathbb{R}^d, \quad f \mapsto \sum_{|\alpha|=n} \langle f|B_\alpha \rangle_{\rho} \frac{1}{(1|B_\alpha \rangle_{\rho}} B_\alpha.$$ 

As far as we know, this operator in this generality has never been studied in the literature. For applications in learning theory, however, it is important to consider this general setting, although the measure $\rho$ will refer to an unknown probability distribution. This generalization was shortly mentioned by Berens and Xu [5] but not further investigated. On the other hand, the Bernstein–Durrmeyer operator (1.4) for the unweighted case (i.e., $d\rho(x) = dx$ up to normalization) and, more general, for Jacobi weights (see (1.5)) is well known and very well studied. It was introduced in the one-dimensional unweighted case by Durrmeyer and, independently, by Lupăș and first studied by Derriennic [7]. It was extended to Jacobi weights by Păltănea [12] and studied by Berens and Xu [5]. The multidimensional case was dealt, e.g., in [8,9]. For details and further references, see the papers mentioned above and [3]. In particular, Ditzian’s paper [9] gives a good overview over properties of the multivariate Bernstein–Durrmeyer operator with Jacobi weights. In this special case, the weighted integrals of
the Bernstein basis polynomials are explicitly known as values of the multivariate beta function

\[ B(\lambda) := \int_{S^d} x^{\lambda - e} \, dx = \frac{\prod_{i=0}^{d} \Gamma(\lambda_i)}{\Gamma(|\lambda|)}, \tag{1.5} \]

where \( \lambda \in \mathbb{R}_{+}^{d+1} \) and \( e = (1, 1, \ldots, 1) \). The normalized Jacobi weight is the measure

\[ d\rho(x) = \frac{1}{B(\mu + e)} \, x^\mu \, dx \tag{1.6} \]

with \( \mu = (\mu_0, \mu_1, \ldots, \mu_d) \in \mathbb{R}^{d+1} \) satisfying the integrability condition \( \mu > -e \). In this case, we will denote the inner product (1.3) by

\[ \langle f | g \rangle_\mu = \frac{1}{B(\mu + e)} \int_{S^d} f(x) \, g(x) \, x^\mu \, dx. \tag{1.7} \]

Due to (1.5),

\[ \langle 1 | B_\alpha \rangle_\mu = \frac{1}{B(\mu + e)} \int_{S^d} B_\alpha(x) \, x^\mu \, dx = \binom{n}{\alpha} \frac{B(\alpha + \mu + e)}{B(\mu + e)}, \tag{1.8} \]

and the Jacobi-weighted Bernstein–Durrmeyer operator is given by

\[ M_n,\mu(f) = \sum_{|\alpha|=n} c_{\alpha,\mu}(f) \, B_\alpha, \quad c_{\alpha,\mu}(f) = \frac{1}{B(\alpha + \mu + e)} \int_{S^d} f(x) \, x^{\alpha + \mu} \, dx. \tag{1.9} \]

A more general operator than the Bernstein–Durrmeyer operator with Jacobi weights was considered by Păltănea. In his book [13, Section 5.2] he studied (univariate) Bernstein–Durrmeyer operators of the form (1.4) with the weight (up to normalization)

\[ d\rho(x) = x^\alpha (1-x)^\beta \, h(x) \, dx, \quad x \in [0, 1], \tag{1.10} \]

where \( h \in C[0, 1], h(t) > 0 \) for all \( t \in [0, 1], \alpha, \beta > -1 \). Such weights are sometimes called generalized Jacobi weights.

2. The kernel functions

Definition 2.1. For a given sequence \( \omega = (\omega_\alpha)_{|\alpha|=n} \) of non-negative numbers, the \( \omega \)-weighted Bernstein basis kernel of degree \( n \) is defined as the function

\[ T_{n,\omega}(x, y) = \sum_{|\alpha|=n} \omega_\alpha \, B_\alpha(x) \, B_\alpha(y), \quad x, y \in S^d. \]

It gives rise to a family of \( \rho \)-weighted integral operators

\[ (L_{n,\omega,\rho}(f))(x) = \int_{S^d} T_{n,\omega}(x, y) \, f(y) \, d\rho(y). \tag{2.1} \]

It is evident that these linear operators are positive, i.e.,

\[ f \geq 0 \implies L_{n,\omega,\rho}(f) \geq 0. \]
and symmetric, i.e.,
\[ \langle L_n,\omega,\rho(f) | g \rangle_\rho = \langle f | L_n,\omega,\rho(g) \rangle_\rho, \quad f, g \in L^2_\rho(\mathbb{S}^d). \tag{2.2} \]

The \( \rho \)-weighted Bernstein–Durrmeyer operator refers to the special case where the \( \omega \)-weights are the canonical weights chosen according to the formula
\[ \frac{1}{\omega_\alpha} = \langle 1 | B_\alpha \rangle_\rho = \int_{\mathbb{S}^d} B_\alpha(x) \, d\rho(x) > 0, \quad |\alpha| = n; \tag{2.3} \]
the integral is positive due to Assumption P. This is the choice of weights which enforces the operator to reproduce constant functions, that is,
\[ M_{n,\rho}(1) = \sum_{|\alpha|=n} B_\alpha = 1, \tag{2.4} \]
see (1.1). The action of the operator on polynomial spaces will be considered in the following sections. In this section, we deal with an interpretation of the \( \omega \)-weighted Bernstein basis kernels as reproducing kernels in an associated polynomial inner product space.

2.1. Properties of the kernel functions

**Lemma 2.2.** The kernel \( T_{n,\omega} \) is non-negative, symmetric, and continuous on \( \mathbb{S}^d \times \mathbb{S}^d \), with bound
\[ \max_{x,y \in \mathbb{S}^d} T_{n,\omega}(x,y) \leq \max_{|\alpha|=n} \omega_\alpha =: \Omega_n. \tag{2.5} \]
Moreover, the kernel is a positive (semi-)definite function on \( \mathbb{S}^d \).

**Proof.** The first statements are clear since
\[ 0 \leq T_{n,\omega}(x,y) \leq \Omega_n \sum_{|\alpha|=n} B_\alpha(x) = \Omega_n \]
for \( x, y \in \mathbb{S}^d \). The positive definiteness follows from
\[ \sum_{i,j=1}^N c_i c_j T_{n,\omega}(x_i, x_j) = \sum_{|\alpha|=n} \omega_\alpha \left\{ \sum_{i=1}^N c_i B_\alpha(x_i) \right\}^2 \geq 0 \]
for all points \( x_1, \ldots, x_N \in \mathbb{S}^d \) and all real numbers \( (c_i)_{i=1}^N \). \( \blacksquare \)

Kernels of this type are called Mercer Kernels. Their properties are used in a standard construction of an associated reproducing kernel Hilbert space (RKHS) that will be considered below, cf. [6, Section 3].

**Remark.** Estimate (2.5) is sharp under the following condition. Let
\[ V_d := \{ e_i : i = 0, \ldots, d \} \]
denote the set of vertices of the simplex \( \mathbb{S}^d \), in barycentric notation. So, all the coordinates of \( e_i \) are zero except at position \( i \), where the entry is 1. Since
\[ B_\alpha(e_i) = \delta_{\alpha,n} e_i \quad \text{for } |\alpha| = n \text{ and } i = 0, \ldots, d, \]
we find
\[
\max_{x, y \in \mathbb{S}^d} T_{n, \omega}(x, y) \geq \max_{i=0, \ldots, d} T_{n, \omega}(e_i, e_i) = \max_{i=0, \ldots, d} \omega_{ne_i}.
\]
This tells that the estimate is sharp whenever \(\max_{|\alpha| = n} \omega_{n\alpha}\) is attained at \(\alpha \in nV_d\). An example will be given in Section 3.2.

2.2. The reproducing kernel Hilbert space

Consider the span of the functions
\[
K_x := T_{n, \omega}(x, \cdot), \quad x \in \mathbb{S}^d,
\]
and the inner product
\[
\left\langle \sum_{i=1}^N c_i K_{x_i} \left| \sum_{j=1}^M d_j K_{y_j} \right. \right\rangle := \sum_{i=1}^N \sum_{j=1}^M c_i T_{n, \omega}(x_i, y_j) d_j.
\]
(2.6)
The span is \(\mathbb{P}_n^d\), and we arrive at the space of algebraic polynomials of degree at most \(n\), imposed with an inner product structure such that the kernel \(T_{n, \omega}\) is a reproducing kernel. Note that since in our case the span is finite dimensional, we do not have to perform the completion process which is usual in such constructions.

**Definition 2.3.** By \(\mathcal{H}_{n, \omega}\) we denote the space \(\mathbb{P}_n^d\) equipped with the inner product (2.6)—which we denote by \(\langle \cdot | \cdot \rangle_{\mathcal{H}_{n, \omega}}\)—and with the corresponding norm \(\|f\|_{\mathcal{H}_{n, \omega}} := \sqrt{\langle f | f \rangle_{\mathcal{H}_{n, \omega}}}\).

The reproducing kernel property refers to the simple identity
\[
f(x) = \langle f | K_x \rangle_{\mathcal{H}_{n, \omega}} = \langle f | T_{n, \omega}(x, \cdot) \rangle_{\mathcal{H}_{n, \omega}}
\]
for \(f \in \mathbb{P}_n^d\) and \(x \in \mathbb{S}^d\).

**Theorem 2.4.** The space \(\mathcal{H}_{n, \omega}\) is compactly embedded into \(C(\mathbb{S}^d)\), with bound
\[
\|f\|_{C(\mathbb{S}^d)} \leq \sqrt{\Omega_n} \|f\|_{\mathcal{H}_{n, \omega}}.
\]
**Proof.** The embedding property is clear from the fact that the space \(\mathcal{H}_{n, \omega}\) has finite dimension. More interesting is the explicit expression for the bound. For reader’s convenience, we repeat the proof given in [6]. We have
\[
|f(x)| = \left| \langle f | K_x \rangle_{\mathcal{H}_{n, \omega}} \right| \leq \|f\|_{\mathcal{H}_{n, \omega}} \|K_x\|_{\mathcal{H}_{n, \omega}} = \|f\|_{\mathcal{H}_{n, \omega}} \sqrt{T_{n, \omega}(x, x)}
\]
for \(x \in \mathbb{S}^d\), and the result follows from Lemma 2.2.

2.3. Orthogonality of the Bernstein basis polynomials

The next result is our main result in this section. It shows that the new structure of the polynomial space results in an orthogonality property for the Bernstein basis functions.
Theorem 2.5. If $\omega_{\alpha} > 0$ for all $\alpha$ with $|\alpha| = n$, then the normalized Bernstein basis polynomials
\[ \tilde{B}_\alpha := \sqrt{\omega_{\alpha}} B_\alpha, \quad |\alpha| = n, \]
are orthonormal in $\mathcal{H}_{n,\omega}$, i.e.,
\[ \langle \tilde{B}_\alpha | \tilde{B}_\beta \rangle_{\mathcal{H}_{n,\omega}} = \delta_{\alpha,\beta} \quad \text{for} \quad |\alpha| = |\beta| = n. \]

Proof. We consider the collocation matrix
\[ M = (\tilde{B}_\alpha(x_\gamma))_{|\alpha|=|\gamma|=n} \]
for the discrete simplicial set of nodes $x_\gamma = \frac{1}{n} \gamma$, with $\gamma \in \mathbb{N}_0^{d+1}$ and $|\gamma| = n$. Here, $\alpha$ refers to columns and $\gamma$ refers to rows of the matrix, for some chosen ordering of the multi-indices (e.g., lexicographic). The matrix $M$ is regular, since
\[ M = (B_\alpha(x_\gamma))_{|\alpha|=|\gamma|=n} D_\omega, \]
where the right-hand factor is the diagonal matrix with diagonal entries $\sqrt{\omega_{\alpha}}$, $|\alpha| = n$, and the left-hand factor is the collocation matrix using the Bernstein basis functions of degree $n$ and the discrete simplicial nodes of order $n$ as interpolation points. It is well known that this collocation matrix is regular.

Denote
\[ \tilde{B}(y) := \begin{pmatrix} \tilde{B}_\alpha(y) \\ \vdots \\ \tilde{B}_{|\alpha|=n}(y) \end{pmatrix} \quad \text{and} \quad C(y) := \begin{pmatrix} \vdots \\ K_{x_\gamma}(y) \\ \vdots \end{pmatrix}_{|\gamma|=n}. \]

The basis transformation formula $C(y) = M \tilde{B}(y)$ yields the identity
\[ \left( T_{n,\omega}(x_\gamma, x_{\gamma'}) \right)_{|\gamma|=|\gamma'|=n} = \left( \cdots C(x_{\gamma'}) \cdots \right)_{|\gamma'|=n} = M \left( \cdots \tilde{B}(x_{\gamma'}) \cdots \right)_{|\gamma'|=n} = MM^T. \]

Since $\tilde{B}(y) = M^{-1} C(y)$, we conclude (with $M^{-T} := (M^{-1})^T$):
\[ \left( \tilde{B}_\alpha | \tilde{B}_\beta \rangle_{\mathcal{H}_{n,\omega}} \right)_{|\alpha|=|\beta|=n} = (\tilde{B}|\tilde{B}^T)_{\mathcal{H}_{n,\omega}} = M^{-1} \langle C|C^T \rangle_{\mathcal{H}_{n,\omega}} M^{-T} = M^{-1} \left( T_{n,\omega}(x_\gamma, x_{\gamma'}) \right)_{|\gamma|=|\gamma'|=n} M^{-T} = M^{-1}MM^T = I. \]

This shows that the Gramian of the normalized Bernstein basis polynomials with respect to the inner product in $\mathcal{H}_{n,\omega}$ is the identity matrix $I$.  

In terms of the (not normalized) Bernstein basis polynomials, the orthogonality property reads as follows:
\[ \{B_\alpha | B_\beta \} = \delta_{\alpha,\beta} \frac{1}{\omega_{\alpha}} \quad \text{for} \quad |\alpha| = |\beta| = n. \quad (2.7) \]
Therefore,

\[ \langle p | q \rangle_{\mathcal{H}_{n,\omega}} = \sum_{|\alpha|=n} \frac{c_{\alpha} d_{\alpha}}{\omega_{\alpha}} \]

for \( p = \sum_{|\alpha|=n} c_{\alpha} B_\alpha \) and \( q = \sum_{|\alpha|=n} d_{\alpha} B_\alpha \).

This shows an interesting connection to the inner product used in Theorem 1 and Corollary 1 of [2], see also [10] for the unweighted case. Denote \( D_{ij} := \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_i} \) for \( 1 \leq i < j \leq d \), \( D_{0j} := \frac{\partial}{\partial x_j} \) for \( 1 \leq j \leq d \), and \( D^k := \prod_{0 \leq i < j \leq d} D_{ij}^{k_{ij}} \), where \( k_{ij} \in \mathbb{N}_0 \) and \( k = (k_{ij})_{0 \leq i < j \leq d} \). Further, denote \( X^k := \prod_{0 \leq i < j \leq d} (x_i x_j)^{k_{ij}} \). It was proved in [10] that

\[ \sum_{|\alpha| = |\beta| = n} \frac{(n - |k|)!}{n! k!} X^k D^k B_\alpha(x) D^k B_\beta(x) = \delta_{\alpha,\beta} B_\alpha(x) \]

for \( |\alpha| = |\beta| = n \) and \( x \in \mathbb{S}^d \). Integrating this identity with respect to the measure \( \rho \), we obtain

\[ \sum_{|\alpha| = |\beta| = n} \frac{(n - |k|)!}{n! k!} \int_{\mathbb{S}^d} X^k D^k B_\alpha(x) D^k B_\beta(x) \, d\rho(x) = \delta_{\alpha,\beta} \langle 1 | B_\alpha \rangle_\rho. \]

Comparing this with (2.7), we arrive at the following statement.

**Corollary 2.6.** Let \( \omega_\alpha \) be the canonical weights as defined in (2.3). Then for all polynomials \( p, q \in \mathbb{P}_n^d \) we have

\[ \langle p | q \rangle_{\mathcal{H}_{n,\omega}} = \sum_{|\alpha| \leq n} \frac{(n - |k|)!}{n! k!} \int_{\mathbb{S}^d} X^k D^k p(x) D^k q(x) \, d\rho(x). \]

If \( d\rho(x) = w(x) \, dx \) with a smooth weight function \( w \) which satisfies certain conditions on the boundary of the simplex \( \mathbb{S}^d \), one can integrate the last identity by parts, as we did in Corollary 1 of [2] for the case of Jacobi weights. We omit the details. This identification of the inner product shows that \( \langle \cdot | \cdot \rangle_{\mathcal{H}_{n,\omega}} \) can be extended to spaces of smooth functions using appropriate differential operators. This gives a hint to a possible construction of a class of ‘Sobolev type’ spaces on the simplex.

### 3. The integral operators

Given the \( \omega \)-weighted Bernstein basis kernel of Definition 2.1, we consider the corresponding integral operator \( L_{n,\omega,\rho} \) defined in (2.1) as a linear operator mapping functions \( f \in L^q_\rho(\mathbb{S}^d) \) onto polynomials \( p \in \mathbb{P}_n^d \). Since the kernel is positive and continuous, we have

\[
\|L_{n,\omega,\rho}\|_{L^1(\mathbb{S}^d) \rightarrow L^1_\rho(\mathbb{S}^d)} = \sup_{f \in L^1_\rho(\mathbb{S}^d), \|f\|_{L^1_\rho(\mathbb{S}^d)} = 1} \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} T_{n,\omega,\rho}(x, y) \, f(y) \, d\rho(y) \, d\rho(x) = \sup_{f \in L^1_\rho(\mathbb{S}^d), \|f\|_{L^1_\rho(\mathbb{S}^d)} = 1} \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} T_{n,\omega,\rho}(x, y) \, |f(y)| \, d\rho(y) \, d\rho(x).
\]
\[
\begin{align*}
\sup_{f \in L^1_\rho(S^d), \|f\|_{L^1_\rho(S^d)} = 1} \left( \sum_{|\alpha| = n} \omega_\alpha \int_{S^d} B_\alpha(x) \, d\rho(x) \right) \int_{S^d} |f(y)| \, d\rho(y) \\
\leq A_{n,\omega,\rho} := \max_{|\alpha| = n} \omega_\alpha \int_{S^d} B_\alpha(x) \, d\rho(x).
\end{align*}
\]

On the other hand,
\[
\|L_{n,\omega,\rho}\|_{L^\infty_\rho(S^d) \rightarrow L^\infty_\rho(S^d)} = \sup_{f \in L^\infty_\rho(S^d), \|f\|_{L^\infty_\rho(S^d)} = 1} \left| \int_{S^d} T_{n,\omega,\rho}(x, y) \, f(y) \, d\rho(y) \right|
\leq \int_{S^d} \sum_{|\alpha| = n} \omega_\alpha B_\alpha(x) \, d\rho(x)
\leq A_{n,\omega,\rho}.
\]

Using the Riesz–Thorin interpolation theorem and taking into account (2.4), we find

**Lemma 3.1.** For general \(\omega\)-weights and \(1 \leq q \leq \infty\), the operators (2.1) are bounded as operators from \(L^q_\rho(S^d)\) into \(L^q_\rho(S^d)\), and
\[
\|L_{n,\omega,\rho}\|_{L^q_\rho(S^d) \rightarrow L^q_\rho(S^d)} \leq A_{n,\omega,\rho}.
\]

If \(\omega_\alpha\) are the canonical weights as in (2.3), we have
\[
\|M_{n,\rho}\|_{L^q_\rho(S^d) \rightarrow L^q_\rho(S^d)} = 1. \tag{3.2}
\]

### 3.1. Spectral properties of the integral operator

In this section we study spectral properties of the operator
\[
L_{n,\omega,\rho} : L^2_\rho(S^d) \rightarrow \mathcal{H}_{n,\omega}.
\]

Here both the domain and the range of the operator carry an inner product structure with inner products (1.3) and (2.6), respectively. Most results of this section follow from those of papers [6,14], where the general theory of compact integral operators was applied to similar problems in a more general context. However, since our operators have finite rank, one can proceed more directly using elementary linear algebra, and we are going to do so. In particular, we will give full proofs.

In this section, we assume that all weights \(\omega_\alpha\), \(|\alpha| = n\), are strictly positive. We use the notation \(\mathbf{B}\) for the vector of the normalized Bernstein basis polynomials of degree \(n\), as in the proof of Theorem 2.5. The kernel of the integral operator can be written as
\[
T_{n,\omega}(x, y) = \mathbf{B}^T(x) \mathbf{B}(y).
\]

From this we obtain the following representation of the operator:
\[
L_{n,\omega,\rho}(f) = \mathbf{B}^T(\mathbf{B} f)_\rho. \tag{3.4}
\]

The right-hand factor here is a column vector where the inner product is taken componentwise. The action of the operator on \(L^2_\rho(S^d)\) can now be easily understood. The orthogonal complement \((\mathbb{P}^d)^\perp\) of \(\mathbb{P}^d\) with respect to the \(\rho\)-product (1.3) is in the kernel of the operator, and the action on \(\mathbb{P}^d\) will be described by specializing \(f\) to a polynomial \(p = \mathbf{B}^T c\), where \(c\) is a vector of
coefficients. Note that, due to the positivity of the weights $\omega_\alpha$, the components of $\widetilde{B}$ form a basis of $\mathbb{P}_n^d$. Now
\[
L_{m,\omega,\rho}(\widetilde{B}^T c) = \widetilde{B}^T G c,
\]
where
\[
G = (\langle \tilde{B}_\alpha | \tilde{B}_\beta \rangle_\rho)_{|\alpha|=|\beta|=n}
\]
is the Gramian of the normalized Bernstein basis polynomials with respect to the $\rho$-product. The spectral decomposition of the Gramian can be used to characterize the spectral properties of the integral operator. Here, the following well-known result will be useful.

**Lemma 3.2.** For $p_1, \ldots, p_N \in L^2_\rho(S^d)$, let $G_k, k = 1, \ldots, N$, denote the Gramian of the system $\{p_1, \ldots, p_k\}$. Then
\[
det G_{k+1} = E^2_{k,\rho}(p_{k+1}) \cdot det G_k, \quad k = 1, \ldots, N - 1,
\]
where
\[
E^2_{k,\rho}(p_{k+1}) := \min_{p \in P_k} \| p_{k+1} - p \|^2_{L^2_\rho(S^d)}, \quad P_k := \text{span}\{p_1, \ldots, p_k\}.
\]
In particular, if $\{p_1, \ldots, p_N\}$ is a linearly independent system of polynomials, then all these determinants are strictly positive.

**Proof.** For the sake of completeness, we give a proof. By the normal equations, the best approximation $p = \sum_{\ell=1}^k a_{\ell,k} p_\ell \in P_k$ to the function $p_{k+1}$ is characterized by the fact that $p_{k+1} - p$ is orthogonal to $P_k$ with respect to the inner product $(1.3)$. If we subtract the $a_{\ell,k}$-multiple of the $\ell$th column, and then the same multiple of the $\ell$th row, $\ell = 1, \ldots, k$, from the last column (the last row, respectively) of $G_{k+1}$, the non-diagonal entries in the last column (and the last row) will vanish, except for the diagonal entry which becomes $\|p_{k+1} - \sum_{\ell=1}^k a_{\ell,k} p_\ell \|^2_{L^2_\rho(S^d)} = E^2_{k,\rho}(p_{k+1})$. From this the recurrence formula for the determinants is clear.

The additional statement follows from the positivity assumption. We have $det G_1 = \|p_1\|^2_{L^2_\rho(S^d)} > 0$ and $E^2_{k,\rho} > 0, k = 1, \ldots, N - 1$. ■

The Gramian (3.6) is regular, hence positive definite, if the weights $\omega_\alpha$ are strictly positive. We choose a set of eigenpairs for $G$,
\[
(\lambda_\gamma, v_\gamma), \quad |\gamma| = n,
\]
with (real) eigenvalues
\[
\lambda_\gamma > 0, \quad |\gamma| = n,
\]
and eigenvectors $v_\gamma \in \mathbb{R}^N, N = \left(\frac{n+d}{d}\right)$, which we assume to be orthonormalized, i.e.,
\[
v_\gamma^T v_{\gamma'} = \delta_{\gamma,\gamma'}, \quad |\gamma| = |\gamma'| = n.
\]
The polynomials
\[
p_{\gamma} = \widetilde{B}^T v_\gamma, \quad |\gamma| = n,
\]
then provide a system of eigenfunctions for our integral operator.
Theorem 3.3. Let $\omega_{\alpha} > 0$ for all $\alpha$ with $|\alpha| = n$. Then the polynomials (3.7) are eigenfunctions of the integral operator (2.1) satisfying

$$L_{n,\omega,\rho}(p) = \lambda_{p} \ p, \quad |p| = n,$$

and

$$(p|p')_{\rho} = \lambda_{p} \ \delta_{p,p'}, \quad |p| = |p'| = n.$$

Moreover, the restriction of the operator to $\mathbb{P}_n^d$ is an isomorphism.

The map

$$\tilde{B}_p \leftrightarrow \tilde{p} := \frac{1}{\sqrt{\lambda_{p}}} \ p, \quad |p| = n,$$

provides an isometry between the two inner product spaces $(\mathbb{P}_n^d, \langle \cdot | \cdot \rangle_{H_{n,\omega}})$ and $(\mathbb{P}_n^d, \langle \cdot | \cdot \rangle_{\rho})$, since

$$\delta_{\alpha,\beta} = \langle \tilde{B}_\alpha | \tilde{B}_\beta \rangle_{H_{n,\omega}} = \langle \tilde{p}_\alpha | \tilde{p}_\beta \rangle_{\rho}, \quad |\alpha| = |\beta| = n.$$

Next we prove one further property of the integral operator.

Corollary 3.4. If $\omega_{\alpha} > 0$ for all $\alpha$ with $|\alpha| = n$, then

$$\langle L_{n,\omega,\rho}(f)|p \rangle_{H_{n,\omega}} = \langle f|p \rangle_{\rho}$$

for $f \in L_{p}^2(S^d)$ and $p \in \mathbb{P}_n^d$.

Proof. With $p = \tilde{B}^T c = \sum_{|\alpha|=n} c_{\alpha} \tilde{B}_{\alpha}$, formula (3.4) yields

$$\langle L_{n,\omega,\rho}(f)|p \rangle_{H_{n,\omega}} = \langle \tilde{B}^T \tilde{B} | f \rangle_{\rho} = \sum_{|\beta|=n} \sum_{|\alpha|=n} c_{\beta} \langle \tilde{B}_{\beta}|f\rangle_{\rho} \langle \tilde{B}_{\alpha}|\tilde{B}_{\beta}\rangle_{H_{n,\omega}} c_{\alpha}$$

$$= \sum_{|\alpha|=n} c_{\alpha} \langle \tilde{B}_{\alpha}|f\rangle_{\rho} = \langle p|f \rangle_{\rho},$$

due to the orthogonality of the normalized Bernstein basis polynomials.

Next we discuss the norm of operator $L_{n,\omega,\rho}$ as an operator from $L_{p}^2(S^d)$ into $H_{n,\omega}$.

Theorem 3.5. If $\omega_{\alpha} > 0$ for all $\alpha$ with $|\alpha| = n$, then the operator (3.3) has the norm

$$\|L_{n,\omega,\rho}\|_{L_{p}^2(S^d) \rightarrow H_{n,\omega}} = \sqrt{\lambda(G)},$$

where $\lambda(G) := \max_{|p| = n} \lambda_{p}$ is the spectral radius of the Gramian (3.6).

Proof. We have, with $p = \tilde{B}^T c \in \mathbb{P}_n^d$, where $c \in \mathbb{R}^N, \ N = \begin{pmatrix} n + d \\ d \end{pmatrix}$,

$$\|L_{n,\omega,\rho}\|_{L_{p}^2(S^d) \rightarrow H_{n,\omega}} = \max_{0 \neq p \in \mathbb{P}_n^d} \frac{\|L_{n,\omega,\rho}(p)\|_{H_{n,\omega}}^2}{\|p\|_{L_{p}^2(S^d)}^2} = \max_{0 \neq c \in \mathbb{R}^N} \frac{c^T G^2 c}{c^T G c}$$

using (3.5) and (3.6). The latter Rayleigh quotient is bounded by the spectral radius of the Gramian, and this bound is sharp.
In the case of the canonical weights we have $\lambda(G) = 1$, since the corresponding $\rho$-weighted Bernstein–Durrmeyer operator is contractive and reproduces constant functions; see (3.2) and (2.4), respectively. Thus, we obtain

**Corollary 3.6.** For the $\rho$-weighted Bernstein–Durrmeyer operator we have

$$\|M_{n,\rho}\|_{L_\rho^2(S^d) \to \mathcal{H}_{n,\omega}} = 1.$$  

In the general setting we arrive at the difficult problem of finding or estimating the spectrum of $G$. However, an easy estimate can be derived as follows. $G$ is a positive definite matrix, with trace

$$\text{trace } G = \sum_{|\alpha| = n} \omega_\alpha \int_{S^d} B_\alpha^2(x) \, d\rho(x) = \int_{S^d} T_{n,\omega}(x, x) \, d\rho(x).$$

By Lemma 2.2 we thus obtain

**Corollary 3.7.** The operator norm in Theorem 3.5 is bounded by

$$\|L_{n,\omega,\rho}\|_{L_\rho^q(S^d) \to \mathcal{H}_{n,\omega}} \leq \sqrt{\Omega_n},$$

with $\Omega_n$ as in (2.5).

**Remark.** The considerations of this section can be also carried out in the case where some of the weights $\omega_\alpha$ vanish. In Theorem 3.3, one has to decompose $L_\rho^2(S^d)$ into the space $\text{span}\{\tilde{B}_\alpha : |\alpha| = n\}$ and its orthogonal complement, which is the kernel of the operator. In Theorem 3.5, we replace $N$ by $N' = \dim \text{span}\{\tilde{B}_\alpha : |\alpha| = n\}$ and $G$ by $G'$, the Gramian of non-vanishing $\tilde{B}_\alpha$'s. In this case, we have $\lambda(G) = \lambda(G')$.

### 3.2. The Jacobi-weighted Bernstein–Durrmeyer operator

In this subsection we consider the Bernstein–Durrmeyer operator (1.9) with respect to the Jacobi weight (1.6). Due to the special structure of the weights $\omega$, the operator $M_{n,\mu}$ has properties that are not valid in the general setting.

Recall that $M_{n,\mu}$ is a positive polynomial operator on $L_\mu^1(S^d)$, with norm

$$\|M_{n,\mu}\|_{L_\mu^q(S^d) \to L_\mu^q(S^d)} = 1, \quad 1 \leq q \leq \infty,$$

which reproduces constant functions.

#### 3.2.1. Sharpness of estimate (2.5)

Estimate (2.5) for the kernel of the operator $M_{n,\mu}$ is sharp whenever the exponents $\mu_i$ in (1.6) are all non-negative, except possibly one. This follows from the following

**Lemma 3.8.** If $\omega$ are the canonical weights as in (2.3) with respect to the Jacobi weight (1.6), and if the exponents $\mu_i$ in (1.6) are all non-negative, except at most one, then $\max_{|\alpha| = n} \omega_\alpha$ is attained at $\alpha \in nV_d$.

**Proof.** According to (1.8),

$$\omega^{-1}_\alpha = \langle 1 | B_\alpha \rangle_\mu = \frac{1}{B(\mu + e)} \frac{n!}{\Gamma(n + |\mu| + d + 1)} \prod_{i=0}^{d} \frac{\Gamma(\alpha_i + \mu_i + 1)}{\alpha_i!}.$$
Denote
\[
\gamma_{\alpha} := \prod_{i=0}^{d} \frac{\Gamma(\alpha_i + \mu_i + 1)}{\alpha_i!} = \frac{\Gamma \left( n - \sum_{i=1}^{d} \alpha_i + \mu_0 + 1 \right)}{\left( n - \sum_{i=1}^{d} \alpha_i \right)!} \prod_{i=1}^{d} \frac{\Gamma(\alpha_i + \mu_i + 1)}{\alpha_i!}.
\]

We need to show that, under the assumptions of the lemma, \(\min_{|\alpha|=n} \gamma_{\alpha}\) is attained at \(\alpha \in nV_d\).

First we prove that \(\min_{|\alpha|=n} \gamma_{\alpha}\) is attained on the boundary of the simplex \(nS^d\). We switch to the Cartesian coordinates \(\alpha = (\alpha_1, \ldots, \alpha_d)\). Let \(j \in \{1, \ldots, d\}\). For the forward difference of \(\gamma_{\alpha}\) in the direction \(e_j\) we have
\[
\Delta_j \gamma_{\alpha} := \gamma_{\alpha+e_j} - \gamma_{\alpha} = \left( \frac{\alpha_j + \mu_j + 1}{\alpha_j + 1} - \frac{n - \sum_{i=1}^{d} \alpha_i + \mu_0}{n - \sum_{i=1}^{d} \alpha_i} \right) \prod_{i=1}^{d} \frac{\Gamma(\alpha_i + \mu_i + 1)}{\alpha_i!} \frac{\Gamma \left( n - \sum_{i=1}^{d} \alpha_i + \mu_0 \right)}{\left( n - \sum_{i=1}^{d} \alpha_i - 1 \right)!} = \left( \mu_j \left( n - \sum_{i=1}^{d} \alpha_i \right) - \mu_0(\alpha_j + 1) \right) \prod_{i=1}^{d} \frac{\Gamma(\alpha_i + \mu_i + 1)}{\alpha_i!} \frac{\Gamma \left( n - \sum_{i=1}^{d} \alpha_i + \mu_0 \right)}{(\alpha_j + 1) \left( n - \sum_{i=1}^{d} \alpha_i \right)!}.
\]

We see that each \(\Delta_j \gamma_{\alpha}\) is a linear function of \(\alpha\) times a positive term. Thus, \(\gamma_{\alpha}\) changes the monotonicity at most once in each variable. We first consider the case when \(|\mu| = 0\). Then there is an index \(j\) with \(1 \leq j \leq d\) such that \(\mu_j \mu_0 \leq 0\). But then \(\Delta_j \gamma_{\alpha}\) has constant sign. Thus, \(\gamma_{\alpha}\) is monotone with respect to \(\alpha_j\) and cannot have a local minimum. Consequently, the minimum is attained on the boundary of \(nS^d\).

Now suppose \(|\mu| \neq 0\). Consider the system of linear equations
\[
\mu_j \left( n - \sum_{i=1}^{d} \alpha_i \right) - \mu_0(\alpha_j + 1) = 0, \quad j = 1, \ldots, d.
\]

If \(|\mu| \neq 0\), this system has the unique solution \(\overline{\alpha} = (\overline{\alpha}_1, \ldots, \overline{\alpha}_d)\) with \(\overline{\alpha}_j = \frac{\mu_j}{|\mu|}(n + d) - 1, \quad j = 1, \ldots, d\). This point provides a local minimum only if the coefficients of \(\alpha_j\) in \(\Delta_j \gamma_{\alpha}\), which are equal to \(-\mu_j - \mu_0\), are positive for every \(j = 1, \ldots, d\), i.e., if
\[
\mu_j + \mu_0 < 0, \quad j = 1, \ldots, d. \tag{3.8}
\]

The minimum is attained strictly inside the simplex only if \(\overline{\alpha}_j > 0, \quad j = 1, \ldots, d\), and \(\sum_{i=1}^{d} \overline{\alpha}_j < n - 1\), which is equivalent to
\[
\frac{\mu_j}{|\mu|}(n + d) > 1, \quad j = 0, \ldots, d. \tag{3.9}
\]

It is not difficult to see that the latter is only possible if all the exponents \(\mu_j, \ j = 0, \ldots, d\) are negative. Indeed, if we suppose that \(|\mu| > 0\), we get from (3.9) that \(\mu_j > \frac{|\mu|}{n+d} > 0, \ j = 0, \ldots, d\), which contradicts (3.8). Thus, \(|\mu| < 0\), but then we get from (3.9) that \(\mu_j < \frac{|\mu|}{n+d} < 0, \ j = 0, \ldots, d\), as desired. At this point we have shown that \(\min_{|\alpha|=n} \gamma_{\alpha}\) is attained at the boundary
of $n \mathbb{S}^d$ whenever at least one of the exponents $\mu_i$, $0 \leq i \leq d$, is non-negative. On the other hand, if the exponents $\mu_i$, $0 \leq i \leq d$, are all negative, it follows from (3.9) that $\min_{|\alpha|=n} \gamma_\alpha$ will be attained strictly inside the simplex $n \mathbb{S}^d$ for $n > \max_{j=0,\ldots,d} \frac{|\alpha|}{\mu_i} - d$, so that the assumption about negativity of the exponents cannot be relaxed for all $n$.

Now let us consider a face of dimension $d-1$, on which the minimum is attained. Without loss of generality, we may assume that this is the face with the equation $\alpha_d = 0$. Repeating the above considerations for the simplex $n \mathbb{S}^{d-1}$ with the barycentric coordinates $\tilde{\alpha} = (\alpha_0, \ldots, \alpha_{d-1})$ and the weight $\tilde{\mu} = (\mu_0, \ldots, \mu_{d-1})$, we see that the minimum is attained on the boundary of $n \mathbb{S}^{d-1}$ whenever at least one of the exponents $\mu_0, \ldots, \mu_{d-1}$ is non-negative. Repeating the process, we obtain the statement of the lemma.  

**Remark.** The assumptions of the lemma cannot be relaxed. As an example, take $d = 2$, $n = 2$, $\mu_0 = \mu_1 = -1/2$, and $\mu_2 = 1/2$. Then $\gamma(0,0,2) = \frac{15}{16} \pi^2$, $\gamma(1,0,1) = \gamma(0,1,1) = \frac{3}{8} \pi^2$, $\gamma(2,0,0) = \gamma(0,2,0) = \frac{3}{16} \pi^2$, and $\gamma(1,1,0) = \frac{1}{8} \pi^2$. Thus, $\min_{|\alpha|=n} \gamma_\alpha$ is attained at $\alpha = (1, 1, 0) \not\in 2 \mathbb{V}_2$.

3.2.2. *Orthogonal polynomials as eigenfunctions*

The Bernstein–Durrmeyer operator $M_{n, \mu}$ has remarkable spectral properties. Because of their importance, we are going to discuss them in detail. All results of this section, except for the remark below, are well known, e.g., [7,8,5,9].

An important property of the Bernstein–Durrmeyer operator with respect to Jacobi weights is that this operator is *degree preserving*. This means that, for a monomial $\varphi_\beta(x) := x^\beta$ with $|\beta| = m \leq n$, we have

$$M_{n, \mu} (\varphi_\beta) = C_\beta \varphi_\beta + p_{m-1}, \quad (3.10)$$

where $p_{m-1} \in \mathbb{P}_m$ and $C_\beta \neq 0$ is a constant. In particular,

$$M_{n, \mu} (\mathbb{P}_m) \subset \mathbb{P}_m, \quad m = 0, 1, \ldots, n. \quad (3.11)$$

One can see this, e.g., as follows. A direct computation shows that the monomials $\varphi_\beta(x)$, $|\beta| = m \leq n$, are mapped by $M_{n, \mu}$ onto the polynomials

$$M_{n, \mu} (\varphi_\beta) = \sum_{|\alpha|=n} c_{\alpha, \beta} B_\alpha$$

with

$$c_{\alpha, \beta} = \frac{B(\alpha + \beta + \mu + \epsilon)}{B(\alpha + \mu + \epsilon)} = \frac{\Gamma(n + |\mu| + d + 1)}{\Gamma(n + m + |\mu| + d + 1)} \prod_{i=0}^d \frac{\Gamma(\alpha_i + \beta_i + \mu_i + 1)}{\Gamma(\alpha_i + \mu_i + 1)}.$$ 

Thus,

$$c_{\alpha, \beta} = q_\beta(\alpha), \quad |\alpha| = n, \quad |\beta| = m \leq n,$$

where $q_\beta$ is a polynomial in $\alpha$ of coordinate degree $\beta$. It follows by a well-known property of Bernstein–Bézier coefficients of a polynomial, cf. Proposition 2.1 in [15] and its proof, that $M_{n, \mu} (\varphi_\beta)$ is a polynomial of coordinate degree $\beta$, which proves (3.10).

Let

$$E_{0, \mu} := \mathbb{P}_0^d; \quad E_{m, \mu} := \mathbb{P}_m^d \cap (\mathbb{P}_{m-1}^d)^\perp, \quad m > 0.$$
denote the spaces of polynomials orthogonal with respect to the inner product (1.7). It follows from the property (3.11) and the symmetry property (2.2) that the kernel of $M_{n,\mu}$ is $(E_{n,\mu})^\perp$, and that the spaces $E_{m,\mu}$, $m = 0, \ldots, n$, are eigenspaces of the operator $M_{n,\mu}$. The corresponding eigenvalues are (e.g., [9])

$$\lambda_{n,m,\mu} = \frac{\binom{n}{m}}{n+|\mu|+d+m}, \quad 0 \leq m \leq n. \tag{3.12}$$

The detailed knowledge of spectral properties of the operators was extensively used for studying approximation properties of $M_{n,\mu}$ like, e.g., in [5] or [3, Theorem 7]. It allows us also to elaborate on results of Section 3.1. Thus, the spectrum of the Gramian (3.6) is given by the values (3.12), where $\lambda_{n,m,\mu}$ has multiplicity 1 for $m = 0$ and multiplicity

$$\dim E_{m,\mu} = \binom{m+d}{m} - \binom{m-1+d}{m-1} = \binom{m+d-1}{m}$$

for $m = 1, \ldots, n$.

**Remark.** Neither the degree preservation property nor the fact that the spaces of orthogonal polynomials are eigenspaces of the operator are valid for the Bernstein–Durrmeyer operator in the general setting (1.4). As an example, consider the measure defined in (3.13). Then the polynomial $p_2(x) = x_1 - x_2$ is an eigenfunction of $M_{1,\rho}$ but not of $M_{2,\rho}$.

A direct calculation shows that the polynomials $p_2$ and $p_5$ are not orthogonal polynomials, i.e., they are not orthogonal to all polynomials of lower degree.

In the case of Jacobi weights, it follows from the spectral properties of the Bernstein–Durrmeyer operator that if $p$ is an eigenfunction of $M_{n,\mu}$ then $p$ is an eigenfunction of $M_{n+1,\mu}$ as well. This property fails to hold in the general setting (1.4), too. As an example, consider again the measure defined in (3.13). Then the polynomial $p = x_1 - x_2$ is an eigenfunction of $M_{1,\rho}$ but not of $M_{2,\rho}$.
4. Convergence

4.1. Convergence for test functions and Korovkin type theorems

We come back to the general case of the operator \( L_{n,\omega,\rho} \) under Assumptions P and N. The \( \rho \)-weighted Bernstein–Durrmeyer operator (1.4) reproduces constant functions, see (2.4). The integral operators (2.1) in general do not reproduce constant functions. We have

\[
1 - L_{n,\omega,\rho}(1) = \sum_{|\alpha|=n} \left( 1 - \omega_\alpha \int_{\mathbb{S}^d} B_\alpha(x) \, d\rho(x) \right) B_\alpha,
\]

and thus the condition

\[
\lim_{n \to \infty} \max_{|\alpha|=n} \left| \omega_\alpha \int_{\mathbb{S}^d} B_\alpha(x) \, d\rho(x) - 1 \right| = 0
\]

implies the convergence

\[
\lim_{n \to \infty} \|1 - L_{n,\omega,\rho}(1)\|_{C(\mathbb{S}^d)} = 0.
\]

Next we give a sufficient condition for convergence in \( C(\mathbb{S}^d) \) for linear functions. Let \( e_i \in V_d \). We consider the linear test functions

\[
\varphi_{e_i}(x) = x^{e_i} = \begin{cases} x_i, & i = 1, \ldots, d, \\ 1 - x_1 - \cdots - x_d, & i = 0. \end{cases}
\] (4.1)

For \( i = 0, \ldots, d \) and \( |\alpha| = n \) define

\[
b_{\alpha,\omega,\rho}^{(i)} := \omega_\alpha \int_{\mathbb{S}^d} x^{e_i} B_\alpha(x) \, d\rho(x).
\] (4.2)

We will denote these coefficients by \( b_{\alpha,\rho}^{(i)} \) in the case of the \( \rho \)-weighted Bernstein–Durrmeyer operator, and by \( b_{\alpha,\mu}^{(i)} \) in the case of the Bernstein–Durrmeyer operator with the Jacobi weight (1.6).

**Lemma 4.1.** Let \( 0 \leq i \leq d \). The condition

\[
\lim_{n \to \infty} \max_{|\alpha|=n} \left| b_{\alpha,\omega,\rho}^{(i)} - \frac{\alpha_i}{n} \right| = 0
\] (4.3)

implies

\[
\lim_{n \to \infty} \|\varphi_{e_i} - L_{n,\omega,\rho}(\varphi_{e_i})\|_{C(\mathbb{S}^d)} = 0.
\] (4.4)

The proof is clear since, by (1.2),

\[
\varphi_{e_i} - L_{n,\omega,\rho}(\varphi_{e_i}) = \sum_{|\alpha|=n} \left( \frac{\alpha_i}{n} - b_{\alpha,\omega,\rho}^{(i)} \right) B_\alpha, \quad i = 0, \ldots, d.
\] (4.5)

**Remark.** In the case of the canonical weights, the coefficients \( b_{\alpha,\rho}^{(i)} \) are given by

\[
b_{\alpha,\rho}^{(i)} = \frac{\int_{\mathbb{S}^d} x^{e_i} B_\alpha(x) \, d\rho(x)}{\int_{\mathbb{S}^d} B_\alpha(x) \, d\rho(x)}.
\]
This allows an interesting interpretation of these coefficients as the expectation of the random variable $X_i$ with respect to the probability measure $\frac{B_\alpha(x) \, d\rho(x)}{\int_{\mathbb{S}^d} B_\alpha(x) \, d\rho(x)}$ supported on the simplex. Our assumption relates these measures to the Dirac measures supported at points $\frac{1}{n} \gamma$, $|\gamma| = n$.

Condition (4.3) is satisfied, for example, in the case of the Jacobi-weighted Bernstein–Durrmeyer operators (1.9). Indeed,

$$b_{\alpha,\mu}^{(i)} = \frac{\int_{\mathbb{S}^d} x_i \, B_\alpha(x) \, x^\mu \, dx}{\int_{\mathbb{S}^d} B_\alpha(x) \, x^\mu \, dx} = \frac{B(\alpha + \mu + e_i + e)}{B(\alpha + \mu + e)} = \frac{\alpha_i + \mu_i + 1}{n + |\mu| + d + 1}, \quad (4.6)$$

and thus

$$b_{\alpha,\mu}^{(i)} - \frac{\alpha_i}{n} = O\left(\frac{1}{n}\right) \quad \text{uniformly in } \alpha. \quad (4.7)$$

We will give a generalization in the next section. On the other hand, it is not difficult to construct an example of an integral operator, even with the canonical weights, for which convergence (4.4) fails. For simplicity, we consider the one-dimensional case, i.e., $d = 1$, $\mathbb{S}^1 = [0, 1]$. Take a number $a \in (0, 1)$, and consider the measure $d\rho(x) = w(x) \, dx$ with

$$w(x) = \begin{cases} \frac{1}{a}, & 0 \leq x \leq a, \\ \frac{a}{0}, & a < x \leq 1. \end{cases}$$

The corresponding Bernstein–Durrmeyer operator has the form

$$(M_{a,\rho}(f))(y) = \sum_{k=0}^{n} \int_0^a f(x) \, x^k \, (1-x)^{n-k} \, dx \, \left(\begin{array}{c} n \\ k \end{array}\right) y^k \, (1-y)^{n-k}, \quad y \in [0, 1].$$

Obviously, for the linear function $\varphi_{e_1}(x) = x$ we have $(M_{a,\rho}(\varphi_{e_1}))(y) \leq a$, $y \in [0, 1]$, so that convergence (4.4) cannot hold.

The example above shows, in particular, that $C(\mathbb{S}^d)$ is not the right space for studying convergence of the $\rho$-weighted Bernstein–Durrmeyer operators, and one should restrict the consideration to the support of the measure $\rho$. Numerical experiments show that convergence holds at least inside the support of $\rho$ for a large class of measures. An interesting question would be to determine conditions on the measure $\rho$ which guarantee convergence

$$\lim_{n \to \infty} \|f - M_{a,\rho}(f)\|_{L^1_\rho(\mathbb{S}^d)} = 0$$

for each $f \in L^1_\rho(\mathbb{S}^d)$.

For the Bernstein–Durrmeyer operator (1.4), a simple test for the convergence in $L^1_\rho(\mathbb{S}^d)$ is given by the following Korovkin type theorem.

**Theorem 4.2.** If condition (4.3) is fulfilled for each $i = 0, 1, \ldots, d$, then

$$\lim_{n \to \infty} \|f - M_{a,\rho}(f)\|_{L^1_\rho(\mathbb{S}^d)} = 0$$

for each $f \in L^1_\rho(\mathbb{S}^d)$.

**Proof.** Condition (4.3) imply convergence for the linear monomials $\varphi_{e_i}$, $i = 1, \ldots, d$, and for the linear function $\varphi_{e_0} = 1 - \sum_{i=1}^d \varphi_{e_i}$. Thus, convergence holds for the constant function $1$ as well.
In addition, the operators \( M_{n, \rho} \) are contractions; see (3.2). Convergence for each \( f \in L^1_\rho(\mathbb{S}^d) \) follows from a Korovkin type theorem of Berens and Lorentz for linear contractions on the space \( L^1 \); see Section 5 of [4].

To establish convergence in \( C(\mathbb{S}^d) \), one has to consider, in addition, quadratic functions

\[
\varphi_{2e_i}(x) = x^{2e_i} = x_i^2, \quad i = 1, \ldots, d.
\]

Taking into account that

\[
\varphi_{2e_i} = \sum_{|\alpha| = n} \frac{\alpha_i (\alpha_i - 1)}{n(n-1)} B_\alpha,
\]

we obtain the condition

\[
\lim_{n \to \infty} \max_{|\alpha| = n} \left| \omega_\alpha \int_{\mathbb{S}^d} x^{2e_i} B_\alpha(x) \, d\rho(x) - \frac{\alpha_i (\alpha_i - 1)}{n(n-1)} \right| = 0.
\] (4.8)

Now we can formulate the result for the general integral operators (2.1).

**Theorem 4.3.** If condition (4.3) is fulfilled for each \( i = 0, 1, \ldots, d \), and condition (4.8) is fulfilled for each \( i = 1, \ldots, d \), then

\[
\lim_{n \to \infty} \| f - L_{n, w, \rho}(f) \|_{C(\mathbb{S}^d)} = 0
\]

for each \( f \in C(\mathbb{S}^d) \).

### 4.2. A class of Jacobi-like measures

In this section we restrict our considerations to the Bernstein–Durrmeyer operators (1.4). We describe a class of measures \( \rho \) with the property that Theorem 4.3 is valid for the corresponding \( \rho \)-weighted Bernstein–Durrmeyer operator. This class generalizes the Jacobi weights and the class (1.10) considered by Păltănea.

**Theorem 4.4.** Suppose that the measure \( \rho \) has the form

\[
d\rho(x) = w(x) \, dx,
\]

and suppose that there are Jacobi exponents \( v \geq \mu > -e \) and constants \( 0 < a, A < \infty \) such that

\[
a \frac{x^v}{B(v + e)} \leq w(x) \leq A \frac{x^\mu}{B(\mu + e)}, \quad x \in \mathbb{S}^d.
\] (4.9)

If \( |v| - |\mu| < 1 \), then (4.3) is fulfilled for each \( i = 0, 1, \ldots, d \), (4.8) is fulfilled for each \( i = 1, \ldots, d \), and, moreover,

\[
\| \varphi_{e_i} - M_{n, \rho}(\varphi_{e_i}) \|_{C(\mathbb{S}^d)} = \mathcal{O} \left( n^{-\frac{1-|v|-|\mu|}{2}} \right), \quad n \to \infty,
\] (4.10)

\[
\| \varphi_{2e_i} - M_{n, \rho}(\varphi_{2e_i}) \|_{C(\mathbb{S}^d)} = \mathcal{O} \left( n^{-\frac{1-2|v|-|\mu|}{2}} \right), \quad n \to \infty.
\] (4.11)

In particular,

\[
\lim_{n \to \infty} \| f - M_{n, \rho}(f) \|_{C(\mathbb{S}^d)} = 0
\] (4.12)

for each \( f \in C(\mathbb{S}^d) \).
Proof. First we prove convergence for the linear test functions (4.1). The coefficients (4.2) in our case have the form
\[
b_{\alpha, \rho}^{(i)} = \frac{\int_{\mathcal{G}_d} x^{\alpha+\epsilon_i} d\rho(x)}{\int_{\mathcal{G}_d} x^\alpha d\rho(x)}.
\]
We will compare them with the coefficients \(b_{\alpha, \mu}^{(i)}\) for the Jacobi-weighted Bernstein-Durrmeyer operator with the weight \(\frac{x^\mu}{B(\mu+\epsilon)}\), which were calculated in (4.6). We have
\[
b_{\alpha, \rho}^{(i)} - b_{\alpha, \mu}^{(i)} = \frac{\int_{\mathcal{G}_d} \left(x_i - \frac{\alpha_i + \mu + 1}{n + |\mu| + d + 1}\right) x^\alpha d\rho(x)}{\int_{\mathcal{G}_d} x^\alpha d\rho(x)},
\]
and, using the Cauchy–Schwarz inequality and (4.9),
\[
\left(b_{\alpha, \rho}^{(i)} - b_{\alpha, \mu}^{(i)}\right)^2 \leq \frac{\int_{\mathcal{G}_d} \left(x_i - \frac{\alpha_i + \mu + 1}{n + |\mu| + d + 1}\right)^2 x^\alpha d\rho(x)}{\int_{\mathcal{G}_d} x^\alpha d\rho(x)} \leq \frac{A}{a} \frac{B(v + e)}{B(\mu + e)} \frac{\int_{\mathcal{G}_d} \left(x_i - \frac{\alpha_i + \mu + 1}{n + |\mu| + d + 1}\right)^2 x^{\alpha+\mu} d\mathbf{x}}{\int_{\mathcal{G}_d} x^{\alpha+\mu} d\mathbf{x}}.
\]
We will estimate the order of (4.13) as \(n \to \infty\). For the factor \(\frac{B(\alpha + \mu + e)}{B(\alpha + v + e)}\) we have
\[
\frac{B(\alpha + \mu + e)}{B(\alpha + v + e)} = \frac{\Gamma(n + |v| + d + 1)}{\Gamma(n + |\mu| + d + 1)} \prod_{i=0}^{d} \frac{\Gamma(\alpha_i + \mu_i + 1)}{\Gamma(\alpha_i + v_i + 1)} \sim n^{|v| - |\mu|} \prod_{i=0}^{d} \frac{1}{(\alpha_i + 1)^{v_i - \mu_i}},
\]
the latter follows from the formula \(\lim_{n \to \infty} n^{-a} \frac{\Gamma(n+a)}{\Gamma(n+b)} = 1\), e.g. (6.1.46) in [1]. Since \(\prod_{i=0}^{d} (\alpha_i + 1)^{v_i - \mu_i} \geq 1\) (and this estimate cannot be improved for all \(\mu, v\) and \(\alpha\)), we finally get
\[
\frac{B(\alpha + \mu + e)}{B(\alpha + v + e)} = \mathcal{O}\left(n^{|v| - |\mu|}\right), \quad n \to \infty.
\]
Now let us consider the last factor in (4.13). We have
\[
\frac{\int_{\mathcal{G}_d} \left(x_i - \frac{\alpha_i + \mu_i + 1}{n + |\mu| + d + 1}\right)^2 x^{\alpha+\mu} d\mathbf{x}}{\int_{\mathcal{G}_d} x^{\alpha+\mu} d\mathbf{x}} = \frac{\int_{\mathcal{G}_d} x^{\alpha+\mu+2\epsilon_i} d\mathbf{x}}{\int_{\mathcal{G}_d} x^{\alpha+\mu} d\mathbf{x}} - 2 \frac{\alpha_i + \mu_i + 1}{n + |\mu| + d + 1} \frac{\int_{\mathcal{G}_d} x^{\alpha+\mu+\epsilon_i} d\mathbf{x}}{\int_{\mathcal{G}_d} x^{\alpha+\mu} d\mathbf{x}} + \left(\frac{\alpha_i + \mu_i + 1}{n + |\mu| + d + 1}\right)^2.
\]
follows from (4.3) is fulfilled for each $i$. Thus, convergence holds by Theorem 4.3. Theorem 4.4 is more general than Păltănea’s class is.

Theorem 4.4

Thus, if $|v| - |\mu| < 1$, then (4.3) is fulfilled for each $i = 0, 1, \ldots, d$. The convergence rate in (4.10) follows from (4.5). For the quadratic test functions, a similar consideration shows that (4.8) is fulfilled for each $i = 1, \ldots, d$ and that the convergence rate in (4.11) holds. Thus, convergence (4.12) holds by Theorem 4.3. ■

An example of a non-Jacobi measure which satisfies the assumptions of Theorem 4.4 is $d\rho(x) = w(x) \, dx$ with $w(x) = x \ln \frac{1}{x}, x \in [0, 1]$. Indeed,

$x (1 - x) \leq w(x) \leq \sqrt{x} (1 - x), \quad 0 \leq x \leq 1.$

In particular, $\|\varphi_{e_1} - M_{n,\rho}(\varphi_{e_1})\|_{C(\mathbb{S}^d)} = O\left(n^{-\frac{1}{2}}\right)$ and $\|\varphi_{e_1} - M_{n,\rho} \varphi_{e_1}\|_{C(\mathbb{S}^d)} = O\left(n^{-\frac{1}{2}}\right), n \to \infty.$

Note that our class (4.9) is more general than Păltănea’s class (1.10). For example, the function considered in the previous paragraph does not belong to the class (1.10). Also, our proof is different from Păltănea’s one in [13], since he essentially used the continuity of the function $h$ in definition (1.10).

4.3. $K$-functional estimates

In the last section of the paper, we give $K$-functional estimates for the $L_p^1$-approximation by the operators (1.4) and (2.1). These estimates generalize one of the estimates of Theorem 2 in [11].

Let $C^1(\mathbb{S}^d)$ be the subspace of functions $g \in C(\mathbb{S}^d)$ with continuous partial derivatives $\partial_i g, i = 1, \ldots, d$, and semi-norm

$$\|\nabla g\|_{C(\mathbb{S}^d)} := \max_{i=1,\ldots,d} \|\partial_i g\|_{C(\mathbb{S}^d)},$$

and let

$$\|g\|_{C^1(\mathbb{S}^d)} := \max \{\|g\|_{C(\mathbb{S}^d)}, \|\nabla g\|_{C(\mathbb{S}^d)}\}.$$  

We will give estimates in terms of the $K$-functionals

$$K(f; t) := \inf_{g \in C^1(\mathbb{S}^d)} \left\{ \|f - g\|_{L_p^1(\mathbb{S}^d)} + t \|\nabla g\|_{C(\mathbb{S}^d)} \right\}$$

and

$$\tilde{K}(f; t) := \inf_{g \in C^1(\mathbb{S}^d)} \left\{ \|f - g\|_{L_p^1(\mathbb{S}^d)} + t \|g\|_{C^1(\mathbb{S}^d)} \right\}.$$
Theorem 4.5. For the integral operator \( L_{n,\omega,\rho} \), we have the error estimate
\[
\| f - L_{n,\omega,\rho}(f) \|_{L^1_\rho(S^d)} \leq (1 + A_{n,\omega,\rho}) \tilde{K} \left( f; \frac{\Delta_{n,\omega,\rho}}{1 + A_{n,\omega,\rho}} \right), \quad f \in L^1_\rho(S^d),
\]
with \( A_{n,\omega,\rho} \) as in (3.1),
\[
\tilde{\Delta}_{n,\omega,\rho} := \| 1 - L_{n,\omega,\rho}(1) \|_{L^1_\rho(S^d)} + \sum_{i=1}^d \| (L_{n,\omega,\rho}(|\varphi_i - x_i 1|))(x) \|_{L^1_\rho(S^d)},
\]
and the test functions \( \varphi_i \) as in (4.1).

For the Bernstein–Durrmeyer operator \( M_{n,\rho} \), we have the error estimate
\[
\| f - M_{n,\rho}(f) \|_{L^1_\rho(S^d)} \leq 2 K \left( f; \frac{\Delta_{n,\rho}}{2} \right), \quad f \in L^1_\rho(S^d),
\]
with
\[
\Delta_{n,\rho} := \sum_{i=1}^d \| (M_{n,\rho}(|\varphi_i - x_i 1|))(x) \|_{L^1_\rho(S^d)}.
\]

In the second term of the definition of \( \tilde{\Delta}_{n,\omega,\rho} \), the operator \( L_{n,\omega,\rho} \) is first applied to the function \( y \mapsto |\varphi_i(y) - x_i 1(y)| \), and then the integral is taken with respect to the variable \( x \). \( \Delta_{n,\rho} \) is defined similarly.

Proof. The proof is standard. Since
\[
f - L_{n,\omega,\rho}(f) = (f - g) - L_{n,\omega,\rho}(f - g) + (g - L_{n,\omega,\rho}(g))
\]
for \( g \in C^1(S^d) \) and
\[
\| (f - g) - L_{n,\omega,\rho}(f - g) \|_{L^1_\rho(S^d)} \leq (1 + A_{n,\omega,\rho}) \| f - g \|_{L^1_\rho(S^d)}, \tag{4.14}
\]
we have to estimate the norm of \( g - L_{n,\omega,\rho}(g) \). For \( g \in C^1(S^d) \) and \( x, y \in S^d \) we have
\[
g(y) - g(x) = \sum_{i=1}^d \partial_i g(\xi) (y_i - x_i)
\]
with \( \xi \in S^d \) a convex combination of \( x \) and \( y \). Therefore,
\[
|g(y) - g(x)| \leq \| \nabla g \|_{C(S^d)} \sum_{i=1}^d |y_i - x_i|
\]
and
\[
|g - g(x) 1| \leq \| \nabla g \|_{C(S^d)} \sum_{i=1}^d |\varphi_i - x_i 1|, \quad x \in S^d. \tag{4.15}
\]

Fix \( x \in S^d \). Since the operator \( L_{n,\omega,\rho} \) is positive, we have
\[
\left| (g - L_{n,\omega,\rho}(g))(x) \right| \leq |g(x) - L_{n,\omega,\rho}(g(x) 1)(x)| + \left| (L_{n,\omega,\rho}(g(x) 1 - g))(x) \right|
\]
\[
\leq |g(x)| \left| (1 - L_{n,\omega,\rho}(1))(x) \right| + \left| (L_{n,\omega,\rho}(|g(x) 1 - g|))(x) \right|
\]
\[ \| g - L_{n,\omega,\rho}(g) \|_{L^1(\mathbb{S}^d)} \leq \| g \|_{C^1(\mathbb{S}^d)} \tilde{\Delta}_{n,\rho}. \] (4.16)

We arrive at the estimate
\[ \| f - L_{n,\omega,\rho}(f) \|_{L^1(\mathbb{S}^d)} \leq (1 + A_{n,\omega,\rho}) \left\{ \| f - g \|_{L^1(\mathbb{S}^d)} + \frac{\tilde{\Delta}_{n,\rho}}{1 + A_{n,\omega,\rho}} \| g \|_{C^1(\mathbb{S}^d)} \right\} \]

for arbitrary \( g \in C^1(\mathbb{S}^d) \), and taking the infimum over all possible \( g \in C^1(\mathbb{S}^d) \) gives the desired bound.

The proof for \( M_{n,\rho} \) repeats the proof for \( L_{n,\omega,\rho} \) with the following changes. Since \( A_{n,\omega,\rho} = 1 \) in this case, estimate (4.14) becomes
\[ \|(f - g) - M_{n,\rho}(f - g)\|_{L^1(\mathbb{S}^d)} \leq 2 \| f - g \|_{L^1(\mathbb{S}^d)}. \]

Since \( M_{n,\rho} \) reproduces constant functions, estimate (4.16) can be replaced by
\[ \| g - M_{n,\rho}(g) \|_{L^1(\mathbb{S}^d)} \leq \| \nabla g \|_{C(\mathbb{S}^d)} \Delta_{n,\rho}. \]

Acknowledgments

We thank Hubert Berens, Paul Nevai, and Joachim Stöckler for helpful and inspiring discussions. We thank Carl de Boor and the anonymous referees for their remarks which helped to improve the presentation of the paper.

References