DISCRETIZATION AND APPROXIMATION ON HIGH DIMENSIONAL DOMAINS

by

Yeli Niu

A thesis submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

Department of Mathematical and Statistical Sciences
University of Alberta

©Yeli Niu, 2021

Abstract

This thesis includes 2 parts. Part 1 is: Discretization on High Dimensional Compact Domains. Part 2 is: Polynomial Approximation on High Dimensional Spheres. A special example for part 1 is the result on the unit sphere of high-dimensional Euclidean spaces.

In chapter 2, we obtained general results about discretization of integration on compact metric domains. Related results on spheres, closed balls and simplexed could be as special examples of our results. The first main result is about regular partitions on compact path-connected metric space equipped with non-atomic Borel probability measure. An example of this result is the regular partition on the unit sphere, which improves the previous results by the uniform absolute constant in the diameter of each partition. Many similar results were obtained with constants depending on the dimension of the Euclidean space (they are an exponential form of the dimension). It is a pity that our method here is not constructable. Resting main results are about numerical integration. Numerical integration plays an important role in approximation theory. To integrate a given function, we sometimes do not know its original function, sometimes it is too complicated to find its original function. Thus in many applied problems, we need to use numerical integration (discrete weighted summation) to asymptotically express it. A main topic is finding fixed nodes and weights to approximately express integration for

a class of function, modifying weights and notes to improve the uniform approximation error for the class of functions. We here used methods from [4] to prove the existence of nodes and weights for numerical integration which result in a better approximation error. One result is about discretization of integration on compact matric spaces that equipped with certain measures. Results here is better than previous results under the following points. First, we reduced the smoothness requirements. Functions here do not need to be differentiable, satisfying Lipschitz condition is enough. Second, example of our result about discretization of integration for piecewise polynomials on the unit sphere gives a better approximation error, somewhat overcome the curse of dimensionality. The last main result is about discretization of integration on finite-dimensional compact domains.

Chapter 3 mainly discusses the Jackson type's inequality and its matching inverse inequality, equivalence of K-functional and modulus of smoothness on the unit sphere $\mathbb{S}^{d-1}$. There are many definitions for K-functional and modulus of smoothness. Here we use the modulus of smoothness defined by Z. Ditzian via rotation operator on the unit sphere. K-functional here we defined through partial derivative in Euler angles. In 1964, D. J. Newman, and H. S. Shapiro [28] proved that for $f \in C(\mathbb{S}^{d-1})$, the constant appear in Jackson type's inequality for $r = 1, p = \infty$ can be a dimension-free constant. Results in this chapter show that this result can be extended to all cases of positive integer r and $p \geq 1$. We also obtain the matching inverse Jackson's inequality, equivalence of $K-$functional and modulus of smoothness with constant in equivalence independent of dimension d. In this sense, we improved Ditzian's results on Jackon's inequality. Jackson type's inequality and its matching inverse inequality connect the rate of polynomial approximation to the smoothness properties of functions on the sphere. Equivalence of K-functional and modulus of smoothness builds the relation between difference with differentiation.

I will describe in detail my current research projects in these two directions and the progress I've made.

Preface

Chapter 2 of this thesis is a joint paper with Professor Feng Dai and Professor Martin D. Buhmann. This paper is published to Adv. Math. on Apr. 2021 as [9]: Martin Buhmann, Feng Dai, Yeli Niu. Discretization of integrals on compact metric measure spaces. Advances in Mathematics 381 (2021).32 pages. Professor Feng Dai is the main designer of this paper. He introduced the problem. Main proofs are given by Professor Feng Dai and me. Professor Martin D. Buhmann gave several examples and references. All of the proofs in this chapter are joint work of Professor Feng Dai and me.

Chapter 3 was a joint work with Professor Feng Dai. He came up with the problem. Proofs of main results in this chapter is solved by both of us.

Acknowledgements

Firstly, I would like to express my deepest appreciation to my supervisor Professor Feng Dai for his patient guidance, constant encouragement these years. His noble personality, profound knowledge, great dedication to work inspired me a lot. His scholarly suggestions and insightful perception had a significant influence on my academic study. His family's friendly and hospitality warmed me in this distant country. Without his tremendous support, this thesis is impossible to be finished.

Secondly, I want to express my sincere gratitude to my master's supervisor Heping Wang. He introduced me to the area of approximation theory and offered vulnerable help and guidance during my study at Capital Normal University.

Thirdly, many thanks to professors who taught me courses, especially Professor Anthony Tony Lau and Professor Bin Han, for these interesting topics in courses and workshops, practical suggestions in academic study. I also want to express my thanks to my committee members, for their time and interesting questions and advice.

Again, I'm very grateful to the teachers and scholars I meet these years. They lead me to beautiful mathematics. Their rigorous spirit for math, studious devotion to research and teaching, admirable personalities are light in my life, prompt me to explore further in this road.

Also, many thanks to the Chinese Scholarship Council for their continuing financial support during the last four years.

Last, but not least, special thanks to my parents and my friends everywhere. For my parents' lasting support these years. I saw so many different possibilities in life from my friends.

Contents

Chapter 1

Introduction

The first part, chapter 2 of this thesis is mainly about discretization of integration on compact metric domains. Given a set $\Omega \subset \mathbb{R}^{d+1}$ with a normalized Borel measure μ, a function f defined on Ω.

$$\Lambda_m(f, \boldsymbol{\xi}) := \sum_{j=1}^{m} \lambda_j f(\xi^j), \quad \boldsymbol{\xi} = (\xi^1, \dots, \xi^m) \in \Omega \times \dots \times \Omega.$$

Such a formula $\Lambda_m(\cdot, \boldsymbol{\xi})$ is called a *cubature formula* (C.F.) with fixed nodes $\boldsymbol{\xi} = (\xi^1, \dots, \xi^m)$ and fixed weights $\Lambda := (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$. A C.F. is called positive if $\lambda_1, \cdots, \lambda_m \geq 0$.

Typically, one assumes that the target function f lies in some class $\mathbf{W}$ of continuous functions on Ω in the numerical approximation

$$\int_\Omega f(\mathbf{x}) d\mu(\mathbf{x}) \approx \Lambda_m(f, \boldsymbol{\xi}) = \sum_{j=1}^{m} \lambda_j f(\xi^j),$$

and then looks for good C.F.s for all functions in $\mathbf{W}$. That is, the interest here is on *good* C.F.s for a given class of functions $\mathbf{W}$ rather than an individual function. The error of such approximation is measured by the following quantity:

$$\Lambda_m(\mathbf{W}, \boldsymbol{\xi}) := \sup_{f \in \mathbf{W}} \left| \int_\Omega f d\mu - \Lambda_m(f, \boldsymbol{\xi}) \right|.$$

One can further optimize the C.F.s and study the quantity with infinimum taken over all possible

$\Lambda, \boldsymbol{\xi}$ in Λ_m

$$\inf \Lambda_m(\mathbf{W}, \boldsymbol{\xi})$$

In classical approximation theory, a typical method of estimating the quantity $\inf_{\Lambda_m} \Lambda_m(\mathbf{W}, \boldsymbol{\xi})$ with $\Omega = \mathbb{S}^d$ involves the following two steps:

(I). Reduce the dimension: find a "good" $X_N \subset L_1(\mathbb{S}^d)$ to approximate functions from $\mathbf{W}$ with $N \approx m$.

(II). Find a good C.F. Λ_m to approximate integrals of functions on X_N:

$$\int_{\mathbb{S}^d} f(\mathbf{x})\, d\mu(\mathbf{x}) = \Lambda_m(f) = \sum_{j=1}^{m} \lambda_j f(\xi^j), \;\; \forall f \in X_N.$$

In the second step, one can use the following theorem:

Tchakaloff's theorem.[17] Let Ω be a compact metric space and μ a Borel probability measure on Ω. If X_N is an N-dimensional space of continuous functions on Ω, then there exist exactly N points $\xi^j \in \Omega$ and numbers $\lambda_j \geq 0$, $1 \leq j \leq N$ such that

$$\int_{\Omega} f(x)\, d\mu(x) = \sum_{j=1}^{N} \lambda_j f(\xi^j), \qquad \forall f \in X_N. \tag{1.0.1}$$

Indeed, using this method, one can show

Corollary 1.0.1. *Let* $\mathbf{W}$ *be a compact subset of* $C(\mathbb{S}^d)$*. Then*

$$\inf_{\substack{\xi_j \in \mathbb{S}^d, \lambda_j \in \mathbb{R} \\ j=1,2,\cdots,m}} \sup_{f \in \mathbf{W}} \Big| \int_{\mathbb{S}^d} f\, d\mu - \sum_{j=1}^{m} \lambda_j f(\xi^j) \Big| \leq 2 d_m(\mathbf{W}, L_\infty),$$

where $d_m(\mathbf{W}, L_\infty)$ *is the Kolmogorov* m*-width of* $\mathbf{W}$ *in the space* $C(\mathbb{S}^d)$*:*

$$d_m(\mathbf{W}, L_\infty) := \inf_{X_m} \sup_{f \in \mathbf{W}} \inf_{g \in X_m} \|f - g\|_\infty$$

In the case when $\mathbf{W}$ is the unit ball $B^r := \{f \in W_\infty^r(\mathbb{S}^d) : \;\; \|f\|_{W_\infty^r} \leq 1\}$ of the usual

Sobolev space W^r_∞ of order r on the sphere $\mathbb{S}^d$, we have the following sharp estimates [5, 18]:

$$\inf_{\substack{\xi_j\in\mathbb{S}^d,\lambda_j\in\mathbb{R}\\ j=1,2,\cdots,N}} \sup_{f\in B^r}\Big|\int_{\mathbb{S}^d} f\,d\mu - \sum_{j=1}^N \lambda_j f(\xi^j)\Big| \asymp N^{-\frac{r}{d-1}}, \tag{1.0.2}$$

where the constants of equivalence depend on d and r.

The problem with the estimate (1.0.2) lies in the following two facts:

(i) if the dimension d is large and the smoothness parameter r is small (say, $r = 1$), then the rate approximation $N^{-\frac{r}{d-1}}$ in (1.0.2) goes to zero very slowly as $N \to \infty$;

(ii) the implied constant of equivalence in (1.0.2) grows exponentially fast to ∞ as $d \to \infty$.

In this part, an example of one main results is the following discritization problem for zonal functions on the sphere:

$$\inf_{\lambda_1,\lambda_2,\cdots,\lambda_N\in\mathbb{R}} \inf_{\xi_1,\cdots,\xi_N\in\mathbb{S}^d} \max_{\mathbf{x}\in\mathbb{S}^d}\Big|\int_{\mathbb{S}^d} \Phi(\mathbf{x}\cdot\mathbf{y})g(\mathbf{y})d\mu(\mathbf{y}) - \sum_{j=1}^N \lambda_j\Phi(\mathbf{x}\cdot\xi_j)\Big|$$

where $g \in L^\infty(\mathbb{S}^d)$ and $\Phi : [-1, 1] \to \mathbb{R}$ is a Lipschit-function satisfying certain conditions. The result shows that the above quantity can be controlled by a better error that goes to zero faster as d goes to infinity (faster than previous results). Our result somewhat overcome the curse of dimensionality. For more details, please read the section 2.5 in the chapter 2.

Now we introduce several useful definitions and main results in the chapter 2. Let (Ω, ρ) be a compact metric space. Open balls and closed balls in Ω will be denoted by $B_\zeta(x) := \{y \in \Omega : \rho(x, y) < \zeta\}$, and $B_\zeta[x] := \{y \in \Omega : \rho(x, y) \leqslant \zeta\}$, respectively. A path connecting two points $x, y \in \Omega$ is a continuous map $\gamma : [0, 1] \to \Omega$ with $\gamma(0) = x$ and $\gamma(1) = y$. A metric space (Ω, ρ) is called path-connected if every two distinct points in Ω can be connected with a path. As is well known, every open connected subset of $\mathbb{R}^{d+1}$ is path-connected. Given a set $A \subset \Omega$ and a point $x \in \Omega$, define

$$\mathrm{dist}(x, A) := \inf_{y\in A} \rho(x, y).$$

A measure μ on Ω is called non-atomic if for any measurable set $A \subset \Omega$ with $\mu(A) > 0$

there exists a measurable subset B of A such that $\mu(A) > \mu(B) > 0$. For non-atomic Borel probability measure μ on Ω, we have the property: If $A_0 \subset A_1 \subset \Omega$, $0 < \mu(A_1)$ and $\mu(A_0) \leq t \leq \mu(A_1)$, then there exists a measurable subset $E_t \subset A_1$ satisfies $\mu(E_t) = t$. Normalized discrete measure is an example of atomic measure. Its subsets only have two possible valued measures: zero or 1.

A partition of Ω consists of finitely many pairwise disjoint subsets of Ω whose union is Ω. We first introduce the following theorem about regular partition on path-connected compact metric space, this theorem will be proved in Section 2.2.

Theorem 1.0.2. *Let (Ω, ρ) be a compact path-connected metric space with diameter $\operatorname{diam}(\Omega) = \sup_{x,y\in\Omega} \rho(x, y) = \pi$. Let μ be a non-atomic Borel probability measure on Ω, and $N \geq 2$ a positive integer. Assume that the inequality*

$$\inf_{x\in\Omega} \mu\Big(B_{\delta/2}(x)\Big) \geqslant \frac{1}{N} \tag{1.0.3}$$

holds for some $\delta > 0$. Then there exists a partition $\{R_1, \dots, R_N\}$ of Ω such that

(i) *the R_j are pairwise disjoint subsets of Ω,*

(ii) *for each $1 \leq j \leq N$, $\mu(R_j) = \frac{1}{N}$ and $\operatorname{diam}(R_j) \leq 4\delta$.*

Theorem 1.0.2 with constants depending on certain geometric parameters of the underlying space (Ω, ρ, μ) (e.g. *dimension*) is known in a more general setting. In result about area-regular partition on the unit sphere $\mathbb{S}^d$, the constant in the diameter of each partition is of an exponential form of the dimension d. The crucial point here lies in the fact that the constant 4 in the estimates of $\operatorname{diam}(R_j)$ is absolute.

Let (X, ρ) be a compact metric space with metric ρ and diameter π. For $x \in X$ and $0 \leq \underline{a} < \underline{b} \leq \pi$, set

$$E(x; \underline{a}, \underline{b}) := \{y \in X : \ \underline{a} \leq \rho(x, y) \leq \underline{b}\}.$$

Definition 1.0.3. *Let $0 = t_0 < t_1 < \cdots < t_\ell = \pi$ be a partition of the interval $[0, \pi]$, and let $r \in \mathbb{N}$. We say $\Phi \in C[0, \pi]$ belongs to the class $\mathcal{S}_r \equiv \mathcal{S}_r(t_1, \dots, t_\ell)$ if there exists an*

r-dimensional linear subspace V_r of $C(X)$ such that for any $x \in X$ and each $1 \le j \le \ell$,

$$\Phi(\rho(x,\cdot))\Big|_{E(x;t_{j-1},t_j)} \in \Big\{ f\Big|_{E(x;t_{j-1},t_j)} : \ f \in V_r \Big\}.$$

Next, let μ be a Borel probability measure on X satisfying the following condition for a parameter $\beta \ge 1$ and some constant $c_1 > 1$:

(a) for each positive integer N, there exists a partition $\{X_1, \dots, X_N\}$ of X such that $\mu(X_j) = \frac{1}{N}$ and $\mathrm{diam}(X_j) \le \delta_N := c_1 N^{-\frac{1}{\beta}}$ for $1 \le j \le N$.

According to Theorem 1.0.2, Condition (a) holds automatically with $c_1 = 20\pi$ if the metric space X is path-connected, and μ is a non-atomic Borel probability measure on X satisfying that for any $0 < t \le 1$,

$$\inf_{x \in X} \mu(B_t(x)) \ge \left(\frac{8}{c_1}\right)^\beta t^\beta. \tag{1.0.4}$$

One main result in this chapter is:

Theorem 1.0.4. *Let $\Phi \in C[0,\pi]$ satisfy*

$$|\Phi(s) - \Phi(s')| \le |s - s'|, \qquad \forall s, s' \in [0,\pi], \tag{1.0.5}$$

and belong to a class $\mathcal{S}_r(t_1, \dots, t_\ell)$ for some compact metric space (X,ρ), where $r \in \mathbb{N}$ and $0 = t_0 < t_1 < \dots < t_\ell = \pi$. Let μ be a Borel probability measure on X satisfying the condition (a) and the following condition:

(b) *for each $x \in X$ and $\delta \in (0,\pi)$,*

$$\mu\Big(E(x; t_j - \delta, t_j + \delta)\Big) \leqslant c_2 \delta, \ \ 1 \le j < \ell, \tag{1.0.6}$$

where $c_2 > 1$ is a constant independent of δ and x.

Then for each positive integer $N \ge 4$, there exist points $y_1, \dots, y_{(r+2)N} \in X$ and nonnegative *numbers* $\lambda_1, \dots, \lambda_{(r+2)N}$ *such that* $\sum_{j=1}^{(r+2)N} \lambda_j = 1$ *and*

$$\max_{x\in X}\left|\int_X \Phi(\rho(x,y))\,\mathrm{d}\mu(y) - \sum_{j=1}^{(r+2)N} \lambda_j \Phi(\rho(x,y_j))\right| \leqslant c_3 N^{-\frac{1}{2}-\frac{3}{2\beta}}\sqrt{\log N},$$

where $c_3 := 8c_1^2\sqrt{c_2\ell}\sqrt{\beta}$.

In the case when the metric space X is path-connected and the borel probability measure μ on X be non-atomic, with the Theorem 1.0.2 we have:

Theorem 1.0.5. *Let* (X,ρ) *be a compact path-connected metric space. Let* $\Phi \in C[0,\pi]$ *satisfy* (1.0.5) *and belong to a class* $\mathcal{S}_r(t_1,\dots,t_\ell)$ *for some* $r \in \mathbb{N}$ *and* $0 = t_0 < t_1 < \cdots < t_\ell = \pi$. *Let* μ *be a non-atomic Borel probability measure on* X *satisfying* (1.0.4). *Assume in addition that the condition (b) in Theorem 1.0.4 is satisfied. Then for any* $g \in L^\infty(X,\mathrm{d}\mu)$ *with* $\|g\|_{L^\infty(\mathrm{d}\mu)} \le 1$, *and each positive integer* $N \ge 20$, *there exist points* $y_1,\dots,y_{2(r+2)N} \in X$ *and real numbers* $\lambda_1,\dots,\lambda_{2(r+2)N}$ *such that*

$$\max_{x\in X}\left|\int_X \Phi(\rho(x,y))g(y)\,\mathrm{d}\mu(y) - \sum_{j=1}^{2(r+2)N} \lambda_j \Phi(\rho(x,y_j))\right| \leqslant 45c_3 N^{-\frac{1}{2}-\frac{3}{2\beta}}\sqrt{\log N}.$$

The proof of this theorem will partly rely on the regular-partition Theorem 1.0.2. Proofs of Theorem 1.0.4 and Theorem 1.0.5 appear in Section 2.3.

Let us give some examples of the metric spaces (X,ρ) and the associated classes $\mathcal{S}_r$ which satisfy the conditions of Theorem 1.0.5 .

Example 1.0.6. *(i) Let* $X = \mathbb{S}^d$ *be the unit sphere of* $\mathbb{R}^{d+1}$ *equipped with the usual geodesic distance* $\rho(x,y) = \arccos x\cdot y$ *for* $x,y \in \mathbb{S}^d$. *If* $\varphi \in C[-1,1]$ *is a piecewise algebraic polynomial of degree at most* n_0 *on* $[-1,1]$, *then the function* $\Phi(\theta) := \varphi(\cos\theta)$, $\theta \in [0,\pi]$ *belongs to a class* $\mathcal{S}_r$ *with* r *being the dimension of the space of all spherical polynomials of degree at most* n_0 *on the sphere* $\mathbb{S}^{d-1}$. *In this case,* $\Phi(\rho(x,y)) = \varphi(x\cdot y)$, *and the condition* (1.0.4) *implies both the condition (a) and the condition (b).*

(ii) Let $X = B_{\frac{\pi}{2}}(0) \subset \mathbb{R}^{d+1}$ *be the Euclidean ball with centre* 0 *and radius* $\frac{\pi}{2}$. *If* $\varphi \in C[0,\infty)$ *is a piecewise algebraic polynomial of degree at most* n_0, *then the function* $\Phi(t) :=$

$\varphi(t^2)$, $t \geq 0$ belongs to a class $\mathcal{S}_r$ with r being the dimension of the space of all algebraic polynomials of degree at most $2n_0$ in $d+1$ variables. In this case, $\Phi(\rho(x,y)) = \varphi(\|x-y\|^2)$, and the condition (1.0.4) *implies both the condition (a) and the condition (b).*

We will discuss these examples in details in section 2.5 and 2.6. The Theorem 1.0.5 along with the Theorem 1.0.4 will be proved in Section 2.3.

We also give a result for discretization on finite-dimensional compact domains. It is an analogue of Theorem 1.0.5 for all $g \in L^1(\mathrm{d}\mu)$ (instead of $g \in L^\infty(\mathrm{d}\mu)$) on finite-dimensional domains. The implied constant in section 2.4 will depend on the dimension and the underlying domain.

Let $(X, \|\cdot\|)$ be a finite-dimensional real normed linear space. Let $B_\zeta(x)$ (resp. $B_\zeta[x]$) denote the open balls (resp. closed balls) with centre $x \in X$ and radius $\zeta > 0$ defined with respect to the metric $\rho(x,y) = \|x-y\|$. Here $\|\cdot\|$ is not necessarily the Euclidean norm. Let $\Omega \subset B_1[0]$ be a compact subset of X (not necessarily connected). Let μ be a Borel probability measure supported on Ω satisfies the following two conditions:

(i) there exist a positive constant $c_4 > 1$ and a parameter $\beta > 1$ such that for any $x \in \Omega$ and $\delta \in (0,2]$

$$c_4^{-1}\delta^\beta \leq \mu\Big(B_\delta(x)\Big) \leq c_4\delta^\beta; \tag{1.0.7}$$

(ii) there exists a constant $c_5 > 0$ such that for any $x \in \Omega$ and $t, s \in (0,2]$,

$$\mu\Big(\{y \in \Omega :\ t \leq \|y-x\| \leq t+s\}\Big) \leq c_5 s. \tag{1.0.8}$$

Under these two conditions, we shall prove

Theorem 1.0.7. *Let $\Phi : [0,\infty) \to \mathbb{R}$ be a function such that*

$$|\Phi(s) - \Phi(s')| \leq |s - s'|,\ \forall s, s' \in [0,2]. \tag{1.0.9}$$

Assume that there exist a partition $0 = t_0 < t_1 < \cdots < t_\ell = 2$ of $[0,2]$ and a translation-invariant linear subspace X_r of $C(\Omega)$ with $\dim X_r = r$ such that with $E_j := \{x \in \mathbb{R}^d :$

$t_{j-1} \le \|x\| \le t_j\}$, $j = 1, 2, \dots, \ell$,

$$\Phi(\|\cdot\|)\Big|_{E_j} \in \Big\{ f\Big|_{E_j} : \ f \in X_r \Big\}.$$

Let $g \in L^1(\Omega, \mu)$ be such that $\|g\|_{L^1(\mathrm{d}\mu)} = 1$. Then for each positive integer $n \ge 2$, there exist points $y_1, \dots, y_n \in \Omega$ and real numbers $\lambda_1, \dots, \lambda_n$, such that

$$\sup_{x \in \Omega} \Big| \int_\Omega \Phi(\|x - y\|) g(y) \, \mathrm{d}\mu(y) - \sum_{k=1}^n \lambda_k \Phi(\|x - y_k\|) \Big|$$
$$\le C(X) \begin{cases} n^{-\frac{1}{2} - \frac{3}{2\beta}} (\log n)^{\frac{1}{2}}, & \text{if } 1 < \beta < 3, \\ n^{-1} (\log n)^{\frac{3}{2}}, & \text{if } \beta = 3, \\ n^{-\frac{\beta+1}{2(\beta-1)}} (\log n)^{\frac{1}{2}}, & \text{if } \beta > 3, \end{cases} \tag{1.0.10}$$

where the constant $C(X)$ depends only on $\dim X$, c_4, c_5, r, ℓ and β.

The proof of this result will be given in section 2.4.

The second main part of this thesis is about: polynomial approximation on high dimensional spheres.

Let $\mathbb{S}^{d-1}$ denote the unit sphere of $\mathbb{R}^d$ equipped with the surface Lebesgue measure normalized by $\int_{\mathbb{S}^{d-1}} 1 \, d\sigma(x) = 1$. Given $1 \le p < \infty$, we denote by $L^p(\mathbb{S}^{d-1})$ the Lebesgue L^p-space on $\mathbb{S}^{d-1}$ equipped with the norm

$$\|f\|_p := \Big(\int_{\mathbb{S}^{d-1}} |f(x)|^p \, d\sigma(x) \Big)^{\frac{1}{p}}.$$

In the limiting case, we identify $L^\infty(\mathbb{S}^{d-1})$ with the space $C(\mathbb{S}^{d-1})$ of all continuous functions on $\mathbb{S}^{d-1}$ equipped with the uniform norm

$$\|f\|_\infty = \max_{x \in \mathbb{S}^{d-1}} |f(x)|.$$

A spherical polynomial of degree at most n on $\mathbb{S}^{d-1}$ is the restriction on $\mathbb{S}^{d-1}$ of an algebraic polynomial in d variables of total degree at most n. Denote by Π_n^d the space of all real spherical polynomials of degree at most n on the sphere $\mathbb{S}^{d-1}$. As is well known (see [19]), Π_n^d is a finite dimensional vector space with

$$\dim \Pi_n^d = \frac{(2n+d-1)\Gamma(n+d-1)}{\Gamma(n+1)\Gamma(d)} = \frac{2n+d-1}{\Gamma(d)} \prod_{j=1}^{d-2} (n+j). \tag{1.0.11}$$

For $1 \le p \le \infty$ and $f \in L^p(\mathbb{S}^{d-1})$, we define

$$E_n(f)_p := \min_{P \in \Pi_n^d} \|f - P\|_p, \quad n = 0, 1, \cdots .$$

Let $SO(d)$ denote the group of rotations on $\mathbb{R}^d$ equipped with the normalized Haar measure dQ. As is well known, for each integrable function f on $SO(d)$ and any $\rho \in SO(d)$,

$$\int_{SO(d)} f(Q)dQ = \int_{SO(d)} f(\rho Q)dQ = \int_{SO(d)} f(Q\rho)dQ = \int_{SO(d)} f(Q^{-1})dQ. \tag{1.0.12}$$

Furthermore, for each $f \in L^1(\mathbb{S}^{d-1})$ and $e \in \mathbb{S}^{d-1}$,

$$\int_{SO(d)} f(Qe)\, dQ = \int_{\mathbb{S}^{d-1}} f(x)\, d\sigma(x). \tag{1.0.13}$$

For each $Q \in SO(d)$, we define the "translation" operator T_Q by $T_Q f(x) = f(Qx)$, where $f \in L^1(\mathbb{S}^{d-1})$ and $x \in \mathbb{S}^{d-1}$. Accordingly, given $Q \in SO(d)$, we define the r-th order difference operator $\triangle_Q^r : L^p(\mathbb{S}^{d-1}) \to L^p(\mathbb{S}^{d-1})$, $1 \le p \le \infty$ by

$$\triangle_Q^r f(x) = (I - T_Q)^r f(x) = \sum_{j=0}^{r} (-1)^j \binom{r}{j} f(Q^j x), \quad x \in \mathbb{S}^{d-1}.$$

where I denotes the identity operator.

Definition 1.0.8. *Given $t > 0$, we define $\mathcal{O}(t)$ to be the class of all rotations $\rho \in SO(d)$ taking* *the form* $\rho = Q^{-1}M_\theta Q$ *for some* $Q \in SO(d)$ *and* $\theta \in [-t, t]$*, where*

$$M_\theta := \begin{Bmatrix} \cos\theta & \sin\theta & & & \\ -\sin\theta & \cos\theta & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{Bmatrix}_{d\times d}, \quad \theta \in \mathbb{R}. \tag{1.0.14}$$

Clearly, the matrix M_θ defined above is a rotation in x_1x_2-plane satisfying $M_\theta^j = M_{j\theta}$ for $j \in \mathbb{Z}$.

Definition 1.0.9. *Let* $1 \le p \le \infty$ *and* $r \in \mathbb{N}$*. The* r*th order modulus of smoothness of* $f \in L^p(\mathbb{S}^{d-1})$ *is defined by*

$$\omega^r(f,t)_p := \sup_{Q\in\mathcal{O}(t)} \|\triangle_Q^r f\|_p, \ \ t > 0. \tag{1.0.15}$$

Clearly, by the definition, for each $Q \in SO(d)$ and $\theta \in \mathbb{R}$,

$$\triangle^r_{Q^{-1}M_\theta Q} = T_Q^{-1}\triangle^r_{M_\theta}T_Q, \tag{1.0.16}$$

and

$$\omega^r(f,t)_p = \sup_{|\theta|\le t}\sup_{Q\in SO(d)} \Big\|\sum_{j=0}^r (-1)^j \binom{r}{j} f(Q^{-1}M_{j\theta}Q\,\cdot)\Big\|_p. \tag{1.0.17}$$

The following properties follow directly from the properties of modulus of smoothness on the unit circle.

Lemma 1.0.10. [22] *Let* $0 < p \le \infty$*. The modulus of smoothness* $\omega^r(f,t)_p$ *defined above have the following properties:*

(i) *For any* $r \in \mathbb{N}$, $\omega^r(f,t)_p \le 2^r\|f\|_p$.

(ii) *For* $0 < s < r$, $\omega^r(f,t)_p \le 2^{r-s}\omega^s(f,t)_p$.

(iii) *For any* $\ell \in \mathbb{N}$, $\omega^r(f, \ell t)_p \leq \ell^r \omega^r(f,t)_p$.

(iv) *(Marchaud inequality)* $m > r$,

$$\omega^r(f,t)_p \leq C_1(m) t^r \int_t^1 \frac{\omega^m(f,u)_p}{u^{r+1}}\, du, \tag{1.0.18}$$

where the constant $C_1(m)$ *depends only on* m.

The modulus of smoothness $\omega^r(f,t)_p$ was introduced by Z. Ditzain [22, Section 10], where the $\omega^r_*(f,t)_p$ was used for $\omega^r(f,t)_p$. It has the advantage that the difference $\triangle^r_Q f(x)$ for $Q \in \mathcal{O}(t)$ is essentially taken over a unit circle (see (1.0.17)) as the class $\mathcal{O}(t)$ consists of two-dimensional rotations on planes only. Note that in [22], the supremum is taken over the set

$$\mathcal{O}'(t) := \Big\{ Q \in SO(d) : \max_{x \in \mathbb{S}^{d-1}} \arccos(x \cdot Qx) \leq t \Big\}.$$

Since each $Q \in \mathcal{O}'(t)$ can be expressed as $Q = A_1 M_\theta A_1^{-1} A_2 M_\theta A_2^{-1} \cdots A_l M_\theta A_l^{-1}$ with some orthogonal matrix $A_1, A_2, \dots A_l$, those two modulis of smoothness are dimension-free equivalent with each other.

Both the Jackson inequality,

$$E_n(f)_p \leq C_{p,d,r} \omega^r(f, n^{-1})_p, \;\; 1 \leq p \leq \infty, \tag{1.0.19}$$

and its matching inverse inequality

$$\omega^r(f, n^{-1})_p \leq C_{p,r,d} n^{-r} \sum_{k=1}^n k^{r-1} E_{k-1}(f)_p \tag{1.0.20}$$

were proved by Ditzian [14, 21]. A relatively simpler proof of the Jackson inequality (1.0.19) can be found in [15]. However, the dependence of the degree of approximation on the number of dimensions was not taken into account in these papers. Indeed, the implicit constants $C_{p,d,r}$ obtained in these papers have the asymptotic behavior d^{d^α} for some $\alpha \in (0,1)$ as $d \to \infty$.

In 1964, D. J. Newman, and H. S. Shapiro [28] proved that for $f \in C(\mathbb{S}^{d-1})$,

$$E_n(f)_\infty \le C\omega(f, \frac{d}{n})_\infty. \tag{1.0.21}$$

The crucial point here lies in the fact that the constant C is absolute and is independent of the dimension d. The authors [28] also raised the question whether similar estimates with constants independent of the dimension d can be extended to higher order moduli of smoothness. They remarked that the technique of the paper [28] does not seem to yield such extensions.

One main aim in this chapter is to show the following Jackson inequality and the matching inverse with dimension-free constants:

Theorem 1.0.11. *If $1 \le p \le \infty$ and $r \in \mathbb{N}$, then there exists a constant $C_r > 0$ depending only on r such that for all $f \in L^p(\mathbb{S}^{d-1})$ and $n = 1, 2, \cdots$,*

$$E_n(f)_p \le C_r \omega^r\Big(f, \frac{d^3}{n}\Big)_p, \tag{1.0.22}$$

and

$$\omega^r(f, n^{-1})_p \le C_r n^{-r} \sum_{k=1}^{n} k^{r-1} E_k(f)_p. \tag{1.0.23}$$

Another main result in this chapter is the dimension free equivalent of K-functional and modulus of smoothness. The partial derivative in Euler angles is given by

Definition 1.0.12. *For $r \in \mathbb{N}$, $f \in C^r(\mathbb{S}^{d-1})$ and $Q \in SO(d)$, we define the r-th order derivative $\mathcal{D}_Q^r f$ to be a function on $\mathbb{S}^{d-1}$ by*

$$\mathcal{D}_Q^r f(x) := \Big(\frac{\partial}{\partial t}\Big)^r \Big(f(Q^{-1}M_t Qx)\Big)\Big|_{t=0}, \quad x \in \mathbb{S}^{d-1}. \tag{1.0.24}$$

The related K-functional is defined as:

Definition 1.0.13. *Let $r \in \mathbb{N}$ and $1 \le p \le \infty$. We define the r-th order K-functional of*

$f \in L^p(\mathbb{S}^{d-1})$ *by*

$$K_r(f,t)_p := \inf\Big\{\|f-g\|_p + t^r \sup_{Q\in SO(d)} \|\mathcal{D}_Q^r g\|_p : \ \ g \in C^r(\mathbb{S}^{d-1})\Big\}, \ \ t>0.$$

This K-functional and its equivalence with modulus of smoothness defined above can be found in [16] and [20], but the independence of dimension appear in the degree of approximation were not considered. Here with the Jackson inequality proved above, we will show dimension free equivalence of modulus of smoothness and K-functional.

Theorem 1.0.14. *If $r \in \mathbb{N}$ and $1 \le p \le \infty$, then there exist a constant $C_r > 0$ depending only on r such that for all $f \in L^p(\mathbb{S}^{d-1})$ and $t \in (0,1)$,*

$$C_r^{-1} K_r(f, d^{-3}t)_p \le \omega^r(f,t)_p \le C_r K_r(f,t)_p. \tag{1.0.25}$$

This chapter is organized in the following way. Section 3.1 introduces De La Vallee Poussin type operator on $L_p(\mathbb{S}^{d-1})$ and the average operator. The most importance of the De La Vallee Poussin type operator is that its norm is bounded by the absolute constant 10 for any p and d. With the average operator and Newman-Shapiro operator, De La Valle Poussin type operator and Newman-Shapiro operators, section 3.2 gives the proof of Theorem 1.0.11. The proof of Marchaud inequality 1.0.18 will be given in section 3.3, as the proof rely on the Jackson inequality 1.0.22. Here we gave a more detailed proof of the Marchaud inequality. Section 3.4 discovers properties of Euler angle partial derivative and presents the proof of Theorem 3.4.3

Chapter 2

Discretization on High Dimensional Domains

2.1 Preliminaries

In this section, we list several basic results from functional analysis and probability that will be needed in later sections. All of the materials in this section can be found in the book [32].

Theorem 2.1.1. *Let X be a real linear topological space with dual space X^*. Then the following statements hold:*

(i) *Let A and B be two nonempty disjoint convex sets in X. If A is open, then there exists $\Lambda \in X^*$ such that*

$$\Lambda x < \inf_{y \in B} \Lambda y, \quad \forall\, x \in A.$$

If A is compact, B is closed and X is locally convex, then there exist $\Lambda \in X^*$ such that

$$\sup_{x \in A} \Lambda x < \inf_{y \in B} \Lambda y.$$

(ii) *If X is an F-space (i.e., a complete vector space with metric that is translation invariant whose multiplications and additions are continuous), then for every compact subset $K \subset X$, the closure of the convex hull of K is compact in X.*

Next, we recall some basic facts on weak and weak*-topologies. A topology τ_1 on a nonempty set X is said to be weaker than another topology τ_2 on X if $\tau_1 \subset \tau_2$.

Theorem 2.1.2. *Let X be a real vector space, and X' a vector space of linear functionals on X which separates points in X (i.e., given any two distinct points $x_1, x_2 \in X$ there exists $\Lambda \in X'$ such that $\Lambda x_1 \neq \Lambda x_2$). If τ denotes the weakest topology on X with respect to which every element in X' is a continuous linear functional on X, then (X, τ) is a locally convex space whose dual is X'.*

Let X be a real, locally convex linear topological space with topology τ and the dual space X^*. Let τ_w denote the weak topology of X, i.e., the weakest topology of X with respect to which every linear functional in X^* is continuous. Then $\tau_w \subset \tau$, and $X_w = (X, \tau_w)$ is a locally convex space whose dual is also X^*. We denote by τ_{w^*} the weak* -topology of X^*; that is, τ_{w^*} is the weakest topology of X^* with respect to which for every $x \in X$, the linear functional $f \in X^* \to f(x)$ is continuous. Then (X^*, τ_{w^*}) is a locally convex linear topological space whose dual is X. If X is separable, then every weak*-compact set K in X^* is metrizable in the weak*-topology.

Theorem 2.1.3. *[Banach-Alaoglu theorem] For every neighborhood V of 0 in X, its polar*

$$K := \{\Lambda \in X^* : |\Lambda x| \leq 1, \forall x \in V\}$$

is weak -compact in X^*. If, in addition, X is separable, then K is sequentially compact in the weak* -topology.*

Third, we review some basic results on vector-valued integration. We start with the following definition:

Definition 2.1.4. *Let X be a real locally convex topological vector space, and let (Q, μ) be a measure space. A vector-valued function $f : Q \to X$ is said to be integrable with respect to μ if*

$$\Lambda(f(\cdot)) = \langle \Lambda, f(\cdot) \rangle \in L^1(Q, \mu), \qquad \forall \Lambda \in X^*$$

and there exists $y \in X$ such that

$$\langle \Lambda, y \rangle = \int_Q \langle \Lambda, f(x) \rangle \, \mathrm{d}\mu(x), \qquad \forall \Lambda \in X^*.$$

If such a vector $y \in X$ exists, it must be unique, and is denoted by $\int_Q f(x)\,\mathrm{d}\mu(x)$.

Recall that a positive Borel measure μ on a topological space Q is regular if

$$\begin{aligned} \mu(E) &= \sup\{\mu(K) : K \subset E \text{ is compact}\} \\ &= \inf\{\mu(G) : E \subset G, G \text{ is open in } X\} \end{aligned}$$

for every Borel set $E \subset Q$. Each Borel probability measure on a locally compact Hausdorff space with a countable base for its topology, or on a compact metric space is regular. If Q is a compact Hausdorff space, and $C(Q)$ is the space of all continuous functions on Q (with the uniform norm), then the dual of $C(Q)$ is the space of all finite regular Borel measures (i.e., Radon measures) on Q (with the norm of total variation).

Theorem 2.1.5. *Suppose that*

(i) *X is a real, locally convex topological vector space;*

(ii) *Q is a compact Hausdorff space;*

(iii) *$f : Q \to X$ is continuous;*

(iv) *$\overline{\mathrm{conv}(f(Q))}$ is compact in X (this is automatically true if X is an F-space).*

Then given any Borel probability measure μ on Q, the function $f : Q \longrightarrow X$ is integrable with respect to μ and moreover,

$$y = \int_Q f \ \mathrm{d}\mu = \int_{f(Q)} z \ \mathrm{d}\mu_f(z) \in \overline{\mathrm{conv}(f(Q))},$$

where μ_f is a Borel probability measure on $f(Q)$ given by

$$\mu_f(E) = \mu(f^{-1}(E)), \ \ E \subset f(Q).$$

Conversely, if $y \in \overline{\mathrm{conv}(f(Q))}$, then there exists a regular Borel probability measure μ_f on $f(Q)$ such that

$$y = \int_{f(Q)} z \, \mathrm{d}\mu_f(z).$$

Theorem 2.1.6. *Suppose that Q is a compact Hausdorff space, X is a Banach space, $f : Q \to X$ is continuous, and μ is a positive Borel measure on Q. Then*

$$\left\| \int_Q f \, \mathrm{d}\mu \right\| \leq \int_Q \|f\| \, \mathrm{d}\mu.$$

We now show some preliminary results that would be of great importance in the proof of main results.

Let Q be a compact metric space equipped with a Borel probability measure μ. Let $M(Q)$ denote the space of all finite signed Borel measures on Q. Then $M(Q)$ is a Banach space with respect to the norm

$$\|\nu\| := |\nu|(Q) = \sup\left\{ \left| \int_Q f \, \mathrm{d}\nu \right| : f \in C(Q), \|f\|_{C(Q)} \leq 1 \right\}.$$

Such a Banach space is the dual space of $C(Q)$. Note that $C(Q)$ is a separable Banach space. Let $M(Q)^{w^*}$ denote the space $M(Q)$ endowed with the weak* -topology τ_{w^*}. Then $M(Q)^{w^*}$ is a locally convex topological space with dual space $C(Q)$.

Next, let X_m denote an m-dimensional linear subspace of $C(Q)$. Let $\Sigma_0 \subset M(Q)$ denote the set of all probability measures $\rho \in M(Q)$ of the form

$$\rho = \sum_{j=1}^{m+2} \lambda_j(\rho)\delta_{y_j(\rho)},$$

where $\lambda_j(\rho) \geq 0$, $y_j(\rho) \in Q$ for $j = 1, 2, \ldots, m+2$ and $\sum_{j=1}^{m+2} \lambda_j(\rho) = 1$.

Let $\Sigma \subseteq \Sigma_0$ denote the set of all probability measures $\rho \in \Sigma_0$ such that

$$\int_Q f(x)\,\mathrm{d}\mu(x) = \int_Q f(x)\,\mathrm{d}\rho(x), \quad \forall f \in X_m.$$

Theorem 2.1.7. *There exists a Borel probability measure ν on the space $M(Q)^{w^*}$ which is supported in the set $\Sigma \subset M(Q)$ and satisfies*

$$\mu = \int_\Sigma \rho\,\mathrm{d}\nu(\rho),$$

where the equality holds in the sense that for any $f \in C(Q)$,

$$\int_Q f(x)\,\mathrm{d}\mu(x) = \int_\Sigma \sum_{j=1}^{m+2} \lambda_j(\rho) f(y_j(\rho))\,\mathrm{d}\nu(\rho)$$

and where μ is the probability measure we wish to discretise.

Lemma 2.1.8. *The set Σ is w^*-compact in $M(Q)$.*

Proof. Define

$$S := \left\{ \lambda = (\lambda_1, \ldots, \lambda_{m+2}) \in \mathbb{R}^{m+2} : \lambda_1, \ldots, \lambda_{m+2} \geq 0, \sum_{j=1}^{m+2} \lambda_j = 1 \right\}.$$

Then $S \times Q^{m+2}$ is a compact topological space with respect to the product topology. Next, consider the mapping $T : S \times Q^{m+2} \to M(Q)^{w*}$ that takes $(\lambda, x) \in S \times Q^{m+2}$ to the measure $\sum_{j=1}^{m+2} \lambda_j \delta_{x_j} \in M(Q)$. Note that for any $f \in C(Q)$, and any $(\lambda, x), (\alpha, y) \in S \times Q^{m+2}$, we have

$$\begin{aligned} \left| \left\langle \sum_{j=1}^{m+2} \lambda_j \delta_{x_j} - \sum_{j=1}^{m+2} \alpha_j \delta_{y_j}, f \right\rangle \right| &\leq \sum_{j=1}^{m+2} |\lambda_j f(x_j) - \alpha_j f(y_j)| \\ &\to 0, \text{as } (\alpha, y) \to (\lambda, x). \end{aligned}$$

This implies that the mapping T is continuous, and hence $\Sigma_0 = T(S \times Q^{m+2})$ is w^*-compact.

Finally, for each $f \in C(Q)$, set $\mu_f := \int_Q f \, \mathrm{d}\mu$. Then

$$\Sigma = \{\rho \in \Sigma_0 : \langle f, \rho \rangle = \mu_f, \forall f \in X_m\}.$$

Since each $X_m \subset C(Q)$ and $C(Q)$ is the dual space of $M(Q)^{w^*}$, it follows that Σ is a w^*-closed subset of the w^*-compact set Σ_0. Thus, Σ is a weak*-compact subset of $M(Q)$. $\square$

Lemma 2.1.9. *The probability measure $\mu \in M(Q)$ is in the weak*-closure of the convex hull K of $\Sigma \subset M(Q)^{w^*}$.*

Proof. Assume to the contrary that $\mu \notin K = \overline{\mathrm{conv}\,\Sigma}^{w^*}$. Then by the convex separation theorem, there exists $g \in C(Q)$ such that

$$\int_Q g \, \mathrm{d}\mu > \sup_{\rho \in \Sigma} \int_Q g \, \mathrm{d}\rho. \tag{2.1.1}$$

Let $X_{m+1} = \mathrm{span}\{X_m, g\}$. By Corollary 4.1 of [17], there exist $x_1, x_2, \ldots, x_{m+2} \in Q$ and $\lambda_1, \ldots, \lambda_{m+2} \geqslant 0$ such that $\sum_{j=1}^{m+2} \lambda_j = 1$ and

$$\int_Q f \, \mathrm{d}\mu = \sum_{j=1}^{m+2} \lambda_j f(x_j), \quad \forall f \in X_{m+1}.$$

This implies that $\rho = \sum_{j=1}^{m+2} \lambda_j \delta_{x_j} \in \Sigma$ and $\int_Q g \, \mathrm{d}\mu = \int_Q g d\rho$, which contradicts (2.1.1).

$\square$

Proof. **Proof of Theorem 2.1.7**

Let $X = C(Q)$. Then $M(Q) = X^*$. By Lemma 3.4.2, μ lies in the w^*-closure of the convex hull of Σ; that is, $\mu \in K := \overline{\mathrm{conv}(\Sigma)}^{w^*}$. By Lemma 2.1.8 , Σ is compact in the space (X^*, w^*). Thus, by Theorem 2.1.5, it is enough to show that K is also compact in the space (X^*, w^*). Note that

$$\Sigma \subset \Sigma_0 \subset B_{X^*} := \{\nu \in X^* : \|\nu\| \leqslant 1\},$$

which also implies that $\mathrm{conv}(\Sigma) \subset B_{X^*}$. Since B_{X^*} is compact in the space (X^*, w^*) (By Theorem 2.1.3), it follows that $K := \overline{\mathrm{conv}(\Sigma)}^{w^*}$ is a closed subset of B_{X^*}, which also implies that K is compact in the space (X^*, w^*). The theorem is proved. $\square$

To prove Theorem 1.0.4 and Theorem 1.0.7 in this chapter, we also need the following Bernstein's inequality.

Theorem 2.1.10. *Let $\{\xi_j\}_{j=1}^{\infty}$ be a sequence of independent random variables such that $\mathbb{E}\xi_j = 0$ and $|\xi_j| \leq 1$ for all j. Then for any $\varepsilon > 0$,*

$$\operatorname{Prob}\Big\{\Big|\frac{1}{n}\sum_{j=1}^{n}\xi_j\Big| > \varepsilon\Big\} \leq 2e^{-\frac{n\varepsilon^2}{2}}.$$

2.2 Regular partitions on path-connected compact metric space

To prove Theorem 1.0.2, we first introduce a lemma:

Lemma 2.2.1. *Let (Ω, ρ) be a compact path-connected metric space with diameter π. Then for each $\delta \in (0, \pi)$, there exist a finite set $\Lambda = \{a_1, \ldots, a_M\} \subset \Omega$ with $M > 1$ such that $\Omega = \bigcup_{j=1}^{M} B_\delta(a_j)$ and*

$$\operatorname{dist}(a_j, \Lambda_{j-1}) = \delta, \quad j = 2, 3, \ldots, M,$$

where $\Lambda_k := \{a_1, a_2, \ldots, a_k\}, k = 1, \ldots, M$.

Proof. Since the metric space Ω is path-connected and has diameter $\pi \geq \delta$, there exist two points $a_1, a_2 \in \Omega$ such that $\rho(a_1, a_2) = \delta$. Assume that $\Lambda_n = \{a_1, \ldots, a_n\}$ is a finite subset of Ω such that

$$\operatorname{dist}(a_j, \Lambda_{j-1}) = \delta, j = 2, \ldots, n,$$

where $\Lambda_j = \{a_1, a_2, \ldots, a_j\}$. If $\Omega = \bigcup_{j=1}^{n} B_\delta(a_j)$, then it is sufficient to use $M = n$. Now assume that, in contrast, $\Omega \neq \bigcup_{j=1}^{n} B_\delta(a_j)$. Then there exists a point $y \in \Omega \setminus \Lambda_n$ such that

$$\operatorname{dist}(y, \Lambda_n) \geqslant \delta.$$

Without loss of generality, we may assume that $\operatorname{dist}(y, \Lambda_n) = \rho(y, a_1)$. Let $\gamma : [0, 1] \to \Omega$ be a path such that $\gamma(0) = y$ and $\gamma(1) = a_1$. Define $f(t) := \mathrm{dist}(\gamma(t), \Lambda_n)$ for $t \in [0, 1]$. Clearly, f is a continuous function on $[0, 1]$ with

$$f(0) = \mathrm{dist}(y, \Lambda_n) \geq \delta \ \text{ and } \ f(1) = \mathrm{dist}(a_1, \Lambda_n) = 0.$$

Thus, there exists a point $a_{n+1} = \gamma(t_n) \in \Omega$ for some $t_n \in [0, 1]$ such that

$$\mathrm{dist}(a_{n+1}, \Lambda_n) = f(t_n) = \delta.$$

We may continue this selection procedure with $\Lambda_{n+1} = \{a_1, \dots, a_{n+1}\}$. Since Ω is compact, this procedure must terminate after a finite number of steps. $\square$

Proof of Theorem 1.0.2. Let

$$\{a_1, \dots, a_M\}$$

be a finite subset of Ω as given in Lemma 2.2.1.

For $1 < j \leq M$, let $1 \leq k_j < j$ be an integer such that

$$\mathrm{dist}(a_j, \Lambda_{j-1}) = \rho(a_j, a_{k_j}) = \delta.$$

For each $1 \leq j \leq M$, define

$$V_j := \left\{ x \in \Omega : \rho(x, a_j) = \mathrm{dist}(x, \Lambda) \ \text{ and } \ \mathrm{dist}(x, \Lambda) < \min_{1 \leq i < j} \rho(x, a_i) \right\}.$$

That is, $x \in V_j$ if and only if j is the smallest positive integer such that $\mathrm{dist}(x, \Lambda) = \rho(x, a_j)$. Clearly, the sets V_j are pairwise disjoint,

$$B_{\frac{\delta}{2}}(a_j) \subset V_j \subset B_\delta[a_j], \ \ j = 1, 2, \dots, M, \tag{2.2.1}$$

and $\Omega = \bigcup_{j=1}^M V_j$. Moreover, using (1.0.3), we have

$$\mu(V_j) \geqslant \frac{1}{N}, \quad \forall 1 \leqslant j \leqslant M.$$

Now we construct the desired partition of Ω as follows via a finite number of steps. In the first step, we write $V_j^0 = V_j$ for $j = 1, \dots, M$, and modify the cells V_M and V_{k_M} slightly so that $N\mu(V_M)$ is an integer. Let $E_M \subset V_M^0$ be such that $\mu(E_M) < \frac{1}{N}$ and $N\mu(V_M^0 \setminus E_M)$ is a positive integer. We then update the cells as follows:

$$V_j^1 := \begin{cases} V_j^0, & \text{if } j \neq M \text{ and } j \neq k_M, \\ V_j^0 \setminus E_M, & \text{if } j = M, \\ V_j^0 \cup E_M, & \text{if } j = k_M. \end{cases}$$

Note that the sets V_j^1 are pairwise disjoint, $\Omega = \bigcup_{j=1}^M V_j^1$, $V_j^0 \subset V_j^1$ for $1 \leq j \leq M-1$ and $V_M^1 \subset V_M^0$.

In the second step, we continue the process with the collection of the first $M-1$ updated cells: V_j^1, $1 \leq j \leq M-1$. More precisely, we choose a subset E_{M-1} of V_{M-1}^0 such that $\mu(E_{M-1}) < \frac{1}{N}$ and $N\mu(V_{M-1}^1 \setminus E_{M-1})$ is a positive integer, and then update the cells as follows:

$$V_j^2 := \begin{cases} V_j^1, & \text{if } j \neq M-1 \text{ and } j \neq k_{M-1}, \\ V_j^1 \setminus E_{M-1}, & \text{if } j = M-1, \\ V_j^1 \cup E_{M-1}, & \text{if } j = k_{M-1}. \end{cases}$$

It is very important here that the set E_{M-1} is selected as a subset of V_{M-1}^0 (rather than a general subset V_{M-1}^1) because this way of selection yields a better control of the diameter of the updated cell $V_{k_{M-1}}^1 := E_{M-1} \cup V_{k_{M-1}}^1$.

In general, at the ℓ-th step with $1 \leq \ell < M$, we modify the cells $V_{M-\ell+1}^{\ell-1}$ and $V_{k_{M-\ell+1}}^{\ell-1}$ in a similar manner. Indeed, let $E_{M-\ell+1} \subset V_{M-\ell+1}^0 \subset V_{M-\ell+1}^{\ell-1}$ be such that $\mu(E_{M-\ell+1}) < \frac{1}{N}$ and $N\mu(V_{M-\ell+1}^{\ell-1} \setminus E_{M-\ell+1})$ is a positive integer. We then define

$$V_j^\ell := \begin{cases} V_j^{\ell-1}, & \text{if } j \neq M-\ell+1 \text{ and } j \neq k_{M-\ell+1}, \\ V_{M-\ell+1}^{\ell-1} \setminus E_{M-\ell+1}, & \text{if } j = M-\ell+1, \\ V_{k_{M-\ell+1}}^{\ell-1} \cup E_{M-\ell+1}, & \text{if } j = k_{M-\ell+1}. \end{cases}$$

Clearly, the sets V_j^ℓ are pairwise disjoint, $\Omega = \bigcup_{j=1}^M V_j^\ell$,

$$V_j^0 \subset V_j^{\ell-1} \subset V_j^\ell \quad \text{for} \quad j = 1, 2, \ldots, M - \ell,$$

and for $j = M - \ell + 1, \ldots, M$,

$$V_j^\ell \subset V_j^{\ell-1} \ \text{ and } \ N\mu(V_j^\ell) \text{ is a positive integer.}$$

Furthermore, by the above construction, it is easily seen that for each $1 \leq j \leq M - \ell$,

$$V_j^\ell \subset \bigcup_{\substack{M-\ell+1 \leq k \leq M \\ \rho(a_k, a_j) = \delta}} (V_j^0 \cup V_k^0),$$

which, using (2.2.1), implies that $V_j^\ell \subset B_{2\delta}[a_j]$ and $\mathrm{diam}(V_j^\ell) \leq 4\delta$ for all $1 \leq j \leq M$.

The above process will be terminated after the $(M-1)$-st step, where we obtain pairwise disjoint subsets V_j^{M-1}, $j = 1, 2, \ldots, M$, of Ω with diameter $\leq 4\delta$ such that $\Omega = \bigcup_{j=1}^M V_j^{M-1}$ and $N\mu(V_j^{M-1})$ is a positive integer for $2 \leq j \leq M$. Since μ is a probability measure, we have

$$N = N\mu(\Omega) = \sum_{j=1}^M N\mu(V_j^{M-1}).$$

This implies that $N\mu(V_1^{M-1})$ is a positive integer as well. Since μ is non-atomic, for each $1 \leq j \leq M$, we may write V_j^{M-1} as a disjoint union

$$V_j^{M-1} = \bigcup_{k=1}^{\ell_j} S_{j,k}$$

such that $\mu(S_{j,k}) = \frac{1}{N}$ and $\mathrm{diam}(S_{j,k}) \leq 4\delta$ for $1 \leqslant k \leqslant \ell_j$. This leads to a partition of Ω with the desired properties:

$$\Omega = \bigcup_{j=1}^M \bigcup_{k=1}^{\ell_j} S_{j,k}.$$

$\square$

2.3 Discretization on compact metric spaces

2.3.1 Proof of Theorem 1.0.4

The proof of Theorem 1.0.4 follows along the same idea as that of [4].

Let $\{X_1, \ldots, X_N\}$ be a partition of X satisfying the condition (a). By the inner regularity of the measure μ, for each $1 \leqslant j \leqslant N$, there exists a compact subset $Q_j \subset X_j$ such that

$$\frac{1}{N} - \mu(Q_j) \leqslant \frac{1}{2}(1 + \|\Phi\|_\infty)^{-1} N^{-\frac{3}{2}-\frac{3}{2\beta}}.$$

Let μ_j denote the probability measure on Q_j given by $\mu_j(E) = \frac{\mu(E)}{\mu(Q_j)}$ for each Borel subset $E \subset Q_j$. Then it is easily seen that

$$\sup_{x\in X} \left| \int_X \Phi(\rho(x,y))\,\mathrm{d}\mu(y) - \frac{1}{N}\sum_{j=1}^N \int_{Q_j} \Phi(\rho(x,y))\,\mathrm{d}\mu_j(y) \right| \leqslant N^{-\frac{1}{2}-\frac{3}{2\beta}}. \tag{2.3.1}$$

Let Σ_j denote the set of all Borel probability measures σ_j on Q_j that take the form

$$\sigma_j = \sum_{i=1}^{r+2} \lambda_i(\sigma_j)\delta_{y_i(\sigma_j)}, \quad \lambda_i(\sigma_j) \geq 0, \quad y_i(\sigma_j) \in Q_j, \quad 1 \leqslant j \leqslant r+2,$$

such that $\sum_{i=1}^{r+2} \lambda_i(\sigma_j) = 1$ and

$$\int_{Q_j} f(y)\ \mathrm{d}\mu_j(y) = \sum_{i=1}^{r+2} \lambda_i(\sigma_j) f(y_i(\sigma_j)), \quad \forall f \in V_r. \tag{2.3.2}$$

According to Theorem 2.1.7, there exists a Borel probability measure ν_j on Σ_j such that

$$\int_{Q_j} f\ \mathrm{d}\mu_j = \int_{\Sigma_j} \sum_{i=1}^{r+2} \lambda_i(\sigma_j) f(y_i(\sigma_j))\,\mathrm{d}\nu_j(\sigma_j), \qquad \forall f \in C(Q_j). \tag{2.3.3}$$

Now we consider the following product probability space:

$$(\widetilde{\Sigma}, \nu) = \prod_{j=1}^{N} (\Sigma_j, \nu_j).$$

We first claim that for each fixed $x \in X$ and parameter $t > \sqrt{\log 2}$, there exists a subset $G(x) \subset \widetilde{\Sigma}$ with $\nu(G(x)) \leq 2e^{-t^2} < 1$ such that for each $\sigma := (\sigma_1, \sigma_2, \ldots, \sigma_N) \in \widetilde{\Sigma} \setminus G(x)$,

$$\begin{aligned} &\left| \frac{1}{N} \sum_{j=1}^{N} \sum_{i=1}^{r+2} \lambda_i(\sigma_j) \Phi(\rho((x, y_i(\sigma_j))) - \frac{1}{N} \sum_{j=1}^{N} \int_{Q_j} \Phi(\rho(x, y)) \, \mathrm{d}\mu_j(y) \right| \\ &\leqslant \frac{4}{\sqrt{3}} c_1 \sqrt{c_1 c_2 \ell} t N^{-\frac{1}{2} - \frac{3}{2\beta}}. \end{aligned} \tag{2.3.4}$$

To show this claim, we consider the following independent random variables on the probability space $(\widetilde{\Sigma}, \nu)$:

$$h_j(\sigma) \equiv h_j(\sigma_j) := \sum_{i=1}^{r+2} \lambda_i(\sigma_j) \Phi\Big(\rho(x, y_i(\sigma_j))\Big) - \int_{Q_j} \Phi(\rho(x, y)) \, \mathrm{d}\mu_j(y),$$

where $\sigma = (\sigma_1, \ldots, \sigma_N) \in \widetilde{\Sigma}$ and $j = 1, \ldots, N$. By (1.0.5) and (2.3.3), we have

$$\mathbb{E} h_j = 0, \quad |h_j| \leq \mathrm{diam}(X_j) \leq \delta_N, \ \ 1 \leq j \leq N.$$

For each $1 \leq j \leq N$, pick a point $y_j \in Q_j$ and set $R_j := B_{\delta_N}[y_j]$ so that $Q_j \subset X_j \subset R_j$. Set

$$S_i(x) := E(x; t_{i-1}, t_i) = \Big\{ y \in X : t_{i-1} \leq \rho(x, y) \leq t_i \Big\}, \quad i = 1, \ldots, \ell.$$

Note that if $R_j \subseteq S_k(x)$ for some $1 \leqslant k \leqslant \ell$ and $1 \leq j \leq N$, then there exists a function $f_{k,x} \in V_r$ such that

$$\Phi(\rho(x, \cdot))\Big|_{Q_j} = f_{k,x}\Big|_{Q_j},$$

which, using (2.3.2), implies that

$$h_j(\sigma_j) = \sum_{i=1}^{r+2} \lambda_i(\sigma_j) f_{k,x}(y_i(\sigma_j)) - \int_{Q_j} f_{k,x}(y)\, \mathrm{d}\mu_j(y) = 0.$$

For $1 \leqslant k \leqslant \ell - 1$ and if $\ell > 1$, let

$$E_k(x) := \{ y \in X : \ t_k - \delta_N \leqslant \rho(x,y) \leqslant t_k + \delta_N \} .$$

Denote by I the set of all positive integers $1 \leqslant j \leqslant N$ such that

$$y_j \in \bigcup_{k=1}^{\ell-1} E_k(x).$$

Let $I^c = \{1,2,\dots,N\} \setminus I$. Note that if $j \in I^c$, then there exists $1 \leq k \leq \ell$ such that $R_j \subset S_k(x)$, which implies $h_j = 0$. Furthermore, since

$$\bigcup_{j\in I} X_j \subseteq \bigcup_{j\in I} R_j \subseteq \bigcup_{k=1}^{\ell-1} \{ y \in X : \quad t_k - 2\delta_N \leqslant \rho(x,y) \leqslant t_k + 2\delta_N \} ,$$

it follows by Condition (b) that

$$\#I \leq 2c_2 \ell N \delta_N = 2c_2 c_1 \ell N^{1-\beta^{-1}}.$$

We shall use this in our next estimate. Now setting

$$\xi_j = \frac{1}{\delta_N} h_j, \quad j = 1,2,\dots,N,$$

and using the Bernstein inequality in probability, we obtain that for any $\varepsilon > 0$,

$$\begin{aligned} \text{Prob}\left\{ \frac{1}{N} \left| \sum_{j=1}^{N} \xi_j \right| > \varepsilon \right\} &= \text{Prob}\left\{ \frac{1}{\#I} \left| \sum_{j\in I} \xi_j \right| > \frac{\varepsilon N}{\#I} \right\} \\ &\leq 2\exp\left(-\frac{3}{8} (\#I) \frac{\varepsilon^2 N^2}{(\#I)^2} \right) \leq 2\exp\left(-\frac{3\varepsilon^2 N^{1+\beta^{-1}}}{16 c_1 c_2 \ell} \right). \end{aligned}$$

It follows that for any $\delta > 0$,

$$\operatorname{Prob}\left\{\frac{1}{N}\left|\sum_{j=1}^{N} h_j\right| > \delta\right\} \leq 2\exp\left(-\frac{3\delta^2 N^{1+3\beta^{-1}}}{16c_1^3 c_2 \ell}\right).$$

Given a parameter $t > 0$, setting

$$\delta := \frac{4}{\sqrt{3}} c_1 \sqrt{c_1 c_2 \ell} N^{-\frac{1}{2}-\frac{3}{2\beta}} t,$$

we conclude that the inequality

$$\frac{1}{N}\left|\sum_{j=1}^{N} h_j\right| \leq \frac{4}{\sqrt{3}} c_1 \sqrt{c_1 c_2 \ell} \cdot t \cdot N^{-\frac{1}{2}-\frac{3}{2\beta}}$$

holds with probability at least $1 - 2e^{-t^2}$ on the probability space $(\widetilde{\Sigma}, \nu)$. This proves the claim (2.3.4).

Now let $t := \sqrt{A \log N} \geq \sqrt{\log 2}$ with $A > 1$ being a parameter to be specified later. By (2.3.1) and (2.3.4), for each $x \in X$, there exists a set $G(x) \subset \widetilde{\Sigma}$ with $\nu(G(x)) \leq 2N^{-A}$ such that for each

$$\sigma = (\sigma_1, \ldots, \sigma_N) \in \widetilde{\Sigma} \setminus G(x),$$

$$\begin{aligned}
&\left|\frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{r+2} \lambda_i(\sigma_j)\Phi(\rho(x, y_i(\sigma_j))) - \Phi_0(x)\right| \\
&\leq \frac{7}{2} c_1 \sqrt{c_1 c_2 \ell}\sqrt{A}\sqrt{\log N} N^{-\frac{1}{2}-\frac{3}{2\beta}},
\end{aligned} \tag{2.3.5}$$

where

$$\Phi_0(x) := \int_X \Phi(\rho(x, y))\, \mathrm{d}\mu(y).$$

Let M be a positive integer such that

$$M - 1 < c_1^{\beta} N^{\frac{3}{2}+\frac{\beta}{2}} \leq M.$$

Then, using Condition (a) with M in place of N, we obtain a partition

$$\{X_1', X_2', \ldots, X_M'\}$$

of X such that $\mu(X_j') = \frac{1}{M}$ and

$$\mathrm{diam}(X_j') \leq \delta_M = c_1 M^{-\beta^{-1}} \leq N^{-\frac{1}{2}-\frac{3}{2\beta}}$$

for each $1 \leq j \leq M$. Choose $z_j \in X_j'$ for each $1 \leq j \leq M$, and let $G = \bigcup_{k=1}^{M} G(z_k)$. Then

$$\nu(G) \leq \sum_{j=1}^{M} \nu(G(z_j)) \leq 2MN^{-A} \leq 3c_1^{\beta} N^{\frac{\beta}{2}+\frac{3}{2}-A}.$$

Thus, setting $A = \frac{1+2c_1}{2}\beta + \frac{3}{2}$, we obtain that for $N \geq 4$, $\nu(G)$ is at most

$$3c_1^{\beta} N^{-c_1\beta} \leq \left(\frac{3c_1}{4^{c_1}}\right)^{\beta} < 1.$$

Finally, using (1.0.5), we have that for each $\sigma = (\sigma_1, \ldots, \sigma_N) \in \widetilde{\Sigma} \setminus G$

$$\begin{aligned}
&\sup_{x \in X} \left| \frac{1}{N} \sum_{j=1}^{N} \sum_{i=1}^{r+2} \lambda_i(\sigma_j) \Phi(\rho(x, y_i(\sigma_j))) - \Phi_0(x) \right| \\
&\leq \max_{1 \leq k \leq M} \left| \frac{1}{N} \sum_{j=1}^{N} \sum_{i=1}^{r+2} \lambda_i(\sigma_j) \Phi(\rho(z_k, y_i(\sigma_j))) - \Phi_0(x) \right| + \delta_M,
\end{aligned}$$

which, using (2.3.5), is estimated from above by

$$\left(\frac{7}{2} c_1^{\frac{3}{2}} (c_2\ell)^{\frac{1}{2}} \sqrt{\frac{2c_1+1}{2}\beta + \frac{3}{2}} + 1\right) N^{-\frac{1}{2}-\frac{3}{2\beta}} \sqrt{\log N} \leq 8c_1^2 (c_2\ell)^{\frac{1}{2}} \sqrt{\beta} N^{-\frac{1}{2}-\frac{3}{2\beta}} \sqrt{\log N}.$$

This completes the proof.

2.3.2 Proof of Theorem 1.0.5

Let $h(x) = b(2 + g(x))$, where b is a normalizing constant so that $\|h\|_{L^1(\mathrm{d}\mu)} = 1$. Clearly,

$$\frac{1}{3} \leq b \leq h(x) \leq 3b \leq 3, \ \ \forall x \in X, \tag{2.3.6}$$

because $\|g\|_\infty \leq 1$. Let τ denote the Borel probability measure given by $\mathrm{d}\tau = h\mathrm{d}\mu$. By (1.0.4), we have that for $N \geq 15$,

$$\tau(B_{\widetilde{\delta}_N/8}(x)) \geq b\mu(B_{\widetilde{\delta}_N/8}(x)) \geq \frac{1}{N}, \ \ x \in X, \tag{2.3.7}$$

where

$$\widetilde{\delta}_N = c_1([Nb])^{-\beta^{-1}} \leq \left(\frac{5}{4b}\right)^{1/\beta} c_1 N^{-\beta^{-1}} \leq \frac{5}{4b} c_1 N^{-\beta^{-1}},$$

because $\beta \geq 1$. Furthermore, by (1.0.6), we have that for each $x \in X$ and $\delta \in (0, \pi)$,

$$\tau\left(\bigcup_{j=1}^{\ell-1} \Big\{ y \in X : \ t_j - \delta \leqslant \rho(x,y) \leqslant t_j + \delta \Big\}\right) \leqslant 3bc_2\ell\delta. \tag{2.3.8}$$

Since X is a compact path-connected metric space, using Theorem 1.0.4 with τ in place of μ, we may find points $y_1, \ldots, y_{(r+2)N} \in X$ and nonnegative real numbers $a_1, \ldots, a_{(r+2)N}$, such that

$$\max_{x\in X} \left| \int_X \Phi(\rho(x,y))h(y)\,\mathrm{d}\mu(y) - \sum_{j=1}^{(r+2)N} a_j \Phi(\rho(x,y_j)) \right| \leqslant \frac{25\sqrt{3}}{16b^{\frac{3}{2}}} c_3 N^{-\frac{1}{2}-\frac{3}{2\beta}} \sqrt{\log N}.$$

On the other hand, using Theorem 1.0.4, we can also find points $z_1, \ldots, z_{(r+2)N} \in X$ and nonnegative real numbers $b_1, \ldots, b_{(r+2)N}$, such that

$$\max_{x\in X} \left| \int_X \Phi(\rho(x,y))\,\mathrm{d}\mu(y) - \sum_{j=1}^{(r+2)N} b_j \Phi(\rho(x,z_j)) \right| \leqslant c_3 N^{-\frac{1}{2}-\frac{3}{2\beta}} \sqrt{\log N}.$$

Since

$$\int_X \Phi(\rho(x,y))g(y)\,\mathrm{d}\mu(y) = \frac{1}{b}\int_X \Phi(\rho(x,y))h(y)\,\mathrm{d}\mu(y) - 2\int_X \Phi(\rho(x,y))\,\mathrm{d}\mu(y)$$

and $\frac{1}{3} \le b \le 1$, it follows that

$$\begin{aligned}
&\sup_{x\in X}\Big|\int_X \Phi(\rho(x,y))g(y)\,\mathrm{d}\mu(y) - \frac{1}{b}\sum_{j=1}^{(r+2)N} a_j\Phi(\rho(x,y_j)) + 2\sum_{j=1}^{(r+2)N} b_j\Phi(\rho(x,z_j))\Big| \\
&\quad\le \left(\frac{25\sqrt{3}}{16b^{\frac{5}{2}}}+2\right)c_3N^{-\frac{1}{2}-\frac{3}{2\beta}}\sqrt{\log N} \le 45c_3N^{-\frac{1}{2}-\frac{3}{2\beta}}\sqrt{\log N}.
\end{aligned}$$

The theorem is proved.

2.4 Discretization on finite-dimensional compact domains

2.4.1 Proof of Theorem 1.0.7

The main idea of our proof comes from the paper [4]. We need the following Besicovitch covering theorem on finite-dimensional normed linear spaces [26]:

Lemma 2.4.1. [24] *Let $E \subset X$ be an arbitrarily given nonempty subset of a finite dimensional normed linear space X. Assume that for each $x \in E$ there exists a closed ball $B_{r(x)}[x]$ with centre x and radius $r(x) > 0$. Assume in addition that $\sup_{x\in E} r(x) < \infty$. Then there exists a sub-collection $\mathcal{R}$ of the closed balls $B_{r(x)}[x]$, $x \in E$, which covers the set E and can be written in the form*

$$\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \cdots \cup \mathcal{R}_m$$

with $m \le \mathcal{N}(X)$, and each $\mathcal{R}_j$ being a collection of pairwise disjoint balls, $1 \le j \le m$. Here $\mathcal{N}(X)$ is a positive constant depending only on the normed space $(X, \|\cdot\|)$.

The best constant $\mathcal{N}(X)$ for the Besicovitch covering theorem has been well studied in literature (see [24, 26] and the references thererin). In the case when $(X, \|\cdot\|) = \mathbb{R}^d$, it was known [26] that $\mathcal{N}(X) \leq 6^d$. The sharp estimate of this constant appears in [33]. A much more general version of the Besicovitch covering theorem can be found in [23].

The proof runs along the same line as that of Theorem 1.0.4. We sketch it as follows.

Without loss of generality, we may assume that $g \geq 0$ since otherwise we may write $g = g^+ - g^-$ with $g^\pm \geq 0$. We may also assume that $N \geq 4\mathcal{N}(X)(r+2)$ since otherwise the stated estimates hold trivially. For the rest of the proof, the letter C denotes a general positive constant depending only on $\mathcal{N}(X)$, c_4, c_5, r, ℓ and β.

Let τ denote the probability measure given by $\mathrm{d}\tau(x) = g(x)\,\mathrm{d}\mu(x)$. Let $n_1 = [\frac{n}{2\mathcal{N}(X)(r+2)}]$. For $x \in \Omega$, let $0 < \theta_x \leq \delta_{n_1} := (c_4/n_1)^{\frac{1}{\beta}}$ be such that

$$\int_{B_{\theta_x}[x]} (1 + g(y))\,\mathrm{d}\mu(y) = \frac{1}{n_1}. \tag{2.4.1}$$

By the Besicovitch covering theorem, we can find finitely many open balls $B_j = B_{\theta_{x_j}}(x_j)$, $j = 1, 2, \ldots, m$, such that $\Omega \subset \bigcup_{j=1}^m B_j$,

$$\{B_1, \ldots, B_m\} = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \cdots \cup \mathcal{R}_{\mathcal{N}(X)} \tag{2.4.2}$$

with each $\mathcal{R}_j$ being a subcollection of pairwise disjoint balls. By (2.4.1) and (2.4.2), we then have $m \leq 2\mathcal{N}(X)n_1 \leq \frac{n}{r+2}$. Note that $\mu(B_r(x)) = \mu(B_r[x])$ for any $x \in \Omega$ and $r > 0$, since by (3.2.6)

$$\mu(B_r[x] \setminus B_r(x)) = \lim_{j\to\infty} \mu\Big\{y \in X : \, r - j^{-1} \leq \rho(x, y) \leq r + j^{-1}\Big\} = 0.$$

Now define $Q_1 = \overline{B_1} \cap \Omega$ and

$$Q_j = \Big(\overline{B_j} \setminus \bigcup_{i=1}^{j-1} B_i\Big) \cap \Omega, \;\; j = 2, \ldots, m.$$

Then $\Omega = \bigcup_{j=1}^m Q_j$, $\tau(Q_i \cap Q_j) = 0$ for $1 \leq i \neq j \leq m$, $Q_j \subset \overline{B_j}$ and $\tau(Q_j) \leq \frac{1}{n_1}$ for $1 \leq j \leq m$. Without loss of generality, we may also assume that $\tau(Q_j) > 0$ for each $1 \leq j \leq m$, since otherwise we remove Q_j from the partition.

For each $1 \le j \le m$, let Σ_j denote the set of all probability measures σ_j on Q_j of the form

$$\sigma_j = \sum_{i=1}^{r+2} \lambda_i(\sigma_j)\delta_{y_i(\sigma_j)}, \;\; \lambda_i(\sigma_j) \ge 0, \;\; y_i(\sigma_j) \in Q_j,$$

such that

$$\frac{1}{\tau(Q_j)} \int_{Q_j} P(x)\, \mathrm{d}\tau(x) = \sum_{i=1}^{r+2} \lambda_i(\sigma_j) P(y_i(\sigma_j)), \;\; \forall P \in X_r.$$

By Theorem 2.1.7, there exists a Borel probability measure ν_j on Σ_j such that

$$\int_{Q_j} f(x)\, \mathrm{d}\tau(x) = \int_{\Sigma_j} \sum_{i=1}^{r+2} \tau(Q_j)\lambda_i(\sigma_j) f(y_i(\sigma_j))\, \mathrm{d}\nu_j(\sigma_j), \;\; \forall f \in C(Q_j).$$

Now we consider the product probability space $(\widetilde{\Sigma}, \nu) = \prod_{j=1}^m (\Sigma_j, \nu_j)$. Fix $x \in \Omega$ temporarily. For $1 \le j \le m$, define

$$h_{j,x}(\sigma_j) = \tau(Q_j) \sum_{i=1}^{r+2} \lambda_i(\sigma_j) \Phi(\|x - y_i(\sigma_j)\|) - \int_{Q_j} \Phi(\|x - y\|)\, \mathrm{d}\tau(y).$$

Then $\mathbb{E} h_{j,x} = 0$,

$$|h_{j,x}(\sigma_j)| \le \tau(Q_j) \cdot \mathrm{diam}(Q_j) \le \tau(B_j)\mathrm{diam}(Q_j) \le C\theta_{x_j} n^{-1}. \tag{2.4.3}$$

For $0 < \theta \le 2\delta_{n_1}$, we denote by $I_\theta := I_\theta(x)$ the set of all integers $1 \le j \le m$ such that $\theta/2 < \theta_{x_j} \le \theta$ and $t_k - \theta \le \|x - x_j\| \le t_k + \theta$ for some $1 \le k \le \ell$. Note that if $\theta/2 < \theta_{x_j} \le \theta$ and $j \notin I_\theta(x)$, then there exists k in the interval $1 \le k \le \ell$ such that $t_{k-1} \le \|x - y\| \le t_k$ for every $y \in Q_j \subset B_j := B_{\theta_{x_j}}(x_j)$, which implies that $h_{j,x} \equiv 0$. Note also that

$$\bigcup_{j \in I_\theta} \Big(B_j \cap \Omega \Big) \subset \Big\{ y \in \Omega : \; t - 2\theta \le \|x - y\| \le t + 2\theta \Big\}.$$

It then follows by (2.4.2) and (3.2.6) that

$$\# I_\theta c_4 \left(\frac{\theta}{2}\right)^\beta \le \sum_{j \in I_\theta} \mu(B_j) \le 4\mathcal{N}(X) c_5 \theta \ell,$$

which implies that

$$\#I_\theta \le C_1 \theta^{1-\beta}. \tag{2.4.4}$$

Note that (2.4.4) holds trivially if

$$\theta \le \left(\frac{(r+2)C_1}{n}\right)^{\frac{1}{\beta-1}} = C_2 n^{-\frac{1}{\beta-1}} \tag{2.4.5}$$

since $\#I_\theta \le m \le \frac{n}{r+2}$. Thus, we will mainly consider those index sets I_θ with

$$C_2 n^{-\frac{1}{\beta-1}} \le \theta \le 2\delta_{n_1} := 2\left(\frac{c_4}{n_1}\right)^{\frac{1}{\beta}}, \tag{2.4.6}$$

the second bound being the bound on θ stated at the beginning of the paragraph.

To be more precise, let k_0, k_1 be integers such that

$$2^{k_0} < 2^{-1}(n_1/c_4)^{\frac{1}{\beta}} \le 2^{k_0+1}$$

and

$$2^{k_1-1} < C_2^{-1} n^{\frac{1}{\beta-1}} \le 2^{k_1}.$$

Define $J_k = J_k(x) := I_{2^{-k}}(x)$ for $k_0 \le k \le k_1$ and

$$J_{k_1+1} \equiv J_{k_1+1}(x) = \bigcup_{k=k_1+1}^{\infty} I_{2^{-k}}(x).$$

Then by (2.4.4) and the remark after (2.4.4), we have

$$\#J_k \le n_k := C_1^{-1} 2^{k(\beta-1)}, \ \ k_0 \le k \le k_1 + 1. \tag{2.4.7}$$

Moreover, by (2.4.3), we have

$$|h_{j,x}| \le C\theta_{x_j} n^{-1} \le C 2^{-k} n^{-1}, \ \ j \in J_k, \ \ k_0 \le k \le k_1 + 1. \tag{2.4.8}$$

Thus, using (2.4.8), (2.4.7), and the Bernstein inequality, we conclude that for each $k_0 \leq k \leq k_1 + 1$ and each $\varepsilon_k > 0$, the inequality

$$\Big|\sum_{j \in J_k} h_{j,x}(\sigma_j)\Big| > \varepsilon_k$$

holds with probability at most

$$2\exp\Big(-C\varepsilon_k^2 n^2 2^{-k(\beta-3)}\Big). \tag{2.4.9}$$

Now we write

$$\sum_{j=1}^{m} h_{j,x}(\sigma_j) = \sum_{k=k_0}^{\infty} \sum_{\{j: 2^{-k} \leq \theta_{x_j} \leq 2^{-k+1}\}} h_{j,x}(\sigma_j) = \sum_{k=k_0}^{k_1+1} \sum_{j \in J_k} h_{j,x}(\sigma_j),$$

where we used the fact that $\theta_{x_j} \leq \left(\frac{c_4}{n_1}\right)^{\frac{1}{\beta}} < 2^{-k_0-1}$ for all $1 \leq j \leq m$ in the first step, and the fact that $h_{j,x} \equiv 0$ if $2^{-k-1} \leq \theta_{x_j} \leq 2^{-k}$ and $j \notin I_{2^{-k}}(\theta)$ in the last step.

Given $\varepsilon > 0$, let $\{\varepsilon_k\}_{k=k_0}^{k_1+1}$ be a sequence of positive numbers such that $\sum_{k=k_0}^{k_1+1} \varepsilon_k \leq \varepsilon$. Then using (2.4.9), we have

$$\begin{aligned}\mathrm{Prob}\Big\{|\sum_{j=1}^{m} h_{j,x}| > \varepsilon\Big\} &\leq \sum_{k=k_0}^{k_1+1} \mathrm{Prob}\Big\{|\sum_{j \in J_k} h_{j,x}| > \varepsilon_k\Big\} \\ &\leq 2\sum_{k=k_0}^{k_1+1} \exp\Big(-C\varepsilon_k^2 n^2 2^{-k(\beta-3)}\Big). \end{aligned} \tag{2.4.10}$$

Noting that $k_0 \sim k_1 \sim \log n$, we may choose for $k_0 \leq k \leq k_1 + 1$,

$$\varepsilon_k = \begin{cases} 2^{\frac{\beta-3}{2}(k-k_1)}\varepsilon, & \text{if } \beta > 3, \\ \frac{\varepsilon}{\log n}, & \text{if } \beta = 3, \\ 2^{(k-k_0)\frac{\beta-3}{2}}\varepsilon, & \text{if } \beta < 3. \end{cases}$$

We use here that $n \neq 1$ so that $\log n \neq 0$.

For simplicity, we shall assume that $\beta > 3$. The proof below with slight modifications works equally well for the case $\beta \le 3$. We then obtain from (2.4.10) that

$$\mathrm{Prob}\Big\{\Big|\sum_{j=1}^{m} h_{j,x}\Big| > \varepsilon\Big\} \le C(\log n)\exp\Big(-Cn^{2-\frac{\beta-3}{\beta-1}}\varepsilon^2\Big).$$

Setting

$$\varepsilon = C^{-\frac{1}{2}} t \cdot n^{-\frac{\beta+1}{2(\beta-1)}} \quad \text{with } \ t > 0,$$

we conclude that for each $x \in \Omega$, the inequality

$$\begin{aligned}&\Big|\sum_{j=1}^{m}\tau(Q_j)\sum_{i=1}^{r+2}\lambda_i(\sigma_j)\Phi(\|x-y_i(\sigma_j)\|) - \int_\Omega \Phi(\|x-y\|)\,\mathrm{d}\tau(y)\Big| \\ &\ge Ctn^{-\frac{\beta+1}{2(\beta-1)}}\end{aligned}$$

holds with probability bounded above by a multiple of $(\log n)e^{-t^2}$. Let

$$t := \sqrt{A\log n} \ \text{ with } \ A = \frac{\beta(\beta+1)}{2(\beta-1)} > 1.$$

The last inequality holds since $\beta^2 - \beta + 2$ has no real zeos.

We further conclude that for each $x \in \Omega$, there exists a set $G(x) \subset \Sigma$ with $\nu(G(x)) \le C_2(\log n)n^{-A}$ such that for any $\sigma = (\sigma_1, \ldots, \sigma_m) \in \Sigma \setminus G(x)$,

$$\begin{aligned}&\Big|\sum_{j=1}^{m}\tau(Q_j)\sum_{i=1}^{r+2}\lambda_i(\sigma_j)\Phi(\|x-y_i(\sigma_j)\|) - \int_\Omega \Phi(\|x-y\|)\,\mathrm{d}\tau(y)\Big| \\ &\le C^{-1/2}\sqrt{A}n^{-\frac{\beta+1}{2(\beta-1)}}(\log n)^{\frac{1}{2}}.\end{aligned}$$

Finally, let $\{z_1, \ldots, z_L\}$ be a maximal ε_1-separated subset of Ω with $\varepsilon_1 := n^{-\frac{\beta+1}{2(\beta-1)}}(\log n)^{\frac{1}{2}}$. By (1.0.7), we have

$$L \le c_4\Big(\frac{2}{\varepsilon_1}\Big)^{\beta} \le C_3 n^{\frac{\beta(\beta+1)}{2(\beta-1)}}(\log n)^{-\frac{1}{2}\beta}.$$

Setting $A = \frac{\beta(\beta+1)}{2(\beta-1)}$, we have that

$$\sum_{j=1}^{L} \nu(G(z_j)) \leq C_2 C_3 (\log n)^{1-\frac{1}{2}\beta}.$$

Since $\beta > 3$, it follows that the following inequality holds with positive probability:

$$\sup_{x\in\Omega}\Big|\sum_{j=1}^{m} \tau(Q_j) \sum_{i=1}^{r+2} \lambda_i(\sigma_j)\Phi(\|x - y_i(\sigma_j)\|) - \int_{\Omega} \Phi(\|x - y\|)\, \mathrm{d}\tau(y)\Big|$$
$$\leq C n^{-\frac{\beta+1}{2(\beta-1)}} (\log n)^{\frac{1}{2}}.$$

The theorem is proved.

2.5 Discretization on the unit sphere $\mathbb{S}^d$

In this section, we will estimate the constants c_1 and c_2 for the unit sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ equipped with the normalized surface Lebesgue measure μ_d and the geodesic distance $\rho(x, y) = \arccos(x \cdot y), x, y \in \mathbb{S}^d$. We will prove on the unit sphere $\mathbb{S}^d$ that

$$c_1 \leq 40\pi, \;\; c_2 \leq \frac{3}{2}\sqrt{d}, \;\; \alpha = \frac{1}{d}. \tag{2.5.1}$$

The main point here lies in the fact that the upper bounds for c_1 and $c_2/\sqrt{d}$ are independent of the dimension d.

By (2.5.1), we also have

$$45c_3 = 45 \cdot 8c_1^2 c_2^{\frac{1}{2}} \sqrt{d} \leq 7 \times 10^6 d^{\frac{3}{4}}. \tag{2.5.2}$$

As a consequence of Theorem 1.0.2 and Lemma 2.5.4, we have that

Theorem 2.5.1. *For each integer $N \geq 1$, there exists a partition $\{R_1, \ldots, R_N\}$ of $\mathbb{S}^d$ such that*

(i) *the R_j are pairwise disjoint subsets of $\mathbb{S}^d$;*

(ii) *for each* $1 \le j \le N$, $\mu_d(R_j) = \frac{1}{N}$ *and diam*$(R_j) \le 40\pi N^{-\frac{1}{d}}$.

Again, the main point here is that the upper bound for $N^{\frac{1}{d}} \max_j \text{diam}(R_j)$ is independent of the dimension d. Theorem 2.5.1 with dimension dependant upper bound for $N^{\frac{1}{d}} \max_j \text{diam}(R_j)$ can be found in [2].

Theorem 2.5.2. *Let* $\Phi : [-1, 1] \to \mathbb{R}$ *be a piecewise polynomial of degree at most* r *with knots* $-1 = s_0 < s_1 < \cdots < s_\ell = 1$ *such that* $|\Phi(s) - \Phi(s')| \le |s - s'|$ *for any* $s, s' \in [-1, 1]$. *Let* $m_r = m_r^d$ *denote the dimension of the space of all spherical polynomials of degree at most* r *on* $\mathbb{S}^d$. *Let* $g \in L^\infty(\mathbb{S}^d)$ *be such that* $\|g\|_\infty \le 1$. *Then for each positive integer* $N \ge 20$, *there exist points* $\xi_1, \ldots, \xi_{2(m_r+2)N} \in \mathbb{S}^d$ *and real numbers* $\lambda_1, \ldots, \lambda_{2(m_r+2)N}$ *such that*

$$\max_{x \in \mathbb{S}^{d-1}} \left| \int_{\mathbb{S}^d} \Phi(x \cdot y) g(y) \, \mathrm{d}\mu_d(y) - \sum_{j=1}^{2(m_r+2)N} \lambda_j \Phi(x \cdot \xi_j) \right|$$
$$\le 7 \cdot 10^6 \sqrt{\ell} \cdot d^{\frac{3}{4}} N^{-\frac{1}{2} - \frac{3}{2d}} \sqrt{\log N}.$$

In the case when $\Phi(t) = |t|$, Theorem 2.5.2, but with constants depending on the dimension of the sphere, was previously obtained in [4].

2.5.1 Proof of (2.5.1)

For $\theta \in (0, \pi)$ and x in the d-dimensional sphere $\mathbb{S}^d$, set

$$B_\theta(x) := \{y \in \mathbb{S}^d : \rho(x, y) < \theta\}, \text{ and } B_\theta[x] := \{y \in \mathbb{S}^d : \rho(x, y) \le \theta\}.$$

Let $\omega_d := \frac{2\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})}$ denote the surface area of $\mathbb{S}^d$. Using the following known estimates on gamma functions [1],

$$x^{1-s} < \frac{\Gamma(x+1)}{\Gamma(x+s)} < (x+1)^{1-s}, \quad x > 0, \quad s \in (0, 1),$$

we have that

$$\pi^{-\frac{1}{2}} \left(\frac{d-1}{2}\right)^{\frac{1}{2}} \le \frac{\omega_{d-1}}{\omega_d} = \frac{\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2})\sqrt{\pi}} \le \pi^{-\frac{1}{2}} \left(\frac{d+1}{2}\right)^{\frac{1}{2}}. \tag{2.5.3}$$

Lemma 2.5.3. *For* $0 < \theta \le \frac{\pi}{4}$ *and* $x \in \mathbb{S}^d$,

$$\frac{1}{\sqrt{2d}} \le \frac{\mu_d(B_\theta(x))}{\sin^d \theta} \le \frac{2}{\sqrt{d}}.$$

Proof. For $\theta \in (0, \pi]$, we have

$$\mu_d(B_\theta(x)) = \frac{\omega_{d-1}}{\omega_d} \int_{\cos\theta}^{1} (1-t^2)^{\frac{d-2}{2}}\,\mathrm{d}t = \frac{\omega_{d-1}}{\omega_d} \int_0^{\sin^2\theta} t^{\frac{d-2}{2}}(1-t)^{-\frac{1}{2}}\,\mathrm{d}t.$$

If $0 < \theta \le \frac{\pi}{4}$, then for any $0 \le t \le \sin^2\theta$, we have

$$1 \le (1-t)^{-\frac{1}{2}} \le \sqrt{2}.$$

Thus,

$$\frac{\omega_{d-1}}{\omega_d}\frac{2}{d}(\sin\theta)^d \le \mu_d(A(x,\theta)) \le \frac{\omega_{d-1}}{\omega_d}\frac{2\sqrt{2}}{d}(\sin\theta)^d,$$

which, using (2.5.3), implies that

$$\frac{1}{2\sqrt{d}} \le \sqrt{\frac{1}{\pi}}d^{-\frac{1}{2}} \le \sqrt{2}\pi^{-\frac{1}{2}}d^{-1}(d-1)^{\frac{1}{2}} \le \frac{\mu_d(A(x,\theta))}{\sin^d\theta} \le 2\pi^{-\frac{1}{2}}d^{-1}(d+1)^{\frac{1}{2}} \le \frac{2}{\sqrt{d}}.$$

$\square$

The following lemma shows that $c_1 \le 40\pi$:

Lemma 2.5.4. *For any positive integer* N,

$$\inf_{x\in\mathbb{S}^d} \mu_d(B_{\delta_N}(x)) \ge \frac{1}{N} \;\; \textit{with} \;\; \delta_N := 5\pi N^{-\frac{1}{d}}. \tag{2.5.4}$$

Proof. We consider the following two cases:

Case 1. $N \ge 2^{\frac{d}{2}+1}\sqrt{d}$.

In this case, set

$$\delta := \min\left\{\theta : \quad 0 \le \theta \le \frac{\pi}{4}, \ \frac{1}{2\sqrt{d}} \sin^d \theta \geqslant \frac{1}{N}\right\}$$

and our condition on the N ensures that δ be well-defined. Using Lemma 2.5.3, we have that

$$\mu_d(B_\delta(x)) \geqslant \frac{1}{N}, \quad \forall x \in \mathbb{S}^d.$$

It remains to estimate the constant δ. By definition of δ, we have that

$$\frac{1}{2\sqrt{d}} \sin^d \delta \geqslant \frac{1}{N} > \frac{1}{2\sqrt{d}} \sin^d \frac{\delta}{2}.$$

This implies that

$$\delta \leqslant \pi \sin \frac{\delta}{2} < \pi \left(\frac{2\sqrt{d}}{N}\right)^{\frac{1}{d}} \leqslant 2\pi \mathrm{e}^{\frac{1}{2\mathrm{e}}} N^{-\frac{1}{d}} < 3\pi N^{-\frac{1}{d}}.$$

Here we have used the fact that the maximum of $(\log y)/y$ is attained at $y = \mathrm{e}$.

Case 2. $1 \le N < 2^{\frac{d}{2}+1}\sqrt{d}$.

In this case,

$$N^{-\frac{1}{d}} > 2^{-\frac{1}{d}-\frac{1}{2}} d^{-\frac{1}{2d}} \ge 2^{-\frac{3}{2}} \mathrm{e}^{-\frac{1}{2\mathrm{e}}} > 0.2,$$

and

$$\delta_N = 5\pi N^{-\frac{1}{d}} \ge \pi.$$

Hence, (2.5.4) holds trivially in this case.

□

The following lemma shows that $c_2 \le \frac{3}{2}\sqrt{d}$:

Lemma 2.5.5. *For any $\delta > 0$, $x \in \mathbb{S}^d$ and $t \in (0, \pi)$,*

$$\mu_d\left(\left\{y \in \mathbb{S}^d :\ t - \delta \leqslant \rho(x,y) \leqslant t + \delta\right\}\right) \leqslant \frac{3}{2}\sqrt{d}\delta. \tag{2.5.5}$$

Proof. Without loss of generality, we may assume that $0 < t \leq \frac{\pi}{2}$. Setting

$$S_\delta(x) := \left\{y \in \mathbb{S}^d :\ t - \delta \leqslant \rho(x,y) \leqslant t + \delta\right\},$$

and using (2.5.3), we have

$$\begin{aligned}\mu_d(S_\delta(x)) &= \frac{\omega_{d-1}}{\omega_d}\int_{\max\{t-\delta,0\}}^{t+\delta} \sin^{d-1} u \,\mathrm{d}u \leq \pi^{-\frac{1}{2}}\left(\frac{d+1}{2}\right)^{\frac{1}{2}} 2\delta \\ &\leq \frac{2}{\sqrt{\pi}}\sqrt{d} < \frac{3}{2}\sqrt{d}.\end{aligned}$$

□

2.6 Further Examples

Further examples for our results stem from the fact that not only piecewise polynomials are suitable for our spaces V_r of dimension r, but also piecewise exponentials [30], [12] and [13], as well as radial basis functions of compact support [7], [6] and [10], [11].

All these function spaces are defined not over piecewise polynomials (splines) with a simple continuity condition, but for instance over piecewise exponentials.

In the most general form, see [30], the exponential splines of compact support are, say, in d dimensions of degree $n-1$ for *equally spaced knots* defined as distributions B that satisfy

$$B(\varphi) = \int_{[0,1]^n} \varphi(\Xi t)\exp(\lambda \cdot t)\,\mathrm{d}t,$$

where φ is a test-function from the Schwartz space S, Ξ is a linear map $\mathbb{R}^n \to \mathbb{R}^d$ and λ is a vector from $\mathbb{R}^n$ to define the exponentials. Alternatively we can write for $\phi \in L^1_{loc}(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} B(x)\phi(x)\,\mathrm{d}x = \int_{[0,1]^n} \phi(\Xi t)\exp(\lambda \cdot t)\,\mathrm{d}t.$$

In the multivariate setting, these functions are called exponential box-splines, in the univariate case they are exponential B-splines. The piecewise polynomial case corresponds to $\lambda \equiv 0$. They may also be conveniently defined by their Fourier transforms

$$\prod_{j=1}^{n} \frac{\exp(\lambda_j - \mathrm{i}\xi_j \cdot x)}{\lambda_j - \mathrm{i}\xi_j \cdot x}.$$

Here, $\lambda = (\lambda_j)_{j=1}^n$ and $\Xi = (\xi_j)_{j=1}^n$.

The univariate piecewise polynomial case corresponds to $\Xi = (1, 1, \ldots, 1) \in \mathbb{R}^n$, $d = 1$. In this case the splines are defined over the interval or cube for $d = 1$ and $d > 1$, respectively, $\Xi[0, h]^n$, e.g., $h = 1/\ell$ as in our cases. The V_r space is here the space of univariate exponential splines spanned by the exponential B-splines with $d + 1$ knots.

More generally, we can define *space of piecewise exponentials* including piecewise polynomials and exponentials as the span of

$$x^{r_1}\exp(\lambda_i x), \quad r_i = 0, 1, \ldots, \tau_i - 1, \quad i = 1, 2, \ldots n,$$

on each subinterval between two knots, now of no longer necessarily equally spaced knots, of dimension $r = \sum_{i=1}^n \tau_i$ when they are required to be continuous. The λ_is may be complex and must be pairwise distinct.

Special cases [12] are $\lambda \equiv 0$ (piecewise polynomials), $\lambda \in i\mathbb{R}$ (V_r containing piecewise trigonometric functions $\sin$, $\cos$ and constants) and $\lambda \in \mathbb{R}$, V_r containing $\sinh$ and $\cosh$ and constants. In fact, it is usual (but not necessary) to restrict the exponents that form the components of λ to $\mathbb{R} \cup \mathrm{i}\mathbb{R}$. Examples for the spaces are the polynomials for some fixed maximal degree (classical spline case) or the spans of, e.g.,

$$1, \cos(\operatorname{Im}\lambda t), \sin(\operatorname{Im}\lambda t), t\cos(\operatorname{Im}\lambda t), t\sin(\operatorname{Im}\lambda t),$$

or

$$1, \cosh(\operatorname{Re}\lambda t), \sinh(\operatorname{Re}\lambda t), t\cosh(\operatorname{Re}\lambda t), t\sinh(\operatorname{Re}\lambda t).$$

These two examples are the suitable generalisations of the $\Phi(t) = |t|$ case (piecewise linears) referred to in the paragraph after the statement of Theorem 2.5.1. For higher powers, larger r and more exponentials, the other piecewise polynomials used in the first sentence of the statement of Theorem 2.5.1 are generalised.

Univariate piecewise polynomial B-splines on equally spaced knots can be generated in a computational useful, recursive way by convolutions [3] but now, for exponential splines we get a weight function, so that, for the B-spline of degree n, the exponential spline

$$\mathrm{e}^{\lambda_i t} H\Big(t\Big) - \mathrm{e}^{\lambda_i} H\Big(t-1\Big)\mathrm{e}^{\lambda_i(t-1)}$$

needs to be convolved with itself n-times, once for the case of piecewise linears multiplied with exponentials. In the display, H denotes the Heaviside function which is identically zero for negative argument and identically one for positive argument.

This results from the identities which we stated already in s dimensions

$$B \star f = \int_{[0,1]^n} \exp(\lambda \cdot t) f(\cdot - \Xi t)\,\mathrm{d}t$$

or

$$B = \int_0^1 \exp(\lambda_\gamma \underline{t})\tilde{B}(\cdot - \xi_\gamma \underline{t})\,\mathrm{d}\underline{t}.$$

Here B is the exponential box-spline as above, $\tilde{B}$ is the same with the direction ξ_γ removed from Ξ.

As with the piecewise polynomials and the special case of piecewise constants above, we consider the special case of piecewise exponentials only (no polynomials as in our example with $\sin, \cos, \cosh, \sinh$).

For this, consider again the vector of exponents λ, set $n = d$ and let $\tilde{\lambda} = \lambda\Xi^{-1}$. Then the spline is

$$B(x) = \frac{1}{|\det \Xi|} \exp(\tilde{\lambda} \cdot x) \chi_{(0,1]^d}\big(\Xi^{-1} x\big), \qquad x \in \mathbb{R}^d.$$

Here, χ is the characteristic function. Starting from this piecewise "constant" function (i.e., one that contains no polynomials, just one exponential), *other* splines can be generated recursively by

$$B(x) = \mathrm{e}^{\mu \cdot x} \int_0^1 \tilde{B}(x - \underline{t}\xi)\, \mathrm{d}\underline{t},$$

where B is the exponential spline with one direction ξ more in the direction set and the μs are chosen arbitrarily from $\mathbb{R}^n$.

The corresponding *radial basis functions* of compact support with exponentials are

$$\Big(1/\mathrm{e} - \exp(-x)\Big)_+^\nu$$

and

$$\Big(1 - \exp(-(1-x)_+^\nu)\Big)^\mu$$

which are positive definite for suitable parameters μ and ν depending on the dimension because they are logarithmically monotone of order μ in the first case and of order $\min(\mu, \nu)$ in the second case [10]. The V_rs are then defined by the translates

$$(1/\mathrm{e} - \exp(-|x|))_+^\nu$$

and

$$(1 - \exp(-(1-|x|)_+^\nu))^\mu,$$

respectively.

Chapter 3

Polynomial Approximation on High Dimensional Spheres.

3.1 Preliminaries

3.1.1 The de la Vallèe Poussin type operators

Recall that the reproducing kernel of the space of Π_n^d is given by

$$G_n(x\cdot y)=\sum_{k=0}^{n}\frac{k+\lambda}{\lambda}C_k^\lambda(x\cdot y),\ \ x,y\in\mathbb{S}^{d-1}.$$

By the addition formula for spherical harmonics, we have

$$G_n(1)=\sum_{k=1}^{n}\frac{k+\lambda}{\lambda}C_k^\lambda(1)=\int_{\mathbb{S}^{d-1}}|G_n(x\cdot y)|^2\,d\sigma(y)=\dim\Pi_n^d,\ \ x\in\mathbb{S}^{d-1}. \tag{3.1.1}$$

Denote by $\mathcal{P}_n$ the space of all univariate algebraic polynomials of degree at most $n.$ Then for each $x\in\mathbb{S}^{d-1}$,

$$P_n(1)=\int_{\mathbb{S}^{d-1}}P_n(x\cdot y)G_n(y\cdot x)\,d\sigma(y)=c_d\int_{-1}^{1}P_n(t)G_n(t)(1-t^2)^{\lambda-\frac{1}{2}}\,dt, \tag{3.1.2}$$

where

$$c_d = \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\Gamma(\frac{d-1}{2})} = \Big(\int_{-1}^{1}(1-t^2)^{\lambda-\frac{1}{2}}\,dt\Big)^{-1}. \tag{3.1.3}$$

Definition 3.1.1. *For $n \in \mathbb{N}$ and $f \in L^1(\mathbb{S}^{d-1})$, we define the de la Vallèe Poussin type operator V_n by*

$$V_n f(x) = \int_{\mathbb{S}^{d-1}} f(y)K_n(x\cdot y)\,d\sigma(y), \quad x \in \mathbb{S}^{d-1},$$

where

$$K_n(t) = \frac{1}{\dim(\Pi^d_{d_1 n})} G_{(1+d_1)n}(t)G_{d_1 n}(t),$$

and d_1 is the largest integer $\le \frac{d-1}{2}$.

Theorem 3.1.2. *The operators V_n defined above have the following properties:*

(i) $V_n f \in \Pi^d_{(1+2d_1)n} \subset \Pi^d_{dn}$ *for each* $f \in L^1(\mathbb{S}^{d-1})$.

(ii) $V_n P = P$ *for each* $P \in \Pi^d_n$.

(iii) *For each* $1 \le p \le \infty$,

$$\|V_n\|_{(p,p)} := \sup_{\|f\|_p=1} \|V_n f\|_p < 10.$$

Proof. (i) By definition, K_n is an algebraic polynomial of degree at most $(1+2d_1)n$. Thus, $V_n f \in \Pi^d_{(1+2d_1)n}$ for each $f \in L^1(\mathbb{S}^{d-1})$.

(ii) Expanding the function K_n in terms of Gegenbauer polynomials, we have

$$K_n(t) = \frac{1}{\dim \Pi^d_{d_1 n}} \sum_{j=0}^{(1+2d_1)n} a_n(j)\frac{j+\lambda}{\lambda}C_j^\lambda(t),$$

where

$$a_n(j) := \frac{c_d}{C_j^\lambda(1)}\int_{-1}^{1} G_{(1+d_1)n}(t)G_{d_1 n}(t)C_j^\lambda(t)d\mu^*(t),$$

$d\mu^*(t) = c_d(1-t^2)^{\lambda-\frac{1}{2}}\,dt$, and the constant c_d is given in (3.1.3).

If $0 \le j \le n$, then $G_{d_1 n}(t)C_j^\lambda(t) \in \mathcal{P}_{(d_1+1)n}$, and hence using (3.1.2), we have

$$a_n(j) = \frac{1}{C_j^\lambda(1)}G_{d_1 n}(1)C_j^\lambda(1) = G_{d_1 n}(1) = \dim \Pi^d_{d_1 n}, \ \ 0 \le j \le n,$$

which implies (ii).

(iii) It is enough to show that

$$L_n := \sup_{x\in\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |K_n(x\cdot y)|\, d\sigma(y) \le 10. \tag{3.1.4}$$

Indeed, using (3.1.1), we obtain

$$\begin{aligned}
(\dim \Pi_{d_1 n}^d) L_n &= \int_{-1}^{1} |G_{(1+d_1)n}(t)||G_{d_1 n}(t)| d\mu^*(t) \\
&\le \Big(\int_{-1}^{1} |G_{(1+d_1)n}(t)|^2 d\mu^*(t)\Big)^{\frac{1}{2}} \Big(\int_{-1}^{1} |G_{d_1 n}(t)|^2 d\mu^*(t)\Big)^{\frac{1}{2}} \\
&= \sqrt{\dim(\Pi_{(1+d_1)n}^d)\dim(\Pi_{d_1 n}^d)}.
\end{aligned}$$

It follows that

$$L_n \le \sqrt{\frac{\dim \Pi_{(1+d_1)n}^d}{\dim \Pi_{d_1 n}^d}}.$$

Let $N_1 = d_1 n$ and $N_2 = N_1 + n$. Then using (1.0.11), we obtain

$$\begin{aligned}
L_n^2 &= \frac{2N_2+d-1}{2N_1+d-1}\prod_{j=1}^{d-2}\Big(\frac{N_2+j}{N_1+j}\Big) = \frac{2N_2+d-1}{2N_1+d-1}\prod_{j=1}^{d-2}\Big(1+\frac{N_1}{d_1(N_1+j)}\Big) \\
&\le \Big(1+\frac{2N_1}{d_1(2N_1+d-1)}\Big)\Big(1+\frac{N_1}{d_1(N_1+1)}\Big)^{d-2} \le \Big(1+\frac{1}{d_1}\Big)^{d-1}
\end{aligned}$$

If $d \ge 8$, then

$$\Big(1+\frac{1}{d_1}\Big)^{d-1} \le \Big(1+\frac{2}{d-2}\Big)^{d-1} \le \frac{4}{3}e^2 < 10.$$

On the other hand, a straightforward calculation shows that

$$\max_{3\le d\le 7}\Big(1+\frac{1}{d_1}\Big)^{d-1} = 2^3 < 10.$$

□

3.1.2 An average operator

Definition 3.1.3. *For $\theta \in \mathbb{R}$, we define the average operator A_θ by*

$$A_\theta f(x) := \int_{SO(d)} f(Q^{-1} M_\theta Q x)\, dQ, \ \ x \in \mathbb{S}^{d-1},$$

where dQ is the normalized Haar measure on $SO(d)$, and M_θ is a $d \times d$ matrix given by (1.0.14).

For $\alpha, \beta \geq 0$, write

$$R_n^{(\alpha\, \beta)}(x) = \frac{P_n^{(\alpha,\beta)}(x)}{P_n^{(\alpha,\beta)}(1)}.$$

Theorem 3.1.4. (i) *For $1 \leq p \leq \infty$, $\sup_{\|f\|_p = 1} \|A_\theta f\|_p = 1$.*

(ii) *For any $\theta \in \mathbb{R}$ and $f \in L^2(\mathbb{S}^{d-1})$,*

$$\operatorname{proj}_n(A_\theta f) = a_n(\cos\theta) \operatorname{proj}_n f, \ \ n = 0, 1, \cdots, \tag{3.1.5}$$

where

$$a_n(x) = \frac{\lambda + \frac{1}{2}}{(1-x)^{\lambda + \frac{1}{2}}} \int_x^1 R_n^{(\lambda - \frac{1}{2}, \lambda - \frac{1}{2})}(y)(y - x)^{\lambda - \frac{1}{2}} dy, \ \ x \in [-1, 1].$$

Lemma 3.1.5. *For each $f \in L^1(\mathbb{S}^{d-1})$ and $1 \leq m \leq d-2$,*

$$\int_{\mathbb{S}^{d-1}} f(x)\, d\sigma(x) = c_m \int_{B^m} \Big[\int_{\mathbb{S}^{d-m-1}} f(u, \sqrt{1 - \|u\|^2} v)\, d\sigma(v) \Big] (1 - \|u\|^2)^{\frac{d-m-2}{2}}\, du,$$

where $B^m := \{y \in \mathbb{R}^m : \ \ \|y\| \leq 1\}$, and

$$c_m = \Big(\int_{B^m} (1 - \|u\|^2)^{\frac{d-m-2}{2}}\, du \Big)^{-1}.$$

Theorem 3.1.4 was proved in [21]. Here we give a relatively simpler proof.

Proof of Theorem 3.1.4. Clearly, (i) follows directly from the Minkowskii inequality and the rotation invariance of $d\sigma(x)$ on $\mathbb{S}^{d-1}$. We only need to prove (ii).

First, it is easily seen from (1.0.12) that

$$A_\theta T_\rho = T_\rho A_\theta, \ \ \forall \rho \in SO(d).$$

This means that A_θ is a rotation-invariant bounded linear operator on $L^2(\mathbb{S}^{d-1})$.

Thus, A_θ is a multiplier operator with respect to the spherical harmonic expansions on $\mathbb{S}^{d-1}$; that is, there exists a bounded sequence $\{\mu_k(\theta)\}_{k=0}^\infty$ of real numbers such that for any $f \in L^2(\mathbb{S}^{d-1})$,

$$\mathrm{proj}_n(A_\theta f)(x) = \mu_n(\theta)\, \mathrm{proj}_n f(x), \ \ \forall x \in \mathbb{S}^{d-1}, \ \ n = 0, 1, \cdots . \tag{3.1.6}$$

Next, we determine the sequence $\{\mu_n(\theta)\}_{n=0}^\infty$. Let $e \in \mathbb{S}^{d-1}$ be an arbitrarily given point on $\mathbb{S}^{d-1}$, and let $f(x) := C_n^\lambda(x \cdot e)$. Then $f \in \mathcal{H}_n$, and (3.1.6) with $x = e$ implies that

$$\begin{aligned}
\mu_n(\theta) C_n^\lambda(1) &= \mathrm{proj}_n(A_\theta f)(e) = \frac{n+\lambda}{\lambda} \int_{\mathbb{S}^{d-1}} A_\theta f(x) C_n^\lambda(x \cdot e)\, d\sigma(x) \\
&= \frac{n+\lambda}{\lambda} \int_{\mathbb{S}^{d-1}} \Big[\int_{SO(d)} C_n^\lambda\Big(\big(Q^{-1} M_\theta Q x\big) \cdot e\Big) dQ\Big] C_n^\lambda(e \cdot x) d\sigma(x) \\
&= \frac{n+\lambda}{\lambda} \int_{\mathbb{S}^{d-1}} \Big[\int_{SO(d)} C_n^\lambda\Big(\big(M_\theta y\big) \cdot Qe\Big) C_n^\lambda(Qe \cdot y) dQ\Big]\, d\sigma(y) \\
&= \frac{n+\lambda}{\lambda} \int_{\mathbb{S}^{d-1}} \Big[\int_{\mathbb{S}^{d-1}} C_n^\lambda\Big(x \cdot M_\theta y\Big) C_n^\lambda(x \cdot y) d\sigma(x)\Big]\, d\sigma(y) \\
&= \int_{\mathbb{S}^{d-1}} C_n^\lambda(y \cdot M_\theta y)\, d\sigma(y),
\end{aligned}$$

where we used (1.0.12) and the change of variable $y = Qx$ in the fourth step, (1.0.13) in the fifth step, and the fact that $\frac{n+\lambda}{\lambda} C_n^\lambda(x \cdot y)$ is the reproducing kernel of the space $\mathcal{H}_n^d$ in the last step.

Now for each $y \in \mathbb{S}^{d-1}$,

$$y \cdot M_\theta y = (y_1^2 + y_2^2) \cos\theta + \sum_{j=3}^d y_j^2 = (1 - \|u\|^2) \cos\theta + \|u\|^2,$$

where $u = (y_3, y_4, \cdots, y_d)$. Thus, using Lemma 3.1.5 with $m = d - 2$, we obtain

$$\int_{\mathbb{S}^{d-1}} C_n^\lambda(y \cdot M_\theta y)\, d\sigma(y) = c \int_{B^{d-2}} C_n^\lambda\big((1 - \|u\|^2)\cos\theta + \|u\|^2\big)\, du$$
$$= c \int_0^1 C_n^\lambda\big((1-t)\cos\theta + t\big) t^{\lambda - \frac{1}{2}}\, dt = \frac{c}{(1-\cos\theta)^{\lambda+\frac{1}{2}}} \int_{\cos\theta}^1 C_n^\lambda(y)(y - \cos\theta)^{\lambda - \frac{1}{2}}\, dy.$$

It follows that

$$\mu_n(\theta) = \frac{c}{(1-\cos\theta)^{\lambda+\frac{1}{2}}} \int_{\cos\theta}^1 R_n^{(\lambda-\frac{1}{2}, \lambda-\frac{1}{2})}(y)(y - \cos\theta)^{\lambda-\frac{1}{2}} dy.$$

Finally, since $\lim_{\theta\to 0+} A_\theta f = f$, we have

$$1 = \lim_{\theta\to 0+} \mu_n(\theta) = \frac{c}{\lambda + \frac{1}{2}}.$$

Thus, $c = \lambda + \frac{1}{2}$. □

3.2 Proof of Theorem 1.0.11

3.2.1 Proof of inequality 1.0.22 for $r = 2$

In this subsection, we shall prove

Theorem 3.2.1. *Let $f \in L^p(\mathbb{S}^{d-1})$ if $1 \le p < \infty$ or $f \in C(\mathbb{S}^{d-1})$ if $p = \infty$. Then*

$$E_n(f)_p \le 5\omega^2\Big(f, \frac{2d}{n}\Big)_p, \quad n = 1, 2, \cdots.$$

For the proof of Theorem 3.2.1, we need to introduce the Newman-Shapiro operators on the sphere $\mathbb{S}^{d-1}$. For $m = 1, 2, \cdots$, we set

$$B_m(t) := \gamma_m \Big[\frac{C_m^\lambda(t)}{t - \eta_m}\Big]^2, \quad t \in [-1, 1],$$

where $\eta_m := \cos\theta_m$ is the largest zero of $C_m^\lambda(t)$, $0 < \theta_m \le \frac{\pi}{2}$ and $\gamma_m > 0$ is a constant normalized by

$$\int_{\mathbb{S}^{d-1}} B_m(x\cdot y)\, d\sigma(y) = 1, \;\; x \in \mathbb{S}^{d-1}.$$

It was shown in [31] that

$$B_m(t) = \sum_{k=0}^{2m-2} a_{m,k} \frac{k+\lambda}{\lambda} C_k^\lambda(t), \;\; t \in [-1,1], \;\; m = 1, 2, \cdots ,$$

where $a_{m,k} > 0$ for $0 \le k \le 2m-2$, and

$$a_{m,0} = 1, \;\; a_{m,1} = 1 - 2\sin^2 \frac{\theta_m}{2} = \cos\theta_m = \eta_m. \tag{3.2.1}$$

Moreover, the zero η_m satisfies (see [27])

$$1 - \eta_m^2 = \sin^2\theta_m < \frac{(2\lambda+1)(2\lambda+5)}{2m(m+2\lambda)+2\lambda+1}, \tag{3.2.2}$$

which in particular implies that

$$0 < 1 - a_{m,1} = 1 - \eta_m < 1 - \eta_m^2 < \frac{(d-1)(d+3)}{2m^2}. \tag{3.2.3}$$

Definition 3.2.2. *For a positive integer m and $f \in L^1(\mathbb{S}^{d-1})$, we define*

$$S_m f(x) := \int_0^\pi A_\theta f(x) B_m(\cos\theta) d\mu_d(\theta), \;\; x \in \mathbb{S}^{d-1},$$

where

$$d\mu_d(\theta) = c_d \sin^{d-2}\theta\, d\theta \;\; \text{and} \;\; c_d = \Big(\int_0^\pi \sin^{d-2}\theta\, d\theta\Big)^{-1}.$$

Lemma 3.2.3. *Let $1 \le p \le \infty$ and $f \in L^p(\mathbb{S}^{d-1})$. Then for any positive integer m, $S_m f \in \Pi_{2m-2}^d$ and*

$$\|f - S_m f\|_p \le 5\omega^2\Big(f, \frac{d}{m}\Big)_p. \tag{3.2.4}$$

Proof. First, it is easily seen from Lemma 3.1.4 that $S_m f \in \Pi^d_{2m-2}$ and $A_\theta f = A_{-\theta} f$. Thus,

$$\begin{aligned}|f(x) - S_m f(x)| &= \Big|\int_0^\pi (f(x) - A_\theta f(x)) B_m(\cos\theta)\, d\mu_d(\theta)\Big| \\ &= \frac{1}{2}\Big|\int_0^\pi \Big[2f(x) - A_\theta f(x)) - A_{-\theta} f(x)\Big] B_m(\cos\theta)\, d\mu_d(\theta)\Big| \\ &= \frac{1}{2}\Big|\int_0^\pi \Big[\int_{SO(d)} T_{Q^{-1}M_{-\theta}Q}(I - T_{Q^{-1}M_\theta Q})^2 f(x)\, dQ\Big] B_m(\cos\theta)\, d\mu_d(\theta)\Big|.\end{aligned}$$

By the Minkowskii inequality, we then obtain

$$\begin{aligned}\|f - S_m f\|_p &\le \frac{1}{2}\int_0^\pi \Big[\int_{SO(d)} \Big\|(I - T_{Q^{-1}M_\theta Q})^2 f\Big\|_p dQ\Big] B_m(\cos\theta) d\mu_d(\theta) \\ &\le \frac{1}{2}\int_0^\pi B_m(\cos\theta)\omega^2(f,\theta)_p d\mu_d(\theta) =: I_{m,1}\end{aligned}$$

For any $m > d$, Choose $n = [\frac{m}{d}]$, then $\frac{m}{d} - 1 < n \le \frac{m}{d}$ and $1 + \frac{\pi^2}{3}(\frac{dn}{m})^2 \le 5$. and using Lemma 1.0.10 (iii), we have

$$\begin{aligned}I_{m,1} &\le \omega^2(f, n^{-1})_p \int_0^\pi (1 + n^2\theta^2) B_m(\cos\theta) d\mu_d(\theta) \\ &\le \Big[1 + \frac{\pi^2}{2} n^2 \int_0^\pi (1 - \cos\theta) B_m(\cos\theta) d\mu_d(\theta)\Big]\omega^2(f, n^{-1})_p =: I_{m,2}.\end{aligned}$$

Since the function $y \to x \cdot y$ is a spherical harmonic of degree one on $\mathbb{S}^{d-1}$ for each fixed $x \in \mathbb{S}^{d-1}$, it follows from (3.2.3) that

$$\begin{aligned}&\int_0^\pi (1 - \cos\theta) B_m(\cos\theta) d\mu_d(\theta) = 1 - \int_{\mathbb{S}^{d-1}} (x \cdot y) B_m(x \cdot y)\, d\sigma(y) \\ &= 1 - a_{m,1} \le \frac{(d-1)(d+3)}{2m^2} \le \frac{2}{3}\Big(\frac{d}{m}\Big)^2.\end{aligned}$$

Thus,

$$I_{m,2} \le \Big(1 + \frac{\pi^2}{3}\Big(\frac{dn}{m}\Big)^2\Big)\omega^2(f, n^{-1})_p \le 5\omega^2\Big(f, \frac{d}{m}\Big)_p.$$

$\square$

Proof of Theore 3.2.1. Let $m \in \mathbb{N}$ be such that $m \le 2m-2 \le n \le 2m-1 \le 2m$. Then $\frac{d}{m} \le 2\frac{d}{n}$

$$E_n(f)_p \le \|f - S_m f\|_p \le 5\omega^2(f, \frac{d}{m})_p \le 5\omega^2(f, \frac{2d}{n})_p.$$

$\square$

3.2.2 Proof of inequality 1.0.22 for $r \ge 3$

The case for $r = 2$ has already been proven in the last subsection. We need to prove that for $r \ge 3$,

$$E_n(f)_p \le C_r \omega^r\Big(f, \frac{d^3}{n}\Big)_p, \quad \forall f \in L^p(\mathbb{S}^{d-1}), \quad n = 1, 2, \cdots, \tag{3.2.5}$$

Let $n_1 \in \mathbb{N}$ be such that $n_1 d^2 \le n < (n_1+1)d^2$, and set $g := f - V_{n_1 d} f$. Since $V_{n_1 d} f \in \Pi^d_{n_1 d^2} \subset \Pi^d_n$, we have $E_n(f)_p \le \|g\|_p$. Thus, it suffices to prove that

$$\|g\|_p \le C_r \omega^r(f, \frac{d^3}{n})_p. \tag{3.2.6}$$

Since $V_{n_1} f \in \Pi^d_{n_1 d}$, we have $V_{n_1} V_{n_1 d} f = V_{n_1 d} V_{n_1} f = V_{n_1} f$, and

$$V_{n_1} g = V_{n_1} f - V_{n_1 d} V_{n_1} f = 0.$$

It follows that

$$\|g\|_p = \|g - V_{n_1} g\|_p \le c E_{n_1}(g)_p \le C\omega^2\Big(g, \frac{d}{n_1}\Big)_p \le C\omega^2\Big(g, \frac{d^3}{n}\Big)_p. \tag{3.2.7}$$

On the other hand, by inequality 3.3.1

$$\begin{aligned}
\omega^2(g,t)_p &\le c_r t^2 \int_t^{2^m t} \frac{\omega^r(g,u)_p}{u^3}\,du + c_r 2^r t^2 \|g\|_p \int_{2^m t}^{\infty} u^{-3}\,du \\
&\le c_r \omega^r(g, 2^m t)_p + c_r 2^{-2m+r}\|g\|_p \le c_r 2^{(m+1)r}\omega^r(g,t)_p + c_r 2^{-2m+r}\|g\|_p
\end{aligned}$$

Hence:

$$\|g\|_p = \|g - V_{n_1} g\|_p \le 2(10+1)E_{n_1}(g)_p \le 22 * 5 * 2^2 \omega^2\Big(g, \frac{d}{n_1}\Big)_p$$
$$\le 440\omega^2\Big(g, \frac{d^3}{n}\Big)_p \le 440c_r 2^{(m+1)r}\omega^r(g, \frac{d^3}{n})_p + 440c_r 2^{-2m+r}\|g\|_p$$

Selecting m: $\frac{1}{4} \le 440c_r 2^{-2m+r} \le \frac{1}{2}$, then

$$\|g\|_p \le 880c_r 2^r (4 * 440c_r 2^r)^{r/2}\omega^r(g, \frac{d^3}{n})_p$$

Thus, to complete the proof of (3.2.6), it suffices to prove that

$$\omega^r(g,t)_p \le 11\omega^r(f,t)_p, \ \ \forall t > 0. \tag{3.2.8}$$

Indeed, by definition, $\omega^r(g,t)_p \le \omega^r(f,t)_p + \omega^r(V_{n_1 d}f,t)_p$. However,

$$\omega^r(V_{n_1 d}f,t)_p = \sup_{Q\in\mathcal{O}(t)} \|(I-T_Q)^r V_{n_1 d} f\|_p = \sup_{Q\in\mathcal{O}(t)} \|V_{n_1 d}(I-T_Q)^r f\|_p$$
$$\le 10 \sup_{Q\in\mathcal{O}(t)} \|(I-T_Q)^r f\|_p = 10\omega^r(f,t)_p.$$

Denote $D_r = 880c_r 2^r (4 * 440c_r 2^r)^{r/2}$. Then

$$E_n(f)_p \le \|g\|_p \le 11D_r\omega^r(f, \frac{d^3}{n})_p$$

3.2.3 Proof of the matching inverse inequality 1.0.23

.

Proof. We need to prove

$$\omega^r(f, n^{-1})_p \le C_r n^{-r}\sum_{k=1}^{n} k^{r-1}E_k(f)_p.$$

The main idea of this proof comes from [22, Theorem 4.1]. Replacing the Delayed mean for spherical harmonics by the de la Vallée Poussin type operator V_n in the last step, with Bernstein's inequality, we can get the proof of this inverse inequality.

It is enough to prove this result for $n = 2^m$. Assume the result is true for cases of $n = 2^m$, then for $n : 2^{m-1} < n < 2^m$.

$$\begin{aligned}\omega^r(f, n^{-1})_p \leq \omega^r(f, 2^{-(m-1)})_p &\leq C_r 2^{-r(m-1)} \sum_{k=1}^{2^{m-1}} k^{r-1} E_k(f)_p \\ &\leq C_r 2^{-rm} \sum_{k=1}^{n} k^{r-1} E_k(f)_p \\ &\leq C_r n^{-r} \sum_{k=1}^{n} k^{r-1} E_k(f)_p.\end{aligned}$$

For $n = 2^m$, given f and each $j : 1 \leq j \leq n$, we choose $P_j \in \Pi_j^d$ satisfies $\|f - P_j\|_p \leq 2E_j(f)_p$. We have

$$\omega^r(f, 1/n)_p \leq \omega^r(f - P_n, 1/n)_p + \omega^r(P_n, 1/n)_p \leq 2^r \|f - P_n\|_p + \omega^r(P_n, 1/n)_p$$

and

$$\omega^r(P_n, 1/n)_p \leq \sum_{j=0}^{m-1} \omega^r(P_{2^{j+1}} - P_{2^j}, 1/n)_p + \omega^r(P_1, 1/n)_p$$

P_1 is a constant, hence $\omega^r(P_1, 1/n)_p = 0$. In order to prove the result, we only need to show that for any spherical harmonic g_k of degree k,

$$\|\triangle_\rho^r g_k\|_p \leq C t^r k^r \|g_k\|_p \; for \; \rho \in \mathcal{O}(t)$$

with C independent of $t,\ k$ and d. We first prove cases for $r = 1$ with $p = \infty$ and $p = 1$. Then use induction and interpolation theorem to get the result as $\triangle_\rho^{r-1} g_k$ is also a spherical harmonic of degree k.

In the case of $r = 1$ and $p = \infty$. For each fixed point $x \in \mathbb{S}^{d-1}$, we may connect ρx and x by the part of circle created by the intersection of $\mathbb{S}^{d-1}$ with the plane generated by vectors x

and ρx. On that circle, $g_k(y) = T_k(\theta)$ is a trigonometric polynomial with $T_k(0) = g_k(x)$ and $T_k(t_1) = g_k(\rho x)$, $\rho x \cdot x = \cos t_1 \geq \cos t$. Hence, with Bernstein's inequality,

$$\begin{aligned}|g_k(\rho x) - g_k(x)| = |T_k(t_1) - T_k(0)| \leq t_1 |T_k'(c)| \leq t_1 k \|T_k\|_{L_\infty(T)} \\ \leq t_1 k \|g_k\|_\infty \leq tk \|g_k\|_\infty\end{aligned}$$

In the case of $r = 1$ and $p = 1$, we have

$$\begin{aligned}\|\triangle_\rho g_k\|_1 &= \sup_{\|g\|_\infty = 1} |\langle \triangle_\rho g_k, g\rangle| = \sup_{\|g\|_\infty = 1} |\langle V_k(\triangle_\rho g_k), g\rangle| = \sup_{\|g\|_\infty = 1} |\langle \triangle_\rho g_k, V_k(g)\rangle| \\ &= \sup_{\|g\|_\infty = 1} |\langle g_k, \triangle_\rho V_k(g)\rangle| \leq \sup_{\|g\|_\infty = 1} \|g_k\|_1 \|\triangle_\rho V_k(g)\|_\infty \leq \sup_{\|g\|_\infty = 1} \|g_k\|_1 tk \|V_k(g)\|_\infty \\ &\leq \sup_{\|g\|_\infty = 1} 10tk \|g_k\|_1 \|g\|_\infty \leq 10tk\|g_k\|_1\end{aligned}$$

Therefore,

$$\begin{aligned}\omega^r(P_n, 1/n)_p &\leq \sum_{j=0}^{m-1} \omega^r(P_{2^{j+1}} - P_{2^j}, 1/n)_p \leq C \sum_{j=0}^{m-1} n^{-r} 2^{r(j+1)} \|P_{2^{j+1}} - P_{2^j}\|_p \\ &\leq C \sum_{j=0}^{m-1} n^{-r} 2^{r(j+1)} \Big(E_{2^{j+1}}(f)_p + E_{2^j}(f)_p \Big) \leq C_r \sum_{j=0}^{m-1} n^{-r} 2^{(r-1)j} 2^j \Big(E_{2^{j+1}}(f)_p + E_{2^j}(f)_p \Big) \\ &\leq C_r \sum_{j=0}^{m-1} n^{-r} 2^{(r-1)j} \Big(E_{2^{j+1}}(f)_p + E_{2^{j+1}-1}(f)_p + \cdots + E_{2^j}(f)_p \Big) \\ &\leq C_r \sum_{j=0}^{m-1} n^{-r} \Bigg(\sum_{l=2^j}^{2^{j+1}} l^{r-1} E_l(f)_p \Bigg) = C_r n^{-r} \sum_{k=1}^{n} k^{r-1} E_k(f)_p\end{aligned}$$

Hence:

$$\omega^r(f, n^{-1})_p \leq C_r n^{-r} \sum_{k=1}^{n} k^{r-1} E_k(f)_p.$$

with $C_r = 10^r 2^r$

□

3.3 Proof of Marchaud inequality 1.0.18

Proof. For the polynomial $g(x) := [1-2^{-k}(x+1)^k]/(x-1)$ of degree $k-1$, we have $(x-1)^k = 2^{-k}(x^2-1)^k + g(x)(x-1)^{k+1}$. Replacing x by the translation operator T_ρ with $\rho \in \mathcal{O}(t)$, we obtain

$$(T_\rho - I)^k = 2^{-k}(T_{\rho^2} - I)^k + g(T_\rho)(T_\rho - I)^{k+1}$$

For $g(T_\rho)$, we have $M := \|g(T_\rho)\|_p \leq k/2$, therefore:

$$\|\triangle_\rho^k f\|_p \leq 2^{-k}\|\triangle_{\rho^2}^k f\|_p + M\|\triangle_\rho^{k+1} f\|_p \leq 2^{-k}\|\triangle_{\rho^2}^k f\|_p + M\omega^{k+1}(f,t)_p$$

Repeating this process, we obtain

$$\|\triangle_\rho^k f\|_p \leq 2^{-km}\|f\|_p + M\sum_{j=0}^{m} 2^{-kj}\omega^{k+1}(f,2^j t)_p$$

Taking $m \to \infty$, we obtain:

$$\omega^k(f,t)_p \leq Mt^k \sum_{j=0}^{\infty}(2^j t)^{-k}\omega^{k+1}(f,2^j t)_p \leq C_0(k)t^k \int_t^\infty \frac{\omega^{k+1}(f,u)_p}{u^{k+1}}\,du \tag{3.3.1}$$

where $C_0(k) = \frac{k^2}{2}\frac{1}{1-2^{-k}}$. Finally, we obtain the marchaud inequality via induction on r. Assume the marchaud inequality is true for r, from (3.3.1), we obtain:

$$\begin{aligned}
\omega^k(f,t)_p &\leq C_0(k)t^k \int_t^\infty \frac{\omega^r(f,u)_p}{u^{k+1}}\,du \\
&\leq C_0(k)t^k \int_t^\infty u^{-k-1}u^r \int_u^\infty \frac{\omega^{r+1}(f,z)_p}{z^{r+1}}\,dz\,du \\
&= C_0(k)t^k \int_t^\infty \frac{\omega^{r+1}(f,z)_p}{z^{r+1}} \int_t^z u^{r-k-1}\,du\,dz \\
&\leq C_0(k)t^k \int_t^\infty \frac{\omega^{r+1}(f,u)_p}{u^{k+1}}\,du
\end{aligned}$$

As there exists $Q \in SO(d)$ and Skew-Symmetric matrix M_θ such that $\rho = Q^{-1}M_\theta Q$ and $\rho^j = Q^{-1}M_{j\theta}Q \in \mathcal{O}(jt)$.

By induction, we have

$$\omega^r(f,t)_p \leq C(m)t^r \int_t^\infty \frac{\omega^m(f,u)_p}{u^{r+1}}\,du \tag{3.3.2}$$

with $C(m) = C_0(r)C_0(r+1)\cdots C_0(m-1)$.

$$\begin{aligned}
\omega^r(f,t)_p &\leq C(m)t^r \int_t^\infty \frac{\omega^m(f,u)_p}{u^{r+1}}\,du \\
&= C(m)t^r \int_t^{2^l t} \frac{\omega^m(f,u)_p}{u^{r+1}}\,du + C(m)t^r \int_{2^l t}^\infty \frac{\omega^m(f,u)_p}{u^{r+1}}\,du \\
&\leq C(m)t^r \int_t^{2^l t} \frac{\omega^m(f,u)_p}{u^{r+1}}\,du + \frac{C(m)}{r} 2^{m-lr} \|f\|_p
\end{aligned}$$

There exists c_0 such that $\|f - c_0\|_p \leq 2\inf_{c\in\mathbb{R}} \|f-c\|_p = E_0(f)_p$. By Jackson's inequality, we have (C_r is a constant from Jackson's inequality.)

$$\begin{aligned}
\|f - c_0\|_p &\leq 2\inf_{c\in\mathbb{R}} \|f-c\|_p = 2E_0(f)_p \\
&\leq C_r\omega^r(f, d^3)_p = C_r\omega^r(f,\pi)_p \\
&\leq C_r\omega^r(f - c_0, \frac{1}{2})_p \\
&\leq C_rC(m)t^r \int_t^{2^l t} \frac{\omega^m(f,u)_p}{u^{r+1}}du + \frac{C_rC(m)}{r} 2^{m-lr} \|f - c_0\|_p
\end{aligned}$$

Choose l large enough such that

$$\frac{1}{4} \leq \frac{C_rC(m)}{r} 2^{m-lr} \leq \frac{1}{2}$$

Then

$$\|f - c_0\|_p \leq 2C_rC(m)t^r \int_t^{2^l t} \frac{\omega^m(f,u)_p}{u^{r+1}}du$$

$$\omega^r(f,t)_p = \omega^r(f-c_0,t)_p \le C(m,r)t^r \int_t^{2^l t} \frac{\omega^m(f,u)_p}{u^{r+1}}du$$

Denote $M_0 = 2^l t > 1$, (if $M_0 \le 1$, proof ended) then for any $s, \frac{1}{2} \le s < 1$

$$\begin{aligned} t^r \int_1^{M_0} \frac{\omega^m(f,u)_p}{u^{r+1}}du &\le Ct^r\omega^m(f,\pi)_p \le C\left(\frac{\pi}{s}\right)^m t^r\omega^m(f,s)_p \\ &\le C(2\pi)^m t^r \omega^m(f,s)_p \le Ct^r \int_{\frac{1}{2}}^1 \frac{\omega^m(f,u)_p}{u^{r+1}}du \end{aligned}$$

Thus, we obtain:

$$\omega^r(f,t)_p \le C_1(m)t^r \int_t^1 \frac{\omega^m(f,u)_p}{u^{r+1}}\,du,$$

□

3.4 Equivalence with K-functionals

Recall that the $r-$th order partial derivative in Euler angles is defined by

$$\mathcal{D}_Q^r f(x) := \left(\frac{\partial}{\partial t}\right)^r \left(f(Q^{-1}M_tQx)\right)\Big|_{t=0}, \quad x \in \mathbb{S}^{d-1}.$$

for $r \in \mathbb{N}$, $f \in C^r(\mathbb{S}^{d-1})$ and $Q \in SO(d)$. r-th order K-functional of $f \in L^p(\mathbb{S}^{d-1})$ is

$$K_r(f,t)_p := \inf\left\{\|f-g\|_p + t^r \sup_{Q\in SO(d)} \|\mathcal{D}_Q^r g\|_p : \; g \in C^r(\mathbb{S}^{d-1})\right\}, \;\; t > 0.$$

for $r \in \mathbb{N}$ and $1 \le p \le \infty$.

Lemma 3.4.1. *Let $r \in \mathbb{N}$ and $Q \in SO(d)$. Then the following statements hold:*

(i) *For each $\theta \in \mathbb{R}$,*

$$\mathcal{D}_Q^r T_{Q^{-1}M_\theta Q} = T_{Q^{-1}M_\theta Q}\mathcal{D}_Q^r.$$

(ii) *For* $x = (x_1, x_2, \cdots, x_d) \in \mathbb{S}^{d-1}$ *and* $f \in C^1(\mathbb{S}^{d-1})$,

$$(T_Q^{-1}\mathcal{D}_Q T_Q)f(x) = D_{12}f(x) := x_2\partial_1 f(x) - x_1\partial_2 f(x). \tag{3.4.1}$$

(iii) *If* $f \in C^{r+1}(\mathbb{S}^{d-1})$, *then*

$$\mathcal{D}_Q^{r+1}f(x) = \mathcal{D}_Q\big(\mathcal{D}_Q^r f\big)(x), \ \ x \in \mathbb{S}^{d-1}.$$

(iv) *The operator* $\mathcal{D}_Q^r$ *is invariant on the spaces of spherical harmonics; that is,* $\mathcal{D}_Q^r\mathcal{H}_n \subset \mathcal{H}_n$ *for* $n = 0, 1, \cdots$. *In particular, this implies* $\mathcal{D}_Q^r V_n = V_n \mathcal{D}_Q^r$.

Proof. (i) For $\theta \in \mathbb{R}$, set $\rho(\theta) := Q^{-1}M_\theta Q$. Clearly, $\rho(\theta_1 + \theta_2) = \rho(\theta_1)\rho(\theta_2)$ for any $\theta_1, \theta_2 \in \mathbb{R}$. For $f \in C^r(\mathbb{S}^{d-1})$, we define

$$F_f(\theta, x) := f\big(\rho(\theta)x\big), \ \ \theta \in \mathbb{R}, \ \ x \in \mathbb{S}^{d-1}.$$

Then

$$\mathcal{D}_Q^r f(x) = \partial_1^r F_f(0, x).$$

Moreover, for any $\theta_1, \theta_2 \in \mathbb{R}$ and $x \in \mathbb{S}^{d-1}$,

$$\begin{aligned} F_f(\theta_1 + \theta_2, x) &= f\big(\rho(\theta_1 + \theta_2)x\big) = f\Big(\rho(\theta_1)\rho(\theta_2)x\Big) = F_f(\theta_1, \rho(\theta_2)x) \\ &= \big(T_{\rho(\theta_2)}f\big)\Big(\rho(\theta_1)x\Big) = F_{T_{\rho(\theta_2)}f}(\theta_1, x). \end{aligned}$$

Differentiating with respect to θ_1 then gives

$$\partial_1^r F_f(\theta_1 + \theta_2, x) = \partial_1^r F_f(\theta_1, \rho(\theta_2)x) = \partial_1^r F_{T_{\rho(\theta_2)}f}(\theta_1, x).$$

Setting $\theta_1 = 0$ and $\theta_2 = \theta$, we deduce

$$\partial_1^r F_f(0, \rho(\theta)x) = \partial_1^r F_{T_{\rho(\theta)}f}(0, x),$$

which implies

$$T_{\rho(\theta)}\mathcal{D}_Q^r f(x) = (\mathcal{D}_Q^r T_{\rho(\theta)} f)(x).$$

This shows (i).

(ii) By definition,

$$\begin{aligned} T_{Q^{-1}}\mathcal{D}_Q T_Q f(x) = \mathcal{D}_Q T_Q f(Q^{-1}x) &= \frac{\partial}{\partial t}\Big[(T_Q f)(Q^{-1}M_t x)\Big]\Big|_{t=0} \\ &= \frac{\partial}{\partial t}\Big(f(M_t x)\Big)\Big|_{t=0} = D_{12} f(x). \end{aligned}$$

(iii) It is known that $D_{12}^{r+1} = D_{12}^r D_{12}$. Thus,

$$\mathcal{D}_Q^{r+1} = T_Q D_{12}^{r+1} T_Q^{-1} = T_Q D_{12} T_Q^{-1} T_Q D_{12}^r T_Q^{-1} = \mathcal{D}_Q \mathcal{D}_Q^r.$$

(iv) It is known that D_{12} is invariant on each space $\mathcal{H}_n$ of spherical harmonics. Since each space $\mathcal{H}_n$ is also rotational invariant, it follows that

$$\mathcal{D}_Q \mathcal{H}_n = T_Q D_{12} T_Q^{-1} \mathcal{H}_n \subset \mathcal{H}_n.$$

Since every multiplier operator restricted on each space $\mathcal{H}_n$ is a constant multiple of the identity operator, it follows by linearity that $\mathcal{D}_Q$ commutes with every multiplier operator. $\square$

Lemma 3.4.2 (Bernstein)**.** *For any* $r \in \mathbb{N}$, $1 \le p \le \infty$ *and* $f \in \Pi_n^d$,

$$\sup_{Q \in SO(d)} \|\mathcal{D}_Q^r f\|_p \le n^r \|f\|_p. \tag{3.4.2}$$

Proof. It is known that for each $1 \le p \le \infty$,

$$\|D_{12}^r f\|_p \le n^r \|f\|_p, \;\; \forall f \in \Pi_n^d.$$

Thus, using (3.4.1), we obtain that for each $1 \le p \le \infty$, $f \in \Pi_n^d$ and $Q \in SO(d)$,

$$\|\mathcal{D}_Q^r f\|_p = \|T_Q D_{12}^r T_{Q^{-1}} f\|_p = \|D_{12}^r T_{Q^{-1}} f\|_p \le n^r \|T_{Q^{-1}} f\|_p = n^r \|f\|_p.$$

$\square$

Theorem 3.4.3. *If $r \in \mathbb{N}$ and $1 \le p \le \infty$, then there exist a constant $C_r > 0$ depending only on r such that for all $f \in L^p(\mathbb{S}^{d-1})$ and $t \in (0,1)$,*

$$C_r^{-1} K_r(f, d^{-3}t)_p \le \omega^r(f,t)_p \le C_r K_r(f,t)_p. \tag{3.4.3}$$

For the proof of Theorem 3.4.3, we need a few lemmas.

Lemma 3.4.4. *Let $1 \le p \le \infty$, $r \in \mathbb{N}$ and $M \in \mathcal{M}$. Then for any $f \in \Pi_n^d$,*

$$n^{-r} \|\mathcal{D}_Q^r f\|_p \le C_r \omega^r(f, n^{-1})_p. \tag{3.4.4}$$

Proof. For the moment, we assume that $p < \infty$. Since $\mathcal{D}_Q^r = T_Q D_{12}^r T_{Q^{-1}}$, setting $g = T_{Q^{-1}} f$, we have

$$\begin{aligned}\|\mathcal{D}_Q^r f\|_p^p = \|D_{12}^r g\|_p^p &= c_d \int_{B^{d-2}} \int_0^{2\pi} |D_{12}^r g(\sqrt{1-\|u\|^2}(\cos\theta, \sin\theta), u)|^p d\theta du \\ &= c_d \int_{B^{d-2}} \int_0^{2\pi} \left| T_u^{(r)}(\theta) \right|^p d\theta du,\end{aligned}$$

where

$$T_u(\theta) := g\Big(\sqrt{1-\|u\|^2}(\cos\theta, \sin\theta), u\Big), \quad \theta \in [0, 2\pi], \quad u \in B^{d-2}.$$

Since $g \in \Pi_n^d$, it is easily seen that for each fixed $u \in B^{d-2}$, T_u is a trigonometric polynomial of degree at most n. Thus, by the Stechkin inequality, we obtain that for each $u \in B^{d-2}$,

$$n^{-r} \Big(\int_0^{2\pi} \left| T_u^{(r)}(\theta) \right|^p d\theta \Big)^{\frac{1}{p}} \le C_r \Big(\int_0^{2\pi} |\triangle_{n^{-1}}^r T_u(\theta)|^p \, d\theta \Big)^{\frac{1}{p}},$$

where

$$\triangle_h^r T_u(\theta) = \sum_{j=0}^{r} \binom{r}{j} (-1)^j T_u(\theta + jh), \ \ \theta, h \in \mathbb{R}.$$

However, by the dfinition, it is easily seen that

$$\triangle_h^r T_u(\theta) = \triangle_{M_h}^r g\Big(\sqrt{1 - \|u\|^2}(\cos\theta, \sin\theta), u\Big).$$

Thus, we get

$$\begin{aligned} n^{-rp}\|\mathcal{D}_Q^r f\|_p^p &\le C_r^p c_d \int_{B^{d-2}} \int_0^{2\pi} \Big|\triangle_{M_{n^{-1}}}^r g(\sqrt{1 - \|u\|^2}(\cos\theta, \sin\theta), u)\Big|^p d\theta du \\ &= C_r^p \int_{\mathbb{S}^{d-1}} \Big|\triangle_{M_{n^{-1}}}^r g(x)\Big|^p dx = C_r^p \int_{\mathbb{S}^{d-1}} \Big|T_Q \triangle_{M_{n^{-1}}}^r T_Q^{-1} f(x)\Big|^p dx \\ &= C_r^p \int_{\mathbb{S}^{d-1}} \Big|\triangle_{Q^{-1}M_{n^{-1}}Q}^r f(x)\Big|^p dx \le C_r^p \omega^r(f, n^{-1})_p^p. \end{aligned}$$

A slight modification of the above proof works equally well for $p = \infty$.

$\square$

Proof of Theorem 3.4.3. First, we prove that for any $g \in C^r(\mathbb{S}^{d-1})$ and $t > 0$,

$$\omega^r(g, t)_p \le t^r \sup_{Q \in SO(d)} \|\mathcal{D}_Q^r g\|_p. \tag{3.4.5}$$

Let $\rho = Q^{-1} M_\theta Q$ with $Q \in SO(d)$ and $\theta \in \mathbb{R}$. For each fixed $x \in \mathbb{S}^{d-1}$, define

$$T_x(t) = g(Q^{-1} M_t Q x), \ \ t \in \mathbb{R}.$$

Then

$$T_x^{(r)}(t) = \Big(\frac{\partial}{\partial t}\Big)^r g(Q^{-1} M_t Q x), \ \ t \in \mathbb{R}.$$

Thus, for any $x \in \mathbb{S}^{d-1}$,

$$\triangle_\rho^r g(x) = \int_{[0,\theta]^r} T_x^{(r)}(u_1 + u_2 + \cdots + u_r) du_1 \cdots du_r$$
$$= \int_{[0,\theta]^r} \Big(\frac{\partial}{\partial t}\Big)^r g\Big(Q^{-1} M_{u_1+\cdots+u_r} Q x\Big)\, du_1 \cdots du_r.$$

It follows that

$$\|\triangle_\rho^r g\|_p \le \int_{[0,\theta]^r} \Big\| \Big(\frac{\partial}{\partial t}\Big)^r g\Big(Q^{-1} M_{u_1+\cdots+u_r} Q \cdot\Big)\Big\|_p du_1 \cdots du_r$$
$$\le \theta^r \|\mathcal{D}_Q^r g\|_p.$$

This proves (3.4.5).

Next, we show that

$$\omega^r(f,t)_p \le C_r K_r(f,t)_p. \tag{3.4.6}$$

Let $g \in C^r(\mathbb{S}^{d-1})$ be such that

$$\|f-g\|_p + t^r \sup_{Q \in SO(d)} \|\mathcal{D}_Q^r g\|_p \le 2K_r(f,t)_p.$$

Indeed, by (3.4.5), we have

$$\omega^r(f,t)_p \le \omega^r(f-g,t)_p + \omega^r(g,t)_p \le C_r\|f-g\|_p + C\omega^r(g,t)_p$$
$$\le C_r\|f-g\|_p + Ct^r \sup_{Q \in SO(d)} \|\mathcal{D}_Q^r g\|_p \le C_r K_r(f,t)_p.$$

Finally, we show that

$$K_r(f, td^{-3})_p \le C_r \omega^r(f,t)_p \tag{3.4.7}$$

Let n be the largest integer s.t. $n-1 \le d^3 t^{-1}$. Let $P_n \in \Pi_n^d$ be such that $E_n(f)_p = \|f-P_n\|_p$. Using Lemma 3.4.4, we have

$$
\begin{aligned}
K_r(f, d^{-3}t)_p &\leq \|f - P_n\|_p + n^{-r} \sup_{Q \in SO(d)} \|\mathcal{D}_Q^r P_n\|_p \\
&\leq E_n(f)_p + C_r \omega^r(P_n, n^{-1})_p \leq C_r E_n(f)_p + C_r \omega^r(f, n^{-1})_p \\
&\leq C_r \omega^r(f, d^3 n^{-1})_p \leq C_r \omega^r(f, t)_p.
\end{aligned}
$$

$\square$

References

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, 1964.

[2] A. Bondarenko, D. Radchenko, and M. Viazovska, *Well-separated spherical designs.*, Constr. Approx. **41** (2015), no. 1, 93-112.

[3] Carl de Boor, *A practical guide to splines*, Springer, 1978.

[4] J. Bourgain and J. Lindenstrauss, *Distribution of points on spheres and approximation by zonotopes*, Israel J. Math. **64** (1988), no. 1, 25–31.

[5] G. Brown and F. Dai, *Approximation of smooth functions on compact two-point homogeneous spaces*, J. Funct. Analysis **220** (2005), no. 2, 401–423.

[6] Martin Buhmann, *Radial functions on compact support*, Proceedings of the Edinburgh Mathematical Society **41** (1998), 33–46.

[7] ______, *A new class of radial functions with compact support*, Mathematics of Computation **70** (2001), 307–318.

[8] ______, *Radial basis functions: theory and implementations*, Cambridge University Press, 2003.

[9] Martin D. Buhmann and Feng Dai and Yeli Niu, *Discretization of integrals on compact metric measure spaces*, Advances in Mathematics **381** (2021), 107602.

[10] Martin Buhmann and Janin Jäger, *Multiply and monotone functions for radial basis function interpolation: extensions and new kernels, preprint, JLU Giessen*.

[11] ______, *Pólya type criteria for conditional strict positive definiteness of functions on spheres, preprint, JLU Giessen*.

[12] Constanza Conti, Mariantonia Cotronei, and Lucia Romani, *Beyond B-splines: exponential pseudo-splines and subdivision schemes reproducing exponential polynomials*, Dolomites Research Notes on Approximation **10** (2017), 31–42.

[13] Constanza Conti, L. Gemignani, and Lucia Romani, *Exponential pseudo-splines: looking beyond exponential B-splines*, Journal of Mathematical Analysis and Applications **439** (2016), 32–56.

[14] Feng Dai, *Jackson-type inequality for doubling weights on the sphere*, Constr. Approx. **24** (2006), no. 1, 91–112.

[15] F. Dai and Z. Ditzian, *Jackson inequality for Banach spaces on the sphere*, Acta Math. Hungar. **118** (2008), no. 1-2, 171–195.

[16] F. Dai, Z. Ditzian, and Hongwei Huang, *Equivalence of measures of smoothness in $L_p(S^{d-1})$, $1 < p < \infty$*, Studia Math. **196** (2010), no. 2, 179–205.

[17] F. Dai, A. Primak, V. N. Temlyakov, and S. Yu. Tikhonov, *Integral norm discretization and related problems*, Uspekhi Mat. Nauk **74** (2019), no. 4(448), 3–58 (Russian, with Russian summary); English transl., Russian Math. Surveys **74** (2019), no. 4, 579–630.

[18] Feng Dai and Heping Wang, *Optimal cubature formulae in weigted Besov spaces with A_∞-weights on multivariate domains*, Constructive Approximation **37** (2013), 167–194.

[19] Feng Dai and Yuan Xu, *Approximation theory and harmonic analysis on spheres and balls*, Springer Monographs in Mathematics, Springer, New York, 2013.

[20] Feng Dai and Xu Yuan, *Moduli of smoothness and approximation on the unit sphere and the unit ball*, Advances in Mathematics **224** (2010), no. 4, 1233 - 1310.

[21] Z. Ditzian, *Jackson-type inequality on the sphere*, Acta Math. Hungar. **102** (2004), no. 1-2, 1–35.

[22] ______, *A modulus of smoothness on the unit sphere*, J. Anal. Math. **79** (1999), 189–200.

[23] Herbert Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969. MR0257325

[24] Zoltán Füredi and Peter A. Loeb, *On the best constant for the Besicovitch covering theorem*, Proc. AmerṀath. Soc. **121** (1994), no. 4, 1063–1073, DOI 10.2307/2161215. MR1249875

[25] Walter Gautschi, *Orthogonal Polynomials*, Numerical Mathematics and Scientific Computation, Oxford University Press, 2004.

[26] Steven G. Krantz, *The Besicovitch covering lemma and maximal functions*, Rocky Mountain J. Math. **49** (2019), no. 2, 539–555, DOI 10.1216/RMJ-2019-49-2-539. MR3973239

[27] Geno Nikolov, *New bounds for the extreme zeros of Jacobi polynomials*, Proc. Amer. Math. Soc. **147** (2019), no. 4, 1541–1550.

[28] D. J. Newman and H. S. Shapiro, *Jackson's theorem in higher dimensions*, On Approximation Theory (Proceedings of Conference in Oberwolfach, 1963), Birkhäuser, Basel., 1964, pp. 208–219.

[29] G. Petrova, *Cubature formulae for spheres, simplices and balls Journal of Computational and Applied Mathematics*, Journal of Computational and Applied Mathematics **162** (2004), 483–496.

[30] Amos Ron, *Exponential box splines*, Constructive Approximation **4** (1988), 357–378.

[31] Manfred Reimer, *Generalized hyperinterpolation on the sphere and the Newman-Shapiro operators*, Constr. Approx. **18** (2002), no. 2, 183–203.

[32] Walter Rudin, *Functional analysis*, 2nd ed., International Series in Pure and Applied Mathematics, McGraw-Hill, Inc., New York, 1991.

[33] John M. Sullivan, *Sphere packings give an explicit bound for the Besicovitch covering theorem*, J. Geom. Anal. **4** (1994), no. 2, 219–231, DOI 10.1007/BF02921548. MR1277507

[34] Yuan Xu, *Orthogonal polynomials and cubature formulae on spheres and on simplices*, SIAM Journal of Mathematical Analysis **29** (1998), no. 3, 779–793.