View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

ACADEMIC PRESS
Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE @ DIRECT®
Journal of Approximation Theory 121 (2003) 199–219
JOURNAL OF Approximation Theory
http://www.elsevier.com/locate/jat

On approximation of functions from Sobolev spaces on metric graphs ☆

Michael Solomyak

Department of Mathematics, Weizmann Institute 76100, Rehovot, Israel

Received 19 June 2002; accepted 18 October 2002

Dedicated to Y. Brudny, outstanding mathematician and an old friend

Communicated by Paul Nevai

Abstract

Some results on the approximation of functions from the Sobolev spaces on metric graphs by step functions are obtained. In particular, we show that the approximation numbers a_n of the embedding operator of the Sobolev space $\mathsf{L}^{1,p}(\mathbf{G})$ on a graph $\mathbf{G}$ of finite length $|\mathbf{G}|$ into the space $\mathsf{L}^p(\mathbf{G},\mu)$, where μ is an arbitrary finite Borel measure on $\mathbf{G}$, satisfy the inequality

$$a_n \leqslant |\mathbf{G}|^{1/p'} \mu(\mathbf{G})^{1/p} n^{-1}, \quad 1<p<\infty.$$

The estimate is sharp for any $n\in\mathbb{N}$.

© 2003 Elsevier Science (USA). All rights reserved.

1. Introduction

A metric graph is a graph whose edges are viewed as non-degenerate line segments, rather than pairs of vertices as in the case of the standard (combinatorial) graphs. This difference is reflected in the nature of functions on the corresponding graph. For a combinatorial graph this is just a family of numbers $\{f(v)\}$ where the argument v runs over the set of all vertices, while a function on a metric graph is a family of functions on its edges, usually subject to some matching conditions at the vertices.

☆Partially supported by the NATO grant PST.CLG.978694.

E-mail address: solom@wisdom.weizmann.ac.il.

0021-9045/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0021-9045(03)00033-9

Sobolev spaces $\mathsf{L}^{1,p}$ on a metric graph $\mathbf{G}$ are defined in a natural way, by analogy with their counterparts for a single interval. The local properties of functions from these spaces outside the vertices are evidently the same as for the case of an interval. However, the global properties may depend on the geometry of a given graph. We establish some results on approximation of functions from $\mathsf{L}^{1,p}$ by step functions. The estimates obtained are uniform with respect to all graphs of a fixed length and do not depend on the structure of the graph. The estimates are sharp with respect to all the parameters involved. We believe that such results are useful for better understanding of function spaces on graphs.

An important phase in the development of analysis on metric graphs was started by Evans and Harris [3]. Embeddings of the Sobolev spaces $\mathsf{W}^{1,p}(\Omega)$ in $L^p(\Omega)$ were studied there for a wide class of domains with irregular boundary. A characteristic feature of these domains is that they have a "ridge", this being a metric tree. In [3] the study of such embeddings was reduced to the investigation of the behavior of the approximation numbers for the weighted Hardy-type integral operators on the ridge. For $p = 2$ approximation numbers coincide with the singular numbers, and the problem can be reformulated in terms of the eigenvalue behavior for the "weighted Laplacian" on the tree. From this point of view the question was analyzed in [5]. Eigenvalue estimates for the weighted Laplacian were obtained there in terms of appropriate partitions of the given tree into a family of segments. Some of the results of [5] were considerably refined by Evans, Harris, and Lang in their recent paper [4]. The main novelty of [4] consisted in replacement of segments, as elements of a partition, with arbitrary compact subtrees. A thorough analysis of the partitions appearing in the process of approximation allowed the authors to obtain important results for arbitrary p, $1 \leqslant p \leqslant \infty$. In particular, for $p \in (1, \infty)$ they established a Weyl-type asymptotic formula for the approximation numbers.

Our goal in this paper is to consider arbitrary graphs, rather than only the trees. The language of Hardy-type integral operators is no more relevant, since such operators are well defined only on trees. Instead, we study embeddings of Sobolev spaces on the graph $\mathbf{G}$ into the space $\mathsf{L}^\infty(\mathbf{G})$ and into the spaces $\mathsf{L}^p(\mathbf{G}, \mu)$ where μ is an arbitrary Borel measure on $\mathbf{G}$. The character of the results obtained makes it apparent that this language is adequate. Following the idea of [4], we use partitions of a given graph into subgraphs, however the way of this usage differs from the one in [4]. We restrict ourselves to the case of compact graphs, since the passage to non-compact ones can be carried out exactly as in [4] and does not require new ideas, as soon as one is interested only in the estimates but not in asymptotics.

Introduce some necessary notations. Let $\mathbf{G}$ be a connected graph with the set of vertices $\mathscr{V} = \mathscr{V}(\mathbf{G})$ and the set of edges $\mathscr{E} = \mathscr{E}(\mathbf{G})$. Compactness of a graph means that $\#\mathscr{E} < \infty$ and hence, also $\#\mathscr{V} < \infty$. The distance $\rho(x, y) = \rho_{\mathbf{G}}(x, y)$ between any two points $x, y \in \mathbf{G}$ (and thus, the metric topology on $\mathbf{G}$), and also the measure dx on $\mathbf{G}$ are introduced in a natural way; see Section 2 for detail. Below $|E| = |E|_{\mathbf{G}}$ stands for the measure of a measurable set $E \subset \mathbf{G}$. If in particular $E = e$ is an edge, then $|e|$ is its length.

Below the symbol $\mathfrak{M}(\mathbf{G})$ stands for the set of all finite Borel measures on $\mathbf{G}$. For $1 \leqslant p \leqslant \infty$, we denote by $||\cdot||_{p,\mu}$ the norm in the space $\mathsf{L}^p(\mathbf{G}, \mu)$, i.e.

$$||u||_{p,\mu} = ||u||_{\mathsf{L}^p(\mathbf{G},\mu)} = \left(\int_{\mathbf{G}} |u|^p d\mu\right)^{1/p}, \quad p < \infty,$$

with the standard change if $p = \infty$. If the measure μ is absolutely continuous, i.e. $d\mu = V\,dx$, then we write V instead of μ in the above notations. We drop the index μ (or V) if $d\mu = dx$.

A function u on $\mathbf{G}$ belongs to the Sobolev space $\mathsf{L}^{1,p} = \mathsf{L}^{1,p}(\mathbf{G})$, if u is continuous on $\mathbf{G}$ and its restriction to each edge e has the distributional derivative u' which is a function from $\mathsf{L}^p(e)$. The functional $||u'||_{\mathsf{L}^p(\mathbf{G})}$ defines on $\mathsf{L}^{1,p}$ a semi-norm vanishing on the one-dimensional subspace of constant functions.

We say that v is a *step function* on $\mathbf{G}$ and write $v \in \mathrm{Step}(\mathbf{G})$, if v takes only a finite number of different values, each one on a connected subset of $\mathbf{G}$. Any function $v \in \mathrm{Step}(\mathbf{G})$ can be represented as a linear combination of characteristic functions of mutually disjoint connected subsets. We write $v \in \mathrm{Step}_n(\mathbf{G})$, if for v there exists a representation with the number of terms less or equal to n.

We are interested in the approximation of functions $u \in \mathsf{L}^{1,p}(\mathbf{G})$ by functions $v \in \mathrm{Step}_n(\mathbf{G})$. More exactly, we study two problems: the uniform approximation (i.e. approximation in the metric $||\cdot||_\infty$) and approximation in the metric $||\cdot||_{p,\mu}$. In the first problem we construct a mapping $Z_p : \mathsf{L}^{1,p}(\mathbf{G}) \to \mathrm{Step}_n(\mathbf{G})$ such that $||u - Z_p u||_\infty \leqslant C_p(\mathbf{G})(n+1)^{-1}||u'||_p$. This problem is elementary for $p = \infty$, when the operator Z_∞ can be chosen linear and $C_\infty(\mathbf{G}) = |\mathbf{G}|$. For $p < \infty$ a linear mapping Z_p with the required properties does not exist but we find a non-linear mapping which gives the same rate of approximation, with $C_p(\mathbf{G}) = |\mathbf{G}|^{1/p'}$. In the second problem we establish a similar result by means of a linear approximation operator; this operator depends on the measure μ.

Below we present formulations of the typical results.

Theorem 1.1. *Let* $\mathbf{G}$ *be a compact graph and* $1 \leqslant p \leqslant \infty$. *Then for any function* $u \in \mathsf{L}^{1,p}(\mathbf{G})$ *and any* $n \in \mathbb{N}$ *there exists a function* $v \in \mathrm{Step}_n(\mathbf{G})$ *such that*

$$||u - v||_\infty \leqslant \frac{|\mathbf{G}|^{1/p'}||u'||_p}{n+1}. \tag{1.1}$$

If $p = \infty$, *the mapping* $u \mapsto v$ *can be chosen linear.*

Theorem 1.2. *Let* $\mathbf{G}$ *be a compact graph and* $\mu \in \mathfrak{M}(\mathbf{G})$.

(i) *Let* $1 \leqslant p < \infty$, *then for any* $n \in \mathbb{N}$ *there exists a linear operator* $P_n : \mathsf{L}^{1,p}(\mathbf{G}) \to \mathrm{Step}(\mathbf{G})$ *such that* $\mathrm{rank}(P_n) \leqslant n$ *and*

$$||u - P_n u||_{p,\mu} \leqslant \frac{|\mathbf{G}|^{1/p'}\mu(\mathbf{G})^{1/p}}{n+1}||u'||_p, \quad \forall u \in \mathsf{L}^{1,p}(\mathbf{G}). \tag{1.2}$$

(ii) *Let* $p=\infty$ *and* $d\mu = V\,dx$ *where* $V\in \mathsf{L}^{\infty}(\mathbf{G})$. *Then for any* $n\in\mathbb{N}$ *there exists a linear operator* $P_n : \mathsf{L}^{1,\infty}(\mathbf{G})\to \mathrm{Step}(\mathbf{G})$ *such that* $\operatorname{rank}(P_n)\leqslant n$ *and*

$$||u-P_nu||_{\infty,V}\leqslant \frac{|\mathbf{G}|||V||_{\infty}}{n+1}||u'||_{\infty},\quad \forall u\in \mathsf{L}^{1,\infty}(\mathbf{G}). \tag{1.3}$$

In Section 6.1 we show in particular that the factor $(n+1)^{-1}$ in (1.2) and (1.3) is the best possible for each n.

The simplest example of a metric graph is the single segment $[0,L]\subset\mathbb{R}$. For this case, above theorems basically turn into the results of Theorems 3.1 and 3.3 of the paper [1] by Birman and the author (more exactly, into the one-dimensional particular case of these results). The most important feature of estimates (1.1)–(1.3) is their uniformity with respect to all graphs of a given length.

Our proofs are based upon Theorem 2.1 on partitioning of a graph. This theorem can be considered as a far going generalization of Theorem 4.1 from [1]. For trees and absolutely continuous measures $d\mu = V\,dx$ Theorem 2.1 was established in [7].

Let us describe the structure of the paper. The auxiliary result about partitioning of graphs is stated in Section 2, its proof is postponed until Section 5. In Section 3 we prove Theorems 1.1 and 1.2, more exactly we are dealing with their generalizations to the Sobolev spaces with weights. In Section 4 we consider Besov spaces of smoothness order $\theta<1$ and prove the corresponding analogs of Theorems 1.1 and 1.2.

The final Section 6 is devoted to discussion of the results obtained. In particular, we interpret our results in terms of approximation numbers of the appropriate embedding operators. We also show that in the case when $\mathbf{G}$ is a tree Theorem 1.2 and its generalization, Theorem 3.2, can be translated into the language of Hardy-type integral operators. The behavior of approximation numbers of such operators was studied in detail in [4], and there are some important intersections between our corresponding results. We discuss them in Section 6.5.

For $p=2$, the results about approximation can be reformulated in terms of the eigenvalue estimates for certain compact operators in a Hilbert space. In the present paper we do not touch upon this problem. For the most important case of Theorem 1.2 and absolutely continuous measures μ this was done in [7], and similar applications of our other results can be obtained in the same way.

2. The key auxiliary result

Let $\mathbf{G}$ be a compact graph. We always consider connected graphs, including the ones with loops and multiple joins. For two vertices v,w the notation $v\sim w$ means that there exists an edge $e\in\mathscr{E}$ whose ends are v and w. Connectedness of the graph means that for any two vertices $v,w\in\mathscr{V}$, $v\neq w$ there exists a sequence $\{v_k\}_{0\leqslant k\leqslant m}$ of vertices, such that $v_0=v$, $v_m=w$ and $v_{k-1}\sim v_k$ for each $k=1,\dots,m$. The *combinatorial distance* $\rho_{\mathrm{comb}}(v,w)$ is defined as the minimal possible m in this construction. We let $\rho_{\mathrm{comb}}(v,v)=0$ for any $v\in\mathscr{V}$.

The *degree* $d(v)$ of a vertex v is the total number of edges incident to v. The graphs $\mathbf{G}$ consisting of a single vertex (i.e. $\#\mathscr{V}(\mathbf{G}) = 1,\ \mathscr{E}(\mathbf{G}) = \emptyset$) are called degenerate. If the (connected) graph $\mathbf{G}$ is non-degenerate, then its vertices v with $d(v) = 1$ form its boundary $\partial\mathbf{G}$.

We say that a graph G is a *subgraph of* $\mathbf{G}$ if G is a closed and connected subset of $\mathbf{G}$. According to this definition, the vertices of a subgraph not necessarily are vertices of the original graph. For this reason, it is often convenient to treat an arbitrary point $x \in \mathbf{G}$ as a vertex. We set $d(x) = 2$ for any $x \notin \mathscr{V}(\mathbf{G})$ and write $v \sim x$ if $v \in \mathscr{V}(\mathbf{G})$ is one of the endpoints of the edge containing x. Given a subgraph G, we denote by $d_G(x)$ the degree of a point $x \in G$ with respect to G. Clearly, always $d_G(x) \leqslant d(x)$. Note also that $\rho_G(x, y) \geqslant \rho_{\mathbf{G}}(x, y)$ for any $x, y \in G$.

Along with subgraphs, our constructions involve arbitrary connected, not necessarily closed subsets $E \subset \mathbf{G}$. Below $\mathscr{C}(\mathbf{G})$ stands for the set of all such subsets. If $E \in \mathscr{C}(\mathbf{G})$, then the closure $\bar{E}$ is a subgraph, and the complement $\bar{E} \setminus E$ is a finite set. The distinction between E and $\bar{E}$ is important only when dealing with measures $\mu \in \mathfrak{M}(\mathbf{G})$ having non-zero point charges.

We denote by $\sqcup$ the union of subsets which are mutually disjoint, and say that the subsets $E_1, \dots, E_k \in \mathscr{C}(\mathbf{G})$ form *a partition*, or *a splitting* of a set $E \in \mathscr{C}(\mathbf{G})$, if $E = E_1 \sqcup \dots \sqcup E_k$. If $E, E_1 \in \mathscr{C}(\mathbf{G})$ and $E_1 \subset E$, then sets $E_2, \dots, E_k \in \mathscr{C}(\mathbf{G})$ can be always found which together with E_1 form a partition of E.

Let Φ be a non-negative function defined on the set $\mathscr{C}(\mathbf{G})$ and taking values in $[0, \infty)$. We call the function Φ *super-additive* if

$$E = \bigsqcup_{j=1}^{k} E_j \Rightarrow \sum_{j=1}^{k} \Phi(E_j) \leqslant \Phi(E). \tag{2.1}$$

It is clear that any super-additive function is monotone:

$$E_1 \subset E \Rightarrow \Phi(E_1) \leqslant \Phi(E). \tag{2.2}$$

We are interested in the class $\mathsf{S}(\mathbf{G})$ consisting of all super-additive functions satisfying some additional properties which are listed below.

(1) Let $\{E^r\},\ r \in \mathbb{N}$ be a family of sets from $\mathscr{C}(\mathbf{G})$. Then

$$\Phi(E^r) \to \Phi\Big(\bigcap_n E^n\Big) \text{ as } r \to \infty \quad \text{if } E^1 \supset E^2 \supset \dots \tag{2.3}$$

$$\Phi(E^r) \to \Phi\Big(\bigcup_n E^n\Big) \text{ as } r \to \infty \quad \text{if } E^1 \subset E^2 \subset \dots\ . \tag{2.4}$$

(2) $\Phi(\{x\}) = 0$ for any $x \in \mathbf{G}$.

Let $\mathfrak{M}_0(\mathbf{G})$ stand for the set of all measures $\mu \in \mathfrak{M}(\mathbf{G})$, such that μ has no points of positive measure. It is clear that $\mathfrak{M}_0(\mathbf{G}) \subset \mathsf{S}(\mathbf{G})$. A more general example is given by the implication

$$\Phi(E) = \mu_1(E)^\alpha \mu_2(E)^{1-\alpha}, \quad \mu_1 \in \mathfrak{M}_0(\mathbf{G}),\ \mu_2 \in \mathfrak{M}(\mathbf{G}),\ 0<\alpha<1\ \Rightarrow \Phi \in \mathsf{S}(\mathbf{G}). \tag{2.5}$$

Indeed, the super-additivity of Φ is implied by Hölder's inequality, (1) follows from the standard properties of measures, and (2) follows from the condition $\mu_1 \in \mathfrak{M}_0(\mathbf{G})$.

It is important for the applications that only one of two measures $\mu_1,\ \mu_2$ has to belong to the set $\mathfrak{M}_0(\mathbf{G})$.

Along with partitions, we shall use *pseudo-partitions*. Let $E, \Gamma_1, \dots, \Gamma_r \in \mathscr{C}(\mathbf{G})$ and $E = \bigcup_{j=1}^r \Gamma_j$. We say that this is a pseudo-partition of E, if $\#(\Gamma_i \cap \Gamma_j) < \infty$ for any $i, j = 1, \dots, r,\ i \neq j$. We call a pseudo-partition *nice* if the intersection $\bigcap_{j=1}^r \Gamma_j$ is not empty. This intersection is necessarily finite.

With each function $\Phi \in \mathsf{S}(\mathbf{G})$ we associate another function $\tilde{\Phi}$ which is defined as follows:

$$\tilde{\Phi}(E) = \inf \max_{j=1,\dots,r} \Phi(\Gamma_j), \tag{2.6}$$

where the infimum is taken over the set of all nice pseudo-partitions of the set E.

All our results on approximation will be derived from the following Theorem 2.1 on super-additive functions on $\mathscr{C}(\mathbf{G})$.

Theorem 2.1. *Let $\mathbf{G}$ be a compact metric graph and $\Phi \in \mathsf{S}(\mathbf{G})$. Then for any $n \in \mathbb{N}$ there exists a partition $\mathbf{G} = E_1 \sqcup \dots \sqcup E_k$ of $\mathbf{G}$ into a family of subsets from $\mathscr{C}(\mathbf{G})$ such that $k \leqslant n$ and*

$$\tilde{\Phi}(E_j) \leqslant (n+1)^{-1} \Phi(\mathbf{G}), \quad \forall j = 1, \dots, k. \tag{2.7}$$

The proof is rather complicated and we postpone it until Section 5. For super-additive functions Φ such that

$$\{E, E_0 \in \mathscr{C}(\mathbf{G}),\ |E \backslash E_0| + |E_0 \backslash E| \to 0\} \ \Rightarrow\ \{\Phi(E) \to \Phi(E_0)\}$$

both the formulation and the proof become much more transparent. This happens due to the fact that then $\Phi(E) = \Phi(\bar{E})$ for any $E \in \mathscr{C}(\mathbf{G})$, and the difference between partitions and pseudo-partitions becomes unimportant. This simplified version of Theorem 2.1 was obtained in [7]. The general result we give here, is necessary only for handling measures $\mu \notin \mathfrak{M}_0(\mathbf{G})$ in Theorem 1.2 and its generalizations, Theorems 3.2 and 4.2.

Now we turn to applications of Theorem 2.1.

3. Approximation of weighted Sobolev spaces

3.1. Weighted Sobolev spaces

Theorems 1.1 and 1.2 are particular cases of similar results for the weighted Sobolev spaces. For this reason we do not present separate proofs of the original theorems but do this for the corresponding general results. We start with the necessary definitions.

Let $\mathbf{G}$ be a compact metric graph, $1 \leqslant p \leqslant \infty$, and $p' = p(p-1)^{-1}$. Let $a(x)$ be a measurable function on $\mathbf{G}$ such that $a(x) > 0$ a.e. It is convenient to associate with $a(x)$ another function,

$$w_a(x) = a(x)^{-1/p}, \ p < \infty; \quad w_a(x) = a(x)^{-1}, \ p = \infty. \tag{3.1}$$

Our basic assumption is $w_a \in \mathsf{L}^{p'}(\mathbf{G})$. For $p < \infty$ this is equivalent to $1/a \in \mathsf{L}^{p'-1}(\mathbf{G})$. A function u on $\mathbf{G}$ belongs to the weighted Sobolev space $\mathsf{L}^{1,p}(\mathbf{G}, a)$ if u is continuous on $\mathbf{G}$, its restriction to each edge $e \in \mathscr{E}$ has the distributional derivative u', and $||u'||_{p,a} < \infty$. The latter functional defines on $\mathsf{L}^{1,p}(\mathbf{G}, a)$ a semi-norm vanishing on the subspace $\mathbf{C}$ of constant functions. It is often convenient to factorize $\mathsf{L}^{1,p}(\mathbf{G}, a)$ over $\mathbf{C}$, on the resulting quotient space $\hat{\mathsf{L}}^{1,p}(\mathbf{G}, a) := \mathsf{L}^{1,p}(\mathbf{G}, a)/\mathbf{C}$ the functional $||u'||_{p,a}$ becomes the norm.

3.2. Uniform approximation

If $a \equiv 1$, the following result turns into Theorem 1.1.

Theorem 3.1. *Let $\mathbf{G}$ be a compact graph and let $a(x)$ be a non-negative function on $\mathbf{G}$, such that $w_a \in \mathsf{L}^{p'}(\mathbf{G})$. Then for any function $u \in \mathsf{L}^{1,p}(\mathbf{G}, a)$ and any $n \in \mathbb{N}$ there exists a function $v \in \mathrm{Step}_n(\mathsf{G})$ such that*

$$||u - v||_\infty \leqslant \frac{||w_a||_{p'} ||u'||_{p,a}}{n+1}.$$

If $p = \infty$, the mapping $u \mapsto v$ can be chosen linear.

Proof. 1. Let first $1 < p < \infty$. Let L be a polygonal path on $\mathbf{G}$ connecting two given points x_0, x and parametrized by the ark length. For any function $u \in \mathsf{L}^{1,p}(\mathbf{G}, a)$,

$$u(x) - u(x_0) = \int_L u'(y)\, dy.$$

Indeed, this is clearly true if x, x_0 lie on the same edge, and due to the continuity of u on the whole of $\mathbf{G}$ the equality extends to any $x, x_0 \in \mathbf{G}$. By Hölder's inequality,

$$|u(x) - u(x_0)| \leqslant \left(\int_L w_a^{p'}\, dx \right)^{1/p'} \left(\int_L a(y)|u'(y)|^p dy \right)^{1/p}. \tag{3.2}$$

Given a function $u \in \mathsf{L}^{1,p}(\mathbf{G}, a)$, define the function of subsets $E \in \mathscr{C}(\mathbf{G})$,

$$\Phi_u(E) = ||w_a||_{\mathsf{L}^{p'}(E)} ||u'||_{\mathsf{L}^p(E,a)}. \tag{3.3}$$

Evidently $\Phi_u \in \mathsf{S}(\mathbf{G})$, and $\Phi_u(\mathbf{G}) = ||w_a||_{p'}||u'||_{p,a}$. It follows from (3.2) that

$$\sup_{x\in E} |u(x) - u(x_0)| \leqslant \Phi_u(E), \quad \forall x_0 \in \bar{E}, \tag{3.4}$$

for any set $E \in \mathscr{C}(\mathbf{G})$.

Let now $E = \Gamma_1 \cup \ldots \cup \Gamma_r$ be a nice pseudo-partition of E. According to the definition, there is a point $x_0 \in \bigcap_{j=1}^r \Gamma_j$. Applying inequality (3.4) to each Γ_j, we come to the inequality

$$\sup_{x\in E} |u(x) - u(x_0)| = \max_{j=1,\ldots,r} \sup_{x\in\Gamma_j} |u(x) - u(x_0)| \leqslant \max_{j=1,\ldots,r} \Phi_u(\Gamma_j).$$

Minimizing the right-hand side over the set of all points $x_0 \in \bigcap_j \Gamma_j$ and then over the set of all nice pseudo-partitions of E and taking into account definition (2.6), we find a point $x_E \in \bar{E}$ such that

$$\sup_{x\in E} |u(x) - u(x_E)| \leqslant \tilde{\Phi}_u(E). \tag{3.5}$$

Suppose that the graph $\mathbf{G}$ is split into the union of subsets $E_1, \ldots, E_k \in \mathscr{C}(\mathbf{G})$. Consider the step function $v = \sum_{1\leqslant j\leqslant k} u(x_{E_j})\chi_j$ where χ_j stands for the characteristic function of the set E_j. Then $v \in \mathrm{Step}_n(\mathbf{G})$ and by (3.5)

$$||u - v||_\infty \leqslant \max_{j=1,\ldots,k} \tilde{\Phi}_u(E_j).$$

Using Theorem 2.1, we find a partition with $k \leqslant n$ such that $\tilde{\Phi}_u(E_j) \leqslant (n+1)^{-1}\Phi_u(\mathbf{G})$ for each $j = 1, \ldots, k$. This gives the desired result for $1 < p < \infty$.

The same argument, with minor changes, goes through for $p = 1$; we skip it.

2. Let now $p = \infty$, then we have instead of (3.2):

$$|u(x) - u(x_0)| \leqslant ||au'||_{\mathsf{L}^\infty(L)} \int_L w_a\, dx \leqslant ||au'||_{\mathsf{L}^\infty(\mathbf{G})} \int_L w_a\, dx.$$

The above argument works if instead of (3.3) we take

$$\Phi(E) = ||u'||_{\mathsf{L}^\infty(\mathbf{G},a)} \int_E w_a\, dx.$$

This function of subgraphs depends on $a(x)$ but does not depend on the choice of the function u. Therefore, also the partition $\mathbf{G} = E_1 \sqcup \cdots \sqcup E_k$ constructed according to Theorem 2.1 does not depend on u, and hence the mapping $u \mapsto v$ is linear. $\square$

3.3. Weighted L^p-approximation

Now we turn to a generalization of Theorem 1.2.

Theorem 3.2. *Let* $\mathbf{G}$ *be a compact graph and let* $a(x)$ *be a non-negative function on* $\mathbf{G}$ *such that* $w_a \in \mathsf{L}^{p'}(\mathbf{G})$.

(i) *Let* $1 \leqslant p < \infty$ *and* $\mu \in \mathfrak{M}(\mathbf{G})$. *Then for any* $n \in \mathbb{N}$ *there exists a linear operator* $P_n = P_{n,\mu} : \mathsf{L}^{1,p}(\mathbf{G}, a) \to \mathrm{Step}(\mathbf{G})$ *such that* $\mathrm{rank}(P_n) \leqslant n$ *and*

$$||u - P_n u||_{p,\mu} \leqslant \frac{||w_a||_{p'} \mu(\mathbf{G})^{1/p}}{n+1} ||u'||_{p,a}, \quad \forall u \in \mathsf{L}^{1,p}(\mathbf{G}, a). \tag{3.6}$$

(ii) *Let* $p = \infty$ *and* $d\mu = V\,dx$ where $V \in \mathsf{L}^\infty(\mathbf{G})$. *Then for any* $n \in \mathbb{N}$ *there exists a linear operator* $P_n = P_{n,a} : \mathsf{L}^{1,\infty}(\mathbf{G}, a) \to \mathrm{Step}(\mathbf{G})$ *such that* $\mathrm{rank}(P_n) \leqslant n$ *and*

$$||u - P_n u||_{\infty,V} \leqslant \frac{||w_a||_1 ||V||_\infty}{n+1} ||u'||_{\infty,a}, \quad \forall u \in \mathsf{L}^{1,\infty}(\mathbf{G}, a).$$

Proof. (i) Let $1 < p < \infty$; we do not discuss minor changes needed in the case $p = 1$. The proof is quite similar to the previous one. This time we use the function

$$\Phi_\mu(E) = ||w_a||_{\mathsf{L}^{p'}(E)} \mu(E)^{1/p}, \quad E \in \mathscr{C}(\mathbf{G}),$$

cf. (3.3). By (2.5), this function also lies in $\mathsf{S}(\mathbf{G})$, and $\Phi_\mu(\mathbf{G}) = ||w_a||_{p'} \mu(\mathbf{G})^{1/p}$.

Let $E = \Gamma_1 \cup \ldots \cup \Gamma_r$ be a nice pseudo-partition of a given subset $E \in \mathscr{C}(\mathbf{G})$ and let $x_0 \in \bigcap_{j=1}^r \Gamma_j$. Then we find, using (3.2):

$$\begin{aligned}
\int_E |u(x) - u(x_0)|^p d\mu(x) &\leqslant \sum_{j=1}^r \sup_{x \in \Gamma_j} |u(x) - u(x_0)|^p \mu(\Gamma_j) \\
&\leqslant \sum_{j=1}^r \left(\int_{\Gamma_j} w_a^{p'}\,dx \right)^{p-1} \mu(\Gamma_j) \int_{\Gamma_j} a(y)|u'(y)|^p dy \\
&\leqslant \left(\max_{j=1,\ldots,r} \Phi_\mu(\Gamma_j) \right)^p \int_E a(y)|u'(y)|^p\,dy.
\end{aligned}$$

Minimizing over the set of all points $x_0 \in \bigcap_j \Gamma_j$ and then over the set of all nice pseudo-partitions of E, we find a point $x_{E,\mu} \in \bar{E}$ such that

$$\int_E |u(x) - u(x_{E,\mu})|^p d\mu(x) \leqslant (\tilde{\Phi}_\mu(E))^p \int_E a(y)|u'(y)|^p\,dy. \tag{3.7}$$

Suppose now that the graph $\mathbf{G}$ is split into the union of subsets $E_1, \ldots, E_k \in \mathscr{C}(\mathbf{G})$ and let v be the step function $v = \sum_{1 \leqslant j \leqslant k} u(x_{E_j,\mu})\chi_j$. Then we derive from (3.7) that

$$\begin{aligned}
\int_{\mathbf{G}} |u(x) - v(x)|^p\,d\mu(x) &\leqslant \sum_{j=1}^k (\tilde{\Phi}_\mu(E_j))^p \int_{E_j} a(y)|u'(y)|^p\,dy \\
&\leqslant \left(\max_{j=1,\ldots,k} (\tilde{\Phi}_\mu(E_j) \right)^p \int_{\mathbf{G}} a(y)|u'(y)|^p\,dy.
\end{aligned}$$

Applying Theorem 2.1 to the function Φ_μ, we find a partition with $k \leqslant n$, for which

$$\max_{j=1,\ldots,k} \tilde{\Phi}_\mu(E_j) \leqslant (n+1)^{-1} \Phi_\mu(\mathbf{G}).$$

This partition depends on μ but does not depend on the choice of the function $u \in \mathsf{L}^p(\mathbf{G}, a)$. This implies that the operator $P_n : u \mapsto v$ is linear, and we arrive at (3.6).

(ii) The result is an immediate consequence of Theorem 3.1. $\square$

4. Approximation of Besov spaces $\mathsf{B}^{\theta,p}(\mathbf{G})$ with $\theta < 1$, $p\theta > 1$

4.1. Spaces $\mathsf{B}^{\theta,p}(\mathbf{G})$

As in the previous section, it is convenient for us to consider the spaces factorized over the subspace $\mathbf{C}$ of constant functions. However, in our notations we do not distinguish between a function u and the corresponding factor-element. In order to simplify our reasonings, we consider only $1 < p < \infty$ and the spaces without weights.

The most natural approach to the spaces $\mathsf{B}^{\theta,p}(\mathbf{G})$ uses interpolation between the space $\hat{\mathsf{L}}^{1,p}(\mathbf{G})$, see Section 3.1, and the quotient space $\hat{\mathsf{L}}^p(\mathbf{G}) = \mathsf{L}^p(\mathbf{G})/\mathbf{C}$. As usual, the norm in $\hat{\mathsf{L}}^p(\mathbf{G})$ is defined by

$$||u||_{\hat{\mathsf{L}}^p(\mathbf{G})} = \min_{c \in \mathbb{C}} ||u - c||_p.$$

The spaces $\hat{\mathsf{L}}^p(\mathbf{G})$ and $\hat{\mathsf{L}}^{1,p}(\mathbf{G})$ form a Banach couple, see e.g. [8], and we define the interpolation space

$$\hat{\mathsf{B}}^{\theta,p}(\mathbf{G}) \coloneqq \hat{\mathsf{B}}^{\theta,p}_p(\mathbf{G}) = (\hat{\mathsf{L}}^p(\mathbf{G}), \hat{\mathsf{L}}^{1,p}(\mathbf{G}))_{\theta,p}, \quad 0 < \theta < 1. \tag{4.1}$$

We write $u \in \mathsf{B}^{\theta,p}(\mathbf{G})$, when it is convenient to view u as an individual function rather than the equivalence class $\{u + \mathbf{C}\}$. We do not discuss here interpolation with the second parameter $q \neq p$ which would lead to the general Besov spaces $\mathsf{B}^{\theta,p}_q$.

There are many ways to define an interpolation norm in $\hat{\mathsf{B}}^{\theta,p}$. For our purposes it is convenient to use the L-method with the parameters $p_0 = p_1 = p$, see e.g. [8, Section 1.4]. So, we define for $0 < t < \infty$:

$$\begin{aligned} L(t, u; \mathbf{G}) = \inf\{ & ||u_0||^p_{\hat{\mathsf{L}}^p(\mathbf{G})} + t||u_1'||^p_{\mathsf{L}^p(\mathbf{G})} : u = u_0 + u_1; \\ & u_0 \in \hat{\mathsf{L}}^p(\mathbf{G}), u_1 \in \hat{\mathsf{L}}^{1,p}(\mathbf{G})\}. \end{aligned} \tag{4.2}$$

A function $u \in \mathsf{L}^p(\mathbf{G}) + L^{1,p}(\mathbf{G})$ belongs to the space $\mathsf{B}^{\theta,p}(\mathbf{G})$ if and only if

$$(||u||_{\mathsf{B}^{\theta,p}(\mathbf{G})})^p = (||u||_{\hat{\mathsf{B}}^{\theta,p}(\mathbf{G})})^p \coloneqq \int_0^\infty t^{-1-\theta} L(t, u; \mathbf{G})\, dt < \infty. \tag{4.3}$$

Replacing in (4.3) the graph $\mathbf{G}$ by its arbitrary subset $E \in \mathscr{C}(\mathbf{G})$ and fixing an element $u \in \mathsf{B}^{\theta,p}(\mathbf{G})$, we obtain the function

$$J_{\theta,u}(E) = (||u||_{\hat{\mathsf{B}}^{\theta,p}(E)})^p, \quad E \in \mathscr{C}(\mathbf{G}). \tag{4.4}$$

Let us show that $J_{\theta,u} \in \mathsf{S}(\mathbf{G})$. First of all, we note that the function $J_{0,u}(E) = ||u||^p_{\hat{\mathsf{L}}^p(E)}$ lies in $\mathsf{S}(\mathbf{G})$. Indeed, the properties (1) and (2) of functions $\Phi \in \mathsf{S}(\mathbf{G})$, cf. Section 2, are evidently fulfilled, and for any constant c and any partition $E = E_1 \sqcup \cdots \sqcup E_k$

we have

$$\int_E |u-c|^p\,dx = \sum_{j=1}^{k} \int_{E_j} |u-c|^p\,dx \geqslant \sum_{j=1}^{k} \inf_{c_j\in\mathbb{C}} \int_{E_j} |u-c_j|^p\,dx$$

which yields super-additivity. The function $J_{1,u}(E) = ||u'||^p_{\mathsf{L}^p(E)}$ also lies in $\mathsf{S}(\mathbf{G})$, therefore the same is true for the function $L(t,u;G)$ defined by equality (4.2) for the set $E\in\mathscr{C}(\mathbf{G})$ substituted for $\mathbf{G}$. Integration in (4.3) does not violate the property of a function to lie in $\mathsf{S}(\mathbf{G})$. Hence, it is proved that $J_{\theta,u}\in\mathsf{S}(\mathbf{G})$. It follows from here and (2.2) that

$$||u||^p_{\hat{\mathsf{B}}^{\theta,p}(E_1)} \leqslant ||u||^p_{\hat{\mathsf{B}}^{\theta,p}(E)}, \quad \forall E, E_1\in\mathscr{C}(\mathbf{G}),\ E_1\subset E. \tag{4.5}$$

If $\theta p>1$, any function $u\in\mathsf{B}^{\theta,p}(\mathbf{G})$ is continuous. This is well known when $\mathbf{G}$ is a single segment. Hence, u is continuous on any polygonal path in $\mathbf{G}$ and thus, on the whole of $\mathbf{G}$.

Denote by $C(\theta,p)$ the sharp constant in the inequality

$$\max_{x,x_0\in[0,l]} |u(x)-u(x_0)| \leqslant C(\theta,p)l^{\theta-1/p}||u||_{\mathsf{B}^{\theta,p}[0,l]}. \tag{4.6}$$

The value of $C(\theta,p)$ does not depend on l, which follows from the homogeneity arguments. Inequality (4.6) automatically extends to the graphs: due to (4.5),

$$\sup_{x,x_0\in E} |u(x)-u(x_0)| \leqslant C(\theta,p)|E|^{\theta-1/p}J_{\theta,u}(E)^{1/p}, \quad \forall E\in\mathscr{C}(\mathbf{G}), \tag{4.7}$$

where the function $J_{\theta,u}(E)$ is defined by (4.4).

4.2. Approximation of $\mathsf{B}^{\theta,p}$

Below are analogs of Theorems 1.1 and 1.2 for the spaces $\mathsf{B}^{\theta,p}(\mathbf{G})$.

Theorem 4.1. *Let* $\mathbf{G}$ *be a compact graph*, $0<\theta<1$, *and* $1/\theta<p<\infty$. *Then for any function* $u\in\mathsf{B}^{\theta,p}(\mathbf{G})$ *and any* $n\in\mathbb{N}$ *there exists a function* $v\in\mathrm{Step}_n(\mathbf{G})$ *such that*

$$||u-v||_\infty \leqslant C(\theta,p)\frac{|\mathbf{G}|^{\theta-1/p}||u||_{\mathsf{B}^{\theta,p}(\mathbf{G})}}{(n+1)^\theta}. \tag{4.8}$$

We only outline the proof; details can be easily reconstructed by analogy with Theorem 3.1.

Together with $J_{\theta,u}(E)$, the function

$$\Phi_u(E) = |E|^{1-1/(p\theta)}J_{\theta,u}(E)^{1/(p\theta)} \tag{4.9}$$

also belongs to $\mathsf{S}(\mathbf{G})$, cf. (2.5). Let subsets $\Gamma_1,\dots,\Gamma_r\in\mathscr{C}(\mathbf{G})$ form a nice pseudo-partition of a set $E\in\mathscr{C}(\mathbf{G})$ and let x_0 be a point from their intersection.

The inequality

$$\sup_{x\in E}|u(x)-u(x_0)|\leqslant C(\theta,p)\left(\max_{j=1,\dots,r}\Phi_u(\Gamma_j)\right)^\theta$$

is easily derived from (4.7). Minimizing over the set of all points from $\bigcap_j \Gamma_j$ and then, over the set of all nice pseudo-partitions of E, we find a point $x_E\in\bar{E}$, such that

$$\sup_{x\in E}|u(x)-u(x_E)|\leqslant C(\theta,p)(\widetilde{\Phi}_u(E))^\theta.$$

The proof is concluded by applying Theorem 2.1 to function (4.9).

Theorem 4.2. *Let* $\mathbf{G}$ *be a compact graph,* $0<\theta<1$, *and* $1/\theta<p<\infty$. *Let* $\mu\in\mathfrak{M}(\mathbf{G})$. *Then for any* $n\in\mathbb{N}$ *there exists a linear operator* $P_n:\mathsf{B}^{\theta,p}(\mathbf{G})\to\mathrm{Step}(\mathbf{G})$ *such that* $\mathrm{rank}(P_n)\leqslant n$ *and*

$$||u-P_nu||_{p,\mu}\leqslant C(\theta,p)\frac{|\mathbf{G}|^{\theta-1/p}\mu(\mathbf{G})^{1/p}}{(n+1)^\theta}||u||_{\mathsf{B}^{\theta,p}},\quad \forall u\in\mathsf{B}^{\theta,p}(\mathbf{G}). \tag{4.10}$$

Again, we only sketch the proof. We make use of the function

$$\Phi_\mu(E)=|E|^{1-1/(\theta p)}\mu(E)^{1/(\theta p)}$$

which by (2.5) belongs to $\mathsf{S}(\mathbf{G})$. For any set $E\in\mathscr{C}(\mathbf{G})$ we find a point $x_{E,\mu}\in\bar{E}$ such that

$$\int_E|u(x)-u(x_{E,\mu})|^p d\mu(x)\leqslant C(\theta,p)^p(\widetilde{\Phi}_\mu(E))^{\theta p}J_{\theta,u}(E), \tag{4.11}$$

cf. (3.4). Let $\mathbf{G}=E_1\sqcup\dots\sqcup E_k$ be an arbitrary partition of the graph $\mathbf{G}$. Let $v=\sum_{j=1}^k u(x_{E_j,\mu})\chi_j$, then we derive from (4.11) using the super-additivity of $J_{\theta,u}$:

$$\begin{aligned}\int_{\mathbf{G}}|u-v|^p d\mu(x)&=\sum_{j=1}^k\int_{E_j}|u-u(x_{E_j,\mu})|^p d\mu(x)\leqslant C(\theta,p)^p(\widetilde{\Phi}_\mu(E_j))^{\theta p}J_{\theta,u}(E_j)\\&\leqslant C(\theta,p)^p\left(\max_{j=1,\dots,k}\widetilde{\Phi}_\mu(E_j)\right)^{\theta p}J_{\theta,u}(\mathbf{G}).\end{aligned}$$

We come to the desired result applying Theorem 2.1 to the function Φ_μ and taking into account that the mapping $P:u\mapsto v$ is linear.

5. Proof of Theorem 2.1

5.1. The case of trees

Let $\mathbf{G} = \mathbf{T}$ be a tree, that is connected graph without cycles, loops and multiple joins. For any two points $x, y \in \mathbf{T}$ there exists a unique simple polygonal path in $\mathbf{T}$ connecting x with y, we denote it by $\langle x, y \rangle$. It is clear that $|\langle x, y \rangle| = \rho(x, y)$.

For trees the notion of nice pseudo-partition simplifies. Indeed, if $T = \Theta_1 \cup \cdots \cup \Theta_r$ is a nice pseudo-partition of a (closed) subtree $T \subset \mathbf{T}$, then the intersection $\Xi = \bigcap_{j=1}^r \Theta_j$ consists of exactly one point. For if $x_1 \neq x_2$ and $x_1, x_2 \in \Xi$, then also $\langle x_1, x_2 \rangle \subset \Xi$ which contradicts the definition of pseudo-partition. So, the point $x \in \bigcap_j \Theta_j$ is uniquely defined by a nice pseudo-partition. Besides, all the subsets Θ_j are necessarily closed, i.e. each of them is a subtree of T.

Conversely, each nice pseudo-partition of T is uniquely determined by the choice of the point x. Indeed, the tree T splits in a unique way into the union of subtrees $\Theta_j \subset T$, $j = 1, \ldots, d_T(x)$, rooted at x and such that $d_{\Theta_j}(x) = 1$ for each j. Evidently this pseudo-partition is nice. We call the pair $\{T, x\}$ a *punctured subtree* and the above constructed partition—its *canonical pseudo-partition*.

Let $\Phi \in \mathsf{S}(\mathbf{T})$. Defining

$$\Phi'(T, x) = \max_{j=1,\ldots,d_T(x)} \Phi(\Theta_j), \tag{5.1}$$

we evidently have

$$\widetilde{\Phi}(T) = \min_{x \in T} \Phi'(T, x). \tag{5.2}$$

Let in particular $T = \mathbf{T}$. Each subtree Θ_j appearing in the canonical pseudo-partition of $\{\mathbf{T}, x\}$ is determined by indication of its initial edge $\langle x, v \rangle$, $v \sim x$ and we denote this subtree by $\Theta_{\langle x,v \rangle}$. For $T = \mathbf{T}$ definition (5.1) takes the form

$$\Phi'(\mathbf{T}, x) = \max_{v \sim x} \Phi(\Theta_{\langle x,v \rangle}).$$

The following lemma is the heart of our proof of Theorem 2.1.

Lemma 5.1. *Let* $\mathbf{T}$ *be a compact metric tree and* $\Phi \in \mathsf{S}(\mathbf{T})$. *Then for any* $\varepsilon \in (0, \Phi(\mathbf{T}))$ *there exists a pseudo-partition* $\mathbf{T} = T \cup T'$, *such that the set* $T' \setminus T$ *is connected* (*that is, belongs to* $\mathscr{C}(\mathbf{T})$) *and for the single point* $x^* \in T \cap T'$ *the inequalities hold*:

$$\Phi'(T, x^*) \leqslant \varepsilon, \tag{5.3}$$

$$\Phi(T' \setminus \{x^*\}) \leqslant \Phi(\mathbf{T}) - \varepsilon. \tag{5.4}$$

Proof. Without loss of generality, we can assume $\Phi(\mathbf{T}) = 1$. Take any vertex $v_0 \in \partial\mathbf{T}$, then $\Phi'(\mathbf{T}, v_0) = \Phi(\mathbf{T}) = 1$. There is a unique vertex $v_1 \sim v_0$. Now we choose the vertices $v_2 \sim v_1, \ldots, v_{k+1} \sim v_k, \ldots$ as follows. If v_k is already chosen, we define v_{k+1} as the vertex different from v_{k-1} and such that

$$\Phi(\Theta_{\langle v_k, v_{k+1}\rangle}) = \max_{w\sim v_k, w\neq v_{k-1}} \Phi(\Theta_{\langle v_k, w\rangle}) = \Phi'(\Theta_{\langle v_k, v_{k+1}\rangle}, v_k). \tag{5.5}$$

If there are several vertices $w\sim v_k$ at which the maximum in the middle term of (5.5) is attained, then any of them can be chosen as v_{k+1}. The described procedure is always finite, it terminates when we arrive at a vertex $v_m \in \partial\mathbf{T}$. On the path $\mathscr{P} = \langle v_0, v_m\rangle$ we introduce the natural ordering, i.e. $y \succeq x$ means that $x \in \langle v_0, y\rangle$. We write $y \succ x$ if $y \succeq x$ and $y \neq x$.

Let $x \in \mathscr{P}$ be not a vertex of $\mathbf{T}$, then $v_{k-1} \prec x \prec v_k$ for some $k = 1, \dots, m$. Denote

$$T_x^+ = \Theta_{\langle x, v_k\rangle}, \quad T_x^- = \Theta_{\langle x, v_{k-1}\rangle}, \quad x \neq v_0, \dots, v_m.$$

We also define the subtrees $T_x^{\pm}$ for $x = v_0, \dots, v_m$. Namely,

$$T_{v_k}^- = T_{\langle v_k, v_{k-1}\rangle}, \quad k = 1, \dots, m,$$

$$T_{v_0}^+ = \mathbf{T}, \quad T_{v_k}^+ = \bigcap_{v_{k-1}\prec x \prec v_k} T_x^+ = \bigcup_{v\sim v_k, v\neq v_{k-1}} T_{\langle v_k, v\rangle}, \quad k = 1, \dots, m-1.$$

Finally, $T_{v_0}^- = \{v_0\}$, $T_{v_m}^+ = \{v_m\}$ are degenerate subtrees. For any $x \in \mathscr{P}$ we have $\mathbf{T} = T_x^+ \cup T_x^-$. Clearly, this is a pseudo-partition of the tree $\mathbf{T}$, and $T_x^+ \cap T_x^- = \{x\}$. Besides, for any $x \in \mathscr{P}$ we have $x \in \partial T_x^-$, and the set $T_x^- \setminus T_x^+ = T_x^- \setminus \{x\}$ is connected, i.e. belongs to $\mathscr{C}(\mathbf{T})$.

The function $F(x) = \Phi(T_x^+)$ is well defined on $\mathscr{P}$ and non-increasing. By (2.3), F is left-continuous with respect to the ordering adopted. By the construction,

$$\Phi'(T_{x_0}^+, x_0) = F(x_0), \quad \forall x_0 \in \mathscr{P}.$$

Further, (2.4) implies that

$$F(x_0+) := \lim_{x\succ x_0, x\to x_0} F(x) = \Phi(T_{x_0}^- \setminus \{x_0\}), \quad \forall x_0 \in \mathscr{P}.$$

We also have

$$0 = F(v_m) < \varepsilon < F(v_0) = 1.$$

Therefore, there exists a point $x^* \in \mathscr{P}$ such that

$$\Phi'(T_{x^*}, x^*) = F(x^*) \geqslant \varepsilon \geqslant F(x^*+).$$

We take $T = T_{x^*}^+$ and $T' = T_{x^*}^-$. Then inequality (5.3) is satisfied and (5.4) is implied by super-additivity:

$$\Phi(T' \setminus \{x^*\}) \leqslant 1 - \Phi(T) = 1 - F(x^*) \leqslant 1 - \varepsilon. \qquad \square$$

5.2. *Proof of Theorem 2.1 for the case of trees*

Let $\mathbf{G} = \mathbf{T}$ be a tree.

1. Let $n = 1$. Apply the result of Lemma 5.1 with $\varepsilon = \Phi(\mathbf{T})/2$. Let $\mathbf{T} = T \cup T'$ be the corresponding pseudo-partition, then $\Phi'(T', x^*) \leqslant \Phi(T') \leqslant \Phi(\mathbf{T})/2$. Consider the canonical pseudo-partition of the punctured tree $\{\mathbf{T}, x^*\}$. Each subtree of this

pseudo-partition is contained either in T or in T', therefore

$$\Phi'(\mathbf{T}, x^*) \leqslant \max(\widetilde{\Phi}(T, x^*), \Phi'(T', x^*)) \leqslant \Phi(\mathbf{T})/2.$$

Taking (5.2) into account, we see that (2.7) with $k = n = 1$ is satisfied if we take $E_1 = \mathbf{T}$.

2. We proceed by induction. Suppose that the result is already proved for $n = n_0 - 1$. Let $\mathbf{T} = T \cup T'$ be the pseudo-partition constructed according to Lemma 5.1 for $\varepsilon = (n_0 + 1)^{-1}\Phi(\mathbf{T})$ and let $T \cap T' = \{x^*\}$. Then

$$\Phi(T' \setminus \{x^*\}) \leqslant n_0(n_0 + 1)^{-1}\Phi(\mathbf{T}).$$

Let us define a function Φ' of subsets $E \in \mathscr{C}(T')$, taking

$$\Phi'(E) = \Phi(E \setminus \{x^*\}), \quad \forall E \in \mathscr{C}(T'),$$

then evidently $\Phi' \in \mathsf{S}(T')$. By the inductive hypothesis, there exists a splitting of T' into the union of subsets $E_j \in \mathscr{C}(T')$, $j = 1, \dots, k$ such that $k \leqslant n_0 - 1$ and for each j

$$\hat{\Phi}'(E_j) \leqslant n_0^{-1}\Phi'(T') = n_0^{-1}\Phi(T' \setminus \{x^*\}) \leqslant (n_0 + 1)^{-1}\Phi(\mathbf{T}).$$

The point x^* lies in only one of the sets E_j, let it be E_k. Since $x^* \in \partial T'$, we conclude that the set $E_k \setminus \{x^*\}$ is connected and therefore belongs to $\mathscr{C}(\mathbf{T})$.

The family $E_1, \dots, E_{k-1}, E_k \setminus \{x^*\}, T$ forms the desired partition of $\mathbf{T}$ for $n = n_0$. For the trees, the proof of Theorem 2.1 is complete.

5.3. General case

Theorem 2.1, for arbitrary graphs, can be easily reduced to the case of trees by means of "cutting cycles". Below we describe the procedure of such reduction.

Let $\mathbf{G}$ be a compact graph and Φ be a function from $\mathsf{S}(\mathbf{G})$. Let e be an edge of $\mathbf{G}$ which is a part of a cycle. Supposing that e is not a loop, we identify e with the segment $[0, |e|]$. Take any point $x \in \mathrm{Int}(e)$ and replace it by the pair x_1, x_2 of new vertices. Respectively, the edge e is replaced by the pair e_1, e_2 of new edges whose total length is equal to $|e|$. As the result, we obtain a new graph, say $\mathbf{G}_1$. Note that the edges e_1, e_2 are parts of no cycle in $\mathbf{G}_1$. Define the mapping $\tau_1 : \mathbf{G}_1 \to \mathbf{G}$ which is identical on $\mathbf{G} \setminus \mathrm{Int}(e)$ and sends isometrically e_1 onto $[0, x]$ and e_2 onto $[x, |e|]$. The mapping τ_1 is one-to-one on $\mathbf{G}_1 \setminus \{x_1, x_2\}$, and $\tau_1(x_1) = \tau_1(x_2) = x$. It is clear that τ_1 is non-expanding and hence, continuous.

The changes in this construction, needed if e is a loop, are evident.

Now, define a function Φ_1 on the set $\mathscr{C}(\mathbf{G})$, namely

$$\Phi_1(E) = \Phi(\tau_1(E)) \text{ if } x_1 \in E, \quad \Phi_1(E) = \Phi(\tau_1(E) \setminus \{x\}) \text{ if } x_1 \notin E.$$

The function Φ_1 is super-additive. Indeed, let $E \in \mathscr{C}(\mathbf{G}_1)$ and $E = \bigsqcup_{j=1}^{k} E_j$. If $x_1 \notin E$, then also $x_1 \notin E_j$ for any j, and if $x_1 \in E$, then $x_1 \in E_{j_0}$ for exactly one value of j. In both cases, inequality (2.1) for Φ_1 is implied by the similar inequality for Φ.

Properties (1) and (2) for the function Φ_1 also follow from the same properties for Φ. Hence, $\Phi_1 \in \mathsf{S}(\mathbf{G}_1)$.

Repeating this procedure, we obtain a sequence of graphs $\mathbf{G}_0 := \mathbf{G}, \mathbf{G}_1, \dots, \mathbf{G}_m$, a sequence of mappings $\tau_j : \mathbf{G}_j \to \mathbf{G}_{j-1}$, $j = 1, \dots, m$, and a family of functions $\Phi_j \in \mathsf{S}(\mathbf{G}_j)$. The procedure stops as soon as we come to a graph without cycles and loops, that is when $\mathbf{G}_m =: \mathbf{T}$ is a compact tree. The mapping $\tau = \tau_m \circ \cdots \circ \tau_1 : \mathbf{T} \to \mathbf{G}$ is continuous and measure preserving. Due to the continuity of τ, $T \in \mathscr{C}(\mathbf{T}) \Rightarrow \tau(T) \in \mathscr{C}(\mathbf{G})$. Besides, τ transforms partition into partition and preserves the property of a partition to be nice. The function Φ_m belongs to $\mathsf{S}(\mathbf{G}_m)$ and $\Phi_m(\mathbf{T}) = \Phi(\mathbf{G})$.

By the result of previous section, for a given $n \in \mathbb{N}$ there exists a partition $\mathbf{T} = \bigsqcup_{j=1}^{k} E_j$ into the union of subsets from $\mathscr{C}(\mathbf{T})$, such that $k \leqslant n$ and $\widetilde{\Phi}_m(E_j) \leqslant (n+1)^{-1}\Phi_m(\mathbf{T}) = (n+1)^{-1}\Phi(\mathbf{G})$ for each j. Taking $E'_j = \tau(E_j)$, we find a partition of $\mathbf{G}$ which meets all the requirements of Theorem 2.1.

6. Complements and concluding remarks

6.1. On the sharpness of estimates

(a) The factor $(n+1)^{-1}$ in inequality (2.7) of Theorem 2.1 is sharp for each n. To see this, consider the star graph $\mathbf{G}_N$ consisting of N edges $e_k = \langle o, v_k \rangle$, $k = 1, \dots, N$ of equal length 1, all emanating from the root o. For any subset $E \in \mathscr{C}(\mathbf{G}_N)$ we define $\Phi(E) = |E|$, then $\Phi \in \mathsf{S}(\mathbf{G}_N)$. Take $n = N-1$, then at least one of the subsets E_j appearing in the conclusion of Theorem 2.1 necessarily contains two edges of $\mathbf{G}_N$. Thus, $\Phi(E_j) \geqslant 2$ and hence, $\widetilde{\Phi}(E_j) \geqslant 1$ for any nice pseudo-partition of E_j. Since $|\mathbf{G}_N| = N = n+1$, we see that inequality (2.7) turns into equality.

(b) The same factor $(n+1)^{-1}$ in inequality (1.2) of Theorem 1.2 is also sharp for each n. Indeed, consider the star graph $\mathbf{G}_N$ and the measure $\mu \in \mathfrak{M}(\mathbf{G}_N)$ defined as $\mu = \delta_{v_1} + \cdots + \delta_{v_N}$. Consider also the subspace $Y \subset \mathsf{L}^{1,p}(\mathbf{G}_N)$ formed by the functions u such that $u \restriction e_k = c_k \rho(o, x)$, $k = 1, \dots, N$. Then

$$\|u'\|_p = \|u\|_{\mathsf{L}^p(\mathbf{G}_N,\mu)} = \|c\|_{\ell^p_N}, \quad c = \{c_k\}_{1 \leqslant k \leqslant N}, \ \forall u \in Y.$$

It follows that for any linear operator $P : \mathsf{L}^{1,p}(\mathbf{G}) \to \mathsf{L}^p(\mathbf{G}_N, \mu)$ with $\operatorname{rank}(P) \leqslant n$ the quantity

$$\inf_{u \in \mathsf{L}^{1,p}(\mathbf{G}):\ \|u'\|_p = 1} \|u - Pu\|_{\mathsf{L}^p(\mathbf{G}_N,\mu)}$$

is no smaller than the n-width in ℓ^p_N of the unit ball of this space. For $n < N$ this n-width is equal to one, see e.g. [6], Proposition 1.3. Since $|\mathbf{G}_N| = \mu(\mathbf{G}_N) = N$, we see that for $n = N-1$ an element $u \in \mathsf{L}^{1,p}(\mathbf{G}) : \|u'\|_p = 1$ can always be found in such a way that

$$||u - Pu||_{\mathsf{L}^p(\mathbf{G}_N,\mu)} \geq 1 = \frac{|\mathbf{G}_N|^{1/p'}\mu(\mathbf{G}_N)^{1/p}}{n+1}.$$

Replacing the above measure μ by a sequence of measures $V_j\,dx$ which $*$-weakly approximate μ, we find that the factor $(n+1)^{-1}$ in (1.2) is the least possible also for absolutely continuous measures. However, for each particular absolutely continuous measure μ inequality in (1.2) is always strict.

(c) The same factor in the inequality (1.1) is sharp for $n = 1$. For $n > 1$ it becomes sharp, provided one passes to the version of Theorem 1.1 (and its generalization, Theorem 3.1) dealing with vector-valued functions. Namely, let X be a Banach space and let $\mathsf{L}^{1,p}(\mathbf{G}; X)$ stand for the space of X-valued functions on $\mathbf{G}$ whose definition is clear by analogy with the case of scalar-valued functions, cf. Section 1. Both mentioned theorems extend to the spaces $\mathsf{L}^{1,p}(\mathbf{G}; X)$, the proof actually remains the same.

Now, take $X = \ell^\infty$. For $k \in \mathbb{N}$, let $\eta_k \in \ell^\infty$ be the element whose kth coordinate is 1 and all the others are equal to zero. On the star graph $\mathbf{G}_N$ consider the function u which is $\eta_k \rho(0, x)$ on the edge $e_k \in \mathbf{G}_N$. Then $u \in \mathsf{L}^{1,p}(\mathbf{G}_N; X)$ for each $p \in [1, \infty]$ and $||u'||_{L^p(\mathbf{G}_N;X)} = 1$. The same reasoning as in (a) shows that for $n = N - 1$ the constant factor $(n+1)^{-1}$ in the vector-valued version of (1.1) is the best possible.

6.2. Graphs and trees:comparison of the corresponding results

Given a compact graph $\mathbf{G}$, let $\mathbf{T}$ and $\tau : \mathbf{T} \to \mathbf{G}$ be the tree and the mapping constructed in Section 5.3. Let $a(x)$ be a non-negative function on $\mathbf{G}$ such that $w_a \in \mathsf{L}^{p'}(\mathbf{G})$ (cf. (3.1)). Define $b(x) = a(\tau(x))$, then $w_b \in \mathsf{L}^{p'}(\mathbf{T})$ and $||w_b||_{\mathsf{L}^{p'}(\mathbf{T})} = ||w_a||_{\mathsf{L}^{p'}(\mathbf{G})}$. Moreover, it is clear from the construction that the mapping $u(x) \mapsto v(x) = u(\tau(x))$ defines an isometry between the space $\mathsf{L}^{1,p}(\mathbf{G}, a)$ and an appropriate subspace of finite codimension in $\mathsf{L}^{1,p}(\mathbf{T}, b)$. Indeed, suppose that the passage from the graph $\mathbf{G}$ to the tree $\mathbf{T}$ consists in replacing the points $x^{(j)} \in \mathbf{G}$, $j = 1, \dots, m$ by the pairs $\{x_1^{(j)}, x_2^{(j)}\} \subset \mathbf{T}$. Then the space $\mathsf{L}^{1,p}(\mathbf{G}, a)$ can be identified with the subspace

$$\{u \in \mathsf{L}^{1,p}(\mathbf{T}, b)\colon\ u(x_1^{(j)}) = u(x_2^{(j)}),\ j = 1, \dots, m.\}$$

The above mapping $u \mapsto v$ defines also the natural isometry between the spaces $\mathsf{L}^p(\mathbf{G}, V)$ and $\mathsf{L}^p(\mathbf{T}, W)$ where $W(x) = V(\tau(x))$. It follows from these remarks that Theorem 3.2 for general graphs reduces to its particular case for trees.

The same is true for Theorem 4.2, though for the spaces $\mathsf{B}^{\theta,p}$ the above mapping $u \mapsto v$ is not necessarily an isometry. But this is always a contraction, so that the constant in estimate (4.10) for a graph $\mathbf{G}$ cannot exceed the one for the corresponding tree $\mathbf{T}$.

6.3. Approximation numbers of embedding operators

Suppose that a point $o\in\mathbf{G}$ is fixed, and define the spaces

$$\mathsf{W}^{1,p}(\mathbf{G},a;o)=\{u\in\mathsf{L}^{1,p}(\mathbf{G},a):\ u(o)=0\}$$

and, for $0<\theta<1$ and $p>1/\theta$,

$$\mathsf{B}^{\theta,p}(\mathbf{G};o)=\{u\in\mathsf{B}^{\theta,p}(\mathbf{G}):\ u(o)=0\}.$$

We take $||u'||_{p,a}$ as the norm in $\mathsf{W}^{1,p}(\mathbf{G},a;o)$ and $||u||_{\mathsf{B}^{\theta,p}}$ (cf. (4.3)) as the norm in $\mathsf{B}^{\theta,p}(\mathbf{G};o)$. It is clear that the spaces $\mathsf{W}^{1,p}(\mathbf{G},a;o)$ and $\mathsf{B}^{\theta,p}(\mathbf{G};o)$ are naturally isometric to the quotient spaces $\hat{\mathsf{L}}^{1,p}(\mathbf{G},a)$ and $\hat{\mathsf{B}}^{\theta,p}(\mathbf{G})$, respectively. For this reason, Theorems of Sections 3 and 4 immediately apply to the spaces $\mathsf{W}^{1,p}(\mathbf{G},a;o)$ and $\mathsf{B}^{\theta,p}(\mathbf{G};o)$.

Given two Banach spaces Y and X and an integer $n\geqslant 0$, let P_n stand for the set of all linear mappings $P:Y\to X$ whose rank does not exceed n. Recall the definition of the approximation numbers $a_n(T)$ of a bounded linear operator $T:X\to Y$, see e.g. [2]:

$$a_n(T)=\inf_{P\in\mathsf{P}_{n-1}(Y,X)}||T-P||_X. \tag{6.1}$$

In particular, this definition applies to the case when Y is embedded in X algebraically and topologically, and $T=J_{Y,X}$ is the corresponding embedding operator. Theorem 3.2 implies that under its assumptions we have, for any $n\in\mathbb{N}$:

$$a_n(J_{\mathsf{W}^{1,p}(\mathbf{G},a;o),\mathsf{L}^p(\mathbf{G},\mu)})\leqslant\frac{||w_a||_{p'}\mu(\mathbf{G})^{1/p}}{n},\quad p<\infty, \tag{6.2}$$

$$a_n(J_{\mathsf{W}^{1,\infty}(\mathbf{G},a;o),\mathsf{L}^\infty(\mathbf{G},V)})\leqslant\frac{||w_a||_1||V||_\infty}{n}. \tag{6.3}$$

In the same way, it follows from Theorem 4.2 that

$$a_n(J_{\mathsf{B}^{\theta,p}(\mathbf{G},a;o),\mathsf{L}^p(\mathbf{G},\mu)})\leqslant C(\theta,p)|\mathbf{G}|^{\theta-1/p}\mu(\mathbf{G})^{1/p}n^{-\theta},\quad \forall n\in\mathbb{N},\ 1<p\theta<\infty.$$

6.4. Hardy-type operators on trees

For the case of trees there is a useful interpretation of estimates (6.2) and (6.3) in terms of approximation numbers of certain integral operators.

Let $\mathbf{T}$ be a compact metric tree on which a point o (the root) is selected. Below we use the notation $\langle x,y\rangle$ introduced is Section 5.1.

The Hardy-type integral operator with weights v, w on the rooted tree $\{\mathbf{T}, o\}$ is defined as

$$g(x) = (H_{v,w}f)(x) = (H_{v,w}(\mathbf{T}, o)f)(x) = v(x) \int_{\langle o,x \rangle} f(y)w(y)dy. \tag{6.4}$$

At first we assume that $w(x) \neq 0$ a.e. and set $a(x) = |w(x)|^{-p}$, then $w = w_a$, cf. (3.1). It is easy to see that the operator

$$Q_w : f(x) \mapsto u(x) = \int_{\langle o,x \rangle} f(y)w(y)dy$$

defines an isometry of the space $\mathsf{L}^p(\mathbf{T})$ onto $\mathsf{L}^{1,p}(\mathbf{T}, a; o)$. Besides, $||g||_p = ||Q_w f||_{p,V}$ where $V = |v|^p$. This shows that

$$a_n(H_{v,w}) = a_n(J_{\mathsf{W}^{1,p}(\mathbf{T},a;o),\mathsf{L}^p(\mathbf{T},V)}), \quad \forall n \in \mathbb{N}.$$

Now we are in a position to justify the following result.

Theorem 6.1. *Let* $\mathbf{T}$ *be a compact metric tree with the root* o *and let* $w \in \mathsf{L}^{p'}(\mathbf{T})$, $v \in \mathsf{L}^p(\mathbf{T})$ *where* $1 < p < \infty$. *Then the operator* $H_{v,w}$ *is compact in* $\mathsf{L}^p(\mathbf{T})$ *and its approximation numbers satisfy the estimate*

$$a_n(H_{v,w}) \leqslant \frac{||v||_p ||w||_{p'}}{n}, \quad \forall n \in \mathbb{N}. \tag{6.5}$$

Proof. If $w(x) \neq 0$ a.e., then (6.5) immediately follows from Theorem 3.2. The result extends to the general case by a standard approximation argument. □

6.5. Comparison with the results of [4]

The techniques of [4] is based upon a careful analysis of the function $A_{v,w}(T)$ of subtrees $T \in \mathscr{C}(\mathbf{T})$ which in the compact case can be defined as follows:

$$A_{v,w}(T) = \min_{o \in T} ||H_{v,w}(T, o) : \mathsf{L}^p(T) \to \mathsf{L}^p(T)||,$$

cf. Theorem 3.8 in [4]. Evidently,

$$A_{v,w}(T) \leqslant ||v||_{\mathsf{L}^p(T)} ||w||_{\mathsf{L}^{p'}(T)}.$$

Up to a change of notations, the expression on the right-hand side is exactly the function Φ appearing in the proof of Theorem 3.2. One may attempt to apply our analysis directly to the function $A_{v,w}(T)$. However, such an attempt fails, since this function is, in general, not super-additive. Note also that the converse inequality $A_{v,w}(T) \geqslant c||v||_{\mathsf{L}^p(T)} ||w||_{\mathsf{L}^{p'}(T)}$ with any $c > 0$ is impossible.

In terms of the function $A_{v,w}(T)$ the authors of [4] found for the approximation numbers $a_n = a_n(H_{v,w}(\mathbf{T}, o))$ some two-sided estimates, see Theorem 3.18 there. Based upon these estimates, they justified the Weyl-type asymptotics for a_n. As it was pointed out to the author by Evans, the inequality

$$a_{n+4} \leqslant 3n^{-1}||v||_p||w||_{p'}$$

which is only slightly rougher than (6.5), can be easily derived from the results of [4].

As we see it, the techniques developed in the present paper gives a direct and unified approach to the upper estimates of approximation numbers for embedding operators of Sobolev spaces on graphs. It cannot give any lower estimates. For integral operator (6.4) our new result consists in finding the upper estimate with the best-possible constant factor.

Some results can be obtained by combination of our both approaches. For example, the reasonings presented in Section 6.2 immediately lead to the following result.

Proposition 6.2. *Under the assumptions of Theorem* 3.2, *the following asymptotic formula for the approximation numbers* a_n *of the embedding operator of the space* $\mathsf{W}^{1,p}(\mathbf{G}, a; o)$ *into* $\mathsf{L}^p(\mathbf{G}, V)$:

$$\lim_{n\to\infty} na_n = \alpha_p \int_{\mathbf{G}} w_a(x)V(x)^{1/p}\,dx, \quad \alpha_p = A_{1,1}([0,1]).$$

Indeed, for trees this is nothing but a reformulation of Corollary 5.4 from [4]. Since the passage to a subspace of finite codimension does not affect the asymptotic behavior of approximation numbers, the desired result for general compact graphs follows.

Lemma 5.9 from [4], which deals with the cases $p = 1$ and ∞, extends to graphs in the same way.

6.6. *More on approximation of spaces* $\mathsf{B}^{\theta,p}(\mathbf{G})$

In Section 4 we used interpolation for the description of these spaces. It is natural to try using interpolation in a somewhat more systematic way, namely for deriving Theorem 4.2 from Theorem 1.2. Unfortunately, this idea does not work. Indeed, the construction of the operator P_n in Theorem 1.2 heavily relies upon continuity of the functions from $\mathsf{L}^{1,p}$. As a consequence, P_n is not well defined as an operator in L^p, and there is no base for interpolation.

It is also unclear whether the result of Theorem 4.2 can be extended to the general Besov spaces $\mathsf{B}_q^{\theta,p}$ with $q \neq p$, and also to the fractional Sobolev spaces $\mathsf{L}^{\theta,p} := [\mathsf{L}^p, \mathsf{L}^{1,p}]_\theta$. The obstacle is basically the same. Indeed, let $0 < \theta_0 < \theta < \theta_1 < 1$ and $p\theta_0 > 1$. Given a measure $\mu \in \mathfrak{M}(\mathbf{G})$, let P_n^0, P_n^1 be the operators P_n constructed in Theorem 4.2 for the spaces $\mathsf{B}^{\theta_0,p}, \mathsf{B}^{\theta_1,p}$, respectively. In general, $P_n^0 \neq P_n^1$, and again there is no base for interpolation.

Acknowledgments

The author expresses his thanks to Professor W.D. Evans for the fruitful discussions. The contents of this subsection is a reply to Referee's remark. I take this opportunity to thank the referee for the careful analysis of the paper.

References

[1] M.Sh. Birman, M. Solomyak, Piecewise polynomial approximations of functions of classes W_p^α, Mat. Sb. (N.S.) 73 (115) (1967) 331–355 (Russian). English transl., Math. USSR-Sb. 73 (115) (1967) 295–317.
[2] D.E. Edmunds, H. Triebel, Function Spaces, Entropy Numbers and Differential Operators, Cambridge Tracts in Mathematics, Vol. 120, Cambridge University Press, Cambridge, 1996.
[3] W.D. Evans, D.J. Harris, Fractals, trees and the Neumann Laplacian, Math. Ann. 296 (1993) 493–527.
[4] W.D. Evans, D.J. Harris, J. Lang, The approximation numbers of Hardy-type operators on trees, Proc. London Math. Soc. 83 (3) (2001) 390–418.
[5] K. Naimark, M. Solomyak, Eigenvalue estimates for the weighted Laplacian on metric trees, Proc. London Math. Soc. 80 (3) (2000) 690–724.
[6] A. Pinkus, *n*-Widths in Approximation Theory, Springer, Berlin, 1985.
[7] M. Solomyak, On the eigenvalue estimates for the weighted Laplacian on metric graphs, in: M.Sh. Birman, S. Hildebrandt, V.A. Solonnikov, N.N. Uraltseva (Eds.), Nonlinear Problems in Mathematical Physics and Related Topics I, In Honor of Professor O.A. Ladyzhenskaya, Kluwer Academic Publishers, Dordrecht, Plenum Press, New York, 2002, pp. 326–347.
[8] H. Triebel, Interpolation Theory, Function Spaces – Differential Operators, North-Holland Publishing Co., Amsterdam, New York, 1978.