Sampling and nonlinear approximation of band limited signals in mean oscillation spaces

Rodolfo H. Torres*

Department of Mathematics
University of Kansas
Lawrence, KS 66045, USA
E-mail: torres@math.ukans.edu

Abstract— This paper deals with the nonlinear approximation of band limited signals using their sampled values on an appropriate discrete set. Some spaces of sequences introduced by the author in previous work are used to measure the oscillation of such signals. These spaces can be viewed as discrete versions of Sobolev or, more generally, Beosv spaces. A discrete wavelet type characterization of signals based on their samples is described and used then to obtain optimal rates of convergence in nonlinear approximations. The work presented provides a rigorous justification for the use of some analytical tools from harmonic analysis and approximation theory in a completely discrete setting.

I. Introduction

The well-known Shannon's sampling theorem for band limited signal with finite energy (i.e., L^2 functions with compactly supported Fourier transform) has been extended to other classes of signals. See the surveys [1], [2], and [3], and the books [4] and [5] for references.

One of the most general version of the sampling theorem is in our work in [6], where we have shown that band limited signals in Besov spaces can be recovered from their samples. Moreover, along the lines of the classical Plancherel–Polya inequality (see [7] and [8]), we have shown that the Besov–norm of a band limited function can be compared to the norm of its sample in an appropriate space of sequences. We have also obtained a discrete wavelet characterization of such spaces and, more recently in [9], we have obtained a mean oscillation characterization. The purpose of this paper is to use all these results to study nonlinear approximations of band limited signals using their samples.

Nonlinear approximation of functions using wavelet type decompositions have been extensively studied. The initial results in [10] in certain Besov spaces have been extended and generalized in, for example, [11] and [12]. References to relevant previous works in nonlinear approximation can be found in [10]. Applications to image compression were developed in [13] and related issues in the context of statistical estimation were recently described in [14].

From an abstract and simplified point of view, the underlying common problem considered in all of the above works is the following. Let X be a normed space of functions and assume that every function f in X admits a wavelet decomposition

$$f = \sum_I a_I(f)\varphi_I,$$

where the functions φ_I are obtained by dyadic dilations and translations of a single function φ in the usual way, and where the numbers $a_I(f)$ are the wavelet coefficients. Given a positive integer N let X_N be the set of all functions of the form $g = \sum_I a_I\varphi_I$, where at most N of the numbers a_I are non-zero. Then, the problem consists in finding a g so that $\|f-g\|_X$ is minimal among all elements in X_N.

The above problem clearly makes sense for any family of functions that generates the space X, not just wavelets, and it has a trivial solution when X is a Hilbert space and the generating

*Supported in part by NSF grant DMS 9623251 and by the University of Kansas General Research Grant allocation 3559-20-0038.

family is an orthonormal basis. Namely, choose N terms from the orthonormal expansion of a given function with coefficients as large as possible. A remarkable thing about wavelets is that the answer is still almost the same for a great variety of function spaces. This depends not only on the fact that wavelets provide frames and even unconditional basis for most function spaces used in analysis, but also on the Littlewood–Paley type characterization of such function spaces in terms of the wavelets coefficients. It is by now well-known that wavelets characterize, Sobolev, Besov, and Triebel–Lizorkin spaces among others (see [15] and [16] for detailed descriptions) and the nonlinear approximation problem have been resolved in those spaces as well. Moreover, in the works mentioned before one can find a precise study of the asymptotic behavior of the error in the nonlinear approximation in terms of the number N and the "smoothness" of the function being approximated. This, in turn, also gives new characterizations of many function spaces. In this article we will carry out a similar analysis in a completely discrete setting. This is relevant because, in applications, a signal is usually given by a discrete set of values or samples. Our work justifies then the used of the analytical results in the discrete setting.

II. Function spaces

We recall some basic facts about Besov spaces and wavelet characterizations. The reader not familiar with Besov spaces can find further details in [17], [18], [19], [20], and [21].

We denote the Fourier transform of f by $\hat{f}(\xi) = \int f(x)e^{-ix\xi}dx$, and we denote the inverse Fourier transform by $\check{f}$. Let φ be a function in the Schwartz space $\mathcal{S}(\mathbf{R})$ satisfying the conditions supp $\hat{\varphi} \subset \{\xi : \pi/4 < |\xi| < \pi\}$ and $|\hat{\varphi}(\xi)| > C$ on $\{\xi : \pi/4 + \epsilon < |\xi| < \pi - \epsilon\}$, for some $C, \epsilon > 0$. For $\nu \in \mathbf{Z}$, let $\varphi_\nu(x) = 2^\nu \varphi(2^\nu x)$. For $\alpha \in \mathbf{R}$ and $0 < p, q < \infty$, the (homogeneous) Besov space $\dot{B}_p^{\alpha,q}(\mathbf{R})$ is the space of all $f \in \mathcal{S}'/\mathcal{P}(\mathbf{R})$ (tempered distributions modulo polynomials on $\mathbf{R}$) equipped with the norm

$$\|f\|_{\dot{B}_p^{\alpha,q}(\mathbf{R})} = \left(\sum_{\nu \in \mathbf{Z}} \left(2^{\nu\alpha} \|f * \varphi_\nu\|_{L^p(\mathbf{R})} \right)^q \right)^{1/q}. \tag{1}$$

The Littlewood–Paley characterization given by (1) is independent of the choice of φ and it has the advantage that it can be used for all possible values of α, p, and q. Although very useful in applications, (1) does not reflect the oscillatory properties of the functions in $\dot{B}_p^{\alpha,q}(\mathbf{R})$ as clearly as some other characterizations do. The Besov spaces can be characterized using differences and mean oscillations. We recall the particular case $p = q$ and $\alpha = 1/p$. Let $\tau_h f(x) = f(x-h)$, let

$$w_p(f,t) = \sup\{\|f - \tau_h f\|_{L^p},\ |h| \le t\},$$

$$I_{\nu k} = [k2^{-\nu}, (k+4)2^{-\nu}],$$

and

$$f_{I_{\nu k}} = \frac{1}{|I_{\nu k}|} \int_{I-\nu k} f(x)dx.$$

For $1 < p < \infty$, the norm of the space $\dot{B}_p^{1/p,p}(\mathbf{R})$ is equivalent to

$$\|f\|^p_{B_p(\mathbf{R})} = \sum_{\nu \in \mathbf{Z}} \sum_{k \in \mathbf{Z}} \left(\frac{1}{|I_{\nu k}|} \int_{I_{\nu k}} |f(x) - f_{I_{\nu k}}| dx \right)^p \tag{2}$$

and to

$$\|f\|_{B^p(\mathbf{R})} = \left(\int_0^\infty w_p(f,t)^p \frac{dt}{t^2} \right)^{1/p}. \tag{3}$$

We have defined the spaces $\dot{B}_p^{\alpha,q}(\mathbf{R})$ as spaces of distributions modulo polynomials of arbitrary order, but they can also be interpreted as spaces of polynomials of degree at most k, where $k = [\alpha - 1/p]$. In particular, the expressions in (2) and (3) are norms modulo constants. Although

$$\|f\|_{\dot{B}_p^{1/p,p}(\mathbf{R})} \le C\|f\|_{B_p(\mathbf{R})},$$

and, if $\|f\|_{B_p(\mathbf{R})}$ is finite, then

$$\|f\|_{B_p(\mathbf{R})} \le C\|f\|_{\dot{B}_p^{1/p,p}(\mathbf{R})},$$

it may happen that for a particular representative f of an equivalence class in $\dot{B}_p^{1/p,p}(\mathbf{R})$, the norm in (2) is not finite. There always exists, however, a polynomial P_f so that

$$\|f\|_{\dot{B}_p^{1/p,p}(\mathbf{R})} \approx \|f - P_f\|_{B_p(\mathbf{R})}. \tag{4}$$

This same interpretation has to be used when the norm in (3) is considered. See [22] for more details. (Here and in what follows, $\approx$ means that each quantity is bounded by a positive constant times the other, independently of f .)

We recall the following nonorthogonal decomposition of Besov spaces. Let φ be an even and real valued function as in the definition of the Besov spaces, and such that $\sum_\nu \hat{\varphi}^2(2^{-\nu}\xi) = 1$ for all $\xi \neq 0$. For ν, $k \in \mathbf{Z}$, let

$$\varphi_{\nu k}(x) = 2^{-\nu/2}\varphi_\nu(x - 2^{-\nu}k) = 2^{\nu/2}\varphi(2^\nu x - k).$$

Let $< \cdot, \cdot >$ denote the usual pairing between distributions and test functions. Then, every $f \in \dot{B}_p^{\alpha,q}(\mathbf{R})$ can be written in the form

$$f = \sum_{\nu\in\mathbf{Z}} \sum_{k\in\mathbf{Z}} < f, \varphi_{\nu k} > \varphi_{\nu k},$$

and

$$\|f\|^q_{\dot{B}_p^{\alpha,q}(\mathbf{R})} \approx \sum_{\nu\in\mathbf{Z}} \left(\sum_{k\in\mathbf{Z}} \left(2^{\nu(\alpha+1/2-1/p)} | < f, \varphi_{\nu k} > | \right)^p \right)^{q/p}.$$

Details, proof, and further references about the above facts can be found in [15] and [16].

Besov spaces on the set of integers were first defined in [6]. Let χ_A be the characteristic function of the set A. For a continuous function f on $\mathbf{R}$, its restriction to the integers is the sequence $Rf = \chi_{\mathbf{Z}} f = \{f(n)\}$. Let φ be a function as before and assume further that $\hat{\varphi} \equiv 1$ in a small neighborhood of $|\xi| = \pi/2$. The space $\dot{B}_p^{\alpha,q}(\mathbf{Z})$ are defined to be the collection of all sequences s (modulo the restrictions of polynomials to $\mathbf{Z}$) equipped with the norm

$$\|s\|_{\dot{B}_p^{\alpha,q}(\mathbf{Z})} = \left(\sum_{\nu\leq 1} \left(2^{\nu\alpha} \|s * \varphi_\nu^d\|_{L^p(\mathbf{Z})} \right)^q \right)^{1/q},$$

where $*$ is now convolution on the group of integers numbers, and $\varphi_\nu^d = R\varphi_\nu$, $\nu \leq 0$, and $\varphi_1^d = R((\chi_{[-\pi,\pi]}\hat{\varphi}_1)\check{})$.

The spaces $\dot{B}_p^{\alpha,q}(\mathbf{Z})$ can be decomposed in terms of discrete nonorthogonal wavelets. For $\nu \leq 0$ and $k \in \mathbf{Z}$, we define $\varphi_{\nu k}^d = R(\varphi_{\nu k})$, while for $\nu = 1$ and $k \in \mathbf{Z}$, $\varphi_{1k}^d = \tau_k \varphi_1^d$. For two sequences s and r let $< s, r >= \sum s(n) r(n)$. Then, every $s \in \dot{B}_p^{\alpha,q}(\mathbf{Z})$ can be written in the form

$$\sum_{\nu\leq 1} \sum_{k\in\mathbf{Z}} < s, \varphi_{\nu k}^d > \varphi_{\nu k}^d, \tag{5}$$

and

$$\|s\|^q_{\dot{B}_p^{\alpha,q}(\mathbf{Z})} \approx \sum_{\nu\leq 1} \left(\sum_{k\in\mathbf{Z}} \left(2^{\nu(\alpha+1/2-1/p)} | < s, \varphi_{\nu k}^d > | \right)^p \right)^{q/p}. \tag{6}$$

We refer again to [6].

For $1 < p < \infty$, the spaces $\dot{B}_p^{\alpha,q}(\mathbf{R})$ and $\dot{B}_p^{\alpha,q}(\mathbf{Z})$ are related as follows. Let E_π be the set of tempered distributions whose Fourier transforms are supported on the interval $[-\pi, \pi]$. The functions in $\dot{B}_p^{\alpha,q}(\mathbf{R})$ are, in general, only defined almost everywhere, but the functions in $\dot{B}_p^{\alpha,q}(\mathbf{R}) \cap E_\pi$ agree almost everywhere with analytic functions and, hence, we can make sense of their restrictions to the integers. In this sense, we can prove (see [6]) that $\dot{B}_p^{\alpha,q}(\mathbf{R}) \cap E_\pi \approx \dot{B}_p^{\alpha,q}(\mathbf{Z})$. The precise statement is as follows.

Theorem 2.1 *The space $\dot{B}_p^{\alpha,q}(\mathbf{Z})$ coincides with the space of restrictions to the integers of functions in $\dot{B}_p^{\alpha,q}(\mathbf{R}) \cap E_\pi$ and for all f in $\dot{B}_p^{\alpha,q}(\mathbf{R})$,*

$$\|f\|_{\dot{B}_p^{\alpha,q}(\mathbf{R})} \approx \|Rf\|_{\dot{B}_p^{\alpha,q}(\mathbf{Z})}. \tag{7}$$

We need to recall other spaces of sequences introduced in [23]; see also [24]. For $1 < p < \infty$ the space $B_p(\mathbf{Z})$ is the space of sequences $\{f(n)\}$ for which

$$\|f\|^p_{B_p(\mathbf{Z})} = \sum_{\nu\leq 0} \sum_{k\in\mathbf{Z}} \left(\frac{1}{|I_{\nu k}|} \sum_{I_{\nu k}} |f(n) - f^d_{I_{\nu k}}| \right)^p$$

is finite. Here, for an interval I, $\sum_I$ stands for $\sum_{n\in I\cap\mathbf{Z}}$, and f_I^d is the discrete average

$$f_I^d = |I|^{-1} \sum_I f(n).$$

In [9] we have proved that the spaces $\dot{B}_p^{1/p,p}(\mathbf{Z})$ and and $B_p(\mathbf{Z})$ are the same. Moreover the following results hold.

Theorem 2.2 *Let f be a function in $B_p(\mathbf{R}) \cap E_\pi$, $1 < p < \infty$. Then, its restriction to the integers, Rf, is a sequence in $B_p(\mathbf{Z})$ and*

$$\|f\|_{B_p(\mathbf{R})} \approx \|Rf\|_{B_p(\mathbf{Z})}.$$

Moreover, every sequence in $B_p(\mathbf{Z})$ is the restriction to the integers of a unique function (modulo polynomials) in $B_p(\mathbf{R}) \cap E_\pi$.

The above results can be viewed as oscillatory versions of the classical Plancherel–Pólya inequality.

III. Nonlinear approximation

In this section we include the statements of some of the results about nonlinear approximation of band limited signals that we are able to prove. We shall provide a complete description for a particular case. The most general case of the results can be easily inferred from the details described below.

By the results in the previous section, band limited signals in Besov spaces can be characterized in terms of their sequence of samples. These, in turn, admit a discrete wavelet decomposition. We want to obtain nonlinear approximation results using such decomposition. In brief, we want to show how to select a finite number of terms in the discrete wavelet expansion of the samples to obtain an approximation of a given signal with a prescribed bound on the error.

For a sequence $s = \{s(n)\}$, a positive integer N, and real numbers α, p, and q, let the error in the approximation of the sequence using at most N terms in its discrete wavelet expansion be

$$E_N(s) = \inf \|s - \sum_{\Gamma} < s, \varphi^d_{\nu k} > \varphi^d_{\nu k}\|_{\dot{B}^{\alpha,q}_p(\mathbf{Z})},$$

where the infimum is taken over all sets of indices Γ with $|\Gamma| \leq N$.

We can show that the error satisfies $E_N(s) = O(N^{-\delta})$, for $\delta > 0$, if s is in $\dot{B}^{\beta,r}_t(\mathbf{Z})$ for appropriate indices β, r and t. In fact, we obtain another characterization of the discrete Besov spaces using the error in the nonlinear approximation. For simplicity in this presentation, we state and prove the following version of such results.

Theorem 3.1 *Let s be a sequence in $\dot{B}^{\alpha,2}_2(\mathbf{Z})$. Select a set Γ_N with N pairs (ν, k) in such a way that $2^{\nu\alpha}| < s, \varphi_{\nu k} > |$ is as large as possible, and let*

$$\mathcal{E}_N(s) = s - \sum_{\Gamma_N} < s, \varphi^d_{\nu k} > \varphi^d_{\nu k}.$$

If $\beta > 0$ and $1/q = 1/2 + \beta$, then

$$\sum_{N \geq 0} \left(N^\beta \|\mathcal{E}_N(s)\|_{\dot{B}^{\alpha,2}_2(\mathbf{Z})} \right)^q \frac{1}{N} < \infty$$

for every s is in $\dot{B}^{\alpha+\beta,q}_q(\mathbf{Z})$.

We note that an appropriate converse to the theorem also holds and that one essentially has that

$$\|\mathcal{E}_N(s)\|_{\dot{B}^{\alpha,2}_2(\mathbf{Z})} \approx O(N^{-\beta}) \iff s \in \dot{B}^{\alpha+\beta,q}_q(\mathbf{Z}).$$

To prove Theorem 3.1, we can use the different characterizations of the discrete spaces to translate the problem into an equivalent nonlinear approximation problem in the spaces of wavelet coefficients. These are the spaces $\dot{b}^{\alpha,q}_p$ of doubly indexed sequence of complex numbers $S = \{s_{\nu k}\}$ with ν, k in $\mathbf{Z}$, $\nu \leq 1$, and norm given by the right hand side of (6). In particular, we have

$$\|S\|_{\dot{b}^{\alpha,2}_2} = \left(\sum_{\nu \leq 1} \sum_{k \in \mathbf{Z}} \left(2^{\nu\alpha}|s_{\nu k}|\right)^2 \right)^{1/2}.$$

Let S_N be obtained from $S = \{< s, \varphi^d_{\nu k} >\}$ by setting to zero the entries with (ν, k) not in Γ_N and let $\mathcal{E}_N(S) = S - S_N$. Also, let l_0 be the space of finite sequence indexed by (ν, k) and, for S_0 in l_0, let $\|S_0\|_{l_0}$ be the number of nonzero entries in S_0. Exploiting the fact that the definition of $\dot{b}^{\alpha,2}_2$ involves only the size of the coefficients we easily see that

$$E_N(S) = \inf \|S - S_0\|_{\dot{b}^{\alpha,2}_2} \approx \|\mathcal{E}_N(S)\|_{\dot{b}^{\alpha,2}_2},$$

where the infimum is taken over all S_0 with $\|S_0\|_{l_0} \leq N$. We can use as in [11] and [12] what have now become standard interpolation techniques in nonlinear approximation using wavelets. In fact, let $l^{2,r}$, $r = 1$ or $r = \infty$, be the Lorentz spaces of sequences (indexed by (ν, k)) with respect to the counting measure. For $S = \{s_{\nu k}\}$, we set $I^\alpha S = \{2^{\nu\alpha} s_{\nu k}\}$. Note that $I^\alpha S$ is in l^q if and only if S is in $\dot{b}^{\alpha+1/q-1/2,q}_q$. In addition, one can check that for an appropriate constant C,

$$C^{-1}\|S\|_{l^{2,\infty}} \leq \|I^{-\alpha}S\|_{\dot{b}^{\alpha,2}_2} \leq C\|S\|_{l^{2,1}}.$$

It follows that

$$C^{-1} \inf \|I^\alpha S - S_0\|_{l^{2,\infty}} \leq \|\mathcal{E}_N S\|_{\dot{b}^{\alpha,2}_2} \leq$$

$$\leq C \inf \|I^\alpha S - S_0\|_{l^{2,1}},$$

where the infimum is taken over all S_0 with $\|S_0\|_{l_0} \leq N$. By real interpolation for Lorentz spaces using the best approximation functional, we have that, for $r = 1$ or $r = \infty$,

$$\sum_{N \geq 0} \left(N^\beta \inf \|I^\alpha S - S_0\|_{l^{2,r}} \right)^q \frac{1}{N} < \infty,$$

if and only if $I^\alpha S$ is in l^q with $1/q = \beta + 1/2$ (see [19] for details). But, as noted before, $I^\alpha S$ is in l^q if and only if S in $\dot{b}_q^{\alpha+\beta,q}$, and the theorem follows.

Finally, using well-known duality and embedding results we obtain the following corollary about nonlinear approximation in L^2. The result states how "good" (measured in L^p–norm) a sequence has to be in order to be nonlinearly approximated using discrete wavelets at a prescribed rate of convergence. It can also be interpreted as a discrete version of the Sobolev embedding theorem.

Corollary 3.2 *If the sequence s satisfies*

$$\|s\|_{L^p} = \left(\sum_{n\in Z} |s(n)|^p \right)^{1/p} < \infty,$$

for $1/p = 1/2 + \beta$, then

$$\|\mathcal{E}(s)\|_{L^2} = O(N^{-\beta}).$$

We remark that the rate of convergence can be improved if one uses discrete versions of the Hardy spaces H^p, $p \le 1$, but we shall not discuss this here any further.

References

[1] A. Jerri, "The Shannon sampling theorem–Its various extensions and applications", *Proc. IEEE* 65 (1977), pp. 1565–1596.

[2] P. Butzer, "A survey of the Whittaker–Shannon sampling theorem and some of its extensions", *J. Math. Res. Expositions* 3 (1983), pp. 185–212.

[3] R. Higgins, "Five short stories about the cardinal series", *Bull. Amer. Math. Soc.* 12 (1985), pp. 45–89.

[4] A. Zayed, "Advances in Shannon's sampling theory", CRC Press, Boca Raton, 1993.

[5] J. Benedetto and M. Frazier, "Wavelets: Mathematics and applications", CRC Press, Boca Raton, 1994.

[6] R. Torres, "Spaces of sequences, sampling theorem and functions of exponential type", *Studia Math.* 100 (1991), pp. 51–74.

[7] M. Plancherel and G. Pólya, "Fonctions entières et intégrales de Fourier multiples", *Comment. Math. Helv.* 10 (1938), pp. 110–163.

[8] R. Boas, "Entire Functions", Academic Press, New York, 1954.

[9] R. Torres, "Mean oscillation of functions and the Paley–Wiener space", to appear in *J. Fourier Anal. and Appl.*

[10] R. DeVore, B. Jawerth, and V. Popov, "Compression of wavelet decompositions", *Amer. J. Math.* 114 (1992), pp. 737–785.

[11] B. Jawerth and M. Milman, "Wavelets and best approximation in Besov spaces", in *Interpolation spaces and related topics*, Bar-Ilan Univ., Ramat Gan, 1992.

[12] C. Hsiao, B. Jawerth, B. Lucier, and X. Yu, "Near optimal compression of orthonormal wavelet expansions", in *Wavelets: Mathematics and Applications*, J. Benedetto and M. Frazier Eds., pp. 425–446, CRC Press, Boca Raton, 1994.

[13] R. DeVore, B. Jawerth, and B. Lucier, "Image compression through wavelet transform coding", *IEEE Trans. Inform. Theory* 38 (1992), pp. 719–746.

[14] D. Donoho, "Unconditional bases are optimal bases for data compression and for statistical estimation", *Appl. Comput. Harmon. Anal.* 1 (1993), pp. 100–115.

[15] Y. Meyer, "Ondelettes et operateurs", Hermann Ed., Paris, 1990.

[16] M. Frazier, B. Jawerth, and G. Weiss, "Littlewood–Paley theory and the study of function spaces", CBMS Regional Conference Series in Mathematics 79, 1991.

[17] E. Stein, "Singular integrals and differentiability properties of functions", Princeton Univ. Press, New Jersey, 1970.

[18] J. Peetre, "New thoughts on Besov spaces", Duke University Math. Series 1, Durham, 1976.

[19] J. Bergh and J. Löfström, "Interpolation spaces", Springer–Verlag, Berlin, 1976.

[20] F. Ricci and M. Taibleson, "Mean oscillation Boundary values of harmonic functions in mixed norm spaces and their atomic structure", *Ann. Scuola Norm. Pisa,* Series IV 10 (1983), pp. 1–54.

[21] H. Triebel, "Theory of function spaces", *Monographs in Mathematics*, vol. 78, Birkhauser Verlag, Basel, 1983.

[22] J. Dorronsoro, "Mean oscillations and Besov spaces", *Canad. Math. Bull.* 28 (1985), pp. 474–488.

[23] R. Rochberg, "Toeplitz and Hankel operators on the Paley-Wiener space", *Integral Eq. and Oper. Th.* 10 (1987), pp. 187–235.

[24] V. Peller, "Wiener-Hopf operators on a finite interval and Schatten-Von Neumann classes", *Proc. Amer. Math. Soc.* (1988), pp. 479–486.