Hindawi
Journal of Function Spaces
Volume 2019, Article ID 1845491, 10 pages
https://doi.org/10.1155/2019/1845491

Research Article

The Role of Nonpolynomiality in Uniform Approximation by RBF Networks of Hankel Translates

Isabel Marrero

Departamento de Análisis Matemático, Universidad de La Laguna, Aptdo. 456, 38200 La Laguna, Tenerife, Spain

Correspondence should be addressed to Isabel Marrero; imarrero@ull.es

Received 22 September 2018; Revised 20 February 2019; Accepted 28 February 2019; Published 1 April 2019

Academic Editor: Yoshihiro Sawano

Given $\mu > -1/2$ and $c \in I =]0, \infty[$, let the space $\mathscr{C}_{\mu,c}$ (respectively, $\mathscr{C}_{\mu}$) consist of all those continuous functions u on $]0, c]$ (respectively, I) such that the limit $\lim_{z \longrightarrow 0+} z^{-\mu-1/2} u(z)$ exists and is finite; $\mathscr{C}_{\mu,c}$ is endowed with the uniform norm $\|u\|_{\mu,\infty,c} = \sup_{z \in [0,c]} |z^{-\mu-1/2} u(z)|$ $(u \in \mathscr{C}_{\mu,c})$. Assume $\phi \in \mathscr{C}_{\mu}$ defines an absolutely regular Hankel-transformable distribution. Then, the linear span of dilates and Hankel translates of ϕ is dense in $\mathscr{C}_{\mu,c}$ for all $c \in I$ if, and only if, $\phi \notin \pi_{\mu}$, where $\pi_{\mu} = \operatorname{span}\{t^{2n+\mu+1/2} : n \in \mathbb{Z}_+\}$.

Dedicated to Professor Fernando Pérez González on the occasion of his retirement

1. Introduction and Motivation

1.1. RBFNNs. The radial basis function (RBF) method is nowadays one of the primary tools for interpolating multidimensional scattered data. Its simple form and ability to accurately approximate an underlying function have made the method increasingly popular in several different types of applications, some of which include cartography, medical imaging, the numerical solution of partial differential equations, and neural networks (see, e.g., [1] and references therein).

Radial basis function neural networks (RBFNNs) as such were introduced in the 1980s by Broomhead and Lowe [2] and soon applied to problems of supervised learning such as regression, classification, and time series prediction [3, 4]. This type of network falls within the general class of nonlinear, single hidden layer feedforward neural networks. Given $d \in \mathbb{N}$, the family of RBFNNs consists of all those functions $v : \mathbb{R}^d \longrightarrow \mathbb{R}$ of the form

$$v(x) = \sum_{i=1}^{m} w_i K\left(\frac{x - z_i}{\sigma_i}\right) = \sum_{i=1}^{m} w_i g\left(\frac{\|x - z_i\|}{\sigma_i}\right), \quad (1)$$

where

(i) $m \in \mathbb{N}$ is the number of kernel nodes in the hidden layer

(ii) $(w_1, \ldots, w_m) \in \mathbb{R}^m$ is the vector of weights from the ith kernel node to the output nodes

(iii) $x \in \mathbb{R}^d$ is an input vector

(iv) K is a radially symmetric kernel function of a unit in the hidden layer

(v) $z_i \in \mathbb{R}^d$ and $\sigma_i \in \mathbb{R}$ are the centroid and smoothing factor (or width) of the ith kernel node ($i \in \mathbb{N}$, $1 \le i \le m$), respectively

(vi) $g : [0, \infty[\longrightarrow \mathbb{R}$ is the so-called activation function, which characterizes the kernel shape, often a Gaussian

The smoothing factors may be the same in all kernel nodes of a RBFNN or may vary across them. Park and Sandberg [5, 6] proved that under mild conditions on the kernel K (or the activation function g) both classes of RBFNNs (with either the same or varying smoothing factors across nodes) have the universal approximation property, meaning that they are dense in suitable spaces of continuous or integrable functions. Chen and Chen [7] considered RBFNNs with a continuous activation function ϕ in the hidden layer defining a tempered distribution in $\mathbb{R}$ and proved that the necessary and sufficient condition for such networks to uniformly approximate every continuous function on compacta is that ϕ is not an even polynomial. Nonpolynomiality is straightforwardly seen to be a necessary condition for these approximations and has been found necessary and sufficient for other types of networks to possess the universal approximation property as well, cf. [8–12]. In this paper we aim to extend the result in [7] to RBFNNs of Hankel translates. The precise meaning of this extension will be clarified in due course.

1.2. The Hankel Transformation and the Hankel Translation. Let $I =]0,\infty[$. The Hankel integral transformation is usually defined by

$$\left(h_\mu \varphi\right)(x) = \int_0^\infty \varphi(t)\,\mathcal{J}_\mu(xt)\,dt \quad (x \in I), \tag{2}$$

where $\mathcal{J}_\mu(z) = z^{1/2} J_\mu(z)$ $(z \in I)$ and J_μ denotes the Bessel function of the first kind and order $\mu \in \mathbb{R}$.

Aiming to obtain a distributional extension of h_μ, Zemanian introduced new spaces of test and generalized functions. The space $\mathcal{H}_\mu$ [13, 14] consists of all those smooth, complex-valued functions $\varphi = \varphi(x)$ $(x \in I)$ such that

$$\begin{aligned}\nu_{\mu,p}(\varphi) = \max_{0\le k\le p} \sup_{x\in I} \left|\left(1+x^2\right)^p \left(x^{-1}D\right)^k x^{-\mu-1/2}\varphi(x)\right| \\ < \infty \quad \left(p \in \mathbb{Z}_+\right).\end{aligned} \tag{3}$$

When topologized by the family of norms $\{\nu_{\mu,p}\}_{p\in\mathbb{Z}_+}$, $\mathcal{H}_\mu$ becomes a Fréchet space where h_μ is an automorphism provided that $\mu \ge -1/2$. Then the generalized Hankel transformation h'_μ, defined by transposition on the dual $\mathcal{H}'_\mu$ of $\mathcal{H}_\mu$, is an automorphism of $\mathcal{H}'_\mu$ when this latter space is endowed with either its weak* or its strong topology. For $\mu \in \mathbb{R}$ and $a \in I$, Zemanian [15] also introduced the space $\mathcal{B}_{\mu,a}$ of all those smooth functions $\varphi = \varphi(x)$ $(x \in I)$ such that $\varphi(x) = 0$ $(x > a)$ and

$$\begin{aligned}\delta_{\mu,n}(\varphi) = \max_{0\le k\le n} \sup_{x\in I} \left|\left(x^{-1}D\right)^k x^{-\mu-1/2}\varphi(x)\right| < \infty \\ \left(n \in \mathbb{Z}_+\right).\end{aligned} \tag{4}$$

Endowed with the topology generated by the family of seminorms $\{\delta_{\mu,k}\}_{k\in\mathbb{Z}_+}$, $\mathcal{B}_{\mu,a}$ becomes a Fréchet space. The strict inductive limit $\mathcal{B}_\mu$ of the family $\{\mathcal{B}_{\mu,a}\}_{a\in I}$ satisfies $\mathcal{B}_\mu \subset \mathcal{H}_\mu$, with continuous embedding. Since $\mathcal{B}_\mu$ is dense in $\mathcal{H}_\mu$, it turns out that $\mathcal{H}'_\mu$ can be regarded as a subspace of $\mathcal{B}'_\mu$, the dual of $\mathcal{B}_\mu$.

The study of the Hankel #-convolution in spaces of generalized functions was initiated by Sousa Pinto [16], only on compactly-supported distributions and for $\mu = 0$. In a series of papers [17–19], Betancor and the author investigated systematically the generalized #-convolution in wider spaces of distributions, allowing $\mu > -1/2$. In this context, the Hankel convolution $\varphi \# \phi \in \mathcal{H}_\mu$ of $\varphi, \phi \in \mathcal{H}_\mu$ is defined as the function

$$(\varphi \# \phi)(x) = \int_0^\infty \varphi(y)\,(\tau_x\phi)(y)\,dy \quad (x \in I), \tag{5}$$

where the Hankel translate $\tau_x\phi \in \mathcal{H}_\mu$ of $\phi \in \mathcal{H}_\mu$ is given by

$$(\tau_x\phi)(y) = \int_0^\infty \phi(z)\,D_\mu(x,y,z)\,dz \quad (x, y \in I). \tag{6}$$

Here, for $x, y, z \in I$,

$$\begin{aligned} D_\mu(x,y,z) &= \int_0^\infty t^{-\mu-1/2}\mathcal{J}_\mu(xt)\,\mathcal{J}_\mu(yt)\,\mathcal{J}_\mu(zt)\,dt \\ &= \begin{cases} \dfrac{\left[z^2-(x-y)^2\right]^{\mu-1/2}\left[(x+y)^2-z^2\right]^{\mu-1/2}}{2^{3\mu-1}\pi^{1/2}\Gamma(\mu+1/2)\,(xyz)^{\mu-1/2}}, & |x-y| < z < x+y \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{7}$$

is the so-called Delsarte kernel. Note that $D_\mu(x,y,z) \ge 0$, $\operatorname{supp} D_\mu(x,y,\cdot) = [|x-y|, x+y]$, $D_\mu(x,y,z)$ is symmetric in x, y, z, and

$$\int_0^\infty D_\mu(x,y,z)\,z^{\mu+1/2}dz = c_\mu^{-1}(xy)^{\mu+1/2}, \tag{8}$$

where $c_\mu = 2^\mu\Gamma(\mu+1)$. Therefore, for any $\phi \in \mathcal{H}_\mu$ we have

$$(\tau_x\phi)(y) = \left(\tau_y\phi\right)(x) \quad (x, y \in I). \tag{9}$$

The formula

$$h_\mu\left(\tau_y\phi\right)(x) = x^{-\mu-1/2}\mathcal{J}_\mu(xy)\left(h_\mu\phi\right)(x) \quad (x, y \in I) \tag{10}$$

and the exchange formula

$$h_\mu(\varphi\#\phi)(x) = x^{-\mu-1/2}\left(h_\mu\varphi\right)(x)\left(h_\mu\phi\right)(x) \quad (x \in I) \tag{11}$$

hold pointwise. The Hankel translation is defined on $\mathcal{H}'_\mu$ by transposition. The Hankel convolution $f\#\varphi \in \mathcal{H}'_\mu$ of $f \in \mathcal{H}'_\mu$ and $\varphi \in \mathcal{H}_\mu$ is defined [19, Definition 3.1] by

$$(f\#\varphi)(x) = \langle f, \tau_x\varphi\rangle \quad (x \in I). \tag{12}$$

The formulas

$$h'_\mu\left(\tau_y f\right)(x) = x^{-\mu-1/2}\mathcal{J}_\mu(xy)\left(h'_\mu f\right)(x) \quad (y \in I) \tag{13}$$

and

$$h'_\mu(f\#\varphi)(x) = x^{-\mu-1/2}(h_\mu\varphi)(x)(h'_\mu f)(x) \quad (14)$$

hold in the sense of equality in $\mathscr{H}'_\mu$ (cf. [19, Proposition 3.5]).

The space $\mathcal{O}$ of all those smooth functions $\theta = \theta(x)$ ($x \in I$) with the property that to every $k \in \mathbb{Z}_+$ there corresponds $n_k \in \mathbb{Z}_+$ satisfying

$$\sup_{x\in I}\left|(1+x^2)^{-n_k}(x^{-1}D)^k\theta(x)\right| < \infty \quad (15)$$

was characterized as the space of multipliers of $\mathscr{H}_\mu$ and $\mathscr{H}'_\mu$ [20, Theorems 2.3 and 2.9].

The generalized Hankel transformation h'_μ establishes an isomorphism between $x^{\mu+1/2}\mathcal{O}$ and the subspace $\mathcal{O}'_{\mu,\#}$ of $\mathscr{H}'_\mu$ consisting of the Hankel convolution operators on $\mathscr{H}_\mu$ and $\mathscr{H}'_\mu$, which is a homeomorphism under the natural topologies of $x^{\mu+1/2}\mathcal{O}$ and $\mathcal{O}'_{\mu,\#}$ [19, Propositions 4.2 and 5.2]. The distribution given by

$$\langle\delta_\mu, \varphi\rangle = c_\mu \lim_{x\to 0+} x^{-\mu-1/2}\varphi(x) \quad (\varphi \in \mathscr{H}_\mu) \quad (16)$$

satisfies $\delta_\mu \in \mathcal{O}'_{\mu,\#}$ and $f\#\delta_\mu = \delta_\mu\# f = f$ ($f \in \mathscr{H}'_\mu$), cf. [19, Proposition 4.7] and [21, Proposition 3].

For the operational rules of the Hankel transformation and further properties of the Hankel translation and Hankel convolution that will be required, in particular those involving the Bessel differential operator

$$\begin{aligned} S_\mu = S_{\mu,x} &= x^{-\mu-1/2}D_x x^{2\mu+1}D_x x^{-\mu-1/2} \\ &= D_x^2 - \frac{4\mu^2-1}{4x^2}, \end{aligned} \quad (17)$$

the reader is mainly referred to [14, 17, 19]. Here we will highlight the following [14, Equation 5.5(8)]:

$$h'_\mu(S_\mu f)(x) = -x^2(h'_\mu f)(x) \quad (f \in \mathscr{H}'_\mu). \quad (18)$$

If $d \in \mathbb{N}$ and $\phi(x) = \phi_0(\|x\|)$ (a.e. $x \in \mathbb{R}^d$) is an integrable radial function, then its d-dimensional Fourier transform

$$\widehat{\phi}(\xi) = (2\pi)^{-d/2}\int_{\mathbb{R}^d}\phi(x)e^{-ix\cdot\xi}dx \quad (\xi \in \mathbb{R}^d) \quad (19)$$

is also radial and becomes a 1-dimensional Hankel transform of order $d/2-1$ [22, Theorem IV.3.3]:

$$\begin{aligned} \widehat{\phi}(\xi) &= \Phi_0(\|\xi\|) \\ &= \|\xi\|^{-(d-1)/2}\int_0^\infty \phi_0(s)s^{(d-1)/2}\mathscr{J}_{d/2-1}(\|\xi\|s)\,ds \\ &\qquad (\xi \in \mathbb{R}^d). \end{aligned} \quad (20)$$

Actually, since

$$\mathscr{J}_{-1/2}(z) = \left(\frac{2}{\pi}\right)^{1/2}\cos z \quad (z \in I), \quad (21)$$

it turns out that, on radial univariate—even—functions, the Fourier transformation, which reduces to a Fourier-cosine transformation, coincides with the Hankel transform of order $\mu = -1/2$; similarly, the Hankel translation and Hankel convolution of order $\mu = -1/2$ can be seen to coincide (modulo a multiplicative constant) with the usual translation and convolution on $\mathbb{R}$ (cf. [23, Example 3.2]). Thus for $2\mu+2 \notin \mathbb{N}$ the Hankel translation and the Hankel convolution provide strict generalizations of the usual translation and convolution operators, inasmuch as arbitrary orders $\mu \geq -1/2$ are allowed.

1.3. RBFNNs of Hankel Translates. Motivated by the fact that the Hankel transformation is best adapted to deal with radial functions, Arteaga and the author [24–27] have proved that the Hankel transformation and the Hankel convolution are suitable tools for the description and analysis of a RBF interpolation scheme by functions of the form

$$u(x) = \sum_{i=1}^n \alpha_i(\tau_{a_i}\phi)(x) + \sum_{j=0}^{m-1}\beta_j p_{\mu,j}(x) \quad (x \in I), \quad (22)$$

where $\mu \geq -1/2$, ϕ is a complex function defined on I (the so-called basis function), $p_{\mu,j}(x) = x^{2j+\mu+1/2}$ ($j \in \mathbb{Z}_+$, $0 \leq j \leq m-1$) is a Müntz monomial, $\tau_z = \tau_{\mu,z}$ ($z \in I$) denotes the Hankel translation operator of order μ, and α_i, β_j ($i, j \in \mathbb{Z}_+$, $1 \leq i \leq n$, $0 \leq j \leq m-1$) are complex coefficients.

In analogy to the standard case (1), we set the family $\mathcal{S}_1(\phi) = \mathcal{S}_{\mu,1}(\phi)$ of RBFNNs of Hankel translates of order $\mu > -1/2$ to consist of all those functions $v : I \longrightarrow \mathbb{R}$ which can be represented as

$$v(x) = \sum_{i=1}^m w_i\tau_{z_i}(\lambda_{\sigma_i}\phi)(x) \quad (x \in I), \quad (23)$$

where $m \in \mathbb{N}$ is the number of kernel nodes in the hidden layer, for $i \in \mathbb{N}$, $1 \leq i \leq m$, $w_i \in \mathbb{R}$ is the weight from the ith kernel node to the output node, and $z_i, \sigma_i \in I$ are, respectively, the centroid and the smoothing factor of the ith kernel node. Further, ϕ is a kernel function of a unit in the hidden layer which, in this case, coincides with the activation function and, as above, τ_z ($z \in I$) denotes the Hankel translation operator, while $(\lambda_r\phi)(t) = \phi(rt)$ ($r, t \in I$) is a dilation operator. Note that, for $d = 1$ and $\mu = -1/2$, (23) becomes (1).

An investigation on the universal approximation capabilities of a closely related class of RBFNNs defined on the nonnegative real axis has been carried out in several papers by Arteaga and the author [28–30]. It should be remarked that the results in the present paper can be derived neither from [24–27], where only the interpolation problem is addressed, nor from [28–30], where RBFNNs are constructed using the Bessel-Kingman hypergroup translation (or Delsarte translation) instead of the Hankel one, and where the universal approximation property, which is studied mainly in spaces of integrable functions, requires in turn integrability of the basis function.

1.4. Objectives. In the sequel we assume $\mu > -1/2$ and consider the following spaces:

(i) Given $c \in I$, $\mathscr{C}_{\mu,c}$ will denote the linear space of all those continuous functions u on $[0,c]$ such that the limit

$$\lim_{z\longrightarrow 0+} z^{-\mu-1/2}u(z) \tag{24}$$

exists and is finite. When endowed with the norm

$$\|u\|_{\mu,\infty,c} = \sup_{z\in[0,c]} \left|z^{-\mu-1/2}u(z)\right| \quad \left(u \in \mathscr{C}_{\mu,c}\right), \tag{25}$$

$\mathscr{C}_{\mu,c}$ becomes a Banach space. In fact, the map

$$u \longmapsto z^{-\mu-1/2}u(z) \tag{26}$$

is an isometry from $\mathscr{C}_{\mu,c}$ onto $C[0,c]$, the space of all continuous functions on $[0,c]$ with the uniform norm.

(ii) The linear space $\mathscr{C}_{\mu}$ consists of all those continuous functions u on I such that the limit (24) exists and is finite. Endowed with the topology generated by the family of seminorms $\{\rho_{\mu,n}\}_{n\in\mathbb{N}}$, where

$$\rho_{\mu,n}(u) = \sup_{z\in[0,n]} \left|z^{-\mu-1/2}u(z)\right| < \infty \qquad \left(u \in \mathscr{C}_{\mu},\ n \in \mathbb{N}\right), \tag{27}$$

$\mathscr{C}_{\mu}$ becomes a Fréchet space. Note that sequential convergence in $\mathscr{C}_{\mu}$ is equivalent to convergence in $\mathscr{C}_{\mu,c}$ for all $c \in I$.

(iii) The space $\mathscr{E}_{\mu}$ consists of all those smooth functions u on I such that the limits

$$\lim_{z\longrightarrow 0+} \left(z^{-1}D\right)^m z^{-\mu-1/2}u(z) \quad (m \in \mathbb{Z}_+) \tag{28}$$

exist and are finite. Endowed with the topology generated by the family of seminorms $\{\kappa_{\mu,n,m}\}_{(n,m)\in\mathbb{N}\times\mathbb{Z}_+}$, where

$$\kappa_{\mu,n,m}(u) = \sup_{z\in[0,n]} \left|\left(z^{-1}D\right)^m z^{-\mu-1/2}u(z)\right| < \infty \qquad \left(u \in \mathscr{E}_{\mu},\ n \in \mathbb{N},\ m \in \mathbb{Z}_+\right), \tag{29}$$

$\mathscr{E}_{\mu}$ becomes a Fréchet space.

Our aim here is to find necessary and sufficient conditions on the basis function ϕ for the family of RBFNNs $\mathscr{S}_1(\phi)$ to have the universal approximation property. More precisely, the above mentioned result in [7] is extended to the Hankel setting in the following way. Given $\phi \in \mathscr{C}_{\mu} \cap \mathscr{H}'_{\mu}$, a necessary and sufficient condition for $\mathscr{S}_1(\phi)$ to be dense in $\mathscr{C}_{\mu,c}$ $(c \in I)$ is nonmembership in the class

$$\pi_{\mu} = \operatorname{span}\left\{t^{2n+\mu+1/2} : n \in \mathbb{Z}_+\right\} \tag{30}$$

of Müntz polynomials generated by $p_{\mu,n}(t) = t^{2n+\mu+1/2}$ $(n \in \mathbb{Z}_+)$. This is the content of Theorem 9 in Section 3. In Section 2 we introduce the concept and give a characterization of zero-supported $\mathscr{B}'_{\mu}$-distributions (Theorem 5), which is used in the proof of Theorem 9 and might be interesting in its own right.

2. Zero-Supported Hankel Distributions

Definition 1. Suppose $\Lambda \in \mathscr{B}'_{\mu}$. We say that the support of Λ is $\{0\}$, in symbols $\operatorname{supp}\Lambda = \{0\}$, if $\Lambda\varphi = 0$ for all $\varphi \in \mathscr{B}_{\mu}$ with $\operatorname{supp}[x^{-\mu-1/2}\varphi(x)] \subset I$, equivalently, if $\Lambda\varphi = 0$ for all $\varphi \in \mathscr{B}_{\mu}$ such that $\varphi(x) = 0$ $(x \in]0,a[)$, for some $a \in I$.

Proposition 2. *Assume* $\operatorname{supp}\Lambda = \{0\}$ *and* $f \in \mathscr{E}_{\mu}$ *satisfies* $x^{-\mu-1/2}f(x) = 1$ $(x \in]0,a[)$, *for some* $a \in I$. *Then* $[x^{-\mu-1/2}f(x)]\Lambda = \Lambda$.

Proof. Let $\varphi \in \mathscr{B}_{\mu}$. Since $\operatorname{supp} x^{-\mu-1/2}[f\varphi - x^{\mu+1/2}\varphi(x)] \subset I$, necessarily

$$\left\langle \Lambda, f\varphi - x^{\mu+1/2}\varphi(x)\right\rangle = 0. \tag{31}$$

Therefore,

$$\begin{aligned}\langle f\Lambda, \varphi\rangle &= \langle \Lambda, f\varphi\rangle = \left\langle \Lambda, x^{\mu+1/2}\varphi(x)\right\rangle \\ &= \left\langle x^{\mu+1/2}\Lambda, \varphi\right\rangle.\end{aligned} \tag{32}$$

This yields the desired conclusion. □

Recall that $\Lambda \in \mathscr{B}'_{\mu}$ if, and only if, the restrictions of Λ to every $\mathscr{B}_{\mu,a}$ $(a \in I)$ are continuous. By (4), this means that to each $a \in I$ there corresponds $C > 0$ and $N \in \mathbb{Z}_+$ such that

$$\left|\langle \Lambda, \varphi\rangle\right| \le C\delta_{\mu,N}(\varphi) \quad \left(\varphi \in \mathscr{B}_{\mu,a}\right). \tag{33}$$

Definition 3. If, in (33), one N will do for all $a \in I$ (not necessarily with the same C), then the smallest such N is called the order of Λ. Otherwise, Λ is said to have infinite order.

Remark 4. Note that every $\Lambda \in \mathscr{B}'_{\mu}$ with $\operatorname{supp}\Lambda = \{0\}$ has finite order. Indeed, fix $r, b \in I$ with $r < b$, and choose $\psi \in \mathscr{B}_{\mu,b}$ such that $x^{-\mu-1/2}\psi(x) = 1$ $(x \in]0,r[)$. By Proposition 2, $[x^{-\mu-1/2}\psi(x)]\Lambda = \Lambda$. Now (33) yields $c_1 > 0$ and $N \in \mathbb{Z}_+$ satisfying

$$\left|\langle \Lambda, \varphi\rangle\right| \le c_1\delta_{\mu,N}(\varphi) \quad \left(\varphi \in \mathscr{B}_{\mu,b}\right). \tag{34}$$

On the other hand, the Leibniz formula gives $c_2 > 0$ such that

$$\delta_{\mu,N}\left[x^{-\mu-1/2}\psi(x)\varphi(x)\right] \le c_2\delta_{\mu,N}(\varphi) \quad \left(\varphi \in \mathscr{B}_{\mu}\right). \tag{35}$$

Thus

$$\begin{aligned}\left|\langle \Lambda, \varphi\rangle\right| &= \left|\left\langle x^{-\mu-1/2}\psi(x)\Lambda, \varphi\right\rangle\right| \\ &= \left|\left\langle \Lambda, x^{-\mu-1/2}\psi(x)\varphi(x)\right\rangle\right| \\ &\le c_1\delta_{\mu,N}\left[x^{-\mu-1/2}\psi(x)\varphi(x)\right] \le c_1c_2\delta_{\mu,N}(\varphi) \\ &\qquad \left(\varphi \in \mathscr{B}_{\mu}\right),\end{aligned} \tag{36}$$

as asserted.

Theorem 5. *Assume* $\Lambda \in \mathscr{B}'_\mu$, $\operatorname{supp}\Lambda = \{0\}$, *and* Λ *has order* N. *Then there are constants* α_j $(j \in \mathbb{Z}_+,\ 0 \le j \le N)$ *such that*

$$\Lambda = \sum_{j=0}^{N} \alpha_j \Lambda_j. \tag{37}$$

Here Λ_j *is the functional defined by*

$$\langle \Lambda_j, \varphi \rangle = (-1)^j c_{\mu+j} \lim_{x\longrightarrow 0+} \left(x^{-1}D\right)^j x^{-\mu-1/2}\varphi(x) \qquad \left(\varphi \in \mathscr{B}_\mu\right), \tag{38}$$

with $c_{\mu+j} = 2^{\mu+j}\Gamma(\mu+j+1)$ $(j \in \mathbb{Z}_+,\ 0 \le j \le N)$ *(cf. (8)). Conversely, every distribution of this form has* $\{0\}$ *for its support, unless* $\alpha_j = 0$ $(j \in \mathbb{Z}_+,\ 0 \le j \le N)$.

Proof. It is clear that $\operatorname{supp}\Lambda_j = \{0\}$ $(j \in \mathbb{Z}_+)$. This establishes the converse.

To prove the nontrivial half of the theorem, consider a $\varphi \in \mathscr{B}_\mu$ that satisfies

$$\langle \Lambda_j, \varphi \rangle = 0 \quad (j \in \mathbb{Z}_+,\ 0 \le j \le N). \tag{39}$$

Our objective is to prove that $\Lambda\varphi = 0$. Since

$$\begin{aligned}\langle \Lambda_N, \varphi \rangle &= (-1)^N c_{\mu+N} \lim_{x\longrightarrow 0+} \left(x^{-1}D\right)^N x^{-\mu-1/2}\varphi(x) \\ &= 0,\end{aligned} \tag{40}$$

given $\varepsilon > 0$, there exists $0 < \delta < 1$ such that $0 < x < \delta$ implies $|(x^{-1}D)^N x^{-\mu-1/2}\varphi(x)| < \varepsilon$. The mean value theorem yields $\xi_1 \in]0, x[$ such that

$$\begin{aligned}&\left|\left(x^{-1}D\right)^{N-1} x^{-\mu-1/2}\varphi(x)\right| \\ &= \left|\left[D\left(t^{-1}D\right)^{N-1} t^{-\mu-1/2}\varphi(t)\right]_{t=\xi_1}\right| x \le \varepsilon x^2 \\ &\qquad (0 < x < \delta).\end{aligned} \tag{41}$$

If $i \in \mathbb{Z}_+$, $0 \le i \le N$, an induction process then shows that, for some $\xi_i \in]0, x[$,

$$\begin{aligned}&\left|\left(x^{-1}D\right)^{N-i} x^{-\mu-1/2}\varphi(x)\right| \\ &= \left|\left[D\left(t^{-1}D\right)^{N-i} t^{-\mu-1/2}\varphi(t)\right]_{t=\xi_i}\right| x \le \varepsilon x^{2i} \\ &\qquad (0 < x < \delta).\end{aligned} \tag{42}$$

Choose $\psi \in \mathscr{B}_{\mu,1}$ such that $x^{-\mu-1/2}\psi(x) = 1$ $(x \in]0, a[)$ for some $0 < a < 1$, and define

$$\psi_r(x) = \left(\frac{x}{r}\right)^{-\mu-1/2} \psi\left(\frac{x}{r}\right) \quad (x \in I,\ r > 0). \tag{43}$$

Fix $0 < r < \delta$ and $n \in \mathbb{Z}_+$. By the Leibniz formula,

$$\begin{aligned}&\left(x^{-1}D\right)^n x^{-\mu-1/2}\left(\psi_r\varphi\right)(x) \\ &= \sum_{k=0}^{n} \binom{n}{k} \left[\left(x^{-1}D\right)^{n-k} \psi_r(x)\right] \\ &\quad \cdot \left[\left(x^{-1}D\right)^k x^{-\mu-1/2}\varphi(x)\right] \quad (x \in I).\end{aligned} \tag{44}$$

On the other hand [14, Equation 5.2(6)],

$$\begin{aligned}&\left(x^{-1}D\right)^{n-k} \psi_r(x) = r^{\mu+1/2}\left(x^{-1}D\right)^{n-k} x^{-\mu-1/2}\psi\left(\frac{x}{r}\right) \\ &= r^{\mu+1/2} x^{-2(n-k)-\mu-1/2} \sum_{j=0}^{n-k} b_{k,j} x^j D^j \left[\psi\left(\frac{x}{r}\right)\right] \\ &= r^{-2(n-k)} \left(\frac{x}{r}\right)^{-2(n-k)-\mu-1/2} \sum_{j=0}^{n-k} b_{k,j} \left(\frac{x}{r}\right)^j \left(D^j\psi\right)\left(\frac{x}{r}\right) \\ &= r^{-2(n-k)} \left[\left(t^{-1}D\right)^{n-k} t^{-\mu-1/2}\psi(t)\right]_{t=x/r} \\ &\qquad (x \in I,\ 0 \le k \le n)\end{aligned} \tag{45}$$

for suitable $b_{k,j} \in \mathbb{R}$ $(k, j \in \mathbb{Z}_+,\ 0 \le j \le n-k)$, with $b_{k,n-k} = 1$. Consequently,

$$\begin{aligned}&\max_{0\le n\le N} \sup_{x\in I} \left|\left(x^{-1}D\right)^n x^{-\mu-1/2}\left(\psi_r\varphi\right)(x)\right| \\ &\le \max_{0\le n\le N} \sum_{k=0}^{n} \binom{n}{k} \sup_{x\in]0,r[} \left|\left(x^{-1}D\right)^{n-k} \psi_r(x)\right| \\ &\quad \cdot \sup_{x\in]0,r[} \left|\left(x^{-1}D\right)^k x^{-\mu-1/2}\varphi(x)\right| \\ &\le \varepsilon 2^N \max_{0\le n\le N} \sup_{x\in I} \left|\left(x^{-1}D\right)^n x^{-\mu-1/2}\psi(x)\right| \\ &\quad \cdot \max_{0\le k\le N} \sup_{x\in]0,r[} \left(\frac{x}{r}\right)^{2(N-k)} \\ &\le \varepsilon 2^N \max_{0\le n\le N} \sup_{x\in I} \left|\left(x^{-1}D\right)^n x^{-\mu-1/2}\psi(x)\right|.\end{aligned} \tag{46}$$

Since Λ has order N, there is a constant $C > 0$ such that

$$\left|\langle \Lambda, \phi \rangle\right| \le C \max_{0\le n\le N} \sup_{x\in I} \left|\left(x^{-1}D\right)^n x^{-\mu-1/2}\phi(x)\right| \qquad \left(\phi \in \mathscr{B}_{\mu,1}\right). \tag{47}$$

And since $\psi_r(x) = 1$ $(x \in]0, ar[)$, from Proposition 2 we infer

$$\begin{aligned}&\left|\langle \Lambda, \varphi \rangle\right| = \left|\langle \psi_r\Lambda, \varphi \rangle\right| = \left|\langle \Lambda, \psi_r\varphi \rangle\right| \\ &\le C \max_{0\le n\le N} \sup_{x\in I} \left|\left(x^{-1}D\right)^n x^{-\mu-1/2}\left(\psi_r\varphi\right)(x)\right| \\ &\le \varepsilon C 2^N \max_{0\le n\le N} \sup_{x\in I} \left|\left(x^{-1}D\right)^n x^{-\mu-1/2}\psi(x)\right|.\end{aligned} \tag{48}$$

The arbitrariness of $\varepsilon > 0$ shows that $\langle \Lambda, \varphi \rangle = 0$. Hence Λ vanishes on the intersection of the null spaces of the functionals Λ_j ($j \in \mathbb{Z}_+$, $0 \leq j \leq N$), and the desired representation follows from [31, Lemma 3.9]. □

Remark 6. Note that the functionals (38) can be written (modulo constant factors) as derivatives of the identity for the Hankel convolution (16). In fact, we have

$$\begin{aligned}\left\langle \Lambda_j, \varphi \right\rangle &= (-1)^j c_{\mu+j} \lim_{x \longrightarrow 0+} \left(x^{-1} D\right)^j x^{-\mu-1/2} \varphi(x) \\ &= (-1)^j c_\mu^{-1} c_{\mu+j} \left\langle \delta_\mu, x^{\mu+1/2} \left(x^{-1} D\right)^j x^{-\mu-1/2} \varphi(x) \right\rangle \\ &= \left\langle c_\mu^{-1} c_{\mu+j} x^{-\mu-1/2} \left(D x^{-1}\right)^j x^{\mu+1/2} \delta_\mu, \varphi \right\rangle \\ &\qquad \left(\varphi \in \mathscr{B}_\mu,\ j \in \mathbb{Z}_+\right).\end{aligned} \tag{49}$$

3. Nonpolynomiality of the Activation Function

We begin by establishing two auxiliary results.

Lemma 7. *Let $r, a \in I$.*

(i) *The dilation operator $\lambda_r : \mathscr{C}_\mu \longrightarrow \mathscr{C}_\mu$, defined by $(\lambda_r \phi)(x) = \phi(rx)$ $(x \in I)$, is continuous.*

(ii) *The translation operator $\tau_a : \mathscr{C}_\mu \longrightarrow \mathscr{C}_\mu$, defined by*

$$\left(\tau_a \phi\right)(x) = \int_0^\infty D_\mu(x, a, z)\, \phi(z)\, dz \quad (x \in I), \tag{50}$$

is continuous.

(iii) *If $\phi \in \mathscr{C}_\mu$, then $\mathscr{S}_1(\phi) \subset \mathscr{C}_\mu$.*

Proof. Given $\phi \in \mathscr{C}_\mu$, it is apparent that $\lambda_r \phi \in \mathscr{C}_\mu$; indeed, the function $(\lambda_r \phi)(x) = \phi(rx)$ $(x \in I)$ is clearly continuous, and the limit

$$\begin{aligned}&\lim_{x \longrightarrow 0+} x^{-\mu-1/2} \left(\lambda_r \phi\right)(x) \\ &\quad = r^{\mu+1/2} \lim_{x \longrightarrow 0+} (rx)^{-\mu-1/2} \phi(rx) \\ &\quad = r^{\mu+1/2} \lim_{z \longrightarrow 0+} z^{-\mu-1/2} \phi(z)\end{aligned} \tag{51}$$

exists and is finite. Similarly,

$$\begin{aligned}\left|\left(\lambda_r \phi\right)(x)\right| = \left|\phi(rx)\right| = (rx)^{\mu+1/2} \left|(rx)^{-\mu-1/2} \phi(rx)\right| \\ (x \in I);\end{aligned} \tag{52}$$

therefore, for any $c \in I$,

$$\begin{aligned}&\sup_{x \in [0,c]} \left|x^{-\mu-1/2} \left(\lambda_r \phi\right)(x)\right| \\ &\quad = r^{\mu+1/2} \sup_{x \in [0,c]} \left|(rx)^{-\mu-1/2} \phi(rx)\right| \\ &\quad = r^{\mu+1/2} \sup_{z \in [0,rc]} \left|z^{-\mu-1/2} \phi(z)\right| \quad \left(\phi \in \mathscr{C}_\mu\right).\end{aligned} \tag{53}$$

Equation (53) proves continuity of λ_r and establishes (i).

To prove (ii), first of all we pick $\phi \in \mathscr{C}_\mu$ and show that $\tau_a \phi$ is well defined. In fact, using (8) we may write

$$\begin{aligned}\left|\left(\tau_a \phi\right)(x)\right| &\leq \int_0^{x+a} D_\mu(x, a, z) \left|\phi(z)\right| dz \\ &\leq \sup_{z \in [0,x+a]} \left|z^{-\mu-1/2} \phi(z)\right| \int_0^\infty D_\mu(x, a, z)\, z^{\mu+1/2} dz \\ &= c_\mu^{-1} (xa)^{\mu+1/2} \sup_{z \in [0,x+a]} \left|z^{-\mu-1/2} \phi(z)\right| \quad (x \in I).\end{aligned} \tag{54}$$

Next, we want to see that $\tau_a \phi$ is continuous on I. To this end, fix $x_0 \in I$. Since

$$\begin{aligned}\lim_{x \longrightarrow x_0} \int_0^\infty \left|D_\mu(x, y, t) - D_\mu(x_0, y, t)\right| t^{\mu+1/2} dt = 0 \\ (y \in I),\end{aligned} \tag{55}$$

given $\varepsilon > 0$ there exists $\delta = \delta(a, x_0, \varepsilon) > 0$ such that $x \in I$ and $|x - x_0| < \delta$ imply

$$\int_0^\infty \left|D_\mu(x, a, t) - D_\mu(x_0, a, t)\right| t^{\mu+1/2} dt < \varepsilon. \tag{56}$$

Therefore, for $|x - x_0| < \delta$ we obtain

$$\begin{aligned}&\left|\left(\tau_a \phi\right)(x) - \left(\tau_a \phi\right)(x_0)\right| \\ &\quad \leq \int_0^\infty \left|D_\mu(x, a, t) - D_\mu(x_0, a, t)\right| \left|\phi(t)\right| dt \\ &\quad \leq \sup_{t \in [0, a+x_0+\delta]} \left|t^{-\mu-1/2} \phi(t)\right| \\ &\quad \cdot \int_0^\infty \left|D_\mu(x, a, t) - D_\mu(x_0, a, t)\right| t^{\mu+1/2} dt \\ &\quad \leq \varepsilon \sup_{t \in [0, a+x_0+\delta]} \left|t^{-\mu-1/2} \phi(t)\right|.\end{aligned} \tag{57}$$

Since the function $f(x) = x^{-\mu-1/2} \phi(x)$ $(x \in I)$ is continuous at a, given $\varepsilon > 0$ there exists $\delta = \delta(a, \varepsilon) > 0$ such that $z \in I$ and $|z - a| < \delta$ imply $|z^{-\mu-1/2} \phi(z) - a^{-\mu-1/2} \phi(a)| < \varepsilon$. Moreover, if $0 < x < \delta$ and $z \in I$ with $|z - a| \geq \delta > x$, then $D_\mu(x, a, z) = 0$. Again by (8), for $0 < x < \delta$ we thus have

$$\begin{aligned}&\left|c_\mu (xa)^{-\mu-1/2} \left(\tau_a \phi\right)(x) - a^{-\mu-1/2} \phi(a)\right| \\ &\quad = \left|c_\mu (xa)^{-\mu-1/2} \int_0^\infty D_\mu(x, a, z)\, \phi(z)\, dz \right. \\ &\quad - c_\mu (xa)^{-\mu-1/2} a^{-\mu-1/2} \phi(a) \\ &\quad \left. \cdot \int_0^\infty D_\mu(x, a, z)\, z^{\mu+1/2} dz \right| \leq c_\mu (xa)^{-\mu-1/2} \\ &\quad \cdot \int_0^\infty D_\mu(x, a, z) \left|z^{-\mu-1/2} \phi(z) - a^{-\mu-1/2} \phi(a)\right| \\ &\quad \cdot z^{\mu+1/2} dz = c_\mu (xa)^{-\mu-1/2} \int_{|z-a|<\delta} D_\mu(x, a, z) \\ &\quad \cdot \left|z^{-\mu-1/2} \phi(z) - a^{-\mu-1/2} \phi(a)\right| z^{\mu+1/2} dz \leq \varepsilon.\end{aligned} \tag{58}$$

Consequently, $\lim_{x\to 0+} x^{-\mu-1/2}(\tau_a\phi)(x) = c_\mu^{-1}\phi(a)$, thus proving that $\tau_a\phi \in \mathscr{C}_\mu$. Now, the continuity of the translation operator can be deduced from (54): given $c \in I$,

$$\begin{aligned} &\sup_{x\in[0,c]} \left|x^{-\mu-1/2}(\tau_a\phi)(x)\right| \\ &\quad \leq c_\mu^{-1} a^{\mu+1/2} \sup_{z\in[0,c+a]} \left|z^{-\mu-1/2}\phi(z)\right| \quad (\phi \in \mathscr{C}_\mu). \end{aligned} \tag{59}$$

Finally, part (iii) derives immediately from (i) and (ii). The proof is complete. □

In what follows, $\phi \in \mathscr{C}_\mu \cap \mathscr{H}'_\mu$ will mean that the function ϕ defines an absolutely regular $\mathscr{H}'_\mu$-distribution (cf. [32, Theorem 3.7]). In other words, $\phi\varphi$ is an integrable function on $]0,\infty[$ whenever $\varphi \in \mathscr{H}_\mu$, and

$$\langle \phi, \varphi \rangle = \int_0^\infty \phi(x)\varphi(x)\,dx \quad (\varphi \in \mathscr{H}_\mu). \tag{60}$$

Lemma 8. *Assume $\phi \in \mathscr{C}_\mu \cap \mathscr{H}'_\mu$ and let $r, a \in I$. The following holds:*

$$S_\mu(\lambda_r\phi) = r^2\lambda_r(S_\mu\phi); \tag{61}$$

$$S_\mu(\tau_a\phi) = \tau_a(S_\mu\phi); \tag{62}$$

$$\begin{aligned} \mathscr{S}_1(\phi) &= \operatorname{span}\{\tau_a(\lambda_r\phi) : a, r \in I\} \\ &= \operatorname{span}\{\lambda_r(\tau_a\phi) : r, a \in I\}. \end{aligned} \tag{63}$$

Proof. The identity in (61) may be easily verified by means of the operational rule (18). Equation (62) is a consequence of the fact that S_μ commutes with Hankel translations on $\mathscr{H}'_\mu$ (cf. [33]). Finally, (63) derives from (23) and the identity

$$\tau_a(\lambda_r\phi) = r^{-\mu-1/2}\lambda_r(\tau_{ar}\phi), \tag{64}$$

which can be checked through (13). □

Theorem 9. *Let $\phi \in \mathscr{C}_\mu \cap \mathscr{H}'_\mu$ and $c > 0$. Then $\mathscr{S}_1(\phi) = \operatorname{span}\{\tau_a(\lambda_r\phi) : a, r \in I\}$ is dense in $\mathscr{C}_{\mu,c}$ if, and only if, $\phi \notin \pi_\mu$.*

Proof. First suppose $\phi \in \pi_\mu$ is such that $\deg[z^{-\mu-1/2}\phi(z)] = m$, so that $S_\mu^{m+1}\phi = 0$. A combination of (61) and (62) gives

$$\begin{aligned} S_\mu^{m+1}[\lambda_r(\tau_a\phi)] = r^{2(m+1)}\lambda_r[\tau_a(S_\mu^{m+1}\phi)] = 0 \\ (r, a \in I). \end{aligned} \tag{65}$$

Bearing in mind (63) and [24, Theorem 2.19], it turns out that, for every $f \in \mathscr{S}_1(\phi)$, one has $f \in \pi_\mu$ and $\deg[z^{-\mu-1/2}f(z)] \leq m$. The subspace formed by such functions being finite-dimensional (hence closed), it cannot be dense in $\mathscr{C}_{\mu,c}$, which prevents $\mathscr{S}_1(\phi)$ from being dense in $\mathscr{C}_{\mu,c}$.

For the converse, suppose that $\mathscr{S}_1(\phi)$ is not dense in $\mathscr{C}_{\mu,c}$. By the Hahn-Banach and Riesz representation theorems, there is a nonzero Radon measure γ, with $\operatorname{supp}\gamma = [0,c]$ and $\int_0^c t^{\mu+1/2}d|\gamma|(t) < \infty$, such that

$$\begin{aligned} f_r(x) = \langle \gamma, \tau_x(\lambda_r\phi)\rangle = \int_0^c \tau_x(\lambda_r\phi)(t)\,d\gamma(t) = 0 \\ (x, r \in I). \end{aligned} \tag{66}$$

The Hankel-Stieltjes transform

$$(h'_\mu\gamma)(t) = \int_0^c \mathscr{J}_\mu(xt)\,d\gamma(x) \quad (t \in I) \tag{67}$$

of the measure γ gives rise to a multiplier of $\mathscr{H}_\mu$. Indeed,

$$\begin{aligned} &(t^{-1}D_t)^k t^{-\mu-1/2}(h'_\mu\gamma)(t) \\ &\quad = \int_0^c (t^{-1}D_t)^k\left[(xt)^{-\mu}J_\mu(xt)\right]x^{\mu+1/2}d\gamma(x) \\ &\quad = \int_0^c (-x^2)^k(xt)^{-\mu-k}J_{\mu+k}(xt)\,x^{\mu+1/2}d\gamma(x) \\ &\qquad (t \in I,\ k \in \mathbb{Z}_+). \end{aligned} \tag{68}$$

Hence

$$\begin{aligned} &\left|(t^{-1}D_t)^k t^{-\mu-1/2}(h'_\mu\gamma)(t)\right| \\ &\quad \leq c^{2k}c_{\mu+k}^{-1}\int_0^c x^{\mu+1/2}d|\gamma|(x) < \infty \\ &\qquad (t \in I,\ k \in \mathbb{Z}_+). \end{aligned} \tag{69}$$

By (15), this proves that $t^{-\mu-1/2}(h'_\mu\gamma)(t) \in \mathscr{O}$ or, equivalently, $\gamma \in \mathscr{O}'_{\mu,\#}$.

Given $r \in I$, we want to show that $f_r \in \mathscr{H}'_\mu$ and $f_r = (\lambda_r\phi)\#\gamma$. To this end, choose $m \in \mathbb{N}$ so that

$$\int_0^\infty \frac{\phi(x)}{(1+x^2)^m}x^{\mu+1/2}dx < \infty \tag{70}$$

[32, Theorem 3.7], and pick an arbitrary $\varphi \in \mathscr{H}_\mu$. Then

$$\begin{aligned} &\int_0^c d|\gamma|(t)\int_0^\infty |(\lambda_r\phi)(x)||\tau_t\varphi(x)|\,dx \leq \int_0^c d|\gamma|(t) \\ &\quad \cdot \int_0^\infty |(\lambda_r\phi)(x)|\,dx \int_{|x-t|}^\infty D_\mu(x,t,z) \\ &\quad \cdot \frac{\left|(1+z^2)^m z^{-\mu-1/2}\varphi(z)\right|}{(1+z^2)^m}z^{\mu+1/2}dz \leq c_\mu^{-1}\nu_{\mu,m}(\varphi) \\ &\quad \cdot \int_0^c t^{\mu+1/2}d|\gamma|(t)\int_0^\infty \frac{|(\lambda_r\phi)(x)|}{(1+|x-t|^2)^m}x^{\mu+1/2}dx. \end{aligned} \tag{71}$$

At this point we apply the following special case of Peetre's inequality (cf. [34, Lemma 5.2])

$$\frac{1}{1+|x-t|^2} \leq 2\frac{1+t^2}{1+x^2} \tag{72}$$

to infer

$$\int_0^c t^{\mu+1/2} d|\gamma|(t) \int_0^\infty \frac{|(\lambda_r\phi)(x)|}{\left(1+|x-t|^2\right)^m} x^{\mu+1/2} dx \le 2^m \left(1+c^2\right)^m \int_0^c t^{\mu+1/2} d|\gamma|(t) \cdot \int_0^\infty \frac{|(\lambda_r\phi)(x)|}{\left(1+x^2\right)^m} x^{\mu+1/2} dx. \quad (73)$$

Now

$$\int_0^\infty \frac{|(\lambda_r\phi)(x)|}{\left(1+x^2\right)^m} x^{\mu+1/2} dx = \int_0^\infty \frac{|\phi(rx)|}{\left(1+x^2\right)^m} x^{\mu+1/2} dx = \int_0^\infty \frac{|\phi(u)|}{\left[1+(u/r)^2\right]^m} \left(\frac{u}{r}\right)^{\mu+1/2} \frac{du}{r} = r^{-\mu-3/2} \int_0^\infty \frac{|\phi(u)|}{\left[1+(u/r)^2\right]^m} u^{\mu+1/2} du. \quad (74)$$

If $0 < r \le 1$, then

$$\frac{1}{\left[1+(u/r)^2\right]^m} \le \frac{1}{\left(1+u^2\right)^m}. \quad (75)$$

If $r \ge 1$, then

$$\frac{1}{\left[1+(u/r)^2\right]^m} = \frac{r^{2m}}{\left(r^2+u^2\right)^m} \le \frac{r^{2m}}{\left(1+u^2\right)^m}. \quad (76)$$

A combination of (71), (73), (74), (75), and (76) finally yields

$$\int_0^c d|\gamma|(t) \int_0^\infty |(\lambda_r\phi)(x)| |\tau_t\varphi(x)| dx \le c_\mu^{-1} \nu_{\mu,m}(\varphi) \cdot 2^m \left(1+c^2\right)^m C(r,m) \int_0^c t^{\mu+1/2} d|\gamma|(t) \cdot \int_0^\infty \frac{|\phi(u)|}{\left(1+u^2\right)^m} u^{\mu+1/2} du < \infty, \quad (77)$$

where $C(r,m)$ denotes a suitable positive constant.

Consequently, $f_r \in \mathscr{H}'_\mu$ and the Fubini theorem, along with [21, Definition 2 and Proposition 3], can be applied to write

$$\begin{aligned} \langle f_r, \varphi \rangle &= \int_0^\infty f_r(x) \varphi(x) dx \\ &= \int_0^\infty \varphi(x) dx \int_0^c \tau_x (\lambda_r\phi)(t) d\gamma(t) \\ &= \int_0^c d\gamma(t) \int_0^\infty (\lambda_r\phi)(x) (\tau_t\varphi)(x) dx \\ &= \langle \gamma, (\lambda_r\phi) \# \varphi \rangle = \langle \gamma \# (\lambda_r\phi), \varphi \rangle \\ &\qquad \left(\varphi \in \mathscr{H}_\mu\right), \end{aligned} \quad (78)$$

thus showing that $f_r = \gamma\#(\lambda_r\phi) = (\lambda_r\phi)\#\gamma$.

Invoking [19, Definition 4.4] and taking into account (66) lead to

$$\langle \lambda_r\phi, \gamma\#(\lambda_r\varphi) \rangle = \langle (\lambda_r\phi)\#\gamma, \lambda_r\varphi \rangle = \langle f_r, \lambda_r\varphi \rangle = 0 \quad \left(\varphi \in \mathscr{H}_\mu,\ r \in I\right), \quad (79)$$

or, equivalently from (14),

$$\frac{1}{r^2} \left\langle \left(h'_\mu\phi\right)\left(\frac{t}{r}\right), t^{-\mu-1/2} \left(h_\mu\varphi\right)\left(\frac{t}{r}\right) \left(h'_\mu\gamma\right)(t) \right\rangle = 0 \quad \left(\varphi \in \mathscr{H}_\mu,\ r \in I\right). \quad (80)$$

Since $\gamma \neq 0$, there exist $t_0 \in I$ and $0 < \delta < t_0$ such that

$$t^{-\mu-1/2} \left(h'_\mu\gamma\right)(t) \neq 0 \quad \left(|t-t_0| < \delta\right). \quad (81)$$

Given $R > 0$, let $t^* = t_0/R$. Then

$$(Rt)^{-\mu-1/2} \left(h'_\mu\gamma\right)(Rt) \neq 0 \quad \left(|t-t^*| < \frac{\delta}{R}\right). \quad (82)$$

Choose any $h_\mu\varphi \in \mathscr{B}_\mu$ with

$$\operatorname{supp}\left[t^{-\mu-1/2}\left(h_\mu\varphi\right)(t)\right] \subset \left]t^* - \frac{\delta}{R}, t^* + \frac{\delta}{R}\right[, \quad (83)$$

and set

$$\left(h_\mu\psi\right)(t) = \frac{R\left(h_\mu\varphi\right)(t)}{(Rt)^{-\mu-1/2}\left(h'_\mu\gamma\right)(Rt)} \quad (t \in I). \quad (84)$$

Clearly, $h_\mu\psi \in \mathscr{B}_\mu$ and

$$\operatorname{supp}\left[t^{-\mu-1/2}\left(h_\mu\psi\right)(t)\right] \subset \left]t^* - \frac{\delta}{R}, t^* + \frac{\delta}{R}\right[. \quad (85)$$

Equation (80) allows us to conclude

$$\begin{aligned} &\langle h'_\mu\phi, h_\mu\varphi \rangle \\ &= \int_0^\infty \left(h'_\mu\phi\right)(t) (Rt)^{-\mu-1/2} \left(h_\mu\psi\right)(t) \left(h'_\mu\gamma\right)(Rt) \frac{dt}{R} \\ &= \int_0^\infty \left(h'_\mu\phi\right)\left(\frac{t}{R}\right) t^{-\mu-1/2} \left(h_\mu\psi\right)\left(\frac{t}{R}\right) \left(h'_\mu\gamma\right)(t) \frac{dt}{R^2} \\ &= \frac{1}{R^2} \left\langle \left(h'_\mu\phi\right)\left(\frac{t}{R}\right), t^{-\mu-1/2} \left(h_\mu\psi\right)\left(\frac{t}{R}\right) \left(h'_\mu\gamma\right)(t) \right\rangle \\ &= 0. \end{aligned} \quad (86)$$

In other words, $\operatorname{supp} h'_\mu\phi = \{0\}$. From Remark 4 and Theorem 5 we get that $h'_\mu\phi \in \operatorname{span}\{\Lambda_j : j \in \mathbb{Z}_+\}$, which means [24, Theorem 2.19] that $\phi \in \pi_\mu$.

The proof is now complete. □

4. Final Remarks

(i) RBFNNs of Hankel translates, as defined in this paper, admit only one-dimensional inputs. In order to allow for multidimensional inputs one should consider the multidimensional Hankel translation, defined by iteration of the one-dimensional translation operator with respect to each of the variables while the others are kept fixed (see, e.g., [35] and references therein). The proof of the above results for the multidimensional case could well be the subject of a forthcoming paper.

(ii) According to Theorem 9 and [32, Theorem 3.7 and Corollary 3.10], any continuous function $\sigma \notin \pi_\mu$ for which there exists $k \in \mathbb{Z}_+$ so that $(1+x^2)^{-k}\sigma(x)$ is bounded on I, or $(1+x^2)^{-k}x^{\mu+1/2}\sigma(x)$ is integrable on I, can be used as an activation function yielding universal approximation. A paradigmatic example is the Gaussian

$$\sigma(x) = x^{\mu+1/2}\exp\left(\frac{-x^2}{2}\right) \quad (x \in I). \tag{87}$$

(iii) By considering RBFNNs of Hankel translates, a new parameter μ is introduced which in practice leaves a greater variety of manageable kernels at our disposal. This could be useful in handling mathematical models built upon a class of radial basis functions depending on the order μ whose performance might be improved by finely tuning μ, without increasing the number of centroids [36, 37].

Data Availability

This research is not based on any experimental data.

Conflicts of Interest

The author declares that she has no conflicts of interest.

References

[1] M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*, vol. 12, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 2003.

[2] D. S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive networks," *Complex Systems*, vol. 2, no. 3, pp. 321–355, 1988.

[3] R. P. Lippmann, "Pattern classification using neural networks," *IEEE Communications Magazine*, vol. 27, no. 11, pp. 47–64, 1989.

[4] S. Renals and R. Rohwer, "Phoneme classification experiments using radial basis functions," in *Proceedings of the International Joint Conference on Neural Networks I*, pp. 461–467, Washington, DC, USA, June 1989.

[5] J. Park and I. W. Sandberg, "Universal approximation using radial basis function networks," *Neural Computation*, vol. 3, no. 2, pp. 246–257, 1991.

[6] J. Park and I. W. Sandberg, "Approximation and radial-basis-function networks," *Neural Computation*, vol. 5, no. 2, pp. 305–316, 1993.

[7] T. Chen and H. Chen, "Approximation capability to functions of several variables, nonlinear functionals, and operators by radial basis function neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 6, no. 4, pp. 904–910, 1995.

[8] T. Chen and H. Chen, "Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 6, no. 4, pp. 911–917, 1995.

[9] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function," *Neural Networks*, vol. 6, no. 6, pp. 861–867, 1993.

[10] H. N. Mhaskar and C. A. Micchelli, "Approximation by superposition of sigmoidal and radial basis functions," *Advances in Applied Mathematics*, vol. 13, no. 3, pp. 350–373, 1992.

[11] A. Pinkus, "TDI-subspaces of $C(\mathbb{R}^d)$ and some density problems from neural networks," *Journal of Approximation Theory*, vol. 85, pp. 269–287, 1996.

[12] S. Sonoda and N. Murata, "Neural network with unbounded activation functions is universal approximator," *Applied and Computational Harmonic Analysis*, vol. 43, no. 2, pp. 233–268, 2017.

[13] A. H. Zemanian, "A distributional Hankel transformation," *SIAM Journal on Applied Mathematics*, vol. 14, pp. 561–576, 1966.

[14] A. H. Zemanian, *Generalized Integral Transformations*, Interscience Publishers, 1968.

[15] A. H. Zemanian, "The Hankel transformation of certain distributions of rapid growth," *SIAM Journal on Applied Mathematics*, vol. 14, pp. 678–690, 1966.

[16] J. de Sousa Pinto, "A generalised Hankel convolution," *SIAM Journal on Mathematical Analysis*, vol. 16, no. 6, pp. 1335–1346, 1985.

[17] J. Betancor and I. Marrero, "The Hankel convolution and the Zemanian spaces $\mathscr{B}_\mu$ and $\mathscr{B}'_\mu$," *Mathematische Nachrichten*, vol. 160, pp. 277–298, 1993.

[18] J. J. Betancor and I. Marrero, "Structure and convergence in certain spaces of distributions and the generalized Hankel convolution," *Mathematica Japonica*, vol. 38, no. 6, pp. 1141–1155, 1993.

[19] I. Marrero and J. J. Betancor, "Hankel convolution of generalized functions," *Rendiconti di Matematica e delle sue Applicazioni, Serie VII*, vol. 15, no. 3, pp. 351–380, 1995.

[20] J. J. Betancor and I. Marrero, "Multipliers of Hankel transformable generalized functions," *Commentationes Mathematicae*, vol. 33, no. 3, pp. 389–401, 1992.

[21] J. J. Betancor and I. Marrero, "On the topology of the space of Hankel convolution operators," *Journal of Mathematical Analysis and Applications*, vol. 201, no. 3, pp. 994–1001, 1996.

[22] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

[23] G. Gigante, "Transference for hypergroups," *Collectanea Mathematica*, vol. 52, no. 2, pp. 127–155, 2001.

[24] C. Arteaga and I. Marrero, "A scheme for interpolation by Hankel translates of a basis function," *Journal of Approximation Theory*, vol. 164, no. 12, pp. 1540–1576, 2012.

[25] C. Arteaga and I. Marrero, "Density in spaces of interpolation by Hankel translates of a basis function," *Journal of Function Spaces and Applications*, vol. 2013, Article ID 813502, 9 pages, 2013.

[26] C. Arteaga and I. Marrero, "Direct form seminorms arising in the theory of interpolation by Hankel translates of a basis

function," *Advances in Computational Mathematics*, vol. 40, no. 1, pp. 167–183, 2014.

[27] C. Arteaga and I. Marrero, "Interpolation by Hankel translates of a basis function: inversion formulas and polynomial bounds," *The Scientific World Journal*, vol. 2014, Article ID 242750, 13 pages, 2014.

[28] C. Arteaga and I. Marrero, "Universal approximation by radial basis function networks of Delsarte translates," *Neural Networks*, vol. 46, pp. 299–305, 2013.

[29] C. Arteaga and I. Marrero, "Approximation in weighted p-mean by RBF networks of Delsarte translates," *Journal of Mathematical Analysis and Applications*, vol. 414, no. 1, pp. 450–460, 2014.

[30] C. Arteaga and I. Marrero, "Wiener's tauberian theorems for the Fourier-Bessel transformation and uniform approximation by RBF networks of Delsarte translates," *Journal of Mathematical Analysis and Applications*, vol. 431, no. 1, pp. 482–493, 2015.

[31] W. Rudin, *Functional Analysis*, McGraw-Hill, 2nd edition, 1991.

[32] I. Marrero, "Regular and absolutely regular Hankel-transformable distributions," *Mathematische Nachrichten*, vol. 263/264, pp. 154–170, 2004.

[33] J. J. Betancor, "A new characterization of the bounded operators commuting with Hankel translation," *Archiv der Mathematik*, vol. 69, no. 5, pp. 403–408, 1997.

[34] J. Barros-Neto, *An Introduction to The Theory of Distributions*, Krieger, 1981.

[35] J. Dziuban'ski, M. Preisner, and B. Wróbel, "Multivariate Hörmander-type multiplier theorem for the Hankel transform," *Journal of Fourier Analysis and Applications*, vol. 19, pp. 417–437, 2013.

[36] H. Corrada, K. Leeb, B. Klein et al., "Examining the relative influence of familial, genetic and environmental covariate information in flexible risk models," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 20, pp. 8128–8133, 2009.

[37] S. H. Javaran, N. Khaji, and A. Noorzad, "First kind Bessel function (J-Bessel) as radial basis function for plane dynamic analysis using dual reciprocity boundary element method," *Acta Mechanica*, vol. 218, no. 3-4, pp. 247–258, 2011.


Advances in
Operations Research
Advances in
Decision Sciences
Journal of
Applied Mathematics
The Scientific
World Journal
Journal of
Probability and Statistics
International
Journal of
Mathematics and
Mathematical
Sciences
Hindawi
Submit your manuscripts at
www.hindawi.com
Journal of
Optimization
International Journal of
Engineering
Mathematics
International Journal of
Analysis
Journal of
Complex Analysis
Advances in
Numerical Analysis
Mathematical Problems
in Engineering
International Journal of
Differential Equations
Discrete Dynamics in
Nature and Society
International Journal of
Stochastic Analysis
Journal of
Mathematics
Journal of
Function Spaces
Abstract and
Applied Analysis
Advances in
Mathematical Physics