Reproducing Kernel Hilbert Spaces and Fractal Interpolation

P. Bouboulis[*,a], M. Mavroforakis[b]

[a]*Department of Informatics and Telecommunications, Telecommunications and Signal Processing, University of Athens, Panepistimiopolis 157 84, Athens, Greece*
[b]*Computational Biomedicine Lab, Department of Computer Science, University of Houston, Houston, Texas 77204, USA*

Abstract

Reproducing Kernel Hilbert Spaces (RKHSs) are a very useful and powerful tool of functional analysis with application in many diverse paradigms, such as multivariate statistics and machine learning. Fractal interpolation, on the other hand, is a relatively recent technique that generalizes traditional interpolation through the introduction of self-similarity. In this work we show that the functional space of any family of (recurrent) fractal interpolation functions ((R)FIFs) constitutes a RKHS with a specific associated kernel function, thus, extending considerably the toolbox of known kernel functions and introducing fractals to the RKHS world. We also provide the means for the computation of the kernel function that corresponds to any specific fractal RKHS and give several examples.

Key words: fractal interpolation, reproducing kernel Hilbert Space, Kernels

1. Introduction

Fractal interpolation, as introduced by Barnsley in [3] (see also [4]), is an alternative to traditional interpolation techniques, which gives a broader set of interpolants. In fact, many traditional interpolation techniques (splines, hermite polynomials etc) are included as special cases. Its main differences include: a) the definition of a functional relation (see (5)), which implies a self similarity in small scales, b) the constructive way (through iterations), which is used to compute the interpolant, instead of the descriptive one (usually a formula) provided by the classical methods and c) the usage of some parameters, which are usually called vertical scaling factors, that are strongly related to the fractal dimension of the interpolant. Fractal interpolation is used in many scientific fields such as image compression, multiwavelets, computer graphics e.t.c. (see for example [16], [10]).

A Reproducing Kernel Hilbert Space (RKHS) on the other hand, introduced in [21], [22], [1], is a rich construct (roughly, a smooth space with a generalized inner product), which has been proved to be a very powerful tool, e.g., in functional analysis and integral equations, information theory and statistical learning ([18], [19], [20], [29], [28]). To every RKHS, an input space and a kernel function can be associated, such that the inner products of any pair of elements of the input space are mapped through the kernel function to the inner product of elements of the RKHS. In this way, problems that are nonlinear in the input space, can be

---

transformed to linear ones in a RKHS, if they can be expressed in terms of inner products, through a suitable kernel function. The main concepts of this procedure can be summarized in the following two steps: 1) Map the finite dimensionality input data from the input space F (usually $F \subset \mathbb{R}^\nu$) into a higher dimensionality (possibly infinite) RKHS $\mathcal{H}$ using the kernel function and 2) Perform a linear processing on the mapped data in $\mathcal{H}$. The procedure is equivalent with a non-linear processing in F (the interested reader may dig more on this subject in [25, 27]).

The notion of the RKHS is a major mathematical tool involved in many paradigms of Machine Learning (ML). The theoretical perspective of these specific methods is mainly the Statistical Learning Theory (SLT) [28], which provides a sound background for the development of learning tools exhibiting guaranteed generalization properties. A main offspring of SLT is the Support Vector Machines (SVMs). SVMs are able to handle classification and regression tasks involving multidimensional, nonlinear data [25, 28]. A variety of kernels with respective RKHS are used in practice (e.g., gaussian kernel, polynomial kernel etc). However, most of these kernel spaces exhibit a common characteristic: they consist of smooth functions.

A major challenge is that, in many real-life situations (e.g., biomedical, environmental, and financial systems), the data involved are highly nonlinear (i.e., not even locally linear). Although the existing ML theory is valuable and provides the tools (i.e., kernel-based classifiers such as SVMs) to handle simple, binary nonlinear cases, there is still an unbridged gap to complex real-life dynamical systems, which obey chaos rules and are accurately modeled by fractals. For example, in regression tasks involving highly nonlinear data, the use of smooth kernel spaces will probably lead to false conclusions.

This work aims to bridge the gap between the two fields (i.e., fractal interpolation and RKHS theory), thus providing the benefits of the one to the other. Specifically, we prove that the space of recurrent fractal interpolation functions (RFIFs) of any given order, constitute a Reproducing Kernel Hilbert Space and provide the means for the computation of the specific kernel functions. Our primary goal is to provide a way to create fractal Reproducing Kernel Hilbert Spaces (RKHSs) and we devised the Fractal Interpolation Functions (FIFs) as the suitable "vehicle" for this purpose. Therefore, the aim of the paper is not to deal with fractal interpolation (as there are not any specific interpolation points given), but, rather, to exploit the mathematical structure of FIFs and construct corresponding RKHSs through them. In particular, we consider three different fractal spaces, i.e., the space of all RFIFs of order r, denoted by $\mathcal{F}_{\mathcal{P},\boldsymbol{u},\boldsymbol{s},r}$, the space of all Hermitian RFIFs of order $2r + 1$, denoted by $\mathcal{H}_{\mathcal{P},\boldsymbol{u},\boldsymbol{s},2r+1}$, and the space of spline RFIFs of order $2r + 1$, denoted by $\mathcal{S}_{\mathcal{P},\boldsymbol{u},\boldsymbol{s},2r+1}$, which we prove that each one is actually a RKHS. Moreover, the corresponding spaces of the simple (non recurrent) fractal interpolation functions are also RKHSs, as specific cases. Furthermore, we compute the dimension of these spaces, which is related to the order of the belonging functions and the number of interpolation points. Finally, we provide the elements for calculating the respective kernel functions κ for these RKHSs as well as the corresponding induced mappings Φ. Thus, we pave the way to the use of new, very broad and rich families of kernel functions, i.e., those related to the (R)FIFs of any kind and order, in very diverse fields like multivariate statistics (e.g., kernel PCA [26]), machine learning (e.g., kernel machines [24], [25]), etc. We should emphasize that all kernel-based methods can directly employ the fractal kernels presented in this paper.

The rest of the material of this paper is organized as follows: the following two Sections

2

are of introductory nature, as they present the fundamental notions of RKHSs and fractal interpolation, thus, providing the necessary background for the development of the results of this work. Most of the material included in these sections may be found in other papers as well ([23], [12], [8], [7], [11], etc). In Sections 4 and 5, the main theoretical results of this work are being deduced. In particular, in Section 4, we show that any functional space of (R)FIFs can be viewed as a RKHS, as it is a finite linear space. Although this result seems rather trivial, we should emphasize that (a) to the best of our knowledge fractals have not been considered before as a tool for kernel learning and (b) several important issues need to be addressed for the computation of the fractal kernels. These issues are been considered in Section 5, where all the necessary steps needed to compute the corresponding kernel functions are described.

2. Reproducing Kernel Hilbert Spaces

We start with some basic definitions regarding the property of positive definite matrices and functions, which play a very important role in the study of RKHS.

Definition 2.1. (Gram Matrix) Given a function $\kappa : X \times X \to \mathbb{R}$ and $x_1, \ldots, x_N \in X$, the matrix[1] $K = (K_{i,j})^N$ with elements $K_{i,j} = \kappa(x_i, x_j)$, for $i, j = 1, \ldots, N$, is called the *Gram matrix* (or *kernel matrix*) of κ with respect to $x_1, \ldots, x_N$.

Definition 2.2. (Positive Definite Matrix) A real symmetric matrix $K = (K_{i,j})^N$ satisfying

$$c^T \cdot K \cdot c = \sum_{i=1, j=1}^{N,N} c_i c_j K_{i,j} \geq 0,$$

for all $c_i \in \mathbb{R}$, $i = 1, \ldots, N$, is called *Positive Definite*.

In matrix analysis literature, this is the definition of the positive semidefinite matrix, but since this is a rather cumbersome term and the distinction between positive definite and positive semidefinite matrices is not important in this paper, we employ the term positive definite in the way introduced here.

Definition 2.3. (Positive Definite Kernel) Let X be a nonempty set. Then a function $\kappa : X \times X \to \mathbb{R}$, which for every specific $N \in \mathbb{N}$ and all $x_1, \ldots, x_N \in X$ gives rise to a positive definite Gram matrix $K = (K_{i,j})^N$, where $K_{i,j} = \kappa(x_i, x_j)$ for $i, j = 1, \ldots, N$, is called a *Positive Definite Kernel*.

In the following we will refer to a positive definite kernel simply as *kernel*. The reason that kernels are so popular is that they can be regarded as a *generalized dot product*. In fact, any dot product is a kernel (of course the opposite is not true). Several properties of dot products (such as the Cauchy-Schwartz inequality) do have natural generalizations to kernels (see [2] and [25]).

Having dealt with the definitions of positivity, we are ready to discuss the main subject of this section. Consider a linear class $\mathcal{H}$ of real valued functions f defined on a set X. Suppose further, that in $\mathcal{H}$ we can define an inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ with corresponding norm $\| \cdot \|_{\mathcal{H}}$ and that $\mathcal{H}$ is complete with respect to that norm, i.e., $\mathcal{H}$ is a Hilbert space. We call

---

[1]The term $(K_{i,j})^N$ denotes a square $N \times N$ matrix.

$\mathcal{H}$ a *Reproducing Kernel Hilbert Space (RKHS)*, if for all $x \in X$ the evaluation functional $T_x : \mathcal{H} \to \mathbb{R} : T_x(f) = f(x)$ is a continuous (or, equivalently, bounded) operator. If this is true, then by the Riesz's representation theorem, for all $x \in X$ there is a function $g_x \in \mathcal{H}$ such that $T_x(f) = f(x) = \langle f, g_x \rangle_{\mathcal{H}}$. The function $\kappa : X \times X \to \mathbb{R} : \kappa(x, y) = g_x(y)$ is called a *reproducing kernel* of $\mathcal{H}$. It can be easily proved that the function κ is a positive definite kernel.

Alternatively, we can define a RKHS as a Hilbert space $\mathcal{H}$ for which there exists a function $\kappa : X \times X \to \mathbb{R}$ with the following two properties:

1. For every $x \in X$, $\kappa(x, \cdot)$ belongs to $\mathcal{H}$.
2. κ has the so called *reproducing property*, i.e.

$$f(x) = \langle f, \kappa(x, \cdot) \rangle_{\mathcal{H}}, \text{ for all } f \in \mathcal{H}, \tag{1}$$

in particular $\kappa(x, y) = \langle \kappa(x, \cdot), \kappa(y, \cdot) \rangle_{\mathcal{H}}$.

It has been proved (see [22], [2]) that to every positive definite kernel κ there corresponds one and only one class of functions $\mathcal{F}$ with a uniquely determined inner product in it, forming a Hilbert space and admitting κ as a reproducing kernel. In fact the kernel κ produces the entire space $\mathcal{H}$, i.e.,

$$\mathcal{H} = \overline{\text{span}\{\kappa(x, \cdot) | x \in X\}}.$$

We note that it is possible to define several different inner products in the same class of functions $\mathcal{H}$, so that $\mathcal{H}$ is complete with respect to each one of the corresponding norms. To each one of the Hilbert spaces $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ there corresponds one and only one kernel function κ. That is to say κ depends not only on the class of functions $\mathcal{H}$, but also on the choice of the inner product that $\mathcal{H}$ admits. Table 1 shows several well known positive definite kernels that are often used in practice. Below we give an important result concerning RKHSs (proved in [1] and [2]) which will be used in our study.

Theorem 2.1 ([2]). *A function $\kappa(x, y)$ is the reproducing kernel of a finite-dimensional class of functions $\mathcal{H}$ if and only if it has the form $\kappa(x, y) = \sum_{i,j=1}^{l} \beta_{i,j} e_i(x) e_j(y)$, where $B = \{\beta_{i,j}\}$ is a positive definite matrix, $l > 0$ and $e_1, \ldots, e_l$ are linearly independent. The corresponding class $\mathcal{H}$ is then generated by the functions $e_1, \ldots, e_l$ (i.e., the functions e_i form a basis of $\mathcal{H}$) and the corresponding inner product of two functions $f = \sum_{i=1}^{l} \zeta_i e_i$ and $\tilde{f} = \sum_{i=1}^{l} \tilde{\zeta}_i e_i$ is given by $\langle f, \tilde{f} \rangle = \sum_{i,j=1}^{l} \alpha_{i,j} \zeta_i \tilde{\zeta}_j$, where $A = \{\alpha_{i,j}\}$ is the inverse of the matrix B and $\alpha_{i,j} = \langle e_i, e_j \rangle$. The converse is also true.*

3. Fractal Interpolation

3.1. Iterated Function System (IFS) - Recurrent IFS (RIFS)

An *Iterated Function System* $\{X; w_{1-N}\}$ (IFS for short, see [17], [14], [5]) is defined as a pair consisting of a complete metric space (X, ρ) together with a finite set of continuous contractive mappings $w_i : X \to X$, with respective contraction factors s_i for $i = 1, 2, \ldots, N$ $(N \geq 2)$. The *attractor* of an IFS is the unique set E for which $E = \lim_{k \to \infty} W^k(A_0)$ for every starting compact set A_0, where

$$W(A) = \bigcup_{i=1}^{N} w_i(A) \text{ for all } A \in \mathfrak{H}(X),$$

4

Kernel	Function
Gaussian	$\kappa(x,y) = exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$, $\sigma > 0$
(Inhomogeneous) Polynomial	$\kappa(x,y) = (\langle x,y \rangle + c)^d$
B$_\mathbf{n}$-Spline of odd order	$\kappa(x,y) = B_{2r+1}(\|x-y\|)$, with $B_n = \bigotimes_{i=1}^{n} I_{\left[-\frac{1}{2},\frac{1}{2}\right]}$

Table 1: Several Kernels that are used in practice. Note that the term $I_{\left[-\frac{1}{2},\frac{1}{2}\right]}$ that appears in the spline Kernel and the term I_i which is used in the following sections are not the same. For a detailed description of the term $I_{\left[-\frac{1}{2},\frac{1}{2}\right]}$ see [25].

is the so called *Hutchinson operator* [17] and $\mathfrak{H}(X)$ is the complete metric space of all nonempty compact subsets of X with respect to the Hausdorff metric h (for the definition of the Hausdorff metric and properties of $(\mathfrak{H}(X), h)$ see [5]).

A more general concept, that allows the construction of even more complicated sets, is that of the *Recurrent Iterated Function System*, or *RIFS* for short (see [4], [5]), which consists of the IFS $\{X; w_{1-N}\}$ together with an irreducible row-stochastic matrix $P = (p_{\nu,\mu})^N$, $(p_{\nu,\mu} \in [0,1] : \nu, \mu = 1, \ldots, N)$, such that $\sum_{\mu=1}^{N} p_{\nu,\mu} = 1$, $\nu = 1, \ldots, N$. The recurrent structure is given by the (irreducible) *connection matrix* $C = (c_{\nu,\mu})^N$ which is defined by

$$c_{\nu,\mu} = \begin{cases} 1, & \text{if } p_{\mu,\nu} > 0 \\ 0, & \text{if } p_{\mu,\nu} = 0 \end{cases},$$

where $\nu, \mu = 1, 2, \ldots, N$. The transition probability for a certain discrete time Markov process is $p_{\nu,\mu}$, which gives the probability of transition into state μ given that the process is in state ν.

In this case the contractive mapping $\boldsymbol{W}$ acts on N-tuples of *l*-dimensional sets: $\boldsymbol{A} = (A_1, ..., A_N)$. In order to define it, a little more effort is needed. The interested reader may found more details on this subject in [4], [5], [9].

3.2. Fractal Interpolation Functions of order r

In this section we briefly describe the construction of fractal interpolation functions based on RIFSs (for details see [5], [9]). The main idea, which will be presented in detail hereafter, is to define a RIFS that acts on the subsets of the input domain that are produced by the partitioning induced by the interpolation points. Let $X = [0,1] \times \mathbb{R}$, $r > 0$ and $\Delta = \{(x_i, y_i) : i = 0, 1, \ldots, N\}$ be an interpolation set with $N+1$ interpolation points such that $0 = x_0 < x_1 < \cdots < x_N = 1$. The interpolation points divide $[0,1]$ into N intervals $I_i = [x_{i-1}, x_i]$, $i = 1, \ldots, N$, which we call *domains*. In addition, let $\hat{\Delta} = \{(\hat{x}_j, \hat{y}_j) : j = 0, 1, \ldots, M\}$ be a subset of Δ, such that $0 = \hat{x}_0 < \hat{x}_1 < \cdots < \hat{x}_M = 1$. We also assume that for every $j = 0, 1, \ldots, M-1$ there is at least one i such that $\hat{x}_j < x_i < \hat{x}_{j+1}$. Thus, the points of $\hat{\Delta}$ divide $[0,1]$ into M intervals $J_j = [\hat{x}_{j-1}, \hat{x}_j]$, $j = 1, \ldots, M$, which we call *regions*. Finally, let $\mathbb{J}$ be a labeling mapping such that $\mathbb{J}: \{1, 2, \ldots, N\} \to \{1, 2, \ldots, M\}$ with $\mathbb{J}(i) = j$ and $\mathbb{I}: \{0, 1, 2, \ldots, M\} \to \{0, 1, 2, \ldots, N\}$ a mapping with the property that $\mathbb{I}(j) = i$, if $\hat{x}_j = x_i$,

5

for $j = 0, \ldots, M$. Since $\hat{\Delta} \subset \Delta$, each one of the values $\hat{x}_j$ will be identical with one x_i. For example $\hat{x}_0 = x_0$, $\hat{x}_1 = x_2$, $\hat{x}_2 = x_4$ etc. The mapping $\mathbb{I}$ is used to change the notation from $\hat{x}_j$ to the corresponding x_i in this manner: $\hat{x}_j = x_{\mathbb{I}(j)}$. In the special case where the interpolation points are equidistant (that is $x_i - x_{i-1} = \delta$, $i = 1, 2, \ldots, N$, and $\hat{x}_j - \hat{x}_{j-1} = \psi$, $j = 1, 2, \ldots, M$), each region contains exactly $\alpha = \psi/\delta \in \mathbb{N}$ domains.

Next, we define N mappings of the form:

$$w_i \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} L_i(x) \\ F_i(x, y) \end{pmatrix}, \quad \text{for } i = 1, 2, \ldots, N, \tag{2}$$

where $L_i(x) = a_i x + b_i$, $F_i(x, y) = s_i y + p_i(x)$ and $p_i(x)$ is a polynomial of order r. Each mapping w_i is constrained to map the endpoints of the region $J_{\mathbb{J}(i)}$ to the endpoints of the domain I_i. That is,

$$w_i \begin{pmatrix} \hat{x}_{\mathbb{J}(i)-1} \\ \hat{y}_{\mathbb{J}(i)-1} \end{pmatrix} = \begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix}, \quad w_i \begin{pmatrix} \hat{x}_{\mathbb{J}(i)} \\ \hat{y}_{\mathbb{J}(i)} \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \quad \text{for } i = 1, 2, \ldots, N. \tag{3}$$

In this manner, vertical segments are mapped to vertical segments scaled by the factor s_i. The parameter s_i is called the *vertical scaling factor* of the mapping w_i. It is easy to show that if $|s_i| < 1$, then there is a metric d equivalent to the Euclidean metric, such that w_i is a contraction (i.e., there is $\hat{s}_i : 0 \leq \hat{s}_i < 1$ such that $d(w_i(\boldsymbol{x}), w_i(\boldsymbol{y})) \leq \hat{s}_i d(\boldsymbol{x}, \boldsymbol{y})$, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$, see [5]). Equations (3) can be recast as:

$$p_i(\hat{x}_{\mathbb{J}(i)-1}) = y_{i-1} - s_i \hat{y}_{\mathbb{J}(i)-1}, \quad p_i(\hat{x}_{\mathbb{J}(i)}) = y_i - s_i \hat{y}_{\mathbb{J}(i)}. \tag{4}$$

The $N \times N$ stochastic matrix $(p_{n,m})^N$ is defined by the labeling function $\mathbb{J}$ as follows:

$$p_{n,m} = \begin{cases} \frac{1}{\gamma_n}, & \text{if } I_n \subseteq J_{\mathbb{J}(m)} \\ 0, & \text{otherwise.} \end{cases},$$

where γ_n is the number of positive entries of the line[2] n, $n = 1, 2, \ldots, N$. This means that $p_{n,m}$ is positive, if the transformation L_m, maps a region containing the n-th domain (i.e., I_n) to the m-th domain (i.e., I_m). If we take a point in $I_n \times \mathbb{R}$, $n = 1, \ldots, N$, we say that we are in state n. The number $p_{n,m}$ shows the probability of applying the mapping w_m to that point, so that the system transits to state m. Sometimes, it is more convenient to describe the matrix P through the connection matrix C or the connection vector $\boldsymbol{v}$, which are defined as follows:

$$c_{n,m} = \begin{cases} 1, & p_{m,n} > 0 \\ 0, & \text{otherwise} \end{cases},$$

$$\boldsymbol{v} = (\mathbb{J}(1), \mathbb{J}(2), \ldots, \mathbb{J}(N)).$$

Next, we consider[3] $(C([x_0, x_N]), \| \cdot \|_\infty)$, where $\|\psi\|_\infty = \max\{|\psi(x)|, \ x \in [x_0, x_N]\}$ for any $\psi \in C([x_0, x_N])$ and the complete metric subspace $\mathbb{F}_\Delta = \{g \in C([x_0, x_N]) : g(x_i) = y_i, \text{ for } i = 0, 1, \ldots, N\}$. The Read-Bajraktarevic operator $T_{\Delta,\hat{\Delta}} : \mathbb{F}_\Delta \to \mathbb{F}_\Delta$ is defined as follows

$$(T_{\Delta,\hat{\Delta}} g)(x) = F_i \left(L_i^{-1}(x), g \left(L_i^{-1}(x) \right) \right), \quad \text{for } x \in I_i, i = 1, 2, \ldots, N.$$

---

[2] Note that this leads to an equiprobable transition matrix

[3] Throughout the paper, $C([x_0, x_N])$ denotes the space of all continuous functions and $C^r([x_0, x_N])$ denotes the space of all r-times differentiable functions.

It is easy to verify that $T_{\Delta,\hat{\Delta}}g$ is well defined and that $T_{\Delta,\hat{\Delta}}$ is a contraction with respect to the $\|\cdot\|_\infty$ metric. According to the Banach fixed-point theorem, there exists a unique $f \in \mathbb{F}_\Delta$ such that $T_{\Delta,\hat{\Delta}}f = f$. If f_0 is any interpolation function and $f_n = T^n_{\Delta,\hat{\Delta}}f_0$, where $T^n_{\Delta,\hat{\Delta}} = T_{\Delta,\hat{\Delta}} \circ T_{\Delta,\hat{\Delta}} \circ \cdots \circ T_{\Delta,\hat{\Delta}}$, then $(f_n)_{n\in\mathbb{N}}$ converges uniformly to f. The graph of the function f is the *attractor* of the RIFS $\{X, w_{1-N}, (p_{i,j})^N\}$ associated with the interpolation points (see [5]). Note that f interpolates the points of Δ for any selection of the parameters of the polynomials p_i that satisfies (3). We will refer to a function of this nature as *Recurrent Fractal Interpolation Function (RFIF) of order r*. In the case where all the elements of the stochastic matrix are equal to 1 (i.e., we have an IFS instead of a RIFS), the function will be simply referred to as *Fractal Interpolation Function (FIF)*. We emphasize that a RFIF is the only function which satisfies the functional relation

$$f(L_i(x)) = s_i f(x) + p_i(x) \tag{5}$$

for all $x \in J_{\mathbb{J}(i)}$, $i = 1, 2, \ldots, N$.

For the following we assume that the polynomials p_i take the form:

$$p_i(x) = c_{i,0} + c_{i,1}x + \sum_{k=2}^{r} c_{i,k}x^k, \tag{6}$$

for $i = 1, 2, \ldots, N$. From the join up conditions (4) one can easily deduce that

$$c_{i,0} = y_{i-1} - s_i\hat{y}_{j-1} - c_{i,1}\hat{x}_{j-1} - \sum_{k=2}^{r} c_{i,k}\hat{x}^k_{j-1}, \tag{7}$$

$$c_{i,1} = \frac{y_i - y_{i-1}}{\hat{x}_j - \hat{x}_{j-1}} - \sum_{k=2}^{r} c_{i,k}\frac{\hat{x}^k_j - \hat{x}^k_{j-1}}{\hat{x}_j - \hat{x}_{j-1}} - s_i\frac{\hat{y}_j - \hat{y}_{j-1}}{\hat{x}_j - \hat{x}_{j-1}}, \tag{8}$$

where $c_{i,2}, \ldots, c_{i,r}$ are free parameters.

3.3. Differentiable Fractal Interpolation Functions

In the previous subsection 3.2, we have dealt with general (continuous) FIFs. Now, we shall restrict the previous results to the differentiable FIFs, which have some interesting properties. In [7] it is shown that the integral of a FIF defined from an IFS is also a FIF defined from a different IFS. A similar result can be proved for the case of the RFIFs (see [11]). This implies that one can construct RFIFs which are differentiable. In this specific case the term "fractal" is somewhat eccentric, since by definition a fractal object has a non-integer dimension, while a differentiable function has always dimension 1. It is used however as a reminder that this function is a product of a RIFS. The following theorem is an extension of the one presented in [7] (for the proof see [11]). For a function g we symbolize $g^{(0)} = g$, $g^{(1)} = g'$ and $g^{(k)}$ as its k-th order derivative. In addition, we consider $C^n([0,1])$ as the space of the functions that have continuous n-th order derivative, equipped with the norm $\|f\| = \|f\|_\infty + \|f'\|_\infty + \cdots + \|f^{(n)}\|_\infty$.

Theorem 3.1 ([11]). *Consider the RIFS $\{\mathbb{R}^2, w_{1-N}, P\}$, whose attractor is the graph of a RFIF associated with the data points Δ, $\hat{\Delta}$ and the labelling mapping $\mathbb{J}$, where*

$$w_i\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} L_i(x) \\ s_i y + p_i(x) \end{pmatrix},$$

$|s_i| < a_i^n$, $p_i \in C^n([0,1])$, for $i = 1, 2, \ldots, N$. We define $F_{k,i}(x,y) = \frac{s_i y + p_i^{(k)}(x)}{a_i^k}$. If for any $k = 0, 1, \ldots, n$, each one of the $2N \times 2N$ linear systems

$$y_{k,i-1} = F_{k,i}(\hat{x}_{j-1}, \hat{y}_{k,j-1}) = \frac{s_i \hat{y}_{k,j-1} + p_i^{(k)}(\hat{x}_{j-1})}{a_i^k},$$

$$y_{k,i} = F_{k,i}(\hat{x}_j, \hat{y}_{k,j}) = \frac{s_i \hat{y}_{k,j} + p_i^{(k)}(\hat{x}_j)}{a_i^k},$$

with $\hat{y}_{k,j} = y_{k,\mathbb{I}(j)}$, $\mathbb{I}(j) = i$ iff $\hat{x}_j = x_i$, $j = \mathbb{J}(i)$ $i = 1, 2, \ldots, N$, has a unique solution for $y_{k,i}$, then the RFIF $f \in C^n([0,1])$ and $f^{(k)}$ is the RFIF defined by the RIFS $\{\mathbb{R}^2;\ w_{k,1-N}, P\}$, where

$$w_{k,i} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} L_i(x) \\ F_{k,i}(x,y) \end{pmatrix}, \quad k = 1, \ldots, n.$$

The above theorem provides the means for the construction of differentiable fractal interpolation functions. In [23] the authors based on the results of [7] introduced the so called Hermite FIFs. This construction was extended in [11] to make use of RIFS instead of IFSs, based on the aforementioned theorems. It is also possible to generalize splines using FIFs. Chand and Kapoor in [12] gave a detailed description of the subject using IFSs. In the following, we shall make use of both these constructions. However, we chose to omit the details of the specific theorems to shorten this introduction. The interested reader can find all the details in the given references.

4. The space of FIFs and RFIFs as a RKHS

In this section we focus our attention on the construction of several Reproducing Kernel Hilbert Spaces (RKHSs) consisting of specific types of FIFs, which is the main result or contribution of this work. For the space of all FIFs of order r we present the following interesting general result:

Theorem 4.1. *Consider* $\mathcal{P} = \{x_0, \ldots, x_N\}$, $N \geq 2$, $\boldsymbol{s} = (s_1, \ldots, s_N)$, $r \geq 2$, $\boldsymbol{v} \in \mathbb{R}^N$ *be fixed. The space* $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ *of all RFIFs of order* r *defined on knots of* $\mathcal{P}$ *with contraction factors* $\boldsymbol{s}$ *and corresponding connection vector* $\boldsymbol{v}$ *($\boldsymbol{v}$ actually defines the function $\mathbb{J}$) is a Reproducing Kernel Hilbert Space (RKHS).*

Proof. The proof is based on the fact that $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ is a finite-dimensional linear space. Consider $f, \tilde{f} \in \mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ two FIFs of order r with corresponding interpolation points Δ, $\tilde{\Delta}$ (where $\hat{\Delta}$, $\hat{\tilde{\Delta}}$ are the corresponding subsets that define the regions) and $\lambda_1, \lambda_2 \in \mathbb{R}$. Then there are polynomials $p_i, \tilde{p}_i$ of order r (the polynomials of the corresponding RIFSs) such that $f(L_i(x)) = s_i f(x) + p_i(x)$ and $\tilde{f}(L_i(x)) = s_i \tilde{f}(x) + \tilde{p}_i(x)$ for $i = 1, \ldots, N$ (see relation (5)). In addition $p_i, \tilde{p}_i$ satisfy the join-up conditions (4). It is a matter of elementary algebra to deduce that

$$(\lambda_1 f + \lambda_2 \tilde{f})(L_i(x)) = s_i(\lambda_1 f + \lambda_2 \tilde{f})(x) + (\lambda_1 p_i + \lambda_2 \tilde{p}_i)(x),$$

where $\lambda_1 p_i + \lambda_2 \tilde{p}_i$ are polynomials of order r that satisfy the join up conditions:

$$(\lambda_1 p_i + \lambda_2 \tilde{p}_i)(\hat{x}_{\mathbb{J}(i)-1}) = (\lambda_1 y_{i-1} + \lambda_2 \tilde{y}_{i-1}) - s_i(\lambda_1 \hat{y}_{\mathbb{J}(i)-1} + \lambda_2 \hat{\tilde{y}}_{\mathbb{J}(i)-1}),$$
$$(\lambda_1 p_i + \lambda_2 \tilde{p}_i)(\hat{x}_{\mathbb{J}(i)}) = (\lambda_1 y_i + \lambda_2 \tilde{y}_i) - s_i(\lambda_1 \hat{y}_{\mathbb{J}(i)} + \lambda_2 \hat{\tilde{y}}_{\mathbb{J}(i)}),$$

8

for $i = 1, \ldots, N$. Hence, $g = \lambda_1 f + \lambda_2 \tilde{f}$ is a RFIF of order r that interpolates the points of the set $\Delta_g = \{(x_i, \lambda_1 y_i + \lambda_2 \tilde{y}_i), \ i = 0, \ldots, N\}$. The coefficients of the polynomials of the corresponding RIFSs are given by the relations $d_{i,k} = \lambda_1 c_{i,k} + \lambda_2 \tilde{c}_{i,k}$, for $i = 1, \ldots, N$, $k = 0, \ldots, r$, where $c_{i,k}$, $\tilde{c}_{i,k}$ are the coefficients of the polynomials of the RIFSs of f and $\tilde{f}$ respectively. Moreover, it is easy to deduce that $\dim(\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}) = r \cdot N + 1$, since each $f \in \mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ can be associated with a vector of $\mathbb{R}^{r \cdot N + 1}$ containing the y-coordinates of the interpolation points and the coefficients of the polynomials of the corresponding RIFS:

$$\Psi(f) = (y_0, \ldots, y_N, c_{1,2}, \ldots, c_{1,r}, c_{2,2}, \ldots, c_{2,r}, \ldots, c_{N,2}, \ldots, c_{N,r}).$$

Evidently Ψ is a one to one and onto mapping. The corresponding kernel depends on the selection of the inner product on $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ as described in theorem 2.1. A basis for $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ is the set of functions $\{e_i = \Psi^{-1}(\boldsymbol{e}_i), \ i = 1, \ldots, N \cdot r + 1\}$, where $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{N \cdot r + 1}\}$ is the standard basis of $\mathbb{R}^{N \cdot r + 1}$. If we choose the simple inner product defined by:

$$\langle f, \tilde{f} \rangle = \sum_{k=1, l=1}^{N \cdot r + 1} (\Psi(f))_k \cdot \left(\Psi(\tilde{f})\right)_l, \tag{9}$$

where for $\boldsymbol{u} \in \mathbb{R}^{N \cdot r + 1}$, $(\boldsymbol{u})_k$ is the k-th coordinate of $\boldsymbol{u}$, then the corresponding kernel is defined by

$$\kappa(x, y) = \sum_{k=1, l=1}^{N \cdot r + 1} e_k(x) e_l(y). \tag{10}$$

Choosing the inner product that induces the L_2-norm:

$$\langle f, \tilde{f} \rangle = \int_{x_0}^{x_N} f(x) \tilde{f}(x) dx, \tag{11}$$

we obtain the kernel:

$$\kappa(x, y) = \sum_{k=1, l=1}^{N \cdot r + 1} \beta_{k,l} e_k(x) e_l(y),$$

where $B = \{\beta_{k,l}\}$ is the inverse of $A = \{\alpha_{k,l}\}$, $\alpha_{k,l} = \int_0^1 e_k(x) e_l(x) dx$. $\qquad \square$

Remark 4.2. The computation of integrals of products of FIFs is not an elementary task. There are, however, formulas which can be used to compute integrals of this type through the r-order moments as it will be shown in section 5. More on the subject can be found in [3], [16], [8].

Accordingly, the following useful theorem can be deduced (the proof of which is similar to the previous one):

Theorem 4.3. *Consider* $\mathcal{P} = \{x_0, \ldots, x_N\}$, $N \geq 2$, $\boldsymbol{s} = (s_1, \ldots, s_N)$, $r \geq 2$, $\boldsymbol{v} \in \mathbb{R}^N$ *be fixed. The space* $\mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1} \subset C^r$ *of all Hermite RFIFs of order* $2r + 1$ *defined on knots of* $\mathcal{P}$ *with contraction factors* $\boldsymbol{s}$ *and corresponding connection vector* $\boldsymbol{v}$ *is a Reproducing Kernel Hilbert Space (RKHS). The same is true for the space* $\mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1} \subset C^{2r}$ *of all spline RFIFs of order* $2r + 1$ *(whenever this is well defined).*

9

Remark 4.4. $\mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$ is a linear space with dimension $(N+1)\cdot(r+1)$. The corresponding mapping $\Psi : \mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1} \to \mathbb{R}^{(N+1)\cdot(r+1)}$ is defined as follows:

$$\Psi(f) = (y_{0,0}, y_{0,1}, \ldots, y_{0,N}, y_{1,0}, y_{1,1}, \ldots, y_{1,N}, \ldots, y_{r,0}, y_{r,1}, \ldots, y_{r,N}),$$

where $y_{k,i} = f^{(k)}(x_i)$, for $k = 0, \ldots, r$, $i = 0, \ldots, N$.

Remark 4.5. Similarly, $\mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$ is a linear space with dimension $(N+1) + 2r$. The corresponding mapping $\Psi : \mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1} \to \mathbb{R}^{(N+1)+2r}$ is defined as follows:

$$\Psi(f) = (y_0, y_1, \ldots, y_N, y_{1,0}, \ldots, y_{r,0}, y_{1,N}, \ldots, y_{r,N}),$$

where $y_{k,0} = f^{(k)}(x_0)$, $y_{k,N} = f^{(k)}(x_N)$, $k = 1, \ldots, r$.

Remark 4.6. Consider the RKHS $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ with the inner product $\langle \cdot, \cdot \rangle$. The corresponding mapping Φ which maps any point of the original space (in our case $\mathbb{R}$) to the feature space $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ (see the discussion in section 2) is:

$$\Phi(x) = \sum_{k=1,l=1}^{N \cdot r+1} \beta_{k,l} e_k(x) e_l(\cdot),$$

where $B = \{\beta_{k,l}\}$ is the inverse of $A = \{\alpha_{k,l}\}$, $\alpha_{k,l} = \langle e_k, e_l \rangle$, for $k, l = 1, \ldots, N \cdot r + 1$. Similar relations hold for the other cases, i.e., $\mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ and $\mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$

Remark 4.7. Consider the RKHSs $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$, $\mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$, $\mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$ with the inner product that induces the L_2-norm $\langle \cdot, \cdot \rangle_{L_2}$ and their respective kernels $\kappa_{\mathcal{F}}$, $\kappa_{\mathcal{H}}$, $\kappa_{\mathcal{S}}$. In general $\kappa_{\mathcal{F}}$ is a fractal function (i.e., its graph has box-counting dimension greater than 2, except in the degenerative case where $\boldsymbol{s} = 0$), whereas $\kappa_{\mathcal{H}}$ is always a C^r function and $\kappa_{\mathcal{S}}$ is a C^{2r} function. Moreover, since $\mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r} \subset \mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ and $\|f\|_{\mathcal{H}} \geq \|f\|_{\mathcal{F}}$ (they are actually equal) for all $f \in \mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$, from a known result, proved in [2] can be stated that $\kappa_{\mathcal{H}} \ll \kappa_{\mathcal{F}}$ (i.e., $\kappa_{\mathcal{F}} - \kappa_{\mathcal{H}}$ is a positive definite kernel). Likewise we have that $\kappa_{\mathcal{S}} \ll \kappa_{\mathcal{H}}$ and, hence, $\kappa_{\mathcal{S}} \ll \kappa_{\mathcal{H}} \ll \kappa_{\mathcal{F}}$.

5. Computation of the Kernel functions

To compute the kernel κ (corresponding to any of the respective spaces $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$, $\mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$, $\mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$) at any point (x, y) one need to consider two basic steps: 1) the computation of the inner product $\langle e_k, e_l \rangle$ for any pair e_k, e_l in the corresponding basis, and 2) the computation of the values $e_k(x)$ and $e_k(y)$ for any e_k. For the latter the *Deterministic Iteration Algorithm - DIA* or *the Random Iteration Algorithm - RIA* can be applied (see [5] for more information on these well known algorithms). On the other hand, the treatment of the former step depends on the selection of the inner product of the corresponding space. For the simple inner product of the form (9) the computation is straightforward. However, if one selects the inner product that induces the L_2 norm (i.e., equation (11)) more effort is needed. The computation of integrals of products of affine FIFs was firstly addressed by Hardin, Kessler and Massopust in [16], based on results provided by Barnsley in [3], since it plays a crucial role in the construction of wavelets. Again, the main idea is to use the functional relation (5). Here we generalize this result for recurrent FIFs (RFIFs) of order r. We consider the case where the interpolation points are equidistant, i.e., all sections I_i have the same diameter $\delta = x_i - x_{i-1}$, $i = 1, \ldots, N$ and all intervals J_j have the same diameter $\psi = \hat{x}_j - \hat{x}_{j-1}$, $j = 1, \ldots, M$.

10

The first thing we need to do is to compute the m-th order moments of an arbitrary RFIF, for $m \in \mathbb{N}$. To this end, consider a RFIF f as defined in section 3.2 and define $P_{m,i} = \int_{x_{i-1}}^{x_i} x^m f(x)dx$, $\hat{P}_{m,j} = \int_{\hat{x}_{j-1}}^{\hat{x}_j} x^m f(x)dx$ and $P_m = \int_{x_0}^{x_N} x^m f(x)dx = \sum_{i=1}^{N} P_{m,i}$. In the following we denote $\mathcal{J}_j = \{i : I_i \subset J_j\}$:

$$P_{m,i} = \int_{x_{i-1}}^{x_i} x^m f(x)dx = \int_{x_{i-1}}^{x_i} x^m \left(s_i f(L_i^{-1}(x)) + p_i(L_i^{-1}(x)) \right) dx$$

$$= a_i \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} (a_i x + b_i)^m (s_i f(x) + p_i(x))dx$$

$$= \sum_{k=0}^{m} \binom{m}{k} a_i^{k+1} b_i^{m-k} \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} x^k (s_i f(x) + p_i(x))dx$$

$$= \sum_{k=0}^{m} \binom{m}{k} a_i^{k+1} b_i^{m-k} s_i \hat{P}_{k,\mathbb{J}(i)} + \sum_{k=0}^{m} \binom{m}{k} a_i^{k+1} b_i^{m-k} \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} x^k p_i(x)dx$$

$$= a_i^{m+1} s_i \hat{P}_{m,\mathbb{J}(i)} + \sum_{k=0}^{m-1} \binom{m}{k} a_i^{k+1} b_i^{m-k} s_i \hat{P}_{k,\mathbb{J}(i)} + \sum_{k=0}^{m} \binom{m}{k} a_i^{k+1} b_i^{m-k} \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} x^k p_i(x)dx$$

$$= a_i^{m+1} s_i \sum_{i' \in \mathcal{J}_{\mathbb{J}(i)}} P_{m,i'} + \sum_{k=0}^{m-1} \binom{m}{k} a_i^{k+1} b_i^{m-k} s_i \hat{P}_{k,\mathbb{J}(i)} + \sum_{k=0}^{m} \binom{m}{k} a_i^{k+1} b_i^{m-k} \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} x^k p_i(x)dx.$$

Assuming that all $P_{k,i}$, for $i = 1, \ldots, N$ and $k = 1, \ldots, m-1$ are known, this is a linear system of the form:

$$\left(I - diag(a_i^{m+1} s_i) \cdot C \right) \cdot \boldsymbol{x} = \boldsymbol{u}, \tag{12}$$

where C is the connection matrix of the respective RIFS, $\boldsymbol{x} = (P_{m,1}, \ldots, P_{m,i}, \ldots, P_{m,n})^T$ and

$$\boldsymbol{u} = \left(\ldots, \sum_{k=0}^{m-1} \binom{m}{k} a_i^{k+1} b_i^{m-k} s_i \hat{P}_{k,\mathbb{J}(i)} + \sum_{k=0}^{m} \binom{m}{k} a_i^{k+1} b_i^{m-k} \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} x^k p_i(x)dx, \ldots \right)^T.$$

Following a similar procedure for the 0-th order moments we obtain the linear system:

$$(I - diag(a_i s_i) \cdot C) \cdot \boldsymbol{x} = \boldsymbol{u}, \tag{13}$$

where $\boldsymbol{x} = (P_{0,1}, \ldots, P_{0,n})^T$ and

$$\boldsymbol{u} = \left(\ldots, a_i \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} p_i(x)dx, \ldots \right)^T.$$

Proposition 5.1. *Each one of the linear systems (12) admits a unique solution, i.e.*

$$\det \left(I - diag(a_i^{m+1} s_i) \right) \cdot C) \neq 0,$$

for all $m \in \mathbb{N}$.

Proof. To prove this claim we will use some properties of square matrices, which we state in the following for the reader's convenience. Consider two $N \times N$ matrices A, B, then:

- $\rho(A) \leq \|A\|$, where $\rho(A)$ is the spectral radius of A (i.e., $\rho(A) = \max\{|\lambda_i|\}$, where λ_i is the eigenvalues of A and $\| \cdot \|$ any matrix-norm.

- $\|A \cdot B\| \leq \|A\| \cdot \|B\|$, for any matrix-norm.

- $\|A\|_p = \sup\{\|Ax\|_p/\|x\|_p\}$, for $1 \leq p < \infty$.

- $\|A\|_\infty = max_{1 \leq i \leq N} \left\{ \sum_{j=1}^{N} |A_{i,j}| \right\}$.

In our case we take:

$$\rho\left(diag(a_i^{m+1}s_i) \cdot C\right) \leq \left\|diag(a_i^{m+1}s_i) \cdot C\right\|_\infty \leq \left\|diag(a_i^{m+1}s_i)\right\|_\infty \cdot \|C\|_\infty,$$

with $\|diag(a_i^{m+1}s_i)\|_\infty = \max_{1 \leq i \leq N}\{|a_i|^{m+1}|s_i|\}$ and $\|C\|_\infty = \max_{1 \leq i \leq N}\{r(i)\}$, where $r(i) = \mathcal{N}\left(\mathcal{J}_{\mathbb{J}(i)}\right)$ (i.e., the number of elements of the set $\mathcal{J}_{\mathbb{J}(i)}$). Clearly, in the case of equidistant interpolation points $r(i) = \psi/\delta$, $a_i = \delta/\psi$, thus $\rho\left(diag(a_i^{m+1}s_i) \cdot C\right) \leq 1$ for all $m \in \mathbb{N}$. Therefore, since the number 1 is greater than the greatest eigenvalue of $diag(a_i^{m+1}s_i) \cdot C$, we conclude that $\det\left(I - \left(diag(a_i^{m+1}s_i)\right) \cdot C\right) \neq 0$, for all $m \in \mathbb{N}$. $\square$

The above equations describe an inductive procedure. We first solve the initial linear system and compute $P_{0,j}$, $j = 1, \ldots, M$. Then we use these values to solve the next linear system (for $m = 1$) and compute $P_{1,i}$, $i = 1, \ldots, N$ and continue until we reach $m = 2r + 1$. To evaluate P_m we simply sum $P_{m,i}$ for all $i = 1, \ldots, N$.

Now, we turn our efforts to the computation of the inner product of the form (11). Consider $f, \tilde{f} \in \mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$, where $P_{m,i}$ and $\tilde{P}_{m,i}$ for $m = 0, \ldots, 2r + 1$, $i = 1, \ldots, N$, are the corresponding moments as defined above and $R_i = \int_{x_{i-1}}^{x_i} f(x)\tilde{f}(x)dx$. Working similarly we take:

$$R_i = \int_{x_{i-1}}^{x_i} f(x)\tilde{f}(x)dx$$

$$= \int_{x_{i-1}}^{x_i} \left(s_i f\left(L_i^{-1}(x)\right) + p_i\left(L_i^{-1}(x)\right)\right)\left(s_i\tilde{f}\left(L_i^{-1}(x)\right) + \tilde{p}_i\left(L_i^{-1}(x)\right)\right) dx$$

$$= a_i \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} \left(s_i f(x) + p_i(x)\right)\left(s_i\tilde{f}(x) + \tilde{p}_i(x)\right) dx$$

$$= a_i s_i^2 \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} f(x)\tilde{f}(x)dx + a_i s_i \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} f(x)\tilde{p}_i(x)dx +$$

$$+ a_i s_i \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} \tilde{f}(x)p_i(x)dx + a_i \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} p_i(x)\tilde{p}_i(x)dx$$

$$= a_i s_i^2 \sum_{i' \in \mathcal{J}_{\mathbb{J}(i)}} R_{i'} + a_i s_i \sum_{i' \in \mathcal{J}_{\mathbb{J}(i)}} \sum_{k=0}^{2r+1} \tilde{c}_{i',k} P_{k,i'} + a_i s_i \sum_{i' \in \mathcal{J}_{\mathbb{J}(i)}} \sum_{l=0}^{2r+1} c_{i',l}\tilde{R}_{l,i'} + a_i \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} p_i(x)\tilde{p}_i(x)dx,$$

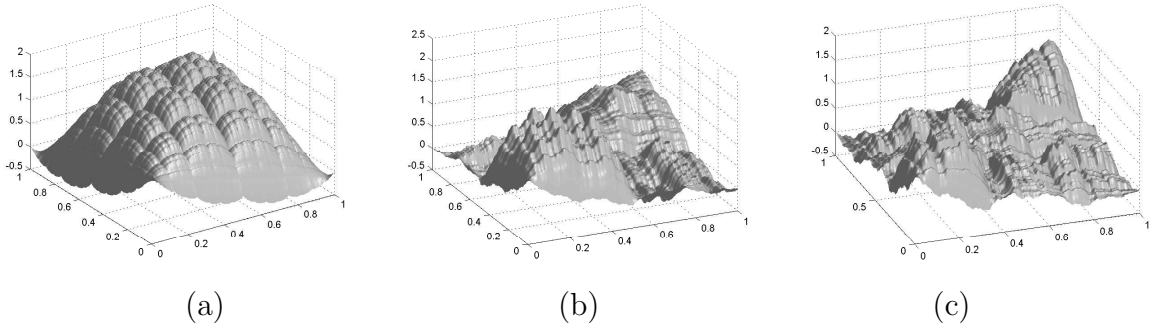Figure 1: The kernel function of the space $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ for (a) $r = 1$, $N = 2$, $\mathcal{P} = \{0, \frac{1}{2}, 1\}$, $\boldsymbol{v} = (1,1)$, $\boldsymbol{s} = (0.5, 0.5)$, (b) $r = 1$, $N = 2$, $\mathcal{P} = \{0, \frac{1}{2}, 1\}$, $\boldsymbol{v} = (1,1)$, $\boldsymbol{s} = (0.5, -0.5)$, (c) $r = 5$, $N = 3$, $\mathcal{P} = \{0, \frac{1}{3}, \frac{2}{3}, 1\}$, $\boldsymbol{v} = (1,1,1)$, $\boldsymbol{s} = (0.4, -0.4, 0.3)$. In all cases the space consists of self-similar FIFs.

for $i = 1, \ldots, N$ (where $c_{i,k}$, $\tilde{c}_{i,l}$ are the coefficients of the polynomials p_i and $\tilde{p}_i$ respectively). Once more, we obtain a $N \times N$ linear system (similar with the one defined for the moments) which has to be solved for R_i, $i = 1, \ldots, N$:

$$\left(I - diag(a_i s_i^2) \cdot C\right) \cdot \boldsymbol{x} = \boldsymbol{u}, \tag{14}$$

where C is the connection matrix of the respective RIFS, $\boldsymbol{x} = (R_1, \ldots, R_i, \ldots, R_n)^T$ and

$$u = \left(\ldots, a_i s_i \sum_{i' \in \mathcal{J}_{\mathbb{J}(i)}} \sum_{k=0}^{2r+1} \tilde{c}_{i',k} P_{k,i'} + a_i s_i \sum_{i' \in \mathcal{J}_{\mathbb{J}(i)}} \sum_{l=0}^{2r+1} c_{i',l} \tilde{R}_{l,i'} + a_i \int_{\hat{x}_{\mathbb{J}(i)-1}}^{\hat{x}_{\mathbb{J}(i)}} p_i(x) \tilde{p}_i(x) dx, \ldots\right).$$

The following proposition can be proved similarly to proposition 5.1.

Proposition 5.2. *The linear system (14) admits a unique solution, i.e.*

$$\det \left(I - diag(a_i s_i^2)\right) \cdot C) \neq 0.$$

Remark 5.1. The inner product (11) is given as the sum of R_i for all i.

To demonstrate the results of the procedure we included some examples. Figures 1 and 2 show the kernel function that corresponds to several spaces of fractal interpolation functions.

6. Conclusions

In this work we consider the spaces of simple and recurrent FIFs of arbitrary order r, denoted by $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$, and prove that they constitute RKHSs. Moreover, we specifically prove that several spaces of differentiable (R)FIFs (spline, denoted by $\mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$, and Hermitian, denoted by $\mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}$), which are special cases of (R)FIFs, are also RKHSs. We calculate the dimensions of these spaces, i.e.,

$$\dim\left(\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}\right) = N \cdot r + 1,$$
$$\dim\left(\mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}\right) = (N+1) \cdot (r+1),$$
$$\dim\left(\mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1}\right) = N + 1 + 2r,$$
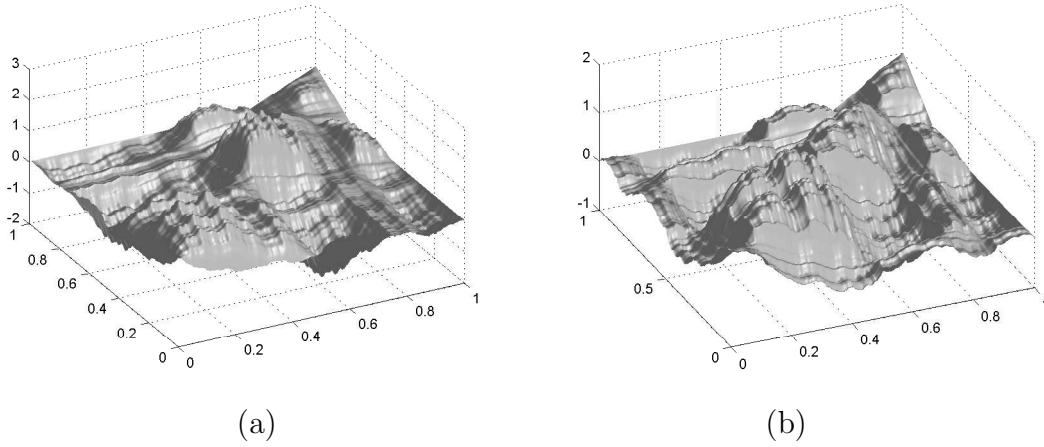
13

Figure 2: The kernel function of the space $\mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ for (a) $r = 1$, $N = 4$, $\mathcal{P} = \{0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1\}$, $\boldsymbol{v} = (1, 2, 1, 2)$, $\boldsymbol{s} = (0.6, -0.4, -0.6, 0.4)$, (b) $r = 2$, $N = 4$, $\mathcal{P} = \{0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1\}$, $\boldsymbol{v} = (2, 1, 1, 2)$, $\boldsymbol{s} = (-0.4, 0.6, -0.4, 0.5)$. In both cases the space consists of RFIFs.

where $N + 1$ is the number of the interpolation points and r (or $2r + 1$ respectively) the order of the corresponding space, and provide the tools for the calculation of the respective kernel functions κ and the corresponding induced functions Φ. Moreover, we prove that the relation $\mathcal{S}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1} \subset \mathcal{H}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},2r+1} \subset \mathcal{F}_{\mathcal{P},\boldsymbol{v},\boldsymbol{s},r}$ holds and thus $\kappa_{\mathcal{S}} \le \kappa_{\mathcal{H}} \le \kappa_{\mathcal{F}}$. Therefore, new, broad and rich families of kernel functions are introduced to the powerful tool of RKHSs and, besides, the methods and results of RKHSs can be applied directly to the spaces of (R)FIFs.

References

[1] N. Aronszajn, La théorie générale des noyaux reproduisants et les applications, Premier Partie, Proc. Cambridge Philos. soc. Math. France vol 67 (1939).

[2] N. Aronszajn, Theory of Reproducing Kernels, Transactions of the American Mathematical Society, vol. 68, no. 3, 1950, pp 337–404.

[3] M. F. Barnsley, Fractal functions and interpolation, Constr. Approx., 2 (1986), 303–329.

[4] M. F. Barnsley, J. H. Elton and D. P. Hardin, Recurrent iterated function systems, Constr. Approx., 5 (1989), 3–31.

[5] M. F. Barnsley, Fractals everywhere, 2nd ed., Academic Press Professional, 1993.

[6] M. F. Barnsley, S. Demko, Iterated function systems and the global construction of fractals, Proc. Roy. Soc. London Ser A, 399 (1985), 243–275.

[7] M. F. Barnsley, A. N. Harrington, The calculus of Fractal Interpolation Functions, J. Approx. Theory, (57),14–43,1989.

[8] P. Bouboulis, Construction of orthogonal multi-wavelets using generalized-affine fractal interpolation functions, IMA Journal of Applied Mathematics, Vol. 74 no. 6 (2009), 904–933.

[9] P. Bouboulis, L. Dalla, A general construction of Fractal Interpolation Functions on grids of $\mathbb{R}^n$, European J. Appl. Math., 18, 449–476, 2007.

[10] P. Bouboulis, L. Dalla, V. Drakopoulos, Construction of Recurrent Bivariate Fractal Interpolation Surfaces and Computation of their Box-Counting Dimension, J. Approx. Theory, 141 (2006), 99–117.

[11] P. Bouboulis, L. Dalla, M. Kostaki-Kosta, Construction of Smooth Fractal Surfaces Using Hermite Fractal Interpolation Functions, Bull. Greek. Math. Soc., 2007.

[12] A. K. B. Chand, G. P. Kapoor, Generalized cubic spline fractal interpolation functions, SIAM J. Numer. Anal., vol 44, no 2, 2006, pp. 655-676.

[13] T. M. Cover, "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition", IEEE Trans. Electronic Computers, vol. 14, pp. 326–334, 1965.

[14] K. Falconer, Fractal Geometry. Mathematical Foundations and Applications, 2nd ed., Wiley, 2004.

[15] F. R. Gantmacher, The Theory of Matrices, vol. II, AMS Chelsea Publishing, 1998.

[16] D. P. Hardin, B. Kessler, P. R. Massopust, Multiresolution Analysis Based on Fractal Functions, J. Approx. Theory, 71 (1992), 104–120.

[17] J. Hutchinson, Fractals and self-similarity, Indiana University Mathematics Journal 30(5), 1981, pp. 713–747.

[18] T. Kailath, RKHS Approach to Detection and Estimation Problems - Part I: Deterministic Signals in Gaussian Noise, IEEE Trans. On Information Theory, 1971, 17, 530-549 .

[19] T. Kailath, R. T. Geesey, H. L. Weinert, Some Relations Among RKHS Norms, Fredholm Equations, and Innovations Representations, IEEE Trans. On Information Theory, 1972, 18, 341–348.

[20] T. Kailath, D. Duttweiler, An RKHS Approach to Detection and Estimation Problems - Part III: Generalized Innovations Representations and a Likelihood-Ratio Formula, IEEE Trans. On Information Theory, 1972, 18, 730-745.

[21] J. Mercer, Sturm-Liuville series of normal functions in the theory of integral equations, Philos. Trans. Roy. Soc. London, Ser. A vol. 211 (1911), 111–198.

[22] E. H. Moore, On properly positive Hermitian matrices, Bull. Amer. Math. Soc., vol. 23 (1916).

[23] M. A. Navascues and M. V. Sebastian, Generalization of Hermite functions by fractal interpolation, J. Approx. Theory, (131), 19–29, 2004.

[24] J. Shawe-Taylor and N. Cristianini, Kernel methods for pattern analysis, Cambridge University Press, 2004.

[25] B. Schölkopf and A. J. Smola, Learning with Kernels, MIT Press, 2002.

[26] B. Schölkopf, A. J. Smola and K. R. Muller, Kernel principal component analysis, Lecture notes in computer science, (1327), 583–588, 1997.

[27] S. Theodoridis and K. Koutroumbas, Pattern Recognition, Academic Press, 2008.

[28] V. N. Vapnik, Statistical Learning Theory, Wiley, NY, 1998.

[29] G. Wahba, Spline Models for Observational Data Series in Applied Mathematics, Vol. 59, SIAM, 1990.