RESEARCH

Real Analysis Exchange
Vol. 37(1), 2011/2012, pp. 61–82

Francesco Calabrò, DAEIMI & LAN, Università di Cassino, Via G. Di Biasio 43 03043 Cassino (FR) Italy. email: calabro@unicas.it

Antonio Corbo Esposito, DAEIMI & LAN, Università di Cassino, Via G. Di Biasio 43 03043 Cassino (FR) Italy. email: corbo@unicas.it

Carmen Perugia, DSGA, Università del Sannio, via dei Mulini 59/A, 82100 Benevento Italy. email: cperugia@unisannio.it

BINOMIAL MEASURES AND THEIR APPROXIMATIONS

Abstract

In this paper we consider the properties of a family of probability (continuous and singular) measures, which will be called Binomial measures because of their relationship with the binomial model in probability. These measures arise in many applications with different notations. Many properties in common with Lebesgue measure hold true for this family, sometimes unexpectedly.

1 Introduction

In this paper we consider the properties of a family of probability measures $\{\mu_\alpha\}_{0<\alpha<1}$ on an interval of the real line ($[0,1]$, for the sake of simplicity) characterized by the following self similarity property. Let I be a dyadic subinterval and let I be bisect into the left and right parts $I = I_L \cup I_R$; then

$$\mu_\alpha(I_R) = \alpha\mu_\alpha(I) \ . \tag{1}$$

When $\alpha = 1/2$ we trivially obtain the Lebesgue measure on $[0,1]$, while in all the other cases we obtain continuous and singular measures, such that $\mu_\alpha(J) > 0$ for every J subinterval of $[0,1]$. For this family many properties in common with the Lebesgue measure hold true, sometimes unexpectedly.

Mathematical Reviews subject classification: Primary: 28A25, 28A80; Secondary: 65D32
Key words: self similar measures, quadrature
Received by the editors April 26, 2010
Communicated by: Emma D'Aniello

This paper intends to highlight several of these properties, since they are often essential for applications. We do not intend present the results in the most general way, since they are often partly present in different environments, all of which are equally important for applications and use completely different notation (cf., for example, [8] for some properties of balanced measures, [10] for integrals used in the computation of wavelet coefficients and refinable functionals, and [5] for Large Deviation Principles).

The measures that satisfy property (1) are strictly related to, but are topologically different from, the so-called Bernoulli trial measures, as considered in [19, 14, 13]. Indeed, it can be easily seen by applying equation (1) k times, that $\mu_\alpha\left(\left[\frac{j}{2^k}, \frac{j+1}{2^k}\right[\right)$ is exactly the probability that in the first k trials of the Bernoulli process we have a number of occurrences equal to the number of 1s in the binary expansion of j (see, too, equation (9)). Therefore, by analogy with the binomial process, we will call these measures binomial, see [2, 3].

The main tools we use in this paper are two different types of approximation for a measure of the family. The first is obtained by absolutely continuous measures with constant density on dyadic intervals whereas the second is obtained from a suitable finite linear combination of Dirac delta measures that is often presented in literature as quadrature formula. The latter was introduced in order to obtain a more rapidly convergent approximation.

In the next section, we consider different environments where the measures μ_α are notable examples. In Section 3 we consider the step constant approximation and prove that this converges both in the weak-star topology and in a natural norm (see Proposition 3.1). With this tool, we prove some basic properties of the measure in Proposition 3.2. In Section 3.1 the strongly mixing property is reconsidered together with a new result on the characterization of the set in which the measures are concentrated (see Propositions 3.4 and 3.5). Finally, in Section 4 we introduce the approximation considered when the numerical procedures are written, as a combination of Dirac measures constructed in order to satisfy some moment equations. In this case it is possible to introduce error estimates. Section 4 also considers the combination of the two approximations leading to the procedure known in numerical analysis as the *composite rule*. The final proposition gives an explicit estimate for the rate of convergence of the latter approximation.

2 Binomial measures

Let us introduce some notation. For latter convenience, we will define a dyadic interval X_j^k as follows:

$$X_j^k \equiv \left[\frac{j}{2^k}, \frac{j+1}{2^k}\right[\quad \forall j,k \ s.t.\ k \in \mathbb{N}\,,\ 0 \le j \le 2^k - 1\ .$$

The integer k will be called the order of the interval. It should be noticed that the set of all possible dyadic intervals is a semi algebra in $[0,1[$ which generates the σ-algebra $\mathcal{B}$ of the Borel subsets of $[0,1[$ (for definitions see [8, 18]).

Following [8, P. 2.5], we will denote by M^1 the set of Borel regular probability measures on $\mathbb{R}$ supported in $[0,1]$. If $\nu \in M^1$, we will denote $\nu(\phi) \equiv \int_0^1 \phi\, d\nu$ for every $\phi \in L^1_\nu$. We introduce the following distance on M^1:

$$L(\nu_1,\nu_2) = \sup\{\nu_1(\phi) - \nu_2(\phi)\ s.t.\ \phi : [0,1] \to \mathbb{R}\,,\ Lip(\phi) \le 1\}\,, \qquad (2)$$

where $\phi(x)$ is a Lipschitz function and $Lip(\phi)$ is its Lipschitz constant.

Theorem 2.1. *There is a unique measure $\mu_\alpha \in \mathrm{M}^1$ which satisfies the following equivalent conditions:*

i

$$\mu_\alpha(E) = (1-\alpha)\mu_\alpha(\mathrm{S}_1^{-1}(E)) + \alpha\mu_\alpha(\mathrm{S}_2^{-1}(E)) \quad \textit{for every } E \in \mathcal{B}\ , \qquad (3)$$

where

$$\mathrm{S}_1 = x/2 \ \textit{and}\ \mathrm{S}_2 = x/2 + 1/2\,.$$

ii

$$\alpha\mu_\alpha(X_j^k) = \mu_\alpha(X_{2j}^{k+1}) \qquad \forall k \in \mathbb{N}\,,\ j = 1,\dots,2^k - 1\ . \qquad (1')$$

iii

$$\mu_\alpha(X_j^k) = \alpha^{n(j)}(1-\alpha)^{k-n(j)} \qquad (4)$$

where:

$$n(j) = \#\{1s\ \textit{of the binary expansion of}\ j\}\ .$$

iv

$$\int_0^1 f(x)\, d\mu_\alpha = (1-\alpha)\int_0^1 f\left(\frac{x}{2}\right) d\mu_\alpha + \alpha\int_0^1 f\left(\frac{x}{2}+\frac{1}{2}\right) d\mu_\alpha\ , \quad \forall f \in C^0([0,1])\ . \qquad (5)$$

Proof. Firstly, we will prove that a measure which satisfies equation (3) exists and is unique. Let us call

$$\rho_\alpha = [(1-\alpha),\alpha]\ , \qquad \alpha \in (0,1)$$

and denote by $(S,\rho_\alpha) : M^1 \to M^1$ the application defined by the following equation:

$$(S,\rho_\alpha)(\nu)(E) = (1-\alpha)\nu(S_1^{-1}(E)) + \alpha\nu(S_2^{-1}(E)) \quad \text{for every } E \in \mathcal{B}\ .$$

¿From [8] (Theorem 1 Section 4.4) we obtain that (S,ρ_α) is a contraction with respect to the metric L defined in (2) and therefore has a unique fixed point for every $\alpha \in (0,1)$.
We shall now prove the following implications (1') $\iff$ (4) ; (4) $\Rightarrow$ (3) ; (3) $\Rightarrow$ (1'); (3) $\iff$ (5).
The equivalence between (1') and the relationship (4) is simply proved by induction in one way and by substitution in the other.
Let us prove (4) $\Rightarrow$ (3). In order to verify (3) we can limit ourselves to dyadic intervals because these generate the σ-algebra of Borel sets. Now, we explicitly notice that $\forall k > 1$:

$$S_1^{-1}(X_j^k) = \begin{cases} \left[\frac{j}{2^{k-1}}, \frac{j+1}{2^{k-1}}\right[& if\ 1 \le j < 2^{k-1} \\ \emptyset & otherwise \end{cases}$$

$$S_2^{-1}(X_j^k) = \begin{cases} \emptyset & if\ 1 \le j < 2^{k-1} \\ \left[\frac{j}{2^{k-1}} - 1, \frac{j+1}{2^{k-1}} - 1\right[& otherwise \end{cases} .$$

Therefore we obtain the required relationship by substitution.
In order to prove that (3) $\Rightarrow$ (1'), we simply note that $S_1^{-1}(X_{2j}^{k+1}) = \emptyset$ and thus $\mu_\alpha(S_1^{-1}(X_{2j}^{k+1})) = 0$.
Finally, the equivalence between (5) and (3) can be proved by observing the following: that (3) is trivially equivalent to (5) for step functions; that step functions can be easily approximated by continuous ones (and viceversa) and that the convergence of integrals is ensured by Lebesgue dominated convergence theorem. □

It should be noticed that equation (5) relates the measures μ_α with the multifractal properties and also indicates that binomial measures are examples of Iterated Function Systems (IFS) measures, see [6, 16] for an introduction and [1] for further developments. In particular, following the notations in [11, 12], μ_α is the attractor of the δ-homogeneous linear IFS balanced measure

with contraction ratio 1/2, probabilities ρ_α and fixed points $\{0,1\}$.
Let us now recall the definition of refinable linear functionals as given in [10] and see how the measures μ_α can be put in this framework. We indicate with P the set of polynomials $p(x)$ with real coefficients.

Definition 2.2. *A linear functional $L : \mathrm{P} \to \mathbb{R}$ is called refinable if there are a $(N+1)$-uple of positive real numbers $(\gamma_0, \ldots, \gamma_N)$ called a mask such that:*

1. $$L[f] = \sum_{j=0}^{N} \frac{\gamma_j}{2} L\left[f\left(\frac{x+j}{2}\right)\right] ;$$
2. $$L[e_0(x)] = 1 \ , \ \text{where } e_0(x) \equiv 1.$$

In [10] (see the Remark in Section 3) it is proved that the functional is positive definite; i.e., $L[f] > 0$ whenever $f \in P$ is nonnegative everywhere and positive on a set of length greater than 0.

Let us consider the case $N = 1$. Property (1) of the definition is concerned only with the properties of the function in $[0,1]$, and for this reason we will consider the polynomials as functions on $[0,1]$. Applying Riesz theorem, the functional L acts as integration w.r.t. a positive Borel regular measure μ. Moreover, from the properties of the definition, μ turns out to be a probability measure with support in $[0,1]$. Thus, we can write:

$$L[f] = \int_0^1 f\, d\mu = 1/2 \sum_{j=0}^{1} \gamma_j \int_0^1 f(E_j(x))\, d\mu \ , \qquad E_j(x) = \frac{x}{2} + \frac{j}{2}$$

In particular, in equation (5), the only choice of positive linear refinable functional with $N = 1$ turns out to be integration w.r.t. a binomial measure, and the possible masks are $(2(1-\alpha), 2\alpha)$ $\forall \alpha \in (0,1)$.
This property relates our family of measures with the study of refinable functions used, for example, in wavelet methods.

3 Approximating measures: measures with constant densities on dyadic level k

For a general $\nu \in M^1$ we can consider an approximating measure ν_k defined as the unique measure which is absolutely continuous with respect to Lebesgue measure and has constant density on the dyadic intervals of level k satisfying:

$$\nu_k(X_j^k) = \nu(X_j^k) \ , \ \forall j = 0, \ldots, 2^k - 1 \ . \tag{6}$$

We observe that:

$$\nu(X_j^k) = \nu_h(X_j^k) = \nu_k(X_j^k)\,,\ \forall h \geq k\ . \tag{7}$$

For these measures we can prove the following proposition:

Proposition 3.1. *Take $\nu \in \mathrm{M}^1$ and define ν_k as in equation* (6)*. Then we have that $\nu_k \to \nu$ in the metric defined in* (2).

PROOF. Let ϕ be in $C^0([0,1])$, take $x_j^k \in X_j^k$ and observe that:

$$\left|\int_0^1 \phi(x)\,d\nu - \int_0^1 \phi(x)\,d\nu_k\right| \leq \sum_{j=0}^{2^k-1}\left|\int_{X_j^k}\phi(x)\,d\nu - \int_{X_j^k}\phi(x)\,d\nu_k\right| =$$

$$= \sum_{j=0}^{2^k-1}\left|\int_{X_j^k}[\phi(x)-\phi(x_j^k)+\phi(x_j^k)]\,d\nu - \int_{X_j^k}[\phi(x)-\phi(x_j^k)+\phi(x_j^k)]\,d\nu_k\right| =$$

$$= \sum_{j=0}^{2^k-1}\left|\int_{X_j^k}[\phi(x)-\phi(x_j^k)]\,d\nu - \int_{X_j^k}[\phi(x)-\phi(x_j^k)]\,d\nu_k\right| \leq$$

$$\leq \sum_{j=0}^{2^k-1}\int_{X_j^k}\left|\phi(x)-\phi(x_j^k)\right|\,d\nu + \int_{X_j^k}\left|\phi(x)-\phi(x_j^k)\right|\,d\nu_k\ . \tag{8}$$

Now, apply definition (2) and equation (8) so we have:

$$L(\nu,\nu_k) = \sup\{\nu(\phi)-\nu_k(\phi)|\phi:[0,1]\to\mathbb{R}, Lip(\phi)\leq 1\}\ \leq$$
$$\leq\ \sup\left\{\sum_{j=0}^{2^k-1}\left[\int_{X_j^k}\left|\phi(x)-\phi(x_j^k)\right|d\nu+\int_{X_j^k}\left|\phi(x)-\phi(x_j^k)\right|d\nu_k\right],\ Lip(\phi)\leq 1\ \right\}.$$

Now we use the Lipschitz condition on the function $\phi(x)$ on each X_j^k and we obtain that $\left|\phi(x)-\phi(x_j^k)\right| \leq \left|x - x_j^k\right| \leq 1/(2^k)$ for $x, x_j^k \in X_j^k$ and thus:

$$L(\nu,\nu_k) \leq \frac{1}{2^k}\sum_{j=0}^{2^k-1} 2\nu(X_j^k) = \frac{1}{2^{k-1}}\ .$$

Therefore the distances converge to 0, as required. □

We can see that, as has already been noticed in [8], the convergence in the L-distance implies the usual measure convergence, commonly denoted as weak convergence, or convergence in the weak-star topology.

When we apply this iterative construction to the measures μ_α, $\mu_{\alpha,k}$ denotes the approximating measure with constant density on dyadic intervals of order k. It can be easily seen, with the notations in theorem 2.1, that this density is:

$$f_{\alpha,k}(x) = 2^k \alpha^{n(j)} (1-\alpha)^{k-n(j)}\,, \qquad x \in X_j^k\,, \quad j = 0 \dots 2^k - 1 \tag{9}$$

In the following we will prove some useful properties of the measures μ_α. Some of these properties are known in a more general framework but here we will give more direct proofs.

Proposition 3.2. *The following hold true:*

1. *Continuity - Fixed $\alpha \in (0,1)$, the measure μ_α is such that $\mu_\alpha(\{x\}) = 0\ \forall x \in [0,1]$.*
2. *Factorization - We have:*
$$\mu_\alpha(X^{k+h}_{2^h j+i}) = \mu_\alpha(X^k_j)\mu_\alpha(X^h_i)$$
$\forall i,j$ such that $i = 0 \dots 2^h - 1, j = 0 \dots 2^k - 1$.
3. *Change of variables - Let f be in $L^1_{\mu_\alpha}$. Then, for each dyadic interval X^k_j we have:*
$$\int_0^1 f(x)\, d\mu_\alpha = \frac{1}{\mu_\alpha(X^k_j)} \int_{X^k_j} f(2^k x - j)\, d\mu_\alpha\ .$$
4. *Recursive calculation of moments - Moments of the measures μ_α are connected by the following recursive relationship:*
$$\int_0^1 x^s\, d\mu_\alpha = \frac{\alpha}{2^s - 1} \sum_{q=1}^{s} \left[\binom{s}{q} \int_0^1 x^{s-q}\, d\mu_\alpha \right] .$$
Moreover, the following holds true:
$$\int_{X^k_j} x^s\, d\mu_\alpha = \frac{\mu_\alpha(X^k_j)}{(2^k)^s} \sum_{q=0}^{s} \binom{s}{q} j^q \int_0^1 x^{s-q}\, d\mu_\alpha\ .$$

PROOF. 1. This is easily proved taking $x \in [0,1]$ and calling J_n the dyadic interval of level n such that $x \in J_n$. Now, $\cap_{n=0}^{\infty} J_n = \{x\}$ thus $\mu_\alpha(\{x\}) = lim_n \mu_\alpha(J_n)$. Call now $\lambda = max\{\alpha, 1-\alpha\}$, we have $\lambda < 1$. Now, from the definition of the measures $\mu_{\alpha,k}$ we have $\mu_\alpha(J_n) = \mu_{\alpha,n}(J_n)$. But by (9)$\mu_{\alpha,n}(J_n) \le \lambda^n$, and thus:

$$\mu_\alpha(\{x\}) = lim_n \mu_\alpha(J_n) \le lim_n \lambda^n = 0$$

as requested.

2. First consider that by (9) we have:

$$\mu_{\alpha,k+h}(X^{k+h}_{2^h j+i}) = \alpha^{n(j2^h+i)}(1-\alpha)^{k+h-n(j2^h+i)} \ .$$

Now, since i runs from $0 \ldots 2^h - 1$ we have:

$$n(j2^h + i) = n(j2^h) + n(i) = n(j) + n(i) \ .$$

Thus:

$$\alpha^{n(j2^h+i)}(1-\alpha)^{k+h-n(j2^h+i)} = [\alpha^{n(j)}(1-\alpha)^{k-n(j)}][\alpha^{n(i)}(1-\alpha)^{h-n(i)}] = \\ = \mu_{\alpha,k}(X^k_j)\mu_{\alpha,h}(X^h_i) \ .$$

Summarizing:

$$\mu_{\alpha,k+h}(X^{k+h}_{2^h j+i}) = \mu_{\alpha,k}(X^k_j)\mu_{\alpha,h}(X^h_i) \ .$$

By equation (7) we immediately get the conclusion.

3. For this result see Lemma 2.3 [2]. Note that the result for the change of variables cannot be extended to any affine change of variables.

4. The calculation of these moments can be done following [11]. For the calculation of the same in the dyadic intervals we can combine these relationships with the previous change of variables.

□

It should be noticed that, due to continuity, closed dyadic intervals have the same μ_α-measure of the one sided open intervals, and we will sometimes use this property implicitly. It should also be noticed that the self-similarity property is taken as a characterization of the measure μ_α in [6].

3.1 Properties related to ergodicity

In this section, we will give an explicit description of a set F^α_∞ so that the measure μ_α is concentrated on it. In order to achieve this aim we need some properties of μ_α related to ergodicity.
We denote the binary shift as T:

$$T : x \in [0,1] \to [0,1] \ , \quad T(x) = 2x - [2x]$$

where we denote the integer part as $[\cdot]$. It is known that T is ergodic with respect to μ_α for every $\alpha \in (0,1)$. Here we give a proof that T is also strongly mixing in the same set. Let us first recall some definitions.

Definition 3.3. *i. A measurable transformation $T : [0,1] \to [0,1]$ is measure preserving if $\nu(T^{-1}(B)) = \nu(B)$ for all $B \in \mathcal{B}$.*

ii. A measure preserving transformation T of $([0,1], \mathcal{B}, \nu)$ is ergodic if the sole sets $B \in \mathcal{B}$ such that $T^{-1}(B) = B$ are such that $\nu(B) = 0$ or $\nu(B) = 1$.

iii. A measure preserving transformation T of $([0,1], \mathcal{B}, \nu)$ is strongly mixing whenever $\forall\, B_1, B_2 \in \mathcal{B}$

$$\lim_{n\to\infty} \nu\left(T^{-n}(B_1) \cap B_2\right) = \nu(B_1)\nu(B_2)$$

It can be easily seen that strongly mixing transformations are ergodic (see [18] p.40). In this section we will also use a special class of step function: let us recall that a function $s(x) : [0,1] \to \mathbb{R}$ is called a level k dyadic step function if it is a finite linear combination of characteristic functions of level k dyadic intervals:

$$s(x) = \sum_{i=0}^{N} \gamma_i \chi_{X_i^k}(x) \qquad \gamma_i \in \mathbb{R}\,, \ \textit{for some } i = 0, \ldots, N \leq 2^k - 1\ .$$

It is trivial to see that each $f(x) \in C^0([0,1])$ can be obtained as the uniform limit of dyadic step functions.

Proposition 3.4. *The transformation of $([0,1], \mathcal{B}, \mu_\alpha)$ given by the binary shift T is strongly mixing for every $\alpha \in (0,1)$.*

Proof. It has to be proved that the application is measure preserving and that the limit property holds true. Both of these can be done via [18] (see Theorem 1.1 and 1.17) restricting ourselves to the case of a semi algebra which generates the Borel sets, thus we will consider dyadic intervals X_j^k.
Considering $T^{-1}(X_j^k)$, it can easily be seen that:

$$T^{-1}(X_j^k) = X_j^{k+1} \cup X_{j+2^k}^{k+1}\ .$$

Now, via proposition 3.2.2:

$$\mu_\alpha(X_j^{k+1}) = \mu_\alpha(X_j^k)\mu_\alpha(X_0^1) = \alpha\mu_\alpha(X_j^k)\ ;$$
$$\mu_\alpha(X_{2^k+j}^{k+1}) = \mu_\alpha(X_j^k)\mu_\alpha(X_1^1) = (1-\alpha)\mu_\alpha(X_j^k)\ .$$

Summing up we obtain:

$$\mu_\alpha(T^{-1}(X_j^k)) = \mu_\alpha(X_j^{k+1}) + \mu_\alpha(X_{2^k+j}^{k+1}) = \mu_\alpha(X_j^k)\ ,$$

and thus T is measure preserving.
Let us now set $B_1 = X_j^k$, $B_2 = X_i^h$, and $m \in \mathbb{N}$; we thus obtain:

$$T^{-m}(X_j^k) = \bigcup_{p=0}^{2^m-1} X_{j+p2^k}^{k+m} \ .$$

Now, let us write X_i^h as a union of level $h+l$ dyadic intervals:

$$X_i^h = \bigcup_{q=2^l i}^{2^l i+2^l-1} X_q^{h+l} \ .$$

Let us now take $m \geq h$, set $l = k+m-h$ (i.e. $k+m = h+l$) and consider $T^{-m}(X_j^k) \cap X_i^h$. We then obtain:

$$T^{-m}(X_j^k) \cap X_i^h = \left(\bigcup_{p=0}^{2^m-1} X_{j+p2^k}^{k+m} \right) \cap \left(\bigcup_{q=2^l i}^{2^l i+2^l-1} X_q^{h+l} \right) = \bigcup_{p=0}^{2^{m-h}-1} X_{i2^{k+m-h}+p2^k+j}^{k+m} \ .$$

Now, call $\lambda \equiv \mu_\alpha(X_j^k)\mu_\alpha(X_i^h)$ so that $\lambda = \alpha^{n(i)+n(j)}(1-\alpha)^{h+k-n(i)-n(j)}$. We thereby obtain:

$$\begin{aligned} \mu_\alpha(T^{-m}(X_j^k) \cap X_i^h) &= \sum_{p=0}^{2^{m-h}-1} \mu_\alpha \left(X_{i2^{k+m-h}+p2^k+j}^{k+m} \right) = \\ &= \sum_{p=0}^{2^{m-h}-1} \alpha^{n(i)+n(p)+n(j)}(1-\alpha)^{k+m-n(i)-n(p)-n(j)} = \\ &= \lambda \sum_{p=0}^{2^{m-h}-1} \alpha^{n(p)}(1-\alpha)^{m-h-n(p)} = \lambda \ , \end{aligned}$$

where the last step can be easily proved by noticing that $\alpha^{n(p)}(1-\alpha)^{m-h-n(p)}$ via (9) are the $\mu_{\alpha,m-h}$-measures of the level $m-h$ dyadic intervals, and thus the sum of all possible choices is one due to the fact that $\mu_{\alpha,m-h}$ is a probability measure.
We have proved that if m is big enough $\mu_\alpha(T^{-m}(B_1) \cap B_2) = \mu_\alpha(B_1)\mu_\alpha(B_2)$ then this property holds for the limit, as requested. □

We can, now, apply the Birkhoff theorem (Theorem 1.14 [18]) and obtain that:

$$\forall f \in C^0([0,1]) \quad \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) = \int_0^1 f\, d\mu_\alpha \quad \mu_\alpha\text{-almost everywhere.} \tag{10}$$

Moreover, via lemma 6.13 [18], there is a set Y_α such that $\mu_\alpha(Y_\alpha) = 1$ and for every $x \in Y_\alpha$ and for every $f \in C^0([0,1])$:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) = \int_0^1 f\, d\mu_\alpha\ . \tag{11}$$

It should be noticed that we can state the same result as the latter for the class of dyadic step functions as shall be seen in the following proposition proof. In particular, if we take the characteristic function $\chi_{[1/2,1]}(x)$ in equation (11), the result is related to the averages of binary digits. There is an extensive bibliography regarding the Hausdorff dimension of sets defined by means of the averages of binary digits, see, for example, [4] and the references therein. Using the notations of this framework, we can better characterize the sets where one has convergence of the limit in equation (10). Let us call:

$$F_k^\alpha = \left\{ x \in (0,1)\ :\ \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} \chi_{X_j^k}(T^i(x)) = \mu_\alpha(X_j^k)\ \forall j\ 0 \le j \le 2^l - 1 \right\}\ .$$

We can again observe that $\int_0^1 \chi_{X_j^k} d\mu_\alpha = \mu_\alpha(X_j^k)$ and thus, more explicitly, F_k^α is the subset of $(0,1)$ such that the limit property of Birkhoff's theorem (10) holds true for level k dyadic step functions.
Now we can state the main result of this section.

Proposition 3.5. *If we call $F_\infty^\alpha = \cap_{k=1}^\infty F_k^\alpha$, we have:*

$$\mu_\alpha(F_\infty^\alpha) = 1$$

$$x_0 \in F_\infty^\alpha \iff \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} f(T^i(x_0)) = \int_0^1 f\, d\mu_\alpha \quad \forall f \in C^0([0,1])\ .$$

Notice that F_∞^α is the biggest possible set satisfying Lemma 6.13 of [18].

PROOF. First of all let us see that if x_0 is a point of $[0,1]$ such that for every $f \in C^0([0,1])$ it holds that:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x_0)) = \int_0^1 f\, d\mu_\alpha\ ;$$

then if $s(x)$ is a dyadic step function it holds that:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} s(T^i(x_0)) = \int_0^1 s\, d\mu_\alpha \ . \tag{12}$$

Given the linearity of the limit and of the integral, we shall restrict ourselves to the case of one characteristic function $s(x) = \chi_{X_j^k}(x)$. For all $\varepsilon > 0$ we can take two continuous piecewise linear functions $\bar{f}^\varepsilon(x)$ and $\underline{f}^\varepsilon(x)$ which coincide with $s(x)$ up to intervals of length $\varepsilon/2$ and such that $\bar{f}^\varepsilon(x) \geq s(x) \geq \underline{f}^\varepsilon(x)$ $\forall x \in [0,1]$. The two functions $\bar{f}^\varepsilon(x)$ and $\underline{f}^\varepsilon(x)$ are continuous and thus:

$$\begin{array}{ll} \forall x_0 \in Y_\alpha & \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \underline{f}^\varepsilon(T^i(x_0)) = \int_0^1 \underline{f}^\varepsilon\, d\mu_\alpha \\ \textit{and} & \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \bar{f}^\varepsilon(T^i(x_0)) = \int_0^1 \bar{f}^\varepsilon\, d\mu_\alpha \end{array}$$

Via the Lebesgue dominated convergence theorem and measure continuity we obtain:

$$\lim_{\varepsilon\to 0} \int_0^1 \underline{f}^\varepsilon\, d\mu_\alpha = \lim_{\varepsilon\to 0} \int_0^1 \bar{f}^\varepsilon\, d\mu_\alpha = \int_0^1 s\, d\mu_\alpha \ . \tag{13}$$

Now:

$$\begin{array}{l} \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \underline{f}^\varepsilon(T^i(x_0)) \leq \liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} s(T^i(x_0)) \\ \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \bar{f}^\varepsilon(T^i(x_0)) \geq \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} s(T^i(x_0)) \ . \end{array} \tag{14}$$

Summarizing, using (13) and (14), the limit $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} s(T^i(x_0))$ exists and:

$$\forall x_0 \in Y_\alpha \ \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} s(T^i(x_0)) = \int_0^1 s\, d\mu_\alpha$$

as requested. From this property we obtain that the set $\mu_\alpha(F_k^\alpha)$ contains the set Y_α provided by the Birkhoff theorem and therefore for every $\alpha \in (0,1)$ and $\forall k \geq 1$ we have $\mu_\alpha(F_k^\alpha) = 1$. Thus the same holds for the intersection and the first statement of the proposition is proved.
The second implication of the second statement ($\Leftarrow$) is proved by (12). Let us prove the first implication ($\Rightarrow$) of the second statement. Take $\epsilon > 0$ and chose two step functions $s_1(x)$ and $s_2(x)$ to be constant on dyadic intervals of level

n_ϵ and such that $f(x)-\epsilon/2 \leq s_1(x) \leq f(x) \leq s_2(x) \leq f(x)+\epsilon/2$. They exist for the uniform continuity of f in $[0,1]$. Now:

$$\int_0^1 f\,d\mu_\alpha - \epsilon/2 = \int_0^1 [f-\epsilon/2]\;d\mu_\alpha \leq \int_0^1 s_1\,d\mu_\alpha = \\ = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} s_1(T^i(x_0)) \ .$$

Notice that this limit exists because $x_0 \in F_\infty^\alpha$ implies $x_0 \in F_{n_\epsilon}^\alpha$.
Analogously:

$$\int_0^1 f\,d\mu_\alpha + \epsilon/2 = \int_0^1 [f+\epsilon/2]\;d\mu_\alpha \geq \int_0^1 s_2\,d\mu_\alpha = \\ = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} s_2(T^i(x_0)) \ .$$

Via monotonicity properties we obtain:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} s_1(T^i(x_0)) \leq \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} f(T^i(x_0))$$
$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} s_2(T^i(x_0)) \geq \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} f(T^i(x_0)) \ .$$

And thus:

$$\int_0^1 f\,d\mu_\alpha - \epsilon/2 \leq \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} f(T^i(x_0)) \\ \leq \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} f(T^i(x_0)) \leq \int_0^1 f\,d\mu_\alpha + \epsilon/2 \ .$$

The conclusion follows since ϵ is arbitrary. □

Finally, notice that the orthogonality of the measures μ_{α_i} can be derived from this proposition in a constructive manner by means of the sets $F_k^{\alpha_i}$ which are disjoint due to the uniqueness of the limit.

4 Quadrature rules

In this section we wish to introduce a discrete approximation of the measures to be used when the numerical integration is considered. We will call an integration rule on a dyadic interval J a choice of $p+1$ distinct points $\zeta_q \in J$ (called nodes) and of values β_q (called weights). Let $f(x) \in L^1_{\mu_\alpha}(J)$ (note that, besides the integrability condition, we will always apply quadrature rules to functions with a finite number of discontinuities and defined everywhere), then we will define $\mathbb{I}_p(f, J)$ as:

$$\mathbb{I}_p(f, J) \equiv \sum_{q=0}^{p} \beta_q \cdot f(\zeta_q) \,.$$

When we consider the whole interval $J \equiv [0, 1]$ we will use the notation $\mathbb{I}_p(f)$. Notice that this is equivalent to considering the following combination of Dirac delta measures $\sum_{q=0}^{p} \beta_q \cdot \delta_{\zeta_q}$ as an approximation of the measure μ_α .
We will call degree of exactness of the formula $\mathbb{I}_p$ with respect to μ_α the greatest positive integer r such that the considered decomposition maintains the same moments up to order r:

$$\int_0^1 x^l \; d\mu_\alpha - \mathbb{I}_p(x^l) = 0 \qquad \forall\, l \leq r \,,\; l \in N_0 \,. \tag{15}$$

This relationship for $r = 0$ gives a re-normalization on the weights: $\sum_{q=0}^{p} \beta_q = 1$. In all cases the moments can be found via Proposition (3.2.5). If we fix the nodes of the quadrature rules, from the linearity of equations (15) with respect to the weights, it is always possible to construct a quadrature rule of at least $r = p$ degree of exactness. A well-known theorem, valid for general positive measures (see [7]), states that quadrature rules can have a degree of exactness up to $r = 2p + 1$.
We want to study the convergence of a general sequence of quadrature rules when applied to a function. This is exactly equivalent to studying the convergence in the weak-$*$ norm of the measure approximation.
Given a function $f \in L^1_\mu$, we will say that a sequence of quadrature rules $\{\mathbb{I}^{(n)}\}_n$ converges to f if $\mathbb{I}^{(n)}(f) \to_n I_\mu(f)$.
Given a function $f \in C^0$, we will denote by $p^*_d(x)$ the polynomial of degree at most d that gives the best approximation to f on I w.r.t. the supremum norm. We will also denote its error by E^*_d, thus $E^*_d \equiv \|f - p^*_d\|_\infty$. With these notations, the following theorem gives the most general error estimate, see [9] Theorem 5.2.2 or [17] Theorem 4.1.

Proposition 4.1. *Let $\mathbb{I}^{(n)}$ be a quadrature rule with weights $w_j, j = 0, \dots, n$ of degree of exactness $d \geq 0$. Then for all $f \in C^0$ we have:*

$$\left| \int_0^1 f(x)\, d\mu_\alpha - \mathbb{I}^{(n)}(f) \right| \leq E^*_d \left[\sum_{j=0}^{n} |w_j| + 1 \right] \,.$$

The result is proved by simply applying the definitions and the triangular inequality.
If we consider a family of rules $\{\mathbb{I}^{(n)}\}_n$ of increasing degrees of exactness d_n and such that $\sum_{j=0}^{n} |w_j| \leq K_n$, then we will obtain that the rule converges if $K_n E^*_{d_n} \to_{n\to\infty} 0$. Notice that for interpolatory quadrature rules the constant K_n is bounded from above by the Lebesgue constant Λ_n (see [15] eq. (8.11)). For further considerations on convergence of families of quadrature rules, see [17, 3].
Notice that if we consider the quadrature rule on the nodes $\zeta_q\,,\ q = 0,\dots,p$ to have $r \geq p$ degree of exactness, we obtain that, in particular, the quadrature rule exactly integrates the unique polynomial $\Pi_{f,\zeta_q}(x)$ of degree p that interpolates the function $f(x)$ on the points $\zeta_q\,,\ q = 0,\dots,p$. Hence all the rules on $p+1$ points of $r \geq p$ degree of exactness are called interpolation-based and for these rules the formula weights can be calculated by integrating, as in the case of the Lebesgue measure, the so-called Lagrange fundamental polynomials (see equation (9.2) in [15]); in our case this integration is to be made with respect to the measure μ_α.
Better estimates of the convergence properties can be attained studying the error estimates. Recall that if $\Pi_{f,\zeta_q}(x)$ is the (unique) polynomial whose degree is exactly p that interpolates the function $f(x) \in C^{p+1}([0,1])$ at the nodes $\zeta_q\,,\ q = 0,\dots,p$ we have that for all $x \in [0,1]$ (see [15] equation 8.7):

$$f(x) - \Pi_{f,\zeta_q}(x) = \frac{f^{(p+1)}(\xi_x)}{(p+1)!}\,\omega_p(x) \ \text{ for some } \xi_x \in [0,1]\ , \qquad (16)$$

where ω_p is the so-called nodal polynomial:

$$\omega_p \equiv \prod_{q=0}^{p} (x - \zeta_q)\,.$$

In order to use this estimate within our quadrature rules, we should explicitly notice that the function

$$f^{(p+1)}(\xi_x) : x \to \mathbb{R}\,,\ f^{(p+1)}(\xi_x) = \frac{f(x) - \Pi_{f,\zeta_q}(x)}{\omega_p(x)}\,(p+1)!$$

in the previous interpolation error estimate can be taken as continuous. This holds true because in each of the points ζ_i we can compute

$$\lim_{x\to\zeta_i} f^{(p+1)}(\xi_x) = \lim_{x\to\zeta_i} \frac{f(x)-\Pi_{f,\zeta_q}(x)}{\omega_p(x)}(p+1)! =$$

$$= (p+1)!\lim_{x\to\zeta_i} \frac{f(x)-f(\zeta_i)-(\Pi_{f,\zeta_q}(x)-f(\zeta_i))}{(x-\zeta_i)}\left[\frac{1}{\prod_{j=0,j\neq i}^{p}(x-\zeta_j)}\right] =$$

$$= (p+1)!\frac{f'(\zeta_i)-\Pi'_{f,\zeta_q}(\zeta_i)}{\prod_{j=0,j\neq i}^{p}(\zeta_i-\zeta_j)} ,$$

thus the function can be extended continuously in all $[0,1]$.
In the next proposition we will prove standard error estimates valid for general positive measures. We will use the following notations:

$$\omega_p^+(x) = \max\{0,\omega_p(x)\} \qquad \omega_p^-(x) = \max\{0,-\omega_p(x)\}$$
$$K^+ = \int_0^1 \omega_p^+(x)\,d\mu \qquad K^- = \int_0^1 \omega_p^-(x)\,d\mu .$$

Proposition 4.2. *Consider $\mathbb{I}_p$ to be a quadrature rule of $r \geq p$ degree of exactness with respect to a positive measure μ, then we have:*

i. If $f(x) \in C^{p+1}([0,1])$, then:

$$\left|\int_0^1 f(x)\,d\mu - \mathbb{I}_p(f)\right| \leq \frac{\left|f^{(p+1)}(\xi)\right|}{(p+1)!}\int_0^1 |\omega_p(x)|\,d\mu$$

for some $\xi \in [0,1]$.

ii. If $f(x) \in C^{p+2}([0,1])$, then:

$$\int_0^1 f(x)\,d\mu - \mathbb{I}_p(f) = \frac{1}{(p+1)!}\left[K^+ f^{(p+2)}(\xi_3)(\xi_1-\xi_2) + f^{(p+1)}(\xi_2)\int_0^1 \omega_p(x)\,d\mu\right]$$

for some $\xi_1, \xi_2, \xi_3 \in [0,1]$.

PROOF. To prove (i) we should notice that:

$$\int_0^1 f(x)\,d\mu - \mathbb{I}_p(f) = \int_0^1 \left[f(x)-\Pi_{f,\zeta_q}(x)\right]\,d\mu ,$$

because, as seen at the beginning of section 3, if the rule $\mathbb{I}_p$ is of $r \geq p$ degree of exactness, then $\mathbb{I}_p$ exactly integrates the interpolating polynomial. Now:

$$\left|\int_0^1 \left[f(x)-\Pi_{f,\zeta_q}(x)\right]\,d\mu\right| \leq \int_0^1 \left|f(x)-\Pi_{f,\zeta_q}(x)\right|\,d\mu =$$

$$= \int_0^1 \left|\frac{f^{(p+1)}(\xi_x)}{(p+1)!}\omega_p(x)\right|\,d\mu = \int_0^1 \frac{\left|f^{(p+1)}(\xi_x)\right|}{(p+1)!}\,|\omega_p(x)|\,d\mu .$$

We can then apply the mean value theorem with respect to the positive measure $|\omega_p(x)|\ d\mu$ to the function $f^{(p+1)}(\xi_x)$ which can be considered a continuous function as noticed previously. We thus obtain:

$$\left|\int_0^1 f(x)\, d\mu - \mathbb{I}_p(f)\right| \leq \frac{\left|f^{(p+1)}(\xi)\right|}{(p+1)!}\int_0^1 |\omega_p(x)|\ d\mu$$

for some $\xi \in [0,1]$. This completes the proof of (i).
Let us now consider (ii). With simple arguments, we obtain:

$$\int_0^1 f(x)\, d\mu - \mathbb{I}_p(f) =$$

$$= \frac{1}{(p+1)!}\left[\int_0^1 f^{(p+1)}(\xi_x)\omega_p^+(x)\, d\mu - \int_0^1 f^{(p+1)}(\xi_x)\omega_p^-(x)\, d\mu\right] = \tag{P4.2a}$$

$$= \frac{1}{(p+1)!}\left[f^{(p+1)}(\xi_1)\int_0^1 \omega_p^+(x)\, d\mu - f^{(p+1)}(\xi_2)\int_0^1 \omega_p^-(x)\, d\mu\right] = \tag{P4.2b}$$

$$= \frac{1}{(p+1)!}\left[K^+[f^{(p+1)}(\xi_1) - f^{(p+1)}(\xi_2)] - \left[K^- - K^+\right] f^{(p+1)}(\xi_2)\right] = \tag{P4.2c}$$

$$= \frac{1}{(p+1)!}\left[K^+ f^{(p+2)}(\xi_3)(\xi_1 - \xi_2) + f^{(p+1)}(\xi_2)\int_0^1 \omega_p(x)\, d\mu\right] \tag{P4.2d}$$

In (P4.2a) we applied the mean value theorem as in the proof of point i.
It should be noticed that in the same way we could use K^- instead of K^+ in equation (P4.2c) and still obtain an analogous estimate. □

The usual way in which the quadrature rules are used in a composite manner is to introduce a partition of the initial interval and to consider on each subinterval the integral to be approximated with a proper quadrature rule. In this section we will see how to do this in the framework of integration with respect to binomial measures. Notice that this new approximation of the measure is the one obtained by the combination of the two previously-considered ones, where we consider taking an atomic measure as an approximation in all the subintervals of the dyadic partition.
Let us consider a partition of the initial interval in 2^k dyadic subintervals of level k: $X_j^k,\ j = 0, \ldots, 2^k - 1$. We know how to rescale integrals on dyadic intervals by means of Proposition 3.2.4. We can thus observe that if we want to gain the same degree of exactness on X_j^k, if ζ_q, β_q are the nodes and weights constructed in the interval $[0,1]$, the nodes and the weights on each subinterval

are of the kind $\dfrac{\zeta_q + j}{2^k}, \mu_\alpha(X_j^k)\beta_q$.
Summarizing, if $\mathbb{I}_p$ is a quadrature rule of nodes ζ_q and weights β_q we will define $\mathbb{I}_p(f, X_j^k)$ as:

$$\mathbb{I}_p(f, X_j^k) \equiv \mu_\alpha(X_j^k) \sum_{q=0}^{p} \beta_q f\left(\frac{\zeta_q + j}{2^k}\right) .$$

In our notation $\mathbb{I}_p^k(f)$ will indicate that we are applying the local quadrature rule $\mathbb{I}_p(f, X_j^k)$ in a composite manner on the 2^k dyadic subintervals of level k to the function $f \in L^1_{\mu_\alpha}$:

$$\mathbb{I}_p^k(f) \equiv \sum_{j=0}^{2^k-1} \mathbb{I}_p(f, X_j^k) .$$

In the following we will write the previous error estimates when applied on a dyadic interval.

Corollary 4.3 (Local error estimates)**.** *Let $\mathbb{I}_p$ be of $r \geq p$ degree of exactness and take $f(x) \in C^{p+2}(X_j^k)$, then:*

$$\int_{X_j^k} f(x)\, d\mu_\alpha - \mathbb{I}_p(f, X_j^k) =$$
$$= \frac{\mu_\alpha(X_j^k)}{2^{k(p+1)}(p+1)!}\left[K_\alpha^+ f^{(p+2)}(\xi_3)(\xi_1 - \xi_2) + f^{(p+1)}(\xi_2)\int_0^1 \omega_p(x)\, d\mu_\alpha\right];$$

for some $\xi_i \in X_j^k$, $i = 1, 2, 3$.

The simplest procedure for using composite integration rules consists in considering an increasing order of partitions. For this algorithm we are interested in convergence properties.

Definition 4.4 (Convergence order)**.** *We will say that the composite rule $\mathbb{I}_p^k$ has convergence of order γ in the function f if*

$$\left|\int_0^1 f(x) d\mu_\alpha - \mathbb{I}_p^k(f)\right| \leq K_{f,p}\left(\frac{1}{2^k}\right)^\gamma .$$

Applying the Taylor expansion of the function with Lagrange's remainder, we find that the formula has an order of convergence that depends on the degree of exactness in sufficiently-regular functions.

Proposition 4.5. *Let $\mathbb{I}_p$ be of r degree of exactness and $f \in C^{r+1}([0,1])$. Then $\mathbb{I}_p^k$ has at least $r+1$ order of convergence in f.*

PROOF. Let us fix a subinterval X_j^k and consider there the Taylor expansion of the function up to the power r:

$$f(x) = P_{x_j}^r(x) + R_{x_j}^{r+1}(x).$$

If we take as initial point the first extreme $x_j = j/2^k$ we have:

$$P_{x_j}^r(x) \equiv \sum_{n=0}^{r} \frac{f^{(n)}(x_j)}{n!} (x - x_j)^n \quad ; \qquad R_{x_j}^{r+1}(x) \equiv \frac{f^{(r+1)}(\xi_x)}{(r+1)!} (x - x_j)^{r+1} \quad ,$$

for every $x \in (x_j, x_{j+1}]$, and with $\xi_x \in (x_j, x)$. Observe that $f^{(r+1)}(\xi_x)$ converges to $f^{(r+1)}(x_j)$ when x converges to x_j; then we can consider the function $x \to f^{(r+1)}(\xi_x)$ defined also for $x = x_j$ and continuous in x_j, and so in all $[0,1]$.
Let us now take the error of formula $\mathbb{I}_p(f, X_j^k)$:

$$\int_{X_j^k} f(x)\, d\mu_\alpha - \mathbb{I}_p(f, X_j^k) =$$

$$= \int_{X_j^k} P_{x_j}^r(x)\, d\mu_\alpha - \mathbb{I}_p(P_{x_j}^r, X_j^k) + \int_{X_j^k} R_{x_j}^{r+1}(x)\, d\mu_\alpha - \mathbb{I}_p(R_{x_j}^{r+1}, X_j^k) =$$

$$= \int_{X_j^k} R_{x_j}^{r+1}(x)\, d\mu_\alpha - \mathbb{I}_p(R_{x_j}^{r+1}, X_j^k)\ .$$

Let us manipulate the first term:

$$\int_{X_j^k} R_{x_j}^{r+1}(x)\, d\mu_\alpha = \int_{X_j^k} \frac{f^{(r+1)}(\xi_x)}{(r+1)!} (x - x_j)^{r+1}\ d\mu_\alpha =$$

$$= \frac{f^{(r+1)}(\xi)}{(r+1)!} \int_{X_j^k} \left(x - j/2^k\right)^{r+1}\ d\mu_\alpha = \frac{f^{(r+1)}(\xi)}{(r+1)!} \frac{\mu_\alpha(X_j^k)}{2^{k(r+1)}} \int_0^1 x^{r+1}\, d\mu_\alpha\ ,$$

where we have applied the mean value theorem and lemma 3.2.4. For the second term we have:

$$\mathbb{I}_p(R_{x_j}^{r+1}, X_j^k) = \sum_{q=0}^{p} \mu_\alpha(X_j^k) \beta_q \frac{f^{(r+1)}(\xi_q)}{(r+1)!} \cdot \left(\frac{\zeta_q + j}{2^k} - \frac{j}{2^k}\right)^{r+1} =$$

$$= \frac{\mu_\alpha(X_j^k)}{2^{k(r+1)}(r+1)!} \sum_{q=0}^{p} \beta_q f^{(r+1)}(\xi_q) \zeta_q^{r+1}$$

Thus:

$$\left|\int_0^1 f(x)d\mu_\alpha - \mathbb{I}_p^k(f)\right| = \left|\sum_{j=0}^{2^k-1}\left[\int_{X_j^k} f(x)d\mu_\alpha - \mathbb{I}_p(f,X_j^k)\right]\right| \leq$$

$$\leq \sum_{j=0}^{2^k-1}\left|\int_{X_j^k} f(x)d\mu_\alpha - \mathbb{I}_p(f,X_j^k)\right| = \sum_{j=0}^{2^k-1}\left|\int_{X_j^k} R_{x_j}^{r+1}(x)\,d\mu_\alpha - \mathbb{I}_p(R_{x_j}^{r+1},X_j^k)\right| =$$

$$\sum_{j=0}^{2^k-1}\left|\frac{f^{(r+1)}(\xi)\mu_\alpha(X_j^k)}{2^{k(r+1)}(r+1)!}\int_0^1 x^{r+1}\,d\mu_\alpha - \frac{\mu_\alpha(X_j^k)}{2^{k(r+1)}(r+1)!}\sum_{q=0}^{p}\beta_q f^{(r+1)}(\xi_q)\zeta_q^{r+1}\right| =$$

$$= \frac{1}{2^{k(r+1)}(r+1)!}\sum_{j=0}^{2^k-1}\mu_\alpha(X_j^k)\left|f^{(r+1)}(\xi)\int_0^1 x^{r+1}\,d\mu_\alpha - \sum_{q=0}^{p}\beta_q f^{(r+1)}(\xi_q)\zeta_q^{r+1}\right| \leq$$

$$\leq \frac{1}{2^{k(r+1)}(r+1)!}\sum_{j=0}^{2^k-1}\mu_\alpha(X_j^k)\left[\left|f^{(r+1)}(\xi)\int_0^1 x^{r+1}\,d\mu_\alpha\right| + \left|\sum_{q=0}^{p}\beta_q f^{(r+1)}(\xi_q)\zeta_q^{r+1}\right|\right] \leq$$

$$\leq \frac{1}{2^{k(r+1)}(r+1)!}\left[\int_0^1 x^{r+1}\,d\mu_\alpha + \sum_{q=0}^{p}|\beta_q|\,\zeta_q^{r+1}\right]\sum_{j=0}^{2^k-1}\mu_\alpha(X_j^k)\|f^{(r+1)}(x)\|_{L^\infty(X_j^k)} \leq$$

$$\leq \frac{\|f^{(r+1)}\|_{L^\infty([0,1])}}{2^{k(r+1)}(r+1)!}\left[\int_0^1 x^{r+1}\,d\mu_\alpha + \sum_{q=0}^{p}|\beta_q|\,\zeta_q^{r+1}\right].$$

This estimate gives the requested property. □

References

[1] M.F. Barnsley and S. Demko, *Iterated function systems and the global construction of fractals*, Proc. Roy. Soc. London Ser. A **399(1857)** (1985), 243–275.

[2] F. Calabrò and A. Corbo Esposito, *An efficient and reliable quadrature algorithm for integration with respect to binomial measures*, BIT **48(3)** (2008), 473–491.

[3] F. Calabrò and A. Corbo Esposito, *An evaluation of Clenshaw-Curtis quadrature rule for integration w.r.t. singular measures*, J. Comput. Appl. Math. **229(1)** (2009), 120– 128.

[4] L. Carbone, G. Cardone, and A. Corbo Esposito, *Binary digits expansion of numbers: Hausdorff dimensions of intersections of level sets of avarages' upper and lower limits*, Scientiae Mathematicae Japonicae **60(2)** (2004), 347–356.

[5] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, second ed., Applications of Mathematics (New York), vol. 38, Springer-Verlag, New York, 1998.

[6] C.J.G. Evertsz and B.B. Mandelbrot, *Multifractal measures*, Chaos and Fractals, ch. Appendix B, 921–953. Springer-Verlag, 1992.

[7] W. Gautschi, *Orthogonal polynomials and quadrature*, ETNA **9** (1999), 65–76.

[8] J.E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30(5)** (1981), 713–747.

[9] Arnold R. Krommer and Christoph W. Ueberhuber, *Computational integration*, Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics, 1998.

[10] D.P. Laurie and J.M. de Villiers, *Orthogonal polynomials for refinable linear functionals*, Math. Comp. **75(256)** (2006), 1891–1903 (electronic).

[11] G. Mantica, *A stable Stieltjes technique for computing orthogonal polynomials and Jacobi matrices associated with a class of singular measures*, Constr. Approx. **12(4)** (1996), 509–530.

[12] G. Mantica, *Fractal measures and polynomial sampling: IFS-Gaussian integration*, Numer. Algor. **45(1-4)** (2007), 269–281.

[13] Y. Pesin and H. Weiss, *The multifractal analysis of Gibbs measures: Motivation, mathematical foundation, and examples*, Chaos **7(1)** (1997), 89–106.

[14] Y. Pesin and H. Weiss, *Global analysis of dynamical systems*, ch. The Maltifractal Analysis of Birkhoff Averages and Large Deviations, IoP Publishing, Bristol, 2001.

[15] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical mathematics*, second ed., Texts in Applied Mathematics, vol. 37, Springer-Verlag, Berlin, 2007.

[16] R.H. Riedi, "Multifractal processes" in *Theory and applications of long-range dependence*, Birkhäuser, ISBN: 0817641688. Switzerland, 2002, 625–715.

[17] L.N. Trefethen, *Is Gauss quadrature better than Clenshaw Curtis?*, SIAM Review **50(1)** (2008), 67–87.

[18] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York, 1982.

[19] Andrew Q. Yingst, *A characterization of homeomorphic Bernoulli trial measures*, Trans. Amer. Math. Soc. **360(2)** (2008), 1103–1131 (electronic).