Uncertainty principle for the Cantor dyadic group

A.V. Krivoshein\textsuperscript{a}, E.A. Lebedeva\textsuperscript{b,c,*}

\textsuperscript{a} Faculty of Applied Mathematics and Control Processes, Saint Petersburg State University, Universitetsky prospekt, 35, Peterhof, Saint Petersburg, 198504, Russia
\textsuperscript{b} Mathematics and Mechanics Faculty, Saint Petersburg State University, Universitetsky prospekt, 28, Peterhof, Saint Petersburg, 198504, Russia
\textsuperscript{c} St. Petersburg Polytechnic University, Department of Calculus, Politekhnicheskaya 29, Saint Petersburg, 195251, Russia

\textbf{Article history:}
Received 3 December 2013
Available online 24 October 2014
Submitted by M. Peloso

\textbf{Keywords:}
Localization
Dyadic analysis
The Cantor group
Uncertainty principle
Scaling function
Wavelet

\textbf{Abstract}
We introduce a notion of localization for functions defined on the Cantor group. Localization is characterized by the functional $UC_d$ that is similar to the Heisenberg uncertainty constant for real-line functions. We are looking for dyadic analogs of quantitative uncertainty principles. To justify our definition we use some test functions including dyadic scaling and wavelet functions.

\textcopyright 2014 Elsevier Inc. All rights reserved.

\section{1. Introduction}

Good time–frequency localization of function $f: \mathbb{R} \rightarrow \mathbb{C}$ means that both function $f$ and its Fourier transform $Ff$ have sufficiently fast decay at infinity. The functional called the Heisenberg uncertainty constant (UC) serves as a quantitative characteristic of this property. Smaller UCs correspond to more localized functions. The uncertainty principle (UP) expresses a fundamental property of nature and can be stated as follows. If $f \neq 0$ then it is impossible for $f$ and $Ff$ to be sharply concentrated simultaneously. In terms of the UC it means that there exists an absolute lower bound for the UC.

There are numerous analogs and extensions of this framework for different algebraic and topological structures. For example, the localization of periodic functions is measured by means of the Breitenberger UC \cite{Breitenberger}. For some particular cases of locally compact groups (namely euclidean motion groups, non-compact semisimple Lie groups, Heisenberg groups) a counterpart of the UC is suggested in \cite{Krivoshein}. The generalization of operator interpretation for the UC is discussed in \cite{Krivoshein}. These and many others related topics are described

* Corresponding author.
\textit{E-mail addresses:} krivosheinav@gmail.com (A.V. Krivoshein), ealebedeva2004@gmail.com (E.A. Lebedeva).

\url{http://dx.doi.org/10.1016/j.jmaa.2014.10.043}
\url{0022-247X/© 2014 Elsevier Inc. All rights reserved}
in the excellent survey [5]. As far as we know, the question of a quantitative UC for the Cantor dyadic group has not been referred in the literature. In this paper we try to understand what “good localization” means for functions defined on the Cantor dyadic group. So, a notion of the dyadic UC is suggested and justified. The existence of a lower bound is proved for the dyadic UC. We calculate this functional for dyadic scaling and wavelet functions and find well-localized dyadic wavelet frames. Some preliminary work is done in [9].

We do not discuss qualitative UPs in this paper. There exists a qualitative UP for a wide class of groups and the Cantor group belongs to the class (see p. 224, (7.1) [5]). It is easy to see that function \( f_0 = \chi_{[0,1]} = \hat{f}_0 \), where \( \hat{f} \) is the Walsh–Fourier transform of \( f \) (see definitions and notations in Section 2), satisfies the extremal equality in this UP. There are a lot of results in this direction (see [8,7] and the references therein).

The paper is organized as follows. First, we introduce necessary notations and auxiliary results. In Section 3, we formulate the definition of the dyadic UC, discuss why the operator approach does not work here, prove the dyadic UP, answer the question how to calculate the dyadic UC in some particular important cases. In Section 4, we calculate the dyadic UC for Lang’s wavelet and look for wavelet frames with small dyadic UCs.

2. Notations and auxiliary results

We use definitions and notations on the Cantor group and Walsh analysis from [6] and [13]. Let \( F \) be the set of sequences \( \bar{x} = (x_k)_{k \in \mathbb{Z}} \), where \( x_k \in \{0, 1\} \) and either there exists \( N(\bar{x}) \in \mathbb{Z} \) such that \( x_{N(\bar{x})} = 1 \) and \( x_k = 0 \) for \( k < N(\bar{x}) \) or \( x_k = 0 \) for all \( k \in \mathbb{Z} \) (the zero element of the group \( F \)). The sum of \( \bar{x} \in F \) and \( \bar{y} \in F \) is defined by

\[
\bar{x} \oplus \bar{y} := (|x_k - y_k|)_{k \in \mathbb{Z}}.
\]

Then \((F, \oplus)\) is an abelian group called the Cantor dyadic group.

Let \( \lambda(\bar{x}) := \sum_{j \in \mathbb{Z}} x_j 2^{-j-1} \), then the map \( \bar{x} \mapsto \lambda(\bar{x}) \) is a one-to-one correspondence taking \( F \setminus \mathbb{Q}_0 \) onto \([0, \infty)\), where \( \mathbb{Q}_0 \) consists of all elements \((x_k)_{k \in \mathbb{Z}}\) such that \( \lim_{k \to \infty} x_k = 1 \). Define the sum of numbers \( \lambda(\bar{x}) \) and \( \lambda(\bar{y}) \) by \( \lambda(\bar{x}) \oplus \lambda(\bar{y}) = \lambda(\bar{x} \oplus \bar{y}) = \sum_{j \in \mathbb{Z}} |x_j - y_j| 2^{-j-1} \). Denote the half-line \([0, \infty)\) equipped with operation \( \oplus \) by \( \mathbb{R}_+ \). The set \( \mathbb{R}_+ \) is the standard interpretation of the group \( F \), although the operation \( \oplus \) is not associative on \( \mathbb{R}_+ \) (see for details [13, Sections 1.3, 9.1], [6, Sections 1.1, 1.2]). Since we do not need the associative property for our purpose, in the sequel, we use \( \mathbb{R}_+ \) instead of \( F \). To simplify notations we denote \( \lambda(\bar{x}) = x \). So, \( x \oplus y = \sum_{j \in \mathbb{Z}} |x_j - y_j| 2^{-j-1} \) for \( x, y \in \mathbb{R}_+ \).

The set \( \mathbb{R}_+ \) is metrizable with the distance between \( x, y \in \mathbb{R}_+ \) defined to be \( x \oplus y \). A function that is continuous from the \( \oplus \)-topology to the usual topology is called \( \mathbb{W} \)-continuous.

The Walsh–Fourier transform of \( f \in L_1(\mathbb{R}_+) \) is defined by

\[
\hat{f}(t) := \int_{\mathbb{R}_+} f(x) w(t, x) \, dx,
\]

where the function \( w(t, x) := (-1)^{\sum_{j \in \mathbb{Z}} t_j x_j} \) is called the generalized Walsh function. It is an \( \mathbb{R}_+ \)-analogue of a character of the group \( F \). The Walsh–Fourier transform inherits many properties from the Fourier transform (see [13, Sections 9.2, 9.3]). For example, the Plancherel theorem holds

\[
\int_{\mathbb{R}_+} f(x) \overline{g(x)} \, dx = \int_{\mathbb{R}_+} \hat{f}(x) \overline{\hat{g}(x)} \, dx,
\]

where the function \( \overline{f(x)} := \chi_{[0,1]}(x) f(x) \) is called the convolution of \( f \) with \( g \).
for $f, g, \hat{f}, \hat{g} \in L_1(\mathbb{R}_+)$ with standard extension to $L_2(\mathbb{R}_+)$. Functions $w(n, x)$, where $n = 0, 1, 2, \ldots$, are called the Walsh functions. They form an orthonormal basis for $L_2([0, 1))$. The Walsh system is a dyadic analog of the trigonometric system.

The fast Walsh-Fourier transform of $x = (x_k)_{k=0,2^m-1} \in \mathbb{R}^{2n}$ is defined by $c = xW$, where $W = 2^{-n/2}(w(m, k/2^n))^{2^n-1}_{k,m=0} = \{\omega^n_{k,m}\}^{2^n-1}_{k,m=0}$ is the normalized Walsh matrix (see [13, Section 9.7] accurate within the normalization). Matrix $W$ is orthogonal, symmetric, and unitary $W^{-1} = W$.

The concept of a dyadic derivative is quite different from its classical counterpart (see [13, Section 1.7], [16, Section 6.3]). The function

$$f^{[1]}(x) := \sum_{j \in \mathbb{Z}} 2^{j-1}(f(x) - f(x \oplus 2^{-j-1}))$$

is called the dyadic derivative of $f$ at $x$. The inherited properties are the following

$$w^{[1]}(n, x) = nw(n, x), \quad \hat{f}^{[1]}(t) = t\hat{f}(t).$$

But unfortunately the dyadic derivative does not support some natural properties such as $(fg)' = fg' + f'g$ and the chain rule.

Let $H$ be a separable Hilbert space. If there exist constants $A, B > 0$ such that for any $f \in H$ the following inequality holds $A\|f\|^2 \leq \sum_{n=1}^{\infty} |(f, f_n)|^2 \leq B\|f\|^2$, then the sequence $(f_n)_{n \in \mathbb{N}}$ is called a frame for $H$. If $A = B (= 1)$, then the sequence $(f_n)_{n \in \mathbb{N}}$ is called a (normalized) tight frame for $H$.

If the set of functions $\psi_{j,k}(x) := 2^{j/2}\psi(2^j x + k)$ forms a frame or a basis of $L_2(\mathbb{R}_+)$, then it is called a dyadic wavelet frame or basis. Using the routine procedure, it can be generated from multiresolution analysis starting with an auxiliary function, that is a scaling function $\varphi$. The concept of dyadic wavelet function and the elements of multiresolution analysis theory for the Cantor dyadic group are developed in [10] and later in [3,2].

3. Localization of functions

The quantitative characteristic of the time–frequency localization is the uncertainty constant (UC). Originally, the concept of an uncertainty constant and principle was introduced for the real line case in 1927. The Heisenberg uncertainty constant of $f \in L_2(\mathbb{R})$ is the functional $UC_H(f) := \Delta_f \Delta_{FF}$ such that

$$\Delta_f^2 := \frac{1}{\|f\|_{L^2(\mathbb{R})}^2} \int \|x - xf\|^2 \|f(x)\|^2 \, dx, \quad \Delta_{FF}^2 := \frac{1}{\|FF\|_{L^2(\mathbb{R})}^2} \int \|t - t_{FF}\|^2 \|FF(t)\|^2 \, dt,$$

$$xf := \frac{1}{\|f\|_{L^2(\mathbb{R})}^2} \int x|f(x)|^2 \, dx, \quad t_{FF} := \frac{1}{\|FF\|_{L^2(\mathbb{R})}^2} \int t|FF(t)|^2 \, dt,$$

where $FF$ denotes the Fourier transform of $f$. It is well known that $UC_H(f) \geq 1/2$ for all functions $f \in L_2(\mathbb{R})$ and the minimum is attained on the Gaussian. Let us make some preliminary remarks to motivate the definition of a localization characteristic for the dyadic case.

**Remark 1.** First, it is easy to see that $xf$ is the solution of the minimization problem

$$\min_x \int \|x - x\|^2 \|f(x)\|^2 \, dx.$$

Second, it is well known that $xf$ is equal to the integral mean value of the function $f$, while $\Delta_f$ means the dispersion with respect to the $xf$. Hence, the squared $UC_H$ takes the form
Now we are ready to introduce the definition of a localization characteristic for the dyadic setup.

**Definition 1.** Suppose \( f \in L_2(\mathbb{R}_+) \) is a complex valued function, then the functional

\[
UC_d(f) := V(f) V(\hat{f}),
\]

where

\[
V(f) := \frac{1}{\|f\|^2_{L^2(\mathbb{R}_+)}} \min_{\dot{x}} \int_{\mathbb{R}_+} (x \oplus \dot{x})^2 |f(x)|^2 \, dx,
\]

\[
V(\hat{f}) := \frac{1}{\|\hat{f}\|^2_{L^2(\mathbb{R}_+)}} \min_{\dot{t}} \int_{\mathbb{R}_+} (t \oplus \dot{t})^2 |\hat{f}(t)|^2 \, dt,
\]

is called the dyadic uncertainty constant (the dyadic UC) of the function \( f \).

**Remark 2.** Suppose \( g \) is a bounded complex-valued function, \( g(x), xg(x) \in L_2(\mathbb{R}_+) \). We denote \( G(y) := \int_{\mathbb{R}_+} (x \oplus y)^2 |g(x)|^2 \, dx \). Since \( g(x), xg(x) \in L_2(\mathbb{R}_+) \) and \( x \oplus y < x + y \) it follows that \( G(y) \) is finite for \( y \in \mathbb{R}_+ \). Then there exists a point \( y^* \) such that \( \min_y G(y) = G(y^*) \). Indeed, it is clear that \( y^* \) cannot be outside the interval \([0, 2^n] \) for some probably large \( n \in \mathbb{N} \) depending on \( g \). It can be checked that \([0, 2^n] \) is compact in the dyadic topology. The function \( x \oplus y \) is \( W \)-continuous, therefore \( G \) is \( W \)-continuous. It is well known that under these conditions, the image \( G([0, 2^n]) \) is compact. Finally, since \( G([0, 2^n]) \subset \mathbb{C} \), it follows that \( G([0, 2^n]) \) is bounded and closed.

**Example 1.** Let \( \chi_M \) be a characteristic function of a set \( M \). Denote \( f_1(x) = \chi_{[0,1/4]}(x) \) and \( g_1(x) = \chi_{[3/4,1]}(x) \). Then it is easy to calculate their Walsh-Fourier transforms \( \hat{f}_1 = \chi_{[0,4]} / 4 \) and \( \hat{g}_1 = w(3, \cdot /4) \chi_{[0,4]} / 4 \). It is natural to characterize “the dispersion” of these functions by means of the diameters of their supports. Thus,\text{diam}[0,1/4] := \sup_{x,y\in[0,1/4]}(x \oplus y) = 1/4, \text{diam}[3/4,1] = 1/4, and \text{diam}[0,4] = 4. So, these functions should have the same localization. On the other side, let us consider the functions \( f_2(x) = \chi_{[0,3/8]}(x) \) and \( g_2(x) = \chi_{[3/4,9/8]}(x) \). Their Walsh-Fourier transforms are \( \hat{f}_2 = \chi_{[0,4]} / 4 + w(1, \cdot /4) \chi_{[0,8]} / 8 \) and \( \hat{g}_2 = w(3, \cdot /4) \chi_{[0,4]} / 4 + w(1, \cdot /4) \chi_{[0,8]} / 8 \). Calculating the diameters we get \text{diam}[0,3/8] = 1/2, \text{diam}[3/4,9/8] = 2, and \text{diam}[0,8] = 8. So, the first function should be more localized. Indeed, Table 1 shows that our suppositions are correct. Columns named \( \tilde{x}_0(f) \) and \( \tilde{t}_0(f) \) mean sets of \( \tilde{x} \) and \( \tilde{t} \) minimizing the functionals \( \int_{\mathbb{R}_+} (x \oplus \tilde{x})^2 |f(x)|^2 \, dx \) and \( \int_{\mathbb{R}_+} (t \oplus \tilde{t})^2 |\hat{f}(t)|^2 \, dt \) respectively.

**Remark 3.** The operator interpretation of the UC does not work for the dyadic setup. Let \( P \) and \( M \) be self-adjoint, symmetric or normal operators defined on a Hilbert space, \( [P, M] := PM - MP \) be...
a commutator of $P$ and $M$, and $[P, M]_+ := PM + MP$ be an anticommutator of $P$ and $M$. The following inequality named the Schrödinger uncertainty principle (see [14]) is a simple consequence of the Cauchy–Bunyakovskii–Schwarz inequality

$$\|Mf - \beta f\|^2\|Pf - \alpha f\|^2 \geq \frac{1}{4} \left( \|([P, M]_f, f)\|^2 + \|([P, M]_f + f, f) - 2\alpha\beta\|f\|^2 \right)^2,$$

where $\beta := (Mf, f)/\|f\|^2$, $\alpha := (Pf, f)/\|f\|^2$. It gives two functionals both used as the UCs: the first one is more traditional, but some authors (see [15]) exploit the second one as well

$$UC_-(f) := \frac{\|Mf - \beta f\|^2\|Pf - \alpha f\|^2}{\|([P, M]_f, f)\|^2} \geq 1/2 \quad (1)$$

$$UC_+(f) := \frac{\|Mf - \beta f\|^2\|Pf - \alpha f\|^2}{\|([P, M]_f + f, f) - 2\alpha\beta\|f\|^2} \geq 1/2. \quad (2)$$

Defying in (1) $Pf(x) = i\psi(x)$ and $Mf(x) = x\psi(x)$, one gets the Heisenberg UC in $L_2(\mathbb{R})$. The dyadic extension of this framework has the following trouble. If the inner product $(PHf, MHf)$ is real-valued then the mean value of the commutator $([P, M]_f, f) = 2i\Im(PHf, MHf)$ vanishes. In classical setup the inner product is pure imaginary for a real-valued $f$. But for the natural choice of dyadic operators on $L_2(\mathbb{R}_+)$, namely $Pf(x) = f^{|\mathbb{N}|}(x)$ and $Mf(x) = x\psi(x)$, it turns out to be real-valued. Thus, one gets the identical zero in the denominator of (1). The reason of this trouble is the difference between the operators $i\psi'$ and $f^{|\mathbb{N}}$. It is caused by the definitions of respective characters and the properties of derivatives, namely $(e^{it})' = ie^{it}$ and $(w(n, t))^{|\mathbb{N}} = n\psi(n, t)$, the imaginary unit appears only in the classical case.

A dyadic counterpart of (2) does not give an adequate characteristic of localization. Indeed, it is equal to infinity for very well localized function $f_0 := \chi_{[0, 1]}$, $\widehat{f}_0 = f_0$, while $UC_d(f_3) = 1/9$.

There is a lower bound for $UC_d$, so we get an uncertainty principle for the dyadic Cantor group.

**Theorem 1.** For any function $f \in L_2(\mathbb{R}_+)$, the following inequality holds

$$UC_d(f) \geq C, \quad \text{where } C \simeq 8.5 \times 10^{-5}.$$

**Proof.** Suppose $f_1(x) := w(t, x)f(x \oplus \bar{x})$, then $\widehat{f}_1(t) := w(t, \bar{x})\widehat{f}(t \oplus \bar{t})$ and it is easy to see from straightforward calculation that

$$\int_{\mathbb{R}_+} (t \oplus \bar{t})^2 |\widehat{f}(t)|^2 dt = \int_{\mathbb{R}_+} \bar{t}^2 |\widehat{f}_1(t)|^2 dt,$$

$$\int_{\mathbb{R}_+} (x \oplus \bar{x})^2 |f(x)|^2 dx = \int_{\mathbb{R}_+} x^2 |f_1(x)|^2 dx. \quad (3)$$

So, it is sufficient to prove

$$\|xg(x)||tg(t)|| \geq \sqrt{C}\|g\|^2.$$

It can be done in the same manner as its classical counterpart (see [12, Theorem 1.1, Corollaries 1.2, 1.3]).

1. Let $E$ be a measurable subset of $\mathbb{R}_+$, $|E|$ be a Lebesgue measure of $E$, and $0 < \theta < 1/2$. Then

$$\left( \int_{E} |\widehat{f}(t)|^2 \right)^{1/2} \leq K_1(\theta)|E|^\theta \|x^\theta f(x)\|_2, \quad \text{where } K_1(\theta) = (2\theta)^{2\theta}(1 - 2\theta)^{1-\theta}.$$
Indeed, suppose $B = [0, b)$, $B' = [b, \infty)$. Then $(\int_E |\hat{f}|^2)^{1/2} \leq (\int_E |\hat{f}X_B|^2)^{1/2} + (\int_E |\hat{f}X_{B'}|^2)^{1/2}$. Using the definition of the Walsh–Fourier transform, the Cauchy–Bunyakovskii–Schwarz inequality, and some elementary properties of integrals we get for the first and the second summands

\[
\left( \int_E |\hat{f}X_B|^2 \right)^{1/2} \leq |E|^{1/2} \sup_E |\hat{f}X_B| \leq |E|^{1/2} \|fX_B\|_1 \leq |E|^{1/2} \|x^{-\theta}X_B(x)\|_2 \|x^\theta f(x)\|_2
\]

\[
= |E|^{1/2}(1 - 2\theta)^{-1/2}b^{-\theta + 1/2}\|x^\theta f(x)\|_2,
\]

\[
\left( \int_E |\hat{f}X_{B'}|^2 \right)^{1/2} \leq \|fX_{B'}\|_2 \leq \sup_{B'} x^{-\theta} \|x^\theta f(x)\|_2 \leq b^{-\theta} \|x^\theta f(x)\|_2.
\]

So,

\[
\left( \int_E |\hat{f}|^2 \right)^{1/2} \leq (|E|^{1/2}(1 - 2\theta)^{-1/2}b^{-\theta + 1/2} + b^{-\theta}) \|x^\theta f(x)\|_2.
\]

It remains to minimize the right-hand side over $b$ ($b_{\min} = 4\theta^2|E|^{-1}(1 - 2\theta)^{-1}$) to get the desired inequality.

2. Let us prove $\|f\|_2^2 \leq 2K_1(\theta)\|x^\theta f(x)\|_2\|t^\theta \hat{f}(t)\|_2$ for $0 < \theta < 1/2$. Denote $E = [0, r)$, $E' = [r, \infty)$. Then using the first item, we obtain

\[
\|f\|_2^2 = \|\hat{f}\|_2^2 = \int_E |\hat{f}|^2 + \int_{E'} |\hat{f}|^2 \leq K_1^2(\theta)r^{2\theta}\|x^\theta f(x)\|_2^2 + r^{-2\theta}\|t^\theta \hat{f}(t)\|_2^2.
\]

Minimizing the last expression over $r$ ($r_{\min} = \|t^\theta \hat{f}(t)\|_2^{1/(4\theta)}(K_1^2(\theta)\|x^\theta f(x)\|_2)^{-1/(4\theta)}$) we get the necessary inequality.

3. Since the function $g(\alpha) := (\|x^\alpha f(x)\|_2\|f\|_2^{-1})^{1/\alpha}$ decreases for $\alpha > 0$ ($g'_\alpha > 0$), then

\[
\|x^\alpha f(x)\|_2 \leq \|f\|_2^{1-\alpha/\beta}\|x^\beta f(x)\|_2^{\alpha/\beta}
\]

for $0 < \alpha < \beta$.

4. Applying the last inequality ($\alpha = \theta$) to item 2 we obtain

\[
\|f\|_2^2 \leq 2K_1(\theta)\|x^\theta f(x)\|_2\|t^\theta \hat{f}(t)\|_2 \leq 2K_1(\theta)\|f\|_2^{2-2\theta/\beta}\|x^\beta f(x)\|_2^{\theta/\beta}\|t^\theta \hat{f}(t)\|_2^{\theta/\beta},
\]

thus

\[
\|f\|_2^2 \leq (2K_1(\theta))^{\beta/\theta}\|x^\beta f(x)\|_2\|t^\beta \hat{f}(t)\|_2.
\]

So, choosing $\beta = 1$ we have

\[
\|xf(x)\|_2\|t\hat{f}(t)\|_2 \geq C(\theta)\|f\|_2^2, \quad \text{where} \quad C(\theta) = (2K_1(\theta))^{-1/\theta}.
\]

To get the dyadic uncertainty principle it remains to maximize $C^2(\theta)$ over $\theta$, $\max_\theta C^2(\theta) \simeq C^2(0.382) \simeq 8.5 \times 10^{-5}$. \hfill \Box

It is not easy to calculate $UC_d$ for an arbitrary function because of the dyadic minimization problem underlying in the definition of $UC_d$. The following result gives a possible way to calculate the dyadic UC on a wide class of functions. The minimization problem adds up to exhaustive search among $2^n$ variants.
Lemma 1. Let \( f(x) = \sum_{k=0}^{\infty} a_k(x) \) be a uniformly convergent series restricted on \([0,1]\), \( f_n(x) = \sum_{k=0}^{2^n-1} a_k(x) \) be its partial sum, \( V(f) < +\infty, V(\hat{f}) < +\infty \). Then the dyadic UC takes the form

\[
UC_d(f) = \lim_{n \to \infty} V(f_n)V(\hat{f}_n),
\]

where

\[
V(f_n) = \min_{k_0=0,2^n-1} \frac{\sum_{k=0}^{2^n-1} |c_k|}{\sum_{k=0}^{2^n-1} |a_k|^2},
\]

\[
V(\hat{f}_n) = \min_{k_1=0,2^n-1} \frac{\sum_{k=0}^{2^n-1} |a_k|}{\sum_{k=0}^{2^n-1} |c_k|^2},
\]

and \( c := (c_k)_{k=0,2^n-1} \) is the fast Walsh–Fourier transform of \( a := (a_k)_{k=0,2^n-1} \).

**Proof.** Suppose \( \Delta_{k,n} := [k2^{-n}, (k+1)2^{-n}] \), \( k = 0, \ldots, 2^n-1, n = 0, 1, \ldots \), is a dyadic interval, \( \xi_{k,n} := \chi_{\Delta_{k,n}} \) is the characteristic function of \( \Delta_{k,n} \), and \( f_n(x) = \sum_{k=0}^{n} b_k \xi_{k,n}(x) \) is a representation of \( f_n \) with respect to the orthogonal system \( \{\xi_{k,n} : k = 0, \ldots, 2^n-1, n = 0, 1, \ldots\} \). It is easy to find a connection between \( a = (a_k)_{k=0,2^n-1} \) and \( b = (b_k)_{k=0,2^n-1} \). Indeed,

\[
\sum_{k=0}^{2^n-1} a_k(x) = f_n(x) = \sum_{k=0}^{2^n-1} b_k \xi_{k,n}(x).
\]

The Walsh–Fourier coefficient of \( f_n \) is

\[
a_k = \int_{[0,1]} f_n(x)w(k,x) \, dx = \int_{[0,1]} \sum_{m=0}^{2^n-1} b_m \xi_{m,n}(x) w(k,x) \, dx
\]

\[
= \sum_{m=0}^{2^n-1} b_m \int_{\Delta_{m,n}} w(k,x) \, dx = \sum_{m=0}^{2^n-1} b_m \frac{1}{2^n} \omega_{k,m}^n,
\]

where \( \omega_{k,m}^n \) is a value of \( w(k,\cdot) \) on \( \Delta_{m,n} \). Let us denote \( c_k := b_k 2^{-n/2}, \omega_{k,m}^n := \omega_{k,m}^n 2^{-n/2} \). Then \( a_k = \sum_{m=0}^{2^n-1} c_m \omega_{k,m}^n \), that is \( a = cW \). Thus, \( c \) is the fast Walsh–Fourier transform of \( a \).

If \( \tilde{x}_n \) minimizes the functional \( \int_{R_{+}} (x \oplus \tilde{x})^2 |f_n(x)|^2 \, dx \) then \( \tilde{x}_n \) cannot be outside the support of \( f_n \). So, \( \tilde{x} \in [0,1] = \bigcup_{k=0,2^n-1} \Delta_{k,n} \). Then, for \( \tilde{x} \in \Delta_{k_0,n} \), we have

\[
\int_{R_{+}} (x \oplus \tilde{x})^2 |f_n(x)|^2 \, dx = \int_{[0,1]} (x \oplus \tilde{x})^2 \left| \sum_{k=0}^{2^n-1} b_k \xi_{k,n}(x) \right|^2 \, dx
\]

\[
= \int_{[0,1]} (x \oplus \tilde{x})^2 \sum_{k=0}^{2^n-1} b_k^2 \xi_{k,n}(x) \, dx = \sum_{k=0}^{2^n-1} b_k^2 \int_{\Delta_{k,n}} (x \oplus \tilde{x})^2 \, dx
\]

\[
= \sum_{k=0}^{2^n-1} b_k^2 \frac{x^3}{3} \Delta_{k,n}(\tilde{x}) = \sum_{k=0}^{2^n-1} b_k^2 \frac{x^3}{3} \Delta_{k,n} = \sum_{k=0}^{2^n-1} c_k^2 \frac{3k^2 + 3k + 1}{3 \times 2^{2n}}.
\]
So, recalling Definition 1, we get

$$V(f_n) := \frac{1}{\|f_n\|_{L_2(\mathbb{R}^+)}^2} \min_{x} \int_{\mathbb{R}^+} (x \oplus \tilde{x})^2 |f(x)|^2 \, dx = \frac{1}{\min_{k_0=0,2^{n-1}-1} \sum_{k=0}^{2^{n-1}-1} |a_k|^2} \min_{k_0=0,2^{n-1}-1} \sum_{k=0}^{2^{n-1}-1} c_k^2 k_0 3k^2 + 3k + 1 \frac{3}{2^{2n}}.$$}

The Walsh–Fourier transform of $f_n$ is

$$\hat{f}_n(t) = \sum_{k=0}^{2^{n-1}-1} a_k \int_{[0,1)} w(x,t)w(x,k) \, dx = \sum_{k=0}^{2^{n-1}-1} a_k \chi_{[k,k+1]}(t).$$

Then repeating the above calculations, we have

$$V(\hat{f}_n) := \frac{1}{\|\hat{f}_n\|_{L_2(\mathbb{R}^+)}^2} \min_{t} \int_{\mathbb{R}^+} (t \oplus \tilde{t})^2 |\hat{f}(t)|^2 \, dt = \frac{1}{\min_{k_1=0,2^{n-1}-1} \sum_{k=0}^{2^{n-1}-1} |c_k|^2} \min_{k_1=0,2^{n-1}-1} \sum_{k=0}^{2^{n-1}-1} a_{k \oplus k_1}^2 \frac{3k^2 + 3k + 1}{3}.$$}

To conclude the proof, it remains to show that $UC_d(f) = \lim_{n \to \infty} UC_d(f_n)$. We denote $V_0(g) := \|g\|_{L_2(\mathbb{R}^+)}^2 V(g) = \min_{x} \int_{\mathbb{R}^+} (x \oplus \tilde{x})^2 |g(x)|^2 \, dx.$

Firstly, we prove $\lim_{n \to \infty} V_0(f_n) = V_0(f)$. Assume that the minimum of the functional $V_0(f_n)$ is achieved at the point $\tilde{x}_n^*$, the minimum of the functional $V_0(f)$ is achieved at the point $\tilde{x}^*$. The functions $f_n$ converge uniformly on $[0,1)$ to $f$, i.e. for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and for all $x \in [0,1)$ we have $||f(x) - f_n(x)|| \leq |f(x) - f_n(x)| < \varepsilon$. Then

$$|f(x)|^2 - |f_n(x)|^2 \leq 2|f_n(x)||f(x) - f_n(x)| + |f(x) - f_n(x)|^2 \leq 2|f_n(x)|\varepsilon + \varepsilon^2 \leq 2(|f(x)| + \varepsilon)\varepsilon + \varepsilon^2.$$

After multiplication by $(x \oplus \tilde{x})^2$ and integration over $[0,1)$ both sides of the above inequality, for all $y \in [0,1)$ and for all $n \geq N$ we get

$$\int_{[0,1)} (x \oplus \tilde{x})^2 |f(x)|^2 \, dx - \int_{[0,1)} (x \oplus \tilde{x})^2 |f_n(x)|^2 \, dx \leq \varepsilon C$$

where $C = \max_{y \in [0,1)} (x \oplus \tilde{x})^2 (2|f(x)| + 3\varepsilon) \, dx$. The last inequality should be valid for $y = \tilde{x}_n^*$

$$\int_{[0,1)} (x \oplus \tilde{x}_n^*)^2 |f(x)|^2 \, dx - V_0(f_n) \leq \varepsilon C \quad \forall n \geq N.$$

Finally, we can decrease the left-hand side of the inequality by taking minimum of the functional over $\tilde{x}_n^*$

$$V_0(f) - V_0(f_n) \leq \varepsilon C.$$

Similarly, we can prove the following inequality

$$V_0(f_n) - V_0(f) \leq \varepsilon C.$$

But it requires to start with

$$|f_n(x)|^2 - |f(x)|^2 \leq 2|f(x)||f(x) - f_n(x)| + |f(x) - f_n(x)|^2 \leq 2|f(x)|\varepsilon + \varepsilon^2 \quad \forall n \geq N, \forall x \in [0,1)$$

and after the integration take $y = \tilde{x}^*$. As a result, we get $\lim_{n \to \infty} V_0(f_n) = V_0(f)$.
Now, let us prove $\lim_{n \to \infty} V_0(\hat{f}_n) = V_0(\hat{f})$. Assume that the minimum of the functional $V_0(\hat{f}_n)$ is achieved at the point $\hat{t}^*_n$, the minimum of the functional $V_0(\hat{f})$ is achieved at the point $\hat{t}^*$. By (4) we conclude that $|\hat{f}_{n+1}(t)|^2 \geq |\hat{f}_n(t)|^2$ for all $t \in \mathbb{R}_+$. After multiplication by $(t \otimes y)^2$ and integration over $\mathbb{R}_+$ both sides of the above inequality, we get

$$
\int_{\mathbb{R}_+} (t \otimes y)^2 |\hat{f}_{n+1}(t)|^2 \, dt \geq \int_{\mathbb{R}_+} (t \otimes y)^2 |\hat{f}_n(t)|^2 \, dt \quad \forall y \in \mathbb{R}_+.
$$

Thus, the last inequality should be valid for $y = \hat{t}^*_{n+1}$

$$
V_0(\hat{f}_{n+1}) = \int_{\mathbb{R}_+} (t \otimes \hat{t}^*_{n+1})^2 |\hat{f}_{n+1}(t)|^2 \, dt \geq \int_{\mathbb{R}_+} (t \otimes \hat{t}^*_n)^2 |\hat{f}_n(t)|^2 \, dt = V_0(\hat{f}_n).
$$

Therefore, $V_0(\hat{f}_{n+1}) \geq V_0(\hat{f}_n)$ for all $n \in \mathbb{N}$, in particular, $V_0(\hat{f}) \geq V_0(\hat{f}_n)$. Let us consider the difference

$$
V_0(\hat{f}) - V_0(\hat{f}_n) = \min_{t} \int_{\mathbb{R}_+} (t \otimes \hat{t}^*_n)^2 |\hat{f}(t)|^2 \, dt - \int_{\mathbb{R}_+} (t \otimes \hat{t}^*_n)^2 |\hat{f}_n(t)|^2 \, dt
$$

$$
\leq \int_{\mathbb{R}_+} (t \otimes \hat{t}^*_n)^2 (|\hat{f}(t)|^2 - |\hat{f}_n(t)|^2) \, dt = \int_{\mathbb{R}_+} (t \otimes \hat{t}^*_n)^2 \sum_{k=2^n}^{\infty} |a_k|^2 \chi_{[k,k+1)}(t) \, dt.
$$

There exists $N \in \mathbb{N}$ such that $\hat{t}^*_n \in [0,2^N)$ and $\hat{t}^* \in [0,2^N)$ for all $n \in \mathbb{N}$ simultaneously. It can be shown by contradiction. Indeed, assume that for any $N \in \mathbb{N}$ there exists $m > N$ such that $\hat{t}^*_m \geq 2^N$. Then the following inequalities

$$
V_0(\hat{f}) \geq V_0(\hat{f}_m) = \int_{[0,2^N)} (t \otimes \hat{t}^*_m)^2 |\hat{f}_m(t)|^2 \, dt + \int_{[2^N,2^{N+1})} (t \otimes \hat{t}^*_m)^2 |\hat{f}_m(t)|^2 \, dt
$$

$$
\geq \int_{[0,2^N)} (t \otimes \hat{t}^*_m)^2 |\hat{f}_m(t)|^2 \, dt \geq 2^N \sum_{k=0}^{2^N-1} |a_k|^2
$$

should be valid for all $N$. This leads to a contradiction. The function $\int_{\mathbb{R}_+} (t \otimes y)^2 |\hat{f}(t)|^2 \, dt$ is bounded on $[0,2^N)$ (see Remark 2). Therefore, for all $\varepsilon > 0$ there exists $M$ such that for all $m > M$, $m \in \mathbb{N}$

$$
\int_{[m,\infty)} (t \otimes y)^2 |\hat{f}(t)|^2 \, dt = \int_{\mathbb{R}_+} (t \otimes y)^2 \sum_{k=m}^{\infty} |a_k|^2 \chi_{[k,k+1)}(t) \, dt < \varepsilon.
$$

Then for $n$ such that $2^n > m$ we have $V_0(\hat{f}) - V_0(\hat{f}_n) < \varepsilon$. Hence, $\lim_{n \to \infty} V_0(\hat{f}_n) = V_0(\hat{f})$. Together with $\lim_{n \to \infty} V_0(f_n) = V_0(f)$, $\lim_{n \to \infty} \|f_n\|_{L_2(\mathbb{R}_+)} = \|f\|_{L_2(\mathbb{R}_+)}$, and $\lim_{n \to \infty} \|\hat{f}_n\|_{L_2(\mathbb{R}_+)} = \|\hat{f}\|_{L_2(\mathbb{R}_+)}$ we get the required statement for $UC_d$. \(\square\)

**Remark 4.** It is easy to extend Lemma 1 to the functions of the following form

$$
g(x) := \chi_{[0,2^N)}(x) \sum_{k=0}^{\infty} a_k w_k(x/2^N).
$$
Indeed, let \( g_n(x) := \chi_{[0,2N]}(x) \sum_{k=0}^{2^n-1} a_k w_k(x/2^N) \) be a partial sum of the above function \( g, f_n(x) = g_n(2^N x) \) the function defined in Lemma 1. Then standard calculations show that \( \|g_n\|_2^2 = 2^N \|f_n\|_2^2, \|\hat{g}_n\|_2^2 = 2^N \int_{\mathbb{R}^+} (x \oplus \hat{x})^2 |g_n(x)|^2 \, dx = 2^{2N} \int_{\mathbb{R}^+} (x \oplus (\hat{x} 2^{-N}))^2 |f_n(x)|^2 \, dx \) and \( \int_{\mathbb{R}^+} (t \oplus \hat{t})^2 |\hat{g}_n(t)|^2 \, dt = 2^{-N} \int_{\mathbb{R}^+} (t \oplus \hat{t})^2 |\hat{f}_n(t)|^2 \, dt \). Hence, recalling the definition of \( UC_d \) we get \( UC_d(g_n) = UC_d(f_n) \). The class of the functions of the form \( g \) is rather large and important since any orthogonal compactly supported dyadic scaling and wavelet functions belong to this set (see [3, Section 5]).

We denote \( q_k := \frac{3k^2 + 3k + 1}{3x^2} \) and suppose \( \|a\| = 1 \), then \( \|c\| = \|aW\| = 1 \) and the \( UC_d(f_n) \) takes the form

\[
UC_d(f_n) = \min_{k_1=0,2^n-1} \sum_{k=0}^{2^n-1} a_k^2 q_k \min_{k_0=0,2^n-1} \sum_{k=0}^{2^n-1} c_k^2 q_k.
\]

Let us fix \( n \). It follows from (3) that the minimization problem

\[
\begin{cases}
UC_d(f_n) \to \min \\
\|a\| = 1
\end{cases}
\]

is equivalent to the following one

\[
\begin{cases}
\sum_{k=0}^{2^n-1} a_k^2 q_k \sum_{k=0}^{2^n-1} c_k^2 q_k \to \min \\
\|a\| = 1
\end{cases}
\]

Using Wolfram Mathematica 8.0 we solve numerically the latter minimization problem for \( n = 2; 3; 4; 5; 6 \). The result is demonstrated in Table 2.

### 4. Examples

#### 4.1. Lang’s wavelet and scaling function

To examine and illustrate the definition of the dyadic UC we use the first nontrivial example of orthogonal wavelets on the Cantor dyadic group (see [10]). The dyadic scaling function is defined by

\[
\varphi_a(x) = \frac{1}{2} \chi_{[0,1)} \left( \frac{x}{2} \right) \left( 1 + a \sum_{j=0}^{\infty} b^j \omega \left( 2^{j+1} - 1, \frac{x}{2} \right) \right), \quad \hat{\varphi}_a = \chi_{[0,1/2)} + a \sum_{j=0}^{\infty} b^j \chi_{[2^{-j-1},2^{-j})},
\]

where \( 0 < |a| \leq 1, |a|^2 + |b|^2 = 1, a, b \in \mathbb{C} \). The corresponding wavelet is defined by

\[
\psi_a(x) = 2a_0 \varphi_a(2x + 1) - 2a_1 \varphi_a(2x) + 2a_2 \varphi_a(2x + 3) - 2a_3 \varphi_a(2x + 2),
\]

where \( a_0 = (1 + a + b)/4, a_1 = (1 + a - b)/4, a_2 = (1 - a - b)/4, a_3 = (1 - a + b)/4 \). Then the wavelet system \( \{ \psi_{j,k} \}_{j \in \mathbb{Z}, k \in \mathbb{Z}^+} \) forms an orthonormal basis in \( L_2(\mathbb{R}^+) \).
The integrals defying the dyadic UC for the scaling and wavelet functions are

$$\int_{\mathbb{R}_+} (x + \tilde{x})^2 |\varphi_a(x)|^2 \, dx = \frac{4}{3} + \frac{1}{4} w \left(1, \frac{\tilde{x}}{2}\right) \left(-4\Re a + \Re(ab)w(1, \tilde{x})\right) + \frac{|a|^2 |b|^2}{16} \sum_{j=0}^{\infty} \left(\frac{|b|^2}{4}\right)^j w(2^j, \tilde{x})$$

$$\int_{\mathbb{R}_+} (t + \tilde{t})^2 |\varphi_a(t)|^2 \, dt = A(0, \tilde{t}) + |a|^2 \sum_{j=0}^{\infty} |b|^{2j} A(2^j - 1/2, \tilde{t})$$

$$\int_{\mathbb{R}_+} (x + \tilde{x})^2 |\psi_a(x)|^2 \, dx = \frac{4}{3} - \Re aw \left(1, \frac{\tilde{x}}{2}\right) - \frac{\Re(ab)}{4} w \left(1, \frac{\tilde{x}}{2}\right) w(1, \tilde{x}) + \frac{\Re(\tilde{a}b)}{2} w(1, \tilde{x}) - \frac{|b|^2 \Re(\tilde{a}^2)}{16} w(3, \tilde{x})$$

$$\int_{\mathbb{R}_+} (t + \tilde{t})^2 |\psi_a(t)|^2 \, dt = |b|^2 A(1/2, \tilde{t}) + |a|^2 \sum_{j=0}^{\infty} |b|^{2j} A(2^j - 1, \tilde{t}) + |a|^4 \sum_{j=1}^{\infty} |b|^{2j} A(2^j - 1/2, \tilde{t})$$

where $A(\xi, \eta) = \frac{1}{3}((\inf\{[\xi, \xi + 1/2] \oplus \eta\}) + 1/2^3 - \frac{1}{3}(\inf\{[\xi, \xi + 1/2] \oplus \eta\})^3$. It turns out that

$$UC_d(\varphi_a), UC_d(\psi_a) < \infty \iff \sqrt{3}/2 < |a| \leq 1.$$
4.2. Dyadic wavelet frames with good localization

1. Let us consider the generators of normalized tight frames [2, Example 3.2] for $L_2(\mathbb{R}^+)$:

$$g_{l,s}(x) = 2^{-s} \chi_{[0,2^s]}(l, 2^{-s} x),$$

where $l \in \mathbb{N}$, $s \in \mathbb{Z}$. The Walsh–Fourier transform of $g_{l,s}$ is $\hat{g}_{l,s} = \chi_{U_{l,s}}$, where $U_{l,s} = 2^{-s}(l \oplus [0, 1))$. Suppose that $\psi = g_{l,s}$. Then $\{\psi_{j,a}\}$ is a normalized tight frame for $L_2(\mathbb{R}^+)$. For all $l \in \mathbb{N}$, $s \in \mathbb{Z}^+$ the dyadic UC is $UC_d(g_{l,s}) = \frac{1}{9}$. So, the functions $g_{l,s}$ have the same localization as the Haar function. One could think that $UC_d = 1/9$ is the least possible. There are some reasons to think so: the Haar function is the extremal function of the qualitative uncertainty principle (see the Introduction). Also, it is the eigenfunction of the Walsh–Fourier operator and it is well known that in the real-line case a function with the analogous properties (the Gaussian) do give the least $UC_H$. But Table 2 shows us that this is not the case in the dyadic setting.

2. As it was noted in Table 2, numerically $\min UC_d(f_n) \approx 0.0891$ for $n = 2$. Let us try to find a frame generator such that its dyadic UC is close to this value. Let $\psi = \chi_{[0,1]}(x) \sum_{k=0}^{3} a_k w(k,x)$. From the frame criteria, we should provide zero moment for the frame generator $\psi$ or, equivalently, $\psi(0) = 0$. Thus, we assume that $a_0 = 0$. Using Wolfram Mathematica 8.0 we solve numerically the minimization problem (5).

The coefficients are $(a_0, a_1, a_2, a_3) = (0, 0.094206, 0.551564, 0.828796)$. Using Theorem 3.2 in [2], we compute the frame bounds for the frame $\{\psi_{jk}\}$. The bounds are $A = 0.313098$ and $B = 0.695777$. The dyadic UC is $UC_d(\psi) = 0.091286$ and it is close to the minimal possible constant for $n = 2$.

The computations can be done for the case $n = 3$. Let $\psi = \chi_{[0,1]}(x) \sum_{k=0}^{7} a_k w(k,x)$. The minimum for $UC_d(\psi)$ is delivered by the coefficients $(a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7) = (0, 0.001335, -0.009155, -0.022170, -0.067567, -0.138436, -0.601657, -0.783391)$. The frame bounds for the frame $\{\psi_{jk}\}$ are $A = 0.004649$, $B = 0.614194$. The dyadic UC is $UC_d(\psi) = 0.0882147$.

Acknowledgments

The work is supported by the RFBR-grant #15-01-05796 and by the grant of the President of the RF #MK-3247.2015.1.

References