SFB Winter School

on Complexity and Discrepancy

Traunkirchen, Austria

November 30 - December 4, 2015

FOUR LECTURES ON

Quasi-Monte Carlo integration, Point distributions on the sphere and the acceptance-rejection sampler

by Josef Dick

School of Mathematics and Statistics

The University of New South Wales

Sydney

Australia

With a big thanks to the organizers of the Winter School Lucia Del Chicca, Gerhard Larcher, Peter Kritzer, Friedrich Pillichshammer and all the participants. I hope you enjoyed the Winter School as much as me!

Table of Contents

(1) Quasi-Monte Carlo (QMC) integration

We want to approximate an integral

$$I(f) = \int_{[0,1]^s} f(\underline{x}) \, dx$$

by an equal weight quadrature rule:

$$Q_P(f) = \frac{1}{N} \sum_{n=0}^{N-1} f(x_n),$$

where $\underline{x}_0, \underline{x}_1, \ldots, \underline{x}_{N-1}$ are suitably chosen quadrature points.

To study the quality of the QMC rule $Q_P(f)$ we consider the worst-case error

$$e(P, \mathcal{H}) = \sup_{\substack{f \in \mathcal{H} \\ \|f\|_{\mathcal{H}} \le 1}} \left| I(f) - Q_P(f) \right|,$$

where $\mathcal{H}$ is a suitable Banach space.

The so-called initial error

$$e(\phi, \mathcal{H}) = \sup_{\substack{f \in \mathcal{H} \\ \|f\|_{\mathcal{H}} \le 1}} \left| I(f) \right|$$

is used in Information Based Complexity as a normalizing factor when one considers the dependence on the dimension.

(1.1) ~~Order one~~ Reproducing kernel Hilbert spaces

(1.1.1) Dimension one

Consider an integrand

$$f : [0,1] \to \mathbb{R}$$

which is absolutely continuous. Hence by the fundamental theorem of calculus we can write

$$f(x) = f(1) - \int_x^1 f'(y)\,dy = f(1) - \int_0^1 f'(y)\, \underbrace{1_{[x,1]}(y)}_{= \begin{cases} 1 & \text{if } y \in [x,1] \\ 0 & \text{otherwise} \end{cases}}\, dy .$$

Thus

$$\int_0^1 f(x)\,dx - \frac{1}{N}\sum_{n=0}^{N-1} f(x_n) = \frac{1}{N}\sum_{n=0}^{N-1} \int_0^1 f'(y)\, 1_{[x_n,1]}(y)\, dy - \int_0^1\int_0^1 f'(y) 1_{[x,1]}(y)\,dy$$

$$= \int_0^1 f'(y) \left[\frac{1}{N}\sum_{n=0}^{N-1} 1_{[0,y]}(x_n) - \int_0^1 1_{[0,y]}(x)\, dx \right] dy$$

$$= \int_0^1 f'(y) \underbrace{\left[\frac{1}{N}\sum_{n=0}^{N-1} 1_{[0,y]}(x_n) - y \right]}_{=: \Delta_P(y) \text{ .. local discrepancy function}} dy .$$

Using Hölder's inequality we get

$$\left| \int_0^1 f(x)\,dx - \frac{1}{N}\sum_{n=0}^{N-1} f(x_n) \right| \le \left(\int_0^1 |f'(y)|^p\, dy \right)^{1/p} \left(\int_0^1 |\Delta_P(y)|^q \right)^{1/q}$$

$$1 \le p, q \le \infty, \ \frac{1}{p} + \frac{1}{q} = 1$$

(1.1.2) Reproducing kernel Hilbert spaces examples

In the previous one-dimensional example we used the representation

$$f(x) = f(1) - \int_0^1 f'(y)\, \mathbb{1}_{[x,1]}(y)\, dy \;\; \forall x \in [0,1]. \qquad (1.1.2.a)$$

In order for this representation to be true, we need $f(1)$ and $f'(y)$ to be well defined. Say $f' \in L_2([0,1])$. Then we can define an (inner product) norm

$$\|f\|^2 = |f(1)|^2 + \int_0^1 |f'(y)|^2\, dy . \qquad (1.1.2.b)$$

If $g: [0,1] \to \mathbb{R}$ is another function for which we have a representation (1.1.2.a), then we can define the inner product

$$\langle f, g \rangle = f(1)\, g(1) + \int_0^1 f'(y)\, g'(y)\, dy. \qquad (1.1.2.c)$$

So we can consider the space

$$\mathcal{H} = \{ f: [0,1] \to \mathbb{R} \mid f \text{ is absolutely continuous and the norm (1.1.2.b) } \|f\| < \infty \}.$$

We now combine (1.1.2.a) and (1.1.2.c):

Consider

$$f(x) = f(1)\,1 - \int_0^1 f'(y)\, 1_{[x,1]}(y)\, dy,$$

$$\langle f, g\rangle = f(1)\, g(1) + \int_0^1 f'(y)\, g'(y)\, dy .$$

If we can choose a function $g : [0,1] \to \mathbb{R}$ such that

$$g(1) = 1 \quad \text{and} \quad g'(y) = -1_{[x,1]}(y) \qquad (1.1.2\,d)$$

then $\quad f(x) = \langle f, g\rangle.$

But (1.1.2 d) implies that

$$g(y) = \begin{cases} c, & 0 \le y < x, \\ 2 - y, & x < y \le 1. \end{cases}$$

If we now choose the constant c such that g becomes continuous, then we even get $g \in \mathcal{H}$, that is,

$$g(y) = \begin{cases} 2 - x & 0 \le y \le x, \\ 2 - y & x \le y \le 1. \end{cases}$$

Define the function

$$K(x,y) = \begin{cases} 2 - x, & 0 \le y \le x, \\ 2 - y, & x \le y \le 1, \end{cases} = 1 + \min(1-x, 1-y),$$

then $\quad \langle f, K(x,\cdot)\rangle = f(x) \quad \forall x \in [0,1], \forall f \in \mathcal{H}$

(1.1.3) Definition of RKHS

The function K is the reproducing kernel of the space $\mathcal{H}$.

Def.: A Hilbert space $\mathcal{H}$ of functions $f: X \to \mathbb{R}$ on a set X with inner product $\langle \cdot, \cdot \rangle$ is called a reproducing kernel Hilbert space if there exists a function

$$K: X \times X \to \mathbb{R}$$

such that:

P1: $K(\cdot, y) \in \mathcal{H}$ for each fixed $y \in \mathcal{H}$, and

P2: $\langle f, K(\cdot, y) \rangle = f(y)$ for each $y \in X$ and $f \in \mathcal{H}$.

The function K also has the following properties:

P3: Symmetry:

$$K(x,y) = \langle K(\cdot,y), K(\cdot,x) \rangle = \langle K(\cdot,x), K(\cdot,y) \rangle = K(y,x).$$

P4: uniqueness: If $\tilde{K}$ is another function satisfying P1 and P2, then

$$\tilde{K}(x,y) = \langle \tilde{K}(\cdot,y), K(\cdot,x) \rangle = \langle K(\cdot,x), \tilde{K}(\cdot,y) \rangle = K(y,x) = K(x,y)$$

P5: positive semi-definite for all $a_1, \dots, a_M \in \mathbb{R}$ and $x_1, \dots, x_M \in X$ we have

$$\sum_{m=1}^{M} \sum_{n=1}^{M} a_n a_m K(x_n, x_m) = \sum_{n,m=1}^{M} a_n a_m \langle K(\cdot, x_n), K(\cdot, x_m) \rangle$$

$$= \left\langle \sum_{n=1}^{M} a_n K(\cdot, x_n), \sum_{m=1}^{M} a_m K(\cdot, x_m) \right\rangle = \left\| \sum_{m=1}^{M} a_m K(\cdot, x_m) \right\|^2 \geq 0.$$

(1.1.4) Integration error

We can express the worst-case integration error in terms of reproducing kernels.

Let
$$K : [0,1]^s \times [0,1]^s \to \mathbb{R}$$
be a reproducing kernel of the Hilbert space $\mathcal{H}$. Then

$$(*) \begin{cases} \int_{[0,1]^s} f(\underline{x})\, d\underline{x} = \int_{[0,1]^s} \langle f, K(\cdot, \underline{x})\rangle\, d\underline{x} = \left\langle f, \int_{[0,1]^s} K(\cdot,\underline{x})\, d\underline{x} \right\rangle, \\ \frac{1}{N}\sum_{n=0}^{N-1} f(x_n) = \frac{1}{N}\sum_{n=0}^{N-1} \langle f, K(\cdot, x_n)\rangle = \left\langle f, \frac{1}{N}\sum_{n=0}^{N-1} K(\cdot, x_n)\right\rangle. \end{cases}$$

Thus, if $(*)$ holds, we have

$$\int_{[0,1]^s} f(\underline{x})\, d\underline{x} - \frac{1}{N}\sum_{n=0}^{N-1} f(x_n) = \Big\langle f, \underbrace{\int_{[0,1]^s} K(\cdot,\underline{x})\, d\underline{x} - \frac{1}{N}\sum_{n=0}^{N-1} K(\cdot, x_n)}_{=: h} \Big\rangle$$

and therefore

$$\left| \int_{[0,1]^s} f(x)\, dx - \frac{1}{N}\sum_{n=0}^{N-1} f(x_n) \right| = \left| \langle f, h\rangle \right|.$$

Then
$$e(P, \mathcal{H}) = \sup_{\substack{f \in \mathcal{H} \\ \|f\| \le 1}} \left| \int_{[0,1]^s} f(x)\, dx - \frac{1}{N}\sum_{n=0}^{N-1} f(x_n) \right| = \frac{|\langle h, h\rangle|}{\|h\|} = \|h\|.$$

Thus

$$e^2(P, \mathcal{H}) = \left\langle \int_{[0,1]^s} K(\cdot, \underline{x})\, d\underline{x} - \frac{1}{N}\sum_{n=0}^{N-1} K(\cdot, \underline{x}_n), \int_{[0,1]^s} K(\cdot, \underline{x})\, d\underline{x} - \frac{1}{N}\sum_{n=0}^{N-1} K(\cdot, \underline{x}_n) \right\rangle$$

$$\overset{(*)}{=} \int_{[0,1]^s} \int_{[0,1]^s} K(\underline{x}, \underline{y})\, d\underline{x}\, d\underline{y} - \frac{2}{N}\sum_{n=0}^{N-1} \int_{[0,1]^s} K(\underline{x}, \underline{x}_n)\, d\underline{x} + \frac{1}{N^2}\sum_{n,m=0}^{N-1} K(\underline{x}_n, \underline{x}_m).$$

Under which assumptions does $(*)$ hold?

We have for any $f \neq 0$, that

$$\left|\frac{f(y)}{\|f\|}\right| = \left|\left\langle \frac{f}{\|f\|}, K(\cdot, y)\right\rangle\right| \leq \left|\left\langle \frac{K(\cdot, y)}{\|K(\cdot, y)\|}, K(\cdot, y)\right\rangle\right|$$

$$= \left|\frac{\langle K(\cdot, y), K(\cdot, y)\rangle}{\|K(\cdot, y)\|}\right| = \sqrt{\langle K(\cdot, y), K(\cdot, y)\rangle} = \sqrt{K(y,y)}.$$

Hence

$$|f(y)| \leq \|f\| \sqrt{K(y,y)}.$$

Thus, if we assume, for instance, that

$$\int_{[0,1]^s} \sqrt{K(y,y)}\, dy < \infty,$$

then $\int_{[0,1]^s} f(x)\, dx$ and $\frac{1}{N}\sum_{n=0}^{N-1} f(x_n)$ are well defined.

Another property which we need is

$$\int_{[0,1]^s} \langle f, K(\cdot, \underline{x})\rangle\, d\underline{x} = \left\langle f, \int_{[0,1]^s} K(\cdot, \underline{x})\, d\underline{x} \right\rangle.$$

In fact, we show this for general linear functionals.

Let T be any bounded linear functional on the rkHs $\mathcal{H}$:

$$T : \mathcal{H} \to \mathbb{R},$$

linear: $T(f+g) = T(f) + T(g)$, $T(\alpha f) = \alpha T(f)$, $\forall f, g \in \mathcal{H}, \alpha \in \mathbb{R}$.

bounded: $|T(f)| \le C \|f\| \quad \forall f \in \mathcal{H}$, $C > 0$ indep. of f.

Riesz representation theorem implies that $\exists R \in \mathcal{H}$ s.t.

$$T(f) = \langle f, R \rangle \quad \forall f \in \mathcal{H}.$$

Then

$$R(\underline{x}) = \langle R, K(\cdot, \underline{x}) \rangle = \langle K(\cdot, x), R \rangle = T(K(\cdot, x)).$$

Hence

$$\underbrace{T(\langle f, K(\cdot, x) \rangle)}_{\text{T with respect to variable } x} = T(f) = \langle f, R \rangle = \langle f, T(K(\cdot, x)) \rangle.$$

Hence we can always change the order of inner products and bounded linear functionals.

(1.1.5) Connection to discrepancy theory

Consider the reproducing kernel

$$K : [0,1]^s \times [0,1]^s \to \mathbb{R}$$

given by

$$K(\underline{x},\underline{y}) = \prod_{j=1}^{s} \left(\min(1-x_j, 1-y_j) \right).$$

$\underline{x} = (x_1 x_2 \dots x_s)$
$\underline{y} = (y_1 y_2 \dots y_s)$

We can write this kernel as

$$K(\underline{x},\underline{y}) = \int_{[0,1]^s} 1_{[\underline{x},\underline{1}]}(\underline{z}) \, 1_{[\underline{y},\underline{1}]}(\underline{z}) \, d\underline{z}.$$

The corresponding r.k.Hs contains all functions of the form

$$f_1(\underline{x}) = \int_{[0,1]^s} g_1(\underline{z}) \, 1_{[\underline{x},\underline{1}]}(\underline{z}) \, d\underline{z}, \quad g_1 \in L_2([0,1]^s)$$

with inner product and norm

$$\langle f_1, f_2 \rangle = \int_{[0,1]^s} g_1(\underline{z}) g_2(\underline{z}) \, d\underline{z}, \quad \|f_1\| = \left(\int |g_1(\underline{z})|^2 d\underline{z} \right)^{1/2}$$

The integration error is then bounded in the following way:

$$\left| \int_{[0,1]^s} f(\underline{x}) \, d\underline{x} - \frac{1}{N} \sum_{n=0}^{N-1} f(\underline{x}_n) \right| = \left| \int_{[0,1]^s} g(\underline{z}) \left[\int_{[0,1]^s} 1_{[\underline{0},\underline{z}]}(\underline{x}) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} 1_{[\underline{0},\underline{z}]}(x_n) \right] d\underline{z} \right|$$

$$\leq \left(\int_{[0,1]^s} |g(\underline{z})|^2 d\underline{z} \right)^{1/2} \underbrace{\left(\int_{[0,1]^s} \Big[\underbrace{\int_{[0,1]^s} 1_{[\underline{0},\underline{z}]}(\underline{x}) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} 1_{[\underline{0},\underline{z}]}(\underline{x}_n)}_{\Delta_P(\underline{z})} \Big]^2 d\underline{z} \right)^{1/2}}_{L_2 \text{ discrepancy.}}$$

Extensions:

- Use Hölder inequality:

$$\left| \int_{[0,1]^s} f(x)\,dx - \frac{1}{N}\sum_{n=0}^{N-1} f(x_n) \right| \le \left(\int_{[0,1]^s} |g(z)|^p\,dz \right)^{1/p} \left(\int_{[0,1]^s} |\Delta_p(z)|^q\,dz \right)^{1/q}$$

$$\frac{1}{p}+\frac{1}{q}=1,\; p,q\ge 1.$$

- Different smoothness:

replace indicator function $1_{[0,z]}(x)$ with the truncated power function $(x-z)_+^{\alpha-1}$, $\alpha > \frac{1}{2}$.

Note that, for instance

$$f(x) = f(0) + x f'(0) + \int_0^1 (x-w)_+ f''(w)\,dw$$

$$(z)_+ = \max(z,0).$$

- Different domain, test sets, general measures, ...

Sphere with spherical caps as test sets;
cube with convex sets with smooth boundary as test sets;

(1.1.6) Some examples

We consider only tensor product spaces with reproducing kernels of the form

$$K(\underline{x},\underline{y}) = \prod_{j=1}^{s} K(x_j, y_j), \quad \underline{x} = (x_1, \dots, x_s), \underline{y} = (y_1, \dots, y_s) \in [0,1)$$

⊙ We have seen already the kernel

$$K(x,y) = 1 + \min(1-x, 1-y).$$

The corresponding inner product is

$$\langle f, g \rangle = f(1)\, g(1) + \int_0^1 f'(x)\, g'(x)\, dx.$$

⊙ Another example is

$$K_\alpha(x,y) = 1 + \frac{B_1(x)}{1!}\frac{B_1(y)}{1!} + \frac{B_2(x)}{2!}\frac{B_2(y)}{2!} + \dots + \frac{B_\alpha(x)}{\alpha!}\frac{B_\alpha(y)}{\alpha!}$$
$$+ (-1)^\alpha \frac{B_{2\alpha}(|x-y|)}{(2\alpha)!},$$

where B_τ is the Beroulli polynomial of degree τ.

The corresponding inner product is

$$\langle f, g \rangle = \int_0^1 f(x)\,dx \int_0^1 g(x)\,dx + \int_0^1 f'(x)\,dx \int_0^1 g'(x)\,dx + \dots + \int_0^1 f^{(\alpha-1)}(x)\,dx \int_0^1 g^{(\alpha-1)}(x)\,dx$$
$$+ \int_0^1 f^{(\alpha)}(x)\, g^{(\alpha)}(x)\,dx.$$

We can obtain another example from Taylor's theorem with integral remainder:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(\alpha-1)}(0)}{(\alpha-1)!}x^{\alpha-1}$$
$$+ \int_0^1 \frac{f^{(\alpha)}(y)}{(\alpha-1)!}(x-y)_+^{\alpha-1}\,dy,$$

where
$$(x-y)_+^{\alpha-1} = \begin{cases} (x-y)^{\alpha-1} & \text{if } 0 \le y \le x \\ 0 & \text{if } x < y \le 1. \end{cases}$$

is the truncated power function.

Define a reproducing kernel by

$$K(x,y) = 1 + xy + \frac{x^2}{2!}\frac{y^2}{2!} + \dots + \frac{x^{\alpha-1}}{(\alpha-1)!}\frac{y^{\alpha-1}}{(\alpha-1)!} + \int_0^1 \frac{(x-z)_+^{\alpha-1}}{(\alpha-1)!}\frac{(y-z)_+^{\alpha-1}}{(\alpha-1)!}\,dz$$

and inner product by

$$\langle f, g\rangle = f(0)g(0) + f'(0)g'(0) + \dots + f^{(\alpha-1)}(0)g^{(\alpha-1)}(0) + \int_0^1 f^{(\alpha)}(z)g^{(\alpha)}(z)\,dz$$

It can be checked that

$$\langle f, K(\cdot,y)\rangle = f(0) + f'(0)y + \dots + \frac{f^{(\alpha-1)}(0)}{(\alpha-1)!}y^{\alpha-1} + \int_0^1 \frac{f^{(\alpha)}(z)}{(\alpha-1)!}(x-z)_+^{\alpha-1}\,d$$
$$= f(y).$$

The kernel based on Bernoulli polynomials is based on the following expansion:

$$f(x) = \int_0^1 f(y)\,dy\; B_0(x) + \int_0^1 f'(y)\,dy\; B_1(x) + \int_0^1 f''(y)\,dy\; \frac{B_2(x)}{2!} + \cdots$$

$$\cdots + \int_0^1 f^{(\alpha-1)}(y)\,dy\; \frac{B_{\alpha-1}(x)}{(\alpha-1)!} + \int_0^1 f^{(\alpha)}(y)\,dy\; \frac{B_\alpha(x)}{\alpha!}$$

$$-(-1)^\alpha \int_0^1 f^{(\alpha)}(z)\, \tilde{b}_\alpha(x-z)\,dz,$$

where

$$\tilde{b}_\alpha(x-y) = \begin{cases} \dfrac{B_\alpha(|x-y|)}{\alpha!} & \text{for } \alpha \text{ even}, \\ (-1)^{1_{x<y}} \dfrac{B_\alpha(|x-y|)}{\alpha!} & \text{for } \alpha \text{ odd} \end{cases}$$

Another example is the following reproducing kernel based on Fourier series

$$K(x,y) = \sum_{\underline{k} \in \mathbb{Z}^s} r_\alpha(\underline{k})\, e^{2\pi i \underline{k}\cdot(\underline{x}-\underline{y})}$$

with inner product

$$\langle f, g \rangle = \sum_{\underline{k} \in \mathbb{Z}^s} \hat{f}(\underline{k})\, \overline{\hat{g}(\underline{k})}\, \frac{1}{r_\alpha(\underline{k})},$$

for Fourier series

$$f(\underline{x}) = \sum_{\underline{k} \in \mathbb{Z}^s} \hat{f}(\underline{k})\, e^{2\pi i \underline{k}\cdot\underline{x}}$$

(1.2) Walsh functions

(1.2.1) Definition

For an integer $b \geq 2$ let $\omega_b = e^{2\pi i/b}$ be the b-th root of unity.

<u>Def</u>: Let $b \geq 2$ be an integer. Let $k \in \mathbb{N}_0$ be an integer (non-negative) with b-adic expansion

$$k = k_0 + k_1 b + \dots + k_{m-1} b^{m-1}.$$

Let $x \in [0,1)$ have b-adic expansion

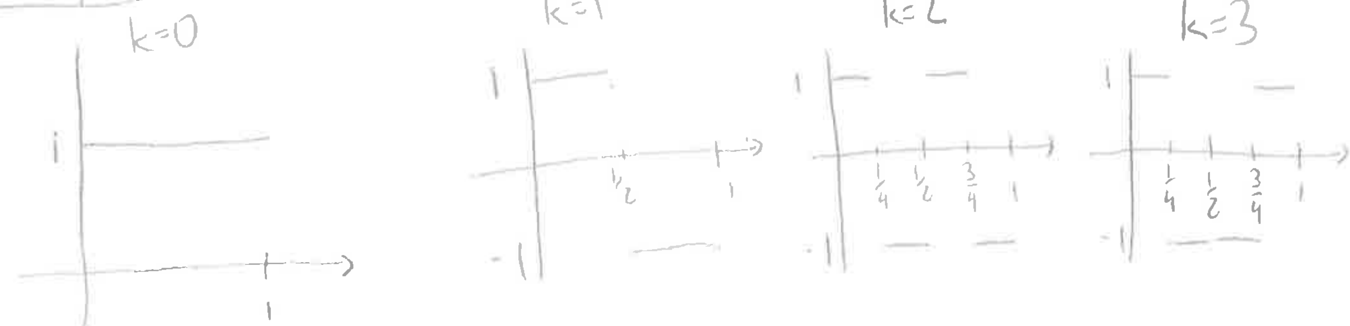
$$x = \frac{\xi_1}{b} + \frac{\xi_2}{b^2} + \dots$$

unique in the sense that infinitely many of the digits ξ_i are different from $b-1$. Then the b-adic Walsh function at x is

$${}_b\mathrm{wal}_k(x) = \omega_b^{k_0 \xi_1 + k_1 \xi_2 + \dots + k_{m-1} \xi_m}.$$

The system $\{ {}_b\mathrm{wal}_k : k \in \mathbb{N}_0 \}$ is called the (b-adic) Walsh function system.

<u>Examples</u>: $b = 2$



For dimension $s \geq 2$ and $\underline{k} = (k_1, k_2, \dots, k_s) \in \mathbb{N}_0^s$, we define the s-dimensional b-adic Walsh functions by

$$ {}_b\mathrm{wal}_{\underline{k}}(\underline{x}) = \prod_{j=1}^{s} {}_b\mathrm{wal}_{k_j}(x_j), $$

where $\underline{x} = (x_1, x_2, \dots, x_s) \in [0,1)^s$.

The system $\{ {}_b\mathrm{wal}_{\underline{k}} : \underline{k} \in \mathbb{N}_0^s \}$ is called the s-dimensional b-adic Walsh function system.

(1.2.2) Basic Properties

For $x = \frac{\xi_1}{b} + \frac{\xi_2}{b^2} + \dots$ and $y = \frac{\eta_1}{b} + \frac{\eta_2}{b^2} + \dots$ ~~in [0,1)~~ we define the (b-adic) digitwise addition modulo b by

$$ x \oplus y = \frac{z_1}{b} + \frac{z_2}{b} + \dots, \quad \text{where } z_i = \xi_i + \eta_i \pmod b. $$

We also define (b-adic) digitwise subtraction modulo b by

$$ x \ominus y = \frac{z_1}{b} + \frac{z_2}{b} + \dots, \quad \text{where } z_i = \xi_i - \eta_i \pmod b. $$

Analoguously we can define these operations also for non-negative integers.

Walsh functions have the following properties:

Let $k = \kappa_0 + \kappa_1 b_1 + \kappa_2 b_2 + \dots$ and $l = \lambda_0 + \lambda_1 b + \lambda_2 b^2 + \dots$ be non-negative integers. Then

(o) $$ \mathrm{wal}_k(x)\,\mathrm{wal}_l(x) = \omega_b^{\kappa_0 \xi_1 + \kappa_1 \xi_2 + \dots} \, \omega_b^{\lambda_0 \xi_1 + \lambda_1 \xi_2 + \dots} $$
$$ = \omega_b^{(\kappa_0 + \lambda_0)\xi_1 + (\kappa_1 + \lambda_1)\xi_2 + \dots} $$
$$ = \mathrm{wal}_{k \oplus l}(x). $$

(c) $\mathrm{wal}_k(x)\,\mathrm{wal}_k(y) = \omega_b^{\kappa_0\xi_1+\kappa_1\xi_2+\cdots}\,\omega_b^{\kappa_0\eta_1+\kappa_1\eta_2+\cdots}$

$$= \omega_b^{\kappa_0(\xi_1+\eta_1)+\kappa_1(\xi_2+\eta_2)+\cdots}$$

$$= \mathrm{wal}_k(x \oplus y).$$

for almost all $x, y \in [0,1)$. Exceptions are of the form $x = 0.010101\ldots$, $y = 0.001010\ldots$, $x \oplus y = 0.1$

We also have

- $\dfrac{1}{\mathrm{wal}_k(x)} = \overline{\mathrm{wal}_k(x)} = \mathrm{wal}_{0 \ominus k}(x) = \mathrm{wal}_k(0 \ominus x).$

- $\mathrm{wal}_{\underline{k}}(\underline{x})\,\overline{\mathrm{wal}_{\underline{k}}(\underline{y})} = \mathrm{wal}_{\underline{k}}(\underline{x} \ominus \underline{y})$

- $\mathrm{wal}_{\underline{k}}(\underline{x})\,\mathrm{wal}_{\underline{k}}(\underline{y}) = \mathrm{wal}_{\underline{k}}(\underline{x} \oplus \underline{y})$.

for all $(x, y) \in [0,1)^2$ except for a set of measure 0.

__Lemma:__ For $1 \le k < b^r$ we have

$$\sum_{a=0}^{b^r-1} \mathrm{wal}_k\left(\frac{a}{b^r}\right) = 0.$$

Proof: Note that for $\kappa \in \{1, 2, \ldots, b-1\}$ we have

$$\sum_{a=0}^{b-1} \omega_b^{\kappa a} = \frac{1-\omega_b^{\kappa b}}{1-\omega_b} = \frac{1-e^{2\pi i \kappa b/b}}{1-\omega_b} = \frac{1-e^{2\pi i \kappa}}{1-\omega_b} = 0.$$

Hence for $k = \kappa_0 + \kappa_1 b + \cdots + \kappa_{r-1} b^{r-1}$ we have

$$\sum_{a=0}^{b^r-1} \mathrm{wal}_k\left(\frac{a}{b^r}\right) = \sum_{a_0,\ldots,a_{r-1}=0}^{b-1} \omega_b^{\kappa_0 a_{r-1}+\cdots+\kappa_{r-1}a_0} = \prod_{i=0}^{r-1} \sum_{a_i=0}^{b-1} \omega_b^{a_i \kappa_{r-1-i}} = 0$$

since $k \in \{1, 2, \ldots, b^r - 1\}$ implies that there is at least one $\kappa_i \neq 0$. $\square$

Lemma: We have

$$\int_0^1 {}_b\mathrm{wal}_k(x)\,dx = \begin{cases} 1 & \text{if } k=0, \\ 0 & \text{if } k\in\mathbb{N}. \end{cases}$$

Proof: We have ${}_b\mathrm{wal}_k(x) = \omega_b^{\kappa_0\xi_1+\kappa_1\xi_2+\dots+\kappa_{r-1}\xi_r}$ for $k\in\mathbb{N}$ with $k=\kappa_0+\kappa_1 b+\dots+\kappa_{r-1}b^{r-1}$ and $x=\frac{\xi_1}{b}+\frac{\xi_2}{b^2}+\dots+\frac{\xi_r}{b^r}+\dots$. Since the Walsh function ${}_b\mathrm{wal}_k$ does not depend on the digits ξ_i with $i>r$, it is constant on intervals of the form $[a b^{-r}, (a+1)b^{-r})$. Therefore, for ~~k~~ $k\in\mathbb{N}$ we have

$$\int_0^1 {}_b\mathrm{wal}_k(x)\,dx = \frac{1}{b^r}\sum_{a=0}^{b^r-1} \mathrm{wal}_k\left(\frac{a}{b^r}\right) = 0 .$$

If $k=0$ we have ${}_b\mathrm{wal}_0 \equiv 1$, hence $\int_0^1 {}_b\mathrm{wal}_0(x)\,dx = 1$. $\square$

Corollary: We have

$$\int_0^1 \mathrm{wal}_k(x)\,\overline{\mathrm{wal}_\ell(x)}\,dx = \begin{cases} 0 & \text{if } k\neq \ell \\ 1 & \text{if } k=\ell. \end{cases}$$

Proof: Follows from previous lemma using

$$\mathrm{wal}_k(x)\,\overline{\mathrm{wal}_\ell(x)} = \mathrm{wal}_{k\ominus\ell}(x) .$$

$\square$

The same result holds in the multidimensional case

$$\int_{[0,1]^s} \mathrm{wal}_{\underline{k}}(\underline{x})\,\overline{\mathrm{wal}_{\underline{\ell}}(\underline{x})}\,d\underline{x} = \begin{cases} 0 & \text{if } \underline{k}\neq\underline{\ell} \\ 1 & \text{if } \underline{k}=\underline{\ell} \end{cases}$$

Thus the Walsh function system is orthonormal in $L_2([0,1]^s)$

The Walsh function system is also complete, i.e. we have Parseval's identity

$$\int_{[0,1]^s} |f(\underline{x})|^2 \, d\underline{x} = \sum_{\underline{k} \in \mathbb{Z}^s} |\hat{f}(\underline{k})|^2, \quad \forall f \in L_2([0,1]^s).$$

When we study numerical integration, we will use Walsh series expansions of the integrand. Since QMC uses function evaluations, we want

$$f(\underline{x}) = \sum_{\underline{k} \in \mathbb{N}_0^s} \hat{f}(\underline{k}) \, \mathrm{wal}_{\underline{k}}(\underline{x}) \quad \forall \underline{x} \in [0,1)^s$$

pointwise.

For functions from the reproducing kernel Hilbert spaces from before, this always holds.

Note that if a function $f : [0,1) \to \mathbb{R}$ is merely continuous, then there are examples where the function f does not coincide with its Walsh series at some point. However, if the function is continuous and has bounded variation, then pointwise convergence holds.

Let $f : [0,1]^s \to \mathbb{R}$ be in $L_2([0,1]^s)$. Then for $\underline{k} \in \mathbb{N}_0^s$ we define the $\underline{k}$-th Walsh coefficient by

$$\hat{f}(\underline{k}) = \int_{[0,1]^s} f(\underline{x}) \overline{\mathrm{wal}_{\underline{k}}(\underline{x})} \, d\underline{x} .$$

The Walsh series of f is given by

$$f(\underline{x}) \sim \sum_{\underline{k} \in \mathbb{N}_0^s} \hat{f}(\underline{k}) \, \mathrm{wal}_{\underline{k}}(\underline{x})$$

(here $\sim$ means we have equality in the L_2 sense

$$\int_{[0,1]^s} |f(\underline{x})|^2 \, d\underline{x} = \sum_{\underline{k} \in \mathbb{N}_0^s} |\hat{f}(\underline{k})|^2).$$

If the function f is in one of the r.k.Hs. discussed above, then we even have

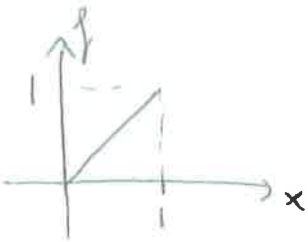
$$f(x) = \sum_{\underline{k} \in \mathbb{N}_0^s} \hat{f}(\underline{k}) \, \mathrm{wal}_{\underline{k}}(\underline{x}) \quad \forall \underline{x} \in [0,1)^s \text{ pointwise.}$$

(1.2.3) Decay of Walsh coefficients of smooth functions.

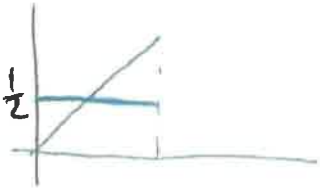
We make some simplifying assumptions: We consider $s=1$ and Walsh functions in base $b=2$.

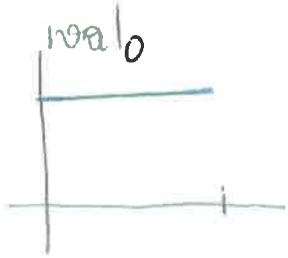
Polynomials:

Consider $f(x) = x$.



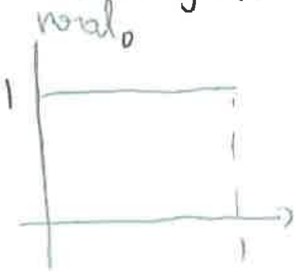
$k=0$

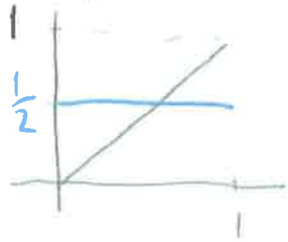




Freq   Walsh fct.   Approximation   Walsh polynomi

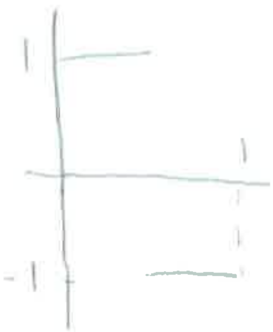
$k=0$:

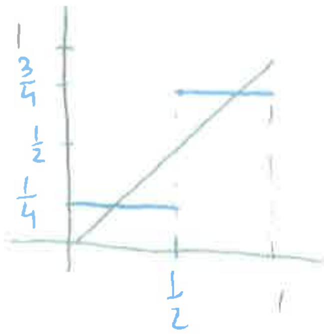




$x \sim \frac{1}{2} \mathrm{wal}_0(x)$

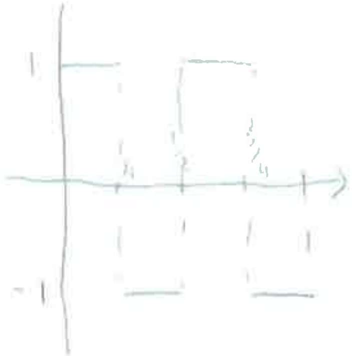
$k=1$:

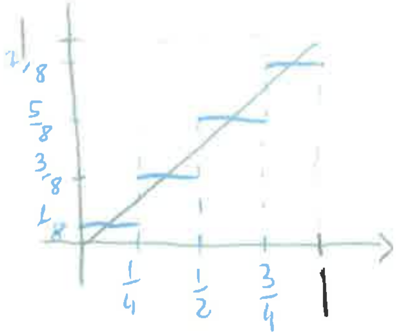




$x \sim \frac{1}{2} \mathrm{wal}_0(x) - \frac{1}{4} \mathrm{wal}_1(x)$

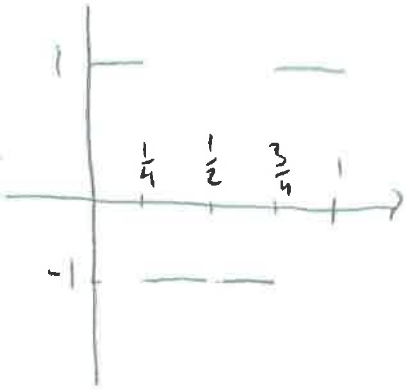
$k=2$:

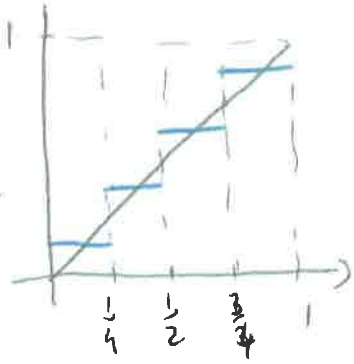




$$x \sim \tfrac{1}{2}\,\mathrm{wal}_0(x) - \tfrac{1}{4}\,\mathrm{wal}_1(x) - \tfrac{1}{8}\,\mathrm{wal}_2(x)$$

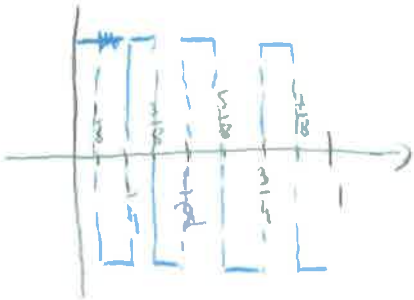
$k=3$:

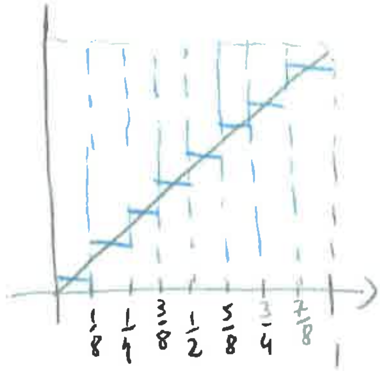




$$x \sim \tfrac{1}{2}\,\mathrm{wal}_0(x) - \tfrac{1}{4}\,\mathrm{wal}_1(x) - \tfrac{1}{8}\,\mathrm{wal}_2(x) + 0\,\mathrm{wal}_3(x)$$

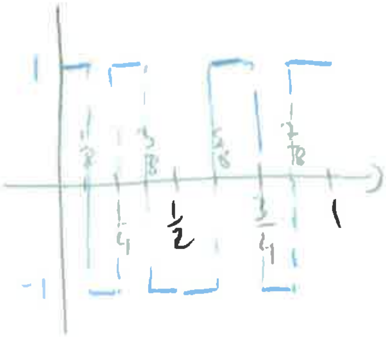
$k=4$:

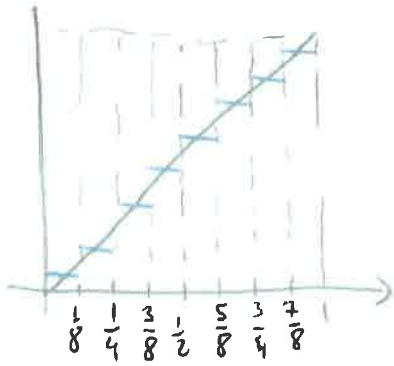




$$x \sim \tfrac{1}{2}\,\mathrm{wal}_0(x) - \tfrac{1}{4}\,\mathrm{wal}_1(x - \tfrac{1}{8}\,\mathrm{wal}_2(x) + 0\,\mathrm{wal}_3(x - \tfrac{1}{16}\,\mathrm{wal}_4(x)$$

$k=5$:





$$x \sim \tfrac{1}{2}\,\mathrm{wal}_0(x) - \tfrac{1}{4}\,\mathrm{wal}_1(x - \tfrac{1}{8}\,\mathrm{wal}_2(x) + 0\,\mathrm{wal}_3(x - \tfrac{1}{16}\,\mathrm{wal}_4(x) + 0\,\mathrm{wal}_5(x$$

In general

$$x = \tfrac{1}{2}\,\mathrm{wal}_0(x) - \tfrac{1}{4}\,\mathrm{wal}_1(x) - \tfrac{1}{8}\,\mathrm{wal}_2(x) - \tfrac{1}{16}\,\mathrm{wal}_4(x) - \cdots$$

$$= \tfrac{1}{2}\,\mathrm{wal}_0(x) - \sum_{a=0}^{\infty} \frac{1}{2^{a+2}}\,\mathrm{wal}_{2^a}(x)$$

23

We have

$$\hat{f}(k) = \begin{cases} \frac{1}{2} & \text{if } k = 0, \\ -\frac{1}{2^{a+2}} & \text{if } k = 2^a, \\ 0 & \text{otherwise}, \end{cases}$$

The function $f(x) = x$ is of course infinitely many times differentiable. However, the Walsh coefficients decay only with rate

$$|\hat{f}(k)| \lesssim \frac{1}{k}. \qquad k \in \mathbb{N}_0$$

If we use the sparsity and ignore all Walsh coefficients which are 0, then

$$|\hat{f}(2^a)| \lesssim \frac{1}{2^a}, \qquad a \in \mathbb{N}_0 \quad (\text{exponential decay})$$

Using the sparsity, the function $f(x) = x$ can be approximated very well using Walsh functions.

Consider now

$$\frac{1}{2} - x = \sum_{a=0}^{\infty} \frac{1}{2^{a+2}} \, \mathrm{wal}_{2^a}(x)$$

Then

$$\left(\frac{1}{2} - x\right)^2 = \sum_{a_1, a_2 = 0}^{\infty} \frac{1}{2^{a_1+2} \, 2^{a_2+2}} \, \mathrm{wal}_{2^{a_1}}(x) \, \mathrm{wal}_{2^{a_2}}(x)$$

$$= \underbrace{\sum_{a_1 = a_2 = 0}^{\infty} \frac{1}{2^{2a_1+4}}}_{= \frac{1}{2^4} \sum_{a_1=0}^{\infty} \frac{1}{4^{a_1}} = \frac{1}{12}} + 2 \sum_{0 \le a_2 < a_1} \frac{1}{2^{a_1+a_2+4}} \, \mathrm{wal}_{2^{a_1}+2^{a_2}}(x)$$

$$= \frac{1}{12} + \frac{1}{2^3} \sum_{a_1 > a_2 \ge 0} \frac{1}{2^{a_1+a_2}} \, \mathrm{wal}_{2^{a_1}+2^{a_2}}(x).$$

Thus for the function $f(x) = \left(\frac{1}{2} - x\right)^2$ we have

$$\hat{f}(k) = \begin{cases} \frac{1}{12} & \text{if } k = 0 \\ \frac{1}{2^3} \frac{1}{2^{a_1+a_2}} & \text{if } k = 2^{a_1} + 2^{a_2}, \; 0 \le a_2 < a_1 \\ 0 & \text{otherwise.} \end{cases}$$

In general: $f(x) = \left(\frac{1}{2} - x\right)^\alpha$

$$|\hat{f}(k)| \lesssim \begin{cases} 1 & \text{if } k = 0 \\ \dfrac{1}{2^{a_1 + \ldots + a_\nu + \nu}} & \text{if } k = 2^{a_1} + 2^{a_2} + \ldots + 2^{a_\nu},\ 1 \le \nu \le \alpha \\ 0 & \text{otherwise.} \end{cases}$$

Consider now functions f in the rkHs with kernel

$$K(x,y) = 1 + xy + \frac{x^2}{2!}\frac{y^2}{2!} + \ldots + \frac{x^{\alpha-1}}{(\alpha-1)!}\frac{y^{\alpha-1}}{(\alpha-1)!} + \int_0^1 \frac{(x-z)_+^{\alpha-1}}{(\alpha-1)!}\frac{(y-z)_+^{\alpha-1}}{(\alpha-1)!}\,dz$$

and inner product

$$\langle f, g\rangle = f(0)g(0) + \ldots + f^{(\alpha-1)}(0)\,g^{(\alpha-1)}(0) + \int_0^1 f^{(\alpha)}(z)\,g^{(\alpha)}(z)\,dz.$$

Functions in this space can be represented by the Taylor series with integral remainder:

$$f(x) = f(0) + f'(0)y + \frac{f''(0)}{2!}y^2 + \ldots + \frac{f^{(\alpha-1)}(0)}{(\alpha-1)!}y^{\alpha-1} + \int_0^1 \frac{f^{(\alpha)}(z)}{\alpha!}(x-z)_+^{\alpha-1}\,dz$$

We have already studied the decay of the Walsh coefficients of polynomials.

Consider now Fine's function

$$J_k(x) = \int_0^x \mathrm{wal}_k(y)\,dy$$

Fine showed that

$$J_k(x) = \frac{1}{2^{a_1+1}}\left[\mathrm{wal}_{k'}(x) - \sum_{c=1}^{\infty} \frac{1}{2^{c+1}}\,\mathrm{wal}_{2^{a_1+c-1}+k}(x)\right] \qquad (1.2.3.a)$$

where $k = 2^{a_1-1} + k'$, $0 \le k' < 2^{a_1-1}$.

We can also define iterates of J_k by $J_k^{(1)}(x) = J_k(x)$ and

$$J_k^{(\eta)}(x) = \int_0^x J_k^{(\eta-1)}(y)\,dy.$$

It is clear that $J_k^{(\eta)}(0) = 0$. Let $k = 2^{a_1-1} + 2^{a_2-1} + \dots + 2^{a_\nu-1}$, $a_1 > a_2 > \dots > a_\nu > 0$. Then by iteratively using (1.2.3.a) we have for $\eta < \nu$ that

$$J_k^{(\eta)}(x) = \frac{1}{2^{a_1+a_2+\dots+a_\eta+\eta}}\,\mathrm{wal}_{2^{a_{\eta+1}-1}+\dots+2^{a_\nu-1}}(x) + \sum_{l > 2^{a_{\eta+1}-1}+\dots+2^{a_{\eta+1}-1}} \hat{c}_l\,\mathrm{wal}_l(x).$$

Hence

$$J_k^{(\eta+1)}(1) = \int_0^1 J_k^{(\eta)}(x)\,dx = 0. \qquad (1.2.3.b)$$

Now we use integration by parts to obtain

$$\hat{f}(k) = \int_0^1 f(x)\, wal_k(x)\, dx = \underbrace{f(x)\, J_k^{(1)}(x)\Big|_0^1}_{= f(1) J_k^{(1)}(1) - f(0) J_k^{(1)}(0) = 0} - \int_0^1 f'(x)\, J_k^{(1)}(x)\, dx$$

$$= - \int_0^1 f'(x)\, J_k^{(1)}(x)\, dx = \int_0^1 f^{(2)} J_k^{(2)}(x)\, dx$$

$$= (-1)^\eta \int_0^1 f^{(\eta)}(x)\, J_k^{(\eta)}(x)\, dx .$$

How large can we choose η? We must have $\eta \le \alpha$ on the one hand and on the other hand we must have $\eta \le \nu$ so that (1.2.3.b) holds. Set $\eta = \min(\alpha, \nu)$. Then

$$|\hat{f}(k)| \le \left| \int_0^1 f^{(\eta)}(x)\, J_k^{(\eta)}(x)\, dx \right|$$

$$\le \left(\int_0^1 |f^{(\eta)}(x)|^p \right)^{1/p} \left(\int_0^1 |J_k^{(\eta)}(x)|^q\, dx \right)^{1/q}, \quad 1 \le p, q \le \infty, \quad \frac{1}{p} + \frac{1}{q} = 1 .$$

To get bounds on the decay of $|\hat{f}(k)|$ we need to estimate $\left(\int_0^1 |J_k^{(\eta)}(x)|^q\, dx \right)^{1/q}$.

We can write

$$J_k(x) = \frac{1}{2^{a_1+1}} \mathrm{wal}_{k'}(x)\left[1 + \underbrace{\mathrm{wal}_{2^{a_1-1}}(x)\left(-\sum_{c=1}^{\infty}\frac{1}{2^c}\mathrm{wal}_{2^{a_1+c-1}}(x)\right)}\right]$$

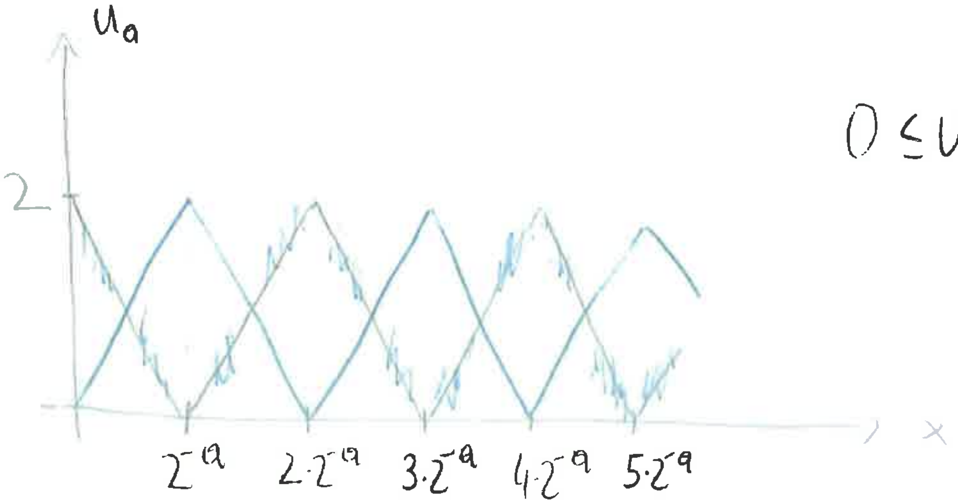
$$= -\sum_{c=1}^{\infty}\frac{1}{2^c}\mathrm{wal}_{2^{c-1}}(\{x2^{a_1}\})$$

$$= -2\sum_{c=0}^{\infty}\frac{1}{2^{c+2}}\mathrm{wal}_{2^c}(\{x2^{a_1}\})$$

$$= +2\left(\{x2^{a_1}\} - \frac{1}{2}\right)$$

$$= -1 + 2\{x2^{a_1}\}$$

Let $u_a = 1 + \mathrm{wal}_{2^{a-1}}(x)\left[-1 + 2\{x2^a\}\right]$.



$$0 \le u_a(x) \le 2$$

$$J_k(x) = \frac{1}{2^{a_1+1}}\mathrm{wal}_{k'}(x)\, u_a(x)$$

We can do a rough estimation of $J_k^{(\eta)}$ for $\eta = 1, 2$;

$$\left| J_k^{(1)}(x) \right| \le \frac{1}{2^{a_1+1}} \underbrace{\left| wal_{k'}(x) \right|}_{=1} \underbrace{\left| u_a(x) \right|}_{\le 2} = \frac{1}{2^{a_1}}$$

$$\left| J_k^{(2)}(x) \right| = \left| \int_0^x J_k^{(1)}(y)\, dy \right| = \frac{1}{2^{a_1+1}} \left| \int_0^x wal_{k'}(y)\, u_{a_1}(y)\, dy \right|$$

$$\le \frac{1}{2^{a_1+1}} \cancel{\frac{1}{2^{a_2+1}}}\, 2 \left| \int_0^x wal_{k'}(y)\, dy \right| \le \frac{1}{2^{a_1+a_2}}$$

In general

$$\left| J_k^{(\eta)}(x) \right| \le C_\eta \frac{1}{2^{a_1+a_2+\dots+a_\eta}}$$

Thus we have

$$\left| \hat{f}(k) \right| \le C_\eta \frac{1}{2^{a_1+\dots+a_\eta}} \left\| f^{(\eta)} \right\|_{L_p}$$

For $k \in \mathbb{N}$ we define

$$\mu_\alpha(k) = a_1 + a_2 + \dots + a_{\min(\alpha, v)} \quad \text{for } k = 2^{a_1-1} + 2^{a_2-1} + \dots + 2^{a_v-1}, \quad a_1 > a_2 > \dots > a_v$$

and

$$\mu_\alpha(0) = 0.$$

For $\underline{k} = (k_1, k_2, \dots, k_s) \in \mathbb{N}_0^s$ we set

$$\mu_\alpha(\underline{k}) = \mu_\alpha(k_1) + \dots + \mu_\alpha(k_s).$$