New Construction of Two Dimensional Low Discrepancy Sequences

Makoto MORI * and Masaki MORI**
(Received October 31, 2011)

Abstract

We construct a two dimensional random sequence of low discrepancy, using the prime field $\mathbb{F}_2$ of characteristic 2, and the irreducible polynomial $\beta^2 + \beta + 1 = 0$ over $\mathbb{F}_2$. This method seems to be able to extend to higher dimensional cases. We have checked in computer experiments for several cases. But until now, we can not yet prove for dimensions greater than 2.

1 Introduction

Let for a 1–dimensional transformation $F$ from the unit interval into itself

$$\xi = \lim_{n \to \infty} \inf_n \frac{1}{n} \text{ess inf}_x \log |F^n'(x)|.$$ 

Then, $e^{-\xi}$ is the radius of essential spectra of the Perron–Frobenius operator restricted to the set of functions with bounded variation. This is the essential point that we can construct a low discrepancy sequences if there exists no eigenvalue except 1 in the annulus $e^{-\xi} < |z| \leq 1$.

For higher dimensional cases, even if we restrict the domain of the Perron–Frobenius operator to a suitable space, the radius of essential spectra usually bigger than $e^{-\xi}$. Here, we change $F^n$ to the Jacobian of $F^n$. Thus to construct low discrepancy sequences, we first need to construct a space which includes all the indicator functions of intervals, and find a transformation whose radius of essential spectra equals $e^{-\xi}$.

1.1 Random Sequences

Let us consider $I = [0, 1]^d$. A random sequence $x_1, x_2, \ldots \in I$ is called uniformly distributed if $\lim_{N \to \infty} D_N(J) = 0$ for any interval $J = \prod_{i=1}^d [a_i, b_i) (0 \leq a_i < b_i < 1)$.
$b_i \leq 1$). Here
\[
D_N(J) = \left| \frac{1}{N} \# \{n \leq N : x_n \in J \} - |J| \right|,
\]
where $|J|$ is the Lebesgue measure of $J$. The discrepancy $D_N$ is defined by
\[
D_N = \sup_J D_N(J),
\]
where sup is taken over all the intervals $J$. When we restrict intervals only for $J = \prod_{i=1}^d [0, b_i)$, we call it $*$-discrepancy and denote it by $D^*_N$. There exists a constant $C_*$ such that
\[
D^*_N \leq D_N \leq C_* D^*_N.
\]
Thus there exists no big difference in both discrepancy.

A random sequence $x_1, x_2, \ldots \in I$ is called of low discrepancy if there exists a constant $C > 0$ such that
\[
D_N \leq C \left( \frac{\log N}{N} \right)^d.
\]
It is proved for $d = 1, 2$
\[
D_N \geq O \left( \frac{\log N}{N} \right)^d,
\]
and is expected the above inequality holds even for $d \geq 3$. Namely, the low discrepancy sequence will be the best uniformly distributed sequence. Thus this is the best sequence to approximate an integration numerically
\[
\frac{1}{N} \sum_{n=1}^{N} f(x_n) \sim \int_I f \, dx
\]
by quasi Monte Carlo method.

1.2 Symbolic Dynamics

Let $F$ be an expanding transformation $I$ into itself, that is, $\xi > 0$. We consider a finite set $A$ and a partition $\{\langle a \rangle\}_{a \in A}$ of $I$. In this article, we only consider the cases $I = [0, 1]^d$ and partitions $\{\langle a \rangle\}_{a \in A}$ into intervals.

For $x \in I$, we define a sequence $a_1^x a_2^x \cdots \in A^\mathbb{N}$ by
\[
F^{n-1}(x) \in \langle a_n^x \rangle.
\]
We call this sequence the expansion of $x$. We define a space $X \subset A^\mathbb{N}$ the closure of all the expansion of $x$. We denote a shift operator $\theta$ on $X$:
\[
\theta(a_1 a_2 a_3 \cdots) = a_2 a_3 \cdots.
\]
We call the dynamical system $(X, \theta)$ the symbolic dynamics associated with $(I, F)$.

We call a finite sequence $w = a_1 \cdots a_n$ a word and define
1. $|w| = n$ (the length of $w$). We define for the empty word $\epsilon$, $|\epsilon| = 0$.

2. $\langle w \rangle = \bigcap_{i=1}^{n} F^{-(i-1)}(\langle a_i \rangle)$. Here, we define $\langle \epsilon \rangle = I$.

3. $W$ is the set of words $w$ with $\langle w \rangle \neq \emptyset$.

4. For $x \in I$, and a word $w$, $wx$ is a point whose expansion equals $a_1 \cdots a_n a_1^2 a_2^2 \cdots$, if it exists.

### 1.3 van der Corput Sequence generated by Dynamical System

On $A$, we consider some order. We fix $x \in I$. On $W$, for words $w = a_1 \cdots a_n$ and $w' = b_1 \cdots b_m$, when both $wx$ and $w'x$ exists, we define $wx < w'x$ ($w, w' \in W$) if one of the following holds:

1. $n < m$,

2. $n = m$, and there exists $k$ such that $a_{k+1} \cdots a_n = b_{k+1} \cdots b_n$ and $a_k < b_k$.

Arrange the points $wx$ which exist in the above order, and denote it by $v_0 x, v_1 x, v_2 x, \ldots$ ($v_0 = \epsilon$). We call this sequence a van der Corput sequence.

**Example 1** We consider the simplest case $I = [0, 1]$. Take a transformation $F(x) = 2x \pmod{1}$ with $A = \{0, 1\}$, $\langle 0 \rangle = [0, \frac{1}{2})$ and $\langle 1 \rangle = [\frac{1}{2}, 1]$. Then $X = \{0, 1\}^\mathbb{N}$. Choose $x = \frac{1}{2}$, then our van der Corput sequence is

$$x, 0x, 1x, 00x, 10x, 01x, 11x, 000x, 100x, \ldots = \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \ldots$$

This is the original van der Corput sequence. Original construction is as follows. Consider a sequence of natural numbers $1, 2, 3 \ldots$, then express them into binary expression $1, 10, 11, 100, \ldots$, then reverse their order and get a sequence $1, 01, 11, 001, \ldots$, finally add ‘.0.’, and get $0.1, 0.01, 0.11, 0.001, \ldots$ This coincides with the above sequence, and one of the most famous sequence of low discrepancy.

In our discussion, for any $x \in [0, 1]$, we have proved the associated sequence is of low discrepancy.

### 1.4 Perron–Frobenius operator

We will consider the Perron–Frobenius operator on $L^1$ defined by

$$Pf(x) = \sum_{y \in I} f(y)|F'(y)|^{-1},$$

where for $d \geq 2$ we consider $F'$ as the Jacobian of $F$. It is well known that

1. $P$ is a contraction and positive operator.
2. If $F$ is expansive and transitive, then 1 is the simple eigenvalue and we can choose a nonnegative eigenfunction. Let us denote by $\rho \geq 0$ the eigenfunction associated with eigenvalue 1 which satisfies $\int \rho \, dx = 1$, and $\mu$ be a probability measure with its density $\rho$. Then $\mu$ is an invariant probability measure, and the associated dynamical system is ergodic.

3. If there exists no eigenvalue on the unit circle except 1, then the dynamical system is mixing. For any $f \in L^1$ and $g \in L^\infty$, the decay rate of correlation is defined by

$$\int f(x)g(F^n(x)) \, d\mu - \int f \, d\mu \times \int g \, d\mu = \int (P^n(f\rho) - \int f \, d\mu \cdot \rho)(x) \, g(x) \, dx.$$ 

As the projection of $f\rho$ to the eigenspace associated with eigenvalue 1 is $\int f \, d\mu \times \rho$, the discrepancy is determined by the second greatest eigenvalue of $P$ in modulus. However, we know that any $z$ ($|z| < 1$) is an eigenvalue. Thus on $L^1$, the decay rate of correlation has no meaning. Hence, we need to restrict the domain of $P$. Usually for $d = 1$, we restrict it to the functions with bounded variations, because all the eigenfunctions associated with eigenvalues on the unit circle are of bounded variation. But we will consider a bit wider class $\mathcal{B}$. On $[0, 1]$, we consider an alphabet $\{0, 1\}$ and define associated intervals $\langle 0 \rangle = [0, \frac{1}{2})$ and $\langle 1 \rangle = [\frac{1}{2}, 1]$. A word is a finite sequence of 0 and 1, and we denote all the set of words by $W_1$. Then we define $W_2 = \{(w, w') : w, w' \in W_1, |w_1|, |w_2|:\text{even}\}$, and for $(w, w') \in W_2$ we define $|w, w'| = |w|+|w'|$, and $\langle w, w' \rangle = \langle w \rangle \times \langle w' \rangle$, where $\langle w \rangle \subset [0, 1]$ is an interval associated with a word $w$. A point $x \in [0, 1]^2$ has a binary expansion of the form:

$$(i_1, j_1)(i_2, j_2)\cdots, \quad (i_k, j_k = 0, 1, k = 1, 2, \ldots).$$

If $x \in (w, w')$, then $w = i_1 \cdots i_{|w|}$ and $w' = j_1 \cdots j_{|w'|}$.

For $0 < r < 1$, we define

$$||f||_r = \inf \sum_{m=0}^{\infty} r^m \sum_{|w, w'| = m} |C_{(w, w')}| < \infty,$$

where infimum is taken over all decompositions of $f$ in the form

$$f = \sum_{(w, w') \in W_2} C_{(w, w')} 1_{\langle w, w' \rangle}.$$ 

We define

$$\mathcal{B} = \{f \in L^1 : \exists \tilde{f} \sim f \text{ such that } ||\tilde{f}||_r < \infty \text{ for all } 0 < r < 1\},$$

where $\tilde{f} \sim f$ means that $\tilde{f}$ is a version of $f$. Then $\mathcal{B}$ is a locally convex space with semi norms $|| \cdot ||_r$ ($0 < r < 1$). For 1-dimensional piecewise monotonic
cases, \( ||f||_r \) is less than or equal to the total variation of \( f \), thus \( \mathcal{B} \) contains all the functions with bounded variation. Hence all the indicator function of an intervals belong to \( \mathcal{B} \).

Now we restrict the domain of \( P \) to \( \mathcal{B} \), and denote it also by \( P \). Our aim is to determine the spectra of \( P \) and also all the indicator functions of intervals are contained in \( \mathcal{B} \).

# 2 Construction of a transformation on \([0,1]^2\)

Now we consider the prime field \( \mathbb{F}_2 \) of characteristic 2, and the irreducible polynomial \( \beta^2 + \beta + 1 = 0 \) over \( \mathbb{F}_2 \). We consider elements of the group generated by \( \beta \) as an operator on \( \{0,1\}^2 \):

\[
0(i,j) = (i,j), \\
1(i,j) = (i+1 \mod 2,j), \\
\beta(i,j) = (i,j+1 \mod 2), \\
(1+\beta)(i,j) = (i+1 \mod 2,j+1 \mod 2).
\]

Let \( \hat{\mathcal{A}} = \{0,1,\beta,1+\beta\} \) a group generated by \( \beta \), and for a sequence \( \alpha = \alpha_1\alpha_2\ldots \in \hat{\mathcal{A}}^N \), we define \( \alpha x \) for \( x \in [0,1]^2 \) whose binary expansion equals \((i_1,j_1)(i_2,j_2)\ldots\), the \( n \)-th coordinate of the binary expansion of \( \alpha x \) equals \( \alpha_n(i_n,j_n) \). We can identify a point \( x \in [0,1]^2 \) as an element of \( \hat{\mathcal{A}}^N \) (0,0) as \( 0 \), (1,0) as \( \beta \), and (1,1) as \( 1+\beta \), because for example \((0,1) = \beta(0,0) \).

Instead of defining a linear transformation \( \hat{F} \) on \( \mathcal{A}^N \), we will construct a linear operator \( \hat{U} \) on \( \hat{\mathcal{A}}^N \).

To determine \( \hat{U} \), we introduce an infinite dimensional matrix \( U \) of the form \( U = (u,0u,00u,0^3u,\ldots) \), where \( u \) is an infinite dimensional vector and the transpose of the vector \( 0^k u \) is given by

\[
t^u(u_1,u_2,\ldots), \quad t(0^k u) = (0,\ldots,0,\underbrace{u_1,u_2,\ldots}_k).
\]

Let \( U^{-1} \hat{F} U \) be a shift operator, that is,

\[
U^{-1} \hat{F} U = \begin{pmatrix}
0 & 1 & 0 & 0 & \cdots \\
0 & 0 & 1 & 0 & \cdots \\
0 & 0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots \\
\end{pmatrix}.
\]

We will determine the components of \( u \) inductively. To make the notations simple, even when we restrict \( U \) and \( \hat{F} \) and so on to finite dimensions, we use the same notations. We want to construct \( \hat{F} \) such that for any \( k \) \( \hat{F}^k \) maps \( \{0,1\}^{2k} \) (or \( \{0,\beta\}^{2k} \)) to \( \hat{\mathcal{A}}^k \) 1 to 1 and onto. Since the number of \( \{0,1\}^{2k} \), \( \{0,\beta\}^{2k} \) and \( \hat{\mathcal{A}}^k \) all equal \( \mathcal{A}^k \), we only need to show \( \hat{F}^k \) is 1 to 1. Thus we need
to show for any $k$, all the vectors which belong to the kernel of $\hat{F}^k$ contain both 1 and $\beta$ as their components. Note that

$$0^l u \xrightarrow{U^{-1}} e_l \xrightarrow{U^{-1} \hat{F}^k U} e_{l-k} \xrightarrow{U} 0^{l-k} u,$$

where for $l < k$, $0^{k-l} u$ is the zero vector. Thus the kernel of $\hat{F}^k$ is generated by $u, 0u, \ldots, 0^{k-1} u$. When we consider $\hat{F}^k$, we restrict the vector space to $2k$ dimension, we need to construct vector $u$ such that all the $2k$ dimensional vectors which belong to the restriction of the subspace generated by $u, 0u, \ldots, 0^{k-1} u$ contain both 1 and $\beta$. Note also for any vector $x$ $\hat{F}^k(0^{k} x) = x$.

When $k = 1$, we put

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \beta \end{pmatrix}.$$  

This is the kernel of $\hat{F}$ from $\hat{A}^2$ to $\hat{A}$, and this contains both 1 and $\beta$.

As a notation, we determine $v^1$ and $v^\beta$ for a vector $v$, where

- $v = v^0 + v^\beta$,
- any component of $v^1$ is either 0 or 1,
- any component of $v^\beta$ is either 0 or $\beta$.

Let $V_k$ be a set of all $2k$-dimensional vectors $v$ of the form

$$v = \sum_{i=1}^{k} a_i (0^{i-1} u),$$

and the first $2k - 2$ components of $v^\beta$ equal 0. We also denote the set of vectors $v \in V_k$ with $a_1 = 1$ by $\tilde{V}_k$. Note that for any $v \in V_k$, $v^1$ is not the zero vector.

We will construct $u$ for which

$$V_k = \left\{ v = \sum_{i=1}^{k} a_i (0^{i-1} u) : v^1 \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, v^\beta = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \beta \\ \beta \end{pmatrix} \right\},$$

and $\tilde{V}_k$ is the set of two vectors in $V_k$. We call a vector $v$ of type $(a, b)_k$ if

$$v^\beta = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a \\ b \end{pmatrix},$$

that is, the vectors in $V_k$ are of type $(\beta, 0)_k$, $(0, \beta)_k$ and $(\beta, \beta)_k$. 

(6)
Now we consider the case $k = 2$. Let \( \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 1 \\ \beta \\ 0 \\ \beta \end{pmatrix} \). Then
\[
\begin{pmatrix} 1 \\ \beta \\ 0 \\ \beta \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ \beta \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1+\beta \\ \beta \end{pmatrix}, \quad \begin{pmatrix} 1 \\ \beta \\ 0 \\ \beta \end{pmatrix} + (\beta + 1) \begin{pmatrix} 0 \\ 1 \\ \beta \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \beta \end{pmatrix},
\]
that is, the vectors which belong to $\tilde{V}_2$ are of type $(\beta, \beta)_2$ and $(0, \beta)_2$. Adding the vectors which belong to $\tilde{V}_2$, we get a vector \( \begin{pmatrix} 0 \\ 1 \\ \beta \\ 0 \end{pmatrix} \) of type $(\beta, 0)_2$. Note that this vector does not belong to $\tilde{V}_2$.

We will inductively construct $u$ such that there exist three types $(0, \beta)_k$, $(\beta, 0)_k$ and $(\beta, \beta)_k$ in $V_k$, and that the vector of type $(0, \beta)_k$ belongs to $\tilde{V}_k$. This assertion holds for $k = 2$, and assume that we have already determined $u_1, \ldots, u_{2k}$. Let us denote by $v, v'$ and $v''$ the vector of type $(0, \beta)_k$, the vector of type $(\beta, 0)_k$ and the vector of type $(\beta, \beta)_k$. Now determine $u_{2k+1}$ to let the $2k+1$ component of $v^{\beta}$ to be 0. Then $2k+1$ element of $v'$ and $v''$ are determined. We denote the $2k+1$ components of $(v')^{\beta}$ and $(v'')^{\beta}$ by $a$ and $a'$. Now determine $u_{2k+2}$ to be $u_{2k+2} + a = \beta$ or $1 + \beta$. Then we get $v + 0v'$ of type $(0, \beta)_{k+1}$, and $v + 0v''$ either of type $(\beta, 0)_{k+1}$ or $(\beta, \beta)_{k+1}$, and
\[
\tilde{V}_{k+1} = \{v + 0v', v + 0v''\}, \quad V_{k+1} = \tilde{V}_{k+1} \cup \{0v' + 0v''\}.
\]
We can denote this procedure as follows
\[
\begin{array}{c|c|c|c}
2k-1 & v^{\beta} & (0v')^{\beta} & (0v'')^{\beta} \\
2k & \beta & \beta & \beta \\
2k+1 & \beta & 0 & \beta \\
\end{array}
\Rightarrow
\begin{array}{c|c|c|c}
2k-1 & v^{\beta} & (0v')^{\beta} & (0v'')^{\beta} \\
2k & \beta & \beta & \beta \\
2k+1 & \beta & a & a' \\
2k+2 & 0 & 0 & 0 \\
\end{array}
\]

**Lemma 1** For even $k$, the restriction of $\tilde{F}^{k+n}$ maps from $0^k \{0,1\}^{2n}$ or from $0^k \{0,\beta\}^{2n}$ to $\tilde{A}^n$ 1 to 1 and onto.

**Proof.** The number of elements of both $0^k \{0,1\}^{2n}$ and $\tilde{A}^n$ equals $2^{2n} = 4^n$. Thus, we only need to prove the map is 1 to 1.

We can take $0^k u, \ldots, 0^{k+2n-1} u$ as a basis of $0^k \tilde{A}^{2n}$. Then since $\tilde{F}^m 0^l u = 0^{l-m} u$, the kernel of $\tilde{F}^{k+n}$ is generated by $0^k u, \ldots, 0^{k+n-1} u$, and its image is generated by $u, \ldots, 0^{n-1} u$. As we construct $u$ such that $k + 2n$ dimensional vectors $0^k u, \ldots, 0^{k+n-1} u$ can generate only vectors with neither $v^\beta$ nor $v^{\beta}$ equal 0. This completes the proof. \( \square \)
Now we will consider a word \( w \in \mathcal{W}_2 \). For words \( w = w_1 \times w_2 \) with length \( |w_1| + |w_2| = 2k \).

1. Assume now \( |w_1| = |w_2| = k \). Then this word can be expressed by \( A_1 \cdots A_k \langle (0,0) \rangle \cdots \langle (0,0) \rangle \) with some \( A_1, \ldots, A_k \in \hat{A} \). Then the \( k \)-dimensional vector which has the form \( A_1 \cdots A_k \) can be expressed as

\[
\sum_{i=1}^{k} a_i (0^{i-1}u),
\]

with some \( a_1, \ldots, a_k \in \hat{A} \). As \( \hat{F}^k(\sum_1^k a_i (0^{i-1}u)) \) is the empty word, we get

\[
F^k(\langle w \rangle) = I.
\]

2. Assume \( |w_1| = l < |w_2| \). Then,

\[
w = (A_1 \cdots A_l) \langle (0,0) \rangle \cdots \langle (0,0) \rangle \langle (0,0) \rangle \langle (0,0) \rangle \langle (1,0) \rangle 2(k-l)
\]

Thus from Lemma 1

\[
\hat{F}^k(\langle 0^l \{0,1\} 2(k-l) \rangle) = \hat{A}^{k-l},
\]

we get

\[
F^k(\langle w \rangle) = I.
\]

3. The case \( |w_1| > |w_2| \), we can prove \( F^k(\langle w \rangle) = I \) in a similar way.

Thus we get:

**Lemma 2** A word \( w \in \mathcal{W}_2 \) with \( |w| = n \), the interval \( \langle w \rangle \) belongs to \( F_n \).

3 Discrepancy

Now we will define a van der Corput sequence generated by the transformation \( F \), and calculate its discrepancy. Consider a partition

\[
(0,0) = [0, \frac{1}{2}] \times [0, \frac{1}{2}], \quad (0,1) = [0, \frac{1}{2}] \times [\frac{1}{2}, 1],
\]

\[
(1,0) = [\frac{1}{2}, 1] \times [0, \frac{1}{2}], \quad (1,1) = [\frac{1}{2}, 1] \times [\frac{1}{2}, 1],
\]

and define some order for example \( (0,0) < (1,0) < (0,1) < (1,1) \) on it. We can define van der Corput sequence by this partition of the form

\[
x, (0,0)x, (1,0)x, (0,1)x, (1,1)x, (0,0)(0,0)x, (1,0)(0,0)x,
\]

\[
(0,1)(0,0)x, (1,1)(0,0)x, (0,0)(1,0)x, \ldots, \quad (x \in I).
\]
Along the notations using 1, β, they are
\[ x, 0x, 1x, βx, (1 + β)x, 00x, 10x, β0x, (1 + β)0x, \\
01x, 11x, β1x, (1 + β)1x, 0βx, 1βx, ββx, \ldots. \]
We express it \( v_0x, v_1x, \ldots \).

We will show:

**Theorem 1** For any \( x \in I \), our van der Corput sequence is of low discrepancy.

We will calculate \( * \)-discrepancy. Let \( J \) be an interval of the form \([0, a] \times [0, b]\). We will divide this interval into subintervals associated with words of \( \mathcal{W}_2 \). Find a word \( u \in \mathcal{W}_2 \) for which \(|u| = k\) for some \( k \), and there exists no word \( u' \in \mathcal{W}_2 \) with \(|u'| < (k - 1)\) and \( \langle u' \rangle \subset J \). We fix one of them. There exists at most \((\#A - 1)^2 = 9\) such words with the same form. We denote the union of them by \( J_1 \). There exists at most two edges and one vertex of such words which belong inside of \( J \). For each edge, we can choose a word which has common vertex with \( J_1 \) with even length, and the interval associated with it contained in \( J \setminus J_1 \). There exist at most three of such copies. We denote these words of type 1. For an edge, we can find a word with even length and the interval associated with it contained in \( J \setminus J_1 \). There exists at most 9 of such copies. We denote this word of type 2. We denote their union with \( J_1 \) by \( J_2 \). We can continue this procedure. We will count the number of edges of \( J_i \) which belongs to the inside of \( I \). A word of type 1 generates one edge, and a word of type 2 generates one vertex and at most two edges. We define a matrix
\[
M = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.
\]
There exist at most \((2^2 - 1)^2 = 9\) words with same form for type 1, and at most \((2^2 - 1) = 3\) words with same form for type 2. Thus, the total number of words with length \( n \) which covers \( J \) equals
\[
(9, 3)M^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 6n + 9.
\]

**Lemma 3** For any interval \( J = [0, a] \times [0, b] \), the indicator function \( 1_J \in B \).

**Proof.** First note that a word \( u \in \mathcal{W}_2 \) with \(|u| = n\) belongs to \( \mathcal{F}_n \). As we showed above,
\[
1_J(x) = \sum_{n=1}^{\infty} \sum_{|u| = n} C_u 1_{\langle u \rangle}(x) \quad (C_u = 0 \text{ or } 1),
\]
and the number \( C_u = 1 \) for \(|u| = n\) is less than or equal to \( 6n + 9 \). Thus for any \( 0 < r < 1 \)
\[
\sum_{n=1}^{\infty} r^n \sum_{|u| = n} |C_u| \leq \sum_{n=1}^{\infty} r^n (6n + 9) < \infty.
\]
This shows \( 1_J \in \mathcal{B} \). □

On the other hand,

\[
P^n 1_J(x) = \sum_{y \colon F^n(y) = x} 1_J(y) 4^{-n}.
\]

We need to consider the spectra of \( P \) restricted to \( \mathcal{B} \). Now we consider a generating function of the form:

\[
s^J(z, x) = \sum_{n=0}^{\infty} z^n P^n 1_J(x).
\]

For \( a \in \mathcal{A} \), we get

\[
s^{(a)}(z, x) = \sum_{n=0}^{\infty} z^n P^n 1_{\langle a \rangle}(x)
\]

\[
= 1_{\langle a \rangle}(x) + \sum_{n=1}^{\infty} z^n P^n 1_{\langle a \rangle}(x)
\]

\[
= 1_{\langle a \rangle}(x) + \sum_{n=0}^{\infty} z^{n+1} \sum_{y \colon F(y) = x} P^n 1_{\langle a \rangle}(y). 4^{-1}
\]

\[
= 1_{\langle a \rangle}(x) + \frac{z}{4} \sum_{n=0}^{\infty} z^n P^n 1_I(x)
\]

\[
= 1_{\langle a \rangle}(x) + \sum_{b \in \mathcal{A}} \frac{z}{4} \sum_{n=0}^{\infty} z^n P^n 1_{\langle b \rangle}(x)
\]

\[
= 1_{\langle a \rangle}(x) + \frac{z}{4} \sum_{b \in \mathcal{A}} s^{(b)}(z, x).
\]

Thus taking

\[
s(z, x) = (s^{(a)}(z, x))_{a \in \mathcal{A}},
\]

we get

\[
s(z, x) = (I - \Phi(z))^{-1}(1_{\langle a \rangle}(z, x))_{a \in \mathcal{A}},
\]

where the Fredholm matrix \( \Phi(z) \) is defined by

\[
\Phi(z)_{a,b} = \frac{z}{4} \quad (a, b \in \mathcal{A}).
\]
Hence, $s(z, x)$ has singularity at $z = 1$:

$$s(z, x) = \frac{1}{4(1-z)} \begin{pmatrix} 4 - 3z & z & z \\ z & 4 - 3z & z \\ z & z & 4 - 3z \end{pmatrix} 1_{(a)}(z, x)_{a \in A}$$

$$= \frac{1}{4(1-z)} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} 1_{(a)}(z, x)_{a \in A}$$

Thus we get

$$s^I(z, x) = \sum_{a \in A} s^{(a)}(z, x) = (1, 1, 1) s(z, x) = \frac{1}{1-z}.$$ 

We get also for $u \in W_2 (|u| = n),$

$$s^{(u)}(z, x) = \sum_{k=0}^{\infty} z^k P^k 1_{(u)}(x)$$

$$= \sum_{k=0}^{n-1} \left( \frac{z}{4} \right)^k 1_{(u)}(x) + \sum_{k=n}^{\infty} \left( \frac{z}{4} \right)^k P^k 1_{(u)}(x)$$

$$= \sum_{k=0}^{n-1} \left( \frac{z}{4} \right)^k 1_{(u)}(x) + \left( \frac{z}{4} \right)^n s^I(z, x).$$

Thus we consider a generating function

$$\sum_{n=0}^{\infty} z^n \# \{ wx \in J : |w| = n \} = \sum_{n=0}^{\infty} (4z)^n 1_f(x) = s^I(4z, x)$$

$$= \sum_{n=0}^{\infty} \sum_{u \in W_2, |u| = n} C_u s^{(u)}(4z, x)$$

$$= \sum_{n=0}^{\infty} \sum_{u \in W_2, |u| = n} C_u \left( \sum_{k=0}^{n-1} z^k 1_{F^k(u)}(x) + z^n s^I(4z, x) \right)$$

$$= \sum_{n=0}^{\infty} \sum_{u \in W_2, |u| = n} C_u \sum_{k=0}^{n-1} z^k 1_{F^k(u)}(x) + \sum_{n=0}^{\infty} \sum_{u \in W_2, |u| = n} C_u z^n \frac{1}{1-4z} (1)$$

The second term of (1) is the main term:

$$\sum_{n=0}^{\infty} \sum_{u \in W_2, |u| = n} C_u z^n \frac{1}{1-4z} = \sum_{n=0}^{\infty} \sum_{u \in W_2, |u| = n} C_u \frac{4^n}{1-4z} + \sum_{n=0}^{\infty} \sum_{u \in W_2, |u| = n} C_u \frac{z^n - 4^n}{1-4z}$$

$$= \frac{|J|}{1-4z} - \sum_{n=0}^{\infty} \sum_{u \in W_2, |u| = n} C_u 4^n \sum_{k=0}^{n-1} (4z)^k.$$ (2)
The latter term of the right hand term of (2):

\[
\left| \sum_{n=0}^{\infty} \sum_{u \in W_2 \mid |u| = n} C_u 4^{-n} \sum_{k=0}^{n-1} (4z)^k \right| \leq \sum_{k=0}^{\infty} |4z|^k \sum_{n=k+1}^{\infty} \sum_{|u|=n} |C_u| 4^{-n}
\]

\[
\leq \sum_{k=0}^{\infty} |4z|^k \sum_{n=k+1}^{\infty} (6n + 9) 4^{-n}
\]

\[
\leq C \sum_{k=0}^{\infty} k |z|^k
\]

with some constant \(C > 0\).

Now we consider the first part of the right hand of (1). First note that there exists only one \(u\) such that \(C_u = 1\) and \(x \in \langle u \rangle\). Also for \(k > 0\), there exists one word of type 1 with length \(k\), and at most \(2k\) words of type 2 with length \(k\) which are mutually disjoint and all the words with length longer than \(k\) are contained in them. So there exists at most \((1, 1) M^k \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 + 2k\) words \(u\) such that \(|u| > k\), \(C_u = 1\) and \(F^k(\langle u \rangle) \ni x\). Thus

\[
\left| \sum_{n=0}^{\infty} \sum_{|u|=n} C_u \sum_{k=0}^{n-1} z^k 1_{F^k(\langle u \rangle)}(x) \right| \leq \sum_{k=0}^{\infty} z^k \sum_{n=k}^{\infty} \sum_{|u|=n} |C_u| 1_{F^k(\langle u \rangle)}(x)
\]

\[
\leq \sum_{k=0}^{\infty} z^k (1 + 2k).
\]

Thus there exists some constant \(C'\) such that

\[
|\#\{wx \in J : |w| = k\} - 4^k |J|| \leq C' k.
\]

Now we consider \(N\) with \(4^k \leq N < 4^{k+1}\), and consider the number of \(w_i x \in J\) until \(i \leq N\). Until \(i \leq 4^k\), we have already calculated and

\[
\#\{v_i x \in J : i \leq 4^k\} = \sum_{n=0}^{k} \#\{w x \in J : |w| = n\}.
\]

\[
= \frac{4^{k+1} - 1}{4 - 1} |J| + O(\sum_{n=0}^{k} n)
\]

\[
= \#\{|w| \leq k\} \times |J| + O(k^2).
\]

As we remarked before, after the \(4^k + 1\) the van der Corput sequence has the form \(w(0, 0)x, w(1, 0)x, w(1, 0)x, w(0, 1)(0, 0)x\) and \(w(1, 1)x\) with \(|w| = k\), and in the words \(w(0, 0)x\) \((|w| = k)\), the order of words are \(w'(0, 0)(0, 0)x, w'(1, 0)x, w'(0, 1)(0, 0)x\) and \(w'(1, 1)(0, 0)x\) with \(|w'| = k - 1\) and so on. Thus we get for
\(4^k < N < 4^{k+1}\)

\[
\#\{v_i x \in J : i \leq N\} = \sum_{n=0}^{k} \#\{w x \in J : |w| = n\} \\
+ \sum_{n=0}^{k-1} \sum_{|u| = k-n} \#\{w u x \in J : |w| = n, w u \leq v_i\},
\]

where we can choose the number of \(|u| = n - k\) in the above formula is at most \(#A - 1 = 3\). Thus we get

\[
\#\{v_i x \in J : i \leq N\} = N|J| + O(k^2).
\]

Since \(k = O(\log N)\), we get the discrepancy

\[
\left| \frac{1}{N} \#\{v_i x \in J : i \leq N\} - |J| \right| \leq O \left( \frac{(\log N)^2}{N} \right).
\]

This proves the theorem.

References


