Complexities of Erez self-dual normal bases

Blondeau Da Silva Stéphane

August 2, 2017

Abstract

The complexities of self-dual normal bases, which are candidates for the lowest complexity basis of some defined extensions, are determined with the help of the number of all but the simple points in well chosen minimal Besicovitch arrangements. In this article, these values are first compared with the expected value of the number of all but the simple points in a minimal randomly selected Besicovitch arrangement in $\mathbb{F}_{d^2}$ for the first 370 prime numbers $d$. Then, particular minimal Besicovitch arrangements which share several geometrical properties with the arrangements considered to determine the complexity will be considered in two distinct cases.

Introduction

Let $q$ be a prime power, $\mathbb{F}_q$ be the field of $q$ elements and $n$ be a positive integer. We consider the Galois group of the extension $\mathbb{F}_{q^n}/\mathbb{F}_q$, which is a cyclic group generated by the Frobenius automorphism $\Phi : x \mapsto x^q$. There exists an $\alpha$ that generates a "normal" basis for $\mathbb{F}_{q^n}/\mathbb{F}_q$, i.e. a basis consisting of the orbit $(\alpha, \alpha^q, ..., \alpha^{q^{n-1}})$ of $\alpha$ under the action of the Frobenius. The difficulty of multiplying two elements of the extension expressed in this basis is measured by the complexity of $\alpha$, namely the number of non-zero entries in the multiplication-by-$\alpha$ matrix: $(\text{Tr}(\alpha^i \alpha^j))_{0 \leq i,j \leq n-1}$, where $\text{Tr}$ is the trace map from $\mathbb{F}_{q^n}$ to $\mathbb{F}_q$ \cite{4.1}. As a large number of zero in this matrix enables faster calculations, finding normal bases with low complexity is a significant issue.

Self-dual normal bases are particular normal bases which verify $\text{Tr}(\alpha^i \alpha^j) = \delta_{i,j}$ (for $0 \leq i, j \leq n - 1$), where $\delta$ is the Kronecker delta. Arnault et al. in \cite{1} have identified the lowest complexity of self-dual normal bases for extensions of low degree and have showed that the best complexity of normal bases is often achieved from a self-dual normal basis. In \cite{7}, Pickett and Vinatier considered cyclotomic extensions of the rationals generated by $d^2$-th roots of unity, where $d$ is a prime. The construction they use yields a candidate for the lowest complexity basis for $\mathbb{F}_{p^d}/\mathbb{F}_p$, where $p \neq d$ is a prime which does not split in the chosen extension. They prove that the multiplication table of this basis can be geometrically interpreted by means of an appropriate minimal Besicovitch arrangement. The complexity of the basis, denoted by $C_d$, is here equal to the number of all but the simple points generated by this arrangement in $\mathbb{F}_{d^2}$.

After a brief overview of the properties this arrangement have, we will compare the complexity $C_d$ with the expected value of the number of all but the simple points in a minimal randomly selected Besicovitch arrangement in $\mathbb{F}_{d^2}$.
for the first 370 prime numbers \( d \). The expectations will be determined using Blondeau Da Silva’s results in [2]. In a third part, we will consider particular minimal Besicovitch arrangements which share several geometrical properties with the arrangements considered to determine the complexity. We will again compare in this part, for the first 370 prime numbers \( d \), \( C_d \) with the expected value of the number of all but the simple points in the randomly selected mentioned above arrangement.

1 The minimal Besicovitch arrangement providing the complexity

Let \( d \) be a prime number and \( \mathbb{F}_d \) be the \( d \)-elements finite field.

A line, in \( \mathbb{F}_d^2 \), is a one-dimensional affine subspace. A Besicovitch arrangement \( B \) is a set of lines that contains at least one line in each direction. A minimal Besicovitch arrangement is a Besicovitch arrangement that is the union of exactly \( d + 1 \) lines in \( \mathbb{F}_d^2 \) (see [2]).

The minimal Besicovitch arrangement considered, brought out by Pickett and Vinatier ([7]), and denoted by \( \mathcal{L} \), is composed of \( d + 1 \) lines with the following equations:

\[
\begin{align*}
  L_a : & \quad ax - (a + 1)y - p(a) = 0 \quad \text{for } a \in \mathbb{F}_d, \\
  L_\infty : & \quad x - y = 0,
\end{align*}
\]

where \( p \) is the following polynomial:

\[
\forall x \in \mathbb{F}_d, \quad p(x) = \frac{(x + 1)^d - x^d - 1}{d}.
\]

For \( d \geq 5 \), Pickett and Vinatier ([7]) have proved that under the action of \( \Gamma = \langle \iota, \theta \rangle \) (a group generated by two elements of \( GL_2(\mathbb{F}_d) \), where \( \iota(x, y) = (y, x) \) and \( \theta(x, y) = (y - x, -x) \) for \( (x, y) \in \mathbb{F}_d^2 \)), this arrangement \( \mathcal{L} \) always has two orbits of cardinal 3: \( \{L_0, L_{-1}, L_\infty\} \) and \( \{L_1, L_{\omega^{-1}}, L_{-2}\} \). They have also stated that:

- if \( d \equiv 1 \mod 3 \), there are one orbit of cardinal 2, \( \{L_\omega, L_{\omega^2}\} \), where \( \omega \) is a primitive cubic root of unity in \( \mathbb{F}_d \) and \( \frac{d-1}{6} \) orbits of cardinal 6;

- if \( d \equiv 2 \mod 3 \), there are \( \frac{d-3}{6} \) orbits of cardinal 6.

The Comp_lib 1.1 package have been implemented in Python 3.4. It provides the complexity \( C_d \) of the basis (by counting all but the simple points in the associated minimal Besicovitch arrangement) and it also enables to determine the points multiplicities distribution in \( \mathbb{F}_d^2 \) of this arrangement. It is available at https://pypi.python.org/pypi/Comp_lib/1.1. Table I in Appendix gathers the first 370 values of \( C_d \).

2 Complexity versus number of all but simple points in randomly selected arrangements

Let us denote by \( A_d \) the expected value of the number of all but the simple points in a randomly chosen minimal Besicovitch arrangement in \( \mathbb{F}_d^2 \). Thanks
to the proof of Theorem 1. in [2], we have:

\[ A_d = d^2 - d(d + 1)(1 - \frac{1}{d})^d \]

\[ = (1 - \frac{1}{e})d^2 - \frac{1}{2e}d + O(1), \quad \text{as } d \to \infty. \]

Figure 1 shows the values of \( \frac{C_d - A_d}{d} \) for the first 370 prime numbers.

2.1 A first test

From the 370 values of Figure 1, we plot the regression line: its slope \( s \) is approximately \( 4.94 \times 10^{-5} \) and its intercept is approximately \( -0.913 \).

Let us consider the following null hypothesis \( H_0: s = 0 \). We have to calculate \( T = \frac{s - 0}{\hat{\sigma}_s} \), where \( \hat{\sigma}_s \) is the estimated standard deviation of the slope. We obtain \( \hat{\sigma}_s \approx 8.74 \times 10^{-5} \) and \( T \approx 0.565 \). \( T \) follows a student’s t-distribution with \((370 - 2)\) degrees of freedom (see [3, Proposition 1.8]). The acceptance region of the hypothesis test with a 5% risk is approximately \([-1.967, 1.967]\). Thus it can be concluded that we cannot reject the null hypothesis: the fact that the slope is not significantly different from zero can not be rejected.

2.2 A second test

Figure 2 below shows the distribution of the values of \( \frac{C_d - A_d}{d} \) for the first 370 prime numbers. In regard to the resulting histogram, one may wonder whether these values are normally distributed or not.
From the result of the first test, we would consider in this part that the function that maps $d$ onto $\frac{C_d - A_d}{d}$ behaves like a random variable with an expected value $\Lambda$ close to $-0.856$. On that assumption we will verify whether the values of $\frac{C_d - A_d}{d}$ are normally distributed for $d \in [2, 2531] \cap N$ (the null hypothesis) or not. For this purpose we use the Shapiro–Wilk test (see [8]). The test statistic $W$ is about 0.991. The associated p-value being about 0.0296, it can be concluded that we can reject the null hypothesis, i.e. the values of $\frac{C_d - A_d}{d}$ are significantly not normally distributed for $d \in [2, 2531] \cap N$.

### 2.3 A third set of tests

Once more, from the result of the first test, we would consider in this part that the function that maps $d$ onto $\frac{C_d - A_d}{d}$ behaves like a random variable with an expected value $\Lambda$ close to $-0.856$ and with a symmetric probability distribution. On that assumption we will verify whether the values higher than $\Lambda$ and those smaller than $\Lambda$ are randomly scattered over the ordered absolute values of $\frac{C_d - A_d}{d}$ (the null hypothesis) or not. To this end we use a non-parametric test, the Mann–Whitney $U$ test: we determine the ranks of $|\frac{C_d - A_d}{d}|$ for each $d$ in the considered interval (see [9] or [5]). The ranks sum of the values higher than $\Lambda$ is approximately normally distributed. The value of $U_1$ is about $-0.911$. The acceptance region of the hypothesis test with a 5% risk being approximately $[-0.960, 1.960]$, it can be concluded that we cannot reject the null hypothesis, i.e. the fact that the greater and smaller than $\Lambda$ values of $\frac{C_d - A_d}{d}$ for $d \in [2, 2531] \cap N$ are randomly scattered: the symmetry of the probability distribution of our potential pseudorandom variable can not be rejected.

Once more, on our first assumption, we will verify whether the values higher than $\Lambda$ and those smaller than $\Lambda$ are randomly scattered over the first 370 prime numbers (the null hypothesis) or not. To this end we use the same test, the Mann–Whitney $U$ test. The prime number ranks sum of the values higher than $\Lambda$ is approximately normally distributed. The value of $U_2$ is about $-0.397$. It can be concluded that we cannot reject the null hypothesis, i.e. the fact that the
greater and smaller than $\Lambda$ values of $C_d - A_d$ for $d \in [2, 2531] \cap \mathbb{N}$ are randomly scattered over the first 370 prime numbers.

2.4 Perspective

Both first test and set of tests could not invalidate the fact that the function that maps $d$ onto $C_d - A_d$ seem to behave like a random variable with $\Lambda$ as expected value. If we succeed in proving such a statement, we could consider the following unbiased estimator of $C_d$, denoted by $\hat{C}_d$:

$$\hat{C}_d = A_d + \Lambda d$$

$$= d^2 - d(d + 1)(1 - \frac{1}{d})^d + \Lambda d$$

$$= (1 - \frac{1}{e})d^2 + (\Lambda - \frac{1}{2e})d + o(d), \quad \text{as } d \to \infty,$$

thanks to the proof of [2, Theorem 1].

3 Complexity versus number of all but simple points in particular arrangements

3.1 Further details on the minimal Besicovitch arrangement providing the complexity

In this part, we will consider particular minimal Besicovitch arrangements which share several geometrical properties with the arrangements considered to determine the complexity and we will compare the expected values of the number of all but simple points in such randomly selected arrangements with $C_d$ values.

Before reviewing the whole cycles highlighted in section 1, let us make a quick remark:

Remark 3.1. If a line in an orbit passes through $(0, 0) \in \mathbb{F}_d^2$ all the other lines of this orbit also pass through this point, the elements of the group $\Gamma$ acting on the lines being in $GL_2(\mathbb{F}_d)$.

In section 1 two cases appear, for $d \geq 5$: the cases where $d \equiv 1 \mod 3$ and those where $d \equiv 2 \mod 3$.

In both cases, the intercepts of the lines in $\{L_0, L_{-1}, L_{\infty}\}$ are 0 (we have $p(0) = 0$ thanks to equality 1, Remark 3.1 allowing us to conclude).

The intercepts of the lines in $\{L_1, L_{-2}, L_{-1}\}$ are non zero values, except for $d = 1093$, the first Wieferich prime number, for which lines intercepts are all zero: $p(2) = 0 \iff \frac{2^{d-1}-1}{d}$ (see equality 1, Remark 3.1 and [4]).

If $d \equiv 1 \mod 3$, the intercepts of the lines in $\{L_\omega, L_{\omega^2}\}$ are 0:

$$p(\omega) = \frac{(\omega + 1)^d - \omega^d - 1}{d} = \frac{-(\omega^d)^2 - \omega^d - 1}{d}$$

$$= -\frac{(\omega)^2 - \omega - 1}{d} = 0,$$

using the fact that $\omega$ is a primitive cubic root of unity in $\mathbb{F}_d$ and using Fermat’s little theorem.
In this part, we will only consider the values of \( d \in [2, 2531] \cap \mathbb{N} \) where all lines in the 6-cycles do not pass through \((0, 0)\); for the 152 values of \( d \) verifying this constraint and also \( d \equiv 1 \mod 3 \), we denote by \( M_d^* \) the expected value of the number of all but the simple points in a randomly chosen arrangement sharing geometrical properties with the arrangement providing the complexity; for the 153 values of \( d \) verifying the same constraint and also \( d \equiv 2 \mod 3 \), we denote by \( M_d^{**} \) the similar expected value. Table 1 shows the values of \( d \) being in either the first or the second case.

### 3.2 Lines intersections of the different cycles

The five functions in \( \Gamma \), other than the identity function \( \text{Id} \), will be denote as in [7]:

\[
\forall (x, y) \in \mathbb{F}_d^2, \\
\iota(x, y) = (y, x) \quad \theta(x, y) = (y - x, -x) \quad \theta^2(x, y) = (-y, x - y) \\
\kappa(x, y) = \theta \circ \iota(x, y) = (x - y, -y) \quad \lambda(x, y) = \iota \circ \theta(x, y) = (-x, y - x).
\]

Note that \( \iota, \kappa \) and \( \lambda \) are of order 2, and \( \theta \) and \( \theta^2 \) are of order 3. We can also easily verify that the fixed points of \( \iota \) are those of the line \( L_\infty \), the fixed points of \( \kappa \) are those of the line \( L_0 \) and the fixed points of \( \lambda \) are those of the line \( L_{-1} \). The following proposition can thus be enunciated:

**Proposition 3.2.** \( \forall \gamma \in \{\iota, \kappa, \lambda\} \) and \( \forall a \in \mathbb{F}_d \setminus \{0, -1\} \), if \( L_a \) and \( \gamma(L_a) \) are two distinct lines, then their intersection point is in line of the fixed points of \( \gamma \).

**Proof.** The image of a point under a function in \( \Gamma \subset \text{GL}_2(\mathbb{F}_d) \) is a point. So, \( \forall \gamma \in \{\iota, \kappa, \lambda\} \) and \( \forall a \in \mathbb{F}_d \setminus \{0, -1\} \), if \( L_a \) and \( \gamma(L_a) \) are two distinct lines, i.e. if their intersection is a point:

\[
\gamma(L_a \cap \gamma(L_a)) = \gamma(L_a) \cap \gamma(\gamma(L_a)) = L_a \cap \gamma(L_a),
\]

each of the considered functions being of order 2. The point \( L_a \cap \gamma(L_a) \) is thus in the fixed line of \( \gamma \). \( \square \)

Let us henceforth denote by \( \mathcal{S} \) the set \( \mathbb{F}_d^2 \setminus \{L_0, L_{-1}, L_\infty\} \). In each 6-cycle, for all \( \gamma \in \Gamma \) and for all \( a \in \mathbb{F}_d \) (such that \( L_a \) is in the considered 6-cycle), \( L_a \) and \( \gamma(L_a) \) are distinct; we can therefore apply Proposition 3.2 in the case where all the lines in a 6-cycle do not pass through \((0, 0)\) (the prevalent selected case in subsection 3.1), there exist 3 intersection points of the 6-cycle lines on each line of \( \{L_0, L_{-1}, L_\infty\} \):

- on \( L_0 \): \( L_a \cap \kappa(L_a), \theta(L_a) \cap \lambda(L_a) \) and \( \theta^2(L_a) \cap \iota(L_a) \);
- on \( L_{-1} \): \( L_a \cap \lambda(L_a), \theta(L_a) \cap \iota(L_a) \) and \( \theta^2(L_a) \cap \kappa(L_a) \);
- on \( L_\infty \): \( L_a \cap \iota(L_a), \theta(L_a) \cap \kappa(L_a) \) and \( \theta^2(L_a) \cap \lambda(L_a) \).

An other proposition can be added:

**Proposition 3.3.** In the case where all the lines in a 6-cycle do not pass through \((0, 0)\), two of the described above 6-cycle intersection points on a line of \( \{L_0, L_{-1}, L_\infty\} \) do not coincide.
Proof. Let us consider a 6-cycle. Its lines do not pass through the origin. \(a \in \mathbb{F}_d\), such that \(L_a\) is in this 6-cycle. We assume that \(L_a \cap \kappa(L_a)\) and \(\theta(L_a) \cap \lambda(L_a)\) coincide on \(L_0\). Knowing that \(\theta(L_0) = L_\infty\) (see [7]), we have:

\[
\theta(L_a \cap \kappa(L_a) \cap \theta(L_a) \cap \lambda(L_a)) \in L_\infty
\]

\[
\theta(L_a) \cap \lambda(L_a) \cap \theta^2(L_a) \cap \iota(L_a) \in L_\infty.
\]

So \(\theta(L_a) \cap \lambda \in L_0 \cap L_\infty = (0, 0)\); it contradicts the hypothesis of the proposition. The considered points do not coincide. All the other cases can be demonstrated in the same way. \(\square\)

Thus the remaining 6 intersection points of the 6-cycle lines are in \(\mathcal{T}\). We can finally prove the following proposition (in the case where \(d \geq 11\), otherwise there is no 6-cycle in the arrangement \(\mathcal{L}\)):

**Proposition 3.4.** The 6 remaining point in \(\mathcal{T}\) (in the case where all the lines in the 6-cycle do not pass through \((0, 0)\)) are distinct.

**Proof.** We first prove the following lemma:

**Lemma 3.5.** \(\theta\) has a single fixed point in \(\mathbb{F}_d^2\) \(\iff\) \(d \neq 3\).

**Proof.** \(\theta \in GL_2(\mathbb{F}_d)\) then \((0, 0)\) is a fixed point of \(\theta\).

For \((x, y) \in \mathbb{F}_d^2\):

\[
\theta(x, y) = (x, y) \iff y - x = x \text{ and } -x = y
\]

\[
\iff 3x = 0 \text{ and } y = -x.
\]

The result follows. \(\square\)

Let us consider a 6-cycle. Its lines do not pass through the origin. \(a \in \mathbb{F}_d\), such that \(L_a\) is in this 6-cycle.

Let us assume that 3 lines in the 6-cycle are concurrent in \(P \in \mathcal{T}\). It is clear from the foregoing that these lines are whether \(L_a\), \(\theta(L_a)\) and \(\theta^2(L_a)\) or \(\iota(L_a)\), \(\kappa(L_a)\) and \(\lambda(L_a)\). We have:

\[
\theta(P) = \theta(L_a \cap \theta(L_a) \cap \theta^2(L_a)) \quad \text{or} \quad \theta(\iota(L_a) \cap \kappa(L_a) \cap \lambda(L_a))
\]

\[
= \theta(L_a) \cap \theta^2(L_a) \cap L_a \quad \text{or} \quad \kappa(L_a) \cap \lambda(L_a) \cap \iota(L_a).
\]

In both cases \(\theta(P) = P\) i.e. \(P\) is a fixed point of \(\theta\). It means that \(P = (0, 0)\) thanks to Lemma 3.5 knowing that \(d \geq 11\); it contradicts the hypothesis of the proposition. The 6 remaining point in \(\mathcal{T}\) are distinct. \(\square\)

The cases of \(\{L_1, L_{ω−1}, L_{−2}\}\) and \(\{L_ω, L_{ω+1}\}\) can be considered as degenerate cases of a 6-cycle. Let us focus on the first arrangement. From [7], we get \(\iota(L_1) = L_{−2}, \lambda(L_{−2}) = L_{ω+1}\) and \(\kappa(L_1) = L_{ω−1}\). Thanks to Proposition 3.2, the 3 intersection points of lines in \(\{L_1, L_{ω−1}, L_{−2}\}\) are:

- on \(L_0\): \(L_1 \cap L_{ω−1}\);
- on \(L_{−1}\): \(L_{−2} \cap L_{ω−1}\);
- on \(L_\infty\): \(L_1 \cap L_{−2}\).
We note that this result is just a particular case of the above result.

The Figure 3 below provides two examples of minimal Besicovitch arrangements leading to the determination of the complexity. For the first one \((d = 7)\), we are in the case where \(d \equiv 1 \mod 3\), for the second one \((d = 11)\) in the case where \(d \equiv 2 \mod 3\). The above results and in particular those of Propositions 3.2, 3.3 and 3.4 are emphasised.

![Figure 3: Lines of the minimal Besicovitch arrangement in \(F_d^2\) providing the complexity \(C_d\) where \(d = 7\) (on the left) and \(d = 11\) (on the right). Red lines are those of \(\{L_0, L_{-1}, L_0\}\), green ones are those of \(\{L_1, L_{d-2}, L_{-2}\}\), blue ones are those of \(\{L_0, L_{\omega^2}\}\) and black ones are lines of a 6-cycle. The number of all but the simple points is 25 for \(d = 7\), and 67 for \(d = 11\); thus \(C_7 = 25\) and \(C_{11} = 67\).](image)

### 3.3 The first model

We first consider the case where \(d \equiv 1 \mod 3\). Let us denote by \(\Omega^*\) the set of minimal Besicovitch arrangements verifying some geometrical constraints similar to those of the considered Besicovitch arrangements. In such arrangements:

- there exist 3 lines of equations \(x = 0, y = 0\) and \(y = x\) (let us denote by \(l_a\) this lines set);
- there exist 2 lines that pass through the origin (let us denote by \(l_2\) this lines set);
- there exist 3 lines that do not pass through the origin, their 3 intersection points being respectively on each of the 3 lines in \(l_a\) (let us denote by \(l_3\) this lines set);
- there exist \(\frac{d-7}{6}\) sets of 6 lines, all verifying the same constraints as in Propositions 3.2, 3.3 and 5.4.

In order to calculate the average number of all but simple points in such arrangements, we will build a probability space: \(\Omega^*\). The \(\sigma\)-algebra chosen here
is the finite collection of all subsets of $\Omega^*$. Our probability measure, denoted by $P$, assigns equal probabilities to all outcomes.

For $Q \in \mathbb{F}_d^2$, let $M_Q$ be the random variable that maps $A \in \Omega^*$ to the multiplicity of $Q$ in $A$.

With the aim of knowing the expected number of simple points in such particular arrangements, we determine $P(M_Q = 1)$, for all $Q \in \mathbb{F}_d^2$. Two cases appear: either $Q$ is in a line of $l_a$ (apart from the origin) or not.

3.3.1 $Q$ is in a line of $l_a$ (apart from the origin)

In this case, for $A \in \Omega^*$, we have:

$M_Q(A) = 1 \iff$ none of the $d-2$ lines of $A$ (other than those of $l_a$) pass through $Q$.

We already know that lines of $l_2$ do not pass through this point.

There is a $\frac{d-2}{d-1} \times \frac{d-2}{d-1}$ probability that the two distinct intersection points between lines of $l_3$ and the considered line of $l_a$ do not coincide with $Q$ (a line is composed of $d$ points and the origin is here not considered).

Similarly, there is a $\frac{d-2}{d-1} \times \frac{d-3}{d-2} \times \frac{d-3}{d-2}$ probability that the three distinct intersection points between lines of a set of 6 lines (verifying the same constraints as in Propositions 3.2 and 3.3) and the considered line of $l_a$ do not coincide with $Q$.

Finally, considering the $\frac{d-7}{6}$ sets of 6 lines and the lines in $l_2$ and $l_3$, we obtain in this case:

$P(M_Q = 1) = \frac{d-3}{d-1} \times \left( \frac{d-4}{d-1} \right)^{\frac{d-7}{6}}$.

3.3.2 $Q$ is not in a line of $l_a$

In this case, for $A \in \Omega^*$, we have:

$M_Q(A) = 1 \iff$ exactly one line of the $d-2$ lines of $A$ (other than those of $l_a$) passes through $Q$.

We will use the following results to study in more detail the different subcases. In $\mathbb{F}_d^2 \setminus l_a$, there are $d^2 - 3 \times (d-1) - 1 = d^2 - 3d + 2$ points. In $\mathbb{F}_d^2 \setminus (l_a \cup l_2)$, there are $2(d-1)$ points of multiplicity 1 and the remaining points of multiplicity 0 ($d^2 - 5d + 4$ points). In $\mathbb{F}_d^2 \setminus (l_a \cup l_3)$, there are $3(d-3)$ points of multiplicity 1 and the remaining points of multiplicity 0 ($d^2 - 6d + 11$ points). In the union of $\mathbb{F}_d^2 \setminus l_a$ and a 6 lines set, there are $6(d-5)$ points of multiplicity 1, 6 points of multiplicity 2 (see Proposition 3.4) and the remaining points of multiplicity 0 ($d^2 - 9d + 26$ points).

This case can be divided in 3 subcases:

- the first one where the line that passes through $Q$ is in $l_2$; then the probability is:

$$\frac{2(d-1)}{d^2 - 3d + 2} \times \frac{d^2 - 6d + 11}{d^2 - 3d + 2} \times \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d-7}{6}};$$
the second one where the line that passes through $Q$ is in $l_3$; then the probability is:

\[
\frac{d^2 - 5d + 4}{d^2 - 3d + 2} \times \frac{3(d - 3)}{d^2 - 3d + 2} \times \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d - 7}{6}}
\]

the third one where the line that passes through $Q$ is in one of the $\frac{d - 7}{6}$ sets of 6 lines; then the probability is:

\[
\frac{d^2 - 5d + 4}{d^2 - 3d + 2} \times \frac{d^2 - 6d + 11}{d^2 - 3d + 2} \times \frac{d - 7}{6} \times \frac{6(d - 5)}{d^2 - 3d + 2} \times \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d - 13}{6}}
\]

Hence we have in this specific case:

\[
P(M_Q = 1) = \frac{2}{d - 2} \times \frac{d^2 - 6d + 11}{d^2 - 3d + 2} \times \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d - 7}{6}}
\]

\[
+ \frac{d - 4}{d - 2} \times \frac{3(d - 3)}{d^2 - 3d + 2} \times \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d - 7}{6}}
\]

\[
+ \frac{d - 4}{d - 2} \times \frac{d^2 - 6d + 11}{d^2 - 3d + 2} \times \frac{d - 7}{6} \times \frac{6(d - 5)}{d^2 - 3d + 2} \times \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d - 13}{6}}
\]

\[\text{3.3.3 The expected value of } M_d^*
\]

Recall that our aim is to determine the expected value $M_d^*$ of the number of all but simple points in arrangements of $\Omega^*$ in order to compare it with the value of the complexity $C_d$.

Thanks to the results of the above section and knowing that the first case concerns $3d - 3$ points and the second one $d^2 - 3d + 2$ points, we get:

\[
M_d^* \equiv \left( \frac{d^2 - 6d + 11}{d^2 - 3d + 2} \right) \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d - 7}{6}}
\]

\[\text{Using the Computer Algebra System Giac/Xcas (Parisse and De Graeve, 2017, } \text{http://www-fourier.ujf-grenoble.fr/~parisse/giac_fr.html}, \text{ version 1.2.3), we obtain:}
\]

\[
M_d^* = (1 - \frac{1}{e})d^2 + \left( \frac{1}{e} - 3 \exp(-\frac{1}{2}) \right)d + O(1), \text{ as } d \to \infty.
\]

\[\text{3.4 The second model}
\]

We henceforth consider the case where $d \equiv 2 \mod 3$. Let us denote by $\Omega^{**}$ the set of minimal Besicovitch arrangements verifying some geometrical constraints similar to those of the considered Besicovitch arrangements. In such arrangements:

- there exist 3 lines of equations $x = 0$, $y = 0$ and $y = x$ (let us denote by $l_a$ this lines set);
there exist 3 lines that do not pass through the origin, their 3 intersection points being respectively on each of the 3 lines in $l_a$ (let us denote by $l_3$ this lines set);

there exist $\frac{d-5}{6}$ sets of 6 lines, all verifying the same constraints as in Propositions 3.3 and 3.4.

In order to calculate the average number of all but simple points in such arrangements, we will build a probability space: $\Omega^{**}$. The $\sigma$-algebra chosen here is the finite collection of all subsets of $\Omega^{**}$. Our probability measure, denoted by $P$, assigns equal probabilities to all outcomes.

For $Q$ in $\mathbb{F}_d^2$, let $M_Q$ be the random variable that maps $A \in \Omega^{**}$ to the multiplicity of $Q$ in $A$.

With the aim of knowing the expected number of simple points in such particular arrangements, we determine $P(M_Q = 1)$, for all $Q$ in $\mathbb{F}_d^2$. Two cases appear: either $Q$ is in a line of $l_a$ (apart from the origin) or not.

3.4.1 $Q$ is in a line of $l_a$ (apart from the origin)

In this case, for $A \in \Omega^{**}$, we have:

$$M_Q(A) = 1 \iff \text{none of the } d-2 \text{ lines of } A \text{ (other than those of } l_a) \text{ pass through } Q.$$  

There is a $\frac{d-2}{d-1} \times \frac{d-3}{d-2}$ probability that the two distinct intersection points between lines of $l_3$ and the considered line of $l_a$ do not coincide with $Q$.

Similarly, there is a $\frac{d-2}{d-1} \times \frac{d-3}{d-2} \times \frac{d-4}{d-3}$ probability that the three distinct intersection points between lines of a set of 6 lines and the considered line of $l_a$ do not coincide with $Q$.

Finally, considering the $\frac{d-5}{6}$ sets of 6 lines and the lines in $l_3$, we obtain in this case:

$$P(M_Q = 1) = \frac{d-3}{d-1} \times \left( \frac{d-4}{d-3} \right)^{\frac{d-5}{6}}.$$  

3.4.2 $Q$ is not in a line of $l_a$

In this case, for $A \in \Omega^{**}$, we have:

$$M_Q(A) = 1 \iff \text{exactly one line of the } d-2 \text{ lines of } A \text{ (other than those of } l_a) \text{ passes through } Q.$$  

We will use the following results to study in more detail the different subcases. In $\mathbb{F}_d^2 \setminus l_a$, there are $d^2-3d+2$ points. In $\mathbb{F}_d^2 \setminus l_a \cup l_3$, there are $3(d-3)$ points of multiplicity 1 and the remaining points of multiplicity 0 ($d^2-6d+11$ points). In the union of $\mathbb{F}_d^2 \setminus l_a$ and a 6 lines set, there are $6(d-5)$ points of multiplicity 1, 6 points of multiplicity 2 (see Proposition 3.4) and the remaining points of multiplicity 0 ($d^2-9d+26$ points).

This case can be divided in 2 subcases:

- the first one where the line that passes through $Q$ is in $l_3$; then the probability is:

$$\frac{3(d-3)}{d^2-3d+2} \times \left( \frac{d^2-9d+26}{d^2-3d+2} \right)^{\frac{d-5}{6}};$$

11
the second one where the line that passes through $Q$ is in one of the $\frac{d-5}{6}$ sets of 6 lines; then the probability is:

$$
\frac{d^2 - 6d + 11}{d^2 - 3d + 2} \times \frac{d - 5}{6} \frac{6(d - 5)}{d^2 - 3d + 2} \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d-5}{6}}
$$

Hence we have in this specific case:

$$
P(M_Q = 1) = \frac{3(d - 3)}{d^2 - 3d + 2} \times \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d-5}{6}}
+ \frac{d^2 - 6d + 11}{d^2 - 3d + 2} \times \frac{d^2 - 10d + 25}{d^2 - 3d + 2} \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d-11}{6}}
$$

3.4.3 The expected value of $M_d^{**}$

Recall that our aim is to determine the expected value $M_d^{**}$ of the number of all but simple points in arrangements of $\Omega^{**}$ in order to compare it with the value of the complexity $C_d$.

Thanks to the results of the above section and knowing that the first case concerns $3d - 3$ points and the second one $d^2 - 3d + 2$ points, we get:

$$
M_d^{**} = d^2 - \left( 3(d - 3) \times \left( \frac{d - 4}{d - 1} \right)^{\frac{d-5}{6}} + 3(d - 3) \times \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d-5}{6}} \right)
+ \frac{d^2 - 6d + 11}{d^2 - 3d + 2} \times \frac{d^2 - 10d + 25}{d^2 - 3d + 2} \left( \frac{d^2 - 9d + 26}{d^2 - 3d + 2} \right)^{\frac{d-11}{6}}
$$

Using the Computer Algebra System Xcas, we obtain:

$$
M_d^{**} = (1 - \frac{1}{e})d^2 + \left( \frac{1}{e} - 3 \exp(-\frac{1}{2}) \right)d + O(1), \quad d \to \infty.
$$

3.5 Results

Figure 4 shows values of both $\frac{C_d - M_d^{**}}{d}$ and $\frac{C_d - M_d^{**}}{d}$ for the selected prime numbers $d$.

Figure 4: Values of both $\frac{C_d - M_d^{**}}{d}$ (on the left) and $\frac{C_d - M_d^{**}}{d}$ (on the right) for the selected prime numbers $d$. 

3.5.1 A first test in each case

From the 152 left plotted values on Figure 4, we draw the regression line: its slope $s^*$ is approximately $1.94 \times 10^{-4}$ and its intercept is approximately 0.352. Let us consider the following null hypothesis $H^*_0$: $s^* = 0$. We have to calculate $T^* = \frac{s^* - 0}{\hat{\sigma}_{s^*}}$, where $\hat{\sigma}_{s^*}$ is the estimated standard deviation of the slope. We obtain $\hat{\sigma}_{s^*} \approx 1.35 \times 10^{-4}$ and $T^* \approx 1.43$. $T^*$ follows a student’s t-distribution with $(152 - 2)$ degrees of freedom [3, Proposition 1.8]. The acceptance region of the hypothesis test with a 5% risk is approximately $[-1.976, 1.976]$. Thus it can be concluded that we cannot reject the null hypothesis, i.e. the fact that the slope $s^*$ is not significantly different from zero.

From the 153 right plotted values on Figure 4, we draw the regression line: its slope $s^{**}$ is approximately $-1.88 \times 10^{-4}$ and its intercept is approximately 0.511. Let us consider the following null hypothesis $H^{**}_0$: $s^{**} = 0$. We again have to calculate $T^{**} = \frac{s^{**} - 0}{\hat{\sigma}_{s^{**}}}$. We here obtain $\hat{\sigma}_{s^{**}} \approx 1.25 \times 10^{-4}$ and $T^{**} \approx -1.50$. $T^{**}$ follows a student’s t-distribution with $(153 - 2)$ degrees of freedom. The acceptance region of the hypothesis test with a 5% risk is approximately $[-1.976, 1.976]$. Thus it can be concluded that we cannot reject the fact that the slope $s^{**}$ is not significantly different from zero.

3.5.2 A set of tests in each case

Figure 5 below shows the distribution of the values of $\frac{C_d - M^*_d}{d}$ (on the left) and $\frac{C_d - M^{**}_d}{d}$ (on the right) for the considered values of $d$.

![Figure 5: Distribution of values of both $\frac{C_d - M^*_d}{d}$ (on the left) and $\frac{C_d - M^{**}_d}{d}$ (on the right) for the selected prime numbers $d$.](image)

From the result of the first test in section 3.5.1 we would consider in this part that the function that maps $d$ onto $\frac{C_d - M^*_d}{d}$ behaves like a random variable with an expected value $\Lambda^*$ close to 0.576 and with a symmetric probability distribution (for the considered values of $d$). On that assumption we will verify whether the values higher than $\Lambda^*$ and those smaller than $\Lambda^*$ are randomly scattered over the ordered absolute values of $\frac{C_d - A^*_d}{d}$ (the null hypothesis) or
not. To this end we use a non-parametric test, the Mann–Whitney $U$ test: we determine the ranks of $\frac{C_d - M_d}{d}$ for each $d$ in the considered interval (see [9] or [5]). The ranks sum of the values higher than $\Lambda^*$ is approximately normally distributed. The value of $U_1^*$ is about $-0.673$. The acceptance region of the hypothesis test with a 5% risk being approximately $[-1.960, 1.960]$, it can be concluded that we cannot reject the null hypothesis, i.e. the fact that the greater and smaller than $\Lambda^*$ values of $\frac{C_d - M_d^*}{d}$ are randomly scattered: the symmetry of the probability distribution of this potential pseudorandom variable can not be rejected.

On the same assumption, we will also verify whether the values higher than $\Lambda^*$ and those smaller than $\Lambda^*$ are randomly scattered over the considered prime numbers (the null hypothesis) or not. To this end we again use the Mann–Whitney $U$ test. The prime numbers ranks sum of the values higher than $\Lambda^*$ is approximately normally distributed. The value of $U_2^*$ is about 0.721. It can be concluded that we cannot reject the null hypothesis, i.e. the fact that the greater and smaller than $\Lambda^*$ values of $\frac{C_d - M_d^*}{d}$ are randomly scattered over the considered prime numbers.

From the result of the second test in 3.5.1, we would consider in this part that the function that maps $d$ onto $\frac{C_d - M_d^*}{d}$ behaves like a random variable with an expected value $\Lambda^{**}$ close to 0.297 and with a symmetric probability distribution (for the considered values of $d$). On that assumption we will verify whether the values higher than $\Lambda^{**}$ and those smaller than $\Lambda^{**}$ are randomly scattered over the ordered absolute values of $\frac{C_d - M_d^*}{d}$ (the null hypothesis) or not. To this end we again use the Mann–Whitney $U$ test. The value of $U_2^{**}$ is here about $-1.08$. It can once more be concluded that we cannot reject the fact that the greater and smaller than $\Lambda^{**}$ values of $\frac{C_d - M_d^{**}}{d}$ are randomly scattered: the symmetry of the probability distribution of this potential pseudorandom variable can not be rejected.

On the same assumption, we will verify whether the values higher than $\Lambda^{**}$ and those smaller than $\Lambda^{**}$ are randomly scattered over the considered prime numbers or not. To this end we again use the Mann–Whitney $U$ test. The value of $U_2^{**}$ is here about $-1.77$. It can once more be concluded that we cannot reject the fact that the greater and smaller than $\Lambda^{**}$ values of $\frac{C_d - M_d^{**}}{d}$ are randomly scattered over the considered prime numbers.

3.5.3 Perspective

Both first test and set of tests could not invalidate the fact that the function that maps $d$ onto $\frac{C_d - M_d}{d}$ and the one that maps $d$ onto $\frac{C_d - M_d^*}{d}$ seem to behave like random variables with respectively $\Lambda^*$ and $\Lambda^{**}$ as expected values. $\Lambda^*$ and $\Lambda^{**}$ are both positive numbers, whereas $\Lambda$ is negative; the added geometrical constraints seem to reduce in average the number of all but the simple points generated by a randomly chosen minimal Besicovitch arrangement. This reduction is slightly higher than expected. Our arrangements cannot obviously be limited to the considered geometrically constrained arrangement. Adding constraints for better modeling the arrangements and finding a way to determine whether the considered functions could be considered as high-quality pseudorandom number generators (PRNG) sketch some avenues for future research on the subject.
## Appendix

| d | 1 | 3.5 | 7 | 11 | 15 | 19 | 23 | 27 | 31 | 35 | 43 | 47 | 53 | 59 | 61 | 67 | 71 |
|---|---|----|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Cd | 1 | 13 | 25 | 61 | 105 | 139 | 173 | 207 | 241 | 275 | 319 | 353 | 387 | 421 | 455 | 489 | 523 | 557 |
| Cd | 1 | 151 | 196 | 241 | 286 | 331 | 376 | 421 | 466 | 511 | 556 | 601 | 646 | 691 | 736 | 781 | 826 | 871 | 916 |

### Table 1: Data for Appendix

| d | 3.5 | 7 | 11 | 15 | 19 | 23 | 27 | 31 | 35 | 43 | 47 | 53 | 59 | 61 | 67 | 71 | 75 | 80 |
|---|----|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Cd | 2 | 29 | 35 | 41 | 47 | 53 | 59 | 65 | 71 | 77 | 83 | 89 | 95 | 101 | 107 | 113 | 119 | 125 | 131 |
| Cd | 3 | 36 | 42 | 48 | 54 | 60 | 66 | 72 | 78 | 84 | 90 | 96 | 102 | 108 | 114 | 120 | 126 | 132 | 138 | 144 |

### Table 2: Additional Data for Appendix
Table 1: The complexities values. Values of $d$ with one asterisk correspond to arrangements where $d \equiv 1 \pmod{3}$ and where all the lines (except those of $\{L_0, L_{-1}, L_{\infty}\}$ and $\{L_{\omega}, L_{\omega^2}\}$) do not pass through the origin, whereas values of $d$ with two asterisks correspond to arrangements where $d \equiv 2 \pmod{3}$ and where all the lines (except those of $\{L_0, L_{-1}, L_{\infty}\}$) do not pass through the origin.

References

[1] F. Arnault, E.J. Pickett, and S. Vinatier. Construction of sel-dual normal bases and their complexity. *Finite Fields and their applications*, 18:458–472, 2012.

[2] S. Blondeau Da Silva. On the arrangements of $q + 1$ lines with different slopes in $\mathbb{P}^2_q$. *Journal of Number Theory*, 180C:533–543, 2017.

[3] P.-A. Cornillon and E. Matzner-Lober. *Régression : Théorie et applications*. Springer, 2007. 302 p.

[4] F. G. Dorais and D. Klyve. A Wieferich prime search up to $6.7 \times 10^{15}$. *Journal of Integer Sequences*, 14(9):Article 11.9.2, 14 p., electronic only, 2011.

[5] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18(1):50–60, 1947.

[6] A. J. Menezes, I. F. Blake, S. Gao, R. C. Mullin, S. A. Vanstone, and T. Yaghoobian, editors. *Applications of Finite Fields*. Kluwer Academic Publishers, 1993.

[7] E.J. Pickett and S. Vinatier. About the complexity of cyclotomic self dual normal bases. *In preparation*.

[8] S. S. Shapiro and Wilk M. B. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4):591–611, 1965.
[9] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945.