AofA'12 DMTCS proc. **AQ**, 2012, 399–412

Analysis of Digital Expansions of Minimal Weight

Florian Heigl[1 †] and Clemens Heuberger[2 ‡]

[1] *TU Graz, Austria*
[2] *Alpen-Adria-Universität, Klagenfurt, Austria*

Extending an idea of Suppakitpaisarn, Edahiro and Imai, a dynamic programming approach for computing digital expansions of minimal weight is transformed into an asymptotic analysis of minimal weight expansions. The minimal weight of an optimal expansion of a random input of length ℓ is shown to be asymptotically normally distributed under suitable conditions.

After discussing the general framework, we focus on expansions to the base of τ, where τ is a root of the polynomial $X^2 - \mu X + 2$ for $\mu \in \{\pm 1\}$. As the Frobenius endomorphism on a binary Koblitz curve fulfils the same equation, digit expansions to the base of this τ can be used for scalar multiplication and linear combination in elliptic curve cryptosystems over these curves.

Keywords: digital expansions, Hamming weight, elliptic curve cryptography, Frobenius endomorphism, minimal expansions, limit distribution, dynamic programming

1 Introduction

Digital expansions are one method for efficient implementations of scalar multiplication (or linear combinations) in Abelian groups, such as the point group of elliptic curves. One application is in public key cryptography, where scalar multiplication is a key ingredient.

Let G be an Abelian group. A positive integer $n \in \mathbb{Z}$ can also be seen as an endomorphism of G by setting $nQ = Q + \cdots + Q$ (with n summands) for $Q \in G$. This immediately carries over to all integers in $\mathbb{Z}$. We consider an algebraic integer τ which also acts as an endomorphism on G. This action can easily be extended to an action of the ring $\mathbb{Z}[\tau]$ on G.

In order to compute nP for some $n \in \mathbb{Z}[\tau]$ and some $P \in G$, we consider a digital expansion

$$n = \sum_{j=0}^{\ell-1} d_j \tau^j \tag{1}$$

[†]Email: `floheigl@sbox.tugraz.at`. Both authors are supported by the Austrian Science Fund (FWF): S09606, that is part of the Austrian National Research Network “Analytic Combinatorics and Probabilistic Number Theory.”

[‡]Email: `clemens.heuberger@aau.at`

to the base τ and digits d_j from a finite digit set $\mathcal{D}$. Then nP can be computed by Horner's scheme as

$$nP = \tau(\tau(\cdots\tau(\tau(d_{\ell-1}P) + d_{\ell-2}P) + \cdots) + d_1P) + d_0P. \tag{2}$$

If dP is precomputed for all digits $d \in \mathcal{D}$, the computation of nP requires $\ell - 1$ applications of τ and $w - 1$ additions in G, where w denotes the weight of the expansion (1). Here, the weight of the expansion is defined to be the number of non-zero digits d_j.

If the digit set $\mathcal{D}$ is sufficiently large, then every $n \in \mathbb{Z}[\tau]$ will have many expansions (1). In order to speed up the computation of nP, we are interested in digital expansions of n of minimal weight.

The basic example are binary expansions with $\tau = 2$ and digits $\mathcal{D} = \{0, 1\}$. In this case, every non-negative $n \in \mathbb{Z}$ can be represented uniquely and (2) corresponds to the so-called double-and-add algorithm, cf. for instance Knuth (1998).

Redundant expansions to the base $\tau = 2$ and digit set $\mathcal{D} = \{0, \pm 1\}$ have also been studied intensively. Reitwiesner (1960) proved that every $n \in \mathbb{Z}$ admits a unique expansion (1) with $\tau = 2$ and $\mathcal{D} = \{0, \pm 1\}$ with the extra syntactic property that $d_j d_{j+1} = 0$ for all j, *i.e.* that there are no adjacent non-zero digits. This expansion is usually called the non-adjacent form (NAF) of n. Reitwiesner showed that the NAF of n has minimal weight among all expansions of n with this digit set. Several larger digit sets for $\tau = 2$ have been considered: see Phillips and Burgess (2004) for the digit set $\mathcal{D} = \{m, m+1, \ldots, u-1, u\}$ for some $m \leq 0$ and $u \geq 1$.

Koblitz (1992) studied the curves

$$E_a\colon y^2 + xy = x^3 + ax^2 + 1 \qquad \text{with} \qquad a \in \{0, 1\} \tag{3}$$

over the field $\mathbb{F}_{2^m}$; these curves are nowadays also known as binary Koblitz curves. The Frobenius endomorphism $\varphi : x \mapsto x^2$ of the field $\mathbb{F}_{2^m}$ can be extended to the curve, mapping points $(x, y) \in E_a(\mathbb{F}_{2^m})$ to $(x^2, y^2) \in E_a(\mathbb{F}_{2^m})$. As the Frobenius endomorphism fulfils

$$\varphi(\varphi(Q)) - \mu\varphi(Q) + 2Q = 0 \quad \text{for} \quad Q \in E_a(\mathbb{F}_{2^m})$$

with $\mu = (-1)^{a+1}$, the imaginary quadratic integer

$$\tau = \frac{\mu + \sqrt{-7}}{2} \tag{4}$$

can be identified with φ, *i.e.*, $\mathbb{Z}[\tau]$ acts on $E_a(\mathbb{F}_{2^m})$ via $\tau(Q) := \varphi(Q)$ for $Q \in E_a(\mathbb{F}_{2^m})$. As the Frobenius endomorphism is computationally cheaper than doubling on the curve, (2) turns into an efficient scalar multiplication algorithm.

For this τ and an integer $w \geq 2$, Solinas (2000) proposed the digit set $\mathcal{D}$ consisting of 0 and the representative of minimal norm of each residue class of $\mathbb{Z}[\tau]$ modulo τ^w which is not divisible by τ. He showed that each element $n \in \mathbb{Z}[\tau]$ admits an expansion (1) with these parameters with the property that each block of w consecutive digits contains at most one non-zero element. This generalisation of the concept of a NAF is called a τ-w-NAF. While the τ-w-NAF has a low Hamming weight, it is not optimal for $w \geq 4$ in the sense that that it does not necessarily minimise the Hamming weight over all expansions of the same $n \in \mathbb{Z}[\tau]$ with digits in $\mathcal{D}$. Even worse, it can be shown that for $w \in \{4, 5, 6\}$, it is impossible to compute an optimal expansion from right to left, *i.e.*, from the least significant digit to the most significant digit, from any other expansion by an online algorithm with finite look-ahead, cf. Heuberger (2010). This is in contrast to the case of base-2 expansions.

It is known that the average weight of a τ-w-NAF of length ℓ to the base τ given in (4) is $\ell/(w+1)+O(1)$. This contribution is devoted to the following question: What is the average weight of the minimal expansions (with the same digit set and base) of all elements of $\mathbb{Z}[\tau]$ given by a τ-1-NAF—*i.e.*, an expansion with digits in $\{0,1\}$ and no syntactic conditions—of length ℓ? We prove that the answer is $28\ell/141+O(1)\approx 0.198582\ell+O(1)$ for $w=4$ and $30\ell/181+O(1)\approx 0.165746\ell+O(1)$ for $w=5$, cf. Theorem 5.

We now turn to the computation of linear combinations $mP+nQ$ for P, $Q\in G$ and m, $n\in\mathbb{Z}[\tau]$. The obvious approach of computing mP and nQ separately and adding the results can be improved by "Shamir's trick", going back to Straus (1964): we consider expansions of the vector $\binom{m}{n}$; the digit set $\mathcal{D}$ now also consists of vectors (leading to the notion of "joint expansions"). A digit $\eta=\binom{\eta_1}{\eta_2}$ translates into an addition of $\eta_1 P+\eta_2 Q$, which has been precomputed for all digits. The weight of such an expansion (also called the joint weight) is the number of digits which are not the zero vector. Therefore, the weight again corresponds to the number of group additions. Obviously, the concept can be generalised to higher dimensions.

Solinas (2001) started the study of joint expansions in base 2 with digit set $\{0,\pm1\}^2$; a generalisation to digit sets $\{m,m+1,\ldots,u-1,u\}^r$ for some $m\le 0$ and $u\ge 1$ and $r\ge 1$ is given in Heuberger and Muir (2007) (see also the references there for intermediate results). As in the one-dimensional case, some syntactically defined expansion minimises the weight.

The situation changes if only zero and odd digits are allowed, *i.e.*, the digit set is $\mathcal{D}=\{0,\pm1,\pm3,\ldots,\pm(2h-1),\pm(2h+1)\}^r$ for $h\ge 0$ and $r\ge 1$. Apart from the cases $r=1$ or $h=0$, no syntactic condition guaranteeing minimality seems to be known. However, Suppakitpaisarn et al. (2010) could give the average weight of minimal weight expansions of r-tuples of integers whose standard binary expansion is of length ℓ for $r=2$ and $1\le h\le 5$ as well as for $(r,h)=(3,1)$.

For the base of the imaginary quadratic number τ given in (4), *i.e.*, the case of binary Koblitz curves, the case of dimension $r=2$ and digit set $\mathcal{D}=\{0,\pm1\}^2$ attracted some attention. Again, there cannot be an online algorithm computing an optimal expansion from right to left, cf. Heuberger (2010) for this result and a survey on earlier results. In this contribution, we show that the average minimal weight of minimal joint expansions of vectors in $\mathbb{Z}[\tau]^2$ whose $\{0,1\}$ expansions to the base of τ have length at most ℓ is given by $0.4649\ldots\ell+O(1)$. Here, the coefficient of the main term is the quotient of two integers with 66 digits (in decimal notation), cf. Theorem 5.

For obtaining their results, Suppakitpaisarn, Edahiro and Imai considered a dynamic programming approach for computing minimal weight expansions from left to right. If only the weight (and not the minimal expansion itself) is of interest, the average weight analysis can be carried out algorithmically for given basis and digit set.

The contributions of this paper are:

- a succinct, but complete, formulation of the method of Suppakitpaisarn, Edahiro and Imai and its generalisation to arbitrary digit sets and bases (Section 4),
- a proof that the underlying directed graph is aperiodic and strongly connected which is independent of the basis (Section 5),
- the random variable W_ℓ, defined to be the minimal weight of the optimal expansions of all inte-

ger (vectors) of "input length" ℓ is shown to be asymptotically normally distributed under some technical conditions,

- specific results for the basis τ given in (4): both a proof that the crucial subadditivity property holds in the one-dimensional case (Section 6) and the main term of the expectation $E(W_\ell)$ (Section 8).

2 Notations and Definitions

The following notations and definitions will be used throughout this paper.

Let τ be a rational integer or a imaginary quadratic integer. We will consider expansions of integers (or integer vectors) in $\mathcal{R} := \mathbb{Z}[\tau]^r$ for some $r \geq 1$. A digit set $\mathcal{D}$ is a finite subset of $\mathcal{R}$.

For a word $d_{\ell-1} \dots d_0$ in $\mathcal{D}^*$, we set

$$\begin{aligned} \mathsf{length}(d_{\ell-1} \dots d_0) &:= \ell, \\ \mathsf{weight}(d_{\ell-1} \dots d_0) &:= \#\{0 \leq j < \ell : d_j \neq 0\}, \\ \mathsf{value}(d_{\ell-1} \dots d_0) &:= \sum_{j=0}^{\ell-1} d_j \tau^j. \end{aligned}$$

The weight is also known as Hamming weight, in dimension > 1, it is usually called the joint (Hamming) weight. A word $\mathbf{d} \in \mathcal{D}^*$ is said to be a $\mathcal{D}$-expansion of $z \in \mathcal{R}$ if $z = \mathsf{value}(\mathbf{d})$. The word $d_{\ell-1} \dots d_0 \in \mathcal{D}^*$ is said to be *trimmed* if it is the empty word or if $d_{\ell-1} \neq 0$.

We say that $(\tau, \mathcal{D})$ is a digit system if for each $z \in \mathcal{R}$, there is a $\mathbf{d} \in \mathcal{D}^*$ with $z = \mathsf{value}(\mathbf{d})$. A digit system is called irredundant, if there is a unique trimmed word $\mathbf{d} \in \mathcal{D}^*$ with $z = \mathsf{value}(\mathbf{d})$ for every $z \in \mathcal{R}$; otherwise, it is called irredundant.

For a digit system $(\tau, \mathcal{D})$ and a $z \in \mathcal{R}$, we set the minimal weight of z with respect to $\mathcal{D}$ to be

$$\mathsf{mw}_{\mathcal{D}}(z) := \min\{\mathsf{weight}(\mathbf{d})\colon \mathbf{d} \in \mathcal{D}^*, \mathsf{value}(\mathbf{d}) = z\}.$$

A word $\mathbf{d} \in \mathcal{D}^*$ is said to be a minimal $\mathcal{D}$-expansion if $\mathsf{mw}_{\mathcal{D}}(\mathsf{value}(\mathbf{d})) = \mathsf{weight}(\mathbf{d})$.

A digit set $(\tau, \mathcal{D})$ is said to fulfil the subadditivity condition with respect to $c \in \mathcal{R}$, if there is a constant $U(c)$ such that

$$|\mathsf{mw}_{\mathcal{D}}(z + c) - \mathsf{mw}_{\mathcal{D}}(z)| \leq U(c)$$

holds for all $z \in \mathcal{R}$. The digit set is said to fulfil the subadditivity condition, if it fulfils the subadditivity condition with respect to all $c \in \mathcal{R}$.

If $(\tau, \mathcal{D})$ is an irredundant digit set, we set

$$\mathsf{length}_{\mathcal{D}}(z) := \mathsf{length}(\mathbf{d})$$

for the unique trimmed word $\mathbf{d} \in \mathcal{D}^*$ with $z = \mathsf{value}(\mathbf{d})$.

In the remainder of this paper, we will consider two digit systems: an irredundant digit system $(\tau, \mathcal{D}_{in})$ and a redundant digit system $(\tau, \mathcal{D}_{out})$. We are interested in the average of $\mathsf{mw}_{\mathcal{D}_{out}}(z)$ over all $z \in \mathcal{R}$ with $\mathsf{length}_{\mathcal{D}_{in}}(z) = \ell$ for some ℓ. In order to avoid double subscripts, we simply write $\mathsf{mw}(z)$ for $\mathsf{mw}_{\mathcal{D}_{out}}(z)$ and $\mathsf{length}(z)$ for $\mathsf{length}_{\mathcal{D}_{in}}(z)$.

3 Carry Set

Assume that we have a $\mathcal{D}_{in}$-expansion $\mathbf{d} \in \mathcal{D}_{in}^*$ and a $\mathcal{D}_{out}$-expansion $\boldsymbol{\eta} \in \mathcal{D}_{out}^*$ of the same integer $z \in \mathcal{R}$, both of length ℓ. For $0 \leq k < \ell$, we have

$$\mathsf{value}(\eta_{\ell-1} \dots \eta_k)\tau^k + \mathsf{value}(\eta_{k-1} \dots \eta_0) = z = \mathsf{value}(d_{\ell-1} \dots d_k)\tau^k + \mathsf{value}(d_{k-1} \dots d_0)$$

and consequently

$$\mathsf{value}(\eta_{\ell-1} \dots \eta_k) = \mathsf{value}(d_{\ell-1} \dots d_k) + c_k$$

for the *carry*

$$c_k = \frac{\mathsf{value}(d_{k-1} \dots d_0) - \mathsf{value}(\eta_{k-1} \dots \eta_0)}{\tau^k}.$$

This motivates the definition of the *carry set* as

$$\mathcal{C} := \left\{ \frac{\mathsf{value}(d_{k-1} \dots d_0) - \mathsf{value}(\eta_{k-1} \dots \eta_0)}{\tau^k} : k \geq 0, d_{k-1} \dots d_0 \in \mathcal{D}_{in}^*, \right.$$
$$\left. \eta_{k-1} \dots \eta_0 \in \mathcal{D}_{out}^*, \mathsf{value}(d_{k-1} \dots d_0) \equiv \mathsf{value}(\eta_{k-1} \dots \eta_0) \pmod{\tau^k} \right\}. \quad (5)$$

Lemma 3.1 *The carry set $\mathcal{C}$ is a finite subset of $\mathcal{R}$.*

Proof: The congruence condition in the definition of $\mathcal{C}$ immediately implies that $\mathcal{C} \subseteq \mathcal{R}$. For any vector norm $\|\cdot\|$, the triangle inequality and summation of the resulting geometric series yields the upper bound

$$\|c\| \leq \frac{\max\{\|d\| : d \in \mathcal{D}_{in}\} + \max\{\|\eta\| : \eta \in \mathcal{D}_{out}\}}{\|\tau\| - 1}$$

for all carries $c \in \mathcal{C}$. This immediately implies that $\mathcal{C}$ is finite. □

In order to compute $\mathcal{C}$, we rewrite the definition of $\mathcal{C}$ slightly:

Lemma 3.2 *Let $c \in \mathcal{R}$. Then $c \in \mathcal{C}$ if and only if*

$$c = 0 \quad \text{or} \quad c = \frac{d - \eta + c'}{\tau}$$

for some $d \in \mathcal{D}_{in}$, $\eta \in \mathcal{D}_{out}$ and $c' \in \mathcal{C}$.

Proof: This follows from (5) by setting $d = d_{k-1}$, $\eta = \eta_{k-1}$ and

$$c' = \frac{\mathsf{value}(d_{k-2} \dots d_0) - \mathsf{value}(\eta_{k-2} \dots \eta_0)}{\tau^{k-1}}.$$

□

This implies that $\mathcal{C}$ is the set of vertices reachable from 0 in the infinite directed graph with set of vertices $\mathcal{R}$ which has an arc from some c' to some $c \in \mathcal{R}$ if and only if $c = (d - \eta + c')/\tau$ for some $d \in \mathcal{D}_{in}$ and $\eta \in \mathcal{D}_{out}$. Due to the finiteness of $\mathcal{C}$, this set can be computed by depth-first search, for instance.

4 Transducer Automaton

In this section, we introduce a transducer automaton which computes $\mathsf{mw}(z)$ for a $z \in \mathcal{R}$ given by its expansion in the input digit set. The underlying idea is dynamic programming, the states (or rather equivalence classes of states) are encoded in the transducer automaton.

We consider the transducer automaton $\mathcal{T}_0$ with set of states

$$V_0 = \{f : \mathcal{C} \to \mathbb{Z} : f(0) = 0\}$$

and transitions

$$f \xrightarrow{d|w} g$$

for $f, g \in V_0$, $d \in \mathcal{D}_{in}$,

$$g(c) = \min\left\{ f\left(\frac{d+c-\eta}{\tau}\right) + \mathsf{weight}(\eta) : \eta \equiv d + c \pmod{\tau}, \eta \in \mathcal{D}_{out} \right\} - w$$

for $c \in \mathcal{C}$, where $w \in \mathbb{Z}$ is chosen such that $g(0) = 0$ holds as required. The initial state f_I is defined as

$$f_I(c) = \mathsf{mw}(c) \quad \text{for} \quad c \in \mathcal{C}.$$

We note that $\mathcal{T}_0$ is a deterministic transducer automaton.

We set V to be the set of states reachable from the initial state f_I and $\mathcal{T}$ to be the transducer automaton arising from restricting $\mathcal{T}_0$ to the set of states V.

The following lemma states the main invariant of the transducer.

Lemma 4.1 *Let $d_{\ell-1} \ldots d_0 \in \mathcal{D}_{in}^*$ and $0 \le k \le \ell$. We consider the unique path in $\mathcal{T}$ from the initial state f_I with input label $d_{\ell-1} \ldots d_k$. It is leading to some state g_k, its output label is denoted by $w_{\ell-1} \ldots w_k$.*

Then we have

$$g_k(c) + \sum_{j=k}^{\ell-1} w_j = \mathsf{mw}(z_k + c) \tag{6}$$

and

$$g_k(c) = \mathsf{mw}(z_k + c) - \mathsf{mw}(z_k) \tag{7}$$

for all $c \in \mathcal{C}$, where $z_k = \mathsf{value}(d_{\ell-1} \ldots d_k)$.

Proof: We prove (6) by backwards induction on k. For $k = \ell$, we have $z_\ell = 0$ (as the value of the empty word), thus (6) is equivalent to the definition of $g_\ell = f_I$.

Assume now that (6) is true for some $k + 1$. For $c \in \mathcal{C}$, we have

$$g_k(c) + \sum_{j=k}^{\ell-1} w_j = \min\Bigg\{ g_{k+1}\left(\frac{c + d_k - \eta}{\tau}\right) + \sum_{j=k+1}^{\ell-1} w_j + \mathsf{weight}(\eta) : \\ \eta \in \mathcal{D}_{out}, \eta \equiv c + d_k \pmod{\tau} \Bigg\}$$

by definition of $g_k(c)$. Inserting the induction hypothesis (6) yields

$$\begin{aligned}g_k(c)+\sum_{j=k}^{\ell-1} w_j &= \min\left\{\mathsf{mw}\left(z_{k+1}+\frac{c+d_k-\eta}{\tau}\right)+\mathsf{weight}(\eta) : \eta\in\mathcal{D}_{out}, \eta\equiv c+d_k \pmod{\tau}\right\}\\ &= \min\left\{\mathsf{mw}\left(\frac{z_k+c-\eta}{\tau}\right)+\mathsf{weight}(\eta) : \eta\in\mathcal{D}_{out}, \eta\equiv z_k+c \pmod{\tau}\right\}\\ &= \mathsf{mw}(z_k+c).\end{aligned}$$

This concludes the proof of (6). Subtracting (6) for c and for 0 yields (7). □

As an immediate consequence of Lemma 4.1, we obtain the following theorem.

Theorem 1 *Let $z=\mathsf{value}(\mathbf{d})$ for some $\mathbf{d}\in\mathcal{D}_{in}^*$. Then $\mathsf{mw}(z)$ is given by the output label $\mathbf{w}$ of the path from f_I with input label $\mathbf{d}$ in $\mathcal{T}$. More precisely, we have*

$$\mathsf{mw}(z)=\sum_{j=0}^{\ell-1} w_j$$

for $\ell=\mathsf{length}(\mathbf{d})=\mathsf{length}(\mathbf{w})$.

If $(\tau,\mathcal{D}_{out})$ fulfils the subadditivity property, then $\mathcal{T}$ is a finite transducer automaton.

Proof: The first part is a direct consequence of Lemma 4.1. The finiteness property follows from (7) and the subadditivity property of $(\tau,\mathcal{D}_{out})$. □

5 Reset Sequence

It is the aim of this section to show that under a suitable condition on the digit set, the transducer $\mathcal{T}$ has a synchronising word of the shape 0^ℓ for sufficiently large ℓ, *i.e.*, from each state of $\mathcal{T}$, the input 0^ℓ leads to the initial state f_I.

We will assume that

$$\text{if } d\in\mathcal{D}_{out} \text{ is divisible by } \tau, \text{ then } d=0. \tag{8}$$

This restriction does not hurt in applications, as it seems reasonable to use 0 as a least significant digit for expanding any $z\in\mathcal{R}$ which is divisible by τ. In fact, this condition is met in all examples we are considering.

Lemma 5.1 *Let $\mathcal{D}_{out}$ fulfil* (8)*. For $z\in\mathcal{R}$ and $k\geq 0$, we set*

$$L(z,k):=\sup\{\mathsf{length}(\boldsymbol{\eta}) : \boldsymbol{\eta}\in\mathcal{D}_{out}^* \textit{ is a trimmed } \mathcal{D}_{out}\textit{-expansion of } z \textit{ with } \mathsf{weight}(\boldsymbol{\eta})\leq k\},$$

where the supremum of an empty set is defined to be $-\infty$.

Then $L(z,k)<\infty$.

Proof: We prove the lemma by induction on k. For $k = 0$, we have $L(0, 0) = 0$ and $L(z, 0) = -\infty$ for $z \neq 0$. We now assume that the assertion has been shown for some $k \geq 0$ and all $z \in \mathcal{R}$.

Let $z \in \mathcal{R}$. If $z = 0$, then $L(z, k+1) = 0$, as the empty word is the only trimmed $\mathcal{D}_{out}$-expansion of $z = 0$ by (8). So we may assume that $z \neq 0$. There is a maximal integer $j \geq 0$ such that z is divisible by τ^j. Then the word $0 \ldots 0$ consisting of j zeros is a suffix of every $\mathcal{D}_{out}$-expansion of z by (8), which implies that $L(z, k+1) = j + L(z/\tau^j, k+1)$. Therefore, we may assume that τ does not divide z. Let now $\boldsymbol{\eta} = \eta_{\ell-1} \ldots \eta_0$ be a trimmed $\mathcal{D}_{out}$-expansion of z of weight at most $k+1$. Then $\eta_0 \neq 0$ and $\eta_{\ell-1} \ldots \eta_1$ is a $\mathcal{D}_{out}$-expansion of $(z - \eta_0)/\tau$ of weight at most k. We conclude that

$$L(z, k+1) \leq 1 + \max \left\{ L\left(\frac{z-\eta}{\tau}, k\right) : \eta \in \mathcal{D}_{out} \text{ with } z \equiv \eta \pmod{\tau} \right\}.$$

By the finiteness of $\mathcal{D}_{out}$ and the induction hypothesis, we conclude that $L(z, k+1) < \infty$. □

Lemma 5.2 *Let $\mathcal{D}_{out}$ fulfil* (8) *and $f \in V$ be a state of $\mathcal{T}$. Then there is a positive integer L such that the path in $\mathcal{T}$ from f with input label* $0 \ldots 0$ *consisting of L zeros leads to the initial state f_I.*

Proof: Let

$$L \geq \max\{L(c, \mathsf{mw}(c) + \mathsf{mw}(c') - f(c')) : c, c' \in \mathcal{C}\} \tag{9}$$

and

$$L \geq \max\{\mathsf{length}(c) : c \in \mathcal{C}\} \tag{10}$$

and let the path in $\mathcal{T}$ from f with input label $0 \ldots 0$ consisting of L zeros lead to some state g. Furthermore, let $\mathbf{d}$ be the input label of a path in $\mathcal{T}$ leading from the initial state f_I to the state f and set $z = \mathsf{value}(\mathbf{d})$.

Let $c \in \mathcal{C}$. By (7), we have

$$g(c) = \mathsf{mw}(z\tau^L + c) - \mathsf{mw}(z\tau^L).$$

As $\mathsf{mw}(z\tau^L) = \mathsf{mw}(z)$ by (8) and $f_I(c) = \mathsf{mw}(c)$ by definition of the initial state, we have to show that

$$\mathsf{mw}(z\tau^L + c) = \mathsf{mw}(z) + \mathsf{mw}(c).$$

By (9) (for $c' = 0$), there is a minimal $\mathcal{D}_{out}$-expansion of c of length at most L, so concatenating a minimal $\mathcal{D}_{out}$-expansion of z with a suitable number of zeros and a minimal $\mathcal{D}_{out}$-expansion of c yields a $\mathcal{D}_{out}$-expansion of $z\tau^L + c$ of weight at most $\mathsf{mw}(z) + \mathsf{mw}(c)$, *i.e.*, we have shown that $\mathsf{mw}(z\tau^L + c) \leq \mathsf{mw}(z) + \mathsf{mw}(c)$.

Let $\boldsymbol{\eta} = \eta_{\ell-1} \ldots \eta_0$ be a (not necessarily trimmed) minimal $\mathcal{D}_{out}$-expansion of $z\tau^L + c$ with $\ell \geq L$. Let $\mathbf{e}$ be a $\mathcal{D}_{in}$-expansion of c of length L (such an expansion exists due to (10)). Set

$$c' = \frac{\mathsf{value}(\mathbf{e}) - \mathsf{value}(\eta_{L-1} \ldots \eta_0)}{\tau^L} = \frac{c - \mathsf{value}(\eta_{L-1} \ldots \eta_0)}{\tau^L} = \mathsf{value}(\eta_{\ell-1} \ldots \eta_L) - z.$$

By construction, we have $c' \in \mathcal{C}$ and as a prefix of a minimal $\mathcal{D}_{out}$-expansion, $\eta_{\ell-1} \ldots \eta_L$ is a minimal expansion of $z + c'$. We conclude that

$$\mathsf{weight}(\eta_{L-1} \ldots \eta_0) = \mathsf{mw}(z\tau^L + c) - \mathsf{mw}(z + c') = \mathsf{mw}(z\tau^L + c) - \mathsf{mw}(z) - f(c'). \tag{11}$$

If $c' = 0$, then $\eta_{L-1}\ldots\eta_0$ is a $\mathcal{D}_{out}$-expansion of c and $\eta_{\ell-1}\ldots\eta_L$ is a $\mathcal{D}_{out}$-expansion of z, which implies that $\mathsf{mw}(z\tau^L + c) = \mathsf{mw}(\boldsymbol{\eta}) \geq \mathsf{mw}(z) + \mathsf{mw}(c)$, as requested.

We may therefore restrict our attention to the case $c' \neq 0$. Then $\eta_{L-1}\ldots\eta_0$ is a $\mathcal{D}_{out}$-expansion of $c - c'\tau^L \neq c$. Let $\eta'_{\ell'-1}\ldots\eta'_L$ be a trimmed minimal $\mathcal{D}_{out}$-expansion of c'. Then $\boldsymbol{\eta}' = \eta'_{\ell'-1}\ldots\eta'_L\eta_{L-1}\ldots\eta_0$ is a trimmed $\mathcal{D}_{out}$-expansion of c of length at least $L+1$ (as $c' \neq 0$). As $L+1 > L(c, \mathsf{mw}(c)+\mathsf{mw}(c') - f(c'))$ by (9), we conclude that

$$\mathsf{mw}(c') + \mathsf{weight}(\eta_{L-1}\ldots\eta_0) = \mathsf{weight}(\boldsymbol{\eta}') > \mathsf{mw}(c) + \mathsf{mw}(c') - f(c').$$

Combining this with (11) yields $\mathsf{mw}(z\tau^L + c) > \mathsf{mw}(z) + \mathsf{mw}(c)$, a contradiction.

$\square$

Theorem 2 *If $\mathcal{D}_{out}$ fulfils* (8)*, then the directed graph underlying $\mathcal{T}$ is strongly connected and aperiodic.*

Proof: By definition of $\mathcal{T}$, there is a path from f_I to every other state. Lemma 5.2 shows that there is a path from every state to the state f_I. This implies that the underlying graph is strongly connected.

Let $g \in V$ be the state which is reached from the initial state with input label 0. Then, by (7), we have

$$g(c) = \mathsf{mw}(0 + c) - \mathsf{mw}(0) = \mathsf{mw}(c) = f_I(c)$$

for all $c \in \mathcal{C}$, *i.e.*, $g = f_I$. Thus the underlying graph has a loop. Hence it is aperiodic. $\square$

6 Minimal Norm Representatives in the Case of Binary Koblitz Curves

In this section, we consider the base $\tau = (\mu \pm \sqrt{-7})/2$ with $\mu \in \{\pm 1\}$ corresponding to binary Koblitz curves as outlined in the introduction. We restrict ourselves to the one-dimensional case, *i.e.*, expansions of elements of $\mathbb{Z}[\tau]$.

For fixed $w \geq 2$, we consider the digit set $\mathcal{D}_w$ consisting of 0 and the representative of minimal norm of all residue classes modulo τ^w which are not divisible by τ. This digit set has been introduced by Solinas (2000). In Avanzi et al. (2011) it has been shown that the choice of minimal norm representative is unique.

An element $z \in \mathbb{Z}[\tau]$ is an element of $\mathcal{D}_w$ if and only if $z/\tau^w \in V$, where V is the Voronoi cell of 0 with respect to the lattice $\mathbb{Z}[\tau]$, cf. Heuberger and Krenn (2010). We have

$$\overline{B\left(0, \frac{1}{2}\right)} \subseteq V \subseteq \overline{B\left(0, \frac{2}{\sqrt{7}}\right)},$$

where $\overline{B(0, r)}$ denotes the closed disc of radius r around the origin.

Lemma 6.1 *Let η, η' be nonzero digits in the digit set $\mathcal{D}_w$ of minimal norm representatives modulo τ^w. Then*

$$\mathsf{mw}(\eta + \eta') \leq 2.$$

Proof: As neither η nor η' is divisible by τ, their sum $\eta+\eta'$ is divisible by τ. We choose $k \geq 1$ maximally such that τ^k divides $\eta + \eta'$. By the definition of $\mathcal{D}_w$, there is a unique $\varepsilon \in \mathcal{D}_w$ with

$$\frac{\eta + \eta'}{\tau^k} \equiv \varepsilon \pmod{\tau^w}.$$

We choose ℓ maximally such that $\tau^{k+\ell}$ divides $\eta + \eta' - \tau^k\varepsilon$. By construction of ε, we have $\ell \geq w$. We set $\gamma = (\eta + \eta' - \tau^k\varepsilon)/\tau^{k+\ell}$.

We now have

$$\left|\frac{\gamma}{\tau^w}\right| \leq \frac{2}{\sqrt{7}} \cdot \frac{2^{w/2} + 2^{w/2} + 2^{(k+w)/2}}{2^{(w+k+\ell)/2}} \leq \frac{2}{\sqrt{7}} \cdot \frac{1+1+\sqrt{2}}{2^{(w+1)/2}} < \frac{1}{2}$$

for $w \geq 4$, which implies that both $\varepsilon \in \mathcal{D}_w$ and $\gamma \in \mathcal{D}_w$ in this case, *i.e.*, we found a $\mathcal{D}_w$-expansion of $\eta + \eta'$ of weight at most 2 for $w \geq 4$.

For $w < 4$, a direct computation leads to the same conclusion. □

Theorem 3 *The digit set $\mathcal{D}_w$ fulfils the subadditivity condition for $w \geq 2$, more precisely, we have*

$$\mathsf{mw}(x+y) \leq \mathsf{mw}(x) + \mathsf{mw}(y)$$

for all x, $y \in \mathbb{Z}[\tau]$.

Proof: Let

$$n = \mathsf{mw}(x) + \mathsf{mw}(y) \quad \text{and} \quad v = \max\{k \geq 0 : \tau^k \text{ divides } x+y\}.$$

We prove the assertion by induction on (n, k), ordered lexicographically. The cases $n < 2$ are trivial.

Let $\boldsymbol{\alpha}$ be a minimal expansion of x and $\boldsymbol{\beta}$ be a minimal expansion of y.

We first consider the case that $v = 0$. Without loss of generality, we have $\alpha_0 = 0$ and $\beta_0 \neq 0$. Applying the induction hypothesis on x/τ and $(y - \beta_0)/\tau$ and appending β_0 yields an expansion of $x+y$ of weight $\leq \mathsf{mw}(x) + \mathsf{mw}(y)$, as required.

Next, we consider the case that both $\alpha_0 = 0$ and $\beta_0 = 0$. Applying the induction hypothesis on x/τ and y/τ and appending a digit 0 again yields a suitable expansion of $x + y$.

Finally, we consider the case that both $\alpha_0 \neq 0$ and $\beta_0 \neq 0$. We apply the induction hypothesis on $(x - \alpha_0)/\tau$ and $(y - \beta_0)/\tau$ to get an expansion $\boldsymbol{\alpha}'$ of $x' = (x - \alpha_0)/\tau + (y - \beta_0)/\tau$ with $\mathsf{weight}(\boldsymbol{\alpha}') \leq \mathsf{mw}(x) + \mathsf{mw}(y) - 2$. On the other hand, we set $y' = (\alpha_0 + \beta_0)/\tau$. By Lemma 6.1, we have $\mathsf{mw}(y') \leq 2$. As $x' + y' = (x+y)/\tau$, we can apply the induction hypothesis on x' and y' and get the required expansion. □

7 Analysis

In this section, we study the minimum weight W_ℓ of a random $z \in \mathcal{R}$ whose $\mathcal{D}_{in}$-expansion has length ℓ, where all of those z are considered to be equally likely. We give expectation $\mathbb{E}(W_\ell)$ and variance $\mathbb{V}(W_\ell)$ in terms of the transducer. Provided that the usual variability condition holds, the random variables W_ℓ tend to normal distribution.

Numbering the states of $\mathcal{T}$ as $\{1, \dots, N\}$ where the initial state f_I is the first state, we consider the matrix $T(q, u)$ where the entry in row i, column j is

$$\frac{1}{\#\mathcal{D}_{in}} z u^w$$

if there is a transition from state i to state j with output label w and 0 otherwise.

Let $C(q, u) = \det(I - T(q, u))$ and set

$$C_{i,j} := \left. \frac{\partial^{i+j}}{\partial q^i \partial u^j} C(q, u) \right|_{\substack{q=1\\u=1}}$$

for $i, j \geq 0$ and $i + j \leq 2$.

Theorem 4 *Assume that $\mathcal{D}_{out}$ fulfils the subadditivity property and that* (8) *holds.*

Consider the random variable $W_\ell = \mathsf{mw}(\mathsf{value}(\mathbf{D}_\ell))$, where $\mathbf{D}_\ell$ is a random word in $\mathcal{D}_{in}^\ell$, where all words in $\mathcal{D}_{in}^\ell$ are assumed to be equally likely.

We set

$$\begin{aligned} E &= \frac{C_{0,1}}{C_{1,0}}, \\ V &= \frac{C_{2,0}C_{0,1}^2 - 2C_{1,1}C_{1,0}C_{0,1} + C_{0,2}C_{1,0}^2 + C_{1,0}^2C_{0,1} + C_{1,0}C_{0,1}^2}{C_{1,0}^3}. \end{aligned}$$

Then

$$\begin{aligned} \mathbb{E}(W_n) &= En + O(1), \\ \mathbb{V}(W_n) &= Vn + O(1). \end{aligned}$$

If the variability condition $V \neq 0$ is fulfilled, then the central limit theorem

$$\lim_{\ell \to \infty} \mathbb{P}\left(\frac{W_\ell - \mathbb{E}(W_\ell)}{\sqrt{\mathbb{V}(W_\ell)}} \leq x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt + O\left(\frac{1}{\sqrt{\ell}} \right)$$

holds.

Proof: The asymptotic analysis of the weight of minimal $\mathcal{D}_{out}$-expansions uses the usual techniques for finite transducer automata, cf. Flajolet and Sedgewick (2009, Section V.6). We consider the generating function

$$F(q, u) = \sum_{\mathbf{d} \in \mathcal{D}_{in}^*} \frac{1}{(\#\mathcal{D}_{in})^{\mathsf{length}(\mathbf{d})}} q^{\mathsf{length}(\mathbf{d})} u^{\mathsf{mw}(\mathsf{value}(\mathbf{d}))}.$$

By the definition of $T(q, u)$, we have

$$F(q, u) = (1, 0, \dots, 0)(I - T(q, u))^{-1}(1, \dots, 1)^t$$

where t denotes transposition.

The central limit theorem follows from Flajolet and Sedgewick (2009, Theorem IX.10). Expectation and variance follow from Flajolet and Sedgewick (2009, (37) in Chapter IX). $\square$

8 Results

In this section, we summarise the results that were obtained for specific digit sets $\mathcal{D}_{out}$. Applying Theorem 4 obviously requires explicit computation of $T(q,u)$, which is not always computationally feasible. In those cases, we performed a simple steady state analysis of the underlying Markov chain and obtained the main term of the expectation only.

Theorem 5 *Let $\tau = (\mu + \sqrt{-7})/2$ with $\mu \in \{\pm 1\}$ and W_ℓ as in Theorem 4 for the digits sets specified below.*

1. *For $2 \leq w \leq 5$ and dimension 1, we consider the input digit set $\mathcal{D}_{in} = \{0,1\}$ and the digit set of minimal norm representatives $\mathcal{D}_w$ as $\mathcal{D}_{out}$. We obtain*
$$\mathbb{E}(W_\ell) = e_{1,w}\ell + O(1),$$
where the constants $e_{1,w}$ are given in Table 1.
2. *In dimension $r = 2$, we consider the the input digit set $\mathcal{D}_{in} = \{0,1\}^2$ and the output digit set $\mathcal{D}_{out} = \{0,\pm 1\}^2$. We have*
$$\mathbb{E}(W_\ell) = e_{2,2}\ell + O(1)$$
with
$$e_{2,2} = \frac{144860476952258069960970532866106253274447934570976220749495791797}{311568669055610401810908730777373617652152489224682841359224538895} \approx 0.4649.$$

Table 1 also lists the cardinalities of the carry sets and the vertex sets of the transducer $\mathcal{T}$.

r	w	$\mu=1$		$\mu=-1$		$\mathbb{E}(W_\ell)$
		$\#\mathcal{C}$	$\#V$	$\#\mathcal{C}$	$\#V$	
1	2	12	54	8	27	$1/3\ell + O(1)$
	3	27	324	28	336	$1/4\ell + O(1)$
	4	85	3746	75	3202	$28/141\ell + O(1)$
	5	159	9065	178	10404	$30/181\ell + O(1)$
2	2	144	730121	64	151593	$e_{2,2}\ell + O(1)$

Tab. 1: Results for $\tau = (\mu + \sqrt{-7})/2$ with $\mu \in \{\pm 1\}$.

References

R. Avanzi, C. Heuberger, and H. Prodinger. Redundant τ-adic expansions I: Non-adjacent digit sets and their applications to scalar multiplication. *Des. Codes Cryptogr.*, 58:173–202, 2011. doi:10.1007/s10623-010-9396-6.

P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.

C. Heuberger. Redundant τ-adic expansions II: Non-optimality and chaotic behaviour. *Math. Comput. Sci.*, 3:141–157, 2010. doi:10.1007/s11786-009-0014-9.

C. Heuberger and D. Krenn. Analysis of width-w non-adjacent forms to imaginary quadratic bases. arXiv:1009.0488v2 [math.NT], 2010.

C. Heuberger and J. A. Muir. Minimal weight and colexicographically minimal integer representations. *J. Math. Cryptol.*, 1:297–328, 2007. doi:10.1515/jmc.2007.015.

D. E. Knuth. *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*. Addison-Wesley, third edition, 1998.

N. Koblitz. CM-curves with good cryptographic properties. In *Advances in cryptology—CRYPTO '91 (Santa Barbara, CA, 1991)*, volume 576 of *Lecture Notes in Comput. Sci.*, pages 279–287. Springer, Berlin, 1992. doi:10.1007/3-540-46766-1_22.

B. Phillips and N. Burgess. Minimal weight digit set conversions. *IEEE Trans. Comput.*, 53:666–677, 2004. doi:10.1109/TC.2004.14.

G. W. Reitwiesner. Binary arithmetic. In *Advances in computers*, volume 1, pages 231–308. Academic Press, New York, 1960.

J. A. Solinas. Efficient arithmetic on Koblitz curves. *Des. Codes Cryptogr.*, 19:195–249, 2000. doi:10.1023/A:1008306223194.

J. A. Solinas. Low-weight binary representations for pairs of integers. Technical Report CORR 2001-41, Centre for Applied Cryptographic Research, University of Waterloo, 2001. available at http://www.cacr.math.uwaterloo.ca/techreports/2001/corr2001-41.ps.

E. Straus. Addition chains of vectors (Problem 5125). *Amer. Math. Monthly*, 71:806–808, 1964.

V. Suppakitpaisarn, M. Edahiro, and H. Imai. Optimal average joint Hamming weight and minimal weight conversion of d integers. Cryptology ePrint Archive, Report 2010/300, 2010. http://eprint.iacr.org/.