Dickson Bases and Finite Fields

Ronald C. Mullin[*]
Department of Mathematical Sciences
Florida Atlantic University
Boca Raton, Fl, 33486
and
Center for Advanced Cryptological research
Department of Combinatorica and Optimization
University of Waterloo
Waterloo ON N2L 3G1

Ayan Mahalanobis[†]
Department of Mathematical Sciences
Florida Atlantic University
Boca Raton, Fl, 33486

Abstract

Finite fields have been used for many applications in electronic communications. In the case of extension fields, the nature of computation depends heavily on the choice of basis used to represent the extension over the base field. The most common choices of basis are polynomial bases although optimal normal bases or some variant of these have also been used despite the fact that such bases exist in only a limited set of cases. Building on these, we develop an alternative class of bases that exist for any extension field. We provide hardware models based on the notion of shift registers for computing with respect to such bases, and investigate some of the properties of these models.

[*] rcmullin@uwaterloo.ca
[†] amahalan@fau.edu

1 Introduction

Since this paper deals with finite fields, we adopt the common practice of using the letter p to denote a prime, and q to denote a prime power. Finite fields have been used for various applications in electronic communications. The most commonly used fields are extension fields of $\mathbb{F}_2$ and large prime fields, although other fields have their own advantages in certain applications [17]. In the case of extension fields, the nature of the computation depends heavily on the choice of the basis used to represent the extension over the base field. The most common choices of bases are the polynomial or normal bases, but clearly others are possible. Normal basis representations are examples of "explicit data" representations as defined by H. Lenstra [11], that is, viewing the field $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ in terms of a specified basis $\{b_1, b_2, \cdots, b_n\}$ where the expression of the "cross terms" $b_i b_j$ in terms of the basis are known, and these are used to compute the product of arbitrary members of the field in its representation with respect to the basis. The formulae for the expanded cross terms determine the algebra uniquely without reference to any other details of its structure, and it is this fact that gives rise to the term "explicit data". In a specific situation, an explicit data representation can have its own advantages. For example, normal bases for $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ have the property that computing a q^{th} power of an element in a coordinate representation can be performed by a cyclic shift of the coordinates. Unfortunately, for even "moderately large" values of n, computation in terms of a randomly chosen normal basis tends to be very inefficient. For certain values of n however there are normal bases for which efficient implementations are possible. In particular there is a class of such bases (called optimal normal bases), which are "best possible" in a very well-defined sense. As discussed in section 5, there are two classes of such bases, called type I and type II, the latter exist in only certain fields of characteristic 2.

Another class of explicit data representations has been discussed in the literature. These bases and some associated implementations are described under the name "permuted optimal normal bases" by A. Jurišić [10], "palindromic" bases by I. F. Blake, R. M. Roth, and G. Seroussi [3] and are described without name by B. Sunar and C.K. Koç [18]. These authors differ slightly in their description of this class of representations, but all of them are based on the same re-ordering of type II optimal normal bases. (This brings us to an issue of convention. As is frequently the case, we omit the term "ordered" and use braces for describing ordered bases. When a distinction between ordered and unordered bases is of significance, it will be explicitly stated.)

Unlike explicit data representations, calculations in polynomial bases $\{1, x, x^2, \cdots, x^{n-1}\}$ are defined with respect to a specific irreducible polynomial f of degree n in $\mathbb{F}_q[x]$ and computation is performed modulo this polynomial, that is, the field is viewed as the quotient ring $\mathbb{F}_q[x]/(f)$. Although the polynomial representation can also be viewed in terms of an explicit data representation, by its nature it allows for alternative methods of calculation by choosing different irreducible polynomials of the same degree. One useful aspect of polynomial basis approach is that for a given degree, one can search the set of irreducible polynomials to try to select one with properties useful in a particular circumstance. For example, if one is trying to compute discrete logarithm in $\mathbb{F}_{2^n}$ using Coppersmith's algorithm [5], then a polynomial of the form $x^n + g(x)$ where the degree of $g(x)$ is "small" with respect to n is required. More importantly for our purposes, "sparse" polynomials are often chosen for computation because of efficiency gained. Since no irreducible binomials of degree greater than one exist in $\mathbb{F}_2$, trinomials (or pentanomials or other low weight polynomials if no trinomials of a given degree exist) are frequently used to extend of $\mathbb{F}_2$, and there are tables of such polynomials available for implementers (for example see [2]).

Type II optimal normal bases and their permuted cousins are relatively scarce. The set $\{n$: there is a type II optimal normal basis of degree n over $\mathbb{F}_2\}$ is generally believed to be infinite, but is easily shown to have asymptotic density 0. Motivated by this fact and the work of Jurišić, Blake-Roth-Seroussi, and Sunar-Koç, it is our purpose here to extend the notion of permuted optimal normal bases from sets of elements in a specific limited set of finite fields of characteristic 2 to obtain a calculus for computation in any finite extension field. To that end we introduce a method for calculation in an algebra from which finite fields can be obtained as quotient rings, being aware that by doing so we must perforce abandon many of the properties that made optimal normal bases and their permuted versions so attractive from a computational point of view in the first place. What will be gained is that this alternative approach shares many of the useful properties of polynomial basis representations. In particular it allows one a choice of irreducible elements, and one can look for those with special properties. For example, it permits irreducible analogues of binomials over $\mathbb{F}_2$ for certain degrees of extension, and perhaps somewhat unexpectedly, it permits irreducible monomial in fields of odd characteristic. As for feasibility of implementation, we present a models of hardware multipliers (based on shift registers) for calculating in terms of this alternative representation and discuss a "hardware complexity measure" in this context. In particular, this yields an efficient shift register model for computing in the original permuted optimal normal basis representation.

2 The algebra

Let $\mathbf{B} = \{\mathbf{I}, \beta_1, \beta_2, \cdots, \beta_n, \cdots\}$ be a set of distinct formal symbols, let $\mathfrak{R} = (\mathfrak{R}, +)$ be the free vector space over $\mathbb{F}_q$ on the symbols of $\mathbf{B}$, and denote the null element of the vector space (the empty word) by $\mathbf{0}$. We define a commutative algebra on $\mathfrak{R}$ (also denoted by $\mathfrak{R}$) as follows. Define multiplication on the basis elements by the following rules:

$$\nu\mathbf{0} = \mathbf{0}\nu = \mathbf{0} \quad \text{for all} \quad \nu \in \mathbf{B}, \tag{1}$$

$$\mathbf{I}\nu = \nu\mathbf{I} = \nu \quad \text{for all} \quad \nu \in \mathbf{B}, \tag{2}$$

$$(a\nu_1)\nu_2 = a(\nu_1\nu_2) = \nu_1(a\nu_2) \ \text{ for all } a \in \mathbb{F}_q,\ \nu_1, \nu_2 \in \mathbf{B}, \tag{3}$$

$$\beta_i\beta_j = \beta_{i+j} + \beta_{|i-j|} \quad \text{for all } i, j \geq 0 \tag{4}$$

where by convention, $\beta_0 = 2\mathbf{I}$.
Note that the set $\{a\mathbf{I}|\ a \in \mathbb{F}_q\}$ is an isomorphic copy of $\mathbb{F}_q$ which lies in $\mathfrak{R}$, so without loss of generality, we may write a for $a\mathbf{I}$ when there is no possibility for ambiguity. Thus we can alternatively use the normal practice of using the symbol 1 to represent the multiplicative identity in both the algebra and the ground field unless we wish to emphasize the context.
We extend multiplication linearly on all elements of $\mathfrak{R}$ which also makes the multiplication distribute over addition. Clearly from the definition, multiplication is commutative and 1 is the identity element. To show that multiplication is associative, we note associativity of multiplication of basis elements, the extension by linearity takes care of the rest.
Now

$$(\beta_i\beta_j)\beta_k = \beta_{i+j+k} + \beta_{|i+j-k|} + \beta_{|i-j|+k} + \beta_{||i-j|-k|},$$

and

$$\beta_i(\beta_j\beta_k) = \beta_{i+j+k} + \beta_{|i-|j+k||} + \beta_{i+|j-k|} + \beta_{|i-|j-k||}.$$

One can do a multi-case analysis to show that the two sets of subscripts are equal, but our colleague, L. Klingler, has pointed out that it is simpler to note that both sets of subscripts are equal to the set

$$\{i + j + k,\ \ | - i + j + k|,\ \ |i - j + k|,\ \ |i + j - k|\}.$$

Recall that each non-null element r of $\mathfrak{R}$ has a unique (canonical) representation $r = a_0(r)\mathbf{I} + B(r)$ where $a_0(r)$ is a field element and either $B(r) = 0$ or $B(r) = \sum_{i=0}^{n(r)} a_i(r)\beta_i$ where $a_1(r), a_2(r), \cdots, a_{n(r)}(r)$ are field elements, and $n(r)$ is the largest subscript of a basis element which occurs with a nonzero coefficient in a representation of r. Again by convention, we consider the canonical representation of $\mathbf{0}$ to be $0\mathbf{I}$.

Definition 1. *We define the degree deg(r) of an element r in $\mathfrak{R}$ as follows:*
if $r = 0$ *then* $\deg(r) = -\infty$*;*
if $r \neq 0$ *and* $B(r) = 0$ *then* $\deg(r) = 0$*;*
otherwise deg$(r) = n(r)$*.*

Definition 2. *An $r \in \mathfrak{R}$ be an element of positive degree is said to be irreducible if it cannot be written as the product of two elements, both of positive degree.*

Remark 3. *By convention, we will use the term "irreducible element" to refer to the its "monic" equivalent, that is, for our purposes we can assume without loss of generality that the coefficient of the term of highest degree is 1.*

It is easily shown that $\mathfrak{R}$ is a principal ideal domain. Therefore, if $\mathfrak{r}$ is irreducible of degree n, then K $= \mathfrak{R}/(\mathfrak{r})$ is a field, and it is easily seen that $\{1, \beta_1, \beta_2, \cdots, \beta_{n-1}\}$ is a basis for K over $\mathbb{F}_q$, so K is the field $\mathbb{F}_{q^n}$.

Theorem 4. *$\mathfrak{R}$ is isomorphic to $\mathbb{F}_q[x]$ as an algebra over $\mathbb{F}_q$.*

Proof. Define $\beta = \beta_1$ and inductively $\beta^n = \beta \cdot \beta^{n-1}$ for $n \geq 2$. Trivially $\mathbb{F}_q[\beta] \subseteq \mathfrak{R}$. Using the relations $\beta_0 = 2, \beta_1 = \beta$ and $\beta_{m+1} = \beta \cdot \beta_m - \beta_{m-1}$ for $m \geq 1$, it follows that β_m is a polynomial of degree m in β, so $\mathfrak{R} \subseteq \mathbb{F}_q[\beta]$. The map $\beta \to x$ induces the required isomorphism. •

Clearly the isomorphism respects irreducibility and degrees. In view of the above, each element of the ring $\mathfrak{R}$ can be written as a unique polynomial in the element $\beta = \beta_1$ and can thus be viewed either simply as a ring element or as a polynomial in $\mathbb{F}_q[\beta]$ (or if one prefers, as a polynomial in $\mathbb{F}_q[x]$). The basis polynomials β_n corresponds to the Dickson polynomial $D_n(\beta, 1)$ over $\mathbb{F}_q$ (For more on Dickson polynomials, see [12]). For this reason we will henceforth denote the ring $\mathfrak{R}$, now viewed as an algebra, by $\mathfrak{D}$. Clearly $\mathfrak{D}$ can be viewed as $\mathbb{F}_q[x]$ represented with respect to the basis consisting of the unit element 1 and the set of Dickson polynomials of positive degree. In keeping with our purposes here we continue to view matters in terms of the algebra $\mathfrak{D}$ as described above, except when noted otherwise.

As mentioned in the introduction, "analogues" of binomials in $\mathfrak{D}$ are of interest because they lead to relatively simple reduction formulae as will be discussed in Section 3. We call an element of $\mathfrak{D}$ of the form $\beta_n + b, b \in \mathbb{F}_q$, a D-binomial. Using the division algorithm in $\mathfrak{D}$, an explicit factorization can be used to establish the reducibility of certain families of D-binomials which are of special interest in the following. Unfortunately the authors do not see any method of establishing criteria for showing irreducibility of such polynomials in the formalism being used here although it is possible to do so in certain special cases. Therefore to show irreducibility, we will use the results of S. Gao and G.L. Mullen [8] who provide a set of necessary and sufficient conditions for the irreducibility of the Dickson polynomials $D(\beta, a) + b$. A first such example follows, in which we separate the approaches.

Lemma 5. *For any finite field $\mathbb{F}_q$, if $n > 1$, then $\beta_n + 1$ is irreducible in $\mathfrak{D}$ only if (i) $n = 3^k$ for some $k \in \mathbb{N}$ and (ii) Char$(\mathbb{F}_q) \neq 3$.*

Proof. For part (i), write $n = 3^k m$ for some $k, m \in \mathbb{N}$ and $(m, 3) = 1$. Then m can be written uniquely as $m = 3s + r$, where $s \in \mathbb{N}$ and $r \in \{1, 2\}$.
If $m = 2$, note that

$$\beta_{2\cdot 3^k} + 1 = \beta_{3^k}^2 - 1 = (\beta_{3^k} - 1)(\beta_{3^k} + 1),$$

so $\beta_n + 1$ is reducible.
If $m \geq 4$, by iteration of the identity

$$\beta_{3^t m} + 1 = (\beta_{3^t} + 1)(\beta_{3^t(m-1)} - \beta_{3^t(m-2)}) + \big(\beta_{3^t(m-3)} + 1\big)\,,$$

we obtain the result

$$\begin{aligned}&\beta_{3^k m} + 1 \\ &\quad = \big(\beta_{3^k} + 1)(\beta_{3^k(m-1)} - \beta_{3^k(m-2)} + \beta_{3^k(m-4)} - \cdots \\ &\qquad\qquad \cdots + \beta_{3^k(r+2)} - \beta_{3^k(r+1)}\big) + \beta_{3^k r} + 1\end{aligned}$$

Since $(\beta_{3^k} + 1)$ divides $\beta_{3^k r} + 1$ for $r \in \{1, 2\}$, then clearly $\beta_{3^k} + 1$ divides $\beta_n + 1$.
For part (ii), note for any positive integer n, that

$$\beta_{3n} \ + 1 = \beta_n^3 - 3\beta_n + 1.$$

So if Char$(\mathbb{F}_q) = 3$ then $\beta_{3n} + 1 = (\beta_n + 1)^3$. •

Theorem 6. *If $n > 1$ then $\beta_n + 1$ is irreducible in $\mathfrak{D}$ over $\mathbb{F}_q$ if and only if (i) $n = 3^k$ for some $k \in \mathbb{N}$ and (ii) $q \equiv 2, 4, 5,$ or 7 mod 9.*

Proof. By the previous lemma, we need only consider the case $n = 3^k$ and $q \equiv 1$ or 2 mod 3.
Suppose that $q \equiv \pm 1$ mod 9. Then $q^2 \equiv 1$ mod 9, so F_{q^2} contains a primitive ninth root of unity α. Let $\gamma = \alpha + \alpha^{-1}$. If $q \equiv -1$ mod 9, then $\alpha^q = \alpha^{-1}$, so in either case, $\gamma \in \mathbb{F}_q$. Let $f(x) = (x^9 - 1)/(x^3 - 1) = x^6 + x^3 + 1$ and $g(x) = x^3 - 3x + 1$. Then the zeroes of $f(x)$ in its splitting field are precisely the primitive ninth roots of unity, so

$$\alpha^3 + \alpha^{-3} = -1,$$

$$\gamma^3 = \gamma - 1,$$

that is, γ is a zero of $g(x)$. In fact

$$g(x) = [x - \gamma][x - (\gamma^2 - 2)][x - (\gamma^4 - 4\gamma^2 + 2)]$$

over $\mathbb{F}_q$. But

$$\begin{aligned} \beta_{3n} \ + 1 &= \beta_n^3 - 3\beta_n + 1 \\ &= [\beta_n - \gamma][\beta_n - (\gamma^2 - 2)][\beta_n - (\gamma^4 - 4\gamma^2 + 2)], \end{aligned}$$

so if $q \equiv \pm 1 \bmod 9$, then $\beta_n + 1$ is always reducible.
For the irreducibility of the cases $q \equiv 2, 4, 5$, or $7 \bmod 9$ we appeal to [8]. •

3 Hardware models for computation in $\mathbb{F}_{2^n}$

In this section we present computational models, based on shift registers, of multipliers for computing the product of two elements $\mathfrak{a}$ and $\mathfrak{b}$ in $\mathbb{F}_{q^n}$ considered as the quotient $\mathfrak{D}/(\mathfrak{r})$ where $\mathfrak{D}$ is the Dickson algebra over $\mathbb{F}_q$ defined earlier and $\mathfrak{r}$ is an irreducible element of degree n in $\mathfrak{D}$. The models are suitable for various types of implementation, for example in terms of a gate array or possibly an application specific integrated circuit (ASIC) possibly as a coprocessor.

For sake of simplicity, we describe the models for the most common ground field $\mathbb{F}_2$, although it is easily extended to other ground fields of small charecteristic by introducing field adders, subtracters and multipliers to the appropriate cells, and adjusting for the field characteristic (see, for example, [17] for a discussion of some merits of using elliptic curve cryptosystems over $\mathbb{F}_3$).

Both models assume that $\mathfrak{a} = \sum_{i=0}^{n-1} a_i\beta_i$ and $\mathfrak{b} = \sum_{i=0}^{n-1} b_i\beta_i$ of $\mathbb{F}_{2^n}$ are expressed in terms of their coordinates $(a_0, a_1, \cdots, a_{n-1})$ and $(b_0, b_1, \cdots, b_{n-1})$ respectively. The objective then is to compute the coordinates of the product $\mathfrak{c} = \mathfrak{ab}$ where $\mathfrak{a}$ and $\mathfrak{b}$ are viewed as elements of $\mathbb{F}_{2^n} = \mathfrak{D}/(\mathfrak{r})$ and $\mathfrak{r}$ is the given irreducible element $\mathfrak{r} = \sum_{i=0}^{n} r_i\beta_i$, where $r_n = 1$. Since $\mathfrak{r}$ is irreducible, then $r_0 \neq 0$, and over $\mathbb{F}_2$ this implies that $r_0 = 1$.

3.0.1 Model 1

Let N be a fixed positive integer greater than equal to 2. The first model accepts an integer n satisfying $2 \leq n \leq N$ and an irreducible element $\mathfrak{r}$ of degree n in $\mathfrak{D}$ as well as $\mathfrak{a}$ and $\mathfrak{b}$ as input. Note that one can select any degree of extension n between 2 and N, and this can be changed from time

to time as desired. This model is described in terms of the parameter n, and is a “two pass” type. The registers are to be interpreted as sub-registers of those used for the maximal case $n = N$. The first pass depends only on the choice of n, and the second depends on the choice of $\mathfrak{r}$. The first pass computes the product $\mathfrak{c} = \mathfrak{a}\mathfrak{b}$ in $\mathfrak{D}$. The second pass “reduces” the results of the first pass modulo $\mathfrak{r}$.
This model utilizes shift registers:

$$\begin{aligned} \mathbf{B}^+ &= [\mathbf{b}^+_{2n-2}\mathbf{b}^+_{2n-1}\cdots\mathbf{b}^+_1] \\ \mathbf{B}^- &= [\mathbf{b}^-_{n-2}\mathbf{b}^-_{n-1}\mathbf{b}^-_1\mathbf{b}^-_0\mathbf{b}^-_{-1}\cdots\mathbf{b}^-_{-n+2}] \end{aligned}$$

and stationary registers

$$\begin{aligned} \mathbf{A} &= [\mathbf{a}_{n-1}\mathbf{a}_{n-2}\cdots\mathbf{a}_0] \\ \mathbf{B} &= [\mathbf{b}_{n-1}\mathbf{b}_{n-2}\cdots\mathbf{b}_0] \\ \mathbf{C} &= [\mathbf{c}_{2n-2}\mathbf{c}_{2n-3}\cdots\mathbf{c}_0] \\ \mathbf{R} &= [\mathbf{r}_{n-1}\mathbf{r}_{n-2}\cdots\mathbf{r}_0] \end{aligned}$$

where $\mathbf{x}_i$ denotes the i^{th} stage (or cell) of shift register $\mathbf{X}$, and the contents of the registers are to be shifted from left to right with zeroes being introduced to the leftmost stage of each shift register any time that its contents are shifted, with the contents of the rightmost stage being lost in the process. We will use $\mathbf{x}_i(t)$ to denote the contents of stage $\mathbf{x}_i$ at time t. Note that $\mathbf{x}_i$ denotes a stage of a register, while $\mathbf{x}_i(t)$ is a field element. If time t is not specified, we will denote the contents of the stage $\mathbf{x}_i$ by $\mathrm{cont}(\mathbf{x}_i)$, and abbreviate $(\mathrm{cont}(\mathbf{x}_i), \mathrm{cont}(\mathbf{x}_j), \cdots, \mathrm{cont}(\mathbf{x}_k))$ to $\mathrm{cont}[\mathbf{x}_i\mathbf{x}_j\cdots\mathbf{x}_k]$.
As a notational convenience, we use the assignment operator ‘$\leftarrow$’ in the following fashion: If $\mathbf{x}$ is the stage of a shift register and y is a field element, the expression $\mathbf{x} \leftarrow y$ means, “replace the contents of register $\mathbf{x}$ by y”, or, more simply, assign y to $\mathbf{x}$.

3.1 Pass 1

Initialization:

Load register A The **A** register is initialized by the assignments $\mathbf{a}_i \leftarrow a_i$, for $i = 0, 1, 2, \cdots, n-1$.

Load register B The **B** register is initialized by the assignments $\mathbf{b}_i \leftarrow b_i$, for $i = 0, 1, 2, \cdots, n-1$.

Load register R The **R** register is initialized by the assignments $\mathbf{r}_i \leftarrow r_i$, for $i = 0, 1, 2, \cdots, n-1$.

Initialize register C If $\text{cont}(\mathbf{b}_0) = 1$, then the **C** register is initialized by the assignments $\mathbf{c}_i \leftarrow \text{cont}(\mathbf{a}_i)$, else $\mathbf{c}_i \leftarrow 0$ for $i = 1, 2, \cdots, n-1$.

Initialize the remaining cells by the assignments $\mathbf{a}_i \leftarrow 0$ for $i = n, n+1, \cdots 2n-2$.

Initialize register $\mathbf{B}^+$ Register $\mathbf{B}^+$ is initialized by the assignments $\mathbf{b}_i^+ \leftarrow 0$ for $i = 0, 1, \cdots, n-1$ and $\mathbf{b}_{n+i}^+ \leftarrow \text{cont}(\mathbf{b}_{i+1})$ for $i = 0, 1, 2, \cdots, n-2$.

Initialize register $\mathbf{B}^-$ Register $\mathbf{B}^-$ is initialized by the assignments $\mathbf{b}_{-n+i}^- \leftarrow 0$ for $i = 1, 2, \cdots, n-1$, and $\mathbf{b}_i^- \leftarrow \text{cont}(\mathbf{b}_{n-1-i})$ for $i = 0, 1, 2, \cdots, n-2$.

Initialize register R Register **R** is initialized by the assignments $\mathbf{r}_i \leftarrow r_i$, for $i = 1, 2, \cdots, n-1$.

Processing:

```
If $\text{cont}(\mathbf{a}_0) = 1$ then $\mathbf{c}_i \leftarrow \text{cont}(\mathbf{c}_i) + \mathbf{b}_{n-1+i}^+(0)$ for $i = 1, 2, \cdots, n-1$
    If $\text{cont}(\mathbf{b}_0) = 1$ then $\mathbf{c}_0 \leftarrow 1$
Endif
At time $t$, $t = 0, 1, 2, \cdots, n-2$
If $\text{cont}(\mathbf{a}_{n-1-t}) = 1$ then
    $\begin{cases} \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}_i^+(t) \text{ for } i = 1, 2, \cdots, 2n-2 \\ \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}_i^-(t) + \mathbf{b}_{-i}^-(t) \text{ for } i = 1, 2, \cdots, n-2 \end{cases}$
    Endif
    Shift registers $\mathbf{B}^+$ and $\mathbf{B}^-$.
Endtime
```

Endprocessing

End of pass 1

After executing the above, the coordinates of the product $\mathfrak{ab}$ in $\mathfrak{D}$ are $\text{cont}[\mathbf{c}_{2n-2}\mathbf{c}_{2n-1} \cdots \mathbf{c}_0]$.

The second pass uses the $\mathbf{B}^+$, $\mathbf{B}^-$, **C** and **R** registers. The contents of register **A** can be retained for other purposes, for example to convert the multiplier to an exponentiator.

Pass 2

Initialization

The contents of the **C** and **R** registers are those retained from pass 1.

Initialize register $\mathbf{B}^+$ Register $\mathbf{B}^+$ is initialized by the assignments $\mathbf{b}_i^+ \leftarrow 0$ for $i = 0, 1, \cdots, n-1$, and $\mathbf{b}_{n+i}^+ \leftarrow \text{cont}(\mathbf{r}_{i+1})$ for $i = 0, 1, 2, \cdots, n-2$.

Initialize register $\mathbf{B}^-$ Register $\mathbf{B}^-$ is initialized by the assignments $\mathbf{b}^-_{-n+i} \leftarrow 0$ for $i = 1, 2, \cdots, n-1$, and $\mathbf{b}^-_i \leftarrow \mathrm{cont}(\mathbf{r}_{n-1-i})$ for $i = 0, 1, 2, \cdots, n-2$.

```
Processing:
At time $t$, $t = 0, 1, \cdots, n-3$,
    Shift registers $\mathbf{B}^+$ and $\mathbf{B}^-$
    If $\mathrm{cont}(\mathbf{c}_{2n-2-t}) = 1$, then
        $\begin{cases} \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}^+_i(t) \text{ for } i = 1, 2, \cdots, 2n-2 \\ \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}^-_i(t) + \mathbf{b}^-_{-i}(t) \text{ for } i = 1, 2, \cdots, n-2 \end{cases}$
        (invert) $\mathbf{c}_{t+2} \leftarrow \mathbf{c}_{t+2}(t) + 1$, $\mathbf{c}_{n-2-t} \leftarrow \mathbf{c}_{n-2-t}(t) + 1$
    Endif
Endtime
If $\mathrm{cont}(\mathbf{c}_n) = 1$, then
Shift register $\mathbf{B}$
    $\mathbf{c}_i \leftarrow \mathrm{cont}(\mathbf{c}_i) + \mathrm{cont}(\mathbf{b}^+_i)$ for $i = 1, 2, \cdots, n-1$.
    (invert) $\mathbf{c}_0 \leftarrow \mathrm{cont}(\mathbf{c}_0) + 1$.
Endif
Endprocessing
```

End of Pass 2

At the end of pass 2, the coordinates of $\mathfrak{a}\mathfrak{b}$ mod $\mathfrak{r}$ are $\mathrm{cont}[\mathbf{c}_{n-1}\mathbf{c}_{n-2}\cdots\mathbf{c}_0]$.

3.1.1 Some observations on the implementation architecture of Model 1

Some important considerations in VLSI implementations are *interconnect* and *fan-out* (*fan-in*). Interconnect deals with the physical wiring (and routing thereof) between components of the array. Fan-out refers to the number of connections leading from a particular source. Too large a number of these has the potential to dilute the electrical current from that source. Fan-in is the opposite, too large a number of connections into a source can overload it. Low interconnect, fan-out, and fan-in are clearly desirable features for architectures used in creating VLSI devices. We consider these aspects of the computational core of the above model, namely the $\mathbf{B}^+, \mathbf{B}^-$, and $\mathbf{C}$ registers.

The main computational aspect of both passes of Model 1 lies in the assignments

$$\begin{cases} \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}^+_i(t) \text{ for } i = 1, \cdots, 2n-2 \\ \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}^-_i(t) + \ \mathbf{b}^-_{-i}(t) \text{ for } i = 1, 2, 3, \cdots, n-2 \end{cases}$$

that at first appearance are temporally ordered, but in fact are carried out in parallel by the physical properties of the implementation. In the first assignment statement, at time t the contents of cell $\mathbf{b}_i^+$ are added to those of cell $\mathbf{c}_i$, $i = 1, 2, \cdots, 2n-2$. This can be accomplished by one connection from each cell of the $\mathbf{B}^+$ register to its corresponding cell of the $\mathbf{C}$ register, producing a fan-out of 1 for each cell of the $\mathbf{B}^+$ register.

The second assignment statement states that at the same time, the contents of $\mathbf{b}_i^-$ and $\mathbf{b}_{-i}^-$ are to be added to the contents of $\mathbf{c}_i$, $i = 1, 2, \cdots, n-2$. This can be accomplished by one connection from each pair of cells $\mathbf{b}_i^-$ and $\mathbf{b}_{-i}^-$ to cell $\mathbf{c}_i$, $i = 1, 2, \cdots, n-2$,so the fan-out from each cell of the $\mathbf{B}^-$ register is 1, with the exception of cell $\mathbf{b}_0^-$, which has fan-out 0 (in analogues over characteristics other than 2, this would not be the case for cell $\mathbf{b}_0^-$). Note that none of the cells $\mathbf{c}_{2n-2}, \mathbf{c}_{2n-1}, \cdots, \mathbf{c}_{n-1}$ is connected to the $\mathbf{B}^-$ register.

Now consider the $\mathbf{C}$ register. The assignments

$$\mathbf{c}_{t+2} \leftarrow \mathbf{c}_{t+2}(t) + 1, \;\; \mathbf{c}_{n-2-t} \leftarrow \mathbf{c}_{n-2-t}(t) + 1$$

(indexed by time t, $t = 0, 1, \cdots, n-3$) require the cells $\mathbf{c}_{n-1}, \mathbf{c}_{n-2}, \cdots, \mathbf{c}_1$ to be equipped with an inverter, that is, a gate that maps a zero to a one or a one to a zero, as the case may be. Because the assignments in this case depend on both n and t, it appears that the inverters should be invoked directly by the "controller", that is, the main control unit that contains the clock and that supervises and initiates all the activities of the device. This is particularly true since the parameter n is a variable, as opposed to the fixed value N of the model. It is evident that the cells of the $\mathbf{C}$ register fall into three classes. The cell $\mathbf{c}_0$ contains only a memory unit and an inverter, and has a fan-in of 1, due to the connection between the inverter and the controller. The cells $\mathbf{c}_{n-1}, \mathbf{c}_{n-2}, \cdots, \mathbf{c}_1$ each contain a memory unit, and cascaded XOR gates (mod 2 adders) which are used to accumulate the sum $\mathbf{b}_i^+(t) + \mathbf{b}_i^-(t) + \mathbf{b}_{-i}^-(t)$ in cell $\mathbf{c}_i$ and add the result to that stored in its memory unit. Each such cell is equipped with an inverter which is invoked by the controller at the appropriate time. Thus the fan-in to such a cell is 3. The remaining cells each have fan-in of one, coming from the corresponding cells of the $\mathbf{B}^+$ register. Thus the features of interconnect, fan-out, and fan-in are amenable to implementation.

3.2 Model 2

In certain applications, a specific fixed field $\mathbb{F}_{q^n}$ is used in order to maximize computational efficiency. In such cases the computation of the product of a pair of elements can be calculated in essentially n clock cycles, since having chosen an irreducible element of degree n in $\mathfrak{D}$, one can precompute the reduction of β_{n+i}, $i = 0, 1, \cdots, 2n-2$ and avoid the second pass of the above multiplier by hard-wiring the multiplier so that the putative

contents of $\mathbf{c}_{n+i}$ as described in pass one are added to the appropriate cells of $[\mathbf{c}_{n-1}\mathbf{c}_{n-2}\cdots\mathbf{c}_0]$ "on the fly". In this regard, it is important to recall that between the $\mathbf{B}^+$ and $\mathbf{C}$ registers, stage $\mathbf{b}_i^+$ is connected precisely to stage $\mathbf{c}_i$, $i = 2n-2, 2n-3, \cdots, n$, and that no stage of register $\mathbf{B}^-$ is connected to any stage in $[\mathbf{c}_{2n-2}, \mathbf{c}_{2n-3}, \cdots, \mathbf{c}_n]$. Therefore we can treat any of the stages in $[\mathbf{c}_{2n-2}\mathbf{c}_{2n-3}\cdots\mathbf{c}_n]$ as a "virtual" stages, and at each time t, we can route its incoming datum to be accumulated in the appropriate stages of $[\mathbf{c}_0\mathbf{c}_1\cdots\mathbf{c}_{n-1}]$ (in addition to data coming from the $\mathbf{B}^-$register or other sources). As a result, this model uses a shorter version of the $\mathbf{C}$ register.
We illustrate this by a particularly convenient example. Suppose that β_n+1 is irreducible, then $\beta_n = 1$ and $\beta_{n+i} = \beta_n\beta_i - \beta_{n-i} = \beta_i - \beta_{n-i} \mod (\beta_n+1)$ for $i = 1, 2, \cdots, n-2$. Again we consider the case in which the ground field is $\mathbb{F}_2$.
In this instance, we replace the set of instructions

$$\begin{cases} \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}_i^+(t) \text{ for } i = 1, 2, \cdots, 2n-2 \\ \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}_i^-(t) + \mathbf{b}_{-i}^-(t) \text{ for } i = 1, 2, \cdots, n-2 \end{cases}$$

in pass 1 of the algorithm in Model 1 by

$$\begin{cases} \mathbf{c}_0 \leftarrow \mathbf{c}_0(t) + \mathbf{b}_n^+(t) \\ \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}_i^+(t) + \mathbf{b}_{n+i}^+(t) \text{ for } i = 1, 2, \cdots, n-1 \\ \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}_{2n-i}^+(t) \text{ for } i = 2, 3, \cdots, n-1 \\ \mathbf{c}_i \leftarrow \mathbf{c}_i(t) + \mathbf{b}_i^-(t) + \mathbf{b}_{-i}^-(t) \text{ for } i = 1, 2, \cdots, n-2 \end{cases}$$

At the end of the modified pass 1, the coordinates of $\mathfrak{ab} \mod \mathfrak{r}$ are cont$[\mathbf{c}_{n-1}\mathbf{c}_{n-2}\cdots\mathbf{c}_0]$.

3.3 Some observations on the implementation of architecture of Model 2

As noted, the interconnect for the $\mathbf{B}^-$ register is identical to that of the $\mathbf{B}^-$ register in Model 1. Also the interconnect for cells $\mathbf{b}_{n-1}^+, \mathbf{b}_{n-2}^+, \cdots, \mathbf{b}_1^+$ of the $\mathbf{B}^-$ register remains unchanged between both models. However cells $\mathbf{b}_{2n-2}^+, \mathbf{b}_{2n-3}^+, \cdots, \mathbf{b}_{n+1}^+$ now each have a fan-out of 2, and cell $\mathbf{b}_n^+$ is now connected to cell $\mathbf{c}_0$ for a fan-out of 1. Also each of cells $\mathbf{c}_{n-1}, \mathbf{c}_{n-2}, \cdots, \mathbf{c}_2$ now has connections from two of the cells $\mathbf{b}_{2n-2}^+, \mathbf{b}_{2n-3}^+, \cdots, \mathbf{b}_{n+1}^+$, and these together with its connection from $\mathbf{b}_{n-1}^+, \mathbf{b}_{n-2}^+, \cdots, \mathbf{b}_1^+$ and its two connections from the $\mathbf{B}^-$ register yield a fan-in of 5, while cell $\mathbf{c}_2$ has a fan-out of 5.

3.3.1 A Connection Complexity Measure for Model 2

It is clearly desirable that in the implementation of Model 2, to minimize (in some sense) the number of connections between the $\mathbf{B}^+$ and $\mathbf{C}$ registers. Let an irreducible element $\mathfrak{r}$ of degree $n \geq 2$ in the Dickson algebra $\mathfrak{D}$ as described in Section 2 be given. For any element $\mathfrak{s}$ of $\mathfrak{D}$, let $W(\mathfrak{s})$ denote the number of non-zero terms in the canonical representation of $\mathfrak{s} \mod (\mathfrak{r})$. We define the average interconnect measure AIM($\mathfrak{r}$) as

$$\mathrm{AIM}(\mathfrak{r}) = \frac{1}{n-1} \sum_{s=0}^{n-2} W(\beta_{n+s}).$$

This measure is a normalized expression for the overall contribution of the reduction mod $\mathfrak{r}$ to the interconnect of the model. Additionally, we define the fan-out complexity measure FCM($\mathfrak{r}$) be defined as

$$\mathrm{FCM}(\mathfrak{r}) = \max\{W(\beta_{n+s}) : s = 0, 1, \cdots, n-2\}.$$

For implementation purposes, given a degree n, we suggest that one choose a set S of irreducible elements of this degree with a "small" value for their AIM, say AIM$\leq$ 6. The set should then searched for candidates which minimize the value of FCM. This strategy is chosen since it gives priority to the size of the interconnect. This raises the question "For fixed degree n, what is the minimum value that the AIM can achieve?" A conjectured answer to this question is given in the next section.

4 Irreducibles of low interconnect complexity

In this section we give some examples of irreducibles with "low" values of AIM.

4.1 Example 1 (*Binomial*)

Recall that if the element $\mathfrak{r} = \beta_n + 1$ is irreducible over $\mathbb{F}_q$, then $\beta_n = 1$ and $\beta_{n+i} = \beta_n\beta_i - \beta_{n-i} = \beta_i - \beta_{n-i} \in \mathfrak{D}/(\mathfrak{r})$ for $i = 1, 2, \cdots, n-2$. So for $n \geq 2, \mathrm{AIM}(\mathfrak{r}) = (2n-3)/(n-1)$ and $\mathrm{FCM}(\mathfrak{r}) = 2$. As noted in theorem 6, if $n \geq 2$, then $\beta_n + 1$ is irreducible in $\mathfrak{D}$ over $\mathbb{F}_q$ if and only if (i) $n = 3^k$ for some $k \in \mathbb{N}$ and (ii) $q \equiv 2, 4, 5,$ or 7 mod 9.

4.2 Example 2 (*Monomial*)

Suppose that $\mathfrak{r} = \beta_n$ is irreducible over $\mathbb{F}_q$. Since $\beta_{n+i} + \beta_{n-i} = \beta_n\beta_i$ for $i \geq 0$, then $\beta_n = 0$ and $\beta_{n+i} = -\beta_{n-i} \in \mathfrak{D}/(\mathfrak{r})$ for $i = 1, 2, \cdots, n-2$. In this case AIM($\mathfrak{r}$) $= (n-2)/(n-1)$. Also FCM($\mathfrak{r}$) $= 0$ in the trivial case $n = 2$ and FCM($\mathfrak{r}$) $= 1$ otherwise.

For a given degree, it seems plausible that these irreducibles (when they exist) minimize the value of AIM, and therefore this case is of special interest. We discuss it below.

Lemma 7. *For any finite field $\mathbb{F}_q$, if $n > 1$ then β_n is irreducible in $\mathfrak{D}$ if and only if (i) $n = 2^k$ for some $k \in \mathbb{N}$ and (ii) $q \equiv 3$ or $5 \mod 8$.*

Proof. Using the recurrence $\beta_{n+2} = \beta_{n+1} - \beta_n$ for $n \geq 0$, and the initial conditions $\beta_0 = 2$, $\beta_1 = \beta_1$ a easy induction establishes the fact that if q is even, or if q is odd and n is odd, then β_1 divides β_n. So we need only consider the case that n is odd and q is even.

Suppose that $n = 2^k m$ for $k \geq 1$ and that $(2, m) = 1$. Let $m = 2t + 1$. If $t \geq 1$ then the identity $\beta_{2^k(2t+1)} = \beta_{2^k}\beta_{2^{k+1}t} - \beta_{2^k(2t-1)}$ yields

$$\beta_{2^k(2t+1)} = \beta_{2^k}(\beta_{2^{k+1}t} - \beta_{2^{k+1}(t-1)} + \beta_{2^{k+1}(t-2)} - \cdots + (-1)^t)$$

so β_{2^k} divides $\beta_{2^k(2t+1)}$. Thus if β_n is irreducible, then $n = 2^k$ for some $k \in \mathbb{N}$, and if $n \geq 2$, then n is even.

Suppose that $n = 2s$ and q is odd. Then $\beta_n = \beta_s^2 - 2$ and therefore β_n is reducible if 2 is a quadratic non-residue in $\mathbb{F}_q$. But 2 is a quadratic non-residue in $\mathbb{F}_q$ only if $q \equiv 3$ or $5 \mod 8$. This establishes the necessity of the conditions. Again we appeal to [8] to establish the sufficiency. •

5 The Dickson algebra and optimal normal bases

5.1 Optimal normal bases.

Normal basis for $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ is a basis of the form $N = \{\alpha, \alpha^q, \alpha^{q^2} \cdots, \alpha^{q^{n-1}}\}$. For convenience, let $\alpha_i = \alpha^{q^i}, i = 0, 1, \cdots, n-1$. Normal basis representations for $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ have the interesting property that calculating q^{th} powers can be accomplished by applying a cyclic shift to coordinate vectors with respect to that basis. Since there is no "degree" associated with a normal basis representation N, there is no reduction of the type in the previous section, instead one can use the expansions of the product terms $\alpha_i\alpha_j$ as expressed in N. Hardware architectures for calculating in normal basis representations appear in [1] and [16].

The analogue of AIM in the case of the normal basis architectures mentioned is the *complexity* of N, denoted by $C(N)$ and defined by

$$C(N) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} W(\alpha_i\alpha_j)$$

where $W(\alpha)$ denotes the number of non-zero terms in the representation of α in the basis N.
Since $\alpha_i\alpha_j = \alpha_0\alpha_{j-i}$, if subscripts are calculated $\mod n$ then $C(N)$ reduces to $\sum_{j=1}^{n} W(\alpha_0\alpha_j)$. It is easily shown that for any normal basis N of degree n over $\mathbb{F}_q$ the inequality $C(N) \geq 2n-1$ is satisfied. In the case of equality N is referred to as an *optimal* normal basis.
In [15] two "types" of optimal normal bases are shown to exist as described below.

Type I bases: Type I bases exist in finite fields of any characteristic. If $n+1$ is prime and q is primitive in $\mathbb{Z}^*_{n+1}$ then any primitive $(n+1)$st root of unity $\gamma \in \mathbb{F}_{q^n}$ generates an optimal normal basis (of type I) for $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ under the action of the Galois group of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$.

Type II bases: Type II bases occur only in fields of characteristic 2. Suppose that $q = 2^v$ for some positive integers v and n such that $(v, n) = 1$ and $2n+1$ is prime and further that 2 and -1 generate the multiplicative group $\mathbb{Z}^*_{2n+1}$. Let γ be any primitive $(2n+1)$st root in $\mathbb{F}_{q2^{2n}}$. Then $\alpha = \gamma + \gamma^{-1}$ lies in $\mathbb{F}_{2^n}$ and α generates an optimal normal basis (of type II) for $\mathbb{F}_{2^{nv}}$ over $\mathbb{F}_{2^v}$ under the action of the Galois group of $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$ and as a *set*

$$\{\alpha, \alpha^q, \alpha^{q^2}, \cdots, \alpha^{q^{n-1}}\} = \{\gamma+\gamma^{-1}, \gamma^2+\gamma^{-2}, \gamma^3+\gamma^{-3}, \cdots, \gamma^n+\gamma^{-n}\}.$$

It is the latter ordering that produces a permuted optimal normal basis.
S. Gao [6] proved that up to a scalar equivalence, these are the only optimal normal bases in finite fields. Subsequently Gao and Lenstra [7] extended the result to any finite Galois extension of an arbitrary field. By these results, optimal normal bases are sparse in finite fields, since for the existence of such objects in $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$, a necessary but not sufficient condition is that either $n+1$ or $2n+1$ be prime.

5.2 Dickson Algebras and Optimal Normal Bases

It was a consideration of optimal normal bases and permuted optimal normal bases that gave rise to the investigation of the Dickson algebra point of view of finite fields discussed in this paper. In this section, we consider optimal normal bases from the viewpoint of Dickson algebras.

5.2.1 Dickson polynomials

We require certain properties of the Dickson polynomials $\beta_n = \beta_n(x) = D_n(x,1)$, which are shown below. The following well-known property is one which is frequently used to define Dickson polynomials.

Lemma 8. *Considered as rational functions, $\beta_n(x+x^{-1}) = x^n + x^{-n}$.*

Proof. The proof is immediate by an induction using the facts that

$$\begin{aligned} \beta_0(x) = 2,\ \beta_1(x) = x \ \text{ and} \\ \beta_{n+2}(x) = x(\beta_{n+1}(x)) - \beta_n(x) \quad \text{for} \quad n \geq 0. \end{aligned}$$

•

Let $\sigma_n = \sigma_n(q)$ denote the element $1 + \sum_{i=1}^{n} \beta_i$ in $\mathfrak{D}$ over $\mathbb{F}_q$. It is easily shown by induction that σ_n satisfies the recurrence

$$\sigma_0 = 1, \sigma_1 = \beta_1, \sigma_{n+1} = \beta_1 \sigma_n - \sigma_{n-1} \text{ for } n \geq 1$$

which is the recursive formula for β_n apart from the initial conditions. The following theorem is an interpretation of the discussion in [8] page 131, modified to fit the current viewpoint.

Lemma 9. *Suppose that $n \geq 1$, and suppose that $\mathbb{K}$ is a field that contains a primitive $2n+1^{st}$ root of unity γ. Then*

$$\{\gamma + \gamma^{-1}, \gamma^2 + \gamma^{-2}, \gamma^3 + \gamma^{-3}, \cdots \gamma^n + \gamma^{-n}\}$$

is the set of zeroes of σ_n in $\mathbb{K}$ when σ_n viewed as a polynomial in $\mathbb{K}[x]$.

Proof. It is easily verified that for $n \geq 1$, the identity $(\beta_1 - \beta_0)\sigma_n = \beta_{n+1} - \beta_n$ is valid. Also, since γ is a primitive $(2n+1)^{st}$ root of unity,

$$(\gamma^i)^{n+1} + (\gamma^{-i})^{n+1} = (\gamma^i)^n + (\gamma^{-i})^n \quad \text{for} \quad i = 0, 1, 2, \cdots, n.$$

Therefore from Lemma 8

$$\begin{aligned} &\beta_{n+1}(\gamma^i + \gamma^{-i}) - \beta_n(\gamma^i + \gamma^{-i}) \\ &= \ ((\gamma^i)^{n+1} + (\gamma^{-i})^{n+1}) - ((\gamma^i)^n + (\gamma^{-i})^n) \\ &= \ 0 \text{ for } i = 0, 1, 2, \cdots, n. \end{aligned}$$

and the result follows. •

If σ_n is irreducible, then the next theorem gives a simple characterization of its zeroes in the field $\mathbb{F}_{q^n}$ in its representation as $\mathfrak{D}/(\sigma_n(q))$.

Theorem 10. *If the polynomial $\sigma_n = 1+\sum_{i=1}^{n}\beta_i$ is irreducible in $\mathfrak{D}$ over $\mathbb{F}_q$, then the elements $\beta_1, \beta_2, \cdots, \beta_n$ in $\mathbb{F}_{q^n} = \mathfrak{D}/(\sigma_n)$ are the zeroes, in that representation of $\mathbb{F}_{q^n}$, of σ_n when σ_n is viewed as a polynomial over $\mathbb{F}_q$.*

Proof. Suppose that σ_n is irreducible, and consider the element β_1 in $\mathbb{F}_{q^n} = \mathfrak{D}/(\sigma_n(q))$. If $n = 1$, the result is trivial, so we assume that $n \geq 2$. Let γ and γ^{-1} denote the zeroes of the quadratic polynomial $x^2 + \beta_1 x + 1$ in $\mathbb{F}_{q^{2n}}$. Then $\gamma + \gamma^{-1} = \beta_1 \in \mathbb{F}_{q^n}$. Also

$$\begin{aligned} \beta_1^2 &= \gamma^2 + \gamma^{-2} + 2 \\ &= \gamma^2 + \gamma^{-2} + \beta_0. \end{aligned}$$

So

$$\begin{aligned} \beta_2 &= \beta_1^2 - \beta_0 \\ &= \gamma^2 + \gamma^{-2}. \end{aligned}$$

A straightforward induction shows that

$$\beta_i = \gamma^i + \gamma^{-i}, \ i = 1, 2, \cdots, n.$$

Since $1 + \sum\limits_{i=1}^{n} \beta_i = 0$ in $\mathbb{F}_{q^n}$, we have

$$1 + \sum_{i=1}^{n} (\gamma^i + \gamma^{-i}) = 0.$$

So

$$\sum_{i=0}^{2n} \gamma^i = 0,$$

and therefore γ is a non-trivial $(2n + 1)^{\text{st}}$ root of unity in $\mathbb{F}_{q^{2n}}$.
Since $\{\beta_1, \beta_2, \cdots, \beta_n\}$ is a basis for $\mathbb{F}_{q^n}$ then the elements $\beta_1, \beta_2, \cdots, \beta_n$ are distinct. Therefore all the elements

$$\gamma, \gamma^2, \gamma^3, \cdots, \gamma^n, \gamma^{-1}, \gamma^{-2}, \gamma^{-3}, \cdots, \gamma^{-n}$$

must be distinct. Hence $\langle\gamma\rangle = \mathbb{Z}_{2n+1}^{*}$. So $2n + 1$ is a prime, and (and all non-trivial $(2n+1)^{\text{st}}$ roots of unity) are primitive $2n+1^{\text{st}}$ roots. The result now follows from the the previous lemma. •

Corollary 11. *The set of elements $N = \{\beta_1, \beta_2, \cdots, \beta_n\}$ is a normal basis for $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. Its complexity $C(N)$ is given by*

$$\begin{aligned} C(N) &= \sum_{i=1}^{n} W(\beta_1 \beta_i) \\ &= \quad W(\beta_1^2) + \sum\nolimits_{i=1}^{n} W(\beta_1 \beta_i) \\ &= \quad W(\beta_2 + \beta_0) + 2n - 2 \end{aligned}$$

which is $2n-1$ if q is even, and $3n-2$ otherwise.

If q is even, then (a permutation of) N is a type II optimal normal basis, as can be seen from the fact that $\beta_i = \gamma^i + \gamma^{-i}$, $i = 1, 2, \cdots, n$, which is used to to define such bases [15].
Bases obtained as permutations of the normal bases as above satisfy (i) the Dickson property $\beta_i \beta_j = \beta_{i+j} + \beta_{|i-j|}$, and also the property (ii) $\sum\limits_{i=1}^{n} \beta_i = 1$. The next theorem shows that any basis that satisfies (i) and (ii) can be obtained in this fashion.

Theorem 12. *Suppose that there exists an element $\zeta \in \mathbb{F}_{q^n}$ such that the set $X = \{\zeta_1, \zeta_2, \cdots, \zeta_n\}$ is a basis for $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$, where*

$$\zeta_0 = 2, \;\; \zeta_1 = \zeta, \;\; \zeta_i = \zeta_1 \zeta_{i-1} - \zeta_{i-2}, \;\; i = 2, 3, \cdots, n. \tag{5}$$

Further suppose that the $\sum\limits_{i=1}^{n} \zeta_i = 1$. Then (i) σ_n is irreducible in $\mathfrak{D}$ over $\mathbb{F}_q$ and $\{\zeta_1, \zeta_2, \cdots, \zeta_n\}$ is the set of zeroes of σ_n in $\mathbb{F}_{q^n}$, and (ii) there are precisely n choices for such a ζ since any of the zeroes of σ_n in $\mathbb{F}_{q^n}$ will satisfy the stated condition.

Proof. As in the previous theorem, let γ and γ^{-1} denote the zeroes of the quadratic polynomial $x^2 + \zeta x + 1$ in $\mathbb{F}_{q^{2n}}$. Then $\gamma + \gamma^{-1} = \zeta_1 \in \mathbb{F}_{q^n}$. Again as above,

$$\zeta_i = \gamma^i + \gamma^{-i}, \;\; i = 1, 2, \cdots, n.$$

Since X is a basis, and $\sum\limits_{i=1}^{n} \zeta_i = 1$, then as above,

$$\{\gamma, \gamma^2, \gamma^3, \cdots, \gamma^n, \gamma^{-1}, \gamma^{-2}, \gamma^{-3}, \cdots, \gamma^{-n}\}$$

must constitute the set of $2n$ nontrivial (and primitive) $(2n+1)^{\text{st}}$ roots of unity in $\mathbb{F}_{q^{2n}}$.
Now consider the action of the automorphism $\psi : \psi(y) = y^q$ in the Galois group of $\mathbb{F}_{q^{2n}}$ over $\mathbb{F}_q$ on the set $X = \{\zeta_1, \zeta_2, \cdots \zeta_n\}$. Since ψ maps non-trivial roots of unity to non-trivial roots of unity and since $\psi(\gamma^i + \gamma^{-i}) =$

$((\psi(\gamma))^i + (\psi(\gamma)^{-i})$, then X is closed under ψ. Suppose that Y, the orbit of ζ_1 under $\langle\psi\rangle$ is of length k. Then $a = \sum_{i=0}^{k-1} \psi^i(\zeta_1)$ is the trace of ζ_1 over $\mathbb{F}_q$, so a belongs to $\mathbb{F}_q$. But $a = \sum_{i=1}^{n} a\zeta_i$. since X is a basis and also $Y \subseteq X$, so $Y = X$ by the uniqueness of the representation of the representation of a in the basis X. Thus m_ζ the minimal polynomial of ζ over $\mathbb{F}_q$ is of degree n. Since $\zeta = \gamma + \gamma^{-1}$ where γ is a primitive $(2n+1)^{\mathrm{st}}$ root of unity, then $m_\zeta = \sigma_n$.

It remains to be shown that any of the zeroes of σ_n in $\mathbb{F}_{q^n}$ generates the entire set under the recursive definition 5. Clearly, by the fact that algebraic conjugates cannot be distinguished by algebraic means, this can be readily established by considering the action of the Galois group. However the result also follows easily from the above considerations. Let $\zeta^{'}$ denote any zero of σ_n. Then $\zeta^{'} = \zeta_i$ for some $i, 1 \leq i \leq n$, and $\zeta^{'} = \omega + \omega^{-1}$ where $\omega = \gamma^i$.

Define

$$\zeta_0^{'} = 2, \zeta_1^{'} = \zeta^{'}, \zeta_i^{'} = \zeta_1^{'}\zeta_{i-1}^{'} - \zeta_{i-2}^{'},\ i = 2, 3, \cdots, n. \tag{6}$$

Then

$$\zeta_i^{'} = \omega^i + \omega^{-i},\ i = 1, 2, \cdots, n.$$

Since ω is a primitive $(2n+1)^{\mathrm{st}}$ root of unity, then

$$\{\zeta_1^{'}, \zeta_2^{'}, \cdots, \zeta_n^{'}\} = \{\zeta_1, \zeta_2, \cdots, \zeta_n\} = X.$$

•

In the above theorem the condition $\sum_{i=1}^{n} \zeta_i = 1$ is a "normalizing" condition in the sense that it could be replaced by $\sum_{i=1}^{n} \zeta_i = a$ for any $a \neq 0$ in $\mathbb{F}_q$, with the effect of obtaining a basis differing from the above only by the scalar factor a.

6 Conclusion

We have developed an alternative calculus for computation in extension fields of finite fields, and have presented models for implementing the calculus in terms of shift-register based hardware. In connection of the latter we have introduced the concept of average interconnect measure AIM. This raises the question whether the monomial bases (Example 2, Section 4.2) are best possible in the sense that for fixed $n \geq 2$ and any prime power q, any (monic) irreducible element ρ of degree n in the Dickson algebra over $\mathbb{F}_q$ satisfies the relation $AIM(\rho) \geq (n-2)/(n-1)$, with equality if and only if ρ is a monomial.

A second, similar question arises from the review of optimal normal bases in Section 5. In particular, in Corollary 11, it was pointed out that there exist normal bases of degree n and complexity $3n-2$ in certain finite fields of odd characteristic. Obviously these bases are not best possible in the context sense of optimal normal basis considerations because of the existence of Type I optimal normal bases in fields of odd characteristic, and the fact that such bases have complexity $2n-1$. We therefore ask if bases of complexity $3n-2$ are next to best possible in the sense that in fields of odd characteristic, there are no normal bases N of degree n whose complexity $C(N)$ satisfies $2n-1 < C(n) < 3n-1$.

References

[1] G. Agnew, R. Mullin, I. Onyzchuk, and S. Vanstone. An implementation of a fast public key cryptosystem. *Journal of Cryptology*, 3:63–79, 1991. (U.S. patent #4,747,568).

[2] I.F. Blake, S. Gao, and R.J. Lambert. Construction and distributation problems for irreducible trinomials over finite fields. In *Application of Finite Fields*, pages 19–32. Clemson Press, Oxford, 1996.

[3] I.F. Blake, R.M. Roth, and G. Seroussi. Efficient arithmetic in GF(2^n) through palindromic representation. Technical report, Hewlett-Packard, 1998.

[4] I.F. Blake, R.M. Roth, and G. Seroussi. *Elliptic Curves in Cryptography.* Cambridge University Press, 1999.

[5] D. Coppersmith. Fast evaluation of logarithms in fields of characteristic two. *IEEE Trans. Info. Th.*, (30):587–594, 1984.

[6] S. Gao. The determination of optimal normal bases over finite fields. Technical report, Faculty of Mathematics, University of Waterloo, Research Report CORR 92-01.

[7] S. Gao and H. W. Lenstra. Optimal normal bases. *Designs, Codes and Cryptography*, 2:315–323, 1992.

[8] S. Gao and G .L. Mullen. Dickson polynomials and irreducible polynomials over finite fields. *J. Number Theory*, 49:118–132, 1994.

[9] Dirk Hachenberger. *Finite Fields, Normal Bases and Completely Free Elements.* Kluwer Academic Publishers, 1997.

[10] A. Jurišić. Computing basic operations in a permuted optimal normal basis. Technical report, CERTICOM, 1996.

[11] H.W. Lenstra Jr. Finding isomorphism between fields. *Mathematics of Computation*, 56(193):329–347, January 1991.

[12] R. Lidl, G.L. Mullen, and G. Turnwald. *Dickson Polynomial.* Pitman Monograph and Survey in Pure and Applied Mathematics, 65, 1993.

[13] Rudolf Lidl and Harald Niederreiter. *Introduction to finite fields and their applications.* Cambridge University Press, 1994.

[14] Alfred J. Menezes, editor. *Applications of Finite Fields.* Kluwer Academic Publishers, 1993.

[15] R.C. Mullin, I. Onyszchuk, S.A. Vanstone, and R. Wilson. Optimal normal bases in $GF(p^n)$. *Discrete Applied Math.*, pages 149–161, 1998/1999.

[16] J. Omura and J. Massey. Computational method and apparatus for finite field arithmetic. *U.S. patent #4,587,627*, May 1998.

[17] K. Rubin and A. Silverberg. Supersingular abelian varieties in cryptology. In *Advances in Cryptology:Crypto 2002, proc. 22nd Annual international Cryptography Conference, Santa Barbara, California.*, 2002.

[18] B. Sunar and C.K. Koç. An efficient optimal normal basis type II multiplier. *IEEE Transactions on Computers*, 50:83–87, 2001.