EFFICIENT ARITHMETIC FOR SOME FINITE FIELDS

ILIA TOLI

Abstract. We propose a class of finite fields where the reduction costs one addition. Their size is in the range of interest for ECC. We extend the idea of OEFs to Optimal Extension Rings, that is $\mathbb{Z}/(2^n \pm 1)$ where $2^n \pm 1 = ap$ with p a big prime and a a small cofactor. In 29 cases $a = 3$, in other 11 cases $a = 17$. We propose several classes of finite fields where the reduction costs at most five additions, at least one. We propose a class of fields that optimizes inversion by Fermat Little Theorem.

1. Introduction

The efficient implementation of arithmetic in finite fields is crucial for the high performance of various cryptographic protocols such as those based on the difficulty of the discrete logarithm in finite fields and on elliptic and hyperelliptic curves. Of all arithmetic operations, inversion is the most expensive, modular reduction and multiplication coming next.

In many situations such as RSA [RSA78] the modulus must be generic, otherwise security is affected. Anyway there exist important exceptions to this rule. In the case of elliptic and hyperelliptic curves for example, the field itself is in general not important to security except for its size. This often allows a suitable choice of the field, with inexpensive reduction. As of now, a further exception to this exception are the composed binary fields, that is fields of the form $\mathbb{F}_{2^{mn}}$ [Sma01].

On binary finite fields (finite fields of characteristic 2) the reduction is relatively inexpensive as the field is constructed by choosing the reduction polynomial to be a trinomial or a pentanomial (if no trinomial available) $x^n + x^m + x^k + x^h + 1$ with $m \leq n/2$ [Doc]. Of course, $k = h = 0$ in the case of the trinomial.

For hardware implementations, binary fields are attractive since the operations involve only shifts and bitwise additions modulo 2. The simplicity is also attractive for software implementations on general-purpose processors. However, the field multiplication is essentially a few bits at a time and can be much slower than prime field arithmetic if hardware multiplier is available. On the other hand, the arithmetic in prime fields can be more difficult to implement efficiently, due in part to the propagation of the carry bits.

toli@posso.dm.unipi.it.

$$2^{192} - 2^{64} - 1$$
$$2^{224} - 2^{96} + 1$$
$$2^{256} - 2^{224} + 2^{192} + 2^{96} - 1$$
$$2^{384} - 2^{128} - 2^{96} + 2^{32} - 1$$
$$2^{521} - 1$$

FIGURE 1. NIST recommended primes.

Much research has been dedicated to constructing the fields in use for ECC. While on binary fields the main question is how to construct a field in order to increase efficiency, on the prime fields the main question is which fields to choose in order to have efficient arithmetic.

There have been many interesting answers to this question, like Generalized Mersenne primes [Sol99], pseudo-Mersenne primes [Bai]. Anyway, there is need to increase the number of fields with good arithmetics and the field is open to research. This need is particularly felt in probabilistic algorithms. One may have to start an algorithm with a given prime (the most efficient available in the needed range). At a certain point it may fail to meet certain other criteria therefore ought to be changed. This is the case, for example, in the generation of good elliptic curves over prime fields [KSZ03, SSK01]. This is one of motivations of this work.

2. Fields with Good Arithmetic

In the following we will need this definition from [HMV04].

Definition 2.1. *A* non-adjacent form *(NAF) of a positive integer k is an expression $k = \sum_{i=0}^{l-1} k_i 2^i$ where $k_i \in \{0, 1, -1\}$, $k_{l-1} \neq 0$ and no two consecutive digits k_i are nonzero. The length of NAF is l.*

In this paper we will denote by wNAF(k) the number of nonzero digits of NAF(k). It stands for *Hamming weight of NAF(k).*

Theorem 2.2 (properties of NAF, [HMV04])**.** *Let k be a positive integer.*

1. *k has a unique NAF denoted by NAF(k).*
2. *NAF(k) has the fewest nonzero digits of any signed digit representation of k.*
3. *The length of NAF(k) is at most one more than the length of the binary representation of k.*
4. *If $NAF(k) = l$, then $2^l/3 < k < 2^{l+1}/3$.*
5. *The average density of nonzero digits among all NAFs of length l is $\approx 1/3$.*

2.1. **Optimal Extension Fields.** Optimal Extension Fields (OEF) were first introduced by Bailey and Paar [Bai]. The general idea in OEFs is to select p, m and the reduction polynomial in the construction of $\mathbb{F}_{p^m}$ to more closely match the underlying hardware characteristics. In particular, the value of p may be selected to fit in a single word, simplifying the handling of the carry.

Definition 2.3. *An* Optimal Extension Field (OEF) *is a finite field* $\mathbb{F}_{p^m}$ *such that:*

(1) $p = 2^n - c$ *for some integers* n *and* c *with* $log_2|c| \leq n/2$ *and*
(2) *an irreducible polynomial* $f(z) = z^m - \omega$ *in* $\mathbb{F}_p[z]$ *exists.*

If $c = \pm 1$, then the OEF is said to be of *Type I*. p is a Mersenne prime if $c = 1$ and a Fermat prime if $c = -1$. If $\omega = 2$, the OEF is said to be of *Type II*. Here we will denote by *Type I+II* the fields that are of Type I and II at the same time.

The primes $p = 2^n - c$ such that $log_2|c| \leq n/2$ are called *pseudo-Mersenne primes*.

The condition $log_2|c| \leq n/2$ assures that reduction will finish within two rounds. Therefore reduction requires two multiplications by c that is in average as expensive as a single multiplication within the field.

In Type I OEFs reduction takes only one addition/subtraction in the subfield $\mathbb{F}_p$.

The reduction over the extension field for Type II OEFs, is very simple also. It requires only $m - 1$ left shifts of one bit (of the coefficients of monomials of degree bigger than $m - 1$) and $m - 1$ additions, all in the base field.

Extending the definition of Type II to $\omega \in \{-2, -1, 2\}$, we have many more fields of Type I+II while the arithmetic only becomes easier. We have observed that very frequently all three $x^2 + 1$ and $x^2 \pm 2$ are irreducible. A counterexample to the second polynomial is the field $\mathbb{F}_{431}$: $x^2 - 2 = (x + 243)(x + 188)$.

As said above, to perform a reduction in a generic OEF $\mathbb{F}_{p^m}$ is at most (and generally) as expensive as a multiplication. In many situations it should be preferable to avoid this one more multiplication. An example of such a situation is ECC, where we transform the addition and doubling formula of a point in order to save a multiplication. For each multiplication there occurs also a reduction. Therefore, avoiding the multiplication within reduction should be equivalent to halving the number of general multiplications.

It is easily seen that the reduction on $\mathbb{F}_p$ heavily depends on $w = wNAF(c)$. Indeed, by a counting argument we observe that for $w =$

$wNAF(c) \geq 3$ we need $3w - 4$ additions/subtractions to perform one reduction.

Of course, one can use other reduction algorithms (Barrett [Bar87], Montgomery [Mon85]) which are all quadratic on n and independent on $NAF(c)$.

Incidentally, all NIST-recomended primes [NIS] 2 which were chosen like GM-primes are also pseudo-Mersenne primes.

3. Optimal Extension Rings

By extension we define Optimal Extension Rings as follows.

Definition 3.1. *An* Optimal Extension Ring (OER) *is a ring* $\mathbb{Z}/d$ *such that:*

(1) $d = 2^n - c$ *for some integers* n *and* c *with* $log_2|c| \leq n/2$ *and*
(2) $d = ap$ *for some prime* p *with* $gcd(a, p) = 1$ *and an irreducible polynomial* $f(z) = z^m - \omega$ *in* $\mathbb{F}_p[z]$ *exists.*

By extension we define Type I, Type II and Type I+II OERs.

The biggest problem with OEFs of Type I or Type I+II is that they are very few. The known Mersenne primes are $2^n - 1$ for $n \in \{1, 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213, 19937, 21701, 23209, 44497, 86243, 110503, 132049, 216091, 756839, 859433, 1257787, 1398269, 2976221, 3021377, 6972593, 13466917, 20996011, 24036583, 25964951, 30402457\}$ [Res].

The only known Fermat primes are $2^n + 1$ for $n \in \{1, 2, 4, 8, 16\}$. Good GM-primes are even rarer. Furthermore, they are not as good as the best OEFs.

The range of interest at present is more than 160 bits and anyway about 200 bits []. There are no primes of this size in the lists of Mersenne and Fermat primes.

Therefore it is natural to move to the "next best thing". For this purpose we may consider the following.

(1) Taking $wNAF(c) = 3$ or bigger if needed,
(2) Optimal Extension Rings.

Consider the first option, primes with $wNAF(c) = 3$. We have checked all primes of the form $2^n \pm 2^k \pm 1$ with $k \leq n/2$ for all $n \leq 3409$. They exist for all $n \neq 3296$ and seem to be well distributed. We have counted up to 21 primes for one n and there are seven for each n in average for $n = 3296$ and $1756 \leq k \leq 3042$.

Taking k the smallest available for fixed n allows for further cost reductions.

The primes of the form $2^n - 2^k - 1$ recur about 40% more frequently than those of the form $2^n + 2^k + 1$ do. For the primes with $gcd(n, k) \neq 1$, the gcd is generally very small; 2 or so. Very big gcd recur also. There can be observed some more patterns, but that falls beyond the scope of this paper. What is most important, there are plenty of these primes in the range of interest for ECCs.

The second approach, OER, is that of considering the rings $\mathbb{Z}/(2^n \pm 1)$ where $2^n \pm 1 = ap$ with p prime and a some tiny factor. All operations should be done in the ring $\mathbb{Z}/(2^n \pm 1)$ and only once in the very end will we have to reduce modulo p in order to mirror values into the field $\mathbb{F}_p$. As p is prime and a small, $gcd(a, p) = 1$ and (from Chinese Remainder Theorem) the ring $\mathbb{Z}/(2^n \pm 1)$ contains the field $\mathbb{F}_p$.

Indeed, still better can be done. We don't need to reduce modulo p at all. We have numbers modulo ap, reduce them all modulo a (which is far easier, as a is tiny) and then find the needed field elements applying Chinese Remainder Theorem.

The most important application of these arithmetic efficiencies is in ECC. In a two-party communication receiver-sender, the sender doesn't need at all to mirror his results into the field $\mathbb{F}_p$. He can take the original values (say, cleartext) from the field $\mathbb{F}_p$, operate with them as elements of the ring $\mathbb{Z}/(2^n \pm 1)$ and send to the receiver a set of elements from $\mathbb{Z}/(2^n \pm 1)$. He treats them as elements of that ring, and only in the end of all his transformations will he have to reduce into the field $\mathbb{F}_p$. By doing so the sender is saved his final reduction.

There are probabilistic algorithms that demand generation of prime fields, for example in elliptic curve generation. If the field in use is not good for some reason at some step of the algorithm, generate another field. This renders OERs furthermore interesting. They add to the list of finite fields with good arithmetic; they are not optimizations of fields already in use, as it is often the case in characteristic 2. If handled with the known generic reduction algorithms these fields should be rather expensive.

The only caveat is that we will deal with n bits in order to perform arithmetic on $log_2\ p$ bit fields. However, fortunately there exist some values within the range of interest such that $log_2\ a = 2$ and $log_2\ p = n - 2$ and some others with a little bit bigger $log_2\ a$.

Particularly interesting are Fermat numbers (numbers of the form 2^n+1) that factor into $3p$. We have checked them exhaustively for $n \leq 268$ and selectively searched for them for the other n. This is the case for $n \in \{3$, 5, 7, 11, 13, 17, 19, 23, 31, 43, 61, 79, 101, 127, 167, 191, 199, 313, 347,

701, 1709, 2617, 3539, 5807, 10501, 10691, 11279, 12391, 14479}. It can be easily proven that all Fermat numbers for n odd are divisible by 3.

Fermat numbers factor into $17p$, p prime for $n \in \{12, 20, 28, 92, 148, 356, 596, 692, 1004, 1228, 1268\}$. We searched for them through $n \leq 10000$, $n = 4q$, q prime.

4. Performance Comparisons

The highest ranking fields with efficient reduction are:

(1) Prime OEFs of Type I,
(2) OEFs of Type I+II.

They are at par; reduction costs only one addition.

There exists only one prime OEF of Type I within the present range of interest for ECC: $GF(2^{521} - 1)$. Even that is too big for many applications. The next candidates are $GF(2^{607} - 1)$, $GF(2^{1279} - 1)$, $GF(2^{2203} - 1)$, $GF(2^{2281} - 1)$ [Res].

For the Hyperelliptic Curves Cryptography few more smaller prime fields should be interesting.

There exist some OEFs of Type I+II.

Let us consider the following fields:

(1) Prime OEFs of wNAF 3,
(2) Type II OEFs of wNAF 3.

In this case the reduction requires 4 additions in general. The sum should in general need to be reduced, therefore need two further additions. Therefore it is about 6 times more expensive than in the fields of the first category above. If optimized for particular (rare) choices of c, it requires only 3 additions, whence about 5 more expensive than fields above. It is obvious that with the growth of the number of additions the reduction modulo p of the sum becomes more complex also. In the Figure in page 2 is given the list of very well optimized primes recommended by NIST [NIS] and their respective reduction complexities.

There are plenty enough prime OEFs of wNAF 3. We have considered this argument in the previous Section.

In the middle of these two categories stand:

(1) Prime OERs of Type I,
(2) OERs of Type I+II.

They have the same reduction cost as the first category: only 1 addition. If they are of the form $3p$ they need only 2 more additions to perform the final reduction. Anyway, the final reduction is never a big burden.

On the other hand, the multiplication (the next most expensive operation) is done on longer integers. In the $3p$ case, the integers are two bits longer.

In the circumstance that we need more primes than OEFs of wNAF 3 can be, we can use OERs of wNAF 3. They seem to be furthermore plentiful, even with very small cofactor.

5. Other Issues

5.1. **The Field Polynomial.** The same arguments discussed above about the characteristic hold for the choice of the irreducible polynomial for constructing a nonprime field. If there is no available irreducible binomial $x^n - \omega$ with $\omega \in \{-2, -1, 2\}$, we may try $\omega = \pm 2^k$ with $|\omega| < p$, where p is the characteristic of the base field. If we find none such binomial, "next best thing" is $\omega = \pm 2^{k_1} \pm 2^{k_2}$ and so on with higher wNAF if needed.

5.2. **Inversion.** Inversion is the most costly field operation. In ECC it is usually traded for multiplications via various types of homogeneous coordinates. Anyway it is unavoidable in many circumstances. For example, in order to dehomogenize the plaintext after decryption we need one inversion. It need not be very optimized, but we need one.

The two main algorithms for performing inversion in a finite field $\mathbb{F}_q$ are by means of Euclidean algorithm and by the formula $a^{-1} = a^{q-2}$ for $a \in \mathbb{F}_q$ (Fermat's Little Theorem). Euclidean algorithm is nearly quadratic in complexity and raising in power $q-2$ is cubic. Anyway, in many situations where we do just a few inversions and for hardware implementation raising in power $q - 2$ turns convenient.

The raising in power $q-2$ is done in general by the square-and-multiply algorithm. Addition chains are used to optimize the number of squarings and multiplications needed. Much depends also on the shape of $q - 2$. In characteristic 2 it is of the worst possible shape: $2^n - 2 = 111 \ldots 110$, with $n - 1$ ones.

So it should be very desirable to have q of a more convenient shape. One thing that may be done is to choose $q - 2$ of the lowest possible Hamming weight. Thus, in calculating a^{q-2} we will have to perform only one or two multiplications, and then only squarings.

For $0 \leq n \leq 10000$, $2^n + 3$ ($q - 2$ of Hamming weight two) is prime for $n \in \{1, 2, 3, 4, 6, 7, 12, 15, 16, 18, 28, 30, 55, 67, 84, 228, 390, 784, 1110, 1704, 2008, 2139, 2191, 2367, 2370, 4002, 4060, 4062, 4552, 5547, 8739\}$.

Choices with $q - 2$ of Hamming weight three are quite plentiful; about 1.3 for each n. That is, primes of the form $2^n + 2^k + 3$. A general feature

observable is that k is generally big and it grows with n. Anyway, no matter how big n is, there do appear some primes with very small k also.

Wishing to have both reduction and inversion as efficient as possible, we should choose primes of the form $2^n + 2^k \pm 1$ for k very small.

Unfortunately we cannot resort (in the fashion of OERs) to numbers of the form $2^n + 3 = ap$ with p a big prime and a a small cofactor. Indeed, raising in power will have again to be done in $p - 2$ and generally $p - 2$ is very dense.

References

[Bai] Daniel Bailey. Computation in Optimal Extension Fields. `http://www.crypto.ruhr-uni-bochum.de/imperia/md/content/texte/theses/ms_bailey.pdf`. Master thesis, Worcester Polytechnic University, MA, USA.

[Bar87] Paul Barrett. Implementing the Rivest Shamir and Adleman public key encryption algorithm on a standard digital signal processor. In *Proceedings on Advances in cryptology—CRYPTO '86*, pages 311–323, London, UK, 1987. Springer-Verlag.

[Doc] Christophe Doche. Redundant Trinomials for Finite Fields of Characteristic 2. Preprint.

[HMV04] Darrel Hankerson, Alfred Menezes, and Scott Vanstone. *Guide to Elliptic Curve Cryptography*. Springer, 2004.

[KSZ03] Elisavet Konstantinou, Yiannis C. Stamatiou, and Christos D. Zaroliagis. On the Efficient Generation of Elliptic Curves over Prime Fields. In *CHES '02: Revised Papers from the 4th International Workshop on Cryptographic Hardware and Embedded Systems*, pages 333–348, London, UK, 2003. Springer-Verlag.

[Mon85] Peter L. Montgomery. Modular multiplication without trial division. *Mathematics of Computation*, 44:519–521, 1985.

[NIS] NIST. `http://www.itl.nist.gov/fipspubs/fip186.htm`.

[Res] TSM Resources. `http://www.tsm-resources.com/alists/mers.html`.

[RSA78] Ron Rivest, Adi Shamir, and Leonard M. Adleman. A method for obtaining digital signatures and public key cryptosystems. In *Communications of the ACM*, 21, pages 120 – 126, 1978. `http://theory.lcs.mit.edu/~rivest/rsapaper.pdf`.

[Sma01] Nigel P. Smart. How secure are elliptic curves over composite extension fields? *Lecture Notes in Computer Science*, 2045:30–??, 2001. `citeseer.ist.psu.edu/smart01how.html`.

[Sol99] J. Solinas. Generalized mersenne numbers, 1999. `http://www.cacr.math.uwaterloo.ca` and `http://citeseer.ist.psu.edu/solinas99generalized.html`.

[SSK01] E. Savaş, T. A. Schmidt, and Çetin Kaya Koç. Generating Elliptic Curves of Prime Order. *Lecture Notes in Computer Science*, 2162:142–??, 2001. `http://citeseer.ist.psu.edu/603197.html`.