PRIME ORDER PRIMITIVE SUBGROUPS IN TORUS-BASED CRYPTOGRAPHY

JASON E. GOWER

ABSTRACT. We use the Bateman-Horn conjecture to study the order of the set of $\mathbb{F}_q$-rational points of primitive subgroups that arise in torus-based cryptography. We provide computational evidence to support the heuristics and make some suggestions regarding parameter selection for torus-based cryptography.

1. BACKGROUND

1.1. **Algebraic Tori and Primitive Subgroups.** Let L/K be a finite and separable field extension with $[L : K] = n$. Let $\mathbb{G}_{\mathrm{m}}$ be the multiplicative algebraic group defined by the following property: Over any field F, the set of F-rational points of $\mathbb{G}_{\mathrm{m}}$, denoted $\mathbb{G}_{\mathrm{m}}(F)$, is the multiplicative group $F^{\times}$ of nonzero elements of the field F. The Weil restriction of scalars of $\mathbb{G}_{\mathrm{m}}$ from L down to K, denoted $\mathrm{Res}_{L/K}\mathbb{G}_{\mathrm{m}}$, enjoys the following property:

$$(\mathrm{Res}_{L/K}\mathbb{G}_{\mathrm{m}})(K) \cong \mathbb{G}_{\mathrm{m}}(L) = L^{\times},$$

where the equality comes from the definition of $\mathbb{G}_{\mathrm{m}}$. In other words the set of K-rational points of $\mathrm{Res}_{L/K}\mathbb{G}_{\mathrm{m}}$ is isomorphic to $L^{\times}$. The algebraic group $\mathrm{Res}_{L/K}\mathbb{G}_{\mathrm{m}}$ is a non-trivial example of an algebraic torus defined over K; that is, an algebraic group T defined over K that over some finite extension field is isomorphic to $(\mathbb{G}_{\mathrm{m}})^d$, where d is the dimension of T.

For any field F with $K \subset F \subsetneq L$, let $N_{L/F} : L \longrightarrow F$ denote the usual norm map defined by $N_{L/F}(\alpha) = \prod_{\sigma \in \mathrm{Gal}(L/F)} \sigma(\alpha)$. Associated with each norm map $N_{L/F}$ there exists a map $\mathcal{N}_{L/F} : \mathrm{Res}_{L/K}\mathbb{G}_{\mathrm{m}} \longrightarrow \mathrm{Res}_{F/K}\mathbb{G}_{\mathrm{m}}$ such that the following diagram commutes.

$$\begin{array}{ccc} (\mathrm{Res}_{L/K}\mathbb{G}_{\mathrm{m}})(K) & \xrightarrow{\ \mathcal{N}_{L/F}\ } & (\mathrm{Res}_{F/K}\mathbb{G}_{\mathrm{m}})(K) \\ \cong \big\downarrow & & \big\downarrow \cong \\ L^{\times} & \xrightarrow{\ N_{L/F}\ } & F^{\times} \end{array}$$

Finally, we define the *primitive subgroup* of the algebraic group $\mathrm{Res}_{L/K}\mathbb{G}_{\mathrm{m}}$ as the intersection

$$T_n = \bigcap_{K \subset F \subsetneq L} \ker \mathcal{N}_{L/F}.$$

Date: December 6, 2006.

Key words and phrases. finite field, primitive subgroup, torus-based cryptography, Bateman-Horn conjecture.

It follows that the K-rational points of T_n can be characterized as follows:

$$T_n(K) \cong \{\alpha \in L^\times \mid N_{L/F}(\alpha) = 1, \text{for all } F \text{ with } K \subset F \subsetneq L\}.$$

It can be shown that T_n is a $\varphi(n)$-dimensional algebraic torus, where φ is the Euler totient function. See [9] for more about the Weil restriction of scalars, algebraic tori, and their related properties.

In this paper, we will be exclusively interested in the case where K is the finite field $\mathbb{F}_q$ with q elements, where q is a prime power p^r for some prime p and positive integer r. Since L is a degree n extension of K, it follows that $L = \mathbb{F}_{q^n}$. From these choices we see that

$$T_n(\mathbb{F}_q) \cong \{\alpha \in \mathbb{F}_{q^n}^\times \mid N_{\mathbb{F}_{q^n}/\mathbb{F}_{q^d}}(\alpha) = 1, \text{for all divisors } d \text{ of } n \text{ with } d \neq n\}.$$

1.2. **Primitive Subgroups in Cryptography.** The group $T_n(\mathbb{F}_q)$ has recently been studied for its usefulness in cryptographic schemes such as Diffie-Hellman key exchange and ElGamal encryption and authentication where the underlying discrete logarithm problem is assumed to be difficult. The following theorem, proved in [2, 5], lists some properties of $T_n(\mathbb{F}_q)$ that make it attractive for use in cryptography.

Theorem 1. *If $\alpha \in T_n(\mathbb{F}_q)$ is an element of prime order not dividing n, then α does not lie in a proper subfield of $\mathbb{F}_{q^n}$. Moreover, $T_n(\mathbb{F}_q) \cong G_{q,n}$, where*

$$G_{q,n} = \{\alpha \in \mathbb{F}_{q^n}^\times \mid \alpha^{\Phi_n(q)} = 1\}$$

and $\Phi_n(x)$ is the n^{th} cyclotomic polynomial in the variable x.

This theorem states that $T_n(\mathbb{F}_q)$ is isomorphic to the cyclic subgroup of $\mathbb{F}_{q^n}^\times$ of order $\Phi_n(q)$, a group which is not contained in any proper subfield of $\mathbb{F}_{q^n}$. As such, $T_n(\mathbb{F}_q)$ is isomorphic to the "cryptographically strongest" subgroup of $\mathbb{F}_{q^n}^\times$ in the sense that an attacker will not be able to successfully use an index calculus algorithm for computing discrete logarithms in $\mathbb{F}_{q^d}$ if d is a proper divisor of n. If we choose q such that $\log_2 q^n \approx 1024$ (for 1024-bit RSA security) and $\Phi_n(q)$ is divisible by a prime with at least 160 bits (so as to thwart "square root" attacks such as the Pollard Rho algorithm for computing discrete logarithms), then it would seem that $T_n(\mathbb{F}_q)$ is a group that can be used to build secure cryptographic schemes.

In addition to the security-related properties, there is another property that makes $T_n(\mathbb{F}_q)$ a particularly attractive group to work with for certain choices of n, as described in the following theorem proved in [9].

Theorem 2. *The torus T_n is rational if n is a prime power or the product of two prime powers.*

This theorem says that since T_n is $\varphi(n)$-dimensional, if n is a prime power or a product of two prime powers, then T_n is birationally isomorphic to $\mathbb{A}^{\varphi(n)}$. As such, most of $T_n(\mathbb{F}_q)$ can be compactly represented with $\varphi(n)$-tuples of elements of $\mathbb{F}_q$, as opposed to the n-tuples of elements of $\mathbb{F}_q$ that are usually needed to represent elements of $\mathbb{F}_{q^n}$. It follows that if we use $T_n(\mathbb{F}_q)$ instead of $\mathbb{F}_{q^n}$, data transmissions will be more efficient by a factor of $n/\varphi(n)$.

Clearly we would like to choose n so as to maximize $n/\varphi(n)$. Since we may as well take n to be squarefree, this leaves us with only two optimal choices for which we are guaranteed that T_n is rational; namely, $n = 2, 6$. Indeed, cryptographic schemes have been built in these groups; see [7, 4, 5]. It has been conjectured in [9] that T_n is rational for all positive integers n. If this were so, then other interesting cases would be $n = 30, 210$, etc.

Though we do not know if T_n is rational for all positive integers n, we do have to following result, also from [9].

Theorem 3. *The torus T_n is stably rational for all positive integers n.*

In other words, for each positive integer n, there exists some positive integer c such that $T_n \times \mathbb{A}^c$ is rational. Currently the best known constructions use $c = 2$ when $n = 30$, and $c = 22$ when $n = 210$; see [8].

1.3. **Order of $T_n(\mathbb{F}_q)$.** Recall that $\#T_n(\mathbb{F}_q) = \Phi_n(q)$. If $\Phi_n(q)$ is divisible only by small primes, then it will be easy to compute discrete logarithms in $T_n(\mathbb{F}_q)$ using the Chinese Remainder Theorem-based Pohlig-Hellman algorithm. To avoid this attack and the Pollard Rho attack mentioned previously, we will need to choose n and q so that $\Phi_n(q)$ is divisible by a prime with at least 160 bits. Recently an index calculus attack that works directly in $T_n(\mathbb{F}_q)$ for $n = 2, 6$ has been proposed [3], though it is not applicable in the case where q is a prime. Therefore we will mainly be interested in the case where both q and $\Phi_n(q)$ are sufficiently large primes, though we will begin with the more general case of $q = p^r$ for prime p and positive integer r.

Let N be a positive integer and define

$$P_{r,n}(N) = \#\{m \mid 2 \leq m \leq N,\ m \text{ and } \Phi_n(m^r) \text{ are primes of } \mathbb{Z}\}.$$

This quantity counts the number of primes $p \in [2, N]$ such that $\Phi_n(q)$ is also prime, where $q = p^r$ and r and n are fixed. Each of these leads to a potentially cryptographically useful group $T_n(\mathbb{F}_q)$, though, again, we are ultimately interested in the case $r = 1$. In the sequel we will study the asymptotic behavior of $P_{r,n}(N)$ as $N \longrightarrow \infty$, restricting our attention to the case where n is the squarefree product of the first few primes. We will provide supporting computational evidence to go along with the heuristics, and also recommend some choices of n and q that result in schemes that provide security against all known attacks.

2. The Bateman-Horn Conjecture

2.1. **Statement of the Conjecture.** We begin our study of $P_{r,n}(N)$ by stating a conjecture of Bateman and Horn from [1]. Let $f_1, \ldots, f_k$ be distinct, irreducible polynomials in $\mathbb{Z}[x]$ with positive leading coefficients. Define $f = \prod_{i=1}^k f_i$ and

$$\mathcal{S}(f) = \{f(m) \mid m \in \mathbb{Z}\}, \tag{2.1}$$

and further suppose that no prime divides every element of $\mathcal{S}(f)$. For each positive integer N define

$$Q(f_1, \ldots, f_k; N) = \#\{m \mid 2 \leq m \leq N,\ f_1(m), \ldots, f_k(m) \text{ are all primes of } \mathbb{Z}\}.$$

The following conjecture describes the asymptotic behavior of $Q(f_1, \ldots, f_k; N)$ as $N \longrightarrow \infty$.

Bateman-Horn Conjecture. *Let $f_1, \ldots, f_k$ and f be as above, $d_i = \deg f_i$, $\mathcal{P}$ be the set of primes of $\mathbb{Z}$, $\omega(p) = \#\{x \mid 1 \leq x \leq p, f(x) \equiv 0 \pmod{p}\}$, and define*

$$C(f_1, \ldots, f_k) = \prod_{p \in \mathcal{P}} \left(1 - \frac{\omega(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}. \tag{2.2}$$

Then

$$Q(f_1, \ldots, f_k; N) \sim \frac{C(f_1, \ldots, f_k)}{d_1 \cdots d_k} \int_2^N (\ln x)^{-k}\, dx, \quad \text{as } N \longrightarrow \infty. \tag{2.3}$$

Note that the Bateman-Horn conjecture reduces to the Prime Number Theorem if $k = 1$, $f_1 = x$, and to Dirichlet's Theorem on primes in an arithmetic progression if $k = 1$, $f_1 = a + bx$, and $\gcd(a, b) = 1$. If $k = 2$, $f_1 = x$, and $f_2 = x + 2$, then we have the Twin Prime Conjecture. See [1] for a heuristic argument supporting (2.3) and a proof that the infinite product in (2.2) converges. Though the supporting computational evidence is overwhelming, there is unfortunately no proof of the Bateman-Horn conjecture. Nonetheless, we will use this conjecture to study the asymptotic behavior of $P_{r,n}(N)$ as $N \longrightarrow \infty$, and provide computational evidence to support our findings.

2.2. **Bateman-Horn and $\#T_n(\mathbb{F}_q)$.** In order to use the Bateman-Horn conjecture in our study of $P_{r,n}(N)$, it is most natural to choose the polynomials $f_1 = x$ and $f_2 = \Phi_n(x^r)$. These distinct polynomials obviously both have positive leading coefficient. If it happens that $\Phi_n(x^r)$ is irreducible and no prime divides every element of the set $\mathcal{S}(f)$ as defined in (2.1) with $f = f_1 \cdot f_2 = x \cdot \Phi_n(x^r)$, then the Bateman-Horn conjecture yields

$$P_{r,n}(N) = Q(x, \Phi_n(x^r); N) \sim \frac{C(x, \Phi_n(x^r))}{r \cdot \varphi(n)} \int_2^N (\ln x)^{-2}\, dx, \quad \text{as } N \longrightarrow \infty. \tag{2.4}$$

We must now study the set $\mathcal{S}(f)$ and the factorization of $\Phi_n(x^r)$.

We begin with $r = 1$. In this case $f_2 = \Phi_n(x)$ is an irreducible polynomial and $f = x \cdot \Phi_n(x)$. A well known fact about cyclotomic polynomials states that:

$$\Phi_n(1) = \begin{cases} \rho, & \text{if } n \text{ is a power of some prime } \rho; \\ 1, & \text{otherwise.} \end{cases}$$

Recall that we are assuming that n is the squarefree product of the first few primes, and thus we will have $f(1) = 1$ except when $n = 2$. Excluding this exceptional case we see that no prime divides every element of the set $\mathcal{S}(f)$.

Now if $n = 2$, then $f_2 = \Phi_2(x) = x + 1$, hence $f = x(x + 1)$. Clearly then the prime 2 divides every element of the set $\mathcal{S}(f)$. In particular we conclude that for $n = 6, 30, 210$, etc., the necessary conditions on $f_1 = x$ and $f_2 = \Phi_n(x)$ for the use of the Bateman-Horn conjecture are satisfied.

The case $r > 1$ is somewhat more complicated. First we must determine whether or not $\Phi_n(x^r)$ is irreducible. Two additional well known facts about cyclotomic polynomials are as follows. If ρ is a prime which does not divide n, then

$$\Phi_n(x^\rho) = \Phi_n(x) \cdot \Phi_{\rho n}(x).$$

In particular, if ρ is a prime dividing r but not n, then define $d = r/\rho$ and substitute in x^d for x in the above identity to see that

$$\Phi_n(x^r) = \Phi_n(x^d) \cdot \Phi_{\rho n}(x^d),$$

from which it follows that $\Phi_n(x^r)$ is reducible. On the other hand, if every prime dividing r also divides n, then

$$\Phi_n(x^r) = \Phi_{rn}(x),$$

an irreducible polynomial. From this and the property of $\Phi_n(1)$ stated above, we conclude that in order to use the Bateman-Horn conjecture in the case $r > 1$, it must be that every prime dividing r also divides n.

Suppose we have fixed suitable r and n such that every prime dividing r also divides n. We would like to have many choices for a prime p such that $\Phi_n(q)$ is prime, where $q = p^r$. From what we have seen above, we can use the Bateman-Horn conjecture to estimate the number of choices for p. We now provide some computational evidence that this is indeed the case, and provide suggested parameters to construct secure torus-based cryptographic schemes.

3. Computations

3.1. **Computational Evidence for Bateman-Horn.** We have seen that the Bateman-Horn conjecture tells us nothing about the case $n = 2$, and so we will present computational evidence for the next few cases $n = 6, 30$ with $r = 1, 2$. Recall that $f_1 = x$ and $f_2 = \Phi_n(x^r)$, and so $d_1 = \deg f_1 = 1$ and $d_2 = \deg f_2 = r \cdot \varphi(n)$. First we made a rough approximation of $C(x, \Phi_n(x^r))$ using the primes up to 2^{15} and found:

$$\frac{C(x, \Phi_n(x^r))}{r \cdot \varphi(n)} \approx \begin{cases} 0.7605, & \text{if } n = 6, r = 1; \\ 1.1086, & \text{if } n = 6, r = 2; \\ 0.6909, & \text{if } n = 30, r = 1; \\ 0.4335, & \text{if } n = 30, r = 2. \end{cases}$$

For simplicity we replaced the integral in the approximation provided by (2.4) with a sum. For each combination of $n = 6, 30$ and $r = 1, 2$, we computed the value of $P_{r,n}(N)$ and the Bateman-Horn prediction $\mathrm{BH}_{r,n}(N)$ for $\log_2 N = 1, 2, \ldots, 30$. Our results, summarized in Tables 1 and 2, reconfirm that the Bateman-Horn conjecture gives very good estimates, even for relatively small values of N.

3.2. **Suggested Parameters for $T_n(\mathbb{F}_q)$.** As was previously mentioned, for the choices $n = 6, 30, 210$, any prime p such that $\Phi_n(p)$ is also prime leads to a group $T_n(\mathbb{F}_p)$ that is resistant to all known discrete logarithm attacks, provided that the following two conditions hold:

$$\log_2 p^n \geq 1024, \tag{3.1}$$

$$\log_2 \Phi_n(p) \geq 160. \tag{3.2}$$

Since $\Phi_n(p) \approx p^{\varphi(n)}$ for large p, it follows that for $n = 6, 30, 210$, condition (3.1) will imply condition (3.2). Following the construction in [6], we identified the smallest ten primes p satisfying condition (3.1) with $n = 6$, $p \equiv 2, 6, 7, 11 \pmod{13}$, and $\Phi_6(p)$ prime. In the interest of conserving space, each of these primes is represented as a sum $p_6 + v$, where

$$p_6 = 2\,375\,668\,978\,229\,576\,954\,621\,987\,172\,734\,316\,848\,349\,556\,051\,596\,973$$

is the smallest prime found, and v and $p_6 + v \pmod{13}$ are given in Table 3.

We also identified small primes suitable for use with schemes based on the conjectured rationality of T_{30} and T_{210}. Table 4 lists the ten smallest such primes for each case.

$\log_2 N$	$P_{1,6}(N)$	$\mathrm{BH}_{1,6}(N)$	$P_{2,6}(N)$	$\mathrm{BH}_{2,6}(N)$
1	1	2	1	2
2	2	3	2	4
3	3	4	3	5
4	4	4	4	7
5	4	6	6	8
6	4	7	8	11
7	6	10	10	14
8	10	13	13	19
9	14	19	21	27
10	22	28	36	40
11	40	42	58	62
12	63	67	94	97
13	100	108	158	157
14	186	178	267	260
15	298	301	453	439
16	500	515	752	751
17	885	894	1296	1304
18	1593	1568	2288	2285
19	2821	2774	4071	4043
20	4959	4945	7175	7208
21	8882	8874	12911	12937
22	16107	16021	23472	23355
23	29212	29075	42455	42384
24	52860	53013	77636	77278
25	97233	97067	142105	141496
26	178915	178412	260834	260075
27	329527	329076	480729	479703
28	609106	608926	889056	887647
29	1129888	1130102	1650290	1647378
30	2103603	2103096	3072103	3065736

TABLE 1. $P_{r,6}(N)$ and $\mathrm{BH}_{r,6}(N)$ for $r = 1, 2$ and $\log_2 N = 1, 2, \ldots, 30$.

$\log_2 N$	$P_{1,30}(N)$	$\mathrm{BH}_{1,30}(N)$	$P_{2,30}(N)$	$\mathrm{BH}_{2,30}(N)$
1	1	1	0	1
2	1	2	1	1
3	2	3	1	2
4	2	4	2	3
5	4	5	2	3
6	6	7	2	4
7	8	9	2	6
8	12	12	4	8
9	18	17	7	11
10	24	25	11	16
11	39	38	20	24
12	54	61	33	38
13	87	98	58	61
14	155	162	97	102
15	291	273	163	172
16	481	468	289	294
17	801	812	514	510
18	1396	1424	922	894
19	2473	2520	1581	1581
20	4463	4492	2818	2819
21	8144	8062	5068	5059
22	14769	14555	9229	9132
23	26724	26414	16967	16574
24	48298	48161	30501	30218
25	88313	88183	55587	55330
26	162218	162084	102108	101698
27	299335	298960	187870	187580
28	553937	553198	348182	347100
29	1027727	1026676	645942	644180
30	1915117	1910623	1201156	1198806

TABLE 2. $P_{r,30}(N)$ and $\mathrm{BH}_{r,30}(N)$ for $r = 1, 2$ and $\log_2 N = 1, 2, \ldots, 30$.

ACKNOWLEDGEMENTS

We thank Mark Bauer, Jesse Brinker, Elisa Gorla, Beth Malmskog, Eric Moorhouse, Vishal Saraswat and Samuel Wagstaff, Jr. for valuable discussions while in

v	$p_6 + v \pmod{13}$
0	7
2418	7
94458	7
202674	11
208584	6
245964	11
248430	7
257820	11
273840	2
344976	2

TABLE 3. Good primes for $T_6(\mathbb{F}_p)$.

$n = 30$ primes	$n = 210$ primes
18 843 310 259	43
18 843 311 363	73
18 843 311 771	409
18 843 314 339	653
18 843 314 821	757
18 843 317 303	1013
18 843 317 483	1153
18 843 318 833	1601
18 843 319 667	2027
18 843 323 479	2153

TABLE 4. Good primes for $T_{30}(\mathbb{F}_p)$ and $T_{210}(\mathbb{F}_p)$.

attendance at the 2006 Summer School on Computational Number Theory and Applications to Cryptography hosted by the University of Wyoming.

REFERENCES

[1] Paul T. Bateman and Roger A. Horn, *A Heuristic Asymptotic Formula Concerning the Distribution of Prime Numbers*, Mathematics of Computation, vol. 16, no. 79. (July, 1962), pp. 363–367, 1962.

[2] W. Bosma, J. Hutton and E. R. Verheul, *Looking Beyond XTR*, Proceedings of ASIACRYPT 2002, Lecture Notes in Computer Science, vol. 2501, pp. 46–63, Springer, 2002.

[3] R. Granger and F. Vercauteren, *On the Discrete Logarithm Problem on Algebraic Tori*, Proceedings of CRYPTO 2005, Lecture Notes in Computer Science, vol. 3621, pp. 66–85, Springer, 2005.

[4] A. K. Lenstra and E. R. Verheul, *An Overview of the XTR Public Key System*, in Public-Key Cryptography and Computational Number Theory (Warsaw, 2000), pp. 151–180, de Gruyter, 2001.

[5] Karl Rubin and Alice Silverberg, *Torus-Based Cryptography*, Proceedings of CRYPTO 2003, Lecture Notes in Computer Science, vol. 2729, pp. 349–365, Springer, 2003.

[6] Karl Rubin and Alice Silverberg, *Using Primitive Subgroups to Do More with Fewer Bits*, Proceedings of ANTS VI, Lecture Notes in Computer Science, vol. 3076, pp. 18–41, Springer, 2004.

[7] P. Smith and C. Skinner, *A Public-Key Cryptosystem and a Digital Signature System Based on the Lucas Function Analogue in Discrete Logarithms*, Proceedings of ASIACRYPT 1994, Lecture Notes in Computer Science, vol. 917, pp. 357–364, Springer, 1995.

[8] Marten van Dijk, Robert Granger, Dan Page, Karl Rubin, Alice Silverberg, Martijn Stam and David Woodruff, *Practical Cryptography in High Dimensional Tori*, Proceedings of EUROCRYPT 2005, vol. 3494, pp. 234–250, Springer, 2005.

[9] V. E. Voskresenskiĭ, Algebraic Groups and their Birational Invariants, Translations of Mathematical Monographs, vol. 179, American Mathematical Society, 1998.

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455-0436 USA

E-mail address: gower@ima.umn.edu