On the Bits of Elliptic Curve Diffie-Hellman Keys

David Jao[1], Dimitar Jetchev[2], and Ramarathnam Venkatesan[3,4]

[1] University of Waterloo, Waterloo ON N2L3G1, Canada
djao@math.uwaterloo.ca
[2] Dept. of Mathematics, University of California at Berkeley, Berkeley, CA 94720
jetchev@math.berkeley.edu
[3] Microsoft Research India Private Limited, "Scientia", No:196/36,
2nd Main Road, Sadashivnagar, Bangalore – 560080, India
[4] Microsoft Research, 1 Microsoft Way, Redmond WA 98052
venkie@microsoft.com

Abstract. We study the security of elliptic curve Diffie-Hellman secret keys in the presence of oracles that provide partial information on the value of the key. Unlike the corresponding problem for finite fields, little is known about this problem, and in the case of elliptic curves the difficulty of representing large point multiplications in an algebraic manner leads to new obstacles that are not present in the case of finite fields. To circumvent this obstruction, we introduce a *small* multiplier version of the hidden number problem, and we use its properties to analyze the security of certain Diffie-Hellman bits. We suggest new character sum conjectures that guarantee the uniqueness of solutions to the hidden number problem, and provide some evidence in support of the conjectures by showing that they hold on average in certain cases. We also present a Gröbner basis algorithm for solving the hidden number problem and recovering the Diffie-Hellman secret key when the elliptic curve is defined over a constant degree extension field and the oracle is a coordinate function in the polynomial basis.

1 Introduction

The Diffie-Hellman scheme is a fundamental protocol for public key exchange between two parties. Its original definition over finite fields is based on the hardness of computing the map $g, g^a, g^b \mapsto g^{ab}$ for $g \in \mathbb{F}_p^*$, while its elliptic curve analogue depends on the difficulty of computing $P, aP, bP \mapsto abP$ for points P on an elliptic curve.

A natural question in this context is whether an adversary can compute some *partial information* about g^{ab} (resp. abP) for the finite field (resp. the elliptic curve) case. In studying this problem for the finite field case, Boneh and Venkatesan [4] formulated the *hidden number problem* (HNP) and showed that a solution to the HNP allows one to reduce the question of computing partial information to the question of computing the key itself (see also [24,15]). For example, using these techniques one can show that computing $\mathrm{MSB}_k(g^{ab})$ is tantamount to

K. Srinathan, C. Pandu Rangan, M. Yung (Eds.): Indocrypt 2007, LNCS 4859, pp. 33–47, 2007.

computing g^{ab} itself for $k \geq 5\sqrt{\log p}$. In addition, the hidden number problem has turned out to be of cryptanalytic interest in its own right. For attacks on cryptosystems using partial information, see [20,23,21,16,24,17,22]. Thus an important motivation for the problem we consider is to find elliptic curve analogues of these attacks.

It is natural to ask the analogous question for elliptic curve Diffie-Hellman bits, namely, can we prove that partial information about elliptic curve Diffie-Hellman keys over a fixed curve E is unpredictable if we assume that the Diffie-Hellman problem for E is hard? Unfortunately, very little is known about this question. If one is allowed to look for a related curve with a hard Diffie-Hellman problem, then Boneh and Shparlinksi [3] provide an affirmative answer. While having formal proofs of the security of Diffie-Hellman bits is the most important application, it is also desirable from a cryptanalytic point of view to have practical algorithms for solving the corresponding hidden number problem (defined in Section 2.1). However, there are two fundamental obstructions which render the question much more difficult in the case of elliptic curves.

In the finite field case, one views elements of $\mathbb{F}_p$ as integers, embeds them in lattices equipped with the Euclidean metric and applies lattice reduction algorithms. In the elliptic curve case, no useful metrics are available; this represents the first fundamental obstruction. Furthermore, point multiplication on elliptic curves transforms the coordinates of a point via rational polynomials whose degrees grow exponentially in the size of the multiplier. This means that in general one can only write down explicit algebraic expressions in the case of small multipliers. This complexity constraint introduces the second fundamental obstruction—it is not even clear if the hidden number problem has a unique solution at all when the random multipliers are constrained to lie within small intervals. (By contrast, if one is allowed to use arbitrary multipliers, it is very easy to establish uniqueness in both the finite field and elliptic curve cases.) To deal with this obstruction, we introduce new character sums, conjecture some non-trivial estimates which are sufficient to prove uniqueness, and prove that our conjecture holds on average in the case of quadratic residuosity of the x coordinate. We also prove an upper bound on the number of solutions for any uniformly distributed output function, under the assumption of the Generalized Riemann Hypothesis. Although this approach falls short of the goal of actually recovering the value of abP via partial information, we feel that it remains a valuable first step given the lack of other results in this area.

We present a complete recovery algorithm for the hidden number problem in the case of curves over constant degree extensions, using Gröbner bases and elimination ideals. At present we are only able to implement our solution using oracles that provide outputs of length approximately 1/3 that of the (compressed) input point itself, e.g. 50 bits of output in the case of a 160-bit base field. Given recent progress and widespread interest in Gröbner bases algorithms, we may in the near future be able to recover Diffie-Hellman keys using less information (e.g. 32 bits of output for a 160-bit base field). However, obtaining results comparable to the

the finite field case (where we output $O(\sqrt{\log p})$ bits) seems to be fundamentally out of reach., and in certain cases is even presumed to be infeasible (see [2]).

2 Preliminaries

Let $q = p^k$ where p is a prime. We view $\mathbb{F}_q$ as a vector space over $\mathbb{F}_p$ and identify $\mathbb{F}_q$ with $\mathbb{F}_p^k$ using a polynomial basis. For a point P on an elliptic curve E over $\mathbb{F}_q$, let $x(P)$ and $y(P)$ be the x and y-coordinates of P, respectively, and let $x_0(P)$ denote the first coordinate in the vector representation of $x(P)$.

2.1 Partial Diffie-Hellman Bits

To extract partial information about points on elliptic curves, we consider a map $\mathbf{Bits}_\ell \colon E(\mathbb{F}_q) \to \{0,1\}^\ell$ which will assume one of the following three types:

1. Algebraic: $\mathbf{Bits}_\ell(P) = x_0(P)$;
2. Analytic: $\mathbf{Bits}_\ell(P) = \chi(x(P))$ for a suitable character $\chi : \mathbb{F}_q^\times \to \mathbb{C}^\times$;
3. MSB: $\mathbf{Bits}_\ell(P) = \mathrm{MSB}_\ell(x(P))$, which is the ℓ most significant bits of $x(P)$ expressed in binary.

Given a point $P \in E(\mathbb{F}_q)$ and two multiples aP and bP, let

$$\mathrm{PDH}_E(P, aP, bP) = \mathbf{Bits}_\ell(abP).$$

To study the security of the function PDH_E, we assume that there is a hidden point Q on E and an oracle $\mathcal{A}$ to compute the function $r \mapsto \mathbf{Bits}_\ell(rQ)$. We refer to r as the *multiplier*. One can then state the general *Multiplier Elliptic Curve Hidden Number Problem* (M-EC-HNP).

Multiplier-EC-Hidden-Number-Problem: Given an oracle $\mathcal{A}$ to compute the map $r \mapsto \mathbf{Bits}_\ell(rQ)$, recover the point Q.

Here the value of r may be chosen either by an adversary or randomly. In our setting, one queries the oracle many times so that one gets a total of $ck \log p$ output bits, for some $c > 1$. The hidden number problem is related to the problem of showing that PDH_E is secure, because a solution to the hidden number problem allows an adversary to determine abP given an oracle for PDH_E.

To solve the M-EC-HNP problem, one needs to address the following two questions:

Uniqueness: Is the underlying solution unique? If not, can the solutions at least be narrowed down to a small list?

Reconstruction: Is there an efficient algorithm to solve M-EC-HNP?

In most cases, one can easily show uniqueness if the queries are allowed to use large multipliers r. Unfortunately, these multipliers lead to division polynomials of exponentially large degree applied to Q, which cannot be handled using the techniques of Section 3. For this reason, any reconstruction algorithm based on these methods will be limited to small multipliers r with $r < O((\log p)^d)$. By contrast,

large multipliers for the analogous HNP over $\mathbb{F}_p$ pose no critical problems, and this difference represents a fundamental new restriction in the elliptic curve context. To analyze the statistical behavior of the output values for general oracles, we can apply the techniques of [18] which make use of the Generalized Riemann Hypothesis (see Section 5). However, these methods turn out to be insufficient for establishing uniqueness. To show uniqueness for the analytic case (only), we present a new character sum conjecture (and supporting evidence). Note that our algebraic map is significantly different from the finite field trace map used for bit extraction (see [13]) because the multipliers act via rational polynomial functions on the hidden point.

In the finite field case, the statistical properties and the pseudorandom number generators can be studied via estimates of character sums over *large* intervals (see [9,10,6,7,8,12]).

Remark 2.1. In the finite field case, one can use metrics and the LLL lattice reduction algorithm (see [19]) to reconstruct the secret efficiently (see [4] and [5]). However, without such metrics, there are no analogous reconstruction algorithms in the elliptic curve case. Nonetheless, we give a reconstruction algorithm using Gröbner bases algorithms in the algebraic case when k is small (see Section 3). In the analytic case, we use our character sum conjectures mentioned above to link the problem of solving M-EC-HNP to that of decoding certain error-correcting codes (see Section 4.2).

Remark 2.2. A detailed account on the general hidden number problem is given in [25]. The slightly more general hidden number problem for elliptic curves (as discussed in [2]) is the following:

EC-HNP: Let E be an elliptic curve over a finite field $\mathbb{F}_q$. Recover a point $P \in E(\mathbb{F}_q)$ given k pairs $(Q_i, \mathrm{MSB}_\ell(x(P + Q_i)))$ for some $\ell > 0$ and for k points $Q_1, \ldots, Q_k \in E(\mathbb{F}_q)$ chosen independently and at random.

3 Algebraic Case with Low Degree Extensions

In this section we outline an efficient reconstruction algorithm for the elliptic curve hidden number problem in the case of $\mathbf{Bits}_\ell(P) = x_0(P)$ over field extensions of constant degree. Since the technique is more transparent in the case of low degree extensions, we first illustrate the algorithm for degree 2 and degree 3 extensions before addressing the general case. Our method makes use of small multipliers and for this reason is limited to constant degree extensions.

3.1 Elliptic Curves over Finite Field Extensions of Degree 2

Suppose E is an elliptic curve over $\mathbb{F}_{p^2}$ given by a Weierstrass equation $y^2 = x^3 + \alpha x + \beta$ with $\alpha, \beta \in \mathbb{F}_{p^2}$. We will solve the M-EC-HNP in the algebraic case where $\mathbf{Bits}_\ell(P) = x_0(P)$ and $\ell = \lfloor \log_2 p \rfloor$.

Proposition 3.1. *Let $\ell = \lfloor \log_2 p \rfloor$ and $\mathbf{Bits}_\ell(P) = x_0(P)$. There exists an efficient algorithm (polynomial in $\log p$) for solving the M-EC-HNP.*

Proof. Let w be a generator for $\mathbb{F}_{p^2}/\mathbb{F}_p$, where $w^2 = u$ for some non-square element $u \in \mathbb{F}_p^\times$. Let $Q \in E(\mathbb{F}_{p^2})$ be the point which we are about to recover. It suffices to recover $\underline{x} = (x_0, x_1)$. The key ingredient for the proof is the observation that the coordinate $x(2Q)$ is expressible as a rational function purely of the coordinate x. More precisely, we have the point doubling formula [26, III.2.3]

$$x(2Q) = \frac{x^4 - 2\alpha x^3 - 8\beta x - \alpha^2}{4(x^3 + \alpha x + \beta)}.$$

We substitute $x = x_0 + wx_1$ into the right hand side and use $w^2 = u$ to write down

$$x(2Q) = \frac{P_0(x_0, x_1) + wP_1(x_0, x_1)}{Q_0(x_0, x_1) + wQ_1(x_0, x_1)},$$

where $P_0(x_0, x_1)$ and $P_1(x_0, x_1)$ are polynomials defined over $\mathbb{F}_p$ of degrees at most 4 and $Q_0(x_0, x_1)$ and $Q_1(x_0, x_1)$ are rational polynomials of degrees 3. Next, we rationalize the denominators to obtain

$$x(2Q) = \frac{P_0Q_0 - uP_1Q_1}{Q_0^2 - uQ_1^2} + w\frac{P_1Q_0 - P_0Q_1}{Q_0^2 - uQ_1^2}.$$

If $x(2Q) = (x_0', x_1')$ for some $x_0' \in \mathbb{F}_p$ and $x_1' \in \mathbb{F}_p$ then

$$x_0' = \frac{P_0(x_0, x_1)Q_0(x_0, x_1) - uP_1(x_0, x_1)Q_1(x_0, x_1)}{Q_0^2(x_0, x_1) - uQ_1^2(x_0, x_1)}.$$

This formula provides a way of patching together the partial data. Indeed, x_0 is recovered directly as $x_0 = \mathrm{MSB}_{\lfloor \log_2 p \rfloor}(x(Q))$. One also knows the value of $x_0' = \mathrm{MSB}_{\lfloor \log_2 p \rfloor}(x(2Q))$, so in order to recover x_1 one needs to find a zero over $\mathbb{F}_p$ of the polynomial

$$F(X) = $$
$$P_0(x_0, X)Q_0(x_0, X) - uP_1(x_0, X)Q_1(x_0, X) - x_0'Q_0(x_0, X)^2 - ux_0'Q_1(x_0, X)^2.$$

The explicit formula for P_0, P_1, Q_0, Q_1 show that F has constant degree (independent of E and p) and non-zero leading coefficient. Since we know that the hidden point Q exists, the polynomial must have a solution over $\mathbb{F}_p$. Computing the $\mathbb{F}_p$-roots can be solved in polynomial time using standard algorithms. This solves the M-EC-HNP in this particular case.

Remark 3.1. The solution of the M-EC-HNP in this case implies the security of the Diffie-Hellman bits for algebraic output functions on degree 2 field extensions. Indeed, if $\mathcal{A}$ is an oracle which computes $x_0(abP)$ from an input (P, aP, bP) then solving M-EC-HNP means that one could reconstruct the secret abP.

3.2 Elliptic Curves over Extensions of Degree 3

Let E be an elliptic curve over $\mathbb{F}_{p^3}$ given by a Weierstrass equation

$$E \ : \ y^2 = x^3 + \alpha x + \beta, \ \alpha, \beta \in \mathbb{F}_{p^3}.$$

We will show how to solve efficiently the M-EC-HNP in the algebraic case. The proof will be similar to the previous case of extensions of degree two, except that it will involve more technicalities. In what follows, $x_0(P)$ may be naturally extended by considering trace$(x(P))$.

Proposition 3.2. *Let $\ell = \lfloor \log_2 p \rfloor$ and $\mathbf{Bits}_\ell(P) = x_0(P)$. There exists an efficient algorithm (polynomial in $\log p$) for solving the M-EC-HNP.*

We first fix some choice for representing elements of the finite field. Let w be a generator for the field extension $\mathbb{F}_{p^3}/\mathbb{F}_p$. Without loss of generality (and to avoid some technical difficulties), choose w so that it is a root of an irreducible polynomial (over $\mathbb{F}_p$) whose quadratic term is zero, i.e., $w^3 - uw - v = 0$.

Proof. Let Q be the hidden point which we wish to recover. We write $x(Q) = (x_0, x_1, x_2)$ and $y(Q) = (y_0, y_1, y_2)$. Let $\mathcal{A}$ be an oracle which computes the function $r \mapsto x_0(rQ)$ for any $P \in E(\mathbb{F}_{p^3})$. We make three queries to $\mathcal{A}$ with $P = Q, 2Q$ and $3Q$, respectively. We use the fact that $x(2Q)$ and $x(3Q)$ are both rational functions of $x = x(Q)$. Let

$$x_0(Q) = s_1,\ x_0(2Q) = s_2,\ x_0(3Q) = s_3.$$

We will show how to put this information together, so that we can recover a finite (constant in p) list of candidates for the point Q.

The query $x_0(3Q)$. According to [26, Ex.3.7], the multiplication-by-3 map on E is given (as a rational function on the coordinates of Q) by

$$x(3Q) = \frac{\phi_3(x,y)}{\psi_3^2(x,y)},$$

where

$$\psi_3(x,y) = 3x^4 + 6\alpha x^2 + 12\beta x - \alpha^2$$

and

$$\begin{aligned}\phi_3(x,y) &= 8y^2(x^6 + 5\alpha x^4 + 20\beta x^3 - 5\alpha^2 x^2 - 4\alpha\beta x - 8\beta^2 - \alpha^3) = \\ &= 8(x^3 + \alpha x + \beta)(x^6 + 5\alpha x^4 + 20\beta x^3 - 5\alpha^2 x^2 - 4\alpha\beta x - 8\beta^2 - \alpha^3).\end{aligned}$$

Writing $\alpha = \alpha_0 + w\alpha_1 + w^2\alpha_2$, $\beta = \beta_0 + w\beta_1 + w^2\beta_2$ and $x = x_0 + wx_1 + w^2x_2$ we can express

$$\frac{\phi_3}{\psi_3^2} = \frac{P_0(x_0,x_1,x_2) + wP_1(x_0,x_1,x_2) + w^2P_2(x_0,x_1,x_2)}{Q_0(x_0,x_1,x_2) + wQ_1(x_0,x_1,x_2) + w^2Q_2(x_0,x_1,x_2)}, \tag{3.1}$$

where the P_i's and Q_i's are polynomials with coefficients in $\mathbb{F}_p$. The next step is to write the above rational function as

$$\frac{\phi_3}{\psi_3^2} = r_0(x_0,x_1,x_2) + wr_1(x_0,x_1,x_2) + w^2r_2(x_0,x_1,x_2),$$

where r_i's are rational functions over $\mathbb{F}_p$ which are explicitly computable in terms of α, β and w. To do this, we need to multiply the numerator and denominator of (3.1) by a suitable factor so that the denominator becomes a polynomial in x_0, x_1 and x_2 with coefficients in $\mathbb{F}_p$. Since w is a root of the polynomial $z^3 - uz - v = 0$ defined over $\mathbb{F}_p$ the rationalizing factor will be

$$\begin{aligned} F &= (Q_0 + w_1Q_1 + w_1^2Q_2)(Q_0 + w_2Q_1 + w_2^2Q_2) \\ &= Q_0^2 + Q_0Q_1(w_1 + w_2) + Q_0Q_2(w_1^2 + w_2^2) + Q_1^2w_1w_2 \\ &\quad + Q_1Q_2w_1w_2(w_1 + w_2) + Q_2^2w_1^2w_2^2 \\ &= (Q_0^2 + 2uQ_0Q_2 + uQ_1^2 + 2vQ_1Q_2 + u^2Q_2^2) \\ &\quad + w(-Q_0Q_1 + 2uQ_1Q_2 + vQ_2^2) + w^2(-Q_0Q_2 + Q_1^2 - uQ_1^2), \end{aligned}$$

where w_1 and w_2 are the other two roots of the above polynomial of degree 3 and for obtaining the last equality we have used $w + w_1 + w_2 = 0$, $w_1w_2w_3 = v$ and $w^3 - uw - v = 0$. Notice that $F\psi_3^2$ is a polynomial in x_0, x_1, x_2 of degree 24 defined over $\mathbb{F}_p$, and $F\phi_3$ (defined over $\mathbb{F}_{p^3}$) has degree 25. Thus, if we write $F\phi_3 = p_0 + wp_1 + w^2p_2$ where p_i's are polynomials in x_0, x_1, x_2 defined over $\mathbb{F}_p$ then we have $r_i = s_i/(F\psi_3^2)$ and the degree of the denominator of r_0 is at most 25, whereas the degree of its numerator is 24. The query $x_0(3Q) = s_3$ gives us the value of the function $r_0(x_0, x_1, x_2)$ at the triple $(x_0, x_1, x_2) \in \mathbb{F}_p^3$ which we are looking for.

The query $x_0(2Q)$. The point doubling formula reads as

$$x(2Q) = \frac{x^4 - 2\alpha x^3 - 8\beta x - \alpha^2}{4(x^3 + \alpha x + \beta)}.$$

Since $\alpha = \alpha_0 + w\alpha_1 + w^2\alpha_2$, $\beta = \beta_0 + w\beta_1 + w^2\beta_2$ and $x = x_0 + wx_1 + w^2x_2$, we can express

$$x(2Q) = \frac{R_0(x_0, x_1, x_2) + wR_1(x_0, x_1, x_2) + w^2R_2(x_0, x_1, x_2)}{T_0(x_0, x_1, x_2) + wT_1(x_0, x_1, x_2) + w^2T_2(x_0, x_1, x_2)}.$$

As in the case of multiplication by 3, we rationalize the above function by multiplying the numerator and denominator by

$$\begin{aligned} F = (T_0^2 + 2uT_0T_2 + uT_1^2 + 2vT_1T_2 + u^2T_2^2)+ \\ + w(-T_0T_1 + 2uT_1T_2 + vT_2^2) + w^2(-T_0Q_2 + T_1^2 - uT_2^2) \end{aligned}$$

and write it in the form

$$q_0(x_0, x_1, x_2) + wq_1(x_0, x_1, x_2) + w^2q_2(x_0, x_1, x_2),$$

where the q_i are $\mathbb{F}_p$-rational functions whose denominators have degree 9 and whose numerators have degree at most 10. As in the previous case, the query $x_0(2Q) = s_2$ gives us the value of the function $q_0(x_0, x_1, x_2)$ at the triple $(x_0, x_1, x_2) \in \mathbb{F}_p^3$.

Recovering $x(Q)$. We recover $x_0 = s_1$ from the query $x_0(Q) = s_1$. The query $x_0(2Q) = s_2$ gives us a polynomial relation $G(x_1, x_2) = 0$ over $\mathbb{F}_p$ between the (yet) unknown x_1 and x_2 coming from

$$q_0(s_1, x_1, x_2) = s_2.$$

Note that the degree of G is at most 10. Similarly, the query $x_0(3Q)$ gives us a polynomial relation $H(x_1, x_2) = 0$ over $\mathbb{F}_p$ coming from

$$r_0(s_1, x_1, x_2) = s_3$$

of degree at most 25. Thus, (x_1, x_2) is a simultaneous solution over $\mathbb{F}_p$ of G and H. We determine the solutions by taking the resultant $\mathrm{Res}(G, H)$, which is a polynomial in a single variable of degree at most 250. Since we are only interested in the $\mathbb{F}_p$-solutions, it suffices to factor the resultant over $\mathbb{F}_p$ and look up the linear factors. Thus, we obtain a finite (constant in p) set of possible solutions (x_1, x_2), which proves the proposition.

3.3 Elliptic Curves over $\mathbb{F}_q$

Let $q = p^k$. Let E be an elliptic curve over $\mathbb{F}_q$ given by a Weierstrass equation

$$E \ : \ y^2 = x^3 + \alpha x + \beta, \ \alpha, \beta \in \mathbb{F}_q.$$

We will describe an algorithm to solve the M-EC-HNP for $\ell = \lfloor \log_2 p \rfloor$ and $\mathbf{Bits}_\ell(P) = x_0(P)$ which will generalize the previous cases of extensions of degrees two and three.

Let w be a generator for the field extension $\mathbb{F}_q/\mathbb{F}_p$ and let

$$f(z) = z^k - u_1 z^{k-1} - \cdots - u_k$$

be the minimal polynomial for w over $\mathbb{F}_p$. As before, suppose that we have an oracle $\mathcal{A}$ which computes $x_0(P')$ given P', aP', bP'. Our goal is to recover $x(Q)$ given

$$\langle x_0(mQ) \ : \ m = 1, 2, \ldots, k \rangle.$$

Let $x = x(Q) = x_0 + wx_1 + \cdots + w^{k-1}x_{k-1}$. The main idea is to interpret the above data as a system of polynomial equations with coefficients in $\mathbb{F}_p$ and degrees bounded independently of $\log p$, and to use a Gröbner basis algorithm to solve the system and thereby compute $x_0, x_1, \ldots, x_{k-1}$. To compute the equations in the system, we use the division polynomials from [26, §III, Ex.3.7] to find $x(mQ) = \varphi_m(Q)/\psi_m(Q)^2$. Next, we observe that φ_m/ψ_m^2 is a rational function on x defined over $\mathbb{F}_q$, and we write it (after rationalizing the denominators) as

$$r_0^{(m)}(x_0, x_1, \ldots, x_{k-1}) + w r_1^{(m)}(x_0, x_1, \ldots, x_{k-1}) + \cdots \\ + w^{k-1} r_{k-1}^{(m)}(x_0, x_1, \ldots, x_{k-1}),$$

where $r_i^{(m)}(x_0, x_1, \ldots, x_{k-1})$ are rational functions defined over $\mathbb{F}_p$ whose denominators have degrees $k \deg(\psi_m^2) = 2k(m^2-1)$ and whose numerators have degrees $m^2 + (k-1)(m^2-1) = 2k(m^2-1)+1$ (here, we are using that the single-variate polynomial ψ_m has degree m^2-1 and ϕ_m has degree m^2). Next, if $x_0(mQ) = s_m$ then we obtain the equation

$$r_0^{(m)}(x_0, x_1, \ldots, x_{k-1}) = s_m, \ \forall m = 1, \ldots, k,$$

which gives us a polynomial equation over $\mathbb{F}_p$

$$g_m(x_0, x_1, \ldots, x_{k-1}) = 0,$$

whose degree is bounded by $2k(m^2-1)+1$. We then use a Gröbner basis algorithm to try to compute a Gröbner basis for the ideal

$$I = \langle g_1, \ldots, g_k \rangle \subset \mathbb{F}_p[x_0, \ldots, x_{k-1}],$$

which will allow us to solve for $x_0, x_1, \ldots, x_{k-1}$.

Wide practical interest in Gröbner bases and continued improvements in algorithmic implementations have pushed the limits of what can be solved by these systems. For a 160-bit EC system it may soon be possible to solve the problem using outputs of $\ell \in [16, 32]$ bits per iteration (for example, by using a degree ten extension, and outputting one co-ordinate of $x(P)$).

4 Analytic Case

We now consider the case of an output function which is equal to a group character (such as the quadratic residuosity character). Let $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ be a nontrivial character of the multiplicative group $\mathbb{F}_q^\times$. Given a point P, we will look at the values $\chi(x(P))$ defined by the character χ. For completeness, if $x(P) = 0$ we set $\chi(0) = 0$.

4.1 Our Conjecture on Character Sums

Our conjecture is the following:

Conjecture 4.1. *Let $P, P' \in E(\mathbb{F}_q)$ be two points, such that $x(P) \neq x(P')$. There exists $\varepsilon > 0$, such that for every $B = \Omega((\log q)^2)$,*

$$\left| \sum_{r \leq B} \chi(x(rP))\chi(x(rP')) \right| = O(B^{1-\varepsilon}).$$

Although this bound suffices for our needs, the actual bound may be closer to $B^{0.5}$. Note that classical character sums over finite fields traditionally take the form $\sum_{x \in \mathbb{F}_q} \chi(f(x))$ for some polynomial $f(x)$. General character sums over $\mathbb{F}_q$ have been considered by Deligne in [11], but very little is known over short intervals. Viewed as a sum over some function field, the above sum does not include all polynomials of small degree, but only those that correspond to the map $P \mapsto rP$.

We obtain the following immediate corollary of the above conjecture.

Corollary 4.2. *Assuming Conjecture 4.1, let P and P' be two points in $E(\mathbb{F}_q)$ with $x(P) \neq x(P')$ and let χ be the quadratic character. Let $B = \Omega((\log q)^2)$ For randomly a chosen $r \in \{1, \ldots, B\}$*

$$\left| \text{Prob}_r \left[\chi(x(rP)) \neq \chi(x(rP')) \right] - \frac{1}{2} \right| = O(B^{-\varepsilon}).$$

Remark 4.1. The purpose of this conjecture is to show that knowing enough of the values of partial bits (namely, character values) of $x(r_i P)$ for small multipliers suffices to uniquely identify the point. If we assume Conjecture 4.1, and choose $P, P' \in E(\mathbb{F}_q)$ such that $x(P) \neq x(P')$, then for B and χ as above and random integers $r_1, \ldots, r_t \in \{1, B\}$, the values $\{\chi(x(r_i P))\}_{i=1}^t$ and $\{\chi(x(r_i P'))\}_{i=1}^t$ will be distinct with high probability.

4.2 Relationship to an Error-Correcting Code

Proofs of many hard-core bit theorems involve problems related to error correcting codes, and the hard core property of the bit is equivalent to finding efficient decoding algorithms for certain codes (see [1]). In our case this connection exists as well, but it is unclear if the code admits an efficient decoding algorithm. However, if decoding should turn out to be intractable then the code may be of independent interest in cryptography. Therefore, it seems worthwhile to mention the resulting code here.

We define a binary code that uses small multiples of points on elliptic curves for encoding. We fix a finite field $\mathbb{F}_q$ and a bound $B \leq O(\log q)^2$. Our choice of code corresponds to a selection of a random sequence $\mathbf{c} = c_1, \ldots, c_t$ with $1 \leq c_i \leq B$ and a character $\chi \colon \mathbb{F}_q^\times \to \{\pm 1\}$. The parameter t will be the length of the code words, and $\mathbb{F}_q$ will roughly correspond to the message space in the following manner. Given an easily invertible map $m \mapsto P$ to map messages into points on elliptic curves, our algorithm to encode m works as follows: let P be its image on the curve and let $P, P_1, \ldots, P_t$ be the sequence of nodes visited by the walk specified by $\mathbf{c}$. Our encoding of m is the sequence $\chi(x(P_1)), \ldots, \chi(x(P_t))$.

Our character sum assumptions imply that the minimum distance of the code is $(\frac{1}{2} - \varepsilon)t$. It is clear that any decoding algorithm that maps an uncorrupted codeword into the point P can be used to solve the hidden number problem using the analytic bit extractor, while correcting a corrupted codeword will yield a proof for the pseudo-randomness of Diffie-Hellman bits.

4.3 Proof of Our Conjecture on Average

We provide some evidence in support of Conjecture 4.1 by showing that the conjecture holds on average for the quadratic character of Corollary 4.2, in the sense that

$$\left| \sum_{P \in E} \sum_{r \leq B} \chi(x(rP)) \chi(x(rP')) \right| \leq (\#E) \cdot O(B^{1/2})$$

for a fixed $P' \in E$ when $\chi\colon \mathbb{F}_q^\times \to \{\pm 1\}$ is the quadratic character. We start with the identity

$$\left|\sum_{P\in E}\sum_{r\le B}\chi(x(rP))\chi(x(rP'))\right| = \left|\sum_{r\le B}\chi(x(rP'))\right| \cdot \left|\sum_{P\in E}\chi(x(P))\right|,$$

and we will prove that

$$\left|\sum_{P\in E}\chi(x(P))\right| = O(\sqrt{\#E}).$$

Clearly this bound is sufficient to finish the proof. To prove it, observe that

$$\left|\sum_{P\in E}\chi(x(P))\right| = \left|\sum_{x\in\mathbb{F}_q}(1+\chi(x^3+\alpha x+\beta))\chi(x)\right| = \left|\sum_{x\in\mathbb{F}_q}\chi(x^4+\alpha x^2+\beta x)\right|,$$

where $y^2 = x^3 + \alpha x + \beta$ is the equation for E. Let C denote the curve $y^2 = x^4 + \alpha x^2 + \beta x$. Then C is singular if and only if $\beta = 0$ or $4\alpha^3 + 27\beta^2 = 0$. The latter possibility may be excluded since E is nonsingular. Hence we are left with two cases to consider. If C is nonsingular, then the claim follows from the work of Deligne [11]. On the other hand, if $\beta = 0$, then $\alpha \neq 0$, and

$$\left|\sum_{x\in\mathbb{F}_q}\chi(x^4+\alpha x^2)\right| = \left|\sum_{x\in\mathbb{F}_q}\chi(x^2+\alpha)\right|.$$

Now the curve C' given by $y^2 = x^2 + \alpha$ is again nonsingular, so [11] again gives the desired bound.

5 Expander Graphs and Character Sums

In this section we formalize the small multiplier hidden number problem in terms of graph theory by defining a graph whose edges correspond to pairs of points related by small multipliers. This graph is isomorphic (as a graph) to a certain Cayley graph for $(\mathbb{Z}/N\mathbb{Z})^\times$. Under the assumption of the Generalized Riemann Hypothesis, we establish its eigenvalue separation. Our graph is directed, but standard techniques allow us to infer the rapid mixing of directed graphs by analyzing its undirected version. The rapid mixing of the graph implies, on average and with high probability, an upper bound on the number of solutions to the small multiplier hidden number problem, for *any* function $\mathbf{Bits}_\ell$ whose output values are uniformly distributed over the points of the elliptic curve.

Let $q = p^k$ and assume that the number of points $N = \#E(\mathbb{F}_q)$ is prime (this is a reasonable cryptographic assumption).

5.1 Constructing the Graph G_E and the Subgraph G'

Let $m = O((\log q)^d)$ for $d > 2$ and some sufficiently large (but absolute) implied constant, and S_m be the set of all prime numbers less than or equal to m. Define a directed graph G_E with nodes consisting of the points $Q \in E(\mathbb{F}_q)$ and edges of the form $\{Q, rQ\}$ for every prime $r \in S_m$.

Consider the subgraph G' with vertices consisting of all points $P \neq O_E$. Since N is prime, this graph is isomorphic (only as a graph) to the Cayley graph of $(\mathbb{Z}/N\mathbb{Z})^\times$ with respect to S_m. To establish such an isomorphism, choose a generator Q of $E(\mathbb{F}_q)$ and a primitive element $g \in (\mathbb{Z}/N\mathbb{Z})^\times$ and map each vertex via $sQ \mapsto g^s$ and each edge via $\{Q, sQ\} \mapsto \{g, g^s\}$. Using the arguments of [18], one shows under GRH that the graph G' is $\#S_m$-regular and connected. Specifically, the eigenvalues of the adjacency matrix are the character sums $\lambda_\chi = \sum_{p \in S_m} \chi(\bar{p}) = \sum_{p \leq m} \chi(\bar{p})$, where $\chi\colon (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ varies over the characters of $(\mathbb{Z}/N\mathbb{Z})^\times$ and $\bar{p}$ denotes the image of the prime p in $(\mathbb{Z}/N\mathbb{Z})^\times$. The eigenvector corresponding to the eigenvalue λ_χ is $e_\chi = (\chi(x))_{x \in (\mathbb{Z}/N\mathbb{Z})^\times}$. The largest eigenvalue, corresponding to the trivial character, is $\lambda_{\text{triv}} = \pi(m)$. Hence, to show that G' has good expansion properties, we need an estimate on λ_χ. Such an estimate can be obtained using the methods of [18]. More precisely, under the Generalized Riemann Hypothesis, one can show that $\lambda_\chi < C(N)\sqrt{\pi(m)}$ for some constant $C(N)$ (depending only on N) with $\lim_{N \to \infty} C(N) = 0$, whenever χ is a non-trivial character.

5.2 Distributional Properties

Consider a pseudorandom number generator that initializes P_0 to some random point on E and then performs the following steps:

1. Choose $r_i \in [1, B]$ at random (where $B = O((\log p)^2)$.
2. Set $P_{i+1} = r_i P_i$.
3. Output $\mathbf{Bits}_\ell(x(P_{i+1}))$.

Given a sequence of ℓ-bit strings $h_1, h_2, \ldots, h_L$, what is the probability that the generator will output this sequence? Using the methods of [14], suitably adapted to our situation, we prove the following proposition, which provides a satisfactory bound as long as the second eigenvalue of the normalized adjacency matrix of the graph G' is small (which is the case for large enough N). This bound implies a corresponding upper bound on the number of solutions to the related hidden number problem under the GRH assumption (although it would take some effort to work out what exactly the corresponding bound is).

Proposition 5.1. *Let $h_1, h_2, \ldots, h_L$ be a sequence of values of the function*

$$\mathbf{Bits}_\ell\colon E(\mathbb{F}_q) \setminus \{O_E\} \to \{0,1\}^\ell,$$

such that the sets $F_i := \mathbf{Bits}_\ell^{-1}(h_i)$ have size $\mu_i v$, where $v = \#V(G_E)$. Let A be the normalized adjacency matrix of G', with second largest eigenvalue λ_2. The number of random walks on G' of length L, such that the i-th node in the walk is equal to h_i is bounded by $\prod_{i=1}^{L} M_i$, where $M_i = \sqrt{\mu_i^2 + \lambda_2^2 + 2\mu_i\sqrt{1-\mu_i}\lambda_2}$.

Proof. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_v$ be the eigenvalues of A. We denote by $e_1, \ldots, e_v$ the corresponding eigenvectors. The eigenvalue $\lambda_1 = 1$ is the trivial eigenvalue and its eigenvector is $e_1 = (1, 1, \ldots, 1)$. Let $V_1 \subset \mathbb{R}^v$ be the subspace spanned by e_1 and $V_2 \subset \mathbb{R}^v$ be the subspace spanned by $e_2, \ldots, e_v$. The spaces V_1 and V_2 are orthogonal to each other and are both preserved by A.

One can then write a given vector $X \in \mathbb{R}^v$ as $X = X_1 + X_2$, where $X_1 \in V_1$ and $X_2 \in V_2$. For each $i = 1, \ldots, L$, denote by P_i the projection operator to the set F_i. In other words, $P_i X$ is the vector $Y \in \mathbb{R}^v$, whose coordinates Y_j for each $1 \leq j \leq v$ are given by $Y_j = X_j$ if the j-th node of G' is a point in F_i and $Y_j = 0$ otherwise.

The proof is based on the observation that if $X = e_1/v$, then the j-th component of the vector $Y = \prod_{i=1}^{L}(PA_i)X$ is exactly the probability that a random walk of length L ends in the j-th node of G' in such a way that for each $i = 1, \ldots, L$ the walk has passed through F_i at the i-th step. Therefore, the probability that a random walk lands in the set F_i at the i-th step is given by

$$P(\text{walk passes through } F_1, \ldots, F_L) = \sum_{j=1}^{v} |Y_j| \leq \sqrt{v}\|Y\| = \sqrt{v} \left\| \prod_{i=1}^{L} (PA_i)X \right\|,$$

where $\|Y\|$ is the L^2-norm of Y, i.e. $\|Y\| = \sqrt{\sum_{i=1}^{v} |Y_i|^2}$.

We will be done if we find an upper bound for $\|P_i AU\|/\|U\|$ for arbitrary vectors $U \in \mathbb{R}^v$ and projection operators P_i. Let $U = U_1 + U_2$ for $U_1 \in V_1$ and $U_2 \in V_2$. Since $AU_1 = U_1$ and $P_i^2 = P_i$, we obtain

$$\|P_i AU\| = \|P_i(P_i U_1 + AU_2)\| \leq \|P_i U_1 + AU_2\|.$$

Our goal is to give an upper bound of $\|P_i U_1 + AU_2\|$ in terms of $\|U\| = \|U_1 + U_2\|$. Since $P_i U_1$ is no longer a vector in $V_1 = (V_2)^{\perp}$, we need to estimate the cosine of the angle between $P_i U_1$ and AU_2 and then use the law of cosines to express the sum in terms of this estimate. Let θ_i be the angle between U_1 and $P_i U_1$. Then

$$\cos\theta_i = \frac{U_1 \cdot P_i U_1}{\|U_1\|\|P_i U_1\|} = \frac{|F_i|}{\sqrt{|F_i|}\sqrt{|G|}} = \sqrt{\frac{|F_i|}{|G|}} = \sqrt{\mu_i}.$$

In particular, we have $0 \leq \theta_i \leq \frac{\pi}{2}$. Let ϕ_i be the angle between $P_i U_1$ and AU_2. Since $\phi_i \leq \frac{\pi}{2} + \theta_i \leq \pi$, it follows that $-\cos\phi_i \leq -\cos\left(\frac{\pi}{2} + \theta_i\right)$ and therefore

$$\begin{aligned} \|P_i U_1 + AU_2\|^2 &= \|P_i U_1\|^2 + \|AU_2\|^2 - 2\|P_i U_1\|\|AU_2\|\cos\phi_i \\ &\leq \|P_i U_1\|^2 + \|AU_2\|^2 - 2\|P_i U_1\|\|AU_2\|\cos\left(\frac{\pi}{2} + \theta_i\right). \end{aligned}$$

But

$$-\cos\left(\frac{\pi}{2} + \theta_i\right) = \sin\theta_i = \sqrt{1 - \cos^2\theta_i} = \sqrt{1 - \mu_i}.$$

Moreover, $\|P_i U_1\| \leq \mu_i |G|$ and $\|AU_2\| \leq \lambda_2 |G|$, so

$$\|P_i U_1 + AU_2\|^2 \leq (\mu_i^2 + \lambda_2^2 + 2\mu_i\sqrt{1 - \mu_i}\lambda_2)|G|.$$

Thus,

$$P(\text{walk passes through } F_1, \ldots, F_L) \leq \prod_{i=1}^{L} \sqrt{\mu_i^2 + \lambda_2^2 + 2\mu_i\sqrt{1-\mu_i}\lambda_2}.$$

References

1. Akavia, A., Goldwasser, S., Safra, S.: Proving hard-core predicates using list decoding. In: FOCS 2003. Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science, p. 146. IEEE Computer Society, Washington, DC (2003)
2. Boneh, D., Halevi, S., Howgrave-Graham, N.: The modular inversion hidden number problem. In: Boyd, C. (ed.) ASIACRYPT 2001. LNCS, vol. 2248, pp. 36–51. Springer, Heidelberg (2001)
3. Boneh, D., Shparlinski, I.: On the unpredictability of bits of the elliptic curve Diffie-Hellman scheme. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 201–212. Springer, Heidelberg (2001)
4. Boneh, D., Venkatesan, R.: Hardness of computing the most significant bits of secret keys in Diffie-Hellman and related schemess. In: Koblitz, N. (ed.) CRYPTO 1996. LNCS, vol. 1109, pp. 129–142. Springer, Heidelberg (1996)
5. Boneh, D., Venkatesan, R.: Rounding in lattices and its cryptographic applications. In: Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 675–681. ACM, New York (1997)
6. Bourgain, J.: New bounds on exponential sums related to the Diffie-Hellman distributions. C.R. Math. Acad. Sci. Paris 338(11), 825–830 (2004)
7. Bourgain, J.: Estimates on exponential sums related to the Diffie-Hellman distributions. Geom. Funct. Anal. 15(1), 1–34 (2005)
8. Bourgain, J.: On an exponential sum related to the Diffie-Hellman cryptosystem. Int. Math. Res. Not., pages Art. ID 61271, 15 (2006)
9. Canetti, R., Friedlander, J., Konyagin, S., Larsen, M., Lieman, D., Shparlinski, I.: On the statistical properties of Diffie-Hellman distributions. Israel J. Math. 120, 23–46 (2000)
10. Canetti, R., Friedlander, J., Shparlinski, I.: On certain exponential sums and the distribution of Diffie-Hellman triples. J. London Math. Soc (2), 59(3), 799–812 (1999)
11. Deligne, P.: Cohomologie étale. In: de Boutot, J.F., Grothendieck, A., Illusie et, L., Verdier, J.L. (eds.) Séminaire de Géométrie Algébrique du Bois-Marie SGA $4\frac{1}{2}$, Avec la collaboration. Lecture Notes in Mathematics, vol. 569, Springer, Berlin (1977)
12. Friedlander, J., Shparlinski, I.: On the distribution of the power generator. Math. Comp (electronic) 70(236), 1575–1589 (2001)
13. Galbraith, S., Hopkins, H., Shparlinski, I.: Secure bilinear Diffie-Hellman bits. In: Wang, H., Pieprzyk, J., Varadharajan, V. (eds.) ACISP 2004. LNCS, vol. 3108, pp. 370–378. Springer, Heidelberg (2004)
14. Goldreich, O., Impagliazzo, R., Levin, L., Venkatesan, R., Zuckerman, D.: Security preserving amplification of hardness. In: 31st Annual Symposium on Foundations of Computer Science, vol. I, II, pp. 318–326. IEEE Comput. Soc. Press, Los Alamitos, CA (1990)
15. González Vasco, M.I., Shparlinski, I.: On the security of Diffie-Hellman bits. In: Cryptography and computational number theory, Progr. Comput. Sci. Appl. Logic, vol. 20, pp. 257–268. Birkhäuser, Basel (2001)

16. González Vasco, M.I., Shparlinski, I.: Security of the most significant bits of the Shamir message passing scheme. Math. Comp (electronic) 71(237), 333–342 (2002)
17. Howgrave-Graham, N., Nguyen, P., Shparlinski, I.: Hidden number problem with hidden multipliers, timed-release crypto, and noisy exponentiation. Math. Comp (electronic) 72(243), 1473–1485 (2003)
18. Jao, D., Miller, S.D., Venkatesan, R.: Do all elliptic curves of the same order have the same difficulty of discrete log? In: Roy, B. (ed.) ASIACRYPT 2005. LNCS, vol. 3788, pp. 21–40. Springer, Heidelberg (2005)
19. Lenstra, A.K., Lenstra Jr., H.W., Lovász, L.: Factoring polynomials with rational coefficients. Math. Ann. 261(4), 515–534 (1982)
20. Nguyen, P.: The dark side of the hidden number problem: lattice attacks on DSA. In: Cryptography and computational number theory, Progr. Comput. Sci. Appl. Logic, Birkhäuser, Basel, vol. 20, pp. 321–330 (2001)
21. Nguyen, P., Shparlinski, I.: The insecurity of the digital signature algorithm with partially known nonces. J. Cryptology 15(3), 151–176 (2002)
22. Nguyen, P., Shparlinski, I.: The insecurity of the elliptic curve digital signature algorithm with partially known nonces. Des. Codes Cryptogr. 30(2), 201–217 (2003)
23. Shparlinski, I.: On the generalised hidden number problem and bit security of XTR. In: Bozta, S., Sphparlinski, I. (eds.) Applied Algebra, Algebraic Algorithms and Error-Correcting Codes. LNCS, vol. 2227, pp. 268–277. Springer, Heidelberg (2001)
24. Shparlinski, I.: Cryptographic applications of analytic number theory. In: Progress in Computer Science and Applied Logic, Complexity lower bounds and pseudorandomness, vol. 22, Birkhäuser Verlag, Basel (2003)
25. Shparlinski, I.: Playing 'hide-and-seek' with numbers: the hidden number problem, lattices and exponential sums. In: Public-key cryptography, Proc. Sympos. Appl. Math., vol. 62, pp. 153–177. Amer. Math. Soc., Providence, RI (2005)
26. Silverman, J.: The arithmetic of elliptic curves. In: Graduate Texts in Mathematics, vol. 106, Springer, New York (1992) Corrected reprint of the 1986 original