

## OBSTACLES TO THE TORSION-SUBGROUP ATTACK ON THE DECISION DIFFIE-HELLMAN PROBLEM

NEAL KOBLITZ AND ALFRED J. MENEZES

ABSTRACT. Cheng and Uchiyama show that if one is given an elliptic curve, depending on a prime  $p$ , that is defined over a number field and has certain properties, then one can solve the Decision Diffie-Hellman Problem (DDHP) in  $\mathbb{F}_p^*$  in polynomial time. We show that it is unlikely that an elliptic curve with the desired properties exists.

### 1. INTRODUCTION

The Discrete Logarithm Problem (DLP) in the multiplicative group  $\mathbb{F}_q^*$  of the field of  $q$  elements, along with the closely related Diffie-Hellman Problem (DHP) and Decision Diffie-Hellman Problem (DDHP), have been the subject of cryptographic research for many years. Recall that the DLP is the problem, given  $g \in \mathbb{F}_q^*$  and  $y$  in the subgroup generated by  $g$ , of finding an integer  $x$  such that  $y = g^x$ ; the DHP is the problem, given  $g, g^{x_1}, g^{x_2} \in \mathbb{F}_q^*$ , of finding  $g^{x_1 x_2}$ ; and the DDHP is the problem, given  $g, g^{x_1}, g^{x_2}, g^{x_3} \in \mathbb{F}_q^*$ , of determining whether or not  $x_3 \equiv x_1 x_2 \pmod{\ell}$ , where  $\ell$  is the multiplicative order of  $g$ .

The authors of [3] consider the DDHP in the multiplicative group  $\mathbb{F}_p^*$  of a prime field. They show that if one is given an elliptic curve, depending on  $p$ , that is defined over a number field and has certain properties, then one can solve any instance of the DDHP in  $\mathbb{F}_p^*$  in polynomial time. They conjecture the existence of elliptic curves with the desired properties. However, the purpose of this paper is to give evidence that the elliptic curves needed in [3] do not exist.

We emphasize that, even if the curves in the conjecture in [3] existed, the result claimed in that paper would not be of practical value unless a reasonable algorithm could be developed to find them. The authors of [3] did not consider this question, and we also will ignore the question of how to find these curves in the unlikely event that they exist.

We now describe the contents of the paper. In §2 we put the purported result of [3] in the context of Maurer's work on the equivalence of the DHP and the DLP. In §3 we introduce elliptic curves over number fields. In order to make everything as self-contained and readable as possible, our treatment of this subject is informal. For details and proofs we refer the reader to Silverman's two volumes [21, 22]. In §4 we list the properties of the elliptic curves that are conjectured in [3] to exist, and we outline the algorithm in [3] for solving the DDHP in  $\mathbb{F}_p^*$ .

---

Received by the editor May 29, 2002 and, in revised form, May 3, 2003.

2000 *Mathematics Subject Classification*. Primary 94A60, 11T71, 14G50.

*Key words and phrases*. Discrete logarithm, Diffie-Hellman Problem, elliptic curve, torsion point, modular curve.

In §5 we discuss the Uniform Boundedness Conjecture (now a theorem), and we investigate the claim in [3] that as the prime  $\ell \rightarrow \infty$  there should exist elliptic curves defined over number fields of degree  $< (\log \ell)^{O(1)}$  having points of order  $\ell$  with coordinates in that number field. From the literature on the subject we conclude that this claim is most likely false, although at present mathematicians are not able to prove that it is false.

In §6 we describe the connection between torsion points on elliptic curves and the so-called modular curves. This section, like §3, is informal and contains no proofs. In §7 we describe the cases  $\ell = 5, 7, 11$ .

In §8 we put aside the question of the degree of the number field that we looked at in §5, and instead we ask about the size of the discriminant of the elliptic curve and the size of the coordinates of the torsion points. We prove that the first of these is polynomially bounded in terms of the second, and we prove a partial converse—that the norm of the  $y$ -coordinate of a torsion point is bounded by the square root of the norm of the discriminant of the curve. We then give reasons why one expects the discriminant of the curve (and therefore also the coordinates of the torsion points) to grow exponentially with  $\ell$ .

In the conclusion we ruminate on the different types of conditional results in cryptography and end with a warning about the danger to the credibility of the research community if we ascribe any validity at all to a result that is contingent upon mathematicians' inability to prove that a certain type of curve does not exist.

## 2. MAURER'S RESULT

The definitions of the DLP, DHP and DDHP given in §1 can be extended to arbitrary finite cyclic groups. It is clear that the DDHP reduces in polynomial time to the DHP, and the DHP reduces in polynomial time to the DLP. Polynomial time reductions of DLP to DHP, or of DHP to DDHP, are not known. There are, however, some groups  $G$  in which the DLP is believed to be intractable but where the DLP can be reduced in polynomial time to the DHP in  $G$ ; examples include arbitrary cyclic groups of order  $n$  where  $\phi(n)$  is smooth (see [1]). There are also some groups in which no polynomial time algorithms for the DHP are known, but where the DDHP can be solved in polynomial time; examples include certain supersingular elliptic curves and certain elliptic curves of trace 2 (see [11]). Finally, we note that Shoup [20] has proved lower bounds of the form  $\Omega(\sqrt{n})$  for the DLP, DHP and DDHP in generic groups of prime order  $n$ . (A generic group is one whose elements have random labelings and which comes equipped with an efficient oracle for performing the group operation.) Shoup's result provides some evidence for the intractability of the DLP, DHP and DDHP in groups that are used in cryptography.

While hardness of the DLP in a group  $G$  is necessary for the security of discrete logarithm cryptographic schemes in  $G$  (since otherwise an adversary can compute private keys from public keys), it is generally not sufficient. For example, the ElGamal public-key encryption scheme [6] in the group  $\mathbb{F}_p^*$  is not semantically secure against passive attacks. This type of security would mean that an adversary is unable to recover any partial information about the plaintext from the public key and ciphertext. In the ElGamal encryption scheme, a user's public key is  $g^a$ , where  $g$  is a generator of  $\mathbb{F}_p^*$  and  $a \in [1, p-1]$  is the user's private key. A message  $m \in \mathbb{F}_p$  is encrypted to  $c = (g^k, mg^{ak})$ , where  $k$  is randomly selected from  $[1, p-1]$ . Even if the DHP (and thus also the DLP) in  $\mathbb{F}_p^*$  is intractable, a passive adversary can deduce

partial information about  $m$  from  $c$  and  $g^a$ —namely, the adversary can compute the Legendre symbol  $(\frac{m}{p})$ . On the other hand, the ElGamal encryption scheme in a prime-order subgroup  $G$  of  $\mathbb{F}_p^*$  can be proven to be semantically secure against passive attacks under the assumption that the DDHP in  $G$  is intractable. Recently, Cramer and Shoup [4] introduced an ElGamal-like encryption scheme which they proved semantically secure against *active* attacks (where the adversary is given the decryptions of polynomially many ciphertexts of her own choosing) assuming the intractability of the DDHP in  $G$ . For other cryptographic applications of the DDHP, see the survey article of Boneh [2].

Maurer [16] proved the polynomial time equivalence of the DLP and DHP in groups  $G$  of prime order  $\ell$  under the assumption that certain elliptic curves exist and can be efficiently found. Namely, if one is given an elliptic curve  $E$  defined over  $\mathbb{F}_\ell$  such that  $E(\mathbb{F}_\ell)$  is cyclic and  $\#E(\mathbb{F}_\ell)$  is  $(\log \ell)^c$ -smooth (i.e., the largest prime factor of  $\#E(\mathbb{F}_\ell)$  is at most  $(\log \ell)^c$ ), then the DLP in  $G$  can be reduced in polynomial time to the DHP in  $G$ . The coefficients of the elliptic curve  $E$  comprise a polynomial size “hint” depending only on  $\ell$  that, once known, allows one to solve *any* instance of the DLP in *any* group  $G$  of order  $\ell$  in polynomial time given a Diffie-Hellman oracle for  $G$ . It is not known if this hint exists; however, heuristic arguments about the distribution of smooth integers in the Hasse interval  $[\ell - 2\sqrt{\ell} + 1, \ell + 2\sqrt{\ell} + 1]$  suggest that such hints do exist. Even if the hint does exist, it is not known how to find it in polynomial time—an exhaustive search would, in general, take  $O(\ell^c)$  time. Nevertheless, Maurer’s result can be viewed as providing some evidence for the equivalence of the DLP and DHP.

A natural question that is still open is whether there exist polynomial size hints, depending only on  $\ell$ , that, once known, allow one in polynomial time to solve any instance of the DLP in any group of order  $\ell$ . Observe that the restriction on the size of the hint is important, since otherwise a table of logarithms for  $\mathbb{F}_p^*$  suffices. The question is also of interest with DLP replaced by DHP or DDHP. In [3], Cheng and Uchiyama consider this question for the DDHP when the group is a subgroup of prime order  $\ell$  in  $\mathbb{F}_p^*$ ; their hint depends on both  $\ell$  and  $p$ . The basic idea of [3] is to use elliptic curves over number fields to reduce these DDHP instances to DDHP instances in certain elliptic curve over finite fields where the DDHP is known to be solvable in polynomial time (see [10], [11] and [24]).

### 3. BACKGROUND ON ELLIPTIC CURVES

Let  $F$  be a field of characteristic not equal to 2 or 3, and let  $a$  and  $b$  be elements of  $F$  such that the *discriminant*  $\Delta = 4a^3 + 27b^2$  is nonzero. By an *elliptic curve*  $E$  defined over  $F$  we mean the equation

$$(1) \quad y^2 = x^3 + ax + b.$$

The condition  $\Delta \neq 0$  says that the cubic polynomial on the right has distinct roots; that means that the curve does not have any singular points and cannot be transformed into a rational curve. By  $E(F)$  we mean the set of all points with coordinates in  $F$  that satisfy (1), along with the so-called *point at infinity*, denoted  $O$ . This set  $E(F)$  forms an abelian group with zero-element  $O$ . More generally, if  $L$  is any field containing  $F$ , we let  $E(L)$  denote the group consisting of points with coordinates in  $L$  that satisfy (1), along with the point at infinity. Then  $E(F)$  is a subgroup of  $E(L)$ .

Now suppose that  $F = K$  is a number field, that is, a finite extension of the field  $\mathbb{Q}$  of rational numbers. A point of finite order on  $E(K)$  is called a *torsion point*. We let  $E_{\text{tors}}(K) \subset E(K)$  denote the subgroup of all torsion points; by the Mordell-Weil theorem, we know that this is always a finite group.

Suppose that we have an elliptic curve (1) defined over  $\mathbb{Q}$  with integer coefficients  $a$  and  $b$ . We can reduce its equation modulo primes  $p$  and consider the resulting equation over the field  $\mathbb{F}_p$  of  $p$  elements. The result is not necessarily an elliptic curve, however, because the discriminant  $\Delta$  might reduce to zero modulo  $p$ . This is called *bad reduction*. In the case when  $\Delta \equiv 0 \pmod{p}$ , the points of the reduced curve (excluding the singular point) still form a group, but the group is isomorphic to one we already knew about before we were working with elliptic curves. Namely, there are three possible groups that occur:

- (a) (*additive bad reduction*) the additive group  $\mathbb{F}_p^+$ ;
- (b) (*split multiplicative bad reduction*) the multiplicative group  $\mathbb{F}_p^*$ ;
- (c) (*nonsplit multiplicative bad reduction*) the subgroup of elements of order  $p + 1$  in the multiplicative group  $\mathbb{F}_{p^2}^*$  of the quadratic extension.

A simple example of these possibilities is given by the elliptic curve  $y^2 = x^3 + ax^2 + p$  (note that this is not quite of the form (1), but it can easily be transformed to the form (1) by a linear change of the  $x$ -variable). If  $a = 0$ , we get additive reduction at  $p$ ; if  $a$  is a quadratic residue in  $\mathbb{F}_p^*$ , we get split multiplicative reduction at  $p$ ; and if  $a$  is a nonresidue, we get nonsplit multiplicative reduction at  $p$ .

Similar definitions apply to an elliptic curve defined over a number field  $K$ , except that instead of reducing modulo a prime number, we reduce modulo a prime ideal of  $K$ , and instead of  $\mathbb{F}_p$  we work with the residue field of that ideal.

Occasionally it is necessary to use a more general form of the equation of an elliptic curve over a field  $F$ , namely

$$(2) \quad y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

If the characteristic of the field  $F$  is not 2 or 3, then this equation can be transformed into (1) by a linear change of variables. However, sometimes one might not want to do this. For example, the equation  $y^2 + y = x^3$  over  $\mathbb{Q}$  is equivalent to the equation  $y^2 = x^3 + 16$ . (In the first equation just replace  $y$  by  $\frac{1}{8}y - \frac{1}{2}$  and  $x$  by  $\frac{1}{4}x$ ; then multiply through by 64.) The second equation has additive bad reduction at the prime 2, whereas the first equation is a perfectly good elliptic curve modulo 2. In general, we do not say that an elliptic curve has bad reduction at  $p$  if it is possible to transform its equation into one that has discriminant prime to  $p$ . In that case we refer to the latter equation and its discriminant as the *minimal equation* and *minimal discriminant* (see p. 172 of [21] for more information on minimal equations and discriminants). Thus, the discriminant of the equation  $y^2 = x^3 + 16$  is  $6912 = 2^8 \cdot 3^3$ , but its minimal discriminant is 27. The elliptic curve given by  $y^2 = x^3 + 16$  has bad reduction only at 3, because its minimal equation has discriminant divisible only by the prime 3.

We next explain what a complex-multiplication (CM) curve is. If  $E$  is an elliptic curve (1) defined over a number field (or any subfield of  $\mathbb{C}$ ), its complex points  $E(\mathbb{C})$  may be regarded as the complex plane modulo a certain lattice  $\mathbb{L}$ ; that is, two complex numbers are considered equivalent if their difference is in the lattice. We write  $E(\mathbb{C}) \approx \mathbb{C}/\mathbb{L}$ . The equivalence classes of the complex plane modulo  $\mathbb{L}$  can be visualized by means of a *fundamental parallelogram* for  $\mathbb{L}$ . This is a parallelogram

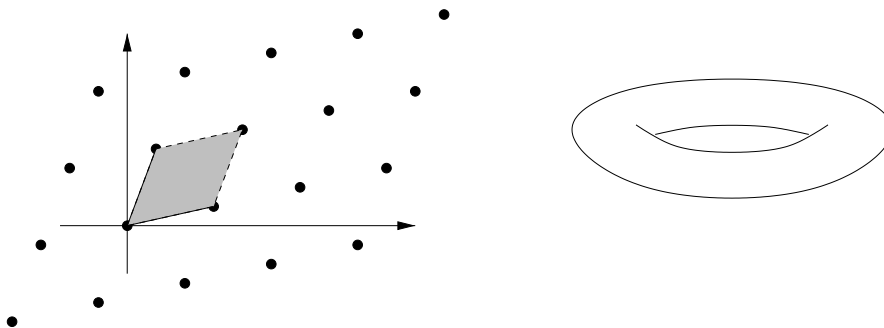


FIGURE 1. Fundamental parallelogram of a lattice, and a torus.

two of whose sides are basis vectors for  $\mathbb{L}$  emanating from the origin. Every complex number is equivalent modulo  $\mathbb{L}$  to a point of the parallelogram, and no two points of the interior of the parallelogram are equivalent to one another. However, the points opposite one another on the perimeter of the parallelogram do differ by a lattice element. Thus, the opposite sides of the parallelogram should be viewed as “glued together” with equivalent points joined. The resulting geometrical shape is a *torus*, depicted in Figure 1.

If we replace  $\mathbb{L}$  by a multiple  $c\mathbb{L}$ , where  $c$  is a nonzero complex number, this does not change the elliptic curve (more precisely, the two elliptic curves are said to be *isomorphic*). Thus, without loss of generality we may suppose that  $\mathbb{L} = \mathbb{Z}\tau + \mathbb{Z}$  is generated by the number 1 and a complex number  $\tau$  having positive imaginary part.

The Weierstrass  $\wp$ -function and its derivative are used to construct a one-to-one correspondence between the points of the complex plane modulo  $\mathbb{L}$  and the complex points of the elliptic curve  $E$  given by (1). Under this correspondence, the group law on the elliptic curve comes from the obvious additive group structure on  $\mathbb{C}/\mathbb{L}$  (namely, vector addition of complex numbers modulo the lattice).

We say that an elliptic curve with lattice  $\mathbb{L}$  has *complex multiplication* if there exist nontrivial complex numbers  $\alpha$  (“nontrivial” means  $\alpha \notin \mathbb{Z}$ ) such that  $\alpha\mathbb{L} \subset \mathbb{L}$ . It is easy to show that if such  $\alpha$  exist, then the set of all  $\alpha$  such that  $\alpha\mathbb{L} \subset \mathbb{L}$  forms a subring of the ring of integers of a certain imaginary quadratic field. This field is called the *CM-field* of the curve. For example, the curve  $y^2 = x^3 + ax$  has complex multiplication by the Gaussian integers; its CM-field is  $\mathbb{Q}(i)$ . Similarly, the curve  $y^2 = x^3 + b$  has CM-field  $\mathbb{Q}(\zeta)$ , where  $\zeta = (-1 + \sqrt{-3})/2$ . It should be noted that CM-curves are rare, in the sense that a random equation (1) almost certainly will give a non-CM curve.

#### 4. THE DDHP ALGORITHM OF CHENG AND UCHIYAMA

In this section we outline the algorithm of Cheng and Uchiyama [3] for solving the DDHP in the multiplicative group of a prime field. We begin by listing the properties that must be satisfied by the number field  $K$  and the elliptic curve  $E$  in [3].

Let  $p$  and  $\ell$  be primes with  $\ell|p-1$ . By “polynomial size” we mean of bitlength  $(\log p)^{O(1)}$ . (Since  $p$  is usually polynomial in  $\ell$ , we shall sometimes write  $(\log \ell)^{O(1)}$  instead.) Concerning  $K$ , one needs both its degree over  $\mathbb{Q}$  and the coefficients

of a generating polynomial to have polynomial size. Next, the elliptic curve  $E$  must be defined over  $K$  and have points of order  $\ell$  with coordinates in  $K$  that all have polynomial size. In [3] the authors claim that the coefficients of their curve  $E$  do not have to satisfy a polynomial bound; however, in §8 we shall prove that their assumption regarding polynomial size torsion points in fact implies that the coefficients of  $E$  also have polynomial size. Finally, all of the prime ideals of  $K$  dividing  $p$  must have degree 1 (that means that the residue field is  $\mathbb{F}_p$ ),  $E$  must have split multiplicative bad reduction at  $p$  (i.e., at some prime ideal  $u$  of  $K$  dividing  $p$ ), and the points of order  $\ell$  must not reduce modulo  $p$  to the singular point.

Now suppose that we wish to solve the DDHP in the subgroup of order  $\ell$  in  $\mathbb{F}_p^*$ . That is, we are given an element  $g \in \mathbb{F}_p^*$  of order  $\ell$ ,  $g^{x_1}$ ,  $g^{x_2}$  and  $g^{x_3}$ , and we wish to determine whether or not  $x_3 \equiv x_1x_2 \pmod{\ell}$ . Let  $\tilde{E}$  denote the singular elliptic curve over  $\mathbb{F}_p$  obtained by reducing  $E$  modulo  $u$ . Since  $\tilde{E}$  has split multiplicative bad reduction at  $u$ ,  $\tilde{E}_{ns}(\mathbb{F}_p) \cong \mathbb{F}_p^*$ , where  $\tilde{E}_{ns}$  denotes the non-singular points on  $\tilde{E}$ . Moreover, an isomorphism  $\phi : \mathbb{F}_p^* \rightarrow \tilde{E}_{ns}(\mathbb{F}_p)$  can be efficiently computed. Next, one lifts the points  $A = \phi(g)$ ,  $A_1 = \phi(g^{x_1})$ ,  $A_2 = \phi(g^{x_2})$ ,  $A_3 = \phi(g^{x_3})$  to  $\ell$ -torsion points  $B, B_1, B_2, B_3$  in  $E(K)$ . (This can be accomplished by first lifting to points over the field of  $p$ -adic numbers—see [3] for details.) Now let  $r$  be a prime such that  $\ell$  divides  $r - 1$ ,  $\ell > 4\sqrt{r}$ , and  $E$  has good reduction at some prime ideal  $v$  dividing  $r$ . Let  $E'$  denote the elliptic curve over  $\mathbb{F}_r$  obtained by reducing  $E$  modulo  $v$ , and let  $C, C_1, C_2, C_3 \in E'(\mathbb{F}_r)$  be the points obtained by reducing  $B, B_1, B_2, B_3$  modulo  $v$ . Since  $\ell > 4\sqrt{r}$ , the only multiple of  $\ell$  in the Hasse interval  $[r - 2\sqrt{r} + 1, r + 2\sqrt{r} + 1]$  is  $r - 1$ . Since  $\ell$  divides  $\#E'(\mathbb{F}_r)$ , we conclude that  $\#E'(\mathbb{F}_r) = r - 1$ . Since  $r < \ell^2$ , it follows that  $\ell^2$  does not divide  $r - 1$ . Thus, as observed by Joux [10], the Tate pairing  $\langle P, P \rangle$  for points  $P$  of order  $\ell$  in  $E'(\mathbb{F}_r)$  is a nontrivial  $\ell$ th root of unity in  $\mathbb{F}_r^*$ . Finally, to solve the original DDHP, one computes  $\langle C_1, C_2 \rangle = \langle C, C \rangle^{x_1x_2}$  and  $\langle C, C_3 \rangle = \langle C, C \rangle^{x_3}$ ; we have  $x_3 \equiv x_1x_2 \pmod{\ell}$  if and only if  $\langle C_1, C_2 \rangle = \langle C, C_3 \rangle$ .

## 5. DEGREE OF THE NUMBER FIELD

One of the great achievements of the theory of elliptic curves over number fields was the complete proof of the Uniform Boundedness Conjecture (UBC). This remarkable conjecture (now a theorem) says that there exists a bound  $B(d)$ , depending only on the degree of the number field, such that the torsion subgroup of *any* elliptic curve  $E$  over *any* number field  $K$  of degree  $d$  has no more than  $B(d)$  elements. The first major result in the direction of proving the UBC was Mazur's theorem for  $d = 1$  (i.e.,  $K = \mathbb{Q}$ ) in 1978 [17]. Mazur proved that  $B(1) = 16$  and that 7 is the largest prime that can be the order of a point in  $E_{\text{tors}}(\mathbb{Q})$ .

The  $d = 2$  case of the UBC was proved fourteen years later by Kamienny [12]. It turns out that when  $\ell$  is a prime greater than 13, there cannot be a point of order  $\ell$  in  $E_{\text{tors}}(K)$  for any quadratic field  $K$ . Soon after, the UBC was proved for some larger values of  $d$  and then finally in [18] for all  $d$ . Merel [18] also proved the bound  $\ell \leq d^{3d^2}$  for  $d > 1$  for primes dividing the order of the torsion subgroup, and Oesterlé soon showed how to improve this bound to  $\ell \leq (3^{d/2} + 1)^2$  (see [8]). The Oesterlé bound is the best result that has been proved without any restriction on the elliptic curves.

Recall that in [3] the authors conjecture that as  $\ell \rightarrow \infty$  there exist number fields  $K$  of degree  $< (\log \ell)^{O(1)}$  and elliptic curves  $E$  over  $K$  having points of order  $\ell$ . The authors cite Oesterlé's bound as support for this conjecture, and say: "Fortunately, the current research seems to indicate that the maximum possible number of torsions over a number field grows exponentially with the degree of the number field."

But a more dispassionate examination of the literature reveals a somewhat different picture. In addition to Oesterlé's bound, specialists have been able to prove that

- (1) if there is a prime of additive bad reduction, then  $\ell < 48d$  (Flexor-Oesterlé [7]);
- (2) for curves with good reduction everywhere one has  $\ell < 1977408d \log d$ ; and, more generally, if there is a bound  $s$  on the number of prime ideals of  $K$  where the curve has bad reduction, then there exists a constant  $c_s$  depending only on  $s$  such that  $\ell < c_s d \log d$  (Hindry-Silverman [8]);
- (3) for a fixed curve  $E$ , considered over varying extension fields of its field of definition, there exists a constant  $c$  depending on  $E$  such that  $\ell < cd \log d$  (Masser [15]).

What about the other direction? What families of  $E(K)$  are known with points of large prime order? The strongest result currently known comes from fixing a curve  $E$  and then considering the same curve over field extensions  $K$  of its field of definition. In this way one can obtain points of prime order  $\ell$  where  $\ell$  is only  $\approx \sqrt{d}$ . (The reason is that the  $x$ -coordinate of a point of order  $\ell$  satisfies the  $\ell$ th division polynomial, which has degree  $(\ell^2 - 1)/2$ ; thus, one usually has to go to a  $O(\ell^2)$ -degree extension to get the coordinates of such a point.) It is an open question [23] whether or not there exists a family of elliptic curves over number fields  $K$  that have a  $K$ -point of prime order  $\ell$  with  $\ell$  growing faster than  $\sqrt{\deg K}$ .

Several experts in the area have suggested that the order  $\ell$  of a torsion point, even if it does not satisfy an upper bound as small as  $\sqrt{d}$ , is likely at least to satisfy a polynomial upper bound. This is stated explicitly by S. David<sup>1</sup> and more tentatively by Hindry and Silverman.<sup>2</sup>

If  $\ell$  is bounded by a polynomial in  $d$ , then the strategy in [3] immediately fails. As we shall see in §8, that is not the only reason for why the torsion-subgroup approach is likely to fail. But first we give some background on modular curves.

## 6. THE MODULAR CURVES $X(\ell)$ , $X_1(\ell)$ , AND $X_0(\ell)$

In §3 we described the connection between lattices  $\mathbb{Z}\tau + \mathbb{Z}$  and elliptic curves. The purpose of this section is to discuss the curves that parameterize the set of all elliptic curves having a point of order  $\ell$ .

Let  $\Gamma$  denote the group of all  $2 \times 2$  matrices with integer entries and determinant 1 ( $\Gamma$  is often denoted  $SL_2(\mathbb{Z})$ ). For each  $\ell$  one can consider the subgroup consisting of all matrices that are congruent modulo  $\ell$  to the identity matrix. This subgroup is denoted  $\Gamma(\ell)$ . We also define two intermediate subgroups, denoted  $\Gamma_0(\ell)$  and  $\Gamma_1(\ell)$ . A matrix  $\gamma \in \Gamma$  is said to be in  $\Gamma_0(\ell)$  if its lower-left entry is divisible by  $\ell$ ,

<sup>1</sup>"Notons toutefois que ces majorations semblent indiquer que l'ordre de la torsion ne devrait dépendre que polynomialement du degré du corps et non exponentiellement comme c'est le cas en ce moment en [18]" ([5], p. 107).

<sup>2</sup>"Il est naturel de demander s'il existe une borne polynomiale en  $d$ " ([8], p. 97).

and it is said to be in  $\Gamma_1(\ell)$  if its lower-left entry is divisible by  $\ell$  and its diagonal entries are congruent to 1 modulo  $\ell$ . In other words, a matrix  $\gamma \in \Gamma$  that reduces modulo  $\ell$  to a matrix of the form

$$\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}, \quad \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

belongs respectively to  $\Gamma_0(\ell)$ ,  $\Gamma_1(\ell)$ ,  $\Gamma(\ell)$ . We obviously have the inclusions

$$\Gamma(\ell) \subset \Gamma_1(\ell) \subset \Gamma_0(\ell) \subset \Gamma.$$

For prime  $\ell$  the index of the first of these groups in the second is  $\ell$ , the index of the second in the third is  $\ell - 1$ , and the index of the third in the fourth is  $\ell + 1$ . In particular,  $\Gamma_1(\ell)$  is a subgroup of  $\Gamma$  of index  $\ell^2 - 1$ . (Note: Sometimes  $\Gamma$  and  $\Gamma_0(\ell)$  are defined as the quotient of the above groups by the 2-element subgroup  $\left\{ \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ ; in that case  $\Gamma_1(\ell)$  is a subgroup of  $\Gamma$  of index  $(\ell^2 - 1)/2$ .)

Given a lattice  $\mathbb{L}$  in the complex plane that is generated by complex numbers  $\omega_1$  and  $\omega_2$ , we define the action of the group  $\Gamma$  on the pair of generators by the usual matrix operation on a vector. That is, for  $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , we get a new basis for the same lattice as follows:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} a\omega_1 + b\omega_2 \\ c\omega_1 + d\omega_2 \end{pmatrix}.$$

In particular, the basis  $\begin{pmatrix} \tau \\ 1 \end{pmatrix}$  for a lattice of the form  $\mathbb{L} = \mathbb{Z}\tau + \mathbb{Z}$  is transformed to the basis  $\begin{pmatrix} a\tau + b \\ c\tau + d \end{pmatrix}$ . If we want our lattice to have one basis element equal to 1, we can scale the lattice  $\mathbb{L}$  by  $\frac{1}{c\tau + d}$  to get the equivalent lattice  $\mathbb{Z}\gamma(\tau) + \mathbb{Z}$ , where  $\gamma(\tau)$  is defined as

$$(3) \quad \gamma(\tau) = \frac{a\tau + b}{c\tau + d}.$$

It is easy to check that  $\gamma(\tau)$  has positive imaginary part if  $\tau$  does; so (3) gives an action of the group  $\Gamma$  on the complex upper half-plane  $\mathcal{H}$ . To summarize: for any  $\gamma \in \Gamma$ , the lattice  $\mathbb{Z}\gamma(\tau) + \mathbb{Z}$  is obtained from the lattice  $\mathbb{Z}\tau + \mathbb{Z}$  by change of basis and then scaling by the second element of the new basis. This lattice gives the same elliptic curve (i.e., an isomorphic one) as the old lattice.

In this way we get a one-to-one correspondence between isomorphism classes of elliptic curves defined over  $\mathbb{C}$  and equivalence classes of numbers in the upper half-plane  $\mathcal{H}$ , where  $\tau'$  is said to be equivalent to  $\tau$  if  $\tau' = \gamma(\tau)$  for some  $\gamma \in \Gamma$ .

Now let us consider torsion points. The complex curve  $E(\mathbb{C})$  has  $\ell^2$  points of order  $\ell$ . Under the correspondence between the elliptic curve and the complex plane modulo the lattice  $\mathbb{Z}\tau + \mathbb{Z}$ , the points of order  $\ell$  correspond to the  $\ell^2$  elements of the lattice  $\mathbb{Z}\left(\frac{\tau}{\ell}\right) + \mathbb{Z}\left(\frac{1}{\ell}\right)$  regarded modulo  $\mathbb{Z}\tau + \mathbb{Z}$ . These are the  $\ell^2$  equally spaced points of a fundamental parallelogram for the lattice  $\mathbb{Z}\tau + \mathbb{Z}$  (see Figure 2).

If our element  $\gamma \in \Gamma$  actually lies in  $\Gamma(\ell)$ , then notice what happens to the basis  $\frac{\tau}{\ell}, \frac{1}{\ell}$  of the lattice of torsion points:

$$(4) \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \frac{\tau}{\ell} \\ \frac{1}{\ell} \end{pmatrix} = \begin{pmatrix} \frac{a\tau}{\ell} + \frac{b}{\ell} \\ \frac{c\tau}{\ell} + \frac{d}{\ell} \end{pmatrix} = \begin{pmatrix} \frac{\tau}{\ell} + \frac{a-1}{\ell}\tau + \frac{b}{\ell} \\ \frac{1}{\ell} + \frac{c}{\ell}\tau + \frac{d-1}{\ell} \end{pmatrix}.$$



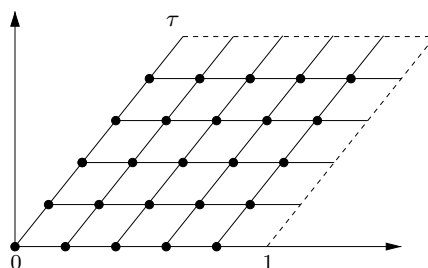


FIGURE 2. Points of order 5 in the fundamental parallelogram of a lattice  $\mathbb{Z}\tau + \mathbb{Z}$ .

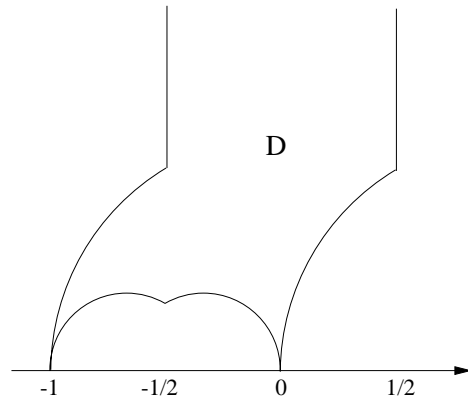
But since  $\ell$  divides  $a - 1, b, c,$  and  $d - 1,$  it follows that modulo the lattice  $\mathbb{Z}\tau + \mathbb{Z}$  these elements are equivalent to  $\begin{pmatrix} \tau/\ell \\ 1/\ell \end{pmatrix}$ . In other words, the group  $\Gamma(\ell)$  preserves not only the elliptic curve, but also a basis of the points of order  $\ell$ .

Similarly, if  $\gamma \in \Gamma$  is actually in  $\Gamma_1(\ell),$  then we see from (4) that one torsion element  $\frac{1}{\ell}$  is taken to itself (more precisely, to the number  $\frac{1}{\ell} + (\frac{c}{\ell}\tau + \frac{d-1}{\ell})$ , which differs from  $\frac{1}{\ell}$  by a lattice element). That is,  $\Gamma_1(\ell)$  preserves not only the elliptic curve, but also a point of order  $\ell$ . Just as  $\Gamma$ -equivalence classes of points in the complex upper half-plane  $\mathcal{H}$  correspond to elliptic curves, the  $\Gamma_1(\ell)$ -equivalence classes of  $\mathcal{H}$  correspond to pairs consisting of an elliptic curve and a point of order  $\ell$  on it.

In the case of  $\Gamma_0(\ell),$  what gets preserved is not a point of order  $\ell$  but rather a subgroup of order  $\ell$ . That is, if  $\gamma \in \Gamma_0(\ell)$  in (4), then the multiples of  $1/\ell$  get permuted, and the subgroup made up of these multiples is preserved. Thus, the  $\Gamma_0(\ell)$ -equivalence classes of  $\mathcal{H}$  correspond to pairs consisting of an elliptic curve and a subgroup of order  $\ell$ .

For any of the groups  $\Gamma, \Gamma(\ell), \Gamma_1(\ell), \Gamma_0(\ell),$  we can construct a *fundamental domain* for its action on  $\mathcal{H}$ . This means that we find a region  $D$  with the property that every point of  $\mathcal{H}$  is equivalent to a point of  $D$  and no two points of the interior of  $D$  are equivalent to one another. Some points on the boundary of  $D$  are equivalent to one another, so we can visualize the fundamental domain as “glued together” by joining the equivalent points on its boundary. The simplest case is  $\Gamma,$  where the fundamental domain  $D_0$  can be taken to be the part of the vertical strip of width 1 centered on the  $y$ -axis that is above the unit circle. A fundamental domain  $D$  for a subgroup of  $\Gamma$  can be constructed by putting together  $D_0$  along with its images  $\gamma_i^{-1}D_0,$  where  $\gamma_i \in \Gamma$  runs through a set of coset elements of  $\Gamma$  modulo the subgroup in question. In the case of  $\Gamma_0(\ell)$  there are  $\ell + 1$  such coset elements, and in the case of  $\Gamma_1(\ell)$  there are  $(\ell^2 - 1)/2$ . In Figure 3 we see a fundamental domain for  $\Gamma_0(3),$  where we have taken coset elements  $\gamma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \gamma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \gamma_3 = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix},$   
 $\gamma_4 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$

It is a basic result of the theory of modular curves that the fundamental domains can be regarded as algebraic curves defined over  $\mathbb{Q}$ . The curves coming from  $\Gamma(\ell), \Gamma_1(\ell),$  and  $\Gamma_0(\ell)$ —after they are “compactified” by adding some “cusps,” which are

FIGURE 3. A fundamental domain  $D$  for  $\Gamma_0(3)$ .

a finite number of points not corresponding to elliptic curves<sup>3</sup>—are denoted  $X(\ell)$ ,  $X_1(\ell)$ , and  $X_0(\ell)$ . We shall be mainly interested in  $X_1(\ell)$ , which, as we have seen, parameterizes elliptic curves along with a torsion point of order  $\ell$ . More precisely, if a (noncusp) point of  $X_1(\ell)$  has coordinates in a number field  $K$ , then it corresponds to an elliptic curve defined over  $K$  that has a point of order  $\ell$  that is also defined over  $K$ . (A  $K$ -point of  $X_0(\ell)$  corresponds to an elliptic curve defined over  $K$  with a subgroup of order  $\ell$  that is defined over  $K$ . What that means is that we have to take an extension  $L$  of  $K$  to get a point  $P$  of order  $\ell$ , but at least we can say that if any automorphism in  $\text{Gal}(L/K)$  is applied to  $P$ , the new point will still be in the subgroup generated by  $P$ .)

In [17] Mazur completely described all torsion subgroups that can occur on an elliptic curve over  $\mathbb{Q}$ . In particular, he proved that for primes  $\ell > 7$  there are no elliptic curves over  $\mathbb{Q}$  with a point of order  $\ell$ . Mazur proved this fundamental result by studying the family of curves  $X_1(\ell)$ . Because of the correspondence between noncusp  $\mathbb{Q}$ -points on  $X_1(\ell)$  and elliptic curves defined over  $\mathbb{Q}$  with a point of order  $\ell$  with coordinates in  $\mathbb{Q}$ , his result can be given as a statement about rational points on the curves  $X_1(\ell)$ : *if  $\ell$  is a prime greater than 7, then  $X_1(\ell)$  has no rational points except cusps*. Thus, Mazur's theorem is the modular curve analogue of Fermat's Last Theorem, which, of course, can be stated in the form: *if  $\ell$  is a prime greater than 2, then the curve  $x^\ell + y^\ell = 1$  has no rational points except the trivial ones  $(1, 0)$  and  $(0, 1)$* .

Recalling the discussion of the Uniform Boundedness Conjecture in §5, we can state the result in [18] (as improved by Oesterlé) as follows: *for  $\ell > (3^{d/2} + 1)^2$  the curve  $X_1(\ell)$  has no noncusp  $K$ -points for any degree- $d$  number field  $K$* . In order to appreciate the power of the UBC, the reader should notice the contrast with the family of Fermat curves. Already when  $d = 2$ , very little is known about points on high-degree Fermat curves with coordinates in quadratic number fields.

The curves  $X_0(\ell)$  and  $X_1(\ell)$  become quite complicated as  $\ell$  increases. One measure of the complexity of a curve is its genus: a rational curve has genus 0, an elliptic curve has genus 1, and a hyperelliptic curve of the form  $y^2 = x^{2g+1} + a_{2g}x^{2g} + \cdots + a_1x + a_0$  has genus  $g$ . The complex points of a genus- $g$  curve form

<sup>3</sup>In Figure 3 the fundamental domain has cusps at 0,  $-1$ , and  $\infty$ .

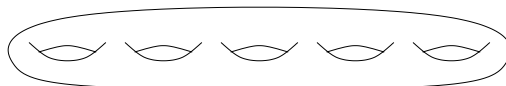


FIGURE 4. A Riemann surface with five handles.

a so-called *Riemann surface* with  $g$  “holes” or “handles.” In Figure 1 we saw a genus-1 curve (an elliptic curve). The complex points of a genus-5 curve are shown in Figure 4.

The exact formulas for the genus of modular curves are a little complicated (see [19]). Asymptotically the genus of  $X_0(\ell)$  is  $\sim \frac{1}{12}\ell$ , and the genus of  $X_1(\ell)$  is  $\sim \frac{1}{24}\ell^2$ . Thus, if  $\ell$  is a 160-bit prime, the curve  $X_1(\ell)$ , considered over the complex numbers, looks like a Riemann surface with about  $2^{315}$  handles! It is this daunting object that would have to have a point of reasonable size if the elliptic curve in the Cheng-Uchiyama algorithm existed.

7. THE EXAMPLES  $\ell = 5, 7, 11$

When  $\ell = 5$  or  $7$ , the modular curve  $X_1(\ell)$  is a rational curve. That is, all elliptic curves over a number field  $K$  with a point of order  $\ell$  in  $E_{\text{tors}}(K)$  can be parameterized by a single variable  $t \in K$ . In each case an explicit equation of the elliptic curve corresponding to  $t$  is given by Silverman. Namely, for  $\ell = 5$  we have (see [22], p. 278 and [14])

$$(5) \quad y^2 + (t + 1)xy + ty = x^3 + tx^2, \quad \Delta = -t^5(t^2 + 11t - 1).$$

For  $\ell = 7$  we have (see [21], p. 223 and [14])

$$(6) \quad y^2 + (1 + t - t^2)xy + (t^2 - t^3)y = x^3 + (t^2 - t^3)x^2, \\ \Delta = t^7(t - 1)^7(t^3 - 8t^2 + 5t + 1).$$

When  $\ell = 11$ , the modular curve  $X_1(11)$  is itself an elliptic curve. It has a very simple equation:

$$(7) \quad s^2 - s = t^3 - t^2.$$

In other words, all elliptic curves over a number field  $K$  with a point of order 11 in  $E_{\text{tors}}(K)$  can be parameterized by two variables  $t, s \in K$  satisfying (7). An equation of the elliptic curve corresponding to a point  $(t, s)$  on the curve (7) is given in [22], p. 279:

$$(8) \quad y^2 + (st + t - s^2)xy + s(s - 1)(s - t)t^2y = x^3 + s(s - 1)(s - t)tx^2.$$

On each curve (5), (6), (8) the point  $P = (0, 0)$  is an  $\ell$ -torsion point.

Since the  $X_1(11)$  case is more typical than the cases  $\ell = 5, 7$  (where the modular curve is rational), we decided to compute the discriminant of the curve (8) as a polynomial in  $t$  and  $s$  and then examine its norm from  $\mathbb{Q}(t, s)$  to  $\mathbb{Q}(t)$ . Note that any polynomial in  $t$  and  $s$  can be reduced to the form  $a(t) + sb(t)$  using (7), and the norm of this element is  $(a(t) + sb(t))(a(t) + (1 - s)b(t)) = a(t)^2 + a(t)b(t) + (t^2 - t^3)b(t)^2$ . We found that the norm of the discriminant factors as a power of  $t$  times a power of  $t - 1$  times an irreducible quintic:

$$\text{Norm}(\Delta) = -t^{34}(t - 1)^{22}(t^5 - 18t^4 + 35t^3 - 16t^2 - 2t + 1).$$

8. COEFFICIENTS, DISCRIMINANTS, COORDINATES —  
HOW BIG ARE THEY?

We first prove that there is a close relation between the size of the coordinates of the torsion points and the size of the coefficients and discriminant in (1).

**Theorem 1.** *Let  $E$  be an elliptic curve with equation (1). If  $E(K)$  has a point of prime order  $\ell$ , then the bitlength of the discriminant of  $E$  is bounded by a constant times the maximal bitlength of the coordinates of the torsion points. In the other direction, if  $\ell > d + 1$  (where  $d = [K : \mathbb{Q}]$ ), then the norm from  $K$  to  $\mathbb{Q}$  of the  $y$ -coordinate of a torsion point is an integer bounded by the square root of the norm of the discriminant.*

*Proof.* Let  $P_1 = (x_1, y_1)$  and  $P_2 = (x_2, y_2)$  be two torsion points such that  $P_1 \neq \pm P_2$ , that is,  $x_1 \neq x_2$ . We have  $y_1^2 = x_1^3 + ax_1 + b$  and  $y_2^2 = x_2^3 + ax_2 + b$ . Subtracting, we immediately solve for  $a$  and  $b$ :

$$a = \frac{y_1^2 - y_2^2}{x_1 - x_2} - (x_1^2 + x_1x_2 + x_2^2), \quad b = y_1^2 - x_1^3 - ax_1.$$

Since  $\Delta = 4a^3 + 27b^2$ , this proves the first assertion. We remark that a similar proof would apply for  $\ell$  large if the equation of the curve were given in the more general form (2) with coefficients  $a_1, a_2, a_3, a_4, a_6$ . We would just have to choose five torsion points  $P_1, P_2, P_3, P_4, P_5$  in general position (that is, no three on a line and  $P_i \neq \pm P_j$ ) and solve a  $5 \times 5$  linear system to express the  $a_i$  in terms of polynomials in the coordinates of the  $P_i$ .

The second assertion in the theorem follows from Exercise 8.11(a) of [21]. In the notation there,  $\nu$  denotes the valuation corresponding to a prime ideal of  $K$ , and  $r_\nu = \left\lceil \frac{1}{\ell-1} \text{ord}_\nu(\ell) \right\rceil$  (here  $\lceil \cdot \rceil$  is the greatest integer function). The assumption  $d < \ell - 1$  implies that  $r_\nu = 0$ , and then the exercise gives for nonzero  $y$ :  $0 \leq \text{ord}_\nu(y^2) \leq \text{ord}_\nu(\Delta)$ . Since this holds for any prime ideal of  $K$ , we conclude that  $y$  is an algebraic integer and  $|\text{Norm}(y^2)| \leq |\text{Norm}(\Delta)|$ , as claimed.  $\square$

We remark that, roughly speaking, the bitlength of the norm of an element  $y \in K$  is at most  $d$  times the bitlength of  $y$ . (A more precise statement would have to take into account the bitlength of the basis elements of  $K$  over  $\mathbb{Q}$  in terms of which  $y$  is written as a  $d$ -tuple.) But it is possible for the norm to have much smaller bitlength. For example, in the field  $K = \mathbb{Q}(\sqrt{2})$  an element  $a + b\sqrt{2}$  has norm  $a^2 - 2b^2$ , which has bitlength bounded by  $1 + 2 \max(\text{bitlength}(a), \text{bitlength}(b))$ . On the other hand, a large power of  $1 + \sqrt{2}$  has large bitlength, but its norm is  $\pm 1$ . Thus, we cannot conclude from Theorem 1 that the bitlength of  $y$  is polynomially bounded whenever the bitlength of the discriminant is.

We claim that the discriminant of the elliptic curve is likely to be at least exponentially large. In the cases  $\ell = 5, 7, 11$  (see §7) we already see the trend toward large discriminants. To find a curve with the properties needed in [3], we must choose  $t$  to be a root modulo  $p$  (more precisely, modulo a prime ideal of  $K$  dividing  $p$ ) of the irreducible quadratic (for  $\ell = 5$ ), cubic (for  $\ell = 7$ ), or quintic (for  $\ell = 11$ ) that appears in the factorization of  $\Delta$  or  $\text{Norm}(\Delta)$ . (Although the discriminant vanishes mod  $p$  when  $t \equiv 0$  or  $1 \pmod{p}$ , those values cannot be used, because the torsion subgroup would collapse into the singular point, as we see from the formulas for  $iP$  given below.) The discriminant of the curve is of order  $t^7$  for  $\ell = 5$ ,  $t^{17}$  for

$\ell = 7$ , and  $t^{30.5}$  for  $\ell = 11$  (here we take the square root of the norm, in view of the remark following the proof of Theorem 1). The powers of  $t$  seem to be growing at least linearly, and perhaps quadratically, with  $\ell$ .

There are two other heuristic arguments that suggest that the discriminant  $\Delta$  grows rapidly with  $\ell$ . First, as mentioned in §5, Hindry and Silverman [8] proved that if the number of prime divisors of  $\Delta$  is bounded, then  $\ell = O(d \log d)$  (with the constant in the big- $O$  depending on the bound). Thus, if  $\ell$  grows much faster than polynomially in  $d$ , as assumed in [3], it is likely that the number of prime divisors of  $\Delta$  would grow rapidly.

In addition, the proof of the second part of Theorem 1 shows that  $\text{Norm}(\Delta)$  is divisible by the square of the norm of the  $y$ -coordinate of the torsion point  $iP = (x_i, y_i)$  for all  $\ell - 1$  values of  $y_i$ ,  $i = 1, \dots, \ell - 1$  (unless  $y_i = 0$ ). This does not, of course, imply that  $\text{Norm}(\Delta)$  has  $\ell - 1$  distinct divisors. It is conceivable, for instance, that  $y_i$  and  $y_j$  frequently differ by a factor that is a unit of  $K$ , in which case they have the same norm. We have very limited data on this question, but what we have seems to indicate that  $|\text{Norm}(y_i)|$  and  $|\text{Norm}(y_j)|$  rarely coincide. In the three cases  $\ell = 5, 7, 11$  where we computed the multiples of  $P$  we have

$$P = (0, 0), 2P = (-t, t^2), 3P = (-t, 0), 4P = (0, -t), 5P = O$$

on the curve (5) of §7;

$$P = (0, 0), 2P = (t^2(t - 1), t^3(t - 1)^2), 3P = (t(t - 1), t(t - 1)^2), \\ 4P = (t(t - 1), t^2(t - 1)^2), 5P = (t^2(t - 1), 0), 6P = (0, t^2(t - 1)), 7P = O$$

on the curve (6); and

$$P = (0, 0), 2P = (t^3(t - 1)(t - s), -st^3(t - 1)(t - s)^2), \\ 3P = (-st(t - s), s^2t^2(t - s)), 4P = (t^4(t - 1), t^6(t - 1)^2), \\ 5P = (t^2(t - 1)(t - s), t^2(t - 1)^2(t - s)^2), 6P = (t^2(t - 1)(t - s), t^4(t - 1)^2(t - s)), \\ 7P = (t^4(t - 1), -st^5(t - 1)), 8P = (-st(t - s), s^2t(t - s)^2), \\ 9P = (t^3(t - 1)(t - s), 0), 10P = (0, t^4(t - 1)(t - s)), 11P = O$$

on the curve (8). In the case  $\ell = 11$  the norms of the  $y$ -coordinates of  $iP$ ,  $1 \leq i \leq 10$ , are

$$0, -t^{10}(t - 1)^7, -t^9(t - 1)^4, t^{12}(t - 1)^4, t^6(t - 1)^8, \\ -t^9(t - 1)^6, -t^{12}(t - 1)^3, t^8(t - 1)^6, 0, -t^9(t - 1)^4.$$

For  $\ell = 5, 7, 11$  we see that, considered as polynomials in  $\mathbb{Q}[t]$ , most of the elements  $\text{Norm}(y_i^2)$  give distinct divisors of  $\text{Norm}(\Delta)$ .

### 9. CONCLUSION

Our analysis of the attempted attack in [3] shows once again that one has to be very careful about using elliptic curves over number fields  $K$ . Any curve that has enough points of the desired sort is likely to be computationally intractable. In the present case the attackers needed a large number of torsion points. In earlier cases, such as the algorithm analyzed in [9], the cryptanalysts needed high rank. Experience over the last few years has given reason for skepticism about the possibility of using the structure of the group of  $K$ -points of an elliptic curve to mount a successful assault on a cryptosystem.

Conditional results in cryptography are of various types. Some are proved rigorously, except in one place where one needs a widely believed and heuristically justified mathematical conjecture (e.g., that there is a  $1/(\log p)$  probability of primality of a number in the Hasse interval around  $p$ ). Another sort of conditional result is based on an intractability assumption (e.g., that factoring is hard). A third type of conditional theorem in cryptography uses an assumed property of a primitive (e.g., randomness of a hash function). Finally, there is the type of conditional result that is based on a conjecture for which the only justification is that it has not yet been proven to be false: “Mathematicians are unable to prove that X does not exist, so I’ll conjecture that X does exist.” If cryptographers start accepting “results” based on assumptions of the last type, then we risk losing credibility.

In some of the nonsciences it is common for mathematics to be used in bizarre ways that would horrify a mathematically educated person. (A famous example of this from political “science” is discussed in [13].) In contrast, in any self-respecting branch of science, researchers are expected to adhere to high standards of reasonableness in their use of mathematics. Thus, when reading a paper in cryptography that depends upon a basic mathematical assumption, we should insist on seeing convincing evidence that that assumption is likely to hold. Of course, it is natural for cryptographers to be intrigued when a paper advertises itself as “the first attempt to apply the . . . Uniform Boundedness Conjecture in cryptography.” However, we must not forget that if the key assumption is mathematically implausible, then the claimed result cannot be accepted as scientifically valid.

#### REFERENCES

1. B. den Boer, Diffie-Hellman is as strong as discrete log for certain primes, *Advances in Cryptology — CRYPTO '88*, Lecture Notes in Computer Science **403** (1990), Springer-Verlag, 530-539.
2. D. Boneh, The decision Diffie-Hellman problem, *Algorithmic Number Theory, Proc. Third Intern. Symp., ANTS-III*, Lecture Notes in Computer Science **1423** (1998), Springer-Verlag, 48-63. MR **2000k**:94024
3. Q. Cheng, S. Uchiyama, Nonuniform polynomial time algorithm to solve decisional Diffie-Hellman problem in finite fields under conjecture, *Topics in Cryptology — CR-RSA 2002*, Lecture Notes in Computer Science **2271** (2002), Springer-Verlag, 290-299.
4. R. Cramer, V. Shoup, A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack, *Advances in Cryptology — CRYPTO '98*, Lecture Notes in Computer Science **1462** (1998), Springer-Verlag, 13-25. MR **99j**:94041
5. S. David, Pointes de petites hauteurs sur les courbes elliptiques, *J. Number Theory* **64** (1997), 104-129. MR **98k**:11067
6. T. ElGamal, A public key cryptosystem and a signature scheme based on discrete logarithms, *IEEE Trans. Information Theory* **31** (1985), 469-472. MR **86j**:94045
7. M. Flexor, J. Oesterlé, Sur les points de torsion des courbes elliptiques, *Astérisque* **183** (1990), 25-36. MR **91g**:11057
8. M. Hindry, J. Silverman, Sur le nombre de points de torsion rationnels sur une courbe elliptique, *C. R. Acad. Sci. Paris* **329** (1999), 97-100. MR **2000g**:11047
9. M. Jacobson, N. Koblitz, J. Silverman, A. Stein, E. Teske, Analysis of the xedni calculus attack, *Designs, Codes and Cryptography* **20** (2000), 41-64. MR **2001b**:14043
10. A. Joux, A one round protocol for tripartite Diffie-Hellman, *Algorithmic Number Theory, Proc. Third Intern. Symp., ANTS-IV*, Lecture Notes in Computer Science **1838** (2000), Springer-Verlag, 385-393. MR **2002i**:14029
11. A. Joux, K. Nguyen, Separating decision Diffie-Hellman from Diffie-Hellman in cryptographic groups, preprint, 2001.
12. S. Kamienny, Torsion points on elliptic curves and  $q$ -coefficients of modular forms, *Invent. Math.* **109** (1992), 221-229. MR **93h**:11054

13. N. Koblitz, A tale of three equations; or the emperors have no clothes, *The Mathematical Intelligencer* **10** (1988), 4-11; and: Reply to unclad emperors, *ibid.*, 14-16.
14. D. Kubert, Universal bounds on the torsion of elliptic curves, *Proc. London Math. Soc.* (3) **33** (1976), 193-237. MR **55**:7910
15. D. Masser, Counting small points on elliptic curves, *Bull. Soc. Math. France* **117** (1989), 247-265. MR **90k**:11068
16. U. Maurer, Towards the equivalence of breaking the Diffie-Hellman protocol and computing discrete logarithms, *Advances in Cryptology — CRYPTO '94*, Lecture Notes in Computer Science **839** (1994), Springer-Verlag, 271-281. MR **95k**:94021
17. B. Mazur, Modular curves and the Eisenstein ideal, *Publ. Math. Inst. Hautes Études Sci.* **47** (1978), 33-186. MR **80c**:14015
18. L. Merel, Borne pour la torsion des courbes elliptiques sur les corps des nombres, *Invent. Math.* **124** (1996), 437-449. MR **96i**:11057
19. G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton Univ. Press, 1971. MR **47**:3318
20. V. Shoup, "Lower bounds for discrete logarithms and related problems", *Advances in Cryptology — EUROCRYPT '97*, Lecture Notes in Computer Science **1233** (1997), Springer-Verlag, 256-266. MR **98j**:94023
21. J. Silverman, *The Arithmetic of Elliptic Curves*, Springer-Verlag, 1986. MR **87g**:11070
22. J. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer-Verlag, 1994. MR **96b**:11074
23. J. Silverman, personal communication, 24 November 2001.
24. E. Verheul, Evidence that XTR is more secure than supersingular elliptic curve cryptosystems, *Advances in Cryptology — EUROCRYPT 2001*, Lecture Notes in Computer Science **2045** (2001), Springer-Verlag, 195-210. MR **2003b**:94062

DEPARTMENT OF MATHEMATICS, BOX 354350, UNIVERSITY OF WASHINGTON, SEATTLE, WASHINGTON 98195

*E-mail address:* [koblitz@math.washington.edu](mailto:koblitz@math.washington.edu)

DEPARTMENT OF COMBINATORICS AND OPTIMIZATION, UNIVERSITY OF WATERLOO, WATERLOO, ONTARIO N2L 3G1 CANADA

*E-mail address:* [ajmeneze@uwaterloo.ca](mailto:ajmeneze@uwaterloo.ca)