Elliptic Curves and Their Applications to Cryptography: An Introduction

Andreas Enge

September 1999

Contents

Foreword

Since the advent of public-key cryptography in 1976 by Diffie and Hellman many public-key schemes have been devised. Almost all have what are generally considered to be "hard" mathematical problems as the basis for their security. In particular, the integer factorization problem and the discrete logarithm problem are at the heart of several of the most well known techniques.

A few public-key technologies are now being widely deployed commercially to secure such activities as electronic payment over the Internet, stock trading from pagers and cell phones, and multi-applications on smart cards. Two of the more well-known methods are the RSA scheme and the DSA (digital signature algorithm). The former bases its security on integer factorization and the latter on the discrete logarithm problem in the multiplicative group of a finite field. For both of these problems there are subexponential time algorithms, which means in practice key sizes are forced to exceed 1000 bits to attain adequate security. For many constrained environments where power, storage and bandwidth are severely limited it becomes impossible to provide public-key cryptography through these methodologies.

In 1985 Neal Koblitz and Victor Miller independently proposed elliptic curve cryptography. The security of this scheme would rest on the difficulty of the discrete logarithm problem in the group formed from the points on an elliptic curve over a finite field. To date the best method for computing elliptic logarithms is fully exponential. This translates into much smaller key sizes permitting one to deploy public-key cryptography on devices where previously it was impossible. Over the past fourteen years elliptic curve cryptography has been gaining popularity and it is now being standardized around the world by agencies such as ANSI, IEEE and ISO. Recently, in January 1999, the elliptic curve version of the DSA (called the ECDSA) became an ANSI X9.62 standard for the US financial sector.

Elliptic curve cryptography relies on the elegant but deep theory of elliptic curves over finite fields. There are, to my knowledge, very few books which provide an elementary introduction to this theory and even fewer whose motivation is the application of this theory to cryptography. Andreas Enge has written a book which addresses these issues. He has developed the basic theory in a simple but thorough manner and in an easily understandable style. I have used a preliminary version of this book from which to teach a senior undergraduate course on elliptic curve cryptography. I was so pleased with the outcome that I encouraged Andreas to publish the manuscript. I firmly believe that this book is a very good starting point for anyone who wants to pursue the theory of elliptic curves over finite fields and their applications to cryptography.

S. A. Vanstone, April 1999

Preface

During the last twenty years the invention of public key cryptosystems, in conjunction with the emerging computer technology, has opened new fields of applications for number theory and algebraic geometry, which were so far considered as the "purest" branches of mathematics. Elliptic curves are among the most promising tools in modern cryptography. This has raised new interest in the topic not only within the mathematical community, but also on the part of engineers and computer scientists, who are concerned with the implementation of new cryptosystems.

My aim is to present a textbook for those who find it hard to learn about elliptic curves from the more advanced treatments, and thus to lay the foundations for studying these more complete references. To follow this book, only undergraduate algebra is needed; the reader should basically have heard of polynomial rings, field extensions and finite fields. Even this elementary approach will eventually guide us towards questions at the forefront of current research, like the problem of counting points on elliptic curves, which was satisfactorily solved only a few years ago.

While many of the fascinating applications of elliptic curves like the factorisation of integers or primality proofs deal with curves over prime fields only, curves over fields of characteristic 2 are especially attractive in the cryptographic context. This textbook treats curves in odd and even characteristic with equal attention, referring to general arguments where possible and falling back on case distinctions where necessary.

I am grateful to Reinhard Schertz, whose enthusiastic undergraduate lectures raised my interest in elliptic curves, to Dieter Jungnickel, who suggested the topic and marvelously supervised the advance of my thesis, from which this book finally emerged, and to Leonard Charlap and David Robbins, whose excellent report on elliptic curves formed the basis for my presentation. I thank Marialuisa de Resmini and Scott Vanstone for encouraging the publication. And I am especially indebted to Dirk Hachenberger, Dieter Jungnickel, Charles Lam and Berit Skjernaa for the time they spent reading the manuscript and for their valuable comments.

I hope that the reader has as much pleasure in reading this book as I had in writing it.

Andreas Enge

Chapter 1

Public Key Cryptography

Today's widespread use of electronic networks in the economic world has raised cryptography from a speciality of the military and secret services to a topic of public interest, which concerns international organisations like the UNO and the EU. Unlike conventional cryptosystems, public key cryptography is applicable on a large scale base, in principle allowing secure and authorised communication between any two persons in the world. In the following chapter we give a brief introduction to the concepts of public key cryptography and present some algorithms. We hereby focus on schemes for encryption and digital signatures which can be generalised to arbitrary groups, especially to elliptic curve groups. A comprehensive treatment of cryptographic issues is given in Stinson (1995) and Menezes, van Oorschoot, Vanstone (1997).

Chapter 2

The Group Law on Elliptic Curves

Elliptic curves can be equipped with an efficiently computable group law, so that they are suited for implementing the cryptographic schemes of the previous chapter, as suggested first in Koblitz (1987) and Miller (1986). They are particularly appealing because they achieve the same level of security as a finite field based cryptosystem with much shorter key lengths, which results in a faster encryption and decryption process. Our aim in this chapter is to prove the group law.

After presenting the necessary definitions we show that there is an intuitive geometric composition law on an elliptic curve, involving lines and their intersection points with the curve. Some elementary computations result in simple algebraic formulae which are suited for computer implementations. The composition law fulfils all group axioms, but strange enough, its associativity is hard to prove. It can be shown in various ways:

The obvious approach is brute force computation, the explicit algebraic formulae for adding two points on a curve being given. Unfortunately there are several formulae, depending on the position of the points to be added, and so an awful lot of case distinctions is needed. What is worse, the proof does not reveal anything about the underlying algebraic and geometric structures and is not only extremely tedious, but also extremely uninstructive. This seems to have deterred most authors, for, to my knowledge, this approach cannot be found in any publication.

Some authors concentrate on elliptic curves over the complex numbers, where the additional analytic structure accounts for particular properties, see Koblitz (1993), Lang (1978) or Lang (1987). But for implementational reasons we are mainly interested in curves over finite fields, to which the analytic proofs do not apply. Hence in this book we concentrate on purely algebraic approaches, which work over any field. It is instructive, however, to relate our algebraic findings to their analytic counterparts, and the reader is invited to take a closer look at the books mentioned above.

Fulton presents a beautiful geometric proof in his book on algebraic curves (see Fulton (1969), p. 125) after developing some general theory. The same proof is reported in Husemöller (1987), Chapter 3. Other arguments use the

Riemann–Roch theorem, which is presented at the end of Fulton's book. These approaches are ideal for specialists in algebraic geometry, in which case the standard references are Silverman (1986) and Silverman (1994). However, elliptic curves are still quite "simple" from the algebraic-geometric point of view and can be understood without knowing much of abstract algebraic geometry.

In this chapter we follow Charlap's and Robbin's elementary proof (1988). On one hand, we explain the basic notions of the theory of algebraic curves, so that the reader gets an introduction to this topic. On the other hand, it is our aim to keep this exposition as elementary and concrete as possible. So we specialise all results to the case of elliptic curves, where many of them can be proved by explicit computations or more elementary arguments than in the general case. Unlike Charlap and Robbins we consistently use the projective point of view when working with the infinite point $\mathcal{O}$, which appears naturally in this setting, and thus avoid seemingly artificial constructions. Furthermore we present a generalised version of the proof, which covers fields of any characteristic, including the case of characteristic 2, which is highly relevant for cryptography.

Chapter 3

Elliptic Curves over Finite Fields

We have verified in the previous chapter that the points on an elliptic curve over an arbitrary field form a group, which can be used to implement the public key cryptosystems presented in the first chapter. Since by the algebraic formulae the group operations eventually amount to computations in the field where the elliptic curve is defined, one has to choose a field with an efficiently implementable arithmetic. Basically, this requirement narrows down to the finite fields. (While the rational numbers and more generally number fields also allow exact computations, they have two drawbacks: First, numbers may become arbitrarily big, which destroys the efficiency of the operations. And more important, the discrete logarithm problem on elliptic curves over these fields is easy to solve.) So during this chapter, we consider the following situation:

Let $k = F_q$ be the finite field with q elements and prime characteristic p, and $K = \overline{k}$ be its algebraic closure. Let E be an elliptic curve which is defined over k, i.e. whose defining coefficients a_1, a_3, a_2, a_4 and a_6 lie in k. As before we denote by E the group of points on the curve with coordinates in K. The group of k-rational points, i.e. the group of points on E with coordinates in k, in which we are eventually interested, is denoted by E_k. Since k is finite, there are only finitely many possibilities for the X- and Y-coordinates of points, and E_k is a finite abelian group. We will see in Chapter 4 that it is mainly the exact cardinality of E_k which determines the security of a cryptosystem built in this group. The biggest part of this chapter is devoted to the proof of Hasse's famous theorem, which gives an estimate on the cardinality of E_k, stating that the elliptic curve group has roughly as many elements as k itself. Precisely,

$$|q + 1 - |E_{F_q}|| \leq 2\sqrt{q}.$$

We follow very closely the excellent report Charlap, Robbins (1988), occasionally putting a different emphasis. However, we take care to prove all theorems for characteristic 2 as well.

While the results of the first seven sections hold in full generality for any finite or infinite field k, we will apply them to finite fields only: namely to prove Hasse's Theorem in Section 3.8 and to compute the exact cardinality of E_k in Chapter 5. In the last sections of this chapter we present some results on special classes of elliptic curves over finite fields and on the group structure of E_k.

Chapter 4

The Discrete Logarithm Problem

The public key cryptosystems presented in Chapter 1 rely on the difficulty of solving the discrete logarithm problem in certain groups: An adversary who could efficiently compute discrete logarithms in the group underlying such a cryptosystem would be able to break the system. So to judge the security of the proposed cryptosystems we must have a closer look at algorithms for solving discrete logarithm problems.

To provide a common framework for the following sections, we reformulate the problem and fix some notations: Let $G = \langle \alpha \rangle$ be a finite, multiplicatively written cyclic group with generator α and known cardinality n and let β be an element of G. The discrete logarithm problem is to compute an integer l (which is denoted by $\log_\alpha \beta$) such that $\beta = \alpha^l$. The integer l is determined uniquely modulo n. The problem can on one hand be solved by a *generic* or *black box algorithm*, which does not take into account the representation of group elements. We only require that it be possible to efficiently multiply and invert group elements and to test them for equality. We then solve the problem by these elementary operations, starting with the given elements α and β. The third requirement may seem surprising since in most groups it is easy to test whether two elements are equal; but it can be an issue in factor groups, which are given modulo an equivalence relation, so that the same element may have different representatives. An example is provided by the divisor class group of an elliptic curve, where the problem is solved by working with the unique representatives given by single points. Other examples are class groups of number fields or divisor class groups of more general curves, in which cases this issue is more serious.

It turns out, however, that the difficulty of the discrete logarithm problem depends heavily on the representation of the group. For instance, it is trivial for $G = Z_n$ and $\alpha = 1$. More generally it is easy to solve for $G = Z_n$ and any generator α of Z_n by the Euclidian algorithm. (Indeed, the discrete logarithm problem in G can be reformulated to the task of computing an explicit isomorphism of G with Z_n.) Hence it is worthwhile to take the concrete representation of the group into account when looking for an efficient solution to the discrete logarithm problem. We will see in Section 4.4, for instance, that there are especially good algorithms for the multiplicative groups of finite fields. Some elliptic curves are also cryptographically insecure, which we will show in Section 4.5, using the preparations made in the previous chapter.

Chapter 5

Counting Points on Elliptic Curves

We have seen in the previous chapter that the security of a discrete logarithm based cryptosystem relies mainly on the order of the underlying group, unless special structures allow more efficient algorithms for breaking the system. If the group order is large enough, then square root attacks like Shanks's baby-step giant-step or Pollard's ρ-methods are not applicable. To make the Pohlig–Hellman attack impractical, two different approaches are conceivable.

On one hand, it is possible to choose a group with unknown order, so that the Pohlig–Hellman algorithm does not work. This is a risky game, however, since for no known type of groups there is a theoretical barrier to compute their orders. For instance, the problem is not known to be NP-complete for any class of groups. Hence, there is a certain chance that an adversary already has an algorithm at hands for determining the group order. Moreover, while this attitude allows to encrypt messages, the signature algorithms of Chapter 1 require that the group order be known.

On the other hand, it is a good strategy to make sure that the group order contains a large prime factor to prevent the Pohlig–Hellman attack. In the case of elliptic curves this can be achieved in various ways. First, by the complex multiplication method, curves with suitable orders can be designed specifically (Atkin, Morain (1993) and Lay, Zimmer (1994)). Second, it is possible to choose special classes of curves whose cardinalities are easy to determine, like supersingular curves (cf. Theorem 3.72), or curves which are defined over a small field, but where the group is chosen over a field extension (cf. Theorem 3.66). While supersingular curves are not recommendable according to Section 4.5, nothing can so far be hold against curves defined over subfields. However, there is a certain reluctance concerning classes of special curves and a widespread belief that the most secure way of selecting a curve is to fix an underlying field, randomly choose a curve, i.e. defining coefficients, and compute the group order until it is divisible by a large prime. This approach is feasible today due to the algorithmic progress made in the past fifteen years.

Index

Errata

The following typos and errors have been found by Wu Ting. Many thanks!

- On p. 15, l. -7, the a_3 in the equation of E should be an a_3Y.
- On p. 31, l. -12, $\mathcal{O}_{\varphi(P)}(E)$ should read $\mathcal{O}_{\varphi(P)}(E')$.
- On p. 35, l. 5, the divisor of the line should be given as

  $$\operatorname{div}(l^*) = \langle P \rangle + \langle \overline{P} \rangle - 2\langle \mathcal{O} \rangle.$$

- On p. 64, l. 1, it is erroneously deduced from $\deg u = 0$ that u is a constant. However, u is a rational function in X and not a polynomial, so the only thing one can say is that $u = \frac{f}{g}$ for polynomials f and g of the same degree n. Then one has $Du = \frac{f'g - fg'}{g^2} DX$. Notice that $\deg(f'g - fg') \leq 2n - 2$: If $p \nmid n$, then $\deg(f'g) = \deg(fg') = 2n - 1$, and the leading terms coincide; if $p | n$, then already $\deg(f'g)$, $\deg(fg') \leq 2n - 2$. Hence $\operatorname{ord}_{\mathcal{O}} \left(\frac{f'g - fg'}{g^2} \right) \geq 4$. Since $\operatorname{ord}_{\mathcal{O}} DX \geq -3$, we obtain $\operatorname{ord}_{\mathcal{O}}(Du) \geq 1$. As before one shows that $\operatorname{ord}_{\mathcal{O}} D(vY) \geq 0$ and concludes that $\operatorname{ord}_{\mathcal{O}} Dr \geq \min\{\operatorname{ord}_{\mathcal{O}} Du, \operatorname{ord}_{\mathcal{O}} D(vY)\} \geq 0$.

  This argumentation makes the proof a bit twisted, and it would be simpler to put $r = \frac{f + gY}{h}$ with polynomials f, g and h right from the beginning. Then $d = \operatorname{ord}_{\mathcal{O}} r = 0$ implies $\deg f = \deg h = n$ for some n and $\deg g \leq n - 2$. One computes

  $$Dr = \frac{(f'h - fh')DX + (g'h - gh')YDX + ghDY}{h^2}.$$

  As above, one concludes that $\operatorname{ord}_{\mathcal{O}} \left(\frac{(f'h - fh')DX}{h^2} \right) \geq 1$; similarly, $\deg(g'h - gh') \leq 2n - 3$ implies that $\operatorname{ord}_{\mathcal{O}} \left(\frac{(g'h - gh')}{h^2} YDX \right) \geq 6 - 3 - 3 = 0$; and we finally have $\deg \frac{ghDY}{h^2} \leq 0$, so that the order of this term in $\mathcal{O}$ is also non-negative.

- On p. 73, the case corresponding to the second square dot should be $(m-1)^3 \equiv 1 \pmod{p}$; this is only a typo, the proof itself is correct.
- On p. 84, all the sums over various $X - X(P)$ in expressions for ψ_m should be products.

And some further errors detected by an attentive reader.

- On p. 87, and sub- and consequently in Schoof's algorithm on pp. 135–138, all occurrences of $3X^2 + 2a_2X + a_4 - a_1Y$ should be garnished with the opposite sign.