Parabola Volume 54, Issue 1 (2018)

Elliptic Curves and an Application in Cryptography

Jeremy Muskat[1]

Abstract

Communication is no longer private, but rather a publicly broadcast signal for the entire world to overhear. Cryptography has taken on the responsibility of securing our private information, preventing messages from being tampered with, and authenticating the author of a message. Since the 1970s, the burden of securing communication has largely rested on the RSA algorithm. Over time, techniques to crack RSA have improved which has forced key sizes to increase in order to maintain security. The size of keys is approaching infeasible, and one possible alternative to RSA uses elliptic curves which have significantly smaller keys.

1 Introduction

Examples of cryptography can be found throughout history as far back as the Ancient Greeks. Cryptography is used when there is a need to secure valuable information. The digital revolution has resulted in half of the world communicating through computers and the Internet [1]. As a result, the primary users of cryptography have shifted from governments to anyone who uses the Internet. As of 2017, Chrome 56 actively warns users when an Internet site is not secure [8]. Google's goal is to have the entire web transition to secure sites. The advent and advances in public key cryptography make this possible.

Question 1. *Investigate different encryption schemes in use on the Internet. You know that you are on a secure website if the URL address begins with HTTPS versus HTTP. Many browsers have a padlock icon to visually represent a secure connection. Click on the padlock and investigate the details of the security being used. Find a website that uses RSA and one that uses ECC encryption.*

Question 1 provides evidence that RSA is the most widely used public key encryption scheme on the Internet. RSA was developed in the late 1970s, and its security is based on the idea that factoring a number into primes is a "hard" problem. Because of the RSA algorithm, research in factoring algorithms increased. As factoring techniques improve, the security behind RSA becomes compromised. The result is that the size of the keys used in RSA have increased exponentially. If this trend continues, then the larger key sizes will result in computations that are infeasible for modern computers.

[1]Jeremy Muskat is an Associate Professor of Mathematics at Western State Colorado University, USA.

Definition 1 (Key)**.** A *cryptographic key* is a set of rules for changing *plain text* into encrypted *cipher text*. Keys are also used to decrypt cipher text back into readable plain text. In the latter usage, the terminology "key" can be thought of as unlocking the encrypted information.

Elliptic curves were introduced in cryptography as a tool used to factor composite numbers in an effort to crack RSA [6]. The consideration of elliptic curves in cryptography eventually led to a suggestion in the 1980s that they could also be used for encryption [5, 7]. The benefit of Elliptic Curve Cryptography (ECC) is that the key sizes are significantly smaller than the key sizes required for RSA of comparable security level. A comparison is shown in Table 1.

RSA	ECC
1024	160
2048	224
3072	256
7680	384
15360	512

Table 1: Bit lengths of keys for RSA and ECC with comparable levels of security

2 A crash course in cryptography

Definition 2 (Cryptography)**.** *Cryptography* is the study of storing and transmitting data in a format that only those for whom it is intended can read and process.

Caesar's cipher is a nice example to get started with the basic concepts in cryptography. Caesar's cipher is a substitution cipher that shifts the standard alphabet by three places. Table 2 shows a symmetric key that can be used to encrypt plain text into cipher text and decrypt cipher text into plain text by reversing the process.

Plain text	A	B	C	D	E	F	G	H	I	J	K	$\cdots$	Y	Z
Cipher text	X	Y	Z	A	B	C	D	E	F	G	H	$\cdots$	V	W

Table 2: Caesar's cipher symmetric key

Question 2. *Use Table 2 to decrypt the cipher text Mxoxylix into plain text.*

It is worth identifying some problems with the simple encryption scheme in Question 2. First, it is easy to crack. If the cipher text is long, or the same key is repeatedly used, then patterns can be identified using frequency analysis. Good encryption can overcome this shortcoming by having the ability to generate multiple keys easily.

Second, in order to decrypt (as in Question 2), the key must be transmitted along with the cipher text. This requires an extra transmission or a pre-established key between parties. Good encryption should be concerned with generating keys, but also should consider how to easily transmit the key. We will eventually describe the Elliptic Curve Diffie-Hellman (ECDH) key exchange, an algorithm that uses elliptic curves to transmit a key over an unsecured channel. ECDH key exchange establishes a private key between two parties even if a third party adversary is observing their communication.

3 Security behind RSA

Caesar's cipher provides evidence that encryption and decryption are straightforward as long as a key is provided. A description of the RSA algorithm can be found in [2]. Instead of the algorithm, we focus on the security that protects RSA and use it to motivate the introduction of elliptic curves. All encryption is based on "hard" problems.

Definition 3 (One-way function)**.** A *one-way function* $f(x)$ is a function that is easy to compute in the forward direction, yet difficult to compute in the other. To clarify, the output $f(x)$ of the function is easily determined from an input x. However, if given a function value $f(x)$, then it is hard to determine the value x.

The security of RSA is based on the following one-way function. Let p and q be prime numbers greater than 2. Define $f(p,q) = pq$ which is just usual multiplication. When considering f in a cryptographic scheme, p and q are approximately a hundred digits each. Attempting to answer Question 3 should clarify the idea behind f being a one-way function.

Question 3. *Use any method available. A computer algebra system like Mathematica is recommended.*

1. *Let* $p = 242596762305237077275763315697698246968 1$
   and $q = 6847944682037444681162770672798288913849$.
   Determine $f(p,q)$.

2. *Find* p *and* q *where* $n = f(p,q)$ *is equal to*

$$15226050279225333605356183781326374297180681149613806886579084945801229632589528976540003506920061 39\ .$$

The 100 digit composite number n in Question 3 is an example of a 330 bit key used for RSA encryption. Cracking the encryption corresponds to finding the factors p and q. The solution was given in 1991 and was the first solution in response to the now-defunct RSA factoring challenge [11]. The most recent solution was in May of 2016 and factored a 704 bit key [11].

1 Why "hard"?

Factoring is described as "hard" because there is empirical evidence of its difficulty. However, the difficulty cannot be proven. The theory behind factoring is simple; the difficulty lies in the length of the calculation. One approach to factoring a composite number n is to check all possible divisors less than or equal to $\lfloor\sqrt{n}\rfloor$.

Question 4. *Suppose that $n = f(p, q)$. Show that either p or q is less than or equal to $\lfloor\sqrt{n}\rfloor$.*

Question 4 provides a list that is guaranteed to contain a divisor of n. Consider the value of n from Question 3, and suppose that it is possible to check a million divisors per second. Completing the task of checking to see if $2|n$, $3|n$, $4|n$, $\ldots$, $\lfloor\sqrt{n}\rfloor|n$ would take 3.2×10^{37} years. Notice it doesn't help much to assume you could check divisors a million times faster.

There are however better factoring algorithms than the one described above [10]. Through the use of such algorithms along with parallel computing, we have seen a 768 bit key factored [11]. The result is an average of a 2048 bit key being used for RSA encryption on the Internet to guarantee security. A natural assumption is that factoring will continue to improve, resulting in a need for longer key bit lengths. The continued growth in key sizes is not sustainable, and one possible solution can be found using elliptic curves.

4 An introduction to elliptic curves

Definition 4 (Elliptic curves)**.** An *elliptic curve* over $\mathbb{R}$ is the set of points (x, y) satisfying an equation of the form $y^2 = x^3 + ax + b$ where $x^3 + ax + b$ has no double roots.

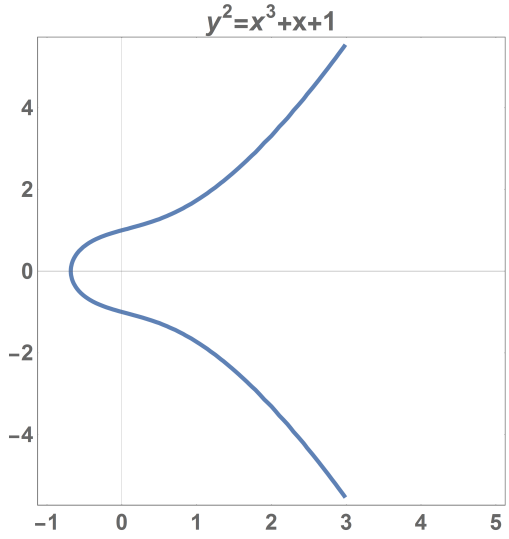

Figure 1: The elliptic curve $E : y^2 = x^3 + x + 1$ over $\mathbb{R}$

When an elliptic curve E is considered with one additional point $\mathcal{O}$, an operation can be defined on the set of points that behaves similarly to addition. The point $\mathcal{O}$ is a point at infinity and acts the same way that zero does when considering the addition of real numbers. The notion of $\mathcal{O}$ can be formalized but can be intuitively considered as a single point that lies at the top and bottom of every vertical line. Figure 2 provides a geometric interpretation of the operation on E. We will refer to the operation as "addition". To add two distinct points P and Q on E, first find the line that contains both P and Q. This line will intersect E at a third point denoted $P * Q$. Reflect $P * Q$ over the x-axis to obtain $P + Q$. Figure 2 depicts the special cases in which P and Q are not distinct and when the line between P and Q does not intersect E.

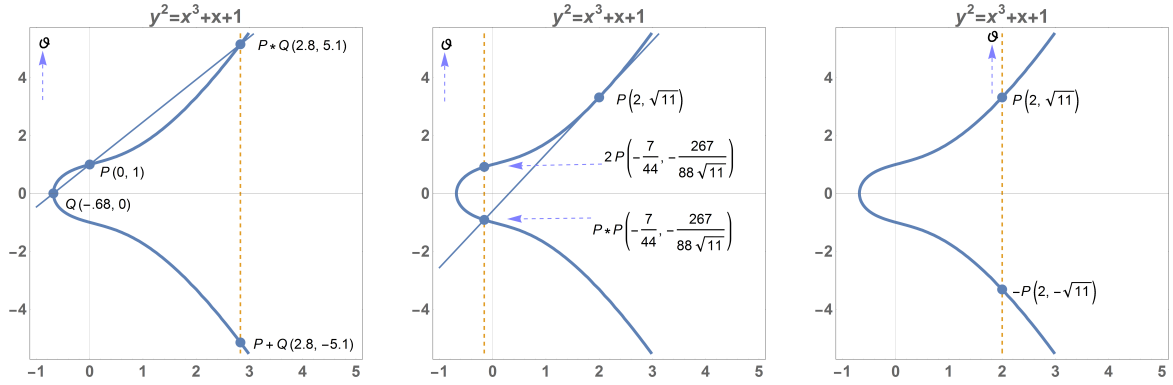

Figure 2: A geometric interpretation of addition on $E : y^2 = x^3 + x + 1$.

Question 5. *Consider the operation of addition that is defined on E in Figure 2.*

1. *Pick two distinct x-coordinates not shown in Figure 2. Find the corresponding points on E and add them together. Verify that your new point satisfies the equation defining E.*

2. *Verify that the coordinates are correct in Figure 2 for the case of doubling P. Instead of a secant line, use the tangent line at P.*

3. *Consider and point P on E. Verify geometrically that $P + \mathcal{O} = P$ and that $-P + P = \mathcal{O}$ where $-P$ is obtained by negating the y-coordinate of P.*

4. *A common additive property is that $(P + Q) + R = P + (Q + R)$. Provide a geometric example that shows this property fails if you replace $+$ with $*$. This is the reason that the reflection was necessary for defining the operation on E.*

1 Elliptic curves over finite fields

Figure 1 depicts $E : y^2 = x^3 + x + 1$ over $\mathbb{R}$. We denote the set of points in Figure 1 along with $\mathcal{O}$ as $E(\mathbb{R})$ and use it to gain intuition on how to add points. Using $E(\mathbb{R})$ for cryptography is possible but the calculations are slow due to unwieldy coordinates. To speed up calculations, we work over a finite field and modify our geometric interpretation accordingly.

Definition 5 (Finite field)**.** The most common examples of *finite fields* are sets of integers where the operation of addition and multiplication are calculated modulo a prime p. That is, numbers are divided by p and we look at the remainders. Table 3 represents the operation of addition and multiplication modulo the prime 5. The set of integers $\{0, 1, 2, 3, 4\}$ along with these operations are denoted as $\mathbb{F}_5$.

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	$\mathbf{2}$
4	4	0	1	2	3

$\cdot$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	$\mathbf{4}$	2
4	4	3	2	1

Table 3: Operations addition (+) and multiplication ($\cdot$) modulo 5

Question 6. *Modular arithmetic is often called* clock arithmetic *because of the "wrapping" effect of a 12-hour clock.*

1. *Suppose that it is 9:00*AM*. Working on a 12-hour clock, what time will it be in* 10 *hours? If you omit* AM *and* PM *from the question and answer, then your calculation is an example of doing addition modulo 12.*

2. *Verify the boldface entries in Table 3 using arithmetic modulo* 5*.*

3. *In order to divide by nonzero elements of a finite field, consider division as the operation that undoes the effect of multiplication. Dividing by* x *is equivalent to multiplying by the element* x^{-1} *where* $x \cdot x^{-1} = 1$*. Explain the following calculation* $3/2 \equiv 3 * 3 = 9 \equiv 4 \mod 5$*.*

4. *Complete Table 4.*

/	1	2	3	4
1	1	2	3	4
2		1	4	2
3			1	
4				1

Table 4: Division (/) modulo 5

Arithmetic modulo 5 allows us to consider solutions to equations representing elliptic curves over $\mathbb{F}_5$.

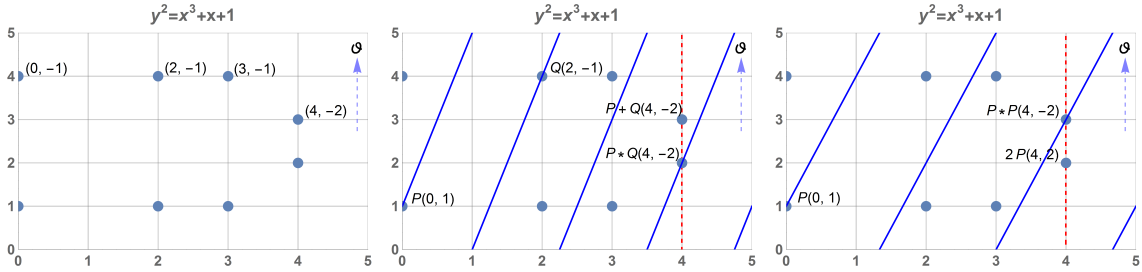

Figure 3: A geometric interpretation of addition on $E : y^2 = x^3 + x + 1$ over $\mathbb{F}_5$.

Question 7. *Consider $E : y^2 = x^3 + x + 1$ defined over $\mathbb{F}_5$ in Figure 3.*

1. *Verify that the eight finite points pictured in Figure 3 lie on $E(\mathbb{F}_5)$. Show there are no others.*

2. *Show that the slope of the line containing $P(0, 1)$ and $Q(2, -1)$ is 4 in $\mathbb{F}_5$. Verify the coordinate found for $P + Q$ in Figure 3.*

3. *There is no geometric interpretation of a tangent line in a discrete setting. Use the formal rules of ordinary calculus to define a tangent line at $P(0, 1)$. Show the tangent line has slope 3, and use it to verify the coordinates found for $2P$ in Figure 3.*

5 Security behind ECC

The security of ECC is based on the following one-way function. Let $k \in \mathbb{Z}$ be an integer and consider a prime $p > 3$ and an elliptic curve E containing the point P. Define $g(P) = kP = Q$ which is adding P to itself k-times on E. Consider Question 8 to familiarize ourselves with g.

P	$(0, 1)$	$(2, 1)$	$(3, 1)$	$(4, 2)$
$2P$	$(4, 2)$	$(2, -1)$	$(0, 1)$	$(3, -1)$
$3P$	$(2, 1)$	$\mathcal{O}$	$(2, -1)$	$(2, -1)$
$4P$	$(3, -1)$	$(2, 1)$	$(4, 2)$	$(0, -1)$
$5P$	$(3, 1)$	$(2, -1)$	$(4, -2)$	$(0, 1)$
$6P$	$(2, -1)$	$\mathcal{O}$	$(2, 1)$	$(2, 1)$
$7P$	$(4, -2)$	$(2, 1)$	$(0, -1)$	$(3, 1)$
$8P$	$(0, -1)$	$(2, -1)$	$(3, -1)$	$(4, -2)$
$9P$	$\mathcal{O}$	$\mathcal{O}$	$\mathcal{O}$	$\mathcal{O}$

Table 5: Scalar multiplication on $E : y^2 = x^3 + x + 1$ over $\mathbb{F}_5$

Question 8. *Consider $E : y^2 = x^3 + x + 1$ defined over $\mathbb{F}_5$.*

1. *Figure 3 demonstrates doubling $P(0, 1)$. Verify the remaining entries in row 2 of Table 5 when the initial point P varies.*
2. *Table 5 contains half of the finite points on E. Explain how the fact that $2(-P) = -2P$ can be used to determine the doubles of the remaining points on E.*
3. *Verify the entries in rows four and eight (i.e powers of two) using the information about doubling in Table 5.*
4. *Use the fact that $-P + P = \mathcal{O}$ to verify the entries in row 9, row 7, and row 5 of Table 5.*
5. *Verify the remaining rows using addition on the elliptic curve E.*
6. *Use Table 5 to solve the inverse of the one-way function. Given $P(0, 1)$ and $Q(2, -1)$ determine k where $kP = Q$.*

Question 8 is an example of how to crack the security behind ECC when working over $\mathbb{F}_5$. We were able to calculate the inverse of g given specific points P and Q on E. The ease of this can be analogously viewed with the security behind RSA. Revisit Question 3 while replacing $n = 323$, and see how easy the inverse of f is to determine. The difficulty of the inverse for both f and g is due to the size of primes you consider. In the case of ECC, we increase the size of the prime p corresponding to the base field $\mathbb{F}_p$. The advantage with ECC is that the size of p required is much smaller than the corresponding size of n in RSA with the same level of security. Attempting to answer Question 9 should clarify the idea behind g being a one-way function. Question 9 works over the finite field $\mathbb{F}_{1223}$. Arithmetic in $\mathbb{F}_{1223}$ is analogous to arithmetic in $\mathbb{F}_5$, except that the modulus is the prime number 1223.

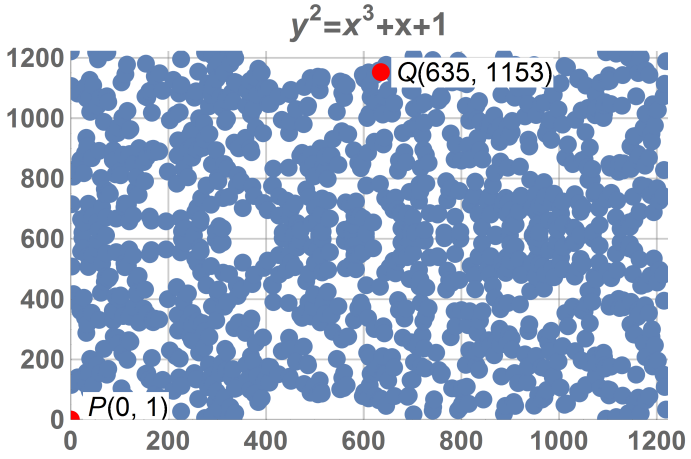

Figure 4: $E : y^2 = x^3 + x + 1$ over $\mathbb{F}_{1223}$ contains 1167 points

Question 9 ("Hard"). *Consider $E : y^2 = x^3 + x + 1$ over $\mathbb{F}_{1223}$ in Figure 4. Given $P(0, 1)$ and $Q(635, 1153)$ on E, determine k where $kP = Q$.*

The key size in Question 9 has bit length 11 corresponding to $p = 1223$. The recommended key size for ECC on the web has bit length 224 [3]. Figure 4 loses all definition with such a large prime so we do not use it for demonstrative purposes.

The answer to Question 9 is $k = 421$. To solve this by brute force, it would have taken 421 additions of P to stumble upon Q. It is difficult to cut down the required number of steps because multiples of P do not seemingly exhibit any pattern. Notice the lack of discernable pattern in the first twenty multiples of P.

$$\begin{aligned} nP = \{&(0,1),(306,1069),(72,611),(884,527),(720,943),(941,796),(315,848),\\ &(452,523),(903,953),(935,1099),(78,755),(559,1118),(579,647),(426,677),\\ &(708,977),(980,562),(900,582),(32,935),(1054,269),(1060,428)\ldots\}\,. \end{aligned}$$

The one-way function g should be easier to calculate than the 421 brute force steps required to answer Question 9. A nice approach considers $421P$ in a base 2 expansion:

$$421P = 256P + 128P + 32P + 4P + P\,.$$

Now $421P$ can be calculated by computing 8 doubles and 4 additions.

6 Elliptic curve Diffie-Hellman key exchange

ECDH key exchange can be used on a public channel to set up a private symmetric key which can be used to encrypt and decrypt transmissions during a communication session. Literature often represents this by way of a story: Alice and Bob set up a secure connection while an adversary Eve eavesdrops. In a modern context, a client connects to a server through a network. The client needs to protect sensitive information being submitted to the server. Also, the server guarantees that the information being sent to the client has not been tampered with.

The ECDH key exchange begins by Alice and Bob individually selecting an integer which each keep private. If Alice and Bob select 421 and 583 respectively, then they calculate the points $A(635, 1153) = 421P$ and $B = (14, 1082) = 583P$. One method to set up a symmetric key is for Alice and Bob to exchange their respective x-coordinates over the unsecured channel. Notice that if Eve intercepts this information, then she needs to solve Question 9 to recover the private keys. In order to set up a symmetric key, Alice and Bob both recover each other's y-coordinates (or its negative; see Question 12).

Question 10. *If Alice is provided $E : y^2 = x^3 + x + 1$ over $\mathbb{F}_{1223}$ and $x = 14$, then use a computer algebra system like* Mathematica *to determine the two points on E with $x = 14$.*

Question 10 provides two points to continue with, and Question 12 shows that it is irrelevant which one you pick for further calculations. After Alice recovers Bob's point $(14, 1082)$, she then uses her private key to calculate $421(14, 1082) = 421 \cdot 583P \equiv 843P \pmod{1223} = (274, 930)$.

Question 11. *If Bob is provided $E : y^2 = x^3 + x + 1$ over $\mathbb{F}_{1223}$ and recovers the point $A(635, 1153)$, then show that, when he calculates $583A$, he determines the same point as Alice.*

Alice and Bob now use the x-coordinate 274 as the symmetric key for all further communication. In practice, the x-coordinate will contain hundreds of digits, so the first 256 bits are used as the key, or, more generally, a so-called *hash function* is applied, which shortens the number of digits.

Question 12. *Use Table 5 to mimic the ECDH key exchange when $p = 5$. Show that the choice of a point from Question 10 is irrelevant by intentionally using the "wrong" point to finish the exchange and observing that the same x-coordinate is recovered.*

References

[1] Enrique de Argaez, *Internet World Stats*, http://www.internetworldstats.com/stats.htm, last retrieved on 27 April 2018.

[2] Joseph A. Gallian, *Contemporary Abstract Algebra*, Sixth edition, Houghton Mifflin, Boston, MA 2006.

[3] Damien Giry, *BlueKrypt, Cryptographic Key Length Recommendation*, https://www.keylength.com/en/4/, last retrieved on 27 April 2018.

[4] Darel W. Hardy, Fred Richman, and Carol L. Walker, *Applied Algebra, Codes, Ciphers, and Discrete Algorithms*, Second edition, CRC Press, Taylor & Francis Group, Boca Raton, FL, 2009.

[5] N. Koblitz, Elliptic curve cryptosystems, *Mathematics of Computation* **48** (1987), 203–209.

[6] H. W. Lenstra, Jr., Factoring integers with elliptic curves, *Annals of Mathematics* **126** (1987), 649–673.

[7] V. Miller, Uses of elliptic curves in cryptography, Advances in Cryptology - CRYPTO '85, *Lecture Notes in Computer Science* **218** (1986), 417–426.

[8] Emily Schechter, *Google Security Blog*, https://security.googleblog.com/2016/09/moving-towards-more-secure-web.html, last retrieved on 27 April 2018.

[9] Joseph H. Silverman and John Tate, *Rational Points on Elliptic Curves*, Springer, New York, NY, 1992.

[10] Lawrence C. Washington, *Elliptic Curves, Number Theory and Cryptography*, Second edition, CRC Press, Taylor & Francis Group, Boca Raton, FL, 2008.

[11] Wikipedia contributiors, RSA Factoring Challenge, *Wikipedia, The Free Encyclopedia*, https://en.wikipedia.org/wiki/RSA_Factoring_Challenge, last retrieved on 27 April 2018.