

Some Facets of Complexity Theory and Cryptography: A Five-Lectures Tutorial

Jörg Rothe*
Abteilung für Informatik
Heinrich-Heine-Universität Düsseldorf
40225 Düsseldorf, Germany
rothe@cs.uni-duesseldorf.de

November 21, 2001

Abstract

In this tutorial, selected topics of cryptology and of computational complexity theory are presented. We give a brief overview of the history and the foundations of classical cryptography, and then move on to modern public-key cryptography. Particular attention is paid to cryptographic protocols and the problem of constructing the key components of such protocols such as one-way functions. A function is one-way if it is easy to compute, but hard to invert. We discuss the notion of one-way functions both in a cryptographic and in a complexity-theoretic setting. We also consider interactive proof systems and present some interesting zero-knowledge protocols. In a zero-knowledge protocol one party can convince the other party of knowing some secret information without disclosing any bit of this information. Motivated by these protocols, we survey some complexity-theoretic results on interactive proof systems and related complexity classes.

*This work was supported in part by grant NSF-INT-9815095/DAAD-315-PPP-g`u-ab.

Contents

Outline of the Tutorial	3
1 Cryptosystems and Perfect Security	4
1.1 Classical Cryptosystems	4
1.2 Conditional Probability and the Theorem of Bayes	7
1.3 Perfect Security: Shannon's Theorem	8
1.4 Vernam's One-Time Pad	11
2 RSA Cryptosystem	11
2.1 Euler and Fermat's Theorems	11
2.2 RSA	12
2.3 Security of RSA and Possible Attacks on RSA	15
3 Protocols for Secret-Key Agreement, Public-Key Encryption, and Digital Signatures	18
3.1 Diffie and Hellman's Secret-Key Agreement Protocol	19
3.2 ElGamal's Public-Key Cryptosystem and Digital Signature Protocol	21
3.3 RSA Digital Signature Protocol	22
3.4 Shamir's No-Key Protocol	24
3.5 Rivest, Rabi, and Sherman's Secret-Key Agreement and Digital Signature Protocols	24
3.6 Discussion of Diffie-Hellman versus Rivest-Sherman	27
4 Strongly Noninvertible Associative One-Way Functions	28
4.1 Creating Strongly Noninvertible, Total, Commutative, Associative One-Way Functions from Any One-Way Function	29
4.2 If $P \neq NP$ then Some Strongly Noninvertible Functions are Invertible	32
5 Interactive Proof Systems and Zero-Knowledge Protocols	34
5.1 Interactive Proof Systems	35
5.2 Zero-Knowledge Protocols	38
5.3 Zero-Knowledge Protocol for the Graph Isomorphism Problem	39
5.4 Fiat and Shamir's Zero-Knowledge Protocol	41
References	44

Outline of the Tutorial

This tutorial consists of five lectures on cryptography, based on the lecture notes for a course on this subject given by the author in August, 2001, at the 11th Jyv^oaskyl^oa Summer School in Jyv^oaskyl^oa, Finland. As the title suggests, a particular focus of this tutorial is on emphasizing the close relationship between cryptography and complexity theory. The material presented here is not meant to be a comprehensive study or a complete survey of (the intersection of) these fields. Rather, five vivid topics from those fields are chosen for exposition, and from each topic chosen, some gems—some particularly important, central, beautiful results—are presented. Needless to say that the choice of topics and of results selected for exposition is based on the author’s personal tastes and biases.

The first lecture sketches the history and the classical foundations of cryptography, introduces a number of classical, symmetric cryptosystems, and briefly discusses by example the main objectives of the two opposing parts of cryptology: cryptography, which aims at designing secure ways of encryption, versus cryptanalysis, which aims at breaking existing cryptosystems. Then, we introduce the notion of perfect security for cryptosystems, which dates back to Claude Shannon’s pioneering work [Sha49] on coding and information theory.

The second lecture presents the historically first public-key cryptosystem, RSA, which was invented by Rivest, Shamir, and Adleman [RSA78]. To this end, some background from number theory is provided, as short as possible and as extensive as necessary to understand the underlying mathematics. In contrast to the information-theoretical approach to perfect security, the security of RSA is based on the assumption that certain problems from number theory are computationally intractable. Potential attacks on the RSA cryptosystem as well as appropriate countermeasures against them are discussed.

The third lecture introduces a number of cryptographic protocols, including the secret-key agreement protocols of Diffie and Hellman [DH76] and of Rivest and Sherman (see [RS93,RS97]), a hybrid version of ElGamal’s public-key cryptosystem [ElG85], Shamir’s no-key protocol, and the digital signature schemes of Rivest, Shamir, and Adleman [RSA78], ElGamal [ElG85], and Rabi and Sherman [RS93,RS97], respectively. Again, the underlying mathematics and, relatedly, security issues of these protocols are briefly discussed.

A remark is in order here. The protocols presented here are among the most central and important cryptographic protocols, with perhaps one exception: While the secret-key agreement protocol of Diffie and Hellman [DH76] is widely used in practice, that of Rivest and Sherman (see [RS93,RS97]) is not (yet) used in applications and, thus, might appear somewhat exotic at first glance. However, in our point of view, there is some hope that this fact, though currently true, might change in the near future. In Section 3.6, we will discuss the state of the art on these two protocols. In particular, we give there—and in more detail in Section 4—an overview on the progress regarding the Rivest-Rabi-Sherman protocols that was recently obtained by Hemaspaandra, Pasanen, and this author [HR99,HPR01]. We also argue how this progress, combined with the progress on the complexity of the shortest lattice vector problem (SVP, for short) obtained by Ajtai [Ajt96], may lead to a significant increase in the cryptographic security and applicability of the Rivest-Rabi-Sherman protocols. Thus, our intention in including these protocols in this tutorial is to provide the students not only with well-known, standard material contained in the textbooks, but to also guide them towards an active field of current research.

The fourth lecture presents in detail the above-mentioned result [HR99] on how to construct, from the assumption that $P \neq NP$, the key building block of the Rivest-Rabi-Sherman protocols: strongly noninvertible, total, commutative, associative one-way functions. In addition, some more recent results [HPR01] on strong noninvertibility are surveyed, including the perhaps somewhat surprising result that if $P \neq NP$ then there exist strongly noninvertible functions that in fact are invertible. These results are obtained only in the *worst-case* complexity model, which is relevant and interesting in a complexity-theoretic setting, but useless in applied cryptography. For cryptographic applications, one would need

to construct such functions based on the *average-case* complexity model, under plausible assumptions. Hence, the most challenging open research question hinted at in the preceding paragraph, is to find some evidence that strongly noninvertible, associative one-way functions exist in the average-case model. As noted there, our hope to obtain such a result is based on the recent progress on the shortest lattice vector problem obtained by Ajtai [Ajt96], who proved that this problem, roughly speaking, is equally hard in the worst-case and in the average-case model. Ajtai's result has already lead to a public-key cryptosystem designed by him and Dwork [AD97]. The reason why Ajtai's breakthrough results and techniques and their cryptographic applications are not presented in this tutorial is that there already exist nice surveys by Cai [Cai99] and, even more recently, by Kumar and Sivakumar [KS01] on the complexity of SVP.

The fifth lecture introduces interactive proof systems and zero-knowledge protocols. This area has rapidly developed and flourished in complexity theory and has yielded a number of powerful results with regard to probabilistically checkable proofs and the nonapproximability of hard optimization problems (see the survey [Gol97]), surprising complexity class characterizations such as Shamir's famous "IP = PSPACE" result [Sha92], and has direct applications in cryptography. In particular, zero-knowledge protocols enable one party to convince the other party of knowing some secret information without conveying any bit of this information. Thus, they are ideal technical tools for authentication purposes. We present two of the classic zero-knowledge protocols: the Goldreich-Micali-Wigderson protocol for graph isomorphism [GMW86,GMW91] and the Fiat-Shamir protocol [FS86] that is based on a number-theoretical problem. For an in-depth treatment of zero-knowledge protocols and many more technical details, the reader is referred to Chapter 4 of Goldreich's book [Gol01b].

The tutorial is suitable for graduate students with some background in computer science and mathematics and may also be accessible to interested undergraduate students. Since it is organized in five pairwise essentially independent, self-contained lectures, it is also possible to present only a proper subset of these lectures. The only dependencies occurring between lectures are that some of the number-theoretical background given Section 2 is also used in Section 3, and that the Rivest-Rabi-Sherman protocols presented in Section 3 motivate the investigations in Section 4.

There are a number of textbooks and monographies on cryptography that cover various parts of the field in varying depth, such as the books by Goldreich [Gol99,Gol01b], Salomaa [Sal96], Stinson [Sti95], and Welsh [Wel98]. Schneier's book [Sch96] provides a very comprehensive collection of literally all notions and concepts known in cryptography, which naturally means that the single notions and concepts cannot be treated in every mathematical detail there, but the interested reader is referred to an extraordinarily large bibliography for such an in-depth treatment. Singh [Sin99] wrote a very charming, easy-to-read, interesting book about the history of cryptography from its ancient roots to its modern and even futuristic branches such as quantum cryptography. We conclude this list, without claiming it to be complete, with some examples of books on this subject that are written in German, such as the books by Bauer [Bau00] and Beutelspacher [Beu01]; note also that Buchmann's book [Buc01] is the author's translation from its German original version.

1 Cryptosystems and Perfect Security

1.1 Classical Cryptosystems

The notion of a cryptosystem is formally defined as follows.

Definition 1.1 (Cryptosystem)

- A cryptosystem is a quintuple $(\mathcal{M}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ such that:

1. \mathcal{M}, \mathcal{C} , and \mathcal{K} are sets, where

\mathcal{M} is the message space (or “plain text space” or “clear text space”);
 \mathcal{C} is the cipher text space;
 \mathcal{K} is the key space.

2. $\mathcal{E} = \{E_k \mid k \in \mathcal{K}\}$ is a family of functions $E_k : \mathcal{M} \rightarrow \mathcal{C}$ that are used for encryption, and $\mathcal{D} = \{D_k \mid k \in \mathcal{K}\}$ is a family of functions $D_k : \mathcal{C} \rightarrow \mathcal{M}$ that are used for decryption.
3. For each key $e \in \mathcal{K}$, there exists a key $d \in \mathcal{K}$ such that for each message $m \in \mathcal{M}$:

$$(1.1) \quad D_d(E_e(m)) = m$$

- A cryptosystem is called symmetric (or “private-key”) if $d = e$, or if d can at least be “easily” computed from e .
- A cryptosystem is called asymmetric (or “public-key”) if $d \neq e$, and it is “practically infeasible” to compute d from e . Here, d is the private key, and e is the public key.

At times, different key spaces are used for encryption and for decryption, which results in a slight modification of the above definition. We now present and discuss some examples of classical cryptosystems.

Example 1.2 (Caesar chiffré, a monoalphabetic symmetric cryptosystem)

Consider the English alphabet $\Sigma = \{A, B, \dots, Z\}$. To carry out the arithmetic modulo 26 with letters as if they were numbers, we identify Σ with $\mathbb{Z}_{26} = \{0, 1, \dots, 25\}$; thus, 0 represents A and 1 represents B, and so on. This encoding of the plain text alphabet by integers and the decoding of \mathbb{Z}_{26} back to Σ is not part of the actual encryption and decryption, respectively.

Let $\mathcal{K} = \mathbb{Z}_{26}$, and let $\mathcal{M} = \mathcal{C} = \Sigma^*$, where Σ^* denotes the set of strings over Σ . The Caesar chiffré encrypts messages by shifting (modulo 26) each bit of the plain text by the same number k of letters in the alphabet, where k is the key. Shifting each bit of the chiffré back using the same key k reveals the original message:

- For each $e \in \mathbb{Z}_{26}$, define the encryption function $E_e : \Sigma^* \rightarrow \Sigma^*$ by

$$E_e(m) = (m + e) \pmod{26},$$

where addition with e modulo 26 is carried out bit-wise, i.e., each bit $m_i \in \Sigma$ of $m \in \Sigma^*$ is shifted by e positions to $m_i + e \pmod{26}$. For example, using the key $e = 11 = L$, the message “SUMMER” will be encrypted as “DFXXPC.”

- For each $d \in \mathbb{Z}_{26}$, define the decryption function $D_d : \Sigma^* \rightarrow \Sigma^*$ by

$$D_d(c) = (c - d) \pmod{26},$$

where subtraction by e modulo 26 again is carried out bit-wise. Hence, $d = e$. For example, decrypting the chiffré “DNSZZW” with the key $d = 11$ reveals the plain text “SCHOOL.”

Since the key space is very small, breaking the Caesar chiffré is very easy. It is vulnerable even to “cipher-text-only attacks,” i.e., an attacker knowing a sample chiffré c can easily check the 26 possible keys to see which one yields a meaningful plain text.

The Caesar chiffré is a monoalphabetic cryptosystem, for it replaces each given plain text letter, wherever in the message it occurs, by the same letter of the chiffré alphabet. In contrast, the French cryptographer and diplomat Blaise de Vigenère (1523–1596) proposed a polyalphabetic cryptosystem, which is much harder to break. Vigenère’s system builds on earlier work by the Italian mathematician

Leon Battista Alberti (born in 1404), the German abbot Johannes Trithemius (born in 1492), and the Italian scientist Giovanni Porta (born in 1535), see [Sin99]. It works like the Caesar chiffré, except that the chiffré letter encrypting any given plain text letter X varies with the position of X in the plain text.

More precisely, one uses for encryption and decryption a *Vigenère square*, which consists of 26 rows with 26 columns each. Every row contains the 26 letters of the alphabet, shifted by one from row to row, i.e., the rows and columns may be viewed as a Caesar chiffré with keys $0, 1, \dots, 25$. Given a message $m \in \Sigma^*$, one first chooses a key $k \in \Sigma^*$, which is written above the message m , symbol by symbol, possibly repeating k if k is shorter than m until the key word has the same length as m . Denoting the i th letter of any string w by w_i , each letter m_i of m is then encrypted as in the Caesar chiffré, using the row of the Vigenère square that starts with k_i , where k_i is the key letter right above m_i . Below, we describe the Vigenère system formally and give an example of a concrete encryption.

Example 1.3 (Vigenère chiffré, a polyalphabetic symmetric cryptosystem)

For fixed $n \in \mathbb{N}$, let $\mathcal{K} = \mathcal{M} = \mathcal{C} = \mathbb{Z}_{26}^n$. Messages $m \in \Sigma^*$, where Σ again is the English alphabet, are split into blocks of length n and are encrypted block-wise. The Vigenère chiffré is defined as follows.

- For each $e \in \mathbb{Z}_{26}^n$, define the encryption function $E_e : \mathbb{Z}_{26}^n \rightarrow \mathbb{Z}_{26}^n$ by

$$E_e(m) = (m \oplus e) \pmod{26},$$

where \oplus denotes bit-wise addition modulo 26.

- For each $d \in \mathbb{Z}_{26}^n$, define the decryption function $D_d : \mathbb{Z}_{26}^n \rightarrow \mathbb{Z}_{26}^n$ by

$$D_d(c) = (c \ominus d) \pmod{26},$$

where \ominus denotes bit-wise subtraction modulo 26. Again, $d = e$.

For example, choose the word $k = \text{ENGLISH}$ to be the key. Suppose we want to encrypt the message $m = \text{FINNISHISALLGREEKTOGERMANS}$,¹ omitting the spaces between words. Table 1 shows how each plain text letter is encrypted, yielding the cipher text c . For instance, the first letter of the message, “F,” corresponds to the first letter of the key, “E.” Hence, the intersection of the “F”-column with the “E”-row of the Vigenère square gives the first letter, “J,” of the cipher text.

k	E	N	G	L	I	S	H	E	N	G	L	I	S	H	E	N	G	L	I	S	H	E	N	G	L	I
m	F	I	N	N	I	S	H	I	S	A	L	L	G	R	E	E	K	T	O	G	E	R	M	A	N	S
c	J	V	T	Y	Q	K	O	M	F	G	W	T	Y	Y	I	R	Q	E	W	Y	L	V	Z	G	Y	A

Table 1: An example of encryption by the Vigenère chiffré.

Our last example of a classical, historically important cryptosystem is the Hill chiffré, which was invented by Lester Hill in 1929. It is based on linear algebra and, like the Vigenère chiffré, is an affine linear block chiffré.

Example 1.4 (Hill chiffré, a symmetric cryptosystem and a linear block chiffré)

For fixed $n \in \mathbb{N}$, the key space \mathcal{K} is the set of all invertible $n \times n$ matrices in $\mathbb{Z}_{26}^{(n \times n)}$ whose determinant is coprime with 26. Again, $\mathcal{M} = \mathcal{C} = \mathbb{Z}_{26}^n$ and messages $m \in \Sigma^*$ are split into blocks of length n and are encrypted block-wise. All arithmetic operations are carried out modulo 26.

The Hill chiffré is defined as follows.

¹From this example we not only learn how the Vigenère chiffré works, but also that using a language such as Finnish, which is not widely used, often makes illegal decryption harder, and thus results in a higher level of security. This is not a purely theoretical observation. During World War II, the US Navy transmitted important messages using the language of the Navajos, a Native American tribe. The “Navajo Code” was never broken by the Japanese code-breakers, see [Sin99].

- For each $K \in \mathcal{K}$, define the encryption function $E_K : \mathbb{Z}_{26}^n \rightarrow \mathbb{Z}_{26}^n$ by

$$E_K(m) = K \cdot m \pmod{26},$$

where \cdot denotes matrix multiplication modulo 26.

- Letting K^{-1} denote the inverse matrix of K , the decryption function $D_{K^{-1}} : \mathbb{Z}_{26}^n \rightarrow \mathbb{Z}_{26}^n$ is defined by

$$D_{K^{-1}}(c) = K^{-1} \cdot c \pmod{26}.$$

Since K^{-1} can easily be computed from K , the Hill chiffrage is a symmetric cryptosystem. It is also the most general linear block chiffrage.

Concrete examples of messages encrypted by the Hill chiffrage can be found in, e.g., [Sal96].

Affine linear block ciphers are easy to break by “known-plain-text attacks,” i.e., for an attacker who knows some sample plain texts with the corresponding encryptions, it is not too hard to find the key used to encrypt these plain texts. They are even more vulnerable to “chosen-plain-text attacks,” where the attacker him- or herself can choose some pairs of corresponding plain texts and encryptions, which may be useful if he or she has certain conjectures about the key used. In 1863, the German cryptanalyst Friedrich Wilhelm Kasiski found a method to break the Vigenère chiffrage.² The books by Salomaa [Sal96] and Singh [Sin99] describe Kasiski’s method. It marks a breakthrough in the history of cryptanalysis, because previously the Vigenère chiffrage was considered unbreakable. In particular, like similar periodic cryptosystems with an unknown period, the Vigenère chiffrage appeared to resist the cryptanalysis by counting and analysing the frequency of letters in the cipher text. Kasiski showed how to determine the period from repetitions of the same substring in the cipher text.

The method of frequency counts is often useful for decrypting messages; it exploits the redundancy of the natural language used for encryption. For example, in many languages the letter “E” occurs, statistically significant, most frequently, with a percentage of 12.31% in English, of 15.87% in French, and even of 18.46% in German, see [Sal96]. Some languages have other letters that occur with the highest frequency; for example, “A” is the most frequent letter in average Finnish texts, with a percentage of 12.06% [Sal96].

In the light of Kasiski’s and Babbage’s achievement, it is natural to ask whether there exist any cryptosystems that are *perfectly secure*. We turn to this question in the next section that describes some of the pioneering work of Claude Shannon [Sha49], who laid the foundations of modern coding and information theory.

1.2 Conditional Probability and the Theorem of Bayes

To discuss perfect security of cryptosystems in mathematical terms, we first need some preliminaries from elementary probability theory.

Definition 1.5 Let A and B be events with $\Pr(B) > 0$.

- The probability that A occurs under the condition that B occurs is defined by

$$\Pr(A | B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

- A and B are independent if $\Pr(A \cap B) = \Pr(A) \Pr(B)$ (equivalently, if $\Pr(A | B) = \Pr(A)$).

²Singh [Sin99] attributes this achievement also to an unpublished work, done probably around 1854, by the British genius and eccentric Charles Babbage.

Lemma 1.6 (Theorem of Bayes) *Let A and B be events with $\Pr(A) > 0$ and $\Pr(B) > 0$. Then,*

$$\Pr(B) \Pr(A | B) = \Pr(A) \Pr(B | A).$$

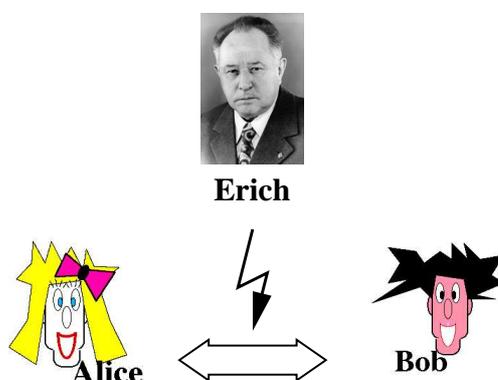
Proof. By definition,

$$\Pr(B) \Pr(A | B) = \Pr(A \cap B) = \Pr(B \cap A) = \Pr(A) \Pr(B | A).$$

■

1.3 Perfect Security: Shannon's Theorem

Consider the following scenario:



- Alice and Bob are communicating over an insecure channel, in the presence of eavesdropper Erich.
- Alice and Bob use a cryptosystem $(\mathcal{M}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$, where \mathcal{M} , \mathcal{C} , and \mathcal{K} are finite sets.
- Erich reads a cipher text c and tries to get some information about the corresponding message m .
- The messages are distributed on \mathcal{M} according to a probability distribution $\Pr_{\mathcal{M}}$ that may depend on the language used.
- For each new message, Alice chooses a new key from \mathcal{K} that is independent of the message to be encrypted. The keys are distributed according to a probability distribution $\Pr_{\mathcal{K}}$ on \mathcal{K} .
- $\Pr_{\mathcal{M}}$ and $\Pr_{\mathcal{K}}$ induce a probability distribution $\Pr = \Pr_{\mathcal{M} \times \mathcal{K}}$ on $\mathcal{M} \times \mathcal{K}$. Thus, for each message m and each key k ,

$$\Pr(m, k) = \Pr_{\mathcal{M}}(m) \Pr_{\mathcal{K}}(k)$$

is the probability that the message m is encrypted with the key k , where m and k are independent.

- For $m \in \mathcal{M}$, let m denote the event $\{(m, k) \mid k \in \mathcal{K}\}$. Then, $\Pr(m) = \Pr_{\mathcal{M}}(m)$ is the probability that the message m will be encrypted.
- For $k \in \mathcal{K}$, let k denote the event $\{(m, k) \mid m \in \mathcal{M}\}$. Then, $\Pr(k) = \Pr_{\mathcal{K}}(k)$ is the probability that the key k will be used.
- For $c \in \mathcal{C}$, let c denote the event $\{(m, k) \mid E_k(m) = c\}$. Then, $\Pr(m | c)$ is the probability that m is encrypted under the condition that c is received.

- Erich knows the cipher text c , and he knows the probability distribution $\Pr_{\mathcal{M}}$, since he knows the language used by Alice and Bob.

Definition 1.7 A cryptosystem $(\mathcal{M}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ is perfectly secure if and only if

$$(\forall m \in \mathcal{M}) (\forall c \in \mathcal{C}) [\Pr(m | c) = \Pr(m)].$$

That is, a cryptosystem achieves perfect security if the event that some message m is encrypted and the event that some cipher text c is received are independent: Erich learns nothing about m from knowing c . The following example of a cryptosystem that is not perfectly secure is due to Buchmann [Buc01].

Example 1.8 (Perfect security) Let \mathcal{M} , \mathcal{C} , and \mathcal{K} be given such that:

- $\mathcal{M} = \{0, 1\}$, where $\Pr(0) = \frac{1}{4}$ and $\Pr(1) = \frac{3}{4}$;
- $\mathcal{K} = \{A, B\}$, where $\Pr(A) = \frac{1}{4}$ and $\Pr(B) = \frac{3}{4}$;
- $\mathcal{C} = \{a, b\}$.

It follows that, for example, the probability that a “1” occurs and is encrypted with the key B is:

$$\Pr(1, B) = \Pr(1) \cdot \Pr(B) = \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}.$$

Let the encryption functions be given by:

$$E_A(0) = a; \quad E_A(1) = b; \quad E_B(0) = b; \quad E_B(1) = a.$$

Hence, the probability that the cipher a occurs is:

$$\Pr(a) = \Pr(0, A) + \Pr(1, B) = \frac{1}{16} + \frac{9}{16} = \frac{5}{8}.$$

Similarly, the probability that the cipher b occurs is:

$$\Pr(b) = \Pr(1, A) + \Pr(0, B) = \frac{3}{16} + \frac{3}{16} = \frac{3}{8}.$$

Then, for each pair $(m, c) \in \mathcal{M} \times \mathcal{C}$, the conditional probability $\Pr(m | c)$ is:

$$\begin{aligned} \Pr(0 | a) &= \frac{\Pr(0, A)}{\Pr(a)} = \frac{\frac{1}{16}}{\frac{5}{8}} = \frac{1}{10}; & \Pr(0 | b) &= \frac{\Pr(0, B)}{\Pr(b)} = \frac{\frac{3}{16}}{\frac{3}{8}} = \frac{1}{2}; \\ \Pr(1 | a) &= \frac{\Pr(1, B)}{\Pr(a)} = \frac{\frac{9}{16}}{\frac{5}{8}} = \frac{9}{10}; & \Pr(1 | b) &= \frac{\Pr(1, A)}{\Pr(b)} = \frac{\frac{3}{16}}{\frac{3}{8}} = \frac{1}{2}. \end{aligned}$$

In particular, it follows that

$$\Pr(0) = \frac{1}{4} \neq \frac{1}{10} = \Pr(0 | a),$$

and thus the given cryptosystem is not perfectly secure: If Erich sees the cipher a , he can be pretty sure that the encrypted message was a “1.”

Theorem 1.9 (Shannon [Sha49]) Let $S = (\mathcal{M}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ be a cryptosystem with $|\mathcal{C}| = |\mathcal{K}|$ and $\Pr(m) > 0$ for each $m \in \mathcal{M}$. Then, S is perfectly secure if and only if

- (1) $\Pr_{\mathcal{K}}$ is the uniform distribution, and

(2) for each $m \in \mathcal{M}$ and for each $c \in \mathcal{C}$, there exists a unique key $k \in \mathcal{K}$ with $E_k(m) = c$.

Proof. Assume that S is perfectly secure. We show that the conditions (1) and (2) hold.

Condition (2): Fix a message $m \in \mathcal{M}$. Suppose that there is a cipher text $c \in \mathcal{C}$ such that for all $k \in \mathcal{K}$, it holds that $E_k(m) \neq c$. Thus,

$$\Pr(m) \neq 0 = \Pr(m | c),$$

which implies that S is not perfectly secure, a contradiction. Hence,

$$(\forall c \in \mathcal{C}) (\exists k \in \mathcal{K}) [E_k(m) = c].$$

Now, $|\mathcal{C}| = |\mathcal{K}|$ implies that each cipher text $c \in \mathcal{C}$ has a unique key k with $E_k(m) = c$.

Condition (1): Fix a cipher text $c \in \mathcal{C}$. For $m \in \mathcal{M}$, let $k(m)$ be the unique key k with $E_k(m) = c$. By the Theorem of Bayes, for each $m \in \mathcal{M}$, we have:

$$(1.2) \quad \Pr(m | c) = \frac{\Pr(c | m) \Pr(m)}{\Pr(c)} = \frac{\Pr(k(m)) \Pr(m)}{\Pr(c)}.$$

Since S is perfectly secure, we have $\Pr(m | c) = \Pr(m)$. By Equation (1.2), this implies $\Pr(k(m)) = \Pr(c)$, and this equality holds independently of m .

Hence, the probabilities $\Pr(k)$ are equal for all $k \in \mathcal{K}$, which implies $\Pr(k) = \frac{1}{|\mathcal{K}|}$. Thus, $\Pr_{\mathcal{K}}$ is the uniform distribution.

Conversely, suppose that the conditions (1) and (2) hold. We show that S is perfectly secure. Let $k = k(m, c)$ be the unique key k with $E_k(m) = c$. By the Theorem of Bayes, it follows that

$$(1.3) \quad \begin{aligned} \Pr(m | c) &= \frac{\Pr(m) \Pr(c | m)}{\Pr(c)} \\ &= \frac{\Pr(m) \Pr(k(m, c))}{\sum_{q \in \mathcal{M}} \Pr(q) \Pr(k(q, c))}. \end{aligned}$$

Since all keys are uniformly distributed, it follows that

$$\Pr(k(m, c)) = \frac{1}{|\mathcal{K}|}.$$

Moreover, we have that

$$\sum_{q \in \mathcal{M}} \Pr(q) \Pr(k(q, c)) = \frac{\sum_{q \in \mathcal{M}} \Pr(q)}{|\mathcal{K}|} = \frac{1}{|\mathcal{K}|}.$$

Substituting this equality in Equation (1.3) gives:

$$\Pr(m | c) = \Pr(m).$$

Hence, S is perfectly secure. ■

1.4 Vernam's One-Time Pad

The Vernam one-time pad is a symmetric cryptosystem that is perfectly secure. It was invented by Gilbert Vernam in 1917, and is defined as follows. Let $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0, 1\}^n$ for some $n \in \mathbb{N}$. For $k \in \{0, 1\}^n$, define

- the encryption function $E_k : \{0, 1\}^n \rightarrow \{0, 1\}^n$ by

$$E_k(m) = m \oplus k \pmod{2}, \text{ and}$$

- the decryption function $D_k : \{0, 1\}^n \rightarrow \{0, 1\}^n$ by

$$d_k(c) = c \oplus k \pmod{2},$$

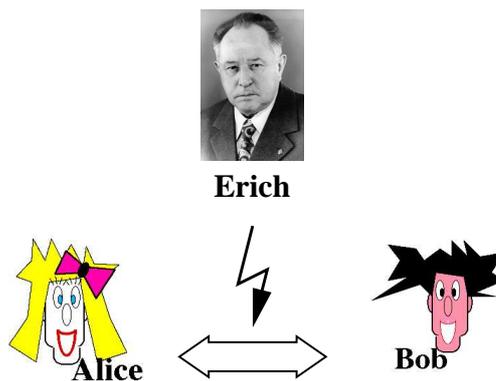
where \oplus denotes bit-wise addition modulo 2. The keys are uniformly distributed on $\{0, 1\}^n$.

By Shannon's Theorem, the one-time pad is perfectly secure, since for each message $m \in \mathcal{M}$ and for each cipher text $c \in \mathcal{C}$, there exists a unique key $k \in \mathcal{K}$ with $c = m \oplus k$, namely the string $k = c \oplus m$.

However, the one-time pad has major disadvantages that make it impractical to use in most concrete scenarios: To obtain perfect security, every key can be used only once, and it must be at least as long as the message to be transmitted. Surely, since for every communication a new secret key at least as long as the message must be transmitted, this results in a vicious circle. Despite these drawbacks, for the perfect security it provides, the one-time pad has been used in real-world applications such as, allegedly, the hotline between Moscow and Washington, see [Sim79, p. 316].

2 RSA Cryptosystem

The RSA cryptosystem, named after its inventors Ron Rivest, Adi Shamir, and Leonard Adleman, is the first public-key cryptosystem [RSA78]. It is still widely used in cryptographic applications today. Again, the scenario is that Alice and Bob want to exchange messages over an insecure channel, which is eavesdropped by Erich:



In order to describe how the RSA cryptosystem works, we first need some preliminaries from elementary number theory.

2.1 Euler and Fermat's Theorems

The *greatest common divisor* of two integers a and b is denoted by $\gcd(a, b)$. For $n \in \mathbb{N}$, define the set

$$\mathbb{Z}_n^* = \{i \mid 1 \leq i \leq n - 1 \text{ and } \gcd(i, n) = 1\}.$$

The *Euler function* Φ is defined by $\Phi(n) = |\mathbb{Z}_n^*|$. Note that \mathbb{Z}_n^* is a group (with respect to multiplication) of order $\Phi(n)$. It follows immediately from the definition that

- $\Phi(m \cdot n) = \Phi(m) \cdot \Phi(n)$ for all $m, n \in \mathbb{N}$ with $\gcd(m, n) = 1$, and
- $\Phi(p) = p - 1$ for all primes p .

Euler's Theorem below is a special case (for the group \mathbb{Z}_n^*) of Lagrange's Theorem, which states that for each element g of a finite multiplicative group G having order $|G|$ and the neutral element 1, it holds that $g^{|G|} = 1$.

Theorem 2.1 (Euler) For each $a \in \mathbb{Z}_n^*$, $a^{\Phi(n)} \equiv 1 \pmod{n}$.

The special case of Euler's Theorem with n being a prime not dividing a is known as Fermat's Little Theorem.

Theorem 2.2 (Fermat's Little Theorem) If p is a prime and $a \in \mathbb{Z}_p^*$, then $a^{p-1} \equiv 1 \pmod{p}$.

2.2 RSA

(1) Key generation:

1. Bob chooses randomly two large primes p and q with $p \neq q$, and computes their product $n = pq$.
2. Bob chooses a number $e \in \mathbb{N}$ with

$$(2.4) \quad 1 < e < \Phi(n) = (p-1)(q-1) \quad \text{and} \quad \gcd(e, \Phi(n)) = 1.$$

3. Bob computes the unique number d satisfying

$$(2.5) \quad 1 < d < \Phi(n) \quad \text{and} \quad e \cdot d \equiv 1 \pmod{\Phi(n)}.$$

That is, d is the inverse of e modulo $\Phi(n)$.

4. The pair (n, e) is Bob's *public key*, and d is Bob's *private key*.

In order to generate two large primes (e.g., primes with 80 digits each) efficiently, one picks large numbers at random and tests them for primality. Since by the Prime Number Theorem, the number of primes not exceeding N is approximately $\frac{N}{\ln N}$, the odds are good to hit a prime after a reasonably small number of trials. To verify the number picked, one usually makes use of randomized polynomial-time primality tests such as the Monte Carlo³ algorithm of Rabin [Rab80] that is related to a deterministic algorithm due to Miller [Mil76]; their primality test is known as the Miller-Rabin test. An alternative, though less popular Monte Carlo algorithm was proposed by Solovay and Strassen [SS77]. These two primality tests, along with a careful complexity analysis and the required number-theoretical background, can be found in, e.g., the books by Stinson [Sti95] and Salomaa [Sal96]. Additional primality test are contained in [Gol01b, Buc01].

We now argue that the keys can be computed efficiently. In particular, the inverse d of e modulo $\Phi(n)$ can be computed efficiently via the extended algorithm of Euclid; see Figure 1.

³A Monte Carlo algorithm is a randomized algorithm whose "yes" answers are reliable, while its "no" answers may be erroneous with a certain error probability, or vice versa. The corresponding complexity classes are called R and coR, respectively, see [Gil77]. In contrast, a Las Vegas algorithm may for certain sequences of coin flips halt without giving an answer at all, but whenever it gives an answer, this answer is correct. The corresponding class, $ZPP = R \cap \text{coR}$, was also defined by Gill [Gil77].

EUCLID'S ALGORITHM (extended)

Input: $b_0 := \Phi(n); b_1 := e.$
begin $x_0 := 1; y_0 := 0; x_1 := 0; y_1 := 1; i := 1;$
 while b_i does not divide b_{i-1} **do**
 begin
 $q_i := \lfloor \frac{b_{i-1}}{b_i} \rfloor;$
 $b_{i+1} := b_{i-1} - q_i \cdot b_i;$
 $x_{i+1} := x_{i-1} - q_i \cdot x_i;$
 $y_{i+1} := y_{i-1} - q_i \cdot y_i;$
 $i := i + 1$
 end
 begin output
 $b := b_i; \quad (* b = \gcd(e, \Phi(n)) = 1 *)$
 $x := x_i;$
 $y := y_i \quad (* y = d \text{ is Bob's private key} *)$
 end output
end

Figure 1: The extended algorithm of Euclid.

Lemma 2.3 *The extended algorithm of Euclid computes, on input e and $\Phi(n)$, in polynomial time integers x and y such that*

$$x \cdot \Phi(n) + y \cdot e \equiv 1 \pmod{\Phi(n)}.$$

Thus, y is the inverse of e modulo $\Phi(n)$, and Bob chooses $d \equiv y \pmod{\Phi(n)}$ as his private key.

Example 2.4 *Bob chooses the primes $p = 11$ and $q = 23$, and computes their product $n = 253$ and $\Phi(253) = 10 \cdot 22 = 220$. The smallest possible e satisfying Equation (2.4) is $e = 3$. The extended algorithm of Euclid yields the following sequence of b_i , x_i , and y_i :*

i	b_i	x_i	y_i	q_i
0	220	1	0	–
1	3	0	1	73
2	1	1	–73	–

Since $1 \cdot 220 + (-73) \cdot 3 = 220 - 219 \equiv 1 \pmod{220}$, the unique value $d = -73 + 220 = 147$ computed by Bob satisfies Equation (2.5) and is the inverse of $e = 3$ modulo 220.

(2) Encryption: We assume that messages over some alphabet Σ are block-wise encoded as positive integers with a fixed block length. Suppose that m , with $1 < m < n$, is the message Alice wants to send to Bob. Alice knows Bob's public key (n, e) and computes the encryption $c = E_{(n,e)}(m)$ of m , where the encryption function is defined by

$$E_{(n,e)}(m) = m^e \pmod{n}.$$

Depending on the choice of e , this computation may require a large number of multiplications. To ensure efficient encryption, we will employ a “fast exponentiation” algorithm called “square-and-multiply,” see Method 2.5 below.

Let the binary expansion of the exponent e be given by

$$e = \sum_{i=0}^k e_i 2^i, \quad \text{where } e_i \in \{0, 1\}.$$

Then, in the arithmetic modulo n , Alice computes:

$$(2.6) \quad m^e = m^{\sum_{i=0}^k e_i 2^i} = \prod_{i=0}^k \left(m^{2^i}\right)^{e_i} = \prod_{\substack{i=0 \\ e_i=1}}^k m^{2^i}.$$

Method 2.5 (Square-and-multiply)

1. Successively compute m^{2^i} , where $0 \leq i \leq k$, using the equality $m^{2^{i+1}} = \left(m^{2^i}\right)^2$.
2. According to Equation (2.6), compute $m^e = \prod_{\substack{i=0 \\ e_i=1}}^k m^{2^i}$.

That is, instead of e multiplications, Alice need compute no more than $2 \log e$ multiplications. Thus, the square-and-multiply method speeds up the encryption exponentially.

Example 2.6 Suppose Alice wants to compute $c = 6^{17} \pmod{100}$. The binary expansion of the exponent is $17 = 1 + 16 = 2^0 + 2^4$.

1. Alice successively computes:

$$\begin{aligned} 6^{2^0} &= 6^1 &&= 6; \\ 6^{2^1} &= 6^2 &&= 36; \\ 6^{2^2} &= 36^2 &&\equiv -4 \pmod{100}; \\ 6^{2^3} &\equiv (-4)^2 \pmod{100} &&\equiv 16 \pmod{100}; \\ 6^{2^4} &\equiv 16^2 \pmod{100} &&\equiv 56 \pmod{100}. \end{aligned}$$

2. Alice computes her cipher text

$$\begin{aligned} c &= 6^{17} \pmod{100} \equiv 6 \cdot 6^{2^4} \pmod{100} \\ &\equiv 6 \cdot 56 \pmod{100} \\ &\equiv 36 \pmod{100}. \end{aligned}$$

Note that only four squarings and one multiplication are needed for her to compute the cipher text.

(3) Decryption: Let c , $0 \leq c < n$, be the cipher text sent to Bob; c is subject to eavesdropping by Erich. Bob decrypts c using his private key d and the following decryption function:

$$D_d(c) = c^d \pmod{n}.$$

Again, the fast exponentiation algorithm described in Method 2.5 ensures that the legal recipient Bob can decrypt the cipher text efficiently. Thus, the RSA protocol is feasible. To prove that it is correct, we show that Equation (1.1) is satisfied.

Theorem 2.7 Let (n, e) and d be Bob's public and private key in the RSA protocol. Then, for each message m with $0 \leq m < n$,

$$(m^e)^d \pmod{n} = m.$$

That is, RSA is a public-key cryptosystem.

Proof. Since $e \cdot d \equiv 1 \pmod{\Phi(n)}$ by Equation (2.5), there exists an integer t such that

$$e \cdot d = 1 + t(p-1)(q-1),$$

where $n = pq$. It follows that

$$\begin{aligned} (m^e)^d &= m^{e \cdot d} = m^{1+t(p-1)(q-1)} \\ &= m \left(m^{t(p-1)(q-1)} \right) \\ &= m \left(m^{p-1} \right)^{t(q-1)}. \end{aligned}$$

Hence, we have

$$(2.7) \quad (m^e)^d \equiv m \pmod{p},$$

since if p divides m then both sides of Equation (2.7) are $0 \pmod{p}$, and if p does not divide m (i.e., $\gcd(p, m) = 1$) then by Fermat's Little Theorem, we have

$$m^{p-1} \equiv 1 \pmod{p}.$$

By a symmetric argument, it holds that

$$(m^e)^d \equiv m \pmod{q}.$$

Since p and q are primes with $p \neq q$, it follows that

$$(m^e)^d \equiv m \pmod{n}.$$

Since $m < n$, the claim follows. ■

2.3 Security of RSA and Possible Attacks on RSA

Figure 2 summarizes the single steps of the RSA protocol and displays the information communicated by Alice and Bob that is subject to eavesdropping by Erich.

The security of the RSA cryptosystem strongly depends on whether factorizing large integers is intractable, as is widely believed: there is no efficient factorization algorithm known, despite considerable efforts in the past to design such algorithms. However, it is not known whether the problem of factorizing large integers and the problem of cracking the RSA system are equally hard.

Here is a list of potential attacks on the RSA system. To preclude these direct attacks, some care must be taken in choosing the primes p and q , the modulus n , the exponent e , and the private key d . For further background on the security of the RSA system and on proposed attacks to break it, the reader is referred to [Sha95, KR95, Moo92]. For each attack on RSA that has been proposed in the literature to date, some practical countermeasures are known, rules of thumb that prevent the success of those attacks or, at least, that make their likelihood of success negligibly small.

Factorization attacks: The aim of the attacker Erich is to use the public key (n, e) to recover the private key d by factorizing n , i.e., by computing the primes p and q with $n = pq$. Knowing p and q , he can just like Bob compute $\Phi(n) = (p-1)(q-1)$ and thus the inverse d of e modulo $\Phi(n)$, using the extended algorithm of Euclid; see Figure 1 and Lemma 2.3. There are various ways in which Erich might mount this type of attack on RSA.

Step	 Alice		 Bob
1			chooses large primes p, q at random, computes $n = pq$ and $\Phi(n) = (p - 1)(q - 1)$, his public key (n, e) with e satisfying Eq. (2.4), and his private key d satisfying Eq. (2.5)
2		(n, e) ←	
3	encrypts message m by computing $c = m^e \pmod n$		
4		c ⇒	
5			decrypts chiffré c by computing $c^d = (m^e)^d \pmod n = m$

Figure 2: The RSA protocol.

- *Brute-force attack*: Erich might try to factorize the modulus n simply by exhaustive search of the complete key space. Choosing n sufficiently large will prevent this type of attack. Currently, it is recommended to use moduli n with at least 768 bits, i.e., the size of 512 bits formerly in use no longer provides adequate protection today. Of course, the time complexity of modular exponentiation grows rapidly with the modulus size, and thus there is a tradeoff between increasing the security of RSA and decreasing its efficiency.

It is also generally accepted that those moduli n consisting of prime factors p and q of roughly the same size are the hardest to factorize.

- *Special-purpose factorization methods*: Depending on the form of the primes p and q , it might be argued that using “special-purpose” factorization methods such as Pollard’s “ $p - 1$ method” [Pol74] may be more effective and more successful than using “general-purpose” factorization methods such as the *general number field sieve* or the older *quadratic sieve*. This potential threat led to the introduction of *strong primes* that resist such special-purpose factorization methods. A strong prime p is required to satisfy certain conditions such as that $p - 1$ has a large factor r and $r - 1$, in turn, has a large factor, etc.
- *Elliptic curve method*: This factorization method was introduced by Lenstra [Len87], and it has some success probability regardless of the form of the primes chosen. Consequently, the most effective countermeasure against the elliptic curve method is to use primes of very large size. This countermeasure simultaneously provides, with a very high probability, protection against all types of special-purpose factorization methods. In short, large primes are more important than strong primes.

Superencryption: Simmons and Norris [SN77] early proposed an attack on RSA called superencryption that is based on the observation that a sufficient number of encryptions may eventually recover the original message. This attack is a threat to the security of RSA, provided that the number of encryptions required is small. Luckily, superencryption is not a practical attack if the primes are large and are chosen at random.

Wiener's attack: Wiener [Wie90] proposed an attack on the RSA system by a continued fraction approximation, using the public key (n, e) to provide sufficient information to recover the private key d . This attack is efficient and practical, and thus is a concern, only if the private key d is chosen to be small relative to the moduli n ; more precisely, only if $d < n^{\frac{1}{4}}$. However, since the exponent e is chosen first, it is unlikely that a small d will be generated: if e is small enough then d will be large enough to resist Wiener's attack.

Small-message attack: The RSA encryption is not effective if both the message m to be encrypted and the exponent e to be used for encryption are small relative to the modulus n . In particular, if $c = m^e < n$ is the cipher text, then m can be recovered from c by ordinary root extraction. Thus, either the public exponent should be large or the messages should always be large. It is this latter suggestion that is more useful, for a small public exponent is often preferred in order to speed up the encryption and to preclude Wiener's attack.

Low-exponent attack: One should be cautious, though, to not choose the public exponent too small. A preferred value of e that has been used often in the past is $e = 3$. However, if three parties participating in the same system encrypt the same message m using the same public exponent 3, although perhaps different moduli n_1, n_2 , and n_3 , then one can easily compute m from the three cipher texts:

$$\begin{aligned}c_1 &= m^3 \pmod{n_1} \\c_2 &= m^3 \pmod{n_2} \\c_3 &= m^3 \pmod{n_3}.\end{aligned}$$

In particular, the message m must be smaller than the moduli, and so m^3 will be smaller than $n_1 n_2 n_3$. Using the Chinese Remainder Theorem [Knu81], one can compute the unique solution

$$c = m^3 \pmod{n_1 n_2 n_3} = m^3.$$

Hence, one can compute m from c by ordinary root extraction.

More generally, suppose that k related plain texts are encrypted with the same exponent e :

$$\begin{aligned}c_1 &= (a_1 m + b_1)^e \pmod{n_1} \\c_2 &= (a_2 m + b_2)^e \pmod{n_2} \\&\vdots \\c_k &= (a_k m + b_k)^e \pmod{n_k},\end{aligned}$$

where a_i and b_i , $1 \leq i \leq k$, are known and $k > \frac{e(e+1)}{2}$ and $\min(n_i) > 2e^2$. Then, an attacker can solve for m in polynomial time using lattice reduction techniques. This observation is due to Johan Håstad [Hås88]. This attack is a concern if the messages are related in a known way. Padding the messages with pseudorandom strings prior to encryption prevents to mount this attack in practice, see, e.g., [KR95]. If the messages are related in a known way, they should not be encrypted with many RSA keys.

A recommended value of e that is commonly used today is $e = 2^{16} + 1$. One advantage of this value for e is that its binary expansion has only two ones, which implies that the square-and-multiply algorithm of Method 2.5 requires very few operations,⁴ and so is very efficient.

⁴How many exactly?

Forging RSA signatures: This attack is based on the fact that the RSA encryption function is a homomorphism: if (n, e) is the public key and m_1 and m_2 are two messages then

$$(2.8) \quad m_1^e \cdot m_2^e \equiv (m_1 \cdot m_2)^e \pmod{n}.$$

Another identity that can easily be verified is:

$$(2.9) \quad (m \cdot r^e)^d \equiv m^d \cdot r \pmod{n}.$$

In particular, these identities can be used to mount an attack on the digital signature scheme based on the RSA algorithm, see Figure 6 and Section 3.3. Given previous message-signature pairs $(m_1, \text{sig}_A(m_1)), \dots, (m_k, \text{sig}_A(m_k))$, Erich can use the congruences (2.8) and (2.9) to compute a new message-signature pair $(m, \text{sig}_A(m))$ by

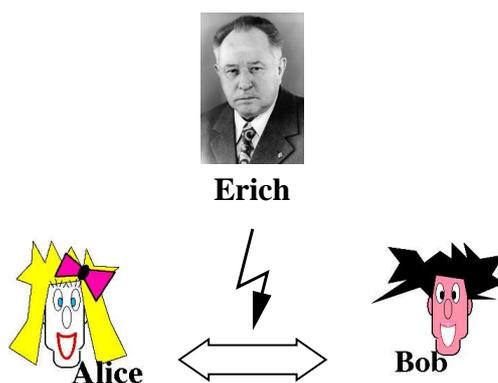
$$m = r^e \prod_{i=1}^k m_i^{e_i} \pmod{n};$$

$$\text{sig}_A(m) = r \prod_{i=1}^k (\text{sig}_A(m_i))^{e_i} \pmod{n},$$

where r and the e_i are arbitrary. Hence, Erich can forge Alice's signature without knowing her private key, and Bob will not detect the forgery, since $m \equiv (\text{sig}_A(m))^e \pmod{n}$. Note that, in Equation (2.8), even if m_1 and m_2 are meaningful plain texts, $m_1 \cdot m_2$ usually is not. Thus, Erich can forge Alice's signature only for messages that may or may not be useful. However, he might choose the messages m_i so as to generate a meaningful message m with a forged digital signature. This *chosen-plain-text attack* can again be avoided by pseudorandom padding techniques that destroy the algebraic relations between messages. Pseudorandom padding is also a useful countermeasure against the following *chosen-cipher-text attack*: Erich intercepts some chiffré c , chooses $r \in \mathbb{N}$ at random, and computes $c \cdot r^e \pmod{n}$, which he sends to the legitimate receiver Bob. By Equation (2.9), Bob will decrypt the string $\hat{c} = c^d \cdot r \pmod{n}$, which is likely to look like a random string. Erich, however, if he were to get his hands on \hat{c} , could obtain the original message m by multiplying by r^{-1} , the inverse of r modulo n , i.e., by computing $m = r^{-1} \cdot c^d \cdot r \pmod{n}$.

3 Protocols for Secret-Key Agreement, Public-Key Encryption, and Digital Signatures

Consider again the scenario that Alice and Bob want to exchange messages over an insecure channel such as a public telephone line, which is eavesdropped by Erich:



That is why Alice and Bob want to encrypt their messages. For efficiency purposes, they decide to use a symmetric cryptosystem in which they both possess the same key for encryption and for decryption; recall Definition 1.1. But then, how can they agree on a joint secret key when they can communicate only over an insecure channel? If they were to send an encrypted message containing the key to be used in subsequent communications, which key should they use to encrypt *this* message?

This paradoxical situation is known as the *secret-key agreement* problem, and it was considered to be unsolvable since the beginning of cryptography. It was quite a surprise when in 1976 Whitfield Diffie and Martin Hellman [DH76] did solve this long-standing, seemingly paradoxical problem by proposing the first secret-key agreement protocol. We describe their protocol in Section 3.1. Interestingly, it was the Diffie-Hellman protocol that inspired Rivest, Shamir, and Adleman to invent the RSA system. That is, Diffie and Hellman’s key idea to solve the secret-key agreement problem opened the door to modern public-key cryptography, which no longer requires sending secret keys over insecure channels.

Section 3.2 shows how to modify the Diffie-Hellman protocol in order to obtain a public-key cryptosystem. This protocol is due to Taher ElGamal [ElG85]. Just like the Diffie-Hellman protocol, ElGamal’s cryptosystem is based on the difficulty of computing discrete logarithms.

Section 3.4 gives an interesting protocol that requires no keys to be agreed upon prior to exchanging encrypted messages. This protocol is due to an unpublished work of Adi Shamir.

Another cryptographic task is the generation of *digital signatures*: Alice wants to sign her encrypted messages to Bob in a way that allows Bob to verify that Alice was indeed the sender of the message. Digital signature protocols are used for the authentication of documents such as email messages. The goal is to preclude Erich from forging Alice’s messages and her signature. Digital signature protocols are described in Section 3.3 and Section 3.5.

3.1 Diffie and Hellman’s Secret-Key Agreement Protocol

Step	 Alice		 Bob
1	Alice and Bob agree upon a large prime p and a primitive root g of p ; p and g are public		
2	chooses a large number a at random, computes $\alpha = g^a \pmod p$		chooses a large number b at random, computes $\beta = g^b \pmod p$
3		$\xrightarrow{\alpha}$ $\xleftarrow{\beta}$	
4	computes her key $k_A = \beta^a \pmod p$		computes his key $k_B = \alpha^b \pmod p$

Figure 3: The Diffie-Hellman secret-key agreement protocol.

Figure 3 shows how the Diffie-Hellman secret-key agreement protocol works. It is based on the modular exponential function with base g and modulus p , where p is a prime and g is a primitive root of p in \mathbb{Z}_p^* , the cyclic group of prime residues modular p ; recall that \mathbb{Z}_p^* has order $\Phi(p) = p - 1$. The formal definition is as follows.

Definition 3.1 Let p be a prime, and let \mathbb{Z}_p^* be the set of all integers a with $1 \leq a \leq p - 1$ and $\gcd(a, p) = 1$.

- A primitive root of p is any element $a \in \mathbb{Z}_p^*$ satisfying that for each d with $1 \leq d < \Phi(p)$

$$a^d \not\equiv 1 \pmod{p}.$$

- Let g be a primitive root of p . The function $\alpha_{(g,p)} : \mathbb{Z}_p^* \rightarrow \mathbb{Z}_p^*$ that is defined by

$$\alpha_{(g,p)}(a) = g^a \pmod{p}.$$

is called the modular exponential function with base g and modulus p . Its inverse function, which for fixed p and g maps $\alpha_{(g,p)}(a)$ to $a = \log_g \alpha \pmod{p}$, is called the discrete logarithm.

Note that every primitive root of p generates the group \mathbb{Z}_p^* , and \mathbb{Z}_p^* has precisely $\Phi(p - 1)$ primitive roots. For example, $\mathbb{Z}_5^* = \{1, 2, 3, 4\}$ and $\mathbb{Z}_4^* = \{1, 3\}$, so $\Phi(4) = 2$, and the two primitive roots of 5 in \mathbb{Z}_5^* are 2 and 3, since

$$\begin{array}{llll} 2^1 = 2; & 2^2 = 4; & 2^3 \equiv 3 \pmod{5}; & 2^4 \equiv 1 \pmod{5}; \\ 3^1 = 3; & 3^2 \equiv 4 \pmod{5}; & 3^3 \equiv 2 \pmod{5}; & 3^4 \equiv 1 \pmod{5}. \end{array}$$

Not every integer has a primitive root: 8 is the smallest such example. It is known from elementary number theory that an integer n has a primitive root if and only if n is 1, 2, 4, or is of the form q^k or $2q^k$ for some odd prime q .

Computing discrete logarithms is considered to be a very hard problem: no efficient algorithms are known for solving it. In contrast, the modular exponential function can be computed efficiently, using the fast exponentiation algorithm “square-and-multiply” described as Method 2.5. That is why exponentiation is considered to be a “one-way function,” i.e., a function that is easy to compute but hard to invert. One-way functions play a key role in cryptography; we will discuss them in more detail in Section 4.

If Erich is listening carefully to Alice and Bob’s communication in the Diffie-Hellman protocol (see Figure 3), he knows p , g , α , and β . He wants to compute their joint secret key $k_A = k_B$. If he could solve the discrete logarithm problem efficiently, Erich could easily compute $a = \log_g \alpha \pmod{p}$ and $b = \log_g \beta \pmod{p}$ and, thus, $k_A = \beta^a \pmod{p}$ and $k_B = \alpha^b \pmod{p}$. Fortunately, as noted above, the discrete logarithm problem is viewed as being intractable, so this attack is very unlikely to be a practical threat. On the other hand, it is the only known attack for computing the keys directly from α and β in the Diffie-Hellman protocol. Note, however, that no proof of security for this protocol has been established up to date.

Note also that computing the keys $k_A = k_B$ directly from α and β is not the only possible attack on the Diffie-Hellman protocol. For example, it is vulnerable to the “Man-in-the-middle” attack. That is, Alice and Bob cannot be certain of the authenticity of their respective partners in the communication. Erich, as the “man in the middle,” might pretend to be Alice when communicating with Bob, and he might pretend to be Bob when communicating with Alice. He could intercept $\alpha = g^a \pmod{p}$ that Alice sends to Bob and he could also intercept $\beta = g^b \pmod{p}$ that Bob sends to Alice, passing on his own values α_E in place of α to Bob and β_E in place of β to Alice. That way Erich could compute two (possibly distinct) keys, one for communicating with Alice, the other one for communicating with Bob, without them having any clue that they in fact are communicating with him. In Section 5, we will introduce *zero-knowledge protocols*, which can be used to ensure proper authentication.

Step	 Alice		 Bob
1	Alice and Bob agree upon a large prime p and a primitive root g of p ; p and g are public		
2			chooses a large number b at random as his private key and computes $\beta = g^b \pmod p$
3		β \leftarrow	
4	chooses a large number a at random, computes $\alpha = g^a \pmod p$, the key $k = \beta^a \pmod p$, and the chiffré $c = E_k(m)$, where m is the message to be sent		
5		α, c \Rightarrow	
6			computes $k = \alpha^b \pmod p$ and $m = D_k(c)$

Figure 4: The ElGamal public-key cryptosystem, which uses the encryption and decryption algorithms E_k and D_k of a given symmetric cryptosystem.

3.2 ElGamal’s Public-Key Cryptosystem and Digital Signature Protocol

By slightly modifying the Diffie-Hellman protocol, it is possible to obtain a public-key cryptosystem. This observation is due to Taher ElGamal [ElG85]. The variant of the ElGamal protocol presented here is not the original one; instead, it is a “hybrid system,” a public-key cryptosystem making use of a given symmetric cryptosystem. Such hybrid systems are often useful in practice, for they combine the advantages of asymmetric and symmetric cryptosystems: symmetric systems are usually more efficient than public-key systems.

The protocol works as follows. Alice and Bob agree on a large prime p and a primitive root g of p , which are public. They also agree on some symmetric cryptosystem $S = (\mathcal{M}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ with encryption functions $\mathcal{E} = \{E_k \mid k \in \mathcal{K}\}$ and decryption functions $\mathcal{D} = \{D_k \mid k \in \mathcal{K}\}$. The subsequent steps of the protocol are shown in Figure 4.

ElGamal’s system modifies the Diffie-Hellman protocol in the following way. While in the Diffie-Hellman scheme Alice and Bob *simultaneously* compute and send their “partial keys” α and β , respectively, they do so *sequentially* in the ElGamal protocol. That is, Alice must wait for Bob’s value β to be able to compute the key k with which she then encrypts her message m via the symmetric cryptosystem S . Moreover, in the ElGamal protocol, Bob generates, once and for all, his public β for possibly several communications with Alice, and also for possibly several users other than Alice who might want to communicate with him. In contrast, Alice has to generate her α anew again and again every time she communicates with Bob, just like in the Diffie-Hellman protocol.

ElGamal’s system can be modified so as to yield a digital signature protocol. A particularly efficient variant of this protocol that is due to an idea of Schnorr [Sch90] is now the United States “Digital Signature Standard” [Nat91, Nat92].

The ElGamal digital signature protocol is presented in Figure 5. Suppose that Bob wants to send a

Step	 Alice	 Bob	 Bob
1	Alice and Bob agree upon a large prime p and a primitive root g of p ; p and g are public		
2			chooses b and $\beta = g^b \pmod p$ as in Fig. 4; chooses a number r with $\gcd(r, p-1) = 1$, computes $\rho = g^r \pmod p$ and s according to Eq. (3.10) and his signature $\text{sig}_B(m) = (\rho, s)$
3		$\beta, m, \text{sig}_B(m) = (\rho, s)$ 	
4	verifies Bob's signature by checking that Eq. (3.11) holds: $g^m \equiv \beta^\rho \cdot \rho^s \pmod p.$		

Figure 5: The ElGamal digital signature protocol.

message m to Alice. Let a large prime p and a primitive root g of p be given as in the ElGamal public-key cryptosystem, see Figure 4. As in that protocol, Bob chooses his private b and computes $\beta = g^b \pmod p$. In addition, he now chooses a number r coprime with $p - 1$, and he computes $\rho = g^r \pmod p$ and a solution s to the congruence

$$(3.10) \quad b \cdot \rho + r \cdot s \equiv m \pmod{p - 1}$$

using the extended algorithm of Euclid, see Figure 1 and Lemma 2.3.

Bob keeps b and r secret, and he sends along with his message m his digital signature $\text{sig}_B(m) = (\rho, s)$ and the value β to Alice.

Alice checks the validity of the signature by verifying the congruence

$$(3.11) \quad g^m \equiv \beta^\rho \cdot \rho^s \pmod p.$$

The protocol is correct, since by Fermat's Little Theorem (see Theorem 2.2) and by Equation 3.10, it holds that

$$g^m \equiv g^{b \cdot \rho + r \cdot s} \equiv \beta^\rho \cdot \rho^s \pmod p.$$

3.3 RSA Digital Signature Protocol

Just as the ElGamal public-key cryptosystem can be modified to a digital signature protocol, also the RSA public-key cryptosystem described in Section 2.2 can be modified so as to yield a digital signature protocol. Figure 6 shows how the RSA digital signature protocol works. A chosen-plain-text attack on the RSA digital signature scheme, and countermeasures to avoid it, are described above in Section 2.3.

Step	 Alice		 Bob
1	chooses $n = pq$, her public key (n, e) , and her private key d as in the RSA protocol, see Section 2.2		
2	computes her signature $\text{sig}_A(m) = m^d \pmod n$ for the message m		
3		$m, \text{sig}_A(m)$ \Rightarrow	
4			verifies Alice's signature by checking the congruence $m \equiv (\text{sig}_A(m))^e \pmod n$

Figure 6: The RSA digital signature protocol.

Step	 Alice		 Bob
1	Alice and Bob agree upon a large prime p , which is public		
2	computes $x = m^a \pmod p$, where m is the message		
3		x \Rightarrow	
4			computes $y = x^b \pmod p$
5		y \Leftarrow	
6	computes $z = y^{a^{-1}} \pmod p$		
7		z \Rightarrow	
8			computes $m = z^{b^{-1}} \pmod p$

Figure 7: Shamir's no-key protocol.

3.4 Shamir's No-Key Protocol

Adi Shamir proposed a symmetric cryptosystem by which Alice and Bob can exchange messages that are encrypted by Alice's and Bob's individual secret keys, yet in which there is no need for Alice and Bob to previously agree on a *joint* secret key. This clever idea is described in an unpublished paper of Shamir, and it is again based on the modular exponentiation function and the difficulty of efficiently computing discrete logarithms that was useful for the Diffie-Hellman secret-key agreement protocol described in Section 3.1.

Figure 7 shows how Shamir's no-key protocol works. Let m be the message that Alice wants to send to Bob. First, Alice and Bob agree on a large prime p . Alice generates a pair (a, a^{-1}) satisfying

$$aa^{-1} \equiv 1 \pmod{p-1},$$

where a^{-1} is the inverse of a modulo $p-1$, see Section 2.2. Similarly, Bob generates a pair (b, b^{-1}) satisfying

$$bb^{-1} \equiv 1 \pmod{p-1},$$

where b^{-1} is the inverse of b modulo $p-1$.

The protocol is correct, since for all messages m , $1 \leq m \leq p$, it holds that:

$$m \equiv m^{aa^{-1}} \pmod{p} \quad \text{and} \quad m \equiv m^{bb^{-1}} \pmod{p}.$$

Hence, looking at Figure 7, we obtain

$$z^{b^{-1}} \equiv y^{a^{-1}b^{-1}} \equiv x^{ba^{-1}b^{-1}} \equiv m^{aba^{-1}b^{-1}} \equiv m \pmod{p},$$

so Step 8 of Figure 7 is correct.

Note that modular exponentiation is used here as a *symmetric* encryption and decryption function. Shamir's no-key protocol is a symmetric cryptosystem, since from a given prime p and a given integer $a \in \mathbb{Z}_p^*$ the inverse a^{-1} of a modulo $p-1$ can easily be computed. The key property for this protocol to work is that modular exponentiation is symmetric in the exponents, i.e., for all a and b , it holds that

$$\alpha_{(g,p)}(a \cdot b) \equiv g^{a \cdot b} \equiv g^{b \cdot a} \pmod{p}.$$

3.5 Rivest, Rabi, and Sherman's Secret-Key Agreement and Digital Signature Protocols

Ron Rivest, Muhammad Rabi, and Alan Sherman developed secret-key agreement and digital signature protocols. The secret-key agreement protocol from Figure 8 is attributed to Rivest and Sherman in [RS93, RS97]. The digital signature protocol from Figure 9 is due to Rabi and Sherman [RS93, RS97].

The key building block of both protocols is a strongly noninvertible, associative one-way function, which we will formally define in the present section. The property of associativity requires that the functions we consider are two-ary (i.e., two-argument) functions; we will use both prefix and infix notation for such functions. In Section 4, we will see how to construct strongly noninvertible, associative one-way functions from any given one-way function [HR99].

The standard notion of one-way functions to be used here is due to Grollmann and Selman [GS88], see also the papers by Ko [Ko85], Berman [Ber77], and Allender [All85, All86], and the surveys [Sel92, BHR99]. Note that this notion is defined in the model of worst-case complexity, as opposed to the average-case notion of one-way functions (see, e.g., [Gol01b, Gol99]), which is often used in cryptographic applications.

Below, the notion of worst-case one-way functions is tailored to the case of two-ary functions. Let FP denote the set of all polynomial-time computable total functions. The length of any integer $n \in \mathbb{N}$, denoted by $|n|$, is the number of bits needed to represent n in binary without leading zeroes.

Step	 Alice		 Bob
1	chooses two large numbers x and y at random, keeps x secret, and computes $x\sigma y$		
2		$y, x\sigma y$ \Rightarrow	
3			chooses a large number z at random, keeps z secret and computes $y\sigma z$
4		$y\sigma z$ \Leftarrow	
5	computes her key $k_A = x\sigma(y\sigma z)$		computes his key $k_B = (x\sigma y)\sigma z$

Figure 8: The Rivest-Sherman secret-key agreement protocol, which uses a strongly noninvertible, associative one-way function σ .

Step	 Alice		 Bob
1	chooses two large numbers x_A and y_A at random, keeps x_A secret, and computes $x_A\sigma y_A$		
2		$y_A, x_A\sigma y_A$ \Rightarrow	
3	computes her signature $\text{sig}_A(m) = m\sigma x_A$ for the message m		
4		$m, \text{sig}_A(m)$ \Rightarrow	
5			verifies Alice's signature by checking the equality $m\sigma(x_A\sigma y_A) = (m\sigma x_A)\sigma y_A$

Figure 9: The Rabi-Sherman digital signature protocol, which uses a strongly noninvertible, associative one-way function σ .

To preclude the notion of noninvertibility from being trivialized, one-way functions are required to be “honest,” i.e., to not shrink their inputs too much. Note that sometimes an even stronger notion of honesty is required; see, e.g., [HRW97,HR00,RH].

Definition 3.2 [GS88,RS97,HR99] *Let $\rho : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ be any two-ary function; ρ may be nontotal and it may be many-to-one.*

1. *We say that ρ is honest if and only if there exists a polynomial q such that:*

$$(\forall z \in \text{image}(\rho)) (\exists (a, b) \in \text{domain}(\rho)) [|a| + |b| \leq q(|z|) \wedge \rho(a, b) = z].$$

2. *We say that ρ is (polynomial-time) invertible if and only if there exists a function $g \in \text{FP}$ such that for all $z \in \text{image}(\rho)$, $\rho(g(z)) = z$.*
3. *We say that ρ is a one-way function if and only if ρ is honest, polynomial-time computable, and noninvertible.*

We now define strong noninvertibility (strongness, for short). We stress that this property of two-ary functions is independent of noninvertibility. That is, strongness does not necessarily imply noninvertibility. Even worse, one can prove that some strongly noninvertible functions are in fact invertible under the plausible assumption that $\text{P} \neq \text{NP}$. This somewhat surprising result was recently obtained by Hemaspaandra, Pasanen, and Rothe [HPR01]. We further discuss this point in Section 4.2.

As with noninvertibility, strongness requires an appropriate notion of honesty so as to not be trivial. To this end, we introduce the notion of “s-honesty” below, and we define *strongly noninvertible one-way functions* to be easily computable, s-honest, strongly noninvertible functions. Let $\langle \cdot, \cdot \rangle : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ be some total, bijective, polynomial-time computable pairing function that has polynomial-time computable inverses, and is nondecreasing in each argument when the other argument is fixed.

Definition 3.3 (see [RS97,HR99]) *Let $\sigma : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ be any two-ary function; σ may be nontotal and it may be many-to-one.*

1. *We say that σ is s-honest if and only if there exists a polynomial q such that both (a) and (b) hold:*
 - (a) $(\forall z, a : (\exists b) [\sigma(a, b) = z]) (\exists b') [|b'| \leq q(|z| + |a|) \wedge \sigma(a, b') = z].$
 - (b) $(\forall z, b : (\exists a) [\sigma(a, b) = z]) (\exists a') [|a'| \leq q(|z| + |b|) \wedge \sigma(a', b) = z].$
2. *We say that σ is (polynomial-time) invertible with respect to its first argument if and only if there exists a function $g_1 \in \text{FP}$ such that for all $z \in \text{image}(\sigma)$ and for all a and b with $(a, b) \in \text{domain}(\sigma)$ and $\sigma(a, b) = z$, it holds that $\sigma(a, g_1(\langle a, z \rangle)) = z$.*
3. *We say that σ is (polynomial-time) invertible with respect to its second argument if and only if there exists a function $g_2 \in \text{FP}$ such that for all $z \in \text{image}(\sigma)$ and for all a and b with $(a, b) \in \text{domain}(\sigma)$ and $\sigma(a, b) = z$, it holds that $\sigma(g_2(\langle b, z \rangle), b) = z$.*
4. *We say that σ is strongly noninvertible if and only if σ is neither invertible with respect to its first argument nor invertible with respect to its second argument.*
5. *We say that σ is a strong one-way function if and only if σ is s-honest, polynomial-time computable, and strongly noninvertible.*

Finally, we present a formal definition of associativity for two-ary functions that is suitable both for total and for nontotal two-ary functions. This definition is due to Hemaspaandra and Rothe [HR99] (see also the survey [BHHR99]) who note that the definition of associativity given by Rabi and Sherman ([RS93,RS97], see Definition 4.1) intuitively is not adequate: Rabi and Sherman’s associativity fails to capture the nontotal function case appropriately. More precisely, Rabi and Sherman’s associativity fails to preclude, for nontotal functions, equations from having a defined value to the left, while being undefined to the right of their equality sign. The two notions of associativity are provably distinct (see [HR99]), and this distinction can be explained (see [HR99]) via Kleene’s careful discussion [Kle52, pp. 327–328] of two distinct notions of equality for partial functions in recursion theory: “Weak equality” between two partial functions explicitly allows “specific, defined function values being equal to undefined” as long as the functions take the same values on their joint domain. In contrast, “complete equality” precludes this unnatural behavior by additionally requiring that two given partial functions be equal only if their domains coincide; i.e., whenever one is undefined, so is the other.

Definition 3.4 [HR99] *Let $\sigma : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ be any two-ary function; σ may be nontotal. Define $\mathbb{N}_\perp = \mathbb{N} \cup \{\perp\}$, and define an extension $\overset{\perp}{\sigma} : \mathbb{N}_\perp \times \mathbb{N}_\perp \rightarrow \mathbb{N}_\perp$ of σ as follows:*

$$\overset{\perp}{\sigma}(a, b) = \begin{cases} \sigma(a, b) & \text{if } a \neq \perp \text{ and } b \neq \perp \text{ and } (a, b) \in \text{domain}(\sigma) \\ \perp & \text{otherwise.} \end{cases}$$

We say that σ is associative if and only if, for all $a, b, c \in \mathbb{N}$, it holds that $(a \overset{\perp}{\sigma} b) \overset{\perp}{\sigma} c = a \overset{\perp}{\sigma} (b \overset{\perp}{\sigma} c)$. We say that σ is commutative if and only if, for all $a, b \in \mathbb{N}$, it holds that $a \overset{\perp}{\sigma} b = b \overset{\perp}{\sigma} a$.

3.6 Discussion of Diffie e-Hellman versus Rivest-Sherman

While the secret-key agreement protocol of Diffie and Hellman [DH76] is widely used in practice, that of Rivest and Sherman (see [RS93,RS97]) is not (yet) used in applications and, thus, might appear somewhat exotic at first glance. Note, however, that neither the Diffie-Hellman nor the Rivest-Sherman protocol has a proof of security up to date. So, let us digress for a moment to compare the state of the art on these two protocols.

- The security of the Diffie-Hellman scheme is based on the (unproven, yet plausible) assumption that computing discrete logarithms is a computationally intractable task.

In contrast, the Rivest-Sherman scheme uses a strongly noninvertible, associative one-way function (see Section 3.5 for the formal definition) as its key building block. Although it is not known whether such functions exist, it has been shown recently by Hemaspaandra and this author [HR99] that they do exist in the worst-case model under the (unproven, yet plausible) assumption that $P \neq NP$. Section 4 below presents this result and its proof in detail.

- Breaking Diffie-Hellman is not even known to be equally hard as computing discrete logarithms, even though some nice progress in this direction has been made recently by Maurer and Wolf [MW99], who established conditions for relating the hardness of breaking Diffie-Hellman to that of computing discrete logarithms. Again, their results rest on unproven assumptions. In particular, let $\nu(p)$ denote the minimum, taken over all numbers d in the interval $[p - 2\sqrt{p} + 1, p + 2\sqrt{p} + 1]$, of the largest prime factors of d . Under the unproven, yet plausible assumption that $\nu(p)$ is polynomial in $\log p$, Maurer and Wolf [MW99] proved that breaking Diffie-Hellman and computing the discrete logarithm are polynomial-time equivalent tasks in the underlying cyclic group G . They also establish a number of related results.

Similarly, even if strongly noninvertible, associative one-way functions were known to exist, one could not conclude that the Rivest-Sherman protocol is secure; rather, strong noninvertibility merely precludes certain types of direct attacks [RS97,HR99]. Moreover, strongly noninvertible, associative one-way functions could be constructed so far only in the *worst-case* complexity model, assuming $P \neq NP$. Although this result is relevant and interesting in a complexity-theoretic setting, it has no implications in applied cryptography. For cryptographic applications, one would need to construct such functions based on the *average-case* complexity model, under plausible assumptions.

As noted in the outline of the tutorial, there is some hope for obtaining such a strong result by combining the worst-case complexity technique of Hemaspaandra and Rothe [HR99] with Ajtai's results on the shortest lattice vector problem, denoted by SVP. Roughly speaking, Ajtai [Ajt96] proved that the problem SVP is equally hard in the worst-case and in the average-case complexity models. More precisely, Ajtai constructed an infinite family $\{\Lambda_n\}_{n \geq 1}$ of lattices, where each Λ_n is represented by a basis as an instance of SVP, and he showed the following result: Suppose one can compute in polynomial time, for each n , an approximately shortest vector in a lattice Λ_i randomly chosen from $\{\Lambda_n\}_{n \geq 1}$, with non-negligible probability. Then, the length of a shortest vector in every lattice from $\{\Lambda_n\}_{n \geq 1}$ can be estimated to within a fixed polynomial factor in polynomial time with probability close to one. However, since the best approximation factor known to be achieved by polynomial-time algorithms is essentially exponential, and since the best algorithms known to achieve polynomial-factor approximations runs in exponential time, it follows that, as mentioned above, SVP is "equally hard in the worst-case and in the average-case models." Based on the worst-case/average-case equivalence of SVP, Ajtai and Dwork [AD97] designed a public-key cryptosystem.

Ajtai [Ajt98] also established the worst-case NP-hardness of SVP under randomized reductions. Since the construction of strongly noninvertible, associative one-way functions in [HR99] is based on the assumption $P \neq NP$, it seems reasonable to consider the NP-hard problem SVP to be a good candidate for achieving strongly noninvertible, associative one-way functions even in the technically more demanding average-case model.

4 Strongly Noninvertible Associative One-Way Functions

Are the protocols of Rabi, Rivest, and Sherman secure? This question has two aspects: (1) Are they secure under the assumption that strongly noninvertible associative one-way functions indeed exist? (2) What evidence do we have for the existence of such functions?

The first question remains an open problem. Security here depends on precisely how "strong non-invertibility" is defined, and the average-case model is certainly more demanding than our worst-case model; see Definitions 3.3 and 3.2 for worst-case noninvertibility notions. As noted by Rabi and Sherman [RS97], no proof of security for the Rabi-Rivest-Sherman protocols is currently known, not even in the weaker worst-case model. Even if strongly noninvertible associative one-way functions were known to exist, this would not in any obvious way imply that the protocols are secure. In that regard, however, the Rabi-Rivest-Sherman protocols behave just like any other protocol that is currently used in practical applications. For example, neither the Diffie-Hellman protocol nor the RSA protocol currently has a proof of security. There are merely heuristic, intuitive arguments about how to avoid certain direct attacks. The "security" of the Diffie-Hellman protocol draws on the assumption that computing discrete logarithms is hard, and the "security" of the RSA protocol draws on the assumption that factorizing large integers is hard. Breaking Diffie-Hellman is not even known to be equally hard as the discrete logarithm problem, and breaking RSA is not even known to be equally hard as the factorization problem.

In a similar vein, Rabi and Sherman [RS93,RS97] only give intuitive arguments for the security of

their protocols, explaining how to employ the strong noninvertibility of associative one-way functions to preclude certain direct attacks.

Let us now turn to the second question raised above: What evidence do we have that strongly non-invertible associative one-way functions exist? Assuming $P \neq NP$, we will show how to construct strong, total, commutative,⁵ associative one-way functions [HR99]. P denotes the class of polynomial-time solvable problems, and NP denotes the class of problems that can be solved nondeterministically in polynomial time; see, e.g., [Pap94,BDG95,BC93] for more background on complexity theory. The question of whether P equals NP is perhaps the most important question in theoretical computer science. It is widely believed that P differs from NP , although this question has remained open for more than thirty years now.

It is well-known that, in the worst-case model adopted here, one-argument one-way functions exist if and only if $P \neq NP$; see, e.g., [Sel92,BDG95]. It is easy to prove the analogous result for two-argument one-way functions, see [HR99,RS97]. The upcoming constructions of one-way functions with enhanced algebraic and security properties will make use of this fact.

Rabi and Sherman [RS93,RS97] showed that $P \neq NP$ if and only if commutative, weakly associative one-way functions exist.⁶ However, they did not achieve strong noninvertibility. They did not achieve totality of their weakly associative one-way functions either, although they presented a construction that they claimed achieves totality of their weakly associative one-way functions [RS93,RS97].⁷ That is, Rabi and Sherman [RS93,RS97] left open the question of whether there are plausible complexity-theoretic conditions sufficient to ensure the existence of strong, total, commutative, associative one-way functions. They also asked whether such functions could be *constructed* from any given one-way function. Section 4.1 presents the answers to these questions.

4.1 Creating Strongly Noninvertible, Total, Commutative, Associative One-Way Functions from Any One-Way Function

Below, we define Rabi and Sherman’s notion of associativity, which henceforth will be called “weak associativity.”

Definition 4.1 [RS93,RS97] *A two-ary function $\circ : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ is said to be weakly associative if and only if $a \circ (b \circ c) = (a \circ b) \circ c$ holds for all $a, b, c \in \mathbb{N}$ for which each of (a, b) , (b, c) , $(y, b \circ c)$, and $(a \circ b, c)$ belongs to the domain of \circ .*

The following proposition explores the relation between the two associativity notions presented respectively in Definition 3.4 and in Definition 4.1. In particular, these are indeed different notions.

Proposition 4.2 *1. Every associative two-ary function is weakly associative.*
2. Every total two-ary function is associative exactly if it is weakly associative.
3. There exist two-ary functions that are weakly associative, yet not associative.

⁵Commutativity is needed to extend the Rabi-Rivest-Sherman protocols from two parties to $m > 2$ parties.

⁶By “weakly associative” we mean Rabi and Sherman’s notion of associativity ([RS93,RS97], see Definition 4.1), which is grounded on Kleene’s “weak equality” between partial functions. As explained in the paragraph preceding Definition 3.4, weak associativity does not adequately fit the nontotal function case.

⁷Hemaspaandra and Rothe [HR99] showed that Rabi and Sherman’s claim is unlikely to be true: Any proof of this claim would imply that $NP = UP$, where UP denotes Valiant’s class “unambiguous polynomial time” [Val76]. Intuitively, the reason that their construction is unlikely to work is precisely that the functions constructed in [RS93,RS97] are not associative in Kleene’s sense of “complete equality” between partial functions, see Definition 3.4. In contrast, the Rabi-Sherman construction indeed is useful to achieve totality of the associative, strongly noninvertible one-way functions constructed in [HR99].

We now introduce the notion of NP certificates, which will be useful later on. Other common names for “certificate” are “witness” and “proof” and “solution.”

Definition 4.3 *Let $A \subseteq \mathbb{N}$ be any set in NP, and let N be a fixed nondeterministic polynomial-time Turing machine for A ; i.e., $A = L(N)$ is the language accepted by N . Let $x \in A$ be an input accepted by N .*

1. A certificate for “ $x \in A$ ” is an integer $z \in \mathbb{N}$ encoding in binary an accepting computation path of N on input x .
2. $\text{Certificates}_N(x)$ denotes the set of all certificates of N on input x .

Note that $\text{Certificates}_N(x) = \emptyset$ exactly if $x \notin A$. A useful property of NP sets is that, for each certificate z for “ $x \in A$,” it holds that:

- (a) the length of z is polynomially bounded in the length of x , and
- (b) z certifies membership of x in A in a way that can be verified deterministically in polynomial time.

For example, if N is the “standard” NP machine that solves the satisfiability problem by guessing and verifying satisfying assignments to a given boolean formula, then for each satisfiable formula F , $\text{Certificates}_N(F)$ is the set of satisfying assignments to F suitably encoded as binary numbers.

Theorem 4.4 ([HR99], see also [BHHR99]) *If $P \neq NP$ then there exist strong, total, commutative, associative one-way functions.*

Theorem 4.4 above is the main result of this section. Since $P \neq NP$ is equivalent to the existence of one-way functions with no additional properties required, the converse of the implication stated in Theorem 4.4 is clearly also true. However, we will state and prove only the interesting implication directions in Theorem 4.4 and in the upcoming Theorem 4.5 and Theorem 4.7.

Proof of Theorem 4.4. Assume $P \neq NP$. Let $A \subseteq \mathbb{N}$ be a set in $NP - P$, and let M be an NP machine accepting A . Without loss of generality, suppose that for each $x \in A$ and for each certificate z for “ $x \in A$,” it holds that $|z| = p(|x|) > |x|$ for some strictly increasing polynomial p depending on M only. For any integers $u, v, w \in \mathbb{N}$, let $\min(u, v)$ denote the minimum of u and v , and let $\min(u, v, w)$ denote the minimum of u, v , and w .

Define a two-ary function $\sigma : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ as follows:

- If $a = \langle x, z_1 \rangle$ and $b = \langle x, z_2 \rangle$ for some $x \in A$ with certificates $z_1, z_2 \in \text{Certificates}_M(x)$ (where, possibly, $z_1 = z_2$), then define $\sigma(a, b) = \langle x, \min(z_1, z_2) \rangle$;
- if there exists some $x \in A$ with certificate $z \in \text{Certificates}_M(x)$ such that either $a = \langle x, x \rangle$ and $b = \langle x, z \rangle$, or $a = \langle x, z \rangle$ and $b = \langle x, x \rangle$, then define $\sigma(a, b) = \langle x, x \rangle$;
- otherwise, $\sigma(a, b)$ is undefined.

What is the intuition behind the definition of σ ? The number of certificates contained in the arguments of σ is decreased by one in a way that ensures the associativity of σ . Moreover, σ is noninvertible, and it is also strongly noninvertible. Why? The intuition here is that, regardless of whether none or either one of its arguments is given in addition to σ ’s function value, the inversion of σ requires information about the certificates for elements of A . However, our assumption that $A \notin P$ guarantees that this information cannot efficiently be extracted.

Firstly, we will show that σ is a commutative, associative one-way function that is strongly noninvertible. Note that σ is not a total function. Therefore, we will secondly show how to extend σ to a total function without losing any of its other properties already established.

It is clear that σ is both honest and s-honest. That σ is in FP can be seen as follows: On input $\langle a, b \rangle$, check whether $\langle a, b \rangle$ is in the domain of σ and, if so, which of $\langle x, x \rangle$ or $\langle x, z \rangle$ for suitable $x \in \mathbb{N}$ and $z \in \text{Certificates}_M(x)$ has to be output as the function value $\sigma(a, b)$. Here, we need our assumption that for each $x \in A$ and for each certificate z for “ $x \in A$,” it holds that $|z| = p(|x|) > |x|$. This assumption ensures that there are no ambiguities in determining whether a and b are of the form $\langle x, x \rangle$ or $\langle x, z \rangle$ for some potential certificate z for “ $x \in A$.”

Since $\bigcup_{x \in \mathbb{N}} \{ \langle x, z \rangle \mid z \in \text{Certificates}_M(x) \}$ is a set in P, it is easy to check whether or not some potential certificate z in pairs of the form $\langle x, z \rangle$ indeed is a certificate for “ $x \in A$.” This property also implies that the domain of σ is decidable in P.

To see that σ is noninvertible, suppose for a contradiction that there were an inverter $g \in \text{FP}$ for σ . In particular, g on input $\langle x, x \rangle$ would yield a pair $\langle a, b \rangle$ such that either a or b contains a certificate z for “ $x \in A$.” Hence, A could be decided in polynomial time, a contradiction. So, σ is noninvertible.

We now show that σ is also strongly noninvertible. Note that proving a function strongly noninvertible does not suffice to conclude it is noninvertible: assuming $P \neq \text{NP}$, we will construct in Section 4.2 strongly noninvertible functions that in fact are invertible.

For a contradiction, suppose there exists a total function $g_2 \in \text{FP}$ such that, for each $w \in \text{image}(\sigma)$ and for each second argument b for which there is an $a \in \mathbb{N}$ with $\sigma(a, b) = w$, it holds that

$$\sigma(g_2(\langle b, w \rangle), b) = w.$$

Then, contradicting our assumption that $A \notin \text{P}$, one could decide A in polynomial time as follows:

On input x , compute $g_2(\langle \langle x, x \rangle, \langle x, x \rangle \rangle)$, compute the integers d and e for which $\langle d, e \rangle$ equals $g_2(\langle \langle x, x \rangle, \langle x, x \rangle \rangle)$, and accept x if and only if $d = x$ and $e \in \text{Certificates}_M(x)$.

Hence, σ is not invertible with respect to its second argument. An analogous argument shows that σ is not invertible with respect to its first argument. Thus, σ is strongly noninvertible.

Next, we prove that σ is associative. Let $\overset{\perp}{\sigma}$ be the total extension of σ as in Definition 3.4. Fix any three elements of \mathbb{N} , say $a = \langle a_1, a_2 \rangle$, $b = \langle b_1, b_2 \rangle$, and $c = \langle c_1, c_2 \rangle$. To show that

$$(4.12) \quad (a \overset{\perp}{\sigma} b) \overset{\perp}{\sigma} c = a \overset{\perp}{\sigma} (b \overset{\perp}{\sigma} c)$$

holds, distinguish two cases.

Case 1: $a_1 = b_1 = c_1$ and $\{a_2, b_2, c_2\} \subseteq \{a_1\} \cup \text{Certificates}_M(a_1)$.

Let $x, y \in \{a, b, c\}$ be any two fixed arguments of σ . As noted above, if x and y together contain i certificates for “ $a_1 \in A$,” where $1 \leq i \leq 2$, then $\sigma(x, y)$ —and thus also $\overset{\perp}{\sigma}(x, y)$ —contains exactly $\max\{0, i - 1\}$ certificates for “ $a_1 \in A$.” In particular, $\overset{\perp}{\sigma}(x, y)$ preserves the minimum certificate if both x and y contain a certificate for “ $a_1 \in A$.”

If exactly one of x and y contains a certificate for “ $a_1 \in A$,” then $\overset{\perp}{\sigma}(x, y) = \langle a_1, a_1 \rangle$.

If none of x and y contains a certificate for “ $a_1 \in A$,” then $\sigma(x, y)$ is undefined, so $\overset{\perp}{\sigma}(x, y) = \perp$.

Let $k \leq 3$ be a number telling us how many of a_2, b_2 , and c_2 belong to $\text{Certificates}_M(a_1)$. For example, if $a_2 = b_2 = c_2 \in \text{Certificates}_M(a_1)$ then $k = 3$. Consequently:

- If $k \leq 1$ then $(a \overset{\perp}{\sigma} b) \overset{\perp}{\sigma} c = \perp = a \overset{\perp}{\sigma} (b \overset{\perp}{\sigma} c)$.

- If $k = 2$ then $(a \overset{\perp}{\sigma} b) \overset{\perp}{\sigma} c = \langle a_1, a_1 \rangle = a \overset{\perp}{\sigma} (b \overset{\perp}{\sigma} c)$.
- If $k = 3$ then $(a \overset{\perp}{\sigma} b) \overset{\perp}{\sigma} c = \langle a_1, \min(a_2, b_2, c_2) \rangle = a \overset{\perp}{\sigma} (b \overset{\perp}{\sigma} c)$.

In each of these three cases, Equation (4.12) is satisfied.

Case 2: Not Case 1.

Then, either $(a_1 \neq b_1 \vee a_1 \neq c_1 \vee b_1 \neq c_1)$, or it holds that $(a_1 = b_1 = c_1 \wedge \{a_2, b_2, c_2\} \not\subseteq \{a_1\} \cup \text{Certificates}_M(a_1))$. By the definition of σ , in both cases it follows that

$$(a \overset{\perp}{\sigma} b) \overset{\perp}{\sigma} c = \perp = a \overset{\perp}{\sigma} (b \overset{\perp}{\sigma} c),$$

which satisfies Equation (4.12) and concludes the proof that σ is associative.

The commutativity of σ follows immediately from its definition. Hence, σ is a commutative, associative one-way function that is strongly noninvertible. To complete the proof, we now show how to extend σ to a *total* commutative, associative one-way function that is strongly noninvertible.

Since $A \not\subseteq P$, there exists an integer, say $x_0 \in \mathbb{N}$, that does not belong to A . Let $a_0 = \langle x_0, 1 + x_0 \rangle$. Note that a_0 is neither of the form $\langle x, x \rangle$ for any $x \in \mathbb{N}$, nor of the form $\langle x, z \rangle$ for any $x \in \mathbb{N}$ and any certificate $z \in \text{Certificates}_M(x)$, since $x_0 \notin A$ and so does not have any certificates. By the definition of σ , for each $b \in \mathbb{N}$, neither (a_0, b) nor (b, a_0) is in the domain of σ . Define a total function $\tau : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ as follows:

$$\tau(a, b) = \begin{cases} \sigma(a, b) & \text{if } (a, b) \in \text{domain}(\sigma) \\ a_0 & \text{otherwise.} \end{cases}$$

We claim that τ is a total, commutative, associative one-way function that is strongly noninvertible. In particular, τ is honest, since for a_0 , which is the only string in the image of τ that is not in the image of σ , it holds that $\tau(a_0, a_0) = a_0$ and $|a_0| + |a_0| \leq 2|a_0|$. Similarly, τ is s-honest. Also, $\tau \in \text{FP}$, since $\sigma \in \text{FP}$ and, as noted above, the domain of σ is in P . That τ is noninvertible and strongly noninvertible follows from the facts that $\text{image}(\sigma) \subseteq \text{image}(\tau)$ and that σ is both noninvertible and strongly noninvertible. To see that τ is associative, note that if $a \overset{\perp}{\sigma} (b \overset{\perp}{\sigma} c) = \perp$ then $a\tau(b\tau c) = a_0$, and otherwise $a\tau(b\tau c) = a \overset{\perp}{\sigma} (b \overset{\perp}{\sigma} c)$. Similarly, if $(a \overset{\perp}{\sigma} b) \overset{\perp}{\sigma} c = \perp$ then $(a\tau b)\tau c = a_0$, and otherwise $(a\tau b)\tau c = (a \overset{\perp}{\sigma} b) \overset{\perp}{\sigma} c$. The associativity of τ now follows immediately from the associativity of σ . The commutativity of τ is immediate from the definition of τ and the commutativity of σ . Hence, τ has all the properties required. ■

4.2 If $P \neq NP$ then Some Strongly Noninvertible Functions are Invertible

Is every strongly noninvertible function noninvertible? Hemaspaandra, Pasanen, and Rothe [HPR01] obtained the surprising result that if $P \neq NP$ then this is not necessarily the case.

Theorem 4.5 [HPR01] *If $P \neq NP$ then there exists a total, honest two-ary function that is strongly one-way but not a one-way function.*

Proof. Assume $P \neq NP$. Then, there exists a total two-ary one-way function, call it ρ . For any integer $n \in \mathbb{N}$, define the notation

$$\text{odd}(n) = 2n + 1 \quad \text{and} \quad \text{even}(n) = 2n.$$

Define a function $\sigma : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ as follows. Let $a, b \in \mathbb{N}$ be any two arguments of σ .

- If $a \neq 0 \neq b$, $a = \langle x, y \rangle$ is odd, and b is even, then define $\sigma(a, b) = \text{even}(\rho(x, y))$.
- If $a \neq 0 \neq b$, a is even, and $b = \langle x, y \rangle$ is odd, then define $\sigma(a, b) = \text{even}(\rho(x, y))$.
- If $a \neq 0 \neq b$, and a is odd if and only if b is odd, then define $\sigma(a, b) = \text{odd}(a + b)$.
- If $a = 0$ or $b = 0$, then define $\sigma(a, b) = a + b$.

It is a matter of routine to check that σ is polynomial-time computable, total, honest, and s-honest, the last property being true regardless of whether or not ρ , which is honest, is s-honest.

For a contradiction, suppose σ were invertible with respect to its first argument via an inverter $g_1 \in \text{FP}$. By the definition of σ , for any $z \in \text{image}(\rho)$ with $z \neq 0$, the function g_1 on input $\langle 2, \text{even}(z) \rangle$ yields an odd integer b from which we can read the pair $\langle x, y \rangle$ with $\rho(x, y) = z$. Hence, using g_1 , one could invert ρ in polynomial time, a contradiction. Thus, σ is not invertible with respect to its first argument. Analogously, one can show that σ is not invertible with respect to its second argument. So, σ is strongly noninvertible.

But σ is invertible. By the fourth line in the definition of σ , every z in the image of σ has a preimage of the form $(0, z)$. Thus, the function g defined by $g(z) = (0, z)$ inverts σ in polynomial time. Hence, σ is not a one-way function. ■

Why don't we use a different notion of strongness that automatically implies noninvertibility? Here is an attempt to redefine the notion of strongness accordingly, which yields a new notion that we will call "overstrongness."

Definition 4.6 [HPR01] *Let $\sigma : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ be any two-ary function; σ may be nontotal and it may be many-to-one. We say that σ is overstrong if and only if for no $f \in \text{FP}$ with $f : \{1, 2\} \times \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ does it hold that for each $i \in \{1, 2\}$ and for each $z, a \in \mathbb{N}$:*

$$((\exists b \in \mathbb{N})[(\sigma(a, b) = z \wedge i = 1) \vee (\sigma(b, a) = z \wedge i = 2)]) \implies \sigma(f(i, z, a)) = z.$$

Note that overstrongness implies both noninvertibility and strong noninvertibility. However, the problem with this new definition is that it completely loses the core of why strongness precludes direct attacks on the Rabi-Rivest-Sherman protocols. To see why, look at Figure 8 and Figure 9, which give the protocols of Rabi, Rivest, and Sherman. In contrast to overstrongness, Rabi, Rivest, and Sherman's original definition of strong noninvertibility (see Definition 3.3) *respects the argument given*. It is this feature that precludes Erich from being able to compute Alice's secret x from the transmitted values $x\sigma y$ and y , which he knows. In short, overstrongness is *not well-motivated* by the protocols of Rabi, Rivest, and Sherman.

We mention without proof some further results of Hemaspaandra, Pasanen, and Rothe [HPR01]. Only for the last item of Theorem 4.7 will we give a short proof sketch.

Theorem 4.7 [HPR01]

1. *If $\text{P} \neq \text{NP}$ then there exists a total, honest, s-honest, two-ary overstrong function. Consequently, if $\text{P} \neq \text{NP}$ then there exists a total two-ary function that is both one-way and strongly one-way.*
2. *If $\text{P} \neq \text{NP}$ then there exists a total, s-honest two-ary one-way function σ such that σ is invertible with respect to its first argument and σ is invertible with respect to its second argument.*
3. *If $\text{P} \neq \text{NP}$ then there exists a total, s-honest two-ary one-way function that is invertible with respect to either one of its arguments (thus, it is not strongly one-way), yet that is not invertible with respect to its other argument.*

4. If $P \neq NP$ then there exists a total, honest, s -honest two-ary function that is noninvertible and strongly noninvertible but that is not overstrong.

Proof Sketch for Part 4 of Theorem 4.7. Assuming $P \neq NP$, let $\widehat{\rho}$ be a total one-argument one-way function that additionally satisfies that, for some $k \geq 2$, $|\widehat{\rho}(x)| = |x|^k + k$ holds for each $x \in \mathbb{N}$.

Define a total one-argument function $\rho : \mathbb{N} \rightarrow \mathbb{N}$ as follows. Let $a \in \mathbb{N}$ be any given argument of ρ .

- If a is odd, then define $\rho(a) = \text{odd}(\widehat{\rho}(a))$;
- if a is even, then define $\rho(a) = a$.

Define a total two-ary function $\sigma : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ as follows. Let $a, b \in \mathbb{N}$ be any two arguments of σ .

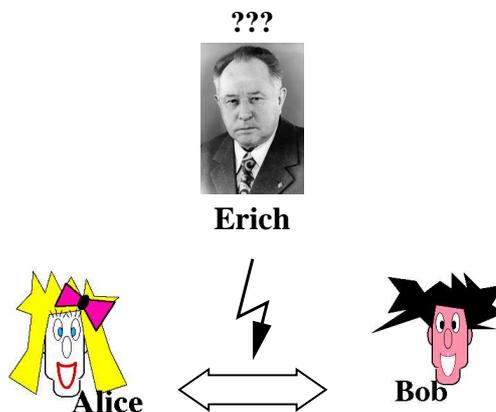
- If there exist $x, y \in \mathbb{N}$ with $|x| = |y|$ such that $a = \text{even}(\langle x, y \rangle) = b$, then define $\sigma(a, b) = \text{odd}(\langle \rho(x), 2^{|y|} - 1 \rangle)$. Note that, in binary representation, $|2^{|y|} - 1| = |y|$; for example, if $y = 5$ is given by 101 in binary, then $|2^{|101|} - 1| = |111| = 3 = |101|$.
- If there exist $x, y \in \mathbb{N}$ with $|x| = |y|$ such that $a = \text{odd}(\langle x, \text{even}(y) \rangle)$ and $b = \text{odd}(\langle x, \text{odd}(\widehat{\rho}(y)) \rangle)$, then define $\sigma(a, b) = \text{odd}(\langle \rho(x), 2^{|y|} - 1 \rangle)$.
- If there exist $x, y \in \mathbb{N}$ with $|x| = |y|$ such that $a = \text{odd}(\langle x, \text{odd}(\widehat{\rho}(y)) \rangle)$ and $b = \text{odd}(\langle x, \text{even}(y) \rangle)$, then define $\sigma(a, b) = \text{odd}(\langle \rho(x), 2^{|y|} - 1 \rangle)$.
- Otherwise, define $\sigma(a, b) = \text{even}(\langle a, b \rangle)$.

We leave it to the reader to verify that σ is polynomial-time computable, honest, s -honest, commutative, noninvertible, and strongly noninvertible, but not overstrong. The last property is the least obvious.

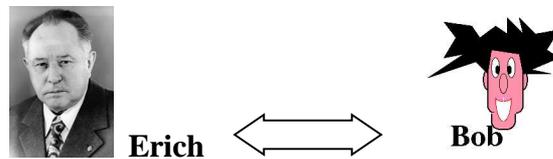
Hint: Do not look at [HPR01] before you have tried to do it yourself. ■

5 Interactive Proof Systems and Zero-Knowledge Protocols

In Section 3.1, we mentioned the “Man-in-the-middle” attack on the Diffie-Hellman secret-key agreement protocol. Imagine that Bob has just agreed with his partner on a joint secret key via a public telephone line. Of course, he assumes it was Alice he was talking to. Bob was so clever to use the Diffie-Hellman protocol, and so he thinks that Erich does not have a clue about what secret key they have chosen:



But Erich was even smarter. Here is what really happened:



This situation raises the issue of *authentication*: How can Bob be certain that it in fact was Alice he was communicating with, and not Erich pretending to be Alice? In other words, how can Alice prove her identity to Bob beyond any doubt?

In Section 3, we have seen how to use digital signatures for the authentication of documents such as email messages. In this section, our goal is to achieve authentication of an *individual* rather than a document. One way to achieve this goal is to assign to Alice’s identity some secret information such as her PIN (“*Personal Identification Number*”) or any other private information that nobody else knows. We refer to the information proving Alice’s identity as Alice’s *secret*.

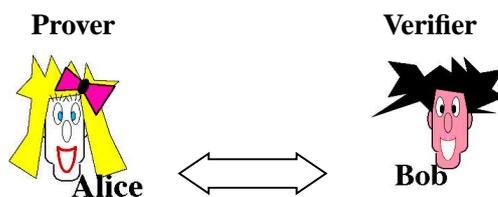
But here’s another catch. Alice would like to convince Bob of her identity by proving that she knows her secret. Ideally, however, she should not disclose her secret because then it wouldn’t be a secret anymore: If Bob, for example, knew Alice’s secret, he could pretend to be Alice when communicating with somebody else. So the question is:

How can one prove to know a secret without telling what the secret is?

That is precisely what zero-knowledge protocols are all about.

5.1 Interactive Proof Systems

Zero-knowledge protocols are a special form of interactive proof systems, which we will describe first. Interactive proof systems were introduced by Shafi Goldwasser, Silvio Micali, and Charles Rack-off [GMR85,GMR89]. As in the previous protocols, we consider the communication between two parties, the “prover” Alice and the “verifier” Bob:



For now, we are not interested in the security aspects that may arise when the communication is eaves-dropped; rather, we are concerned with the following communication problem: Alice and Bob want to jointly solve a given problem L , i.e., they want to decide whether or not any given instance belongs to L . For concreteness, consider the graph isomorphism problem.

Definition 5.1 *The vertex set of any graph G is denoted by $V(G)$, and the edge set of G is denoted by $E(G)$. Let G and H be undirected, simple graphs, i.e., graphs with no reflexive or multiple edges.*

An isomorphism between G and H is a bijective mapping π from $V(G)$ onto $V(H)$ such that, for all $i, j \in V(G)$,

$$\{i, j\} \in E(G) \iff \{\pi(i), \pi(j)\} \in E(H).$$

Graph-Isomorphism *denotes the set of all pairs of isomorphic graphs.*

The graph isomorphism problem is to determine whether or not any two given graphs are isomorphic. This problem belongs to NP, but is not known to be complete for NP, i.e., it is not known whether this problem belongs to the hardest NP problems.⁸ On the other hand, there is no efficient algorithm known for solving it, and it is widely considered to be a hard, intractable problem.

Returning to Alice and Bob’s communication problem, their task is to decide whether or not any given pair (G, H) of graphs is isomorphic. Alice, the prover, tries to *prove* them isomorphic by providing Bob with an isomorphism π between G and H . She intends to convince Bob *no matter whether or not G and H in fact are isomorphic*. But Bob is impatient. To accept the input, he wants to be convinced with overwhelming probability that the proof provided by Alice indeed is correct. Even worse, he is convinced only if *every potential prover strategy* Alice might come up with yields an overwhelming success probability. If Alice can accomplish this then Bob accepts the input, otherwise he rejects it.

To formalize this intuition, imagine Alice and Bob to be Turing machines. Alice, the prover, is an all-powerful Turing machine with no computational limitation whatsoever. Bob, the verifier, is a randomized Turing machine working in polynomial time, but capable of making random moves by flipping an unbiased coin. In Definition 5.2 below, in case of acceptance, it is enough that Alice finds one sufficient strategy to convince Bob. In case of rejection, however, rather than considering every potential prover strategy of Alice, it is useful to quantify over all possible provers that may replace Alice.

Definition 5.2 (Interactive Proof System) [GMR85,GMR89]

1. An interactive proof system (or “IP protocol”) (A, B) is a protocol between Alice, the prover, and Bob, the verifier. Alice runs a Turing machine A with no limit on its resources, while Bob runs a polynomial-time randomized Turing machine B .⁹ Both access the same input on a joint input tape, and they are equipped with private work tapes for internal computations. They also share a read-write communication tape to exchange messages. Alice does not see Bob’s random choices. Let $\Pr((A, B)(x) = 1)$ denote the probability (according to the random choices made in the communication) that Bob accepts the input x ; i.e., for a particular sequence of random bits, “ $(A, B)(x) = 1$ ” denotes the event that Bob is convinced by Alice’s proof for x and accepts.

2. An interactive proof system (A, B) accepts a set L if and only if for each x :

$$(5.13) \quad x \in L \implies (\exists A) [\Pr((A, B)(x) = 1) \geq \frac{3}{4}];$$

$$(5.14) \quad x \notin L \implies (\forall \hat{A}) [\Pr((\hat{A}, B)(x) = 1) \leq \frac{1}{4}],$$

where \hat{A} is any prover (i.e., any Turing machine of unlimited computational power) that may replace Alice.

3. IP denotes the class of all sets that can be accepted by an interactive proof system.

Note that the acceptance probabilities of at least $\frac{3}{4}$ if $x \in L$ (respectively, of at most $\frac{1}{4}$ if $x \notin L$) are chosen at will. By probability amplification techniques [Pap94,BDG95,BC93], one can use any constants

⁸Due to its “lowness” properties [Sch87,KST92] it is doubted that the graph isomorphism problem is NP-complete. It is even conjectured that the graph isomorphism problem might be neither in P nor NP-complete, and this is what makes this problem so interesting for complexity theoreticians. Of course, proving this conjecture would immediately prove P different from NP; so, such a proof seems beyond current techniques. For more complexity-theoretic background on the graph isomorphism problem, we refer to the book by Köbler, Schöning, and Torán [KST93].

⁹For the definition of randomized Turing machines, we refer to any text book on complexity theory such as [Pap94,BDG95,BC93]. Essentially, every nondeterministic Turing machine can be viewed as a randomized Turing machine by defining a suitable probability measure on the computation trees of the machine.

$\frac{1}{2} + \epsilon$ and $\frac{1}{2} - \epsilon$, respectively, where $\epsilon > 0$. It is even possible to make the error probability as small as $2^{-p(|x|)}$, for any fixed polynomial p . Better yet, Goldreich, Mansour, and Sipser [GMS87] have shown that one can even require the acceptance probability of exactly 1 if $x \in L$, without changing the class IP.

In the literature, verifier and prover are sometimes referred to as *Arthur* and *Merlin*. In fact, the Arthur-Merlin games introduced by Babai [Bab85] (see also [BM88]) are nothing else than the interactive proof systems of Goldwasser et al. [GMR85,GMR89]. One difference between Definition 5.2 and the definition of Arthur-Merlin games is that the random bits chosen by Arthur are public (i.e., they are known to Merlin), while they are private to Bob in Definition 5.2. However, Goldwasser and Sipser [GS89] have shown that the privacy of the verifier’s random bits does not matter: Arthur-Merlin games are equivalent to interactive proof systems.

What if Bob has run out of coins? That is, what if he behaves deterministically when verifying Alice’s proof for “ $x \in L$ ”? Due to her unlimited computational power, Alice can provide proofs of unlimited length, i.e., of length not bounded by any function in the length of x . However, since Bob is a polynomial-time Turing machine, it is clear that he can check only proofs of length polynomially in $|x|$. It follows that IP, when restricted to deterministic polynomial-time verifiers, is just a cumbersome way of defining the class NP. Hence, since **Graph-Isomorphism** belongs to NP, it must also belong to the (unrestricted) class IP. We omit presenting an explicit IP protocol for **Graph-Isomorphism** here, but we refer to Section 5.3, where in Figure 11 an IP protocol for **Graph-Isomorphism** with an additional property is given: it is a zero-knowledge protocol.

But what about the complement of **Graph-Isomorphism**? Does there exist an interactive proof system that decides whether or not two given graphs are *non-isomorphic*? Note that even though Alice is all-powerful computationally, she may run into difficulties when she is trying to prove that the graphs are non-isomorphic. Consider, for example, two non-isomorphic graphs with 1000 vertices each. A proof of that fact seems to require Alice to show that none of the 1000! possible permutations is an isomorphism between the graphs. Not only would it be impossible for Bob to check such a long proof in polynomial time, also for Alice it would be literally impossible to write this proof down. After all, 1000! is approximately $4 \cdot 10^{2567}$. This number exceeds the number of atoms in the entire visible universe,¹⁰ which is currently estimated to be around 10^{77} , by a truly astronomical factor.

That is why the following result of Goldreich, Micali, and Wigderson [GMW86,GMW91] was a bit of a surprise.

Theorem 5.3 [GMW86,GMW91] $\overline{\text{Graph-Isomorphism}}$ is in IP.

Proof. Figure 10 shows the interactive proof system for the graph non-isomorphism problem.

Let us check that the implications (5.13) and (5.14) from Definition 5.2 do hold. Suppose that G_1 and G_2 are non-isomorphic. Then, it is easy for Alice to determine that graph G_b , $b \in \{1, 2\}$, to which H is isomorphic. So she sends $a = b$, and Bob accepts with probability 1. That is,

$$(G_1, G_2) \in \overline{\text{Graph-Isomorphism}} \implies (\exists A) [\Pr((A, B)(G_1, G_2) = 1) = 1].$$

Now suppose that G_1 and G_2 are isomorphic. Then, no matter what clever strategy Alice applies, her chance of answering correctly (i.e., with $a = b$) is no better than $\frac{1}{2}$ because she does not see Bob’s random bit b and so can do no better than guessing. That is,

$$(G_1, G_2) \notin \overline{\text{Graph-Isomorphism}} \implies (\forall \hat{A}) [\Pr((\hat{A}, B)(G_1, G_2) = 1) \leq \frac{1}{2}].$$

Note that the acceptance probability of $\leq \frac{1}{2}$ above is not yet the acceptance probability of $\leq \frac{1}{4}$ required in (5.14) of Definition 5.2. However, as mentioned above, standard probability amplification techniques yield an error probability as close to zero as one desires. We leave the details to the reader. ■

¹⁰Dark matter excluded.

Step	 Alice		 Bob
	Input: Two graphs G_1 and G_2		
1			randomly chooses a permutation π on $V(G_1)$ and a bit $b \in \{1, 2\}$, and computes $H = \pi(G_b)$
2		\xleftarrow{H}	
3	determines $a \in \{1, 2\}$ such that G_a and H are isomorphic		
4		\xrightarrow{a}	
5			accepts if and only if $a = b$

Figure 10: The Goldreich-Micali-Wigderson IP protocol for Graph-Isomorphism.

The above result shows that IP contains not only all of NP but also a problem from coNP, the class of complements of NP problems, which is unlikely to be in NP. So, the question arises of how big the class IP actually is. A famous result of Adi Shamir [Sha92] settled this question: IP equals PSPACE, the class of problems that can be decided in polynomial space.

5.2 Zero-Knowledge Protocols

Recalling the issue of authentication mentioned at the beginning of this section, we are now ready to define zero-knowledge protocols.

As mentioned above, Graph-Isomorphism is in IP. To prove that the two given graphs are isomorphic, Alice simply sends an isomorphism π to Bob, which he then checks deterministically in polynomial time. Suppose, however, that Alice wants to keep the isomorphism π secret. On the one hand, she does not want to disclose her secret; on the other hand, she wants to prove to Bob that she knows it. What she needs is a very special IP protocol that conveys nothing about her secret isomorphism, and yet proves that the graphs are isomorphic. The next section will present such a zero-knowledge protocol for Graph-Isomorphism.

But what is a zero-knowledge protocol and how can one formalize it? The intuition is this. Imagine that Alice has a twin sister named Malice who looks just like her. However, Malice does not know Alice's secret. Moreover, Malice does not have Alice's unlimited computational power; rather, just as the verifier Bob, she only operates like a randomized polynomial-time Turing machine. Still, she tries to simulate Alice's communication with Bob. An IP protocol has the *zero-knowledge property* if the information communicated in Malice's simulated protocol cannot be distinguished from the information communicated in Alice's original protocol. Malice, not knowing the secret, cannot put any information about the secret into her simulated protocol, and yet she is able to generate that clone of the original protocol that looks just like the original to an independent observer. Consequently, the verifier Bob (or any other party such as Erich) cannot extract any information from the original protocol. In short, if there's nothing in there, you can't get anything out of it.

Definition 5.4 (Zero-Knowledge Protocols) [GMR85,GMR89] *Let (A, B) be an interactive proof system accepting a problem L . We say (A, B) is a zero-knowledge protocol for L if and only if there*

exists a simulator Malice such that the following holds:

- Malice runs a randomized polynomial-time Turing machine M to simulate the prover Alice in her communication with Bob, thus yielding a simulated protocol (M, B) ;
- for each $x \in L$, the tuples (a_1, a_2, \dots, a_k) and (m_1, m_2, \dots, m_k) representing the communication in (A, B) and in (M, B) , respectively, are identically distributed over the coin tosses of A and B in (A, B) and of M and B in (M, B) , respectively.

The above definition is called “perfect zero-knowledge” in the literature. There are several other, weaker notions of zero-knowledge such as “statistical zero-knowledge” and “computational zero-knowledge.” In the latter model, Goldreich, Micali, and Wigderson [GMW86,GMW91] showed what is considered by far the most important result on zero-knowledge: Every problem in NP has a computational zero-knowledge protocol (under a plausible cryptographic assumption). The key idea is a computational zero-knowledge protocol for the NP-complete graph three-colorability problem. In contrast, it seems unlikely [BC89] that such a strong claim can be proven for the perfect zero-knowledge model presented in Definition 5.4.

For more information about interactive proof systems and zero-knowledge, we refer to the books by Goldreich [Gol01b, Chapter 4], Köbler et al. [KST93, Chapter 2], Papadimitriou [Pap94, Chapter 12.2], Balcázar et al. [BDG90, Chapter 11], and Bovet et al. [BC93, Chapter 10] and to the surveys by Oded Goldreich [Gol88], Shafi Goldwasser [Gol89], and Joan Feigenbaum [Fei92].

5.3 Zero-Knowledge Protocol for the Graph Isomorphism Problem

Oded Goldreich, Silvio Micali, and Avi Wigderson [GMW86,GMW91] proposed a zero-knowledge protocol for the graph isomorphism problem. This result was quite a surprise, since previously zero-knowledge protocols were known only for problems contained both in NP and coNP. It is considered to be unlikely that NP equals coNP; in particular, it is considered to be unlikely that **Graph-Isomorphism** is in coNP.

Theorem 5.5 [GMW86,GMW91] *Graph-Isomorphism has a zero-knowledge protocol.*

Proof. Figure 11 shows the Goldreich-Micali-Wigderson protocol. One difference to the protocol for the graph non-isomorphism problem in Figure 10 is that now Alice too makes random choices.

Alice’s secret is the isomorphism π she has chosen. The protocol is correct, since Alice knows her secret π and also her random permutation ρ . Hence, she can easily compute the isomorphism σ with $\sigma(G_b) = H$ to prove her identity to Bob. When doing so, she does not have to disclose her secret π to Bob in order to convince him of her identity. In particular,

$$(G_1, G_2) \in \mathbf{Graph-Isomorphism} \implies (\exists A) [\Pr((A, B)(G_1, G_2) = 1) = 1],$$

so the implication (5.13) from Definition 5.2 holds. Since Alice herself has chosen two isomorphic graphs, the case $(G_1, G_2) \notin \mathbf{Graph-Isomorphism}$ does not occur, so the implication (5.14) from Definition 5.2 trivially holds if the protocol is implemented properly. Thus, the protocol is an interactive proof system for **Graph-Isomorphism**.

Recall that Alice wants to prove her identity via this protocol. Suppose that Erich or Malice want to cheat by pretending to be Alice. They do not know her secret isomorphism π , but they do know the public isomorphic graphs G_1 and G_2 . They want to convince Bob that they know Alice’s secret, which corresponds to (G_1, G_2) . If, by coincidence, Bob’s bit b equals their previously chosen bit a , they win. However, if $b \neq a$, computing $\sigma = \rho \circ \pi$ or $\sigma = \rho \circ \pi^{-1}$ requires knowledge of π . Without

knowing π , computing π from the public graphs G_1 and G_2 seems to be impossible for them, since Graph-Isomorphism is a hard problem, too hard even for randomized polynomial-time Turing machines. Thus, they will fail provided that the graphs are chosen large enough.

Step	 Alice		 Bob
Generation of isomorphic graphs and a secret isomorphism			
1	chooses a large graph G_1 , a random permutation π on G_1 's vertices, and computes the graph $G_2 = \pi(G_1)$; (G_1, G_2) are public, π is private		
Protocol			
2	randomly chooses a permutation ρ on $V(G_1)$ and a bit $a \in \{1, 2\}$, computes $H = \rho(G_a)$		
3		\xRightarrow{H}	
4			chooses a bit $b \in \{1, 2\}$ at random and wants to see an isomorphism between G_b and H
5		\xleftarrow{b}	
6	computes the permutation $\sigma = \begin{cases} \rho & \text{if } b = a \\ \rho \circ \pi & \text{if } 1 = b \neq a = 2 \\ \rho \circ \pi^{-1} & \text{if } 2 = b \neq a = 1 \end{cases}$ satisfying $\sigma(G_b) = H$		
7		$\xRightarrow{\sigma}$	
8			verifies that indeed $\sigma(G_b) = H$ and accepts accordingly

Figure 11: The Goldreich-Micali-Wigderson zero-knowledge protocol for graph isomorphism.

Since they cannot do better than guessing the bit b , they can cheat with probability at most $\frac{1}{2}$. Of course, they can always guess the bit b , which implies that their chance of cheating successfully is exactly $\frac{1}{2}$. Hence, if Bob demands, say, k independent rounds of the protocol to be executed, he can make the cheating probability as small as 2^{-k} , and thus is very likely to detect any cheater. Note that after only 20 rounds the odds of malicious Malice to get away with it undetected are less than one to one million. Hence, the protocol is correct.

It remains to show that the protocol in Figure 11 is zero-knowledge. Figure 12 shows a simulated protocol with Malice, who does not know the secret π , replacing Alice. The information communicated in one round of the protocol is given by a triple of the form (H, b, σ) . Whenever Malice chooses a bit

Step	 Malice		 Bob
Simulated generation of isomorphic graphs			
1	knows the public pair (G_1, G_2) of isomorphic graphs, does not know Alice's secret π		
Simulated Protocol			
2	randomly chooses a permutation ρ on $V(G_1)$ and a bit $a \in \{1, 2\}$, computes $H = \rho(G_a)$		
3		\xRightarrow{H}	
4			chooses a bit $b \in \{1, 2\}$ at random and wants to see an isomorphism between G_b and H
5		\xleftarrow{b}	
6	if $b \neq a$ then M deletes all messages transmitted in this round and repeats; if $b = a$ then M sends $\sigma = \rho$		
7		$\xRightarrow{\sigma}$	
8			$b = a$ implies that indeed $\sigma(G_b) = H,$ so Bob accepts "Alice's" identity

Figure 12: How to simulate the Goldreich-Micali-Wigderson protocol without knowing the secret π .

a with $a = b$, she simply sends $\sigma = \rho$ and wins: Bob, or any independent observer, will not detect that she in fact is Malice. Otherwise, whenever $a \neq b$, Malice fails. However, that's no problem at all: She simply deletes this round from the simulated protocol and repeats. Thus, she can produce a sequence of triples of the form (H, b, σ) that is indistinguishable from the corresponding sequence of triples in the original protocol between Alice and Bob. It follows that the Goldreich-Micali-Wigderson protocol is zero-knowledge. ■

5.4 Fiat and Shamir's Zero-Knowledge Protocol

Amos Fiat and Adi Shamir [FS86] proposed a zero-knowledge protocol for a number-theoretical problem. It is based on the assumption that computing square roots in \mathbb{Z}_n^* is practically infeasible. Due to its properties, the Fiat-Shamir protocol is particularly suitable for authentication of individuals in large computer networks. It is a public-key protocol, it is more efficient than other public-key protocols such as the RSA algorithm, it can be implemented on a chip card, and it is zero-knowledge. These advantages resulted in a rapid deployment of the protocol in practical applications. The Fiat-Shamir protocol

is integrated in the “Videocrypt” Pay-TV system [Eur91]. The theory of zero-knowledge may also become important in future internet technologies. To prevent confusion, we note that Zero-Knowledge Systems, Inc., a Montréal-based company that was founded in 1997 and provides products and services enabling users to protect their privacy on-line on the world wide web, is not a commercial fielding of zero-knowledge protocols [Gol01a].

Step	 Alice		 Bob
Key generation			
1	chooses two large primes p and q and a secret $s \in \mathbb{Z}_n^*$, $n = pq$, and computes $v = s^2 \pmod n$; p , q , and s are kept secret, whereas n and v are public		
Protocol			
2	chooses $r \in \mathbb{Z}_n^*$ at random and computes $x = r^2 \pmod n$		
3		\xRightarrow{x}	
4			chooses a bit $b \in \{0, 1\}$ at random
5		\xleftarrow{b}	
6	computes $y = r \cdot s^b \pmod n$		
7		\xRightarrow{y}	
8			verifies that indeed $y^2 \equiv x \cdot v^b \pmod n$ and accepts accordingly

Figure 13: The Fiat-Shamir zero-knowledge protocol.

Theorem 5.6 [FS86] *The Fiat-Shamir procedure given in Figure 13 is a zero-knowledge protocol.*

Proof. Look at Figure 13. The protocol is correct, since Alice knows the secret $s \in \mathbb{Z}_n^*$ that she has chosen, and thus she can compute $y = r \cdot s^b$, where b is the bit that Bob has chosen at random. Hence, it holds in \mathbb{Z}_n^* that

$$y^2 \equiv (r \cdot s^b)^2 \equiv r^2 \cdot s^{2b} \equiv r^2 \cdot v^b \equiv x \cdot v^b \pmod n,$$

so Bob accepts Alice’s identity.

Suppose now that Erich or Malice want to cheat by pretending to be Alice. They do not know her secret s , nor do they know the primes p and q , but they do know the public $n = pq$ and $v = s^2 \pmod n$. They want to convince Bob that they know Alice’s secret s , the square root of v modulo n . If, by coincidence, Bob’s bit b equals zero then $y = r \cdot s^0 = r$ and they win. However, if $b = 1$, computing a y that satisfies $y^2 \equiv x \cdot v^b \pmod n$ requires knowledge of the secret s , assuming that computing square roots modulo n is hard. Without knowing s , if Malice or Erich were able to compute the correct answer

Step	 Malice		 Bob
Simulated key generation			
1	knows the public $n = pq$ and $v = s^2 \pmod n$; does not know the private primes p and q and Alice's secret s		
Simulated Protocol			
2	randomly chooses $r \in \mathbb{Z}_n^*$ and a bit $c \in \{0, 1\}$, computes $x = r^2 \cdot v^{-c} \pmod n$		
3		\xRightarrow{x}	
4			chooses a bit $b \in \{0, 1\}$ at random
5		\xleftarrow{b}	
6	if $b \neq c$ then M deletes all messages transmitted in this round and repeats; if $b = c$ then M sends $y = r$		
7		\xRightarrow{y}	
8			$b = c$ implies that indeed $y^2 = r^2 = r^2 v^{-c} v^b \equiv x \cdot v^b \pmod n,$ so Bob accepts "Alice's" identity

Figure 14: How to simulate the Fiat-Shamir protocol without knowing the secret s .

for both $b = 0$ and $b = 1$, say y_b with $y_b^2 \equiv x \cdot v^b \pmod n$, they could efficiently compute square roots modulo n as follows: $y_0^2 \equiv x \pmod n$ and $y_1^2 \equiv x \cdot v \pmod n$ implies $(\frac{y_1}{y_0})^2 \equiv v \pmod n$; hence, $\frac{y_1}{y_0}$ is a square root of v modulo n .

It follows that they can cheat with probability at most $\frac{1}{2}$. Of course, they can always guess the bit b in advance and prepare the answer accordingly. Choosing $x = r^2 \cdot v^{-b} \pmod n$ and $y = r$ implies that

$$(5.15) \quad y^2 \equiv r^2 \equiv r^2 \cdot v^{-b} \cdot v^b \equiv x \cdot v^b \pmod n.$$

Thus, Bob will not detect any irregularities and will accept. Hence, their chance to cheat successfully is exactly $\frac{1}{2}$. Again, if Bob demands, say, k independent rounds of the protocol to be executed, he can make the cheating probability as small as desired and is very likely to detect any cheater.

It remains to show that the Fiat-Shamir protocol in Figure 13 is zero-knowledge. Figure 14 shows a simulated protocol with Malice, who does not know the secret s , replacing Alice. The information communicated in one round of the protocol is given by a triple of the form (x, b, y) . In addition to the randomly chosen $r \in \mathbb{Z}_n^*$, Malice guesses a bit $c \in \{0, 1\}$ and computes $x = r^2 \cdot v^{-c} \pmod n$, which she

sends to Bob. Whenever c happens to be equal to Bob's bit b , Malice simply sends $y = r$ and wins. By an argument analogous to Equation (5.15) above, neither Bob nor any independent observer will detect that she actually is Malice:

$$y^2 \equiv r^2 \equiv r^2 \cdot v^{-c} \cdot v^b \equiv x \cdot v^b \pmod{n}.$$

Otherwise, whenever $c \neq b$, Malice fails. However, that's no problem at all: She simply deletes this round from the simulated protocol and repeats. Thus, she can produce a sequence of triples of the form (x, b, y) that is indistinguishable from the corresponding sequence of triples in the original protocol between Alice and Bob. It follows that the Fiat-Shamir protocol is zero-knowledge. ■

Acknowledgments. I am grateful to Peter Widmayer for reading a preliminary draft of this paper, for his insightful comments, and for interesting discussions that helped much to improve the presentation of this paper. I thank Pekka Orponen for inviting me to be a lecturer of the 11th Jyvaskylä Summer School that was held in August, 2001, at the University of Jyvaskylä. I thank Kari Pasanen for being a great tutor of this tutorial, for carefully proofreading a preliminary draft of this paper, and in particular for subletting his summer house on an island of scenic Lake Keitele to me and my family during the summer school. I am grateful to Pekka and Kari for their hospitality, and I thank my 33 summer school students from 16 countries for making this course so much fun and pleasure. Godmar Back's helpful comments on the English style of this paper are gratefully acknowledged. I also thank Eric Allender, Harald Baier, Lane Hemaspaandra, Eike Kiltz, Alan Selman, and Gerd Wechsung for their insightful advice and helpful comments and for their interest in this paper.

References

- [AD97] M. Ajtai and C. Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *Proceedings of the 29th ACM Symposium on Theory of Computing*, pages 284–293. ACM Press, 1997.
- [Ajt96] M. Ajtai. Generating hard instances of lattice problems. In *Proceedings of the 28th ACM Symposium on Theory of Computing*, pages 99–108. ACM Press, 1996.
- [Ajt98] M. Ajtai. The shortest vector problem in L_2 is NP-hard for randomized reductions. In *Proceedings of the 30th ACM Symposium on Theory of Computing*, pages 10–19. ACM Press, 1998.
- [All85] E. Allender. Invertible functions, 1985. PhD thesis, Georgia Institute of Technology.
- [All86] E. Allender. The complexity of sparse sets in P. In *Proceedings of the 1st Structure in Complexity Theory Conference*, pages 1–11. Springer-Verlag *Lecture Notes in Computer Science* #223, June 1986.
- [Bab85] L. Babai. Trading group theory for randomness. In *Proceedings of the 17th ACM Symposium on Theory of Computing*, pages 421–429, April 1985.
- [Bau00] F. Bauer. *Entzifferte Geheimnisse*. Springer-Verlag, third edition, 2000. In German.
- [BC89] Gilles Brassard and Claude Crepeau. Sorting out zero-knowledge. In *Advances in Cryptology—EUROCRYPT 89*, pages 181–191. Springer-Verlag *Lecture Notes in Computer Science* #434, April 1989.

- [BC93] D. Bovet and P. Crescenzi. *Introduction to the Theory of Complexity*. Prentice Hall, 1993.
- [BDG90] J. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity II*. EATCS Monographs on Theoretical Computer Science. Springer-Verlag, 1990.
- [BDG95] J. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity I*. EATCS Monographs on Theoretical Computer Science. Springer-Verlag, second edition, 1995.
- [Ber77] L. Berman. *Polynomial Reducibilities and Complete Sets*. PhD thesis, Cornell University, Ithaca, NY, 1977.
- [Beu01] A. Beutelspacher. *Moderne Verfahren der Kryptographie*. Vieweg, 4th edition, 2001. In German.
- [BHHR99] A. Beygelzimer, L. Hemaspaandra, C. Homan, and J. Rothe. One-way functions in worst-case cryptography: Algebraic and security properties are on the house. *SIGACT News*, 30(4):25–40, December 1999.
- [BM88] L. Babai and S. Moran. Arthur-Merlin games: A randomized proof system, and a hierarchy of complexity classes. *Journal of Computer and System Sciences*, 36(2):254–276, 1988.
- [Buc01] J. Buchmann. *Introduction to Cryptography*. Undergraduate Texts in Mathematics. Springer-Verlag, 2001.
- [Cai99] J. Cai. Some recent progress on the complexity of lattice problems. In *Proceedings of the 14th Annual IEEE Conference on Computational Complexity*, pages 158–179. IEEE Computer Society Press, May 1999.
- [DH76] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, IT-22(6):644–654, 1976.
- [ElG85] T. ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, IT-31(4):469–472, 1985.
- [Eur91] European Patent Application 0 428252 A2. A system for controlling access to broadcast transmissions, 1991.
- [Fei92] J. Feigenbaum. Overview of interactive proof systems and zero-knowledge. In G. Simmons, editor, *Contemporary Cryptology: The Science of Information Integrity*, pages 423–439. IEEE Computer Society Press, 1992.
- [FS86] A. Fiat and A. Shamir. How to prove yourself: Practical solutions to identification and signature problems. In *Advances in Cryptology—CRYPTO '86*, pages 186–194. Springer-Verlag *Lecture Notes in Computer Science #263*, 1986.
- [Gil77] J. Gill. Computational complexity of probabilistic Turing machines. *SIAM Journal on Computing*, 6(4):675–695, 1977.
- [GMR85] S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proof systems. In *Proceedings of the 17th ACM Symposium on Theory of Computing*, pages 291–304, April 1985.
- [GMR89] S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal on Computing*, 18(1):186–208, February 1989.

- [GMS87] O. Goldreich, Y. Mansour, and M. Sipser. Interactive proof systems: Provers that never fail and random selection. In *Proceedings of the 28th IEEE Symposium on Foundations of Computer Science*, pages 449–461. IEEE Computer Society Press, October 1987.
- [GMW86] O. Goldreich, S. Micali, and A. Wigderson. Proofs that yield nothing but their validity and a methodology of cryptographic protocol design. In *Proceedings of the 27th IEEE Symposium on Foundations of Computer Science*, pages 174–187, April 1986.
- [GMW91] O. Goldreich, S. Micali, and A. Wigderson. Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. *Journal of the ACM*, 38(3):691–729, July 1991.
- [Gol88] O. Goldreich. Randomness, interactive proofs, and zero-knowledge—A survey. In R. Herken, editor, *The Universal Turing Machine: A Half-Century Survey*, pages 377–405. Oxford University Press, Oxford, 1988.
- [Gol89] S. Goldwasser. Interactive proof systems. In J. Hartmanis, editor, *Computational Complexity Theory*, pages 108–128. AMS Short Course Lecture Notes: Introductory Survey Lectures, Proceedings of Symposia in Applied Mathematics, Volume 38, American Mathematical Society, 1989.
- [Gol97] O. Goldreich. A taxonomy of proof systems. In L. Hemaspaandra and A. Selman, editors, *Complexity Theory Retrospective II*, pages 109–134. Springer-Verlag, 1997.
- [Gol99] O. Goldreich. *Modern cryptography, probabilistic proofs, and pseudorandomness*, volume 17 of *Algorithms and Combinatorics*. Springer-Verlag, 1999.
- [Gol01a] I. Goldberg, November 13, 2001. Personal Communication.
- [Gol01b] O. Goldreich. *Foundations of Cryptography*. Cambridge University Press, 2001.
- [GS88] J. Grollmann and A. Selman. Complexity measures for public-key cryptosystems. *SIAM Journal on Computing*, 17(2):309–335, 1988.
- [GS89] S. Goldwasser and M. Sipser. Private coins versus public coins in interactive proof systems. In S. Micali, editor, *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 73–90. JAI Press, Greenwich, 1989. A preliminary version appeared in *Proc. 18th Ann. ACM Symp. on Theory of Computing*, 1986, pp. 59–68.
- [Hås88] J. Håstad. Solving simultaneous modular equations of low degree. *SIAM Journal on Computing*, 17(2):336–341, April 1988. Special issue on cryptography.
- [HPR01] L. Hemaspaandra, K. Pasanen, and J. Rothe. If $P \neq NP$ then some strongly noninvertible functions are invertible. In *Proceedings of the 13th International Symposium on Fundamentals of Computation Theory*, pages 162–171. Springer-Verlag *Lecture Notes in Computer Science #2138*, August 2001.
- [HR99] L. Hemaspaandra and J. Rothe. Creating strong, total, commutative, associative one-way functions from any one-way function in complexity theory. *Journal of Computer and System Sciences*, 58(3):648–659, 1999.
- [HR00] L. Hemaspaandra and J. Rothe. Characterizing the existence of one-way permutations. *Theoretical Computer Science*, 244(1–2):257–261, 2000.

- [HRW97] L. Hemaspaandra, J. Rothe, and G. Wechsung. On sets with easy certificates and the existence of one-way permutations. In *Proceedings of the Third Italian Conference on Algorithms and Complexity*, pages 264–275. Springer-Verlag *Lecture Notes in Computer Science* #1203, March 1997.
- [Kle52] S. Kleene. *Introduction to Metamathematics*. D. van Nostrand Company, Inc., New York and Toronto, 1952.
- [Knu81] D. Knuth. *The Art of Computer Programming: Seminumerical Algorithms*, volume 2 of *Computer Science and Information*. Addison-Wesley, second edition, 1981.
- [Ko85] K. Ko. On some natural complete operators. *Theoretical Computer Science*, 37(1):1–30, 1985.
- [KR95] B. Kaliski Jr. and M. Robshaw. The secure use of RSA. *CryptoBytes*, 1(3):7–13, 1995.
- [KS01] R. Kumar and D. Sivakumar. Complexity of SVP—a reader’s digest. *SIGACT News*, 32(3):40–52, June 2001.
- [KST92] J. Köbler, U. Schöning, and J. Torán. Graph isomorphism is low for PP. *Computational Complexity*, 2:301–330, 1992.
- [KST93] J. Köbler, U. Schöning, and J. Torán. *The Graph Isomorphism Problem: Its Structural Complexity*. Birkhäuser, 1993.
- [Len87] H. Lenstra Jr. Factoring integers with elliptic curves. *Annals of Mathematics*, 126:649–673, 1987.
- [Mil76] G. Miller. Riemann’s hypothesis and tests for primality. *Journal of Computer and System Sciences*, 13:300–317, 1976.
- [Moo92] J. Moore. Protocol failures in cryptosystems. In G. Simmons, editor, *Contemporary Cryptology: The Science of Information Integrity*, pages 541–558. IEEE Computer Society Press, 1992.
- [MW99] U. Maurer and S. Wolf. The relationship between breaking the Diffie-Hellman protocol and computing discrete logarithms. *SIAM Journal on Computing*, 28(5):1689–1721, 1999.
- [Nat91] National Institute of Standards and Technology (NIST). Digital signature standard (DSS). *Federal Register*, 56(169), August 1991.
- [Nat92] National Institute of Standards and Technology (NIST). The Digital Signature Standard, proposed by NIST. *Communications of the Association for Computing Machinery*, 35(7):36–40, July 1992.
- [Pap94] C. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.
- [Pol74] J. Pollard. Theorems on factorization and primality testing. *Proc. Cambridge Philos. Soc.*, 76:521–528, 1974.
- [Rab80] M. Rabin. Probabilistic algorithms for testing primality. *Journal of Number Theory*, 12:128–138, 1980.
- [RH] J. Rothe and L. Hemaspaandra. On characterizing the existence of partial one-way permutations. *Information Processing Letters*. To appear.

- [RS93] M. Rabi and A. Sherman. Associative one-way functions: A new paradigm for secret-key agreement and digital signatures. Technical Report CS-TR-3183/UMIACS-TR-93-124, Department of Computer Science, University of Maryland, College Park, Maryland, 1993.
- [RS97] M. Rabi and A. Sherman. An observation on associative one-way functions in complexity theory. *Information Processing Letters*, 64(5):239–244, 1997.
- [RSA78] R. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signature and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.
- [Sal96] A. Salomaa. *Public-Key Cryptography*, volume 23 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, second edition, 1996.
- [Sch87] U. Schöning. Graph isomorphism is in the low hierarchy. *Journal of Computer and System Sciences*, 37:312–323, 1987.
- [Sch90] C. Schnorr. Efficient identification and signature schemes for smart cards. In *CRYPTO '89*, pages 239–251. Springer-Verlag *Lecture Notes in Computer Science #435*, February 1990.
- [Sch96] B. Schneier. *Applied Cryptography: Protocols, Algorithms, and Source Code in C*. John Wiley and Sons, New York, second edition, 1996.
- [Sel92] A. Selman. A survey of one-way functions in complexity theory. *Mathematical Systems Theory*, 25(3):203–221, 1992.
- [Sha49] C. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28(4):657–715, 1949.
- [Sha92] A. Shamir. $IP=PSPACE$. *Journal of the ACM*, 39(4):869–877, 1992.
- [Sha95] A. Shamir. RSA for paranoids. *CryptoBytes*, 1(3):1–4, 1995.
- [Sim79] G. Simmons. Symmetric and asymmetric encryption. *ACM Computing Surveys*, 11(4):305–330, 1979.
- [Sin99] S. Singh. *The Code Book. The Science of Secrecy from Ancient Egypt to Quantum Cryptography*. Fourth Estate, London, 1999.
- [SN77] G. Simmons and M. Norris. Preliminary comments on the MIT public-key cryptosystem. *Cryptologia*, 1(4):406–414, 1977.
- [SS77] R. Solovay and V. Strassen. A fast Monte Carlo test for primality. *SIAM Journal on Computing*, 6:84–85, 1977. Erratum appears in the same journal, 7(1):118, 1978.
- [Sti95] D. Stinson. *Cryptography Theory and Practice*. CRC Press, Boca Raton, 1995.
- [Val76] L. Valiant. The relative complexity of checking and evaluating. *Information Processing Letters*, 5(1):20–23, 1976.
- [Wel98] D. Welsh. *Codes and Cryptography*. Oxford science publications. Clarendon Press, Oxford, 6th edition, 1998. Reprinted with corrections.
- [Wie90] M. Wiener. Cryptanalysis of short RSA secret exponents. *IEEE Transactions on Information Theory*, IT-36(3):553–558, 1990.