

CrypTool Script*

Mathematics and Cryptography

(c) The authors, 1998-2003
Frankfurt am Main

July 17, 2003

In this CrypTool script you will find predominantly mathematically oriented information on using cryptographic procedures. The main chapters have been written by various **authors** (see appendix **B**) and are therefore independent from one another. At the end of most chapters you will find literature and web links.

You will obtain information about the principles of symmetrical and asymmetrical **encryption**. A large section of this script is dedicated to the fascinating topic of **prime numbers**. Using numerous examples, the **elementary number theory** and **modular arithmetic** are introduced and applied in an exemplary manner for the **RSA procedure**. By reading the following chapter you'll gain an insight into the mathematical ideas behind **modern cryptography**.

A further chapter is devoted to **digital signatures**, which are an essential component of e-business applications. The last chapter describes **elliptic curves**: they could be used as an alternative to RSA and in addition are extremely well suited for implementation on smartcards.

Whereas the *program* CrypTool teaches you how to use cryptography in practice, the *script* provides those interested in the subject with a deeper understanding of the mathematical algorithms used – trying to do it in an instructive way.

The authors Bernhard Esslinger, Matthias Büger, Bartol Filipovic, Henrik Koy, Roger Oyono and Jörg Cornelius Schneider would like to take this opportunity to thank their colleagues in the company and at the universities of Frankfurt, Gießen, Siegen, Karlsruhe and Darmstadt. They are particularly indebted to Dr. Peer Wichmann from the Karlsruhe computer science research centre (Forschungszentrum Informatik, FZI) for his down-to-earth support.

As at CrypTool, the quality of the script is enhanced by your suggestions and ideas for improvement. We look forward to your feedback.

You will find the current version of CrypTool under
<http://www.cryptool.org>, <http://www.cryptool.com> or <http://www.cryptool.de>.

The contact people for this free open-source tool are listed in the “readme” file delivered within the CrypTool package.

*Background reading for the CrypTool program

Contents

Contents	2
Preface to the 6th Edition of the CrypTool Script	7
Introduction – How do the Script and the Program Play together?	8
1 Encryption Procedures	9
1.1 Symmetric encryption	9
1.1.1 New results about cryptanalysis of AES	10
1.1.2 Current status of brute-force attacks on symmetric algorithms (RC5)	12
1.2 Asymmetric encryption	12
1.3 Hybrid procedures	13
1.4 Further details	14
Bibliography	15
Web links	16
2 Prime Numbers	17
2.1 What are prime numbers?	17
2.2 Prime numbers in mathematics	18
2.3 How many prime numbers are there?	19
2.4 The search for extremely large primes	20
2.4.1 Special number types – Mersenne numbers	21
2.4.2 EFF	24
2.5 Prime number tests	24
2.6 Overview special number types and the search for a formula for primes	26
2.6.1 Mersenne numbers $f(n) = 2^n - 1$ for n prime	26
2.6.2 Generalized Mersenne numbers $f(k, n) = k \cdot 2^n \pm 1$	27
2.6.3 Generalized Mersenne numbers $f(b, n) = b^n \pm 1$ / Cunningham project	27
2.6.4 Fermat numbers $f(n) = 2^{2^n} + 1$	27
2.6.5 Generalized Fermat numbers $f(b, n) = b^{2^n} + 1$	28
2.6.6 Carmichael numbers	29
2.6.7 Pseudo prime numbers	29
2.6.8 Strong pseudo prime numbers	29
2.6.9 Idea based on Euclid's proof $p_1 \cdot p_2 \cdots p_n + 1$	29

2.6.10	As above but -1 except $+1$: $p_1 \cdot p_2 \cdots p_n - 1$	29
2.6.11	Euclidean numbers $e_n = e_0 \cdot e_1 \cdots e_{n-1} + 1$	29
2.6.12	$f(n) = n^2 + n + 41$	30
2.6.13	$f(n) = n^2 - 79 \cdot n + 1,601$	31
2.6.14	Polynomial functions $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x^1 + a_0$	31
2.6.15	Catalan's conjecture	32
2.7	Density and distribution of the primes	32
2.8	Notes	35
2.8.1	Number of prime numbers in various intervals	37
2.8.2	Indexing prime numbers (n -th prime number)	38
2.8.3	Orders of magnitude / dimensions in reality	39
2.8.4	Special values in the binary and decimal systems	39
	Bibliography	40
	Web links	42
	Acknowledgments	43
3	Introduction to Elementary Number Theory with Examples	44
3.1	Mathematics and cryptography	44
3.2	Introduction to number theory	45
3.2.1	Convention	47
3.3	Prime numbers and the first fundamental theorem of elementary number theory	48
3.4	Divisibility, modulus and remainder classes	50
3.4.1	The modulo operation – working with congruence	50
3.5	Calculations with finite sets	52
3.5.1	Laws of modular calculations	52
3.5.2	Patterns and structures	53
3.6	Examples of modular calculations	54
3.6.1	Addition and multiplication	54
3.6.2	Additive and multiplicative inverses	55
3.6.3	Raising to the power	58
3.6.4	Fast calculation of high powers	59
3.6.5	Roots and logarithms	60
3.7	Groups and modular arithmetic in \mathbb{Z}_n and \mathbb{Z}_n^*	61
3.7.1	Addition in a group	62

3.7.2	Multiplication in a group	62
3.8	Euler function, Fermat's little theorem and Euler-Fermat	64
3.8.1	Patterns and structures	64
3.8.2	The Euler function	64
3.8.3	The theorem of Euler-Fermat	65
3.8.4	Calculation of the multiplicative inverse	66
3.8.5	Fixpoints modulo 26	66
3.9	Multiplicative order and primitive roots	68
3.10	Proof of the RSA procedure with Euler-Fermat	71
3.10.1	Basic idea of public key cryptography	71
3.10.2	How the RSA procedure works	72
3.10.3	Proof of requirement 1 (invertibility)	73
3.11	Considerations regarding the security of the RSA algorithm	75
3.11.1	Complexity	75
3.11.2	Security parameters because of progress in factorisation	76
3.11.3	Further current research in factorisation and number theory	77
3.11.4	Status regarding factorisation of concrete large numbers	80
3.12	Further applications of number theory in cryptography	83
3.12.1	One way functions	83
3.12.2	The Diffie-Hellman key exchange protocol	84
3.13	The RSA procedure with actual numbers	86
3.13.1	RSA with small prime numbers and with a number as message	87
3.13.2	RSA with slightly larger primes and an upper-case message	88
3.13.3	RSA with even larger primes and a text made up of ASCII characters	89
3.13.4	A small RSA cipher challenge (1)	91
3.13.5	A small RSA cipher challenge (2)	93
	Bibliography	95
	Web links	98
	Acknowledgments	99
	Appendix A: gcd of whole numbers and the two Algorithms of Euclid	100
	Appendix B: Forming closed sets	102
	Appendix C: Comments on modulo subtraction	102
	Appendix D: Base representation and base transformation of numbers, estimation of length of digits	103

Appendix E: Examples using Mathematica and Pari-GP	105
Appendix F: List of the formulated definitions and theorems	109
4 The Mathematical Ideas behind Modern Cryptography	110
4.1 One way functions with trapdoor and complexity classes	110
4.2 Knapsack problem as a basis for public key procedures	112
4.2.1 Knapsack problem	112
4.2.2 Merkle-Hellman knapsack encryption	113
4.3 Decomposition into prime factors as a basis for public key procedures	115
4.3.1 The RSA procedure	115
4.3.2 Rabin public key procedure (1979)	117
4.4 The discrete logarithm as a basis for public key procedures	117
4.4.1 The discrete logarithm in \mathbb{Z}_p	117
4.4.2 Diffie-Hellman key agreement	118
4.4.3 ElGamal public key encryption procedure in \mathbb{Z}_p^*	119
4.4.4 Generalised ElGamal public key encryption procedure	119
Bibliography	122
Web links	122
5 Hash Functions and Digital Signatures	123
5.1 Hash functions	124
5.2 RSA signatures	125
5.3 DSA signatures	125
5.4 Public key certification	126
5.4.1 Impersonation attacks	126
5.4.2 X.509 certificate	127
Bibliography	128
6 Elliptic Curves	129
6.1 Elliptic curve cryptography – a high-performance substitute for RSA?	129
6.2 Elliptic curves – history	131
6.3 Elliptic curves – mathematical basics	132
6.3.1 Groups	132
6.3.2 Fields	133
6.4 Elliptic curves in cryptography	135

6.5	Operating on the elliptic curve	137
6.6	Security of elliptic-curve-cryptography: The ECDLP	140
6.7	Encryption and signing with elliptic curves	141
6.7.1	Encryption	141
6.7.2	Signing	142
6.7.3	Signature verification	142
6.8	Factorisation using elliptic curves	143
6.9	Implementing elliptic curves	143
6.10	Elliptic curves in use	144
	Bibliography	145
	Web links	145
Appendix		147
A	CrypTool Menus	147
B	Authors of this script	150
Index		151

Preface to the 6th Edition of the CrypTool Script

Starting in the year 2000 this script became a part of the CrypTool package. It is designed to accompany the program CrypTool by explaining some mathematical topics in more detail, but still in a way which is easy to understand.

In order to also enable developers/authors to work together independently the topics have been split up and for each topic an extra chapter has been written which can be read on its own. The later editorial work in TeX added cross linkages between different sections and footnotes describing where you can find the according functions within the CrypTool program (see menu tree in appendix A). Naturally there are many more interesting topics in mathematics and cryptography which could be discussed in greater depth – therefore this is only one of many ways to do it.

The rapid spread of the Internet has also lead to intensified research in the technologies involved, especially within the area of cryptography where a good deal of new knowledge has arisen.

This edition of the script mainly updates the summaries of the following topical research areas:

- the search for the largest prime numbers (generalized Mersenne and Fermat primes) (chap. 2.6, 2.4.1),
- progress in number theory (“Primes in P”) (chap. 3.11.3),
- the factorisation of big numbers (TWIRL, RSA-160) (chap. 3.11.3, 3.11.4),
- progress in cryptanalysis of the AES standard (chap. 1.1.1) and
- progress in brute-force attacks against symmetric algorithms (chap. 1.1.2).

Chapter 5 (more about hash algorithms) and chapter 6 (now we explain in greater detail how elliptic curves are defined over different number fields) of this script have also been enhanced.

The first time the document was delivered with CrypTool was in version 1.2.01. Since then it has been expanded and revised in every new version of CrypTool (1.2.02, 1.3.00, 1.3.02, 1.3.03 and now 1.3.04).

I’d be more than happy if this also continues in the further open-source versions of CrypTool (they will be delivered by the University of Darmstadt).

I am deeply grateful to all the people helping with their impressive commitment who have made this global project so successful. Especially I would like to acknowledge the English language proof-reading of this script version done by Richard Christensen and Lowell Montgomery.

I hope that many readers have fun with this script and that they get out of it more interest and greater understanding of this modern but also very ancient topic.

Bernhard Esslinger

Hofheim, May 2003

Introduction – How do the Script and the Program Play together?

This script

This document is delivered together with the program CrypTool.

The articles in this script are largely self-contained and can also be read independently of CrypTool.

Chapters 4 and 6 require a deeper knowledge in mathematics, while the other chapters should be understandable with a school leaving certificate.

The authors have attempted to describe cryptography for a broad audience – without being mathematically incorrect. We believe that this didactical pretension is the best way to promote the awareness for IT security and the readiness to use standardised modern cryptography.

The program CrypTool

CrypTool is a program with an extremely comprehensive online help enabling you to use and analyse cryptographic procedures within a unified graphical user interface.

CrypTool was developed during the end-user awareness program at Deutsche Bank in order to increase employee awareness of IT security and provide them with a deeper understanding of the term security. A further aim has been to enable users to understand the cryptographic procedures. In this way, using CrypTool as a reliable reference implementation of the various encryption procedures (because of using the industry-proven Secude Library), you can test the encryption implemented in other programs.

CrypTool is currently been used for training in companies and teaching at school and universities, and moreover several universities are helping to further develop the project.

Acknowledgment

At this point I'd like to thank explicitly 3 people who particularly contributed to CrypTool. Without their talents and engagement CrypTool would not be what it is today:

- Mr. Henrik Koy
- Mr. Jörg Cornelius Schneider and
- Dr. Peer Wichmann.

Also I want to thank all the many people not mentioned here for their hard work (mostly carried out in their spare time).

Bernhard Esslinger

Hofheim, May 2003

1 Encryption Procedures

(Bernhard Esslinger, May 1999, Updates Dec. 2001, Feb. 2003)

This chapter introduces the topic in a more descriptive way without using too much mathematics.

The purpose of encryption is to change data in such a way that only an authorised recipient is able to reconstruct the plaintext. This allows us to transmit data without worrying about it getting into unauthorised hands. Authorised recipients possess a piece of secret information – called the key – which allows them to decrypt the data while it remains hidden from everyone else.

One encryption procedure has been mathematically proved to be secure, the *One Time Pad*. However, this procedure has several practical disadvantages (the key used must be randomly selected and must be at least as long as the message being protected), which means that it is hardly used except in closed environments such as for the hot wire between Moscow and Washington.

For all other procedures there is a (theoretical) possibility of breaking them. If the procedures are good, however, the time taken to break them is so long that it is practically impossible to do so, and these procedures can therefore be considered (practically) secure.

The book of Bruce Schneier [[Schneier1996](#)] offers a very good overview of the different algorithms. We basically distinguish between symmetric and asymmetric encryption procedures.

1.1 Symmetric encryption¹

For *symmetric* encryption sender and recipient must be in possession of a common (secret) key which they have exchanged before actually starting to communicate. The sender uses this key to encrypt the message and the recipient uses it to decrypt it.

All classical methods are of this type. Examples can be found within CrypTool or in [[Nichols1996](#)]. Now we want to consider more modern mechanisms.

The advantages of symmetric algorithms are the high speed with which data can be encrypted and decrypted. One disadvantage is the need for key management. In order to communicate with one another confidentially, sender and recipient must have exchanged a key using a secure channel before actually starting to communicate. Spontaneous communication between individuals who have never met therefore seems virtually impossible. If everyone wants to communicate with everyone else spontaneously at any time in a network of n subscribers, each subscriber must have previously exchanged a key with each of the other $n - 1$ subscribers. A total of $n(n - 1)/2$ keys must therefore be exchanged.

The most well-known symmetric encryption procedure is the DES-algorithm. The DES-algorithm

¹With CrypTool v1.3 you can execute the following modern symmetric encryption algorithms (using the menu **Crypt \ Symmetric**):

IDEA, RC2, RC4, DES (ECB), DES (CBC), Triple-DES (ECB), Triple-DES (CBC), MARS (AES candidate), RC6 (AES candidate), Serpent (AES candidate), Twofish (AES candidate), Rijndael (official AES algorithm).

has been developed by IBM in collaboration with the National Security Agency (NSA), and was published as a standard in 1975. Despite the fact that the procedure is relatively old, no effective attack on it has yet been detected. The most effective way of attacking consists of testing (almost) all possible keys until the right one is found (*brute-force-attack*). Due to the relatively short key length of effectively 56 bits (64 bits, which however include 8 parity bits), numerous messages encrypted using DES have in the past been broken. Therefore, the procedure can now only be considered to be conditionally secure. Symmetric alternatives to the DES procedure include the IDEA or Triple DES algorithms.

Up-to-the-minute procedure is the symmetric AES standard. The associated Rijndael algorithm was declared winner of the AES award on October 2nd, 2000 and thus succeeds the DES procedure.

More details about the AES algorithms and the AES candidates of the last round can be found within the online help of CrypTool².

1.1.1 New results about cryptanalysis of AES

Below you will find some results, which have recently called into question the security of the AES algorithm – from our point of view these doubts practically still remain unfounded. The following information is based on the original papers and the articles [Wobst-iX2002] and [Lucks-DuD2002].

AES with a minimum key length of 128 bit is still in the long run sufficiently secure against brute-force attacks – as long as the quantum computers aren't powerful enough. When announced as new standard AES was immune against all known crypto attacks, mostly based on statistical considerations and earlier applied to DES: using pairs of clear and cipher texts expressions are constructed, which are not completely at random, so they allow conclusions to the used keys. These attacks required unrealistically large amounts of intercepted data.

Cryptanalysts already label methods as “academic success” or as “cryptanalytic attack” if they are theoretically faster than the complete testing of all keys (brute force analysis). In the case of AES with the maximal key length (256 bit) exhaustive key search on average needs 2^{255} encryption operations. A cryptanalytic attack needs to be better than this. At present between 2^{75} and 2^{90} encryption operations are estimated to be performable only just for organizations, for example a security agency.

In their 2001-paper Ferguson, Schneier and Whiting [Ferguson2001] presented a new method of symmetric codes cryptanalysis: They described AES with a closed formula (in the form of a continued fraction) which was possible because of the “relatively” clear structure of AES. This formula consists of around 1000 trillion terms of a sum - so it does not help concrete practical cryptanalysis. Nevertheless curiosity in the academic community was awakened. It was already known, that the 128-bit AES could be described as an over-determined system of about 8000 quadratic equations (over an algebraic number field) with about 1600 variables (some of them are the bits of the wanted key) – equation systems of that size are in practice not solvable. This

²CrypTool online help: the index head-word **AES** leads to the 3 help pages: **AES candidates**, **The AES winner Rijndael** and **The Rijndael encryption algorithm**.

special equation system is relatively sparse, so only very few of the quadratic terms (there are about 1,280,000 possible quadratic terms in total) appear in the equation system.

The mathematicians Courtois and Pieprzyk [Courtois2002] published a paper in 2002, which got a great deal of attention amongst the crypto community: The pair had further developed the XL-method (eXtended Linearization), introduced at Eurocrypt 2000 by Shamir et al., to create the so called XSL-method (eXtended Sparse Linearization). The XL-method is a heuristic technique, which in some cases manages to solve big non-linear equation systems and which was till then used to analyze an asymmetric algorithm (HFE). The innovation of Courtois and Pieprzyk was, to apply the XL-method on symmetric codes: the XSL-method can be applied to very specific equation systems. A 256-bit AES could be attacked in roughly 2^{230} steps. This is still a purely academic attack, but also a direction pointer for a complete class of block ciphers. The major problem with this attack is that until now nobody has worked out, under what conditions it is successful: the authors specify in their paper necessary conditions, but it is not known, which conditions are sufficient. There are two very new aspects of this attack: firstly this attack is not based on statistics but on algebra. So attacks seem to be possible, where only very small amounts of cipher text are available. Secondly the security of a product algorithm³ does not exponentially increase with the number of rounds.

Currently there is a large amount of research in this area: for example Murphy and Robshaw presented a paper at Crypto 2002 [Robshaw2002a], which could dramatically improve cryptanalysis: the burden for a 128-bit key was estimated at about 2^{100} steps by describing AES as a special case of an algorithm called BES (Big Encryption System), which has an especially "round" structure. But even 2^{100} steps are beyond what is achievable in the foreseeable future. Using a 256 bit key the authors estimate that a XSL-attack will require 2^{200} operations.

More details can be found at:

<http://www.cryptosystem.net/aes>
<http://www.minrank.org/aes/>

So for 256-AES the attack is much more effective than brute-force but still far more away from any computing power which could be accessible in the short-to-long term.

The discussion is very controversial at the moment: Don Coppersmith (one of the inventors of DES) for example queries the practicability of the attack because XLS would provide no solution for AES [Coppersmith2002]. This implies that then the optimization of Murphy and Robshaw [Robshaw2002b] would not work.

³A cipher text can be used as input for another encryption algorithm. A cascade cipher is build up as a composition of different encryption transformations. The overall cipher is called product algorithm or cascade cipher (sometimes depending whether the used keys are statistically dependent or not).

Cascading does not always improve the security.

This process is also used *within* modern algorithms: They usually combine simple and, considered at its own, cryptologically relatively unsecure single steps in several rounds into an efficient overall procedure. Most block ciphers (e.g. DES, IDEA) are cascade ciphers.

Also serial usage of the same cipher with different keys (like with Triple-DES) is called cascade cipher.

1.1.2 Current status of brute-force attacks on symmetric algorithms (RC5)

The current status of brute-force attacks on symmetric encryption algorithms can be explained with the block cipher RC5.

Brute-force (exhaustive search, trial-and-error) means to completely examine all keys of the key space: so no special analysis methods have to be used. Instead, the cipher text is decrypted with all possible keys and for each resulting text it is checked, whether this is a meaningful clear text. A key length of 64 bit means at most $2^{64} = 18,446,744,073,709,551,616$ or about 18 trillion (GB) / 18 quintillion (US) keys to check⁴.

Companies like RSA Security provide so-called cipher challenges in order to quantify the security offered by well-known symmetric ciphers as DES, Triple-DES or RC5⁵. They offer prizes for those who manage to decipher cipher texts, encrypted with different algorithms and different key lengths, and to unveil the symmetric key (under controlled conditions). So theoretical estimates can be confirmed.

It is well-known, that the “old” standard algorithm DES with a fixed key length of 56 bit is no more secure: this was demonstrated already in January 1999 by the Electronic Frontier Foundation (EFF). With their specialized computer Deep Crack they cracked a DES encrypted message within less than a day⁶.

The current record for strong symmetric algorithms unveiled a key 64 bit long. The algorithm used was RC5, a block cipher with variable key size.

The RC5-64 challenge has been solved by the distributed.net team after 5 years⁷. In total 331,252 individuals co-operated over the internet to find the key. More than 15 trillion (GB) / 15 quintillion (US) keys were checked, until they found the right key.

This makes clear, that symmetric algorithms (even if they have no cryptographical weakness) using keys of size 64 bit are no more appropriate to keep sensible data private.

Similar cipher challenges are there for asymmetric algorithms (please see chapter 3.11.4).

1.2 Asymmetric encryption⁸

In the case of *asymmetric* encryption each subscriber has a personal pair of keys consisting of a *secret* key and a *public* key. The public key, as its name implies, is made public, e.g. in a key directory on the Internet.

If Alice⁹ wants to communicate with Bob, then she finds Bob’s public key in the directory and

⁴With CrypTool v1.3 you can also try brute-force attacks of modern symmetric algorithms (using the menu **Analysis \ Ciphertext-Only**): to achieve a result in an appropriate time with a single PC you should mark not more than 20 bit of the key as unknown.

⁵<http://www.rsasecurity.com/rsalabs/challenges/secretkey/index.html>

⁶<http://www.rsasecurity.com/rsalabs/challenges/des3/index.html>

⁷<http://distributed.net/pressroom/news-20020926.html>

⁸With CrypTool v1.3 you can execute RSA encryption and decryption (using the menu **Crypt \ Asymmetric**).

⁹In order to describe cryptographic protocols participants are often named Alice, Bob, ... (see [Schneier1996, p. 23]). Alice and Bob perform all 2-person-protocols. Alice will initiate all protocols and Bob answers. The attackers are

uses it to encrypt her message to him. She then sends this cipher text to Bob, who is then able to decrypt it again using his secret key. As only Bob knows his secret key, only he can decrypt messages addressed to him. Even Alice who sends the message cannot restore plaintext from the (encrypted) message she has sent. Of course, you must first ensure that the public key cannot be used to derive the private key.

Such a procedure can be demonstrated using a series of thief-proof letter boxes. If I have composed a message, I then look for the letter box of the recipient and post the letter through it. After that, I can no longer read or change the message myself, because only the legitimate recipient has the key for the letter box.

The advantage of asymmetric procedures is the easy key management. Let's look again at a network with n subscribers. In order to ensure that each subscriber can establish an encrypted connection to each other subscriber, each subscriber must possess a pair of keys. We therefore need $2n$ keys or n pairs of keys. Furthermore, no secure channel is needed before messages are transmitted, because all the information required in order to communicate confidentially can be sent openly. In this case, you simply have to pay attention to the accuracy (integrity and authenticity) of the public key. Disadvantage: Pure asymmetric procedures take a lot longer to perform than symmetric ones.

The most well-known asymmetric procedure is the RSA algorithm¹⁰, named after its developers Ronald Rivest, Adi Shamir and Leonard Adleman. The RSA algorithm was published in 1978. The concept of asymmetric encryption was first introduced by Whitfield Diffie and Martin Hellman in 1976. Today, the ElGamal procedures also play a decisive role, particularly the Schnorr variant in the DSA (Digital Signature Algorithm).

1.3 Hybrid procedures¹¹

In order to benefit from the advantages of symmetric and asymmetric techniques together, hybrid procedures are usually used (for encryption) in practice.

In this case the data is encrypted using symmetric procedures: the key is a session key generated by the sender randomly¹² that is only used for this message. This session key is then encrypted using the asymmetric procedure and transmitted to the recipient together with the message.

named Eve (eavesdropper) and Mallory (malicious active attacker).

¹⁰The RSA algorithm is extensively described in chapter 3.10 and later within this script. The RSA cryptosystem can be executed in many variations with CrypTool v1.3 (using the menu **Individual Procedures \ RSA Demonstration \ RSA Cryptosystem**). The topical research results concerning RSA are described in chapter 3.11.

¹¹Within CrypTool v1.3 you can get a visualization of this technique using the menu **Crypt \ Hybrid Demonstration**: this dialogue shows the single steps and its dependencies with concrete numbers.

¹²An important part of cryptographically secure techniques is to generate random numbers. Within CrypTool v1.3 you can check out different random number generators using the menu **Indiv. Procedures \ Generate Random Numbers**. Using the menu **Analysis \ Random Tests** you can apply different test methods for random data to binary documents.

Up to now CrypTool has concentrated on cryptographically strong pseudo random number generators. Only the integrated Secude generator involves a "pure" random source.

Recipients can determine the session key using their secret keys and then use the session key to encrypt the message. In this way, we can benefit from the easy key management of asymmetric procedures and encrypt large quantities of data quickly and efficiently using symmetric procedures.

1.4 Further details

Beside the information you can find in the following chapters, many other books and on a good number of websites the online help of CrypTool also offers very many details about the symmetric and asymmetric encryption methods.

References

- [Nichols1996] Randall K. Nichols,
Classical Cryptography Course, Volume 1 and 2,
Aegean Park Press 1996.
- [Schmeh2003] Klaus Schmeh,
Cryptography and Public Key Infrastructures on the Internet,
John Wiley & Sons Ltd., Chichester 2003.
A considerable, up-to-date, easy to read book, which also considers practical problems
such as standardisation or real existing software.
- [Schneier1996] Bruce Schneier,
Applied Cryptography, Protocols, Algorithms, and Source Code in C,
Wiley 1994, 2nd edition 1996.
- [Coppersmith2002] Don Coppersmith,
Re: Impact of Courtois and Pieprzyk results,
2002-09-19, “AES Discussion Groups” at
<http://aes.nist.gov/aes/>
- [Courtois2002] Nicolas Courtois, Josef Pieprzyk,
Cryptanalysis of Block Ciphers with Overdefined Systems of Equations,
received 10 Apr 2002, last revised 9 Nov 2002.
A different version, so called compact version of the first XSL attack, was published at
Asiacrypt Dec 2002.
<http://eprint.iacr.org/2002/044>
- [Ferguson2001] Niels Ferguson, Richard Schroepel, Doug Whiting,
A simple algebraic representation of Rijndael, Draft 2001/05/1,
<http://www.xs4all.nl/~vorpai/pubs/rdalgeq.html>
- [Lucks-DuD2002] Stefan Lucks, Rüdiger Weis,
Neue Ergebnisse zur Sicherheit des Verschlüsselungsstandards AES, in DuD Dec. 2002.
- [Robshaw2002a] S.P. Murphy, M.J.B. Robshaw,
Essential Algebraic Structure within the AES, June 5, 2002, Crypto 2002,
<http://www.isg.rhul.ac.uk/~mrobshaw/rijndael/rijndael.html>
- [Robshaw2002b] S.P. Murphy, M.J.B. Robshaw,
Comments on the Security of the AES and the XSL Technique, September 26, 2002,
<http://www.isg.rhul.ac.uk/~mrobshaw/rijndael/rijndael.html>
- [Wobst-iX2002] Reinhard Wobst,
Angekratzt - Kryptoanalyse von AES schreitet voran, in iX Dec. 2002,
plus the reader’s remark by Johannes Merkle in iX Feb. 2003.

Web links

1. AES or Rijndael Cryptosystem
<http://www.cryptosystem.net/aes>
<http://www.minrank.org/aes/>
2. AES Discussion Groups at NIST
<http://aes.nist.gov/aes>
3. distributed.net: “RC5-64 has been solved”
<http://distributed.net/pressroom/news-20020926.html>
4. RSA: “The RSA Secret Key Challenge”
<http://www.rsasecurity.com/rsalabs/challenges/secretkey/index.html>
5. RSA: “DES Challenge”
<http://www.rsasecurity.com/rsalabs/challenges/des3/index.html>
6. Further Links can be found at the CrypTool Homepage
<http://www.cryptool.org>

2 Prime Numbers

(Bernhard Esslinger, May 1999, Updates Nov. 2000, Dec. 2001, June 2003)

*Albert Einstein*¹³:
Progress requires exchange of knowledge.

2.1 What are prime numbers?

Prime numbers are whole, positive numbers greater than or equal to 2 that can only be divided by 1 and themselves. All other natural numbers greater than or equal to 2 can be formed by multiplying prime numbers.

The *natural* numbers $\mathbb{N} = \{1, 2, 3, 4, \dots\}$ thus comprise

- the number 1 (the unit value)
- the primes and
- the composite numbers.

Prime numbers are particularly important for 3 reasons:

- In number theory, they are considered to be the basic components of natural numbers, upon which numerous brilliant mathematical ideas are based.
- They are of extreme practical importance in modern cryptography (public key cryptography). The most common public key procedure, invented at the end of the 1970's, is RSA encryption. Only using (large) prime numbers for particular parameters can you guarantee that an algorithm is secure, both for the RSA procedure and for even more modern procedures (digital signature, elliptic curves).
- The search for the largest known prime numbers does not have any practical usage known to date, but requires the best computers, is an excellent benchmark (possibility for determining the performance of computers) and leads to new calculation methods on many computers (see also: <http://www.mersenne.org/prime.htm>).

Many people have been fascinated by prime numbers over the past two millennia. Ambition to make new discoveries about prime numbers has often resulted in brilliant ideas and conclusions. The following section provides an easily comprehensible introduction to the basics of prime numbers. We will also explain what is known about the distribution (density, number of prime numbers in particular intervals) of prime numbers and how prime number tests work.

¹³Albert Einstein, German physicist and Nobel Prize winner, Mar 14, 1879 – Apr 14, 1955.

2.2 Prime numbers in mathematics

Every whole number has a factor. The number 1 only has one factor, itself, whereas the number 12 has the six factors 1, 2, 3, 4, 6, 12. Many numbers can only be divided by themselves and by 1. With respect to multiplication, these are the “atoms” in the area of numbers. Such numbers are called prime numbers.

In mathematics, a slightly different (but equivalent) definition is used.

Definition 2.1. *A whole number $p \in \mathbf{N}$ is called prime if $p > 1$ and p only possesses the trivial factors ± 1 and $\pm p$.*

By definition, the number 1 is not a prime number. In the following sections, p will always denote a prime number.

The sequence of prime numbers starts with

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, \dots

The first 100 numbers include precisely 25 prime numbers. After this, the percentage of primes constantly decreases. Prime numbers can be factorised in a uniquely *trivial* way:

$$5 = 1 \cdot 5, \quad 17 = 1 \cdot 17, \quad 1,013 = 1 \cdot 1,013, \quad 1,296,409 = 1 \cdot 1,296,409.$$

All numbers that have 2 or more factors not equal 1 are called *composite* numbers. These include

$$4 = 2 \cdot 2, \quad 6 = 2 \cdot 3$$

as well as numbers that *look like primes*, but are in fact composite:

$$91 = 7 \cdot 13, \quad 161 = 7 \cdot 23, \quad 767 = 13 \cdot 59.$$

Theorem 2.1. *Each whole number m greater than 1 possesses a lowest factor greater than 1. This is a prime number p . Unless m is a prime number itself, then: p is less than or equal to the square root of m .*

All whole numbers greater than 1 can be expressed as a product of prime numbers — in a unique way. This is the claim of the 1st fundamental theorem of number theory (= fundamental theorem of arithmetic = fundamental building block of all positive integers).

Theorem 2.2. *Each element n of the natural numbers greater than 1 can be written as the product $n = p_1 \cdot p_2 \dots p_m$ of prime numbers. If two such factorisations*

$$n = p_1 \cdot p_2 \cdot \dots \cdot p_m = p'_1 \cdot p'_2 \cdot \dots \cdot p'_{m'}$$

are given, then they can be reordered such that $m = m'$ and for all i : $p_i = p'_i$. (p_1, p_2, \dots, p_m are called the prime factors of n).

In other words: each natural number other than 1 can be written as a product of prime numbers in precisely one way, if we ignore the order of the factors. The factors are therefore unique (the *expression as a product of factors* is unique)! For example,

$$60 = 2 \cdot 2 \cdot 3 \cdot 5 = 2^2 \cdot 3^1 \cdot 5^1.$$

And this — other than changing the order of the factors — is the only way in which the number 60 can be factorised. If you allow numbers other than primes as factors, there are several ways of factorising integers and the uniqueness is lost:

$$60 = 1 \cdot 60 = 2 \cdot 30 = 4 \cdot 15 = 5 \cdot 12 = 6 \cdot 10 = 2 \cdot 3 \cdot 10 = 2 \cdot 5 \cdot 6 = 3 \cdot 4 \cdot 5 = \dots$$

The following section is aimed more at those familiar with mathematical logic: The 1st fundamental theorem only appears to be obvious. We can construct numerous other sets of numbers (i.e. other than positive whole numbers greater than 1), for which numbers in the set cannot be expressed uniquely as a product of the prime numbers of the set: In the set $M = \{1, 5, 10, 15, 20, \dots\}$ there is no equivalent to the fundamental theorem under multiplication. The first five prime numbers of this sequence are 5, 10, 15, 20, 30 (note: 10 is prime, because 5 is not a factor of 10 in this set — the result is not an element of the given basic set M). Because the following applies in M :

$$100 = 5 \cdot 20 = 10 \cdot 10$$

and 5, 10, 20 are all prime numbers in this set, the expression as a product of prime factors is not unique here.

2.3 How many prime numbers are there?

For the natural numbers, the primes can be compared to elements in chemistry or the elementary particles in physics (see [Blum1999, p. 22]).

Although there are only 92 natural chemical elements, the number of prime numbers is unlimited. Even the Greek, Euclid¹⁴ knew this in the third century B.C.

Theorem 2.3 (Euclid¹⁵). *The sequence of prime numbers does not discontinue. Therefore, the quantity of prime numbers is infinite.*

¹⁴Euclid, a Greek mathematician of 4th and 3rd century B.C. He worked at the Egyptian academy of Alexandria and wrote “The Elements”, the most well known systematically textbook of the Greek mathematics.

¹⁵The common usage of the term does not denote Euclid as the inventor of the theorem rather; the true inventor is merely not as prominent. The theorem has already been distinguished and proven in Euclid’s Elements (Book IX, theorem 20). The phraseology is remarkable due to the fact that the word infinite is not used. The text reads as followed

Οί πρώτοι ἀριθμοὶ πλείους εἰσὶ παντὸς τοῦ προτεθέντος πλήθους πρώτων ἀριθμῶν,

the English translation of which is: the prime numbers are more than any previously existing amount of prime numbers.

His proof that there is an infinite number of primes is still considered to be a brilliant mathematical consideration and conclusion today (proof by contradiction). He assumed that there is only a finite number of primes and therefore a largest prime number. Based on this assumption, he drew logical conclusions until he obtained an obvious contradiction. This meant that something must be wrong. As there were no mistakes in the chain of conclusions, it could only be the assumption that was wrong. Therefore, there must be an infinite number of primes!

Euclid's proof by contradiction goes as follows:

Proof

Assumption: There is a *finite* number of primes.

Conclusion: Then these can be listed $p_1 < p_2 < p_3 < \dots < p_n$, where n is the (finite) number of prime numbers. p_n is therefore the largest prime. Euclid now looks at the number $a = p_1 \cdot p_2 \cdot \dots \cdot p_n + 1$. This number cannot be a prime number because it is not included in our list of primes. It must therefore be divisible by a prime, i.e. there is a natural number i between 1 and n , such that p_i divides the number a . Of course, p_i also divides the product $a - 1 = p_1 \cdot p_2 \cdot \dots \cdot p_n$, because p_i is a factor of $a - 1$. Since p_i divides the numbers a and $a - 1$, it also divides the difference of these numbers. Thus: p_i divides $a - (a - 1) = 1$. p_i must therefore divide 1, which is impossible.

Contradiction: Our assumption was false.

Thus there is an *infinite* number of primes (Cross-reference: [overview](#) under [2.8.1](#) of the number of prime numbers in various intervals). \square

Here we should perhaps mention yet another fact which is initially somewhat surprising. Namely, in the prime numbers sequence p_1, p_2, \dots , gaps between prime numbers can have an individually determined length n . It is undeniable that under the n succession of natural numbers

$$(n+1)! + 2, \dots, (n+1)! + (n+1),$$

none of them is a prime number since in order, the numbers $2, 3, \dots, (n+1)$ are comprised respectively as real divisors. ($n!$ means the product of the first n natural numbers therefore $n! = n * (n-1) * \dots * 2 * 1$).

2.4 The search for extremely large primes

The largest prime numbers known today have several millions digits (M-39), which is too big for us to imagine. The number of elementary particles in the universe is “only” estimated to be a 80-digit number ([See: overview under 2.8.3 of various orders of magnitude / dimensions](#)).

2.4.1 Special number types – Mersenne numbers

Almost all known huge prime numbers are special candidates, called *Mersenne numbers*¹⁶ of the form $2^p - 1$, where p is a prime. Not all Mersenne numbers are prime:

$$\begin{aligned} 2^2 - 1 &= 3 && \Rightarrow \text{prime} \\ 2^3 - 1 &= 7 && \Rightarrow \text{prime} \\ 2^5 - 1 &= 31 && \Rightarrow \text{prime} \\ 2^7 - 1 &= 127 && \Rightarrow \text{prime} \\ 2^{11} - 1 &= 2,047 = 23 \cdot 89 && \Rightarrow \text{NOT prime!} \end{aligned}$$

Even Mersenne knew that not all Mersenne numbers are prime (see exponent $p = 11$). A prime Mersenne number is called Mersenne prime number.

However, he is to be thanked for the interesting conclusion that a number of the form $2^n - 1$ cannot be a prime number if n is a composite number:

Theorem 2.4 (Mersenne). *If $2^n - 1$ is a prime number, then n is also a prime number.*

Proof

The theorem of Mersenne can be proved by contradiction. We therefore assume that there exists a composite natural number n (with real factorisation) $n = n_1 \cdot n_2$, with the property that $2^n - 1$ is a prime number.

From

$$\begin{aligned} (x^r - 1)((x^r)^{s-1} + (x^r)^{s-2} + \dots + x^r + 1) &= ((x^r)^s + (x^r)^{s-1} + (x^r)^{s-2} + \dots + x^r) \\ &\quad - ((x^r)^{s-1} + (x^r)^{s-2} + \dots + x^r + 1) \\ &= (x^r)^s - 1 = x^{rs} - 1, \end{aligned}$$

we conclude

$$2^{n_1 n_2} - 1 = (2^{n_1} - 1)((2^{n_1})^{n_2-1} + (2^{n_1})^{n_2-2} + \dots + 2^{n_1} + 1).$$

Because $2^n - 1$ is a prime number, one of the above two factors on the right-hand side must be equal to 1. This is the case if and only if $n_1 = 1$ or $n_2 = 1$. But this contradicts our assumption. Therefore the assumption is false. This means that there exists no composite number n , such that $2^n - 1$ is a prime. \square

Unfortunately this theorem only applies in one direction (the inverse statement does not apply, no equivalence): that means that there exist prime exponent for which the Mersenne number is **not** prime (see the above example $2^{11} - 1$, where 11 is prime, but $2^{11} - 1$ not).

Mersenne claimed that $2^{67} - 1$ is a prime number. There is also a mathematical history behind this claim: it first took over 200 years before Edouard Lucas (1842-1891) proved that this number

¹⁶Marin Mersenne, French priest and mathematician, Sep 08, 1588 – Sep 01, 1648.

is composite. However, he argued indirectly and did not name any of the factors. Then Frank Nelson Cole¹⁷ showed in 1903 which factors make up this composite number:

$$2^{67} - 1 = 147, 573, 952, 589, 676, 412, 927 = 193, 707, 721 \cdot 761, 838, 257, 287.$$

He admitted to having worked 20 years on the factorisation (expression as a product of prime factors)¹⁸ of this 21-digit decimal number!

Due to the fact that the exponents of the Mersenne numbers do not use all natural numbers, but only the primes, the *experimental space* is limited considerably. The currently known Mersenne prime numbers have the exponents

2; 3; 5; 7; 13; 17; 19; 31; 61; 89; 107; 127; 521; 607; 1,279; 2,203; 2,281; 3,217; 4,253;
4,423; 9,689; 9,941; 11,213; 19,937; 21,701; 23,207; 44,497; 86,243; 110,503; 132,049;
216,091; 756,839; 859,433; 1,257,787; 1,398,269; 2,976,221; 3,021,377; 6,972,593;
13,466,917.

Thus 39 Mersenne prime numbers are currently known. For the first 38 Mersenne prime numbers we know that this list is complete. The exponents until the 39th Mersenne prime number have not yet been checked completely (see chapter 2.5, prime number tests).

The 19th number with the exponent 4,253 was the first with at least 1,000 digits in decimal system (the mathematician Samuel Yates coined the expression *titanic* prime for this; it was discovered by Hurwitz in 1961); the 27th number with the exponent 44,497 was the first with at least 10,000 digits in the decimal system (Yates coined the expression *gigantic* prime for this. These names are now long outdated).

These numbers can be found at the following URLs:

http://reality.sgi.com/chongo/prime/prime_press.html
<http://www.utm.edu/>

M-37 – January 1998

The 37th Mersenne prime,

$$2^{3,021,377} - 1$$

was found in January 1998 and has 909,526 digits in the decimal system, which corresponds to 33 pages in the newspaper!

M-38 – June 1999

The 38th Mersenne prime, called M-38,

$$2^{6,972,593} - 1$$

¹⁷Frank Nelson Cole, American mathematician, Sep. 20, 1861 – May 26, 1926.

¹⁸Using CrypTool v1.3 you can factorize numbers in the following way: menu **Indiv. Procedures \ RSA Demonstration \ Factorisation of a Number**.

was discovered in June 1999 and has 2,098,960 digits in the decimal system (that corresponds to around 77 pages in the newspaper).

Right now (Feb. 2003) all exponents smaller than 6,972,593 have been tested and double-checked (see home page of the GIMPS project: <http://www.mersenne.org>): so we can be certain, that this is really the 38th Mersenne prime number and that there are no smaller undiscovered Mersenne primes.

M13466917 – December 2001 – M-39 ?

This number was discovered as 39th Mersenne prime (and already called M-39, despite it has not been proven yet, whether no further Mersenne prime numbers between M-38 und M13466917 do indeed exist),

$$2^{13,466,917} - 1,$$

at December 6, 2001 – more exactly, the verification of this number, found at November 14, 2001 by the Canadian student Michael Cameron, was successfully completed. This number has about 4 million decimal digits (exactly 4,053,946 digits). Trying only to print this number

$$(924947738006701322247758 \cdots 1130073855470256259071)$$

would require around 200 pages in the Financial Times.

40th Mersenne prime not found – June 2003

On June 1st a possible Mersenne prime was reported to the GIMPS server, which was checked afterwards as usual, before it was to be published. Unfortunately mid June the initiator and GIMPS project leader George Woltman had to tell, that two independent verification runs proved the number was composite. This was the first false positive report in 7 years.

GIMPS

Discovering the 39th Mersenne prime the GIMPS project (Great Internet Mersenne Prime Search) founded in 1996 already discovered for the 5th time the greatest Mersenne number to be proven being prime (<http://www.mersenne.org>).

Now more than 130,000 volunteers, amateurs and experts, are working for the GIMPS project. They connect their computers into the so called “primenet”, organized by the company entropia, to find such numbers using distributed computer programs.

Currently the 4 largest known primes are Mersenne prime numbers. Only the 5th biggest known prime is another number type: it belongs to the so called **generalized Fermat primes**.

2.4.2 EFF

This search is also spurred on by a competition started by the non-profit organisation EFF (Electronic Frontier Foundation) using the means of an unknown donator. The participants are rewarded with a total of 500,000 USD if they find the longest prime number. In promoting this project, the unknown donator is not looking for the quickest computer, but rather wants to draw people's attention to the opportunities offered by *cooperative networking*

<http://www.eff.org/coopawards/prime-release1.html>

The discoverer of M-38 received 50,000 USD from the EFF for discovering the first prime with more than 1 million decimal digits.

The next prize of 100,000 USD offered by EFF is for a proven prime with more than 10 million decimal digits.

According to the EFF rules for their prizes they offer in the next stage 150,000 USD for a proven prime with more than 100 million decimal digits.

Edouard Lucas (1842-1891) held the record for the longest prime number for over 70 years by proving that $2^{127} - 1$ is prime. No new record is likely to last that long.

2.5 Prime number tests

In order to implement secure encryption procedures we need extremely large prime numbers (in the region of $2^{2,048}$, i.e. numbers with 600 digits in the decimal system!).

Up to now we have looked for the prime factors in order to decide whether a number is prime. However, even if the smallest prime factor is enormous, the search takes too long. Factorising numbers using systematic computational division or using the *sieve of Eratosthenes* is only feasible using current computers for numbers with up to around 20 digits in the decimal system. The biggest number factorized into its 2 almost equal prime factors has 158 digits (see [chapter 3.11.4](#)).

However, if we know something about the *construction* of the number in question, there are extremely highly developed procedures that are much quicker. These procedures can determine the primality attribute of a number, but they cannot determine the prime factors of a number, if it is compound.

In the 17th century, Fermat¹⁹ wrote to Mersenne that he presumed that all numbers of the form

$$f(n) = 2^{2^n} + 1$$

are prime for all whole numbers $n \geq 0$ ([see below](#), chapter [2.6.4](#)).

As early as in the 19th century, it was discovered that the 29-digit number

$$f(7) = 2^{2^7} + 1$$

is not prime. However, it was not until 1970 that Morrison/Billhart managed to factorise it.

$$\begin{aligned} f(7) &= 340, 282, 366, 920, 938, 463, 463, 374, 607, 431, 768, 211, 457 \\ &= 59, 649, 589, 127, 497, 217 \cdot 5, 704, 689, 200, 685, 129, 054, 721 \end{aligned}$$

¹⁹Pierre de Fermat, French mathematician, Aug 17, 1601 – Jan 12, 1665.

Despite Fermat was wrong with this supposition, he is the originator of an important theorem in this area: Many rapid prime number tests are based on the (little) Fermat theorem put forward by Fermat in 1640 (see chapter 3.8.3).

Theorem 2.5 (“little” Fermat). *Let p be a prime number and a be any whole number, then for all a*

$$a^p \equiv a \pmod{p}.$$

This could also be formulated as follows:

Let p be a prime number and a be any whole number that is not a multiple of p (also $a \not\equiv 0 \pmod{p}$), then $a^{p-1} \equiv 1 \pmod{p}$.

If you are not used to calculating with remainders (modulo), please simply accept the theorem or first read chapter 3 “Introduction to Elementary Number Theory with Examples”. What is important here is that this sentence implies that if this equation is not met for any whole number a , then p is not a prime! The tests (e.g. for the first formulation) can easily be performed using the *test basis* $a = 2$.

This gives us a criterion for non-prime numbers, i.e. a negative test, but no proof that a number a is prime. Unfortunately Fermat’s theorem does not apply — otherwise we would have a simple proof of the prime number property (or to put it in other words, we would have a simple prime number criterion).

Pseudo prime numbers

Numbers n that have the property

$$2^n \equiv 2 \pmod{n}$$

but are not prime are called *pseudo prime numbers* (i.e. the exponent is not a prime). The first pseudo prime number is

$$341 = 11 \cdot 31.$$

Carmichael numbers

There are numbers that pass the Fermat test with all bases and yet are not prime: these numbers are called *Carmichael numbers*. The first of these is

$$561 = 3 \cdot 11 \cdot 17.$$

Strong pseudo prime numbers

A stronger test is provided by Miller/Rabin²⁰: it is only passed by so-called *strong pseudo prime numbers*. Again, there are strong pseudo prime numbers that are not primes, but this is much less often the case than for (simple) pseudo prime numbers or for Carmichael numbers. The smallest strong pseudo prime number base 2 is

$$15,841 = 7 \cdot 31 \cdot 73.$$

If you test all 4 bases, 2, 3, 5 and 7, you will find only one strong pseudo prime number up to $25 \cdot 10^9$, i.e. a number that passes the test and yet is not a prime number.

More extensive mathematics behind the Rabin test delivers the probability that the number examined is prime (such probabilities are currently around 10^{-60}).

Detailed descriptions of tests for finding out whether a number is prime can be found on Web sites such as:

<http://www.utm.edu/research/primes/mersenne.shtml>

<http://www.utm.edu/research/primes/prove/index.html>

2.6 Overview special number types and the search for a formula for primes

There are currently no useful, open (i.e. not recursive) formulae known that only deliver prime numbers (recursive means that in order to calculate the function the same function is used with a smaller variable). Mathematicians would be happy if they could find a formula that leaves gaps (i.e. does not deliver all prime numbers) but does not deliver any composite (non-prime) numbers.

Ideally, we would like, for the number n , to immediately be able to obtain the n -th prime number, i.e. for $f(8) = 19$ or for $f(52) = 239$.

Ideas for this can be found at

http://www.utm.edu/research/primes/notes/faq/p_n.html.

Cross-reference: [the table under 2.8.2](#) contains the precise values for the n th prime numbers for selected n .

The following enumeration contains the most common ideas for “prime number formulae” and what the status is for big special number types, trying to either prove primality or determine the prime factors:

2.6.1 Mersenne numbers $f(n) = 2^n - 1$ for n prime

As shown [above](#), this formula seems to deliver relatively large prime numbers but - as for $n = 11$ [$f(n) = 2,047$] - it is repeatedly the case that the result even with prime exponents is [not](#) prime.

²⁰In 1976 an efficient probabilistic primality test was published by Prof. Rabin, based on a number theoretic result of Prof. Miller from the year before.

Prof. Miller worked at the Carnegie-Mellon University, School of Computer Science. Prof. Rabin, born in 1931, worked at the Harvard and Hebrew University.

Today, all the Mersenne primes having less than around 2,000,000 digits are known (M-38):

<http://perso.wanadoo.fr/yves.gallot/primes/index.html>

2.6.2 Generalized Mersenne numbers $f(k, n) = k \cdot 2^n \pm 1$ for n prime and k small prime

For this first generalisation of the Mersenne numbers there are (for small k) also extremely quick prime number tests (see [Knuth1981]). This can be performed in practice using software such as the Proths software from Yves Gallot

<http://www.prothsearch.net/index.html>.

2.6.3 Generalized Mersenne numbers $f(b, n) = b^n \pm 1$ / The Cunningham project

This is another possible generalisation of the Mersenne numbers. The **Cunningham project** determines the factors of all composite numbers that are formed as follows:

$$f(b, n) = b^n \pm 1 \quad \text{for } b = 2, 3, 5, 6, 7, 10, 11, 12$$

(b is not equal to multiples of bases already used, such as 4, 8, 9).

Details of this can be found at:

<http://www.cerias.purdue.edu/homes/ssw/cun>

2.6.4 Fermat numbers²¹ $f(n) = 2^{2^n} + 1$

As mentioned **above** in chapter 2.5, Fermat wrote to Mersenne regarding his assumption, that all numbers of this type are primes. This assumption was disproved by Euler (1732). The prime 641 divides $f(5)$ ²².

$f(0) = 2^{2^0} + 1 = 2^1 + 1$	$= 3$	\mapsto prime
$f(1) = 2^{2^1} + 1 = 2^2 + 1$	$= 5$	\mapsto prime
$f(2) = 2^{2^2} + 1 = 2^4 + 1$	$= 17$	\mapsto prime
$f(3) = 2^{2^3} + 1 = 2^8 + 1$	$= 257$	\mapsto prime
$f(4) = 2^{2^4} + 1 = 2^{16} + 1$	$= 65,537$	\mapsto prime
$f(5) = 2^{2^5} + 1 = 2^{32} + 1$	$= 4,294,967,297 = 641 \cdot 6,700,417$	\mapsto NOT prime!
$f(6) = 2^{2^6} + 1 = 2^{64} + 1$	$= 18,446,744,073,709,551,617$	
	$= 274,177 \cdot 67,280,421,310,721$	\mapsto NOT prime!
$f(7) = 2^{2^7} + 1 = 2^{128} + 1$	$=$ (see page 24)	\mapsto NOT prime!

Within the project “Distributed Search for Fermat Number Dividers” offered by Leonid Durman there is also progress in finding new monster primes:

²¹The Fermat prime numbers play a role in circle division. As proven by Gauss a regular p -edge can only be constructed with the use of a pair of compasses and a ruler, when p is a Fermat prime number.

²²Surprisingly this number can easily be found by using Fermat’s theorem (see e.g. [Scheid1994, p. 176])

<http://www.fermatsearch.org/>

This website links to other webpages in Russian, Italian and German.

The discovered factors can be compound integers or primes.

On February 22, 2003 John Cosgrave discovered

- the largest composite Fermat number to date and
- the largest prime non-simple Mersenne number so far with 645,817 decimal digits.

The Fermat number

$$f(2, 145, 351) = 2^{(2^{2,145,351})} + 1$$

is divisible by the prime

$$p = 3 * 2^{2,145,353} + 1$$

This prime p is the largest known prime generalized Mersenne number and the 5th largest known prime number today. It was discovered only a few days after Michael Angel has come across the largest generalized Fermat number ([see below](#)). This moved GIMPS' first prime discovery, M-35 = M1398269, into 7th place.

This work was done using NewPGen from Paul Jobling's, PRP from George Woltman's, Proth from Yves Gallot's programs and also the Proth-Gallot group at St. Patrick's College, Dublin.

More details are in

http://www.fermatsearch.org/history/cosgrave_record.htm/

2.6.5 Generalized Fermat numbers²³ $f(b, n) = b^{2^n} + 1$

Generalized Fermat numbers are more numerous than Mersenne numbers of a equal size and many of them are waiting to be discovered to fill the big gaps between the Mersenne primes already found or still undiscovered. Progress in number theory made it possible that numbers, where the representation is not limited to the base 2, can be tested at almost the same speed than a Mersenne number.

Yves Gallot wrote the program Proth.exe to investigate generalized Fermat numbers.

Using this program at February 16, 2003 Michael Angel discovered the largest of them till then with 628,808 digits, which so became the 5th largest known prime number:

$$b^{2^{17}} + 1 = 62,722^{131,072} + 1.$$

This moved GIMPS' first prime discovery, M-35 = M1398269, into 6th place.

Today (date June 2003) this generalized Fermat number is the 6th largest known prime number.

More details are in

<http://www.prothsearch.net/index.html>

<http://perso.wanadoo.fr/yves.gallot/primes/index.html>

²³The base of this power is no longer restricted to 2 .
Even more generic would be: $f(b, c, n) = b^{c^n} \pm 1$

2.6.6 Carmichael numbers

As mentioned [above](#) in chapter [2.5](#) not all Carmichael numbers are prime.

2.6.7 Pseudo prime numbers

See [above](#) in chapter [2.5](#).

2.6.8 Strong pseudo prime numbers

See [above](#) in chapter [2.5](#).

2.6.9 Idea based on Euclid's proof $p_1 \cdot p_2 \cdots p_n + 1$

This idea is based on [Euclid's proof](#) that there are infinite many prime numbers.

$2 \cdot 3 + 1$	$= 7$	\mapsto prime
$2 \cdot 3 \cdot 5 + 1$	$= 31$	\mapsto prime
$2 \cdot 3 \cdot 5 \cdot 7 + 1$	$= 211$	\mapsto prime
$2 \cdot 3 \cdots 11 + 1$	$= 2,311$	\mapsto prime
$2 \cdot 3 \cdots 13 + 1$	$= 59 \cdot 509$	\mapsto NOT prime!
$2 \cdot 3 \cdots 17 + 1$	$= 19 \cdot 97 \cdot 277$	\mapsto NOT prime!

2.6.10 As above but -1 except $+1$: $p_1 \cdot p_2 \cdots p_n - 1$

$2 \cdot 3 - 1$	$= 5$	\mapsto prime
$2 \cdot 3 \cdot 5 - 1$	$= 29$	\mapsto prime
$2 \cdot 3 \cdots 7 - 1$	$= 11 \cdot 19$	\mapsto NOT prime!
$2 \cdot 3 \cdots 11 - 1$	$= 2,309$	\mapsto prime
$2 \cdot 3 \cdots 13 - 1$	$= 30,029$	\mapsto prime
$2 \cdot 3 \cdots 17 - 1$	$= 61 \cdot 8,369$	\mapsto NOT prime!

2.6.11 Euclidean numbers $e_n = e_0 \cdot e_1 \cdots e_{n-1} + 1$ with $n \geq 1$ and $e_0 := 1$

e_{n-1} is not the $(n-1)$ th prime number, but the number previously found here. Unfortunately this formula is not open but recursive. The sequence starts with

$e_1 = 1 + 1$	$= 2$	\mapsto prime
$e_2 = e_1 + 1$	$= 3$	\mapsto prime
$e_3 = e_1 \cdot e_2 + 1$	$= 7$	\mapsto prime
$e_4 = e_1 \cdot e_2 \cdot e_3 + 1$	$= 43$	\mapsto prime
$e_5 = e_1 \cdot e_2 \cdots e_4 + 1$	$= 13 \cdot 139$	\mapsto NOT prime!
$e_6 = e_1 \cdot e_2 \cdots e_5 + 1$	$= 3, 263, 443$	\mapsto prime
$e_7 = e_1 \cdot e_2 \cdots e_6 + 1$	$= 547 \cdot 607 \cdot 1, 033 \cdot 31, 051$	\mapsto NOT prime!
$e_8 = e_1 \cdot e_2 \cdots e_7 + 1$	$= 29, 881 \cdot 67, 003 \cdot 9, 119, 521 \cdot 6, 212, 157, 481$	\mapsto NOT prime!

e_9, \dots, e_{17} are also composite, which means that this formula is not particularly useful.

Remark: However, what is special about these numbers is that any pair of them does not have a common factor other than 1²⁴. Therefore they are *relatively prime*.

2.6.12 $f(n) = n^2 + n + 41$

This sequence starts off very *promisingly*, but is far from being a proof.

$f(0) = 41$	\mapsto prime
$f(1) = 43$	\mapsto prime
$f(2) = 47$	\mapsto prime
$f(3) = 53$	\mapsto prime
$f(4) = 61$	\mapsto prime
$f(5) = 71$	\mapsto prime
$f(6) = 83$	\mapsto prime
$f(7) = 97$	\mapsto prime
\vdots	
$f(39) = 1, 601$	\mapsto prime
$f(40) = 11 \cdot 151$	\mapsto NOT prime!
$f(41) = 41 \cdot 43$	\mapsto NOT prime!

The first 40 values are prime numbers (which have the obvious regularity that their difference starts with 2 and increases by 2 each time), but the 41th and 42th values are not prime numbers. It is easy to see that $f(41)$ cannot be a prime number: $f(41) = 41^2 + 41 + 41 = 41(41 + 1 + 1) = 41 \cdot 43$.

²⁴This can easily be shown via the following *greatest common divisor* (*gcd*) rule $\gcd(a, b) = \gcd(b - \lfloor b/a \rfloor \cdot a, a)$ (see page 100): We have for $i < j$:

$$\gcd(e_i, e_j) \leq \gcd(e_1 \cdots e_i \cdots e_{j-1}, e_j) = \gcd(e_j - e_1 \cdots e_i \cdots e_{j-1}, e_1 \cdots e_i \cdots e_{j-1}) = \gcd(1, e_1 \cdots e_i \cdots e_{j-1}) = 1.$$

2.6.13 $f(n) = n^2 - 79 \cdot n + 1,601$

This function delivers prime numbers for all values from $n = 0$ to $n = 79$. Unfortunately $f(80) = 1,681 = 11 \cdot 151$ is not a prime number. To this date, no function has been found that delivers more prime numbers in a row. On the other hand, each prime occurs twice (first in the decreasing then in the increasing sequence), which means that the algorithm delivers a total of 40 difference prime values (the same ones as delivered by the function in chapter 2.6.12).

$f(0) = 1,601$	\mapsto prime	$f(28) = 173$	\mapsto prime
$f(1) = 1,523$	\mapsto prime	$f(29) = 151$	\mapsto prime
$f(2) = 1,447$	\mapsto prime	$f(30) = 131$	\mapsto prime
$f(3) = 1,373$	\mapsto prime	$f(31) = 113$	\mapsto prime
$f(4) = 1,301$	\mapsto prime	$f(32) = 97$	\mapsto prime
$f(5) = 1,231$	\mapsto prime	$f(33) = 83$	\mapsto prime
$f(6) = 1,163$	\mapsto prime	$f(34) = 71$	\mapsto prime
$f(7) = 1,097$	\mapsto prime	$f(35) = 61$	\mapsto prime
$f(8) = 1,033$	\mapsto prime	$f(36) = 53$	\mapsto prime
$f(9) = 971$	\mapsto prime	$f(37) = 47$	\mapsto prime
$f(10) = 911$	\mapsto prime	$f(38) = 43$	\mapsto prime
$f(11) = 853$	\mapsto prime	$f(39) = 41$	\mapsto prime
$f(12) = 797$	\mapsto prime	$f(40) = 41$	\mapsto prime
$f(13) = 743$	\mapsto prime	$f(41) = 43$	\mapsto prime
$f(14) = 691$	\mapsto prime	$f(42) = 47$	\mapsto prime
$f(15) = 641$	\mapsto prime	$f(43) = 53$	\mapsto prime
$f(16) = 593$	\mapsto prime	\dots	
$f(17) = 547$	\mapsto prime	$f(77) = 1,447$	\mapsto prime
$f(18) = 503$	\mapsto prime	$f(78) = 1,523$	\mapsto prime
$f(19) = 461$	\mapsto prime	$f(79) = 1,601$	\mapsto prime
$f(20) = 421$	\mapsto prime	$f(80) = 11 \cdot 151$	\mapsto NOT prime!
$f(21) = 383$	\mapsto prime	$f(81) = 41 \cdot 43$	\mapsto NOT prime!
$f(22) = 347$	\mapsto prime	$f(82) = 1,847$	\mapsto prime
$f(21) = 383$	\mapsto prime	$f(83) = 1,933$	\mapsto prime
$f(22) = 347$	\mapsto prime	$f(84) = 43 \cdot 47$	\mapsto NOT prime!
$f(23) = 313$	\mapsto prime		
$f(24) = 281$	\mapsto prime		
$f(25) = 251$	\mapsto prime		
$f(26) = 223$	\mapsto prime		
$f(27) = 197$	\mapsto prime		

2.6.14 Polynomial functions $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x^1 + a_0$ (a_i in \mathbb{Z} , $n \geq 1$)

There exists no such polynomial that for all x in \mathbb{Z} only delivers prime values. For a proof of this, please refer to [Padberg1996, p. 83 f.], where you will also find further details about prime number formulae.

This means there is no hope in looking for further formulae (functions) similar to that in chap. 2.6.12 or chap. 2.6.13.

2.6.15 Catalan's conjecture²⁵

Catalan conjectured that C_4 is a prime:

$$\begin{aligned}C_0 &= 2, \\C_1 &= 2^{C_0} - 1, \\C_2 &= 2^{C_1} - 1, \\C_3 &= 2^{C_2} - 1, \\C_4 &= 2^{C_3} - 1, \dots\end{aligned}$$

(see <http://www.utm.edu/research/primes/merenne.shtml> under Conjectures and Unsolved Problems).

This sequence is also defined recursively and increases extremely quickly. Does it only consist of primes?

$C(0) = 2$	\mapsto prime
$C(1) = 2^2 - 1 = 3$	\mapsto prime
$C(2) = 2^3 - 1 = 7$	\mapsto prime
$C(3) = 2^7 - 1 = 127$	\mapsto prime
$C(4) = 2^{127} - 1 = 170, 141, 183, 460, 469, 231, 731, 687, 303, 715, 884, 105, 727$	\mapsto prime

It is not (yet) known whether C_5 and all higher elements are prime, but this is not very likely. In any case, it has not been proved that this formula delivers only primes.

2.7 Density and distribution of the primes

As Euclid discovered, there is an infinite number of primes. However, some infinite sets are *denser* than others. Within the set of natural numbers, there is an infinite number of even, uneven and square numbers.

The following proves that there are more even numbers than square ones:

- the size of the n th element:

The n th element of the even numbers is $2n$; the n th element of the square numbers is n^2 . Because for all $n > 2$: $2n < n^2$, the n th even number occurs much earlier than the n th square number. Thus the even numbers are distributed more densely and we can say that there are more even numbers than square ones.

²⁵Eugene Charles Catalan, Belgian mathematician, May 5, 1814–Feb 14, 1894.

After him the so-called *Catalan numbers* $A(n) = (1/(n+1)) * (2n)!/(n!)^2$
 $= 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, 208012, 742900, 2674440, 9694845, \dots$ are named.

- the number of values that are less than or equal to a certain *maximum value* x in \mathbb{R} is:
There are $\lfloor x/2 \rfloor$ such even numbers and $\lfloor \sqrt{x} \rfloor$ square numbers. Because for large x the value $x/2$ is much greater than the square root of x , we can again say that there are more even numbers.

The value of the n -th prime $P(n)$

Theorem 2.6. *For large n : The value of the n -th prime $P(n)$ is asymptotic to $n \cdot \ln(n)$, i.e. the limit of the relation $P(n)/(n \cdot \ln n)$ is equal to 1 if n tends to infinity.*

For $n > 5$, $P(n)$ lies between $2n$ and n^2 . This means that there are fewer prime numbers than even natural numbers but more prime numbers than square numbers²⁶.

The number of prime numbers $PI(x)$

The definition is similar for the number of prime numbers $PI(x)$ that do not exceed the maximum value x :

Theorem 2.7. *$PI(x)$ is asymptotic to $x/\ln(x)$.*

This is the **prime number theorem**. It was put forward by Legendre²⁷ and Gauss²⁸ but not proved until over 100 years later.

Cross-reference: The overview under 2.8.1 shows the number of prime numbers in various intervals.

These formulae, which only apply when n tends to infinity, can be replaced by more precise formulae. For $x \geq 67$:

$$\ln(x) - 1,5 < x/PI(x) < \ln(x) - 0,5$$

Given that we know $PI(x) = x/\ln x$ only for very large x (x tending towards infinity), we can create the following overview:

x	$\ln(x)$	$x/\ln(x)$	$PI(x)$ (counted)	$PI(x)/(x/\ln(x))$
10^3	6.908	144	168	1.160
10^6	13.816	72,386	78,498	1.085
10^9	20.723	48,254,942	50,847,534	1.054

For a binary number²⁹ x of the length of 250 bits (2^{250} is approximately $= 1.809251 \cdot 10^{75}$) it is:

$$PI(x) = 2^{250}/(250 \cdot \ln 2) \text{ is approximately } = 2^{250}/173.28677 = 1.045810 \cdot 10^{73}.$$

²⁶Please refer to [the table 2.8.2](#)

²⁷Adrien-Marie Legendre, French mathematician, Sep 18, 1752 – Jan 10, 1833.

²⁸Carl Friedrich Gauss, German mathematician and astronomer, Apr 30, 1777–Feb 23, 1855.

²⁹Number written in the binary system consists only of the digits 0 and 1.

We can therefore expect that the set of numbers with a bit length of less than 250 contains approximately 10^{73} primes (a reassuring result?!).

We can also express this as follows: Let us consider a *random* natural number n . Then the probability that this number is prime is around $1/\ln(n)$. For example, let us take numbers in the region of 10^{16} . Then we must consider $16 \cdot \ln 10 = 36,8$ numbers (on average) until we find a prime. A precise investigation shows: There are 10 prime numbers between $10^{16} - 370$ and $10^{16} - 1$.

Under the heading *How Many Primes Are There* at

<http://www.utm.edu/research/primes/howmany.shtml>

you will find numerous other details.

Using the following Web site:

<http://www.math.Princeton.EDU/~arbooker/nthprime.html>

you can easily determine $PI(x)$.

The **distribution** of primes displays several irregularities for which no system has yet been found: On the one hand, many occur closely together, like 2 and 3, 11 and 13, 809 and 811, on the other hand large gaps containing no primes also occur. For example, no primes lie between 113 and 127, 293 and 307, 317 and 331, 523 and 541, 773 and 787, 839 and 853 as well as between 887 and 907.

For details, please see:

<http://www.utm.edu/research/primes/notes/gaps.html>

This is precisely part of what motivates mathematicians to discover its secrets.

Sieve of Eratosthenes

An easy way of calculating all $PI(x)$ primes less than or equal to x is to use the sieve of Eratosthenes. In the 3rd century B.C., he found an extremely easy, automatic way of finding this out. To begin with, you write down all numbers from 2 to x , circle 2, then cross out all multiples of 2. Next, you circle the lowest number that hasn't been circled or crossed out (3) and again cross out all multiples of this number, etc. You only need to continue until you reach the largest number whose square is less than or equal to x .

Apart from 2, prime numbers are never even. Apart from 2 and 5, prime numbers never end in 2, 5 or 0. So you only need to consider numbers ending in 1, 3, 7, 9 anyway (there are infinite primes ending in these numbers; see [Tietze1973, vol. 1, p. 137]).

You can now find a large number of finished programs on the Internet - often complete with source code - allowing you to experiment with large numbers yourself (see chapter 2.6). You also have access to large databases that contain either a large number of primes or the factorisation of numerous composite numbers.

2.8 Notes

Proven statements / theorems about primes

- For each number n in \mathbf{N} there are n consecutive natural numbers that are not primes. A proof of this can be found in [Padberg1996, p. 79].
- Paul Erdős³⁰ proved: Between each random number not equal to 1 and its double, there is at least one prime. He was not the first to prove this theorem, but proved it in a much simpler manner than those before him.
- There is a real number a such that the function $f : \mathbf{N} \rightarrow \mathbb{Z}$ where $n \mapsto a^{3^n}$ only delivers primes for all n (see [Padberg1996, p. 82]). Unfortunately, problems arise when we try to determine a (see below).

Unproven statements / conjectures about primes

- Christian Goldbach³¹ conjectured: Every even natural number greater than 2 can be represented as the sum of two prime numbers. Computers have verified³² the Goldbach conjecture for all even numbers up to $4 \cdot 10^{14}$ but no general proof has yet been found³³.
- Bernhard Riemann³⁴ put forward a formula for the distribution of primes that would further improve the estimate. However, this has neither been proved nor disproved so far.

³⁰Paul Erdős, Hungarian mathematician, Mar 26, 1913–Sep 20, 1996.

³¹Christian Goldbach, German mathematician, Mar 18, 1690–Nov 20, 1764.

³²It is generally accepted today, that the Goldbach Conjecture is true, i. e. valid for all even natural numbers greater than 2. In 1999, mathematician Jörg Richstein from the computer sciences institute at the University of Giessen, studied even numbers up to 400 billion and found no contradictory example (see <http://www.informatik.uni-giessen.de/staff/richstein/de/Goldbach.html>). Nevertheless, this does not provide us with general proof.

The fact is that despite all efforts, Goldbach's conjecture has to date not been proven. This leads one to believe that since the pioneer work of the Austrian mathematician Kurt Gödel is well-known, not every true mathematical theorem is provable (see <http://www.mathematik.ch/mathematiker/goedel.html>). Perhaps Goldbach's conjecture was correct, but in any case the proof will never be found. Conversely, that will presumably also remain unproven.

³³The English publisher *Faber* and the American publisher *Bloomsbury* issued in 2000 the 1992 published book "Uncle Petros and Goldbach's Conjecture" by Apostolos Doxiadis. It's the story of an old maths professor who fails to prove a more than 250 year old puzzle. To boost the sales figures the English and American publishers have offered a prize of 1 million USD, if someone can prove the conjecture – which should be published by 2004 in a well-known mathematical journal.

Surprisingly only British and American citizens are allowed to participate.

The theorem which has come closest so far to Goldbach's conjecture was proved by Chen Jing-Run in 1966 in a way which is somewhat hard to understand: Each even integer greater than 2 is the sum of one prime and of the product of two primes. E.g.: $20 = 5 + 3 \cdot 5$.

Most of the research about the Goldbach conjecture is collected in the book: "Goldbach Conjecture", ed. Wang Yuan, 1984, World scientific Series in Pure Maths, Vol. 4.

Especially this conjecture makes it clear, that even today we do not have a complete understanding of the deeper connections between addition and multiplication of natural numbers.

³⁴Bernhard Riemann, German mathematician, Sep 17, 1826–Jul 20, 1866.

Open questions

Twin primes are prime numbers whose difference is 2. Examples include 5 and 7 or 101 and 103. Triplet primes, however, only occur once: 3, 5, 7. For all other sets of three consecutive uneven numbers, one of them is always divisible by 3 and thus not a prime.

- The number of twin primes is an open question: infinite or limited number? The largest twin primes known today are $1,693,965 \cdot 2^{66,443} \pm 1$.
- Does a formula exist for calculating the number of twin primes per interval?
- The [above](#) proof of the function $f : N \rightarrow Z$ with $n \mapsto a^{3^n}$ only guarantees the existence of such a number a . How can we determine this number a and will it have a value, making the function also of some practical interest?
- Is there an infinite number of Mersenne prime numbers?
- Is there an infinite number of Fermat prime numbers?
- Does a polynomial time algorithm exist for calculating the prime factors of a number (see [Klee1997](#), p. 167)? This question can be divided into the three following questions:
 - Does a polynomial time algorithm exist that decides whether a number is prime? This question has been answered by the AKS algorithm (see chapter [3.11.3](#), “Primes in P”: Primality testing is polynomial).
 - Does a polynomial time algorithm exist that calculates for a composite number from how many prime factors it is made up (without calculating these factors)?
 - Does a polynomial time algorithm exist that calculates for a composite number n a non-trivial (i.e. other than 1 and n) factor of n ?^{[35](#)}

At the end of chapter [3.11.4](#), section [RSA-160](#) you can see the dimensions for which numbers the current algorithms deliver results (test for primality, factorization).

Further interesting topics regarding prime numbers

This chapter doesn't consider other number theory topics such as divisibility rules, modulus calculation, modular inverses, modular powers and roots, Chinese remainder theorem, Euler PHI function, perfect numbers. Some of these are considered in the next chapter.

³⁵Please compare chapters [3.11.3](#) and [3.11.4](#).

2.8.1 Number of prime numbers in various intervals

Ten-sized intervals		Hundred-sized intervals		Thousand-sized intervals	
Interval	Number	Interval	Number	Interval	Number
1-10	4	1-100	25	1-1000	168
11-20	4	101-200	21	1001-2000	135
21-30	2	201-300	16	2001-3000	127
31-40	2	301-400	16	3001-4000	120
41-50	3	401-500	17	4001-5000	119
51-60	2	501-600	14	5001-6000	114
61-70	2	601-700	16	6001-7000	117
71-80	3	701-800	14	7001-8000	107
81-90	2	801-900	15	8001-9000	110
91-100	1	901-1000	14	9001-10000	112

Further intervals:

Interval	Number	Average number per 1000
1 - 10,000	1,229	122.900
1 - 100,000	9,592	95.920
1 - 1,000,000	78,498	78.498
1 - 10,000,000	664,579	66.458
1 - 100,000,000	5,761,455	57.615
1 - 1,000,000,000	50,847,534	50.848
1 - 10,000,000,000	455,052,512	45.505

2.8.2 Indexing prime numbers (n -th prime number)

Index	Precise value	Rounded value	Comment
1	2	2	
2	3	3	
3	5	5	
4	7	7	
5	11	11	
6	13	13	
7	17	17	
8	19	19	
9	23	23	
10	29	29	
100	541	541	All prime numbers up to 1E+07 were known at the beginning of the 20th century.
1,000	7,917	7,917	
664,559	9,999,991	9.99999E+06	
1E+06	15,485,863	1.54859E+07	
6E+06	104,395,301	1.04395E+08	
1E+07	179,424,673	1.79425E+08	
1E+09	22,801,763,489	2.28018E+10	
1E+12	29,996,224,275,833	2.99962E+13	
			This prime was discovered in 1959.

Comment: With gaps, extremely large prime numbers were discovered at an early stage.

Web links:

<http://www.math.Princeton.EDU/~arbooker/nthprime.html>.

Output of the n -th prime number

See http://www.utm.edu/research/primes/notes/by_year.html.

2.8.3 Orders of magnitude / dimensions in reality

In the description of cryptographic protocols and algorithms, numbers occur that are so large or so small that they are inaccessible to our intuitive understanding. It may therefore be useful to provide comparative numbers from the real world around us so that we can develop a feeling for the security of cryptographic algorithms. Some of the numbers listed below originate from [Schwenk1996] and [Schneier1996, p.18].

Probability that you will be hijacked on your next flight	$5.5 \cdot 10^{-6}$	
Annual probability of being hit by lightning	10^{-7}	
Probability of 6 correct numbers in the lottery	$7.1 \cdot 10^{-8}$	
Risk of being hit by a meteorite	$1.6 \cdot 10^{-12}$	
<hr/>		
Time until the next ice age (in years)	14,000	(2^{14})
Time until the sun dies (in years)	10^9	(2^{30})
Age of the earth (in years)	10^9	(2^{30})
Age of the universe (in years)	10^{10}	(2^{34})
Number of the earth's atoms	10^{51}	(2^{170})
Number of the sun's atoms	10^{57}	(2^{190})
Number of atoms in the universe (without dark material)	10^{77}	(2^{265})
Volume of the universe (in cm^3)	10^{84}	(2^{280})

2.8.4 Special values in the binary and decimal systems

Dual system	Decimal system
2^{10}	1024
2^{40}	$1.09951 \cdot 10^{12}$
2^{56}	$7.20576 \cdot 10^{16}$
2^{64}	$1.84467 \cdot 10^{19}$
2^{80}	$1.20893 \cdot 10^{24}$
2^{90}	$1.23794 \cdot 10^{27}$
2^{112}	$5.19230 \cdot 10^{33}$
2^{128}	$3.40282 \cdot 10^{38}$
2^{150}	$1.42725 \cdot 10^{45}$
2^{160}	$1.46150 \cdot 10^{48}$
2^{250}	$1.80925 \cdot 10^{75}$
2^{256}	$1.15792 \cdot 10^{77}$
2^{320}	$2.13599 \cdot 10^{96}$
2^{512}	$1.34078 \cdot 10^{154}$
2^{768}	$1.55252 \cdot 10^{231}$
2^{1024}	$1.79769 \cdot 10^{308}$
2^{2048}	$3.23170 \cdot 10^{616}$

Calculation using GMP, for example: <http://www.gnu.ai.mit.edu>.

References

- [Aaronson2003] Scott Aaronson,
The Prime Facts: From Euclid to AKS,
<http://www.cs.berkeley.edu/~aaronson/prime.ps>.
Only after I had completed this article, did I come across the extremely well-written paper by Scott Aaronson, which also offers a didactically well done introduction to this topic. It is humorous and easy to read but at the same time precise and erudite.
- [Bartholome1996] A. Bartholom, J. Rung, H. Kern,
Zahlentheorie für Einsteiger, Vieweg 1995, 2nd edition 1996.
- [Blum1999] W. Blum,
Die Grammatik der Logik, dtv, 1999.
- [Bundschuh1998] Peter Bundschuh,
Einführung in die Zahlentheorie, Springer 1988, 4th edition 1998.
- [Doxiadis2000] Apostolos Doxiadis,
Uncle Petros and the Goldbach's Conjecture,
Faber/Bloomsbury, 2000.
- [Graham1989] R.E. Graham, D.E. Knuth, O. Patashnik,
Concrete Mathematics, Addison-Wesley, 1989.
- [Klee1997] V. Klee, S. Wagon,
Ungelöste Probleme in der Zahlentheorie und der Geometrie der Ebene,
Birkhäuser Verlag, 1997.
- [Knuth1981] Donald E. Knuth,
The Art of Computer Programming, vol 2: Seminumerical Algorithms,
Addison-Wesley, 1969, 2nd edition 1981.
- [Lorenz1993] F. Lorenz,
Algebraische Zahlentheorie, BI Wissenschaftsverlag, 1993.
- [Padberg1996] F. Padberg,
Elementare Zahlentheorie, Spektrum Akademischer Verlag 1988, 2nd edition 1996.
- [Pieper1983] H. Pieper,
Zahlen aus Primzahlen, Verlag Harri Deutsch 1974, 3rd edition 1983.
- [Richstein1999] J. Richstein,
*Verifying the Goldbach Conjecture up to $4 * 10^{14}$* , Mathematics of Computation.
- [Scheid1994] Harald Scheid,
Zahlentheorie, BI Wissenschaftsverlag, 2nd edition, 1994.

- [Schneier1996] Bruce Schneier,
Applied Cryptography, Protocols, Algorithms, and Source Code in C,
Wiley and Sons, 2nd edition 1996.
- [Schroeder1999] M.R. Schroeder,
Number Theory in Science and Communication,
Springer 1984, 3rd edition 1997, Corrected Printing 1999.
- [Schwenk1996] J. Schwenk
Conditional Access, in taschenbuch der telekom praxis 1996,
Hrgb. B. Seiler, Verlag Schiele und Schön, Berlin.
- [Tietze1973] H. Tietze,
Gelöste und ungelöste mathematische Probleme,
Verlag C. H. Beck 1959, 6th edition 1973.

Web links

1. GIMPS (Great Internet Mersenne-Prime Search)
www.mersenne.org is the home page of the GIMPS project,
<http://www.mersenne.org/prime.htm>
2. The Proth Search Page with the Windows program by Yves Gallot
<http://prothsearch.net/index.html>
3. Generalized Fermat Prime Search
<http://perso.wanadoo.fr/yves.gallot/primes/index.html>
4. Distributed Search for Fermat Number Dividers
<http://www.fermatsearch.org/>
5. At the University of Tennessee you will find extensive research results about prime numbers.
<http://www.utm.edu/>
6. The best overview about prime numbers is offered from my point of view by “The Prime Pages” from Chris Caldwell.
<http://www.utm.edu/research/primes>
7. Descriptions e.g. about prime number tests
<http://www.utm.edu/research/primes/mersenne.shtml>
<http://www.utm.edu/research/primes/prove/index.html>
8. Showing the n -th prime number
http://www.utm.edu/research/primes/notes/by_year.html
9. The supercomputer manufacturer SGI Cray Research not only employed brilliant mathematicians but also used the prime number tests as benchmarks for their machines.
http://reality.sgi.com/chongo/prime/prime_press.html
10. The Cunningham Project,
<http://www.cerias.purdue.edu/homes/ssw/cun/>
11. <http://www.eff.org/coop-awards/prime-release1.html>
12. <http://www.math.Princeton.EDU/~arbooker/nthprime.html>
13. <http://www.cerias.purdue.edu/homes/ssw/cun>
14. <http://www.informatik.uni-giessen.de/staff/richtstein/de/Goldbach.html>
15. <http://www.mathematik.ch/mathematiker/goedel.html>
16. <http://www.mscs.dal.ca/~dilcher/goldbach/index.html>

Acknowledgments

I would like to take this opportunity to thank Mr. Henrik Koy and Mr. Roger Oyono for their very constructive proof-reading of this article.

3 Introduction to Elementary Number Theory with Examples

(Bernhard Esslinger, July 2001, Updates: Dec. 2001, June 2002, May 2003)

This “introduction” is for people with a mathematical interest. There is no more pre-knowledge necessary than what you learn in the secondary school.

We intentionally had “beginners” in mind; we did not take the approach of mathematical textbooks, called “introduction”, which cannot be understood at the first reading further than page 5 and which have the real purpose to deliver all information that special monographs can be read.

3.1 Mathematics and cryptography

A large proportion of modern, asymmetric cryptography is based on mathematical knowledge – on the properties (“laws”) of whole numbers, which are investigated in elementary number theory. Here, the word “elementary” means that questions raised in number theory are essentially rooted in the set of natural and whole numbers.

Further mathematical disciplines currently used in cryptography include (see [Bauer1995, p. 2], [Bauer2000, p. 3]) :

- Group theory
- Combination theory
- Complexity theory
- Ergodic theory
- Information theory.

Number theory or arithmetic (the emphasis here is more on the aspect of performing calculations with numbers) was established by Carl Friedrich Gauss³⁶ as a special mathematical discipline. Its elementary features include the greatest common divisor³⁷ (gcd), congruence (remainder classes), factorisation, the Euler-Fermat theorem and primitive roots. However, the most important aspect is prime numbers and their multiplicative operation.

For a long time, number theory was considered to be the epitome of pure research, the ideal example of research in the ivory tower. It delved into “the mysterious laws of the realm of numbers”, giving rise to philosophical considerations as to whether it described elements that exist everywhere in nature or whether it artificially constructed elements (numbers, operators and properties).

³⁶Carl Friedrich Gauss, German mathematician and astronomer, Apr 30, 1777–Feb 23, 1855.

³⁷This article deals with the gcd (greatest common divisor) in [Appendix A of this chapter](#).

We now know that patterns from number theory can be found everywhere in nature. For example, the ratio of rotating counterclockwise and rotating clockwise spirals in a sunflower is equal to two consecutive Fibonacci numbers³⁸, for example 21 : 34.

Also, at the latest when number theory was applied in modern cryptography, it became clear that a discipline that had been regarded as purely theoretical for centuries actually had a practical use. Today, experts in this field are in great demand on the job market.

Applications in (computer) security now use cryptography because this mathematical discipline is simply better and easier to prove than all other "creative" substitution procedures that have been developed over the course of time and better than all sophisticated physical methods such as those used to print bank notes [Beutelspacher1996, p. 4].

This article explains the basics of elementary number theory in a way that you can easily understand. It provides numerous examples and very rarely goes into any proofs (these can be found in mathematical textbooks).

The goal is not to exhaustively explain the number theory findings, but to show the essential procedures. The volume of the content is so oriented that the reader can understand and apply the RSA method.

For this purpose we will use both theory and examples to explain how to perform calculations in finite sets and describe how these techniques are applied in cryptography. Particular attention will be paid to the traditional Diffie-Hellman (DH) and RSA public key procedures.

Additionally I added some qualified statements about the security of the RSA algorithm.

Carl Friedrich Gauss:

Mathematics is the queen of sciences and number theory is the queen of mathematics.

3.2 Introduction to number theory

Number theory arose from interest in positive whole numbers $1, 2, 3, 4, \dots$, also referred to as the set of natural numbers *natural numbers* \mathbb{N} . These are the first mathematical constructs used by human civilisation. According to Kronecker³⁹, they are a creation of God. In Dedekind's⁴⁰ opinion, they are a creation of the human intellect. Dependent upon one's ideology, this is an unsolvable contradiction or one and the same thing.

In ancient times, no distinction was made between number theory and numerology, which attributed a mystical significance to specific numbers. In the same way as astronomy and chem-

³⁸The sequence of Fibonacci numbers $(a_i)_{i \in \mathbb{N}}$ is defined by the "recursive" rule $a_1 := a_2 := 1$ and for all numbers $n = 1, 2, 3, \dots$ we define $a_{n+2} := a_{n+1} + a_n$. This historical sequence can be found in many interesting forms in nature (for example, see [Graham1994, p. 290 ff] or the website of Ron Knott, which is devoted to Fibonacci numbers). A lot is known about the Fibonacci sequence and it is used today as an important tool in mathematics.

³⁹Leopold Kronecker, German mathematician, Dec 7, 1823 – Dec 29, 1891

⁴⁰Julius Wilhelm Richard Dedekind, German mathematician, Oct 6, 1831 – Feb 12, 1916.

istry gradually detached themselves from astrology and alchemy during the Renaissance (from the 14th century), number theory also separated itself from numerology.

Number theory has always been a source of fascination – for both amateurs and professional mathematicians. In contrast to other areas of mathematics, many of the problems and theorems in number theory can be understood by non-experts. On the other hand, mathematicians often take a long time to find solutions to the problems or prove the theorems. It is therefore one thing to pose good questions but quite another matter to find the answer. One example of this is what is known as Fermat’s Last (or large) theorem⁴¹.

Up until the mid 20th century, number theory was considered to be the purest area of mathematics, an area that had no practical use in the real world. This changed with the development of computers and digital communication, as number theory was able to provide several unexpected solutions to real-life tasks. At the same time, advances in information technology allowed specialists in number theory to make huge progress in factorising large numbers, finding new prime numbers, testing (old) conjectures and solving numerical problems that were previously impossible to solve. Modern number theory is made up of areas such as:

- Elementary number theory
- Algebraic number theory
- Analytic number theory
- Geometric number theory
- Combinatorial number theory
- Numeric number theory
- Probability theory.

All of the different areas are concerned with questions regarding whole numbers (both positive and negative whole numbers plus zero). However, they each have different methods of dealing with them.

This article only deals with the area of elementary number theory.

⁴¹One of the things you learn in mathematics at school is Pythagoras theorem, which states the following for a right-angle triangle: $a^2 + b^2 = c^2$, where a and b are the lengths of the sides containing the right angle and c is the length of the hypotenuse. Fermat famously proposed that $a^n + b^n \neq c^n$ for whole-number exponents $n > 2$. Unfortunately, the letter in which Fermat made the claim did not have enough space for him to prove it. The theorem was not proved until over 300 years later [Wiles1994, p. 433-551].

3.2.1 Convention

Unless stated otherwise:

- The letters $a, b, c, d, e, k, n, m, p, q$ are used to present whole numbers.
- The letters i and j represent natural numbers.
- The letters p always represents a prime number.
- The sets $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ are the *natural* and *whole* numbers respectively.

Joanne K. Rowling⁴²

This isn't magic – it's logic – a puzzle. A lot of the greatest wizards haven't got an ounce of logic.

3.3 Prime numbers and the first fundamental theorem of elementary number theory

Many of the problems in elementary number theory are concerned with prime numbers.

Every whole number has divisors or factors. The number 1 has just one itself, whereas the number 12 has the six factors 1, 2, 3, 4, 6 and 12⁴³. Many numbers are only divisible by themselves and by 1. When it comes to multiplication, these can be regarded as the “atoms” in the realm of numbers.

Definition 3.1. *Prime numbers are natural numbers greater than 1 that can only be divided by 1 and themselves.*

By definition, 1 is not a prime number.

If we write down the prime numbers in ascending order (prime number sequence), then we get:

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, \dots

The first 100 numbers include precisely 25 prime numbers. After this, the percentage of primes decreases, but never reaches zero.

We come across whole numbers that are prime fairly often. In the last decade only, three years were prime: 1993, 1997 and 1999. If they were rare, cryptography would not be able to work with them to the extent it does.

Prime numbers can be factorised in a unique (“*trivial*”) way:

$$\begin{aligned}5 &= 1 * 5 \\17 &= 1 * 17 \\1,013 &= 1 * 1,013 \\1,296,409 &= 1 * 1,296,409.\end{aligned}$$

Definition 3.2. *Natural numbers greater than 1 that are not prime are called **composite numbers**. These have at least two factors other than 1.*

⁴²Joanne K. Rowling, “Harry Potter and the Philosopher’s Stone”, Bloomsbury, (c) 1997, chapter “Through the trapdoor”, p. 307, by Hermine.

⁴³Due to the fact that 12 has so many factors, this number – and multiples of this number – is often found in everyday life: the 12-hour scale on clocks, the 60 minutes in an hour, the 360-degree scale for measuring angles, etc. If we divide these scales into segments, the segments often turn out to be whole numbers. These are easier to use in mental arithmetic than fractions.

Examples of the decomposition of such numbers into prime factors:

$$\begin{aligned} 4 &= 2 * 2 \\ 6 &= 2 * 3 \\ 91 &= 7 * 13 \\ 161 &= 7 * 23 \\ 767 &= 13 * 59 \\ 1,029 &= 3 * 7^3 \\ 5,324 &= 22 * 11^3. \end{aligned}$$

Theorem 3.1. *Each composite number a has a lowest factor greater than 1. This factor is a prime number p and is less than or equal to the square root of a .*

All whole numbers greater than 1 can be expressed as a product of prime numbers — in a *unique* way.

This is the claim of the 1st *fundamental theorem of number theory* (= fundamental theorem of arithmetic = fundamental building block of all positive integers). This was formulated precisely for the first time by Carl Friedrich Gauss in his *Disquisitiones Arithmeticae* (1801).

Theorem 3.2. Gauss 1801 *Every even natural number greater than 1 can be written as the product of prime numbers. Given two such decompositions $a = p_1 * p_2 * \dots * p_n = q_1 * q_2 * \dots * q_m$, these can be resorted such that $n = m$ and, for all i , $p_i = q_i$.*

In other words: Each natural number other than 1 can be written as a product of prime numbers in precisely one way, if we ignore the order of the factors. The factors are therefore unique (the “expression as a product of factors” is unique)!

For example, $60 = 2 * 2 * 3 * 5 = 2^2 * 3 * 5$. And this — other than changing the order of the factors — is the only way in which the number 60 can be factorised.

If you allow numbers other than primes as factors, there are several ways of factorising integers and the *uniqueness* is lost:

$$60 = 1 * 60 = 2 * 30 = 4 * 15 = 5 * 12 = 6 * 10 = 2 * 3 * 10 = 2 * 5 * 6 = 3 * 4 * 5 = \dots$$

The 1st fundamental theorem only appears to be obvious. We can construct numerous other sets of numbers⁴⁴ for which numbers in the set *cannot* be expressed uniquely as a product of the prime numbers of the set.

In order to make a mathematical statement, therefore, it is important to state not only the operation for which it is defined but also the basic set on which the operation is defined.

For more details on prime numbers (e.g. how “Fermat’s Little Theorem” can be used to test extremely large numbers to determine whether they are prime), please refer to the article on [prime numbers, chapter 2](#) in this script.

⁴⁴These sets are formed especially from the set of natural numbers. An example of this can be found in this [script](#) on page [19](#) at the end of chapter [2.2](#).

3.4 Divisibility, modulus and remainder classes

If whole numbers are added, subtracted or multiplied, the result is always another whole number.

The division of two whole numbers does not always result in a whole number. For example, if we divide 158 by 10 the result is the decimal number 15.8, which is not a whole number!

If, however, we divide 158 by 2 the result 79 is a whole number. In number theory we express this by saying that 158 is *divisible* by 2 but not by 10. In general, we say:

Definition 3.3. *A whole number n is **divisible** by a whole number d if the quotient n/d is a whole number c such that $n = c * d$.*

n is called a *multiple* of d , whereas d is called a *divisor* or *factor* of n .

The mathematical notation for this is $d|n$ (read “ d divides n ”). The notation $d \nmid n$ means that d does not divide the number n .

In our example therefore: $10 \nmid 158$ but $2|158$.

3.4.1 The modulo operation – working with congruence

When we investigate divisibility, it is only the remainder of the division that is important. When dividing a number n by m , we often use the following notation:

$$\frac{n}{m} = c + \frac{r}{m},$$

where c is a whole number and r is a number with the values $0, 1, \dots, m-1$. This notation is called division with remainder, whereby c is called the whole-number “quotient” and r is the “remainder” of the division.

Example:

$$\frac{19}{7} = 2 + \frac{5}{7} \quad (m = 7, \quad c = 2, \quad r = 5)$$

What do the numbers 5, 12, 19, 26, \dots have in common for division by 7? The remainder is always $r = 5$. For division by 7, only the following remainders are possible:

$$r = 0, 1, 2, \dots, 6.$$

The numbers that result in the same remainder r when divided by 7 are combined to form the “remainder class r modulo 7”. Two numbers a and b belonging to the same remainder class modulo 7 are said to be “congruent modulo 7”. Or in general:

Definition 3.4. *The **remainder class r modulo m** is the set of all whole numbers a that have the same remainder r when divided by m .*

Examples:

Remainder class 0 modulo 4 =

$$\{x|x = 4 * n; n \in \mathbb{N}\} = \{\dots, -16, -12, -8, -4, 0, 4, 8, 12, 16, \dots\}$$

Remainder class 3 modulo 4 =

$$\{x|x = 4 * n + 3; n \in \mathbb{N}\} = \{\dots, -13, -9, -5, -1, 3, 7, 11, 15, \dots\}$$

As only the remainders $0, 1, 2, \dots, m-1$ are possible for division modulo m , modular arithmetic works with finite sets. For each modulo m there are precisely m remainder classes.

Definition 3.5. *Two numbers $a, b \in \mathbb{N}$ are said to be congruent modulo $m \in \mathbb{N}$ if and only if they have the same remainder when divided by m .*

We write: $a \equiv b \pmod{m}$ (read a is congruent b modulo m), which means that a and b belong to the same remainder class. The modulo is therefore the divisor. This notation was introduced by Gauss. Although the divisor is usually positive, a and b can also be any whole numbers.

Examples:

$19 \equiv 12 \pmod{7}$, because the remainders are equal: $19/7 = 2$ remainder 5 and $12/7 = 1$ remainder 5.

$23103 \equiv 0 \pmod{453}$, because $23103/453 = 51$ remainder 0 and $0/453 = 0$ remainder 0.

Theorem 3.3. *$a \equiv b \pmod{m}$ if and only if, the difference $(a - b)$ is divisible by m , i.e. if $q \in \mathbb{Z}$ exists with $(a - b) = q * m$.*

These two statements are therefore equivalent.

Therefore: If m divides the difference, there exists a whole number q such that: $a = b + q * m$. As an alternative to the congruence notation, we can also use the divisibility notation: $m|(a - b)$.

Example of equivalent statements:

$35 \equiv 11 \pmod{3} \iff 35 - 11 \equiv 0 \pmod{3}$, where $35 - 11 = 24$ is divisible by 3 without remainder while $35 : 3$ and $11 : 3$ leave the remainder 2.

Comment:

The above equivalence does not apply to the sum $(a + b)$!

Example:

$11 \equiv 2 \pmod{3}$, therefore $11 - 2 \equiv 9 \equiv 0 \pmod{3}$; but $11 + 2 = 13$ is not divisible by 3. The statement in theorem 3.3 does not even apply to sums in one direction. It is correct for sums only if the remainder is 0 and only in the following direction: if a divisor divides both summands with no remainder, it also divides the sum with no remainder.

We can apply the above equivalence in theorem 3.3 if we need a quick and easy method of determining whether large numbers are divisible by a certain number.

Example:

Is 69,993 divisible by 7?

The number can be written in the form of a difference in which it is clear that each operand is divisible by 7: $69,993 = 70,000 - 7$. Therefore, the difference is also divisible by 7.

Although these considerations and definitions may seem to be rather theoretical, we are so familiar with them in everyday life that we no longer think about the formal procedure. For example, the 24 hours on a clock are represented by the numbers $1, 2, \dots, 12$. We obtain the hours after 12 noon as the remainder of a division by 12 and know immediately that 2 o'clock in the afternoon is the same as 14.00.

This “modular” arithmetic (based on division remainders) forms the basis of asymmetric encryption procedures. Cryptographic calculations are therefore not based on real numbers, as the calculations you performed at school, but rather on character strings with a limited length, in other words on positive whole numbers that cannot exceed a certain value. This is one of the reasons why we choose a large number m and “calculate modulo m ”. That is, we ignore whole-number multiples of m and, rather than working with a number, we only work with the remainder when this number is divided by m . The result is that all results are in the range 0 to $m - 1$.

3.5 Calculations with finite sets

3.5.1 Laws of modular calculations

From algebra theorems it follows that essential parts of the conventional calculation rules are kept when we proceed to modular calculations over a basic set \mathbb{Z} . For example, addition remains commutative. The same goes for multiplication modulo m . The result of a division⁴⁵ is not a fraction but rather a whole number between 0 and $m - 1$.

The known laws apply:

1. Associative law:

$$\begin{aligned} ((a + b) + c) \pmod{m} &\equiv (a + (b + c)) \pmod{m}. \\ ((a * b) * c) \pmod{m} &\equiv (a * (b * c)) \pmod{m}. \end{aligned}$$

2. Commutative law:

$$\begin{aligned} (a + b) \pmod{m} &\equiv (b + a) \pmod{m}. \\ (a * b) \pmod{m} &\equiv (b * a) \pmod{m}. \end{aligned}$$

The associative law and commutative law apply to both addition and multiplication.

3. Distributive law:

$$(a * (b + c)) \pmod{m} \equiv (a * b + a * c) \pmod{m}.$$

4. Reducibility:

$$\begin{aligned} (a + b) \pmod{m} &\equiv (a \pmod{m} + b \pmod{m}) \pmod{m}. \\ (a * b) \pmod{m} &\equiv (a \pmod{m} * b \pmod{m}) \pmod{m}. \end{aligned}$$

⁴⁵When dividing modulo m we cannot use every number because some numbers have the same property as zero. See footnote 49 in chapter 3.6.1.

The order in which the modulo operation is performed is irrelevant.

5. Existence of an identity (neutral element):

$$(a + 0) \pmod{m} \equiv (0 + a) \pmod{m} \equiv a \pmod{m}.$$

$$(a * 1) \pmod{m} \equiv (1 * a) \pmod{m} \equiv a \pmod{m}.$$

6. Existence of an inverse element:

For all whole numbers a and m there exists a whole number $-a$ such that:

$$(a + (-a)) \pmod{m} \equiv 0 \pmod{m} \quad (\text{additive inverse}).$$

For each a ($a \not\equiv 0 \pmod{p}$) where p is prime there exists a whole number a^{-1} , such that:

$$(a * a^{-1}) \pmod{p} \equiv 1 \pmod{p} \quad (\text{multiplicative inverse}).$$

7. Closeness⁴⁶:

$$a, b \in G \implies (a + b) \in G.$$

$$a, b \in G \implies (a * b) \in G.$$

8. Transitivity:

$$[a \equiv b \pmod{m}, b \equiv c \pmod{m}] \implies [a \equiv c \pmod{m}].$$

3.5.2 Patterns and structures

In general mathematicians investigate “Structures”. They ask e.g. at $a * x \equiv b \pmod{m}$, which values x can take for given values of a , b , m .

Especially the case is investigated, where the result b of this operation is the neutral element. Then x is the inverse of a regarding this operation.

⁴⁶The property of closeness is always defined in relation to an operation in a set. See [Appendix B of this chapter](#).

*Seneca*⁴⁷:

The way of theory is long, it is short and effective by examples.

3.6 Examples of modular calculations

As we have already seen:

For two natural numbers a and m , $a \bmod m$ denotes the remainder obtained when we divide a by m . This means that $a \bmod m$ is always a number between 0 and $m - 1$.

For example, $1 \equiv 6 \equiv 41 \equiv 1 \pmod{5}$ because the remainder is always 1. Another example is: $2000 \equiv 0 \pmod{4}$ because 4 divides 2000 with no remainder.

Modular arithmetic only contains a limited quantity of non-negative numbers. The number of these is specified by a modulus m . If the modulo is $m = 5$, then only the 5 numbers in the set $\{0, 1, 2, 3, 4\}$ are used.

A calculation result larger than 4 is then reduced “modulo 5”. In other words, it is the remainder when the result is divided by 5. For example, $2 * 4 \equiv 8 \equiv 3 \pmod{5}$ because 3 is the remainder when we divide 8 by 5.

3.6.1 Addition and multiplication

The following shows

- the addition table⁴⁸ $\pmod{5}$ and
- the multiplication tables⁴⁹ for mod 5 and mod 6.

Example of an addition table

The result when we add 3 and 4 $\pmod{5}$ is calculated as follows: Calculate $3 + 4 = 7$ and keep subtracting 5 from the result until the result is less than the modulo: $7 - 5 = 2$. Therefore: $3 + 4 \equiv 2 \pmod{5}$.

⁴⁷Lucius Annaeus Seneca, philosophical writer and poet, 4 B. C. – 65 A. D.

⁴⁸Comment on subtraction modulo 5:

$2 - 4 \equiv -2 \equiv 3 \pmod{5}$.

It is therefore not true modulo 5 that $-2 = 2$ (see also [Appendix C of this chapter](#)).

⁴⁹Comment on division modulo 6:

Due to the special role of zero as the identity for addition, division by zero is not permitted:

for all a it is $a * 0 = 0$, because $a * 0 = a * (0 + 0) = a * 0 + a * 0$. Obviously 0 has no inverse regarding the multiplication, because if there would be one, it must be $0 = 0 * 0^{-1} = 1$. Also see footnote [45](#).

Addition table modulo 5:	+	0	1	2	3	4
	0	0	1	2	3	4
	1	1	2	3	4	0
	2	2	3	4	0	1
	3	3	4	0	1	2
	4	4	0	1	2	3

Example of an multiplication table:

The result of the multiplication $4 * 4 \pmod{5}$ is calculated as follows: $4 * 4 = 16$ and subtract 5 until the result is less than the modulus.

$$16 - 5 = 11; 11 - 5 = 6; 6 - 5 = 1.$$

The table directly shows that $4 * 4 \equiv 1 \pmod{5}$ because $16 : 5 = 3$ remainder 1. Multiplication is defined on the set \mathbb{Z} excluding 0.

Multiplication table modulo 5:	*	1	2	3	4
	1	1	2	3	4
	2	2	4	1	3
	3	3	1	4	2
	4	4	3	2	1

3.6.2 Additive and multiplicative inverses

You can use the tables to read the inverses for each number in relation to addition and multiplication.

The inverse of a number is the number that gives the result 0 when the two numbers are added and 1 when they are multiplied. Thus, the inverse of 4 for addition mod 5 is 1 and the inverse of 4 for multiplication mod 5 is 4 itself, because

$$4 + 1 = 5 \equiv 0 \pmod{5};$$

$$4 * 4 = 16 \equiv 1 \pmod{5}.$$

The inverse of 1 for multiplication mod 5 is 1, while the inverse modulo 5 of 2 is 3 and, since multiplication is commutative, the inverse of 3 is again 2.

If we take a random number and add or multiply another number (here 4) and then add⁵⁰ or multiply the corresponding inverse (1 or 4) to the interim result (1 or 3), then the end result is the same as the initial value.

Example:

$$2 + 4 \equiv 6 \equiv 1 \pmod{5}; \quad 1 + 1 \equiv 2 \pmod{5},$$

$$2 * 4 \equiv 8 \equiv 3 \pmod{5}; \quad 3 * 4 \equiv 12 \equiv 2 \pmod{5}.$$

⁵⁰In general $x + y + (-y) \equiv x \pmod{m}$ [$(-y) =$ additive inverse of $y \pmod{m}$].

In the set $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ for the addition and in the set \mathbb{Z}_5^* for the multiplication, all numbers have a **unique** inverse modulo 5.

In the case of modular addition, this is true for every modulo (not just for 5).

This is not the case, however, for modular multiplication.

Theorem 3.4. *A natural number a from the set $\{1, \dots, m-1\}$ has one inverse if and only if it and the modulo m are co-prime⁵¹, in other words if a and m have no common prime factors.*

Since $m = 5$ is prime, the numbers 1 to 4 are relatively prime to 5 and **each** of these numbers has a multiplicative inverse in mod 5.

A counterexample shows the multiplication table for mod 6 (since the modulus $m = 6$ is not prime, not all elements from $\mathbb{Z}_6 \setminus \{0\}$ are relatively prime to 6):

Multiplication table modulo 6:		*	1	2	3	4	5
	1		1	2	3	4	5
	2		2	4	0	2	4
	3		3	0	3	0	3
	4		4	2	0	4	2
	5		5	4	3	2	1

In addition to 0, the numbers 2, 3 and 4 also have no unique inverse (we can also say they have **no** inverse, because the elementary property of an inverse is uniqueness).

The numbers 2, 3 and 4 have the factor 2 or 3 in common with the modulus 6. Only the numbers 1 and 5, which are relatively prime to 6, have multiplicative inverses, namely themselves.

The number of numbers that are relatively prime to the modulus m is the same as the number of numbers that have a multiplicative inverse (see the **Euler function** $J(m)$ below).

For the two moduli 5 and 6 used in the multiplication tables, this means: the modulus 5 is a prime number itself. In mod 5, therefore, there are exactly $J(5) = 5 - 1 = 4$ numbers that are relatively prime to the modulus, that is all numbers from 1 to 4.

Since 6 is not a prime number, we write it as a product of its factors: $6 = 2 * 3$. In mod 6, therefore, there are exactly $J(6) = (2 - 1) * (3 - 1) = 1 * 2 = 2$ numbers that have a multiplicative inverse, that is 1 and 5.

Although it may seem difficult to calculate the table of multiplicative inverses for large moduli (this only applies to the areas of the table shaded dark grey), we can use Fermat's Little Theorem

⁵¹Two whole numbers a and b are co-prime if and only if $\gcd(a, b) = 1$.

If p is prime and a is a random whole number that is not a multiple of p , then p and a are co-prime.

Further name to the topic co-prime (with $a_i \in \mathbb{Z}, i = 1, \dots, n$):

1. a_1, a_2, \dots, a_n are *relatively prime*, if $\gcd(a_1, \dots, a_n) = 1$.

2. An even stronger request for more than two numbers is :

a_1, \dots, a_n are *in pairs relatively prime*, if for all $i = 1, \dots, n$ and $j = 1, \dots, n$ with $i \neq j$: $\gcd(a_i, a_j) = 1$.

Example: 2, 3, 6 are relatively prime, because $\gcd(2, 3, 6) = 1$. They are not in pairs relatively prime, because $\gcd(2, 6) = 2 > 1$.

to create a simple algorithm for this [Pfleeger1997, p. 80]. Quicker algorithms are described, for instance, in [Knuth1998]⁵².

Cryptographically not only the unique nature of the inverse is important, but also that the set of possible values has been exhausted.

Theorem 3.5. *For $a, i \in \{1, \dots, m-1\}$ with $\gcd(a, m) = 1$, then the product $a * i \bmod m$ takes for a certain number a all values from $\{1, \dots, m-1\}$ (exhaustive permutation of the length $m-1$)⁵³.*

The following three examples⁵⁴ illustrate the properties of multiplicative inverses.

In the multiplication table mod 17, the following was calculated for $i = 1, 2, \dots, 18$:

$$(5 * i)/17 = a \text{ remainder } r \text{ and high-lighted } 5 * i \equiv 1 \pmod{17},$$

$$(6 * i)/17 = a \text{ remainder } r \text{ and high-lighted } 6 * i \equiv 1 \pmod{17}.$$

We need to **find** the i for which the product remainder $a * i$ modulo 17 with $a = 5$ or $a = 6$ has the value 1.

Table 1: Multiplication table modulo 17 (for $a = 5$ and $a = 6$)

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$5 * i$	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90
remainder	5	10	15	3	8	13	1	6	11	16	4	9	14	2	7	12	0	5
$6 * i$	6	12	18	24	30	36	42	48	54	60	66	72	78	84	90	96	102	108
remainder	6	12	1	7	13	2	8	14	3	9	15	4	10	16	5	11	0	6

Between $i = 1, \dots, m$, all values between $0, \dots, m-1$ occur for the remainders, because both 5 and 6 are also relatively prime to the modulus $m = 17$.

The multiplicative inverse of 5 (mod 17) is 7, while the inverse of 6 (mod 17) is 3.

Table 2: Multiplication table modulo 13 (for $a = 5$ and $a = 6$)

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$5 * i$	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90
remainder	5	10	2	7	12	4	9	1	6	11	3	8	0	5	10	2	7	12
$6 * i$	6	12	18	24	30	36	42	48	54	60	66	72	78	84	90	96	102	108
remainder	6	12	5	11	4	10	3	9	2	8	1	7	0	6	12	5	11	4

⁵²Using Euclid's extended theorem (extended gcd), we can calculate the multiplicative inverse and determine whether numbers have an inverse (see [appendix A of this chapter](#)). Alternatively, we can also use the primitive roots.

⁵³See also theorem 3.14 in [chapter 3.9, Multiplicative order and primitive roots](#).

⁵⁴See [Appendix E of this chapter](#) for the source code to compute the tables using Mathematica and Pari-GP.

Between $i = 1, \dots, m$, all values between $0, \dots, m - 1$ occur for the remainders, because both 5 and 6 are relatively prime to the modulus $m = 13$.

The multiplicative inverse of 5 (mod 13) is 8, while the inverse of 6 (mod 13) is 11.

The following table contains an example, where the modulus m and the number $a = 6$ are *not* relatively prime.

Table 3: Multiplication table modulo 12 (for $a = 5$ and $a = 6$)

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$5 * i$	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90
remainder	5	10	3	8	1	6	11	4	9	2	7	0	5	10	3	8	1	6
$6 * i$	6	12	18	24	30	36	42	48	54	60	66	72	78	84	90	96	102	108
remainder	6	0	6	0	6	0	6	0	6	0	6	0	6	0	6	0	6	0

We have calculated $(5 * i) \pmod{12}$ and $(6 * i) \pmod{12}$. Between $i = 1, \dots, m$, not all values between $0, \dots, m - 1$ occur and 6 does not have an inverse mod 12, because 6 and the modulus $m = 12$ are not co-prime.

The multiplicative inverse of 5 (mod 12) is 5. The number 6 has no inverse (mod 12).

3.6.3 Raising to the power

In modular arithmetic, raising to the power is defined as repeated multiplication – as usual except that multiplication is now slightly different. We can even apply the usual rules, such as:

$$\begin{aligned} a^{b+c} &= a^b * a^c, \\ (a^b)^c &= a^{b*c} = a^{c*b} = (a^c)^b. \end{aligned}$$

Modular powers work in the same way as modular addition and modular multiplication:

$$3^2 \equiv 9 \equiv 4 \pmod{5}.$$

Even consecutive powers work in the same way:

Example 1:

$$(4^3)^2 \equiv 64^2 \equiv 4096 \equiv 1 \pmod{5}.$$

- (1) We can speed up⁵⁵ the calculation by reducing the **interim results** modulo 5 but we need to take care because not everything will then work in the same way as

⁵⁵The time required to calculate the multiplication of two numbers normally depends on the length of the numbers. We can observe this if we use the school method to calculate, for instance, $474 * 228$. The time required increases in a quadratic square manner, because we need to multiply $3 * 3$ numbers. The numbers become considerably smaller if we reduce the interim result.

in standard arithmetic.

$$\begin{aligned}
(4^3)^2 &\equiv (4^3 \pmod{5})^2 \pmod{5} \\
&\equiv (64 \pmod{5})^2 \pmod{5} \\
&\equiv 4^2 \pmod{5} \\
&\equiv 16 \equiv 1 \pmod{5}.
\end{aligned}$$

(2) In standard arithmetic, consecutive powers can be reduced to a single power by multiplying the exponents:

$$(4^3)^2 = 4^{3*2} = 4^6 = 4096.$$

This is not quite as simple in modular arithmetic because this would give:

$$(4^3)^2 \equiv 4^{3*2 \pmod{5}} \equiv 4^6 \pmod{5} \equiv 4^1 \equiv 4 \pmod{5}.$$

But as we saw above, the correct result is 1!!

(3) Therefore, the rule is slightly different for consecutive powers in modular arithmetic: we do not multiply the exponents in (\pmod{m}) but rather in $(\pmod{J(m)})$.

Using $J(5) = 4$ gives:

$$(4^3)^2 \equiv 4^{3*2 \pmod{J(5)}} \equiv 4^6 \pmod{4} \equiv 4^2 \equiv 16 \equiv 1 \pmod{5}.$$

This delivers the correct result.

Theorem 3.6. $(a^b)^c \equiv a^{b*c \pmod{J(m)}} \pmod{m}.$

Example 2:

$$3^{28} \equiv 3^{4*7} \equiv 3^{4*7 \pmod{10}} \equiv 3^8 \equiv 6561 \equiv 5 \pmod{11}.$$

3.6.4 Fast calculation of high powers

RSA encryption and decryption⁵⁶ entails calculating high powers modulo m . For example, the calculation $(100^5) \pmod{3}$ exceeds the 32-bit long integer number range provided we calculate a^n by actually multiplying a with itself n times in line with the definition. In the case of extremely large numbers, even a fast computer chip would take longer than the age of the universe to calculate a single exponential. Luckily, there is an extremely effective shortcut for calculating exponentials (but not for calculating logarithms).

If the expression is divided differently using the rules of modular arithmetic, then the calculation does not even exceed the 16-bit short integer number range:

$$(a^5) \equiv (((a^2 \pmod{m})^2 \pmod{m}) * a) \pmod{m}.$$

⁵⁶See chapter 3.10 (Proof of the RSA procedure with Euler-Fermat) and chapter 3.13 (The RSA procedure with actual numbers).

We can generalise this by representing the exponent as a binary number. For example, the naive method would require 36 multiplications in order to calculate a^n for $n = 37$. However, if we write n in the binary representation as $100101 = 1 * 2^5 + 1 * 2^2 + 1 * 2^0$, then we can rewrite the expression as: $a^{37} = a^{2^5+2^2+2^0} = a^{2^5} * a^{2^2} * a^1$

Example 3: $87^{43} \pmod{103}$.

Since $43 = 32 + 8 + 2 + 1$, 103 is prime, $43 < J(103)$

and the squares $\pmod{103}$ can be calculated beforehand

$$\begin{aligned} 87^2 &\equiv 50 \pmod{103}, \\ 87^4 &\equiv 50^2 \equiv 28 \pmod{103}, \\ 87^8 &\equiv 28^2 \equiv 63 \pmod{103}, \\ 87^{16} &\equiv 63^2 \equiv 55 \pmod{103}, \\ 87^{32} &\equiv 55^2 \equiv 38 \pmod{103}. \end{aligned}$$

we have⁵⁷:

$$\begin{aligned} 87^{43} &\equiv 87^{32+8+2+1} \pmod{103} \\ &\equiv 87^{32} * 87^8 * 87^2 * 87 \pmod{103} \\ &\equiv 38 * 63 * 50 * 87 \equiv 85 \pmod{103}. \end{aligned}$$

The powers $(a^2)^k$ can be determined easily by means of repeated squaring. As long as a does not change, a computer can calculate them beforehand and – if enough memory is available – save them. In order to then find a^n in each individual case, it now only needs to multiply those $(a^2)^k$ for which there is a one in the k -th position of the binary representation of n . The typical effort is then reduced from 2^{600} to $2 * 600$ multiplications! This frequently used algorithm is called “Square and Multiply”.

3.6.5 Roots and logarithms

The inverses of the powers are also defined. The roots and logarithms are again whole numbers. Yet in contrast to the usual situation, they are not only difficult to calculate but, in the case of large numbers, cannot be calculated at all within a reasonable amount of time.

Let us take the equation $a \equiv b^c \pmod{m}$.

a) Taking the logarithm (determining c) – Discrete logarithm problem :

If we know a and b of the three numbers a , b and c that meet this equation, then every known method of finding c is approximately just as time-consuming as trying out all m possible values for c one after the other. For a typical m of the order of magnitude of 10^{180} for 600-digit binary numbers, this is a hopeless task. More precisely, for suitably

⁵⁷See [Appendix E of this chapter](#) for source code implementing the square and multiply method in Mathematica and Pari-GP, which can be used to reproduce the calculations above.

large numbers m , the time required according to current knowledge is proportional to $\exp(C * (\log m [\log \log m]^2)^{1/3})$ with a constant $C > 1$.

b) Calculating the root (determining b):

The situation is similar if b is the unknown variable and we know the values of a and c :
If we know the Euler function of m , $J(m)$, then we can easily⁵⁸ calculate d with $c * d \equiv 1 \pmod{J(m)}$ and use theorem 3.6 to obtain:

$$a^d \equiv (b^c)^d \equiv b^{c*d} \equiv b^{c*d \pmod{J(m)}} \equiv b^1 \equiv b \pmod{m}$$

the c -th root b of a .

If $J(m)$ cannot be determined⁵⁹, it is difficult to calculate the c -th root. This forms the basis for the security assumption used by the RSA encryption system (see Sub-section 4.3.1: [the RSA procedure](#) and [Sub-section 3.10 Proof of the RSA procedure](#)).

The time required for inverting addition and multiplication, on the other hand, is simply proportional to $\log m$ or $(\log m)^2$. Powers (for a number x calculate x^a with a fixed) and exponents (for a number x calculate a^x with a fixed) are therefore typical one way functions (See Overview of the one way functions in this [Script](#) and [article](#)).

3.7 Groups and modular arithmetic in \mathbb{Z}_n and \mathbb{Z}_n^*

Mathematical “groups” play a decisive role in number theory and cryptography. We only talk of groups if, for a defined set and a defined relation (an operation such as addition or multiplication), the following properties are fulfilled:

- The set is closed
- A neutral element exists
- An inverse element exists for each element
- The associative law applies.

The abbreviated mathematical notation is $(G, +)$ or $(G, *)$.

Definition 3.6. \mathbb{Z}_n :

\mathbb{Z}_n comprises all numbers from 0 to $n - 1$: $\mathbb{Z}_n = \{0, 1, 2, \dots, n - 2, n - 1\}$.

\mathbb{Z}_n is an often used finite group of the natural numbers. It is sometimes also called the *remainder set R modulo n* .

For example, 32-bit computers (standard PCs) only directly work with whole numbers in a finite set, that is the value range $0, 1, 2, \dots, 2^{32} - 1$.

This value range is equivalent to the set $\mathbb{Z}_{2^{32}}$.

⁵⁸See [Appendix A of this chapter](#): the greatest common divisor (gcd) of whole numbers.

⁵⁹According to the first fundamental theorem of number theory and theorem 3.11, we determine $J(m)$ by reducing m to prime factors.

3.7.1 Addition in a group

If we define the operation $\text{mod}+$ on such a set where

$$a \text{ mod } + b := (a + b) \pmod{n},$$

then the set \mathbb{Z}_n together with the relation $\text{mod}+$ is a group because the following properties of a group are valid for all elements in \mathbb{Z}_n :

- $a \text{ mod } + b$ is an element of \mathbb{Z}_n (the set is closed),
- $(a \text{ mod } + b) \text{ mod } + c \equiv a \text{ mod } + (b \text{ mod } + c) \pmod{n}$ ($\text{mod}+$ is associative),
- the neutral element is 0.
- each element $a \in \mathbb{Z}_n$ has an inverse for this operation, namely $n - a$ (because $a \text{ mod } + (n - a) \equiv a + (n - a) \pmod{n} \equiv n \equiv 0 \pmod{n}$).

Since the operation is commutative, i.e. $(a \text{ mod } + b) = (b \text{ mod } + a)$, this structure is actually a “commutative group”.

3.7.2 Multiplication in a group

If we define the operation mod^* on the set \mathbb{Z}_n where

$$a \text{ mod } * b := (a * b) \pmod{n},$$

then \mathbb{Z}_n together with this operation is **usually not a group** because not all properties are fulfilled for each n .

Examples:

- In \mathbb{Z}_{15} , for example, the element 5 does not have an inverse. That is to say, there is no a with $5 * a \equiv 1 \pmod{15}$. Each modulo product with 5 on this set gives 5, 10 or 0.
- In $\mathbb{Z}_{55} \setminus \{0\}$, for example, the elements 5 and 11 do not have multiplicative inverses. That is to say, there is no $a \in \mathbb{Z}_{55}$ such that $5 * a \equiv 1 \pmod{55}$ and no a such that $11 * a \equiv 1 \pmod{55}$. This is because 5 and 11 are not relatively prime to 55. Each modulo product with 5 on this set gives 5, 10, 15, ..., 50 or 0. Each modulo product with 11 on this set gives 11, 22, 33, 44 or 0.

On the other hand, there are subsets of \mathbb{Z}_n that form a group with the operation mod^* . If we choose all elements in \mathbb{Z}_n that are relatively prime to n , then this set forms a group with the operation mod^* . We call this set \mathbb{Z}_n^* .

Definition 3.7. \mathbb{Z}_n^* :

$$\mathbb{Z}_n^* = \{a \in \mathbb{Z}_n \mid \gcd(a, n) = 1\}.$$

\mathbb{Z}_n^* is sometimes also called the reduced remainder set R' modulo n .

Example: For $n = 10 = 2 * 5$ the following applies:

full remainder set $R = \mathbb{Z}_n = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

reduced remainder set $R' = \mathbb{Z}_n^* = \{1, 3, 7, 9\} \longrightarrow J(n) = 4$.

Comment: R' or \mathbb{Z}_n^* is always a genuine subset of R or \mathbb{Z}_n because 0 is always an element of R but never an element of R' . Since 1 and $n - 1$ are always relatively prime to n , they are always elements of both sets.

If we select a random element in \mathbb{Z}_n^* and multiply it by every other element in \mathbb{Z}_n^* , then the products⁶⁰ are all in \mathbb{Z}_n^* , and the results are also a unique permutation of the elements in \mathbb{Z}_n^* . Since 1 is always an element of \mathbb{Z}_n^* , there is a unique “partner” in this set such that the product is 1. In other words:

Theorem 3.7. *Each element in \mathbb{Z}_n^* has a multiplicative inverse.*

Example for $a = 3$ modulo 10 with $\mathbb{Z}_n^* = \{1, 3, 7, 9\}$:

$$\begin{aligned} 3 &\equiv 3 * 1 \pmod{10}, \\ 9 &\equiv 3 * 3 \pmod{10}, \\ 1 &\equiv 3 * 7 \pmod{10}, \\ 7 &\equiv 3 * 9 \pmod{10}. \end{aligned}$$

The unique invertibility is an essential condition for cryptography (see section 3.10).

⁶⁰This is due to the fact that \mathbb{Z}_n^* is closed with respect to the multiplication and due to the gcd property:
 $[a, b \in \mathbb{Z}_n^*] \Rightarrow [((a * b) \pmod{n}) \in \mathbb{Z}_n^*]$, exactly:
 $[a, b \in \mathbb{Z}_n^*] \Rightarrow [\gcd(a, n) = 1, \gcd(b, n) = 1] \Rightarrow [\gcd(a * b, n) = 1] \Rightarrow [((a * b) \pmod{n}) \in \mathbb{Z}_n^*]$.

Eric Berne⁶¹:

Mathematical game theory postulates players who respond rationally. Transactional game theory, on the other hand, deals with games that are not rational, perhaps even **irrational and thereby closer to reality**.

3.8 Euler function, Fermat's little theorem and Euler-Fermat

3.8.1 Patterns and structures

As mathematicians investigate the structure $a * x \equiv b \pmod{m}$ (see [chapter 3.5.2](#)), so they are interested in the structure $x^a \equiv b \pmod{m}$.

Again here they are interested in the case, if $b = 1$ (value of the multiplicative inverse) and if $b = x$ (the function has a fixpoint).

3.8.2 The Euler function

Given n , the number of numbers from the set $\{1, \dots, n-1\}$ that are relatively prime to n is equal to the value of the Euler⁶² function $J(n)$.

Definition 3.8. The Euler function⁶³ $J(n)$ specifies the number of elements in \mathbb{Z}_n^* .

$J(n)$ also specifies how many whole numbers have multiplicative inverses in \pmod{n} . $J(n)$ can be calculated if we know the prime factors of n .

Theorem 3.8. For a prime number, the following is true: $J(p) = p - 1$.

Theorem 3.9. If m is the product of two distinct primes, then:

$$J(p * q) = (p - 1) * (q - 1) \quad \text{or} \quad J(p * q) = J(p) * J(q).$$

This case is important for the RSA procedure.

Theorem 3.10. If $n = p_1 * p_2 * \dots * p_k$ where p_1 to p_k are distinct prime numbers (i.e. no factor occurs more than once), then the following is true (as a generalisation of theorem [3.9](#)):

$$J(n) = (p_1 - 1) * (p_2 - 1) * \dots * (p_k - 1).$$

Theorem 3.11. In general, the following is true for every prime number p and every n in \mathbb{N} :

$$1. \quad J(p^n) = p^{n-1} * (p - 1).$$

⁶¹Eric Berne, "Games People Play", rororo, (c) 1964, page 235.

⁶²Leonhard Euler, Swiss mathematician, Apr 15, 1707 – Sep 18, 1783

⁶³Often written as the Euler phi function $\Phi(n)$.

2. If $n = p_1^{e_1} * p_2^{e_2} * \dots * p_k^{e_k}$, where p_1 to p_k are distinct prime numbers, then:

$$J(n) = [(p_1^{e_1-1}) * (p_1 - 1)] * \dots * [(p_k^{e_k-1}) * (p_k - 1)] = n * ([(p_1 - 1)/p_1] * \dots * [(p_k - 1)/p_k]).$$

Examples:

- $n = 70 = 2 * 5 * 7 \implies$ using theorem 3.10: $J(n) = 1 * 4 * 6 = 24$.
- $n = 9 = 3^2 \implies$ using theorem 3.11: $J(n) = 3^1 * 2 = 6$, because $\mathbb{Z}_9^* = \{1, 2, 4, 5, 7, 8\}$.
- $n = 2,701,125 = 3^2 * 5^3 * 7^4 \implies$ using theorem 3.11:

$$J(n) = [3^1 * 2] * [5^2 * 4] * [7^3 * 6] = 1,234,800.$$

3.8.3 The theorem of Euler-Fermat

In order to prove the RSA procedure, we need Fermat's theorem and its generalisation (Euler-Fermat theorem) – please see chapter 2.5.

Theorem 3.12. Fermat's Little Theorem⁶⁴ Let p be a prime number and a be a random whole number, then:

$$a^p \equiv a \pmod{p}.$$

An alternative formulation of Fermat's Little Theorem is as follows: Let p be a prime number and a be a random whole number that is relatively prime to p , then:

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

Theorem 3.13. Euler-Fermat theorem (generalisation of Fermat's Little Theorem)

For all elements a in the group \mathbb{Z}_n^* (i.e. a and n are natural numbers that are co-prime):

$$a^{J(n)} \equiv 1 \pmod{n}.$$

This theorem states that if we raise a group element (here a) to the power of the order of the group (here $J(n)$), we always obtain the neutral element for multiplication (the number 1).

The 2nd formulation of Fermat's Little Theorem is derived directly from Euler's theorem if n is a prime number.

If n is the product of two prime numbers, we can - in certain cases - use Euler's theorem to calculate the result of a modular power very quickly. We have: $a^{(p-1)*(q-1)} \equiv 1 \pmod{pq}$.

Examples for calculating a modular power:

- With $2 = 1 * 2$ and $6 = 2 * 3$ where 2 and 3 are both prime; $J(6) = 2$ because only 1 and 5 are relatively prime to 6, we obtain the equation $5^2 \equiv 5^{J(6)} \equiv 1 \pmod{6}$, without having to calculate the power.
- With $792 = 22 * 36$ and $23 * 37 = 851$ where 23 and 37 are both prime, it follows that $31^{792} \equiv 31^{J(23*37)} \equiv 31^{J(851)} \equiv 1 \pmod{851}$.

⁶⁴Pierre de Fermat, French mathematician, Aug 17, 1601 – Jan 12, 1665.

3.8.4 Calculation of the multiplicative inverse

Another interesting application is a special case of determining the multiplicative inverses using the Euler-Fermat theorem (multiplicative inverses are otherwise determined using the extended Euclidean algorithm).

Example:

Find the multiplicative inverse of 1579 modulo 7351.

According to Euler-Fermat: $a^{J(n)} \equiv 1 \pmod{n}$ for all a in \mathbb{Z}_n^* . If we divide both sides by a , we get: $a^{J(n)-1} \equiv a^{-1} \pmod{n}$. For the special case that the modulo is prime, we have $J(n) = p-1$. Therefore, the modular inverse is

$$a^{-1} = a^{J(n)-1} \equiv a^{(p-1)-1} \equiv a^{p-2} \pmod{p}.$$

For our example, this means:

$$\begin{aligned} \text{Since the modulus 7351 is prime, } p-2 &= 7349. \\ 1579^{-1} &\equiv 1579^{7349} \pmod{p}. \end{aligned}$$

By cleverly breaking down the exponent, we can calculate this power relatively easily (see Section 3.6.4 Fast calculation of high powers):

$$\begin{aligned} 7349 &= 4096 + 2048 + 1024 + 128 + 32 + 16 + 4 + 1 \\ 1579^{-1} &\equiv 4716 \pmod{7351}. \end{aligned}$$

3.8.5 Fixpoints modulo 26

According to theorem 3.6, the arithmetic operations of modular expressions are performed in the exponents modulo $J(n)$ rather than modulo n ⁶⁵.

In $a^{e*d} \equiv a^1 \pmod{n}$, if we wish to determine the inverses for the factor e in the exponent, we need to calculate modulo $J(n)$.

Example (with reference to the RSA algorithm):

If we calculate modulo 26, which set can e and d come from?

Solution: we have $e * d \equiv 1 \pmod{J(26)}$.

The reduced remainder set $R' = \mathbb{Z}_{26}^* = \{1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25\}$ are the elements in \mathbb{Z}_{26} , which have a multiplicative inverse, that is which are relatively prime to 26.

The reduced remainder set R'' contains only the elements of R' that are relatively prime to $J(n) = 12$: $R'' = \{1, 5, 7, 11\}$.

For every e in R'' there exists a d in R'' such that $a \equiv (a^e)^d \pmod{n}$.

⁶⁵For the following example, we will adopt the usual practice for the RSA procedure of using “ n ” rather than “ m ” to denote the modulus.

For every e in R'' , there exists therefore precisely one element (not necessarily different from e) such that $e * d \equiv 1 \pmod{J(26)}$.

For all e that are relatively prime to $J(n)$ we could calculate d as follows using the Euler-Fermat theorem: For $a^{J(n)} \equiv 1 \pmod{n}$ is the same as saying $a^{J(n)-1} \equiv a^{-1} \pmod{n}$. Therefore

$$d \equiv e^{-1} \pmod{J(n)} \equiv e^{J(n)-1} \pmod{J(n)}.$$

The problems of factorising $n = pq$ with $q \neq p$ and finding $J(n)$ have a similar degree of difficulty and if we find a solution for one of the two problems, we also have a solution for the other⁶⁶ (please compare requisition 3 in [section 3.10.1](#)).

⁶⁶If we know the factors of $n = p * q$ with $p \neq q$, then $J(n) = (p - 1) * (q - 1) = n - (p + q) + 1$. Additionally the factors p and q are solutions of the quadratic equation $x^2 - (p + q)x + pq = 0$.

If only n and $J(n)$ are known, then it is: $pq = n$ and $p + q = n - J(n) + 1$. So you get p and q by solving the equation

$$x^2 + (J(n) - n - 1)x + n = 0.$$

3.9 Multiplicative order and primitive roots

Mathematicians often ask, in which conditions the repeated application of an operation results in the neutral element (compare patterns and structures before).

For the i -times successive modular multiplication of a number a with $i = 1, \dots, m-1$ the product is the neutral element of the multiplication (1) if and only if a and m are relatively prime. The value of i , for which the product $a^i = 1$, is called the multiplicative order of a .

The multiplicative order and the primitive root are two useful constructs (concepts) in elementary number theory.

Definition 3.9. The **multiplicative order** $\text{ord}_m(a)$ of a whole number $a \pmod{m}$ (where a and m are co-prime) is the smallest whole number e for which $a^e \equiv 1 \pmod{m}$.

The following table shows that in a multiplicative group (here \mathbb{Z}_{11}^*) not all numbers necessarily have the same order. The orders in this case are 1, 2, 5 and 10 and we notice that:

1. The orders are all factors of 10.
2. The numbers $a = 2, 6, 7$ and 8 have the order 10 - we say that these numbers have the **maximum order** in \mathbb{Z}_{11}^* .

Example 1:

The following table⁶⁷ shows the values $a^i \pmod{11}$ for the exponents $i = 1, 2, \dots, 10$ and for the bases $a = 1, 2, \dots, 10$ as well as the resulting value $\text{ord}_{11}(a)$ for each a :

Table 4: Values of $a^i \pmod{11}$, $1 \leq a, i < 11$ and according order of $a \pmod{m}$:

	i=1	i=2	i=3	i=4	i=5	i=6	i=7	i=8	i=9	i=10	$\text{ord}_{11}(a)$
$a = 1$	1	1	1	1	1	1	1	1	1	1	1
$a = 2$	2	4	8	5	10	9	7	3	6	1	10
$a = 3$	3	9	5	4	1	3	9	5	4	1	5
$a = 4$	4	5	9	3	1	4	5	9	3	1	5
$a = 5$	5	3	4	9	1	5	3	4	9	1	5
$a = 6$	6	3	7	9	10	5	8	4	2	1	10
$a = 7$	7	5	2	3	10	4	6	9	8	1	10
$a = 8$	8	9	6	4	10	3	2	5	7	1	10
$a = 9$	9	4	3	5	1	9	4	3	5	1	5
$a = 10$	10	1	10	1	10	1	10	1	10	1	2

The table shows, for example, that the order of 3 modulo 11 has the value 5.

Definition 3.10. If a and m are co-prime and if $\text{ord}_m(a) = J(m)$ (i.e. a has maximum order), then we say that a is a **primitive root** of m .

⁶⁷See [Appendix E of this chapter](#) for the source code to generate the table using Mathematica and Pari-GP.

A number a is not a primitive root for every modulo m . In the above table, only $a = 2, 6, 7$ and 8 is a primitive root with respect to mod 11 ($J(11) = 10$).

Using the primitive roots, we can clearly establish the conditions for which powers modulo m have a unique inverse and the calculation in the exponents is manageable.

The following two tables show the multiplicative orders and primitive roots modulo 45 and modulo 46.

Example 2:

The following table⁶⁸ shows the values $a^i \bmod 45$ for the exponents $i = 1, 2, \dots, 12$ and for the bases $a = 1, 2, \dots, 12$ as well as the resulting value $\text{ord}_{45}(a)$ for each a :

Table 5: Values of $a^i \bmod 45, 1 \leq a, i < 13$:

$a \setminus i$	1	2	3	4	5	6	7	8	9	10	11	12	$\text{ord}_{45}(a)$	$J(45)$
1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
2	2	4	8	16	32	19	38	31	17	34	23	1	12	24
3	3	9	27	36	18	9	27	36	18	9	27	36	—	24
4	4	16	19	31	34	1	4	16	19	31	34	1	6	24
5	5	25	35	40	20	10	5	25	35	40	20	10	—	24
6	6	36	36	36	36	36	36	36	36	36	36	36	—	24
7	7	4	28	16	22	19	43	31	37	34	13	1	12	24
8	8	19	17	1	8	19	17	1	8	19	17	1	4	24
9	9	36	9	36	9	36	9	36	9	36	9	36	—	24
10	10	10	10	10	10	10	10	10	10	10	10	10	—	24
11	11	31	26	16	41	1	11	31	26	16	41	1	6	24
12	12	9	18	36	27	9	18	36	27	9	18	36	—	24

$J(45)$ is calculated using theorem 3.11: $J(45) = J(3^2 * 5) = 3^1 * 2 * 4 = 24$.

Since 45 is not a prime, there is no “multiplicative order” for all values of a (e. g. for the numbers that are not relatively prime to $45 : 3, 5, 6, 9, 10, 12, \dots$, because $45 = 3^2 * 5$).

Example 3:

Is 7 a primitive root modulo 45?

The requirement/condition $\gcd(7, 45) = 1$ is fulfilled. The table ‘values of $a^i \bmod 45$ ’ shows that the number 7 is not a primitive root of 45, because $\text{ord}_{45}(7) = 12 \neq 24 = J(45)$.

Example 4:

The following table⁶⁹ answers the question as to whether the number 7 is a primitive root of 46. The requirement/condition $\gcd(7, 46) = 1$ is fulfilled.

⁶⁸See [Appendix E of this chapter](#) for the source code to generate the table using Mathematica and Pari-GP.

⁶⁹See [Appendix E of this chapter](#) for the source code to generate the table using Mathematica and Pari-GP.

Table 6: Values of $a^i \bmod 46, 1 \leq a, i < 23$:

$a \backslash i$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	ord
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	4	8	16	32	18	36	26	6	12	24	2	4	8	16	32	18	36	26	6	12	24	2	–
3	3	9	27	35	13	39	25	29	41	31	1	3	9	27	35	13	39	25	29	41	31	1	3	11
4	4	16	18	26	12	2	8	32	36	6	24	4	16	18	26	12	2	8	32	36	6	24	4	–
5	5	25	33	27	43	31	17	39	11	9	45	41	21	13	19	3	15	29	7	35	37	1	5	22
6	6	36	32	8	2	12	26	18	16	4	24	6	36	32	8	2	12	26	18	16	4	24	6	–
7	7	3	21	9	17	27	5	35	15	13	45	39	43	25	37	29	19	41	11	31	33	1	7	22
8	8	18	6	2	16	36	12	4	32	26	24	8	18	6	2	16	36	12	4	32	26	24	8	–
9	9	35	39	29	31	3	27	13	25	41	1	9	35	39	29	31	3	27	13	25	41	1	9	11
10	10	8	34	18	42	6	14	2	20	16	22	36	38	12	28	4	40	32	44	26	30	24	10	–
11	11	29	43	13	5	9	7	31	19	25	45	35	17	3	33	41	37	39	15	27	21	1	11	22
12	12	6	26	36	18	32	16	8	4	2	24	12	6	26	36	18	32	16	8	4	2	24	12	–
13	13	31	35	41	27	29	9	25	3	39	1	13	31	35	41	27	29	9	25	3	39	1	13	11
14	14	12	30	6	38	26	42	36	44	18	22	32	34	16	40	8	20	4	10	2	28	24	14	–
15	15	41	17	25	7	13	11	27	37	3	45	31	5	29	21	39	33	35	19	9	43	1	15	22
16	16	26	2	32	6	4	18	12	8	36	24	16	26	2	32	6	4	18	12	8	36	24	16	–
17	17	13	37	31	21	35	43	41	7	27	45	29	33	9	15	25	11	3	5	39	19	1	17	22
18	18	2	36	4	26	8	6	16	12	32	24	18	2	36	4	26	8	6	16	12	32	24	18	–
19	19	39	5	3	11	25	15	9	33	29	45	27	7	41	43	35	21	31	37	13	17	1	19	22
20	20	32	42	12	10	16	44	6	28	8	22	26	14	4	34	36	30	2	40	18	38	24	20	–
21	21	27	15	39	37	41	33	3	17	35	45	25	19	31	7	9	5	13	43	29	11	1	21	22
22	22	24	22	24	22	24	22	24	22	24	22	24	22	24	22	24	22	24	22	24	22	24	22	–
23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	–

$J(46)$ is calculated using theorem 3.9: $J(46) = J(2 * 23) = 1 * 22 = 22$. The number 7 is a primitive root of 46, because $\text{ord}_{46}(7) = 22 = J(46)$.

Theorem 3.14. ^{70,71} For a modulus n and a relative prime to n the following holds:
 $\{a^i \bmod n | i = 1, \dots, J(n)\}$ equals the multiplicative group Z_n^* if and only if $\text{ord}_n(a) = J(n)$.

⁷⁰For prime moduli p all a with $0 < a < p$ are of order $J(p) = p - 1$. Compare table 5 for an example. In this case $a^i \bmod n$ goes through all the values $1, \dots, p - 1$. Exhausting all possible values of the set is an important cryptographic proposition (compare theorem 3.5). This determines a permutation $\pi(p - 1)$.

⁷¹Table 6 demonstrates that for composite moduli n not all a are of maximal order $J(n)$. In this example only 5, 7, 11, 15, 17, 19 and 21 are of order 22.

3.10 Proof of the RSA procedure with Euler-Fermat

Using the Euler-Fermat theorem, we can “prove” the RSA^{72,73} procedure in the group \mathbb{Z}_n^* .

3.10.1 Basic idea of public key cryptography

The basic idea behind public key cryptography is that all participants possess a different pair of keys (P and S) and the public keys for all recipients are published. You can retrieve the public key P for a recipient from a directory just as you would look up someone’s phone number in the phone book. Furthermore, each recipient has a secret key S that is needed in order to decrypt the message and that is not known to anyone else. If the sender wishes to send a message M , he encrypts it using the public key P of the recipient before sending it:

The cipher text C is determined as $C = E(P; M)$, where E (encryption) is the encryption rule. The recipient uses his private key S to decrypt the message with the decryption rule $D : M = D(S; C)$.

In order to ensure that this system works for every message M , the following four **requirements** must be met:

1. $D(S; E(P; M)) = M$ for every M (invertibility) and M takes “very many” of its possible values.
2. All (S, P) pairs are different for all participants (i.e. lots of them are needed).
3. The time required to derive S from P is at least as high as the time required to decrypt M with no knowledge of S .
4. Both C and M can be calculated relatively easily.

The 1st requirement is a general condition for all cryptographic encryption algorithms.

The 2nd requirement can easily be met because there is a “very” large number of prime numbers⁷⁴ and because this can be ensured by a central office that issues certificates.

⁷²The RSA procedure is the most common asymmetric cryptography procedure. Developed in 1978 by Ronald Rivest, Adi Shamir and Leonard Adleman, it can be used both for signatures and for encryption. Cryptographers always associate this procedure with the abbreviation “RSA”. However, please note that the abbreviation may be used in other fields to mean something completely different.

⁷³In literature and movies not only classical methods are used (like the secret message solved by Sherlock Holmes in “The Adventure of the Dancing Men” by Arthur Conan Doyle), but also modern schemes. Examples are

- the movie “House of Cards”, 1992, where autistic children communicate via primes,
- the story “The Dialogue of the Sisters” by Dr. C. Elsner , 1999 (included in the CrypTool package as PDF file), where the sisters communicate confidentially using a variant of RSA.
- the novel “Das Museum der gestohlenen Erinnerungen (Museum of the stolen memories)” by Ralf Isau , Thiene-mann 2003, where the last part of the oracle can only be solved with the joined help of the computer community.

⁷⁴According to the prime number theorem of Legendre and Gauss there are approximately $n/\ln(n)$ prime numbers up to the number n . This means, for example, that there are $6.5 \cdot 10^{74}$ prime numbers under $n = 2^{256}$ ($= 1.1 \cdot 10^{77}$) and $3.2 \cdot 10^{74}$ prime numbers under $n = 2^{255}$. Between 2^{255} and 2^{256} there are therefore $3.3 \cdot 10^{74}$ prime numbers with precisely 256 bits. This large number is also the reason why we cannot simply save them all.

It is this last requirement that makes the procedure actually usable. This is because it is possible to calculate the powers in a linear amount of time (because there is a restriction on the length of the numbers).

Although Whitfield Diffie and Martin Hellman formulated the general method as early as 1976, the actual procedure that met all four requirements was only discovered later by Rivest, Shamir and Adleman.

3.10.2 How the RSA procedure works

The individual steps for implementing the RSA procedure can be described as follows (see [Eckert2003, p. 213 ff] and [Sedgewick1990, p. 338 ff]). Steps 1 to 3 constitute key generation, steps 4 and 5 are the encryption, and steps 6 and 7 are the decryption:

1. Select two distinct random prime numbers^{75,76} p and q and calculate $n = p * q$ ⁷⁷. The value n is called the RSA modulus⁷⁸.
2. Select a random $e \in \{2, \dots, n - 1\}$ such that:
 e is relatively prime to $J(n) = (p - 1) * (q - 1)$. For example, we can select e such that:
 $\max(p, q) < e < J(n) - 1$ ⁷⁹. We can then “throw away” p and q .
3. Select $d \in \{1, \dots, n - 1\}$ with $e * d \equiv 1 \pmod{J(n)}$, i.e. d is the multiplicative inverse of e modulo $J(n)$ ^{80,81}. We can then “throw away” $J(n)$.
 $\rightarrow (n, e)$ is the public key P .
 $\rightarrow (n, d)$ is the secret key S (only d must be kept secret).

⁷⁵Compaq introduced the so-called multi-prime method with high marketing effort in 2000. n was the product of two big and one relative small prime: $n = o * p * q$. With theorem 3.10 we get: $J(n) = (o - 1) * (p - 1) * (q - 1)$. This method did not assert itself yet.

One reason probably is, that Compaq claimed a patent on it. Generally there is less understanding in Europe and with the Open Source Initiative, that one can claim patents on algorithms. But there is really no understanding outside the U.S., that one can get a patent for a special case (3 factors) of an algorithm (RSA), although the patent for the general case is almost expired.

⁷⁶If the two primes p and q are equal then $(m^e)^d \equiv m \pmod{n}$ is not true for all $m < n$ (although $e * d \equiv 1 \pmod{J(n)}$ is fulfilled). Example: If $n = 5^2$ then according to theorem 3.11 it is $J(n) = 5 * 4 = 20$, $e = 3$, $d = 7$, $e * d = 21 \equiv 1 \pmod{J(n)}$. But it is $(5^3)^7 \equiv 0 \pmod{25}$.

⁷⁷The GISA (German Information Security Agency) recommends, to choose the prime factors p and q almost the same, but not too close:

$$0.5 < |\log_2(p) - \log_2(q)| < 30.$$

They recommend to generate the primes independently and check that the restriction is fulfilled (see [GISA2002]).

⁷⁸In CrypTool the RSA modulo is denoted with a capital “N”.

⁷⁹The procedure also allows us to select d freely and then calculate e . However, this has practical disadvantages. We usually want to be able to encrypt messages “quickly”, which is why we choose a public exponent e such that it has a short bit length compared to the modulus n and as few binary ones as possible. A number often used is $65537 = 2^{16} + 1$, or in binary: $10 \dots 0 \dots 01$. We want to select the publicly known e to be an advantageous value that allows the exponential calculation to be performed quickly during encryption. The prime numbers 3, 17 and 65537 have proved to be particularly practical for this purpose.

⁸⁰For reasons of security, d should not be too small.

⁸¹We start by determining either d or e depending on the implementation.

4. For encryption, the message represented as a (binary) number is divided into parts such that each part of the number is less than n .

5. Encryption of the plaintext (or the parts of it) $M \in \{1, \dots, n-1\}$:

$$C = E((n, e); M) := M^e \pmod{n}.$$

6. For decryption, the cipher text represented as a binary number is divided into parts such that each part of the number is less than n .

7. Decryption of the cipher text (or the parts of it) $C \in \{1, \dots, n-1\}$:

$$M = D((n, d); C) := C^d \pmod{n}.$$

The numbers d, e and n are usually extremely large (e. g. d and e 300 bits, n 600 bits).

Comment:

The security of the RSA algorithm depends as with all public key methods on the difficulty to calculate the private key d from the public key (n, e) .

Concrete for the RSA method does this mean:

1. it is hard to calculate $J(n)$ for big compounds n and
2. it is hard to calculate the prime factors of big compounds n (Factorisation!factorisation problem).

3.10.3 Proof of requirement 1 (invertibility)

For pairs of keys (n, e) and (n, d) that possess fixed properties in steps 1 to 3 of the RSA procedure, the following must be true for all $M < n$:

$$M \equiv (M^e)^d \pmod{n} \quad \text{with} \quad (M^e)^d = M^{e*d}.$$

This means that the deciphering algorithm above works correctly.

We therefore need to show that:

$$M^{e*d} \equiv M \pmod{n}.$$

We will show this in 3 steps (see [Beutelspacher1996, p. 131ff]).

Step 1:

In the first step we show that: $M^{e*d} \equiv M \pmod{p}$. This results from the requirements and from Euler-Fermat (theorem 3.13). Since $n = p * q$ and $J(p * q) = (p - 1) * (q - 1)$ and since e and

d are selected in such a way that $e * d \equiv 1 \pmod{J(n)}$, there is a whole number k such that:
 $e * d = 1 + k * (p - 1) * (q - 1)$.

$$\begin{aligned} M^{e*d} &\equiv M^{1+k*J(n)} \equiv M * M^{k*J(n)} \equiv M * M^{k*(p-1)*(q-1)} \pmod{p} \\ &\equiv M * (M^{p-1})^{k*(q-1)} \pmod{p} \quad \text{based on little Fermat : } M^{p-1} \equiv 1 \pmod{p} \\ &\equiv M * (1)^{k*(q-1)} \pmod{p} \\ &\equiv M \pmod{p} \end{aligned}$$

The requirement for using the simplified Euler-Fermat theorem (theorem 3.12) was that M and p are relatively prime.

Since this is not true in general, we need to consider the case when M and p are not relatively prime. Since p is a prime number, this implies that p is a factor of M . But this means:

$$M \equiv 0 \pmod{p}.$$

If p is a factor of M , then p is also a factor of M^{e*d} . Therefore:

$$M^{e*d} \equiv 0 \pmod{p}.$$

Since p is a factor of both M and M^{e*d} , it is also a factor of their difference:

$$(M^{e*d} - M) \equiv 0 \pmod{p}.$$

And therefore our conjecture is also true in this special case.

Step 2:

In exactly the same way we prove that: $M^{e*d} \equiv M \pmod{q}$.

Step 3:

We now combine the conjectures from (a) and (b) for $n = p * q$ to show that:

$$M^{e*d} \equiv M \pmod{n} \text{ for all } M < n.$$

From (a) and (b) we have $(M^{e*d} - M) \equiv 0 \pmod{p}$ and $(M^{e*d} - M) \equiv 0 \pmod{q}$. Therefore, p and q are both factors of the same number $z = (M^{e*d} - M)$. Since p and q are **distinct** prime numbers, their product must also be a factor of this number z . Thus:

$$(M^{e*d} - M) \equiv 0 \pmod{p * q} \quad \text{or} \quad M^{e*d} \equiv M \pmod{p * q} \quad \text{or} \quad M^{e*d} \equiv M \pmod{n}.$$

□

1st comment:

We can also condense the three steps if we use the theorem 3.13 (Euler-Fermat) - i.e. not the simplified theorem where $n = p$ and which corresponds to Fermat's Little Theorem:

$$\begin{aligned} (M^e)^d \equiv M^{e*d} &\equiv M^{(p-1)(q-1)*k+1} \equiv (\underbrace{M^{(p-1)(q-1)}}_{\equiv M^{J(n)} \equiv 1 \pmod{n}})^k * M \equiv 1^k * M \equiv M \pmod{n}. \\ &\equiv M^{J(n)} \equiv 1 \pmod{n} \end{aligned}$$

2nd comment:

When it comes to signing messages, we perform the same operations but first use the secret key d , followed by the public key e . The RSA procedure can also be used to create digital signatures, because:

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

3.11 Considerations regarding the security of the RSA algorithm⁸²

There have always been discussions about the suitability of the RSA algorithm for digital signatures and encryption, e. g. after publications of breakthroughs in factorisation. Nevertheless the RSA algorithm has become a de-facto standard since it was published more than 20 years ago (compare 6.1).

The security of the RSA algorithm rests — as with all cryptographic methods — on the following 4 central pillars:

- the complexity of the number theoretical problem on which the algorithm is based (here factorisation of big numbers),
- the election of fitting parameters (here the length of the module N),
- the adequate usage of the algorithm and key generation and
- the correct implementation of the algorithm.

Usage and key generation are well understood today. Implementation based on long integer arithmetic is very easy.

The following sections examine the RSA algorithm with respect to the first two points.

3.11.1 Complexity

Successful decryption or forgery of a signature — without knowing the private key — requires calculating the e -th root mod n . The private key, this is the multiplicative inverse of e mod $J(n)$, can be easily determined if $J(n)$ is known. $J(n)$ again can be calculated from the prime factors of n . Breaking of RSA therefore cannot be more difficult than factorisation of the module n .

The best factorisation method known today is a further development of the General Number Field Sieve (GNFS), which was originally devised to factor only numbers of a special form (like Fermat numbers). The complexity of solving the factorisation problem with the GNFS is asymptotically

$$O(l) = e^{c \cdot (l \cdot \ln 2)^{1/3} \cdot (\ln(l \cdot \ln 2))^{2/3} + o(l)}$$

Please refer to:

- A. Lenstra, H. Lenstra: *The development of the Number Field Sieve* [Lenstra1993].

⁸²Major parts of chapters 3.11.1 and 3.11.2 follow the article “Vorzüge und Grenzen des RSA-Verfahrens” written by F. Bourseau, D. Fox and C. Thiel [Bourseau2002]

- Robert D. Silverman: *A Cost-Based Security Analysis of Symmetric and Asymmetric Key Lengths* [Silverman2000].

This formula shows, that the factorisation problem belongs to the class of problems with sub-exponential time complexity (i. e. time complexity grows asymptotically not as fast as exponential functions like e^l or 2^l , but strictly slower, e. g. like $e^{\sqrt{l}}$). This classification is all that is currently known; it does not preclude the possibility that the factorisation problem can be solved in polynomial time (see 3.11.3).

$O(l)$ is the average number of processor steps depending on the bit length l of the number n to be factorised. For the best currently known factorisation algorithm the constant $c = (64/9)^{1/173} = 1,923$.

The inverse proposition, that the RSA algorithm can be broken only by factorisation of n , is still not proven. Most number theorists consider the “RSA problem” and the factorisation problem equivalent in terms of time complexity.

Please refer to: *Handbook of Applied Cryptography* [Menezes2001].

3.11.2 Security parameters because of progress in factorisation

The complexity is basically determined by the length l of the module n .

- In 1994 a 129-digit RSA module (428 bit), published in 1977, was factorised by a distributed implementation of the Quadratic Sieve algorithm (QS), developed 1982 by Pomerance. This effort took 8 months.

Please refer to:

C. Pomerance: *The quadratic sieve factoring algorithm* [Pomerance1984].

- In 1999 a 155-digit module (512 bit) was factored with an implementation of the General Number Field Sieve algorithm (GNFS), developed by Buhler, Lenstra and Pomerance. The GNFS is more efficient than QS if n is longer than about 116 decimal digits. This effort took 5 months.

Please refer to:

J.P. Buhler, H.W. Lenstra, C. Pomerance: *Factoring integers with the number field sieve* [Buhler1993].

This made evident that a module length of 512 bit no longer prevents from attackers.

Within the last 20 years a lot of progress has been made. Estimations about the future development of the ability to factor RSA modules vary and depend on some assumptions:

- progression in computing performance (Moore’s law: every 18 month the computing power will double) and in grid computing.

- development of new algorithms.

Within the last years the module bit length feasible for factorisation increased — even without new algorithms — by 10 bit per year. Larger numbers require not only more time to be factored, but also huge RAM storage for the solutions matrix being used by the best algorithms known today. This need for storage grows like the square root of the computation time, i. e. also sub-exponentially. Because RAM availability increased exponentially in the recent decades, it seems that this should not be the limiting factor.

Within the above mentioned article [Bourseau2002] Dirk Fox⁸³ prognosticates an almost linear progression with the factorisation, if all factors are included: Each year the module length feasible for factorisation increases by 20 bit on average.

3.11.3 Further current research in factorisation and number theory

Prime numbers are part of very many topical research areas in number theory and computer science. Progress made with factorisation is bigger than was estimated 5 years ago – this is not only due to faster computers but also new knowledge.

The security of the RSA algorithm is based on the empirical observation that factoring large numbers is a hard problem. A module n (typically, 1024 bit) can be easily constructed as the product of two large primes p , q (typically, 500–600 bit each), by calculating $n = pq$. However, it is a hard problem to extract p , q from n . Without knowing p or q , the private key cannot be calculated.

Thus, any progress in efficiency of factorising large integers will effect the security of the RSA. As a consequence, the underlying primes p , q and, thus, the module n (1024 bit as of today) have to be increased. In case of a quantum leap in factorisation, the RSA algorithm might be compromised.

Bernstein’s paper and its implication on the security of the RSA algorithm

In his paper “Circuits for integer factorisation: a proposal” (<http://cr.yp.to/djb.html>), published November 2001, D. J. Bernstein [Bernstein2001] addresses the problem of factorising large integers. Therefore, his results are of relevance from a RSA point of view. As a main result Bernstein claims that the implementation of the General Number Field Sieve algorithm (GNFS) can be improved to factor, with the same effort as before, integers with three times more digits.

We note that the definition of *effort* is a crucial point: Bernstein claims that effort is the product of time and costs of the machine (including the memory used). The gist of the paper lies in the fact that he can reduce a big part of factorising to sorting. Using Schimmler’s scheme, sorting

⁸³His company Secorvo Ltd made a statement on the recommendation for key length selection published by the GISA (German Information Security Agency). Chapter 2.3.1 of this statement contains a very competent and understandable discussion of RSA security (this document exists – to my knowledge – only in German): <http://www.secorvo.de/publikat/stellungnahme-algorithmenempfehlung-020307.pdf>

can be optimized by massive parallel computing. At the end of section 3 Bernstein explains this effect: The costs of m^2 parallel computers with a constant amount of memory is a constant times m^2 . The costs of a computer with a single processor and memory of size m^2 is also of the order of m^2 , but with a different constant factor. With m^2 processors in parallel, sorting of m^2 numbers (with Schimmler's scheme) can be achieved in time m , while a m^2 -memory computer needs time of the order of m^2 . Decreasing memory and increasing the number of processors, the computing time can be reduced by a factor $1/m$ without additional effort in terms of total costs. In section 5 it is said that massive parallel computing can also increase efficiency of factorising using Lenstra's elliptic-curve-method (a search algorithm has costs that increase in a quadratic square manner instead of cubically).

We note that all results achieved so far are asymptotic results. This means that they only hold in the limit n to infinity. Unfortunately, there is no upper limit for the residual error (i.e. the difference between the real and the asymptotic value) for finite n — a problem which has already been addressed by the author. As a consequence, one cannot conclude whether the costs (in the sense of Bernstein) for factorising 1024–2048-bit RSA modules can be significantly reduced.

There is no doubt that Bernstein's approach is innovative. However, the reduction of computing time under constant costs comes along with a massive use of parallel computing — a scenario which seems not to be realistic yet. For example, formally 1 sec computing time on one machine and $1/1,000,000$ sec time parallel computing time on 1,000,000 machines might have same costs. In reality, it is much harder to realize the second situation, and Bernstein does not take into account the fixed costs, in particular for building a network between all these computers.

Although distributed computing over a large network might help to overcome this problem, realistic costs for data transfer have to be taken into account — a point which was not addressed in Bernstein's proposal.

As long as there is neither (low cost) hardware nor a distributed computing approach (based on Bernstein's ideas), there should not be a problem for RSA. It has to be clarified from which magnitude of n on Bernstein's method could lead to a significant improvement (in the sense of the asymptotic result).

Arjen Lenstra, Adi Shamir et. al. analyzed the paper of Bernstein [Lenstra2002]. In summary they expect a factorisation improvement on how much longer the bit length of the keys could be with a factor of 1.17 (instead of factor 3 as proposed by Bernstein).

The abstract of their paper "Analysis of Bernstein's Factorization Circuit" says:

"... Bernstein proposed a circuit-based implementation of the matrix step of the number field sieve factorisation algorithm. We show that under the non-standard cost function used in [1], these circuits indeed offer an asymptotic improvement over other methods but to a lesser degree than previously claimed: for a given cost, the new method can factor integers that are 1.17 times larger (rather than 3.01). We also propose an improved circuit design based on a new mesh routing algorithm, and show that for factorisation of 1024-bit integers the matrix step can, under an optimistic assumption about the matrix size, be completed within a day by a device that costs a few thousand dollars. We conclude that from a practical standpoint, the security of RSA relies exclusively on the hardness of the relation collection step of the number field sieve."

RSA Security's analysis of the Bernstein paper [RSA Security 2002] from April, 8 2002 also – as expected – concludes, that RSA is still not compromised.

This is still an ongoing discussion.

When this section was written (June 2002) nothing was publicly known about, how far there exist implementations of his theoretical onsets and how much financing there was for his research project.

Links:

<http://cr.yp.to/djb.html>

<http://www.counterpane.com/crypto-gram-0203.html#6>

<http://www.math.uic.edu>

The TWIRL device

In January 2003 Adi Shamir and Eran Tromer from the Weizmann Institute of Science published a preliminary draft called “*Factoring Large Numbers with the TWIRL Device*” raising concerns about the security of key sizes till 1024 bits [Shamir2003].

Their abstract summarizes their results very well: “The security of the RSA cryptosystem depends on the difficulty in factoring large integers. The best current factoring algorithm is the Number Field Sieve (NFS), and its most difficult part is the sieving step. In 1999 a large distributed computation involving thousands of workstations working for many months managed to factor a 512-bit RSA key, but 1024-bit keys were believed to be safe for the next 15-20 years. In this paper we describe a new hardware implementation of the NFS sieving step ... which is 3-4 orders of magnitude more cost effective than the best previously published designs Based on a detailed analysis of all the critical components (but without an actual implementation), we believe that the NFS sieving step for 1024-bit RSA keys can be completed in less than a year with a \$10M device, and that the NFS sieving step for 512-bit RSA keys can be completed in less than ten minutes with a \$10K device. Coupled with recent results about the difficulty of the NFS matrix step ... this raises some concerns about the security of these key sizes.”

The 3-page article in the DuD issue of June 2003 [Weis2003] contains a very good explanation, how the attack using the Generalized Number Field Sieve (GNFS) works and which progress is made, to factorize numbers. At GNFS we can distinguish 2 general steps: The sieve step (relation collecting) and the matrix reduction. Besides the sieve step is highly parallelizable, it dominates the overall calculation burden. Shamir and Tromer haven't built a TWIRL device yet, but the estimated costs of 10 till 50 million Euro (in order to factorise a 1024-bit number) is not prohibitive for secret agencies or big criminal organizations, because the “costs for a single espionage satellite is estimated e.g. to be several billion USD”. The authors therefore recommend, to get as soon as possible rid of today used sensible RSA, Diffie-Hellman or ElGamal keys up to 1024 bit and to use then keys of at least 2048 bit length. The planned TPCA/Palladium hardware will use 2048-bit RSA keys!

So recommendations like the ones from the GISA (German Information Security Agency) to use higher key lengths are very valid.

“Primes in P”: Primality testing is polynomial

In August 2002 the three Indian researchers M. Agrawal, N. Kayal and N. Saxena published the paper “*PRIMES in P*” about a new primality testing algorithm called AKS [Agrawal2002]. They discovered a polynomial time deterministic algorithm for determining if a number is prime or not.

The importance of this discovery is that it provides number theorists with new insights and opportunities for further research. Lots of people over centuries have been looking for a polynomial time test for primality, and this result is a major theoretic breakthrough. It shows that new results can be generated from already known facts.

But even its authors note that other known algorithms may be faster (for example ECPP). The new algorithm works on any integer. For example the GIMPS project uses the Lucas-Lehmer primality test which takes advantage of the special properties of Mersenne numbers. This makes the Lucas-Lehmer test much faster, allowing to test numbers with millions of digits while general purpose algorithms are limited to numbers with a few thousand digits.

Current research results on this topic can be found at:

<http://www.mersenne.org/>

<http://fatphil.org/maths/AKS/> Original paper in English

<http://ls2-www.cs.uni-dortmund.de/lehre/winter200203/kt/material/primes.ps>

Good explanation in German by Thomas Hofmeister.

*Hermann Hesse*⁸⁴:

To let the possible happen, you again and again have to try the impossible.

3.11.4 Status regarding factorisation of concrete large numbers

The web page <http://www.crypto-world.com> contains an excellent overview about the factoring records of composed integers using different methods.

The current record (as of May 2003) obtained using the GNFS method (General Number Field Sieve) factorised a 160-digit into its both prime factors.

The last records with factorisation algorithms for general numbers are:

RSA-155

On August 22, 1999 researchers from the Netherlands found the solution of the RSA challenge⁸⁵. They factorised a 155-digit number into its both 78-digit primes (see chapter 3.11.2).

This 512 bit RSA-155 meant to reach a kind of *magic* border.

⁸⁴German/Swiss writer and Nobel Prize winner, July 2, 1877 – August 9, 1962.

⁸⁵<http://www.rsasecurity.com/rsalabs/challenges/factoring/numbers.html>

C158

On January 18, 2002 researchers at the German University of Bonn⁸⁶ factorised a 158-digit number into its both prime factors (these are build with 73 and 86 decimal digits) using the GNFS method (General Number Field Sieve).

This record got much less attention within the press than the solution of RSA-155.

The task of the researchers from Bonn was not initiated by a challenge, but they wanted to find the last prime factors of the integer $2^{953} - 1$ (see “Wanted List” of the Cunningham Project ⁸⁷).

The 6 smaller prime factors, already found before have been:

3, 1907, 425796183929,
1624700279478894385598779655842584377,
3802306738549441324432139091271828121 and
128064886830166671444802576129115872060027.

The first 3 factors can be easily computed⁸⁸. The next three prime factors were found by P. Zimmermann⁸⁹, T. Grandlund⁹⁰ and R. Harley during the years 1999 and 2000 using the elliptic curve factorisation method.

The last remaining factor, called “C158”, was known to be composite by then, but its factors where not known (the following 3 lines contain one number):

39505874583265144526419767800614481996020776460304936
45413937605157935562652945068360972784246821953509354
4305870490251995655335710209799226484977949442955603

The factorisation of C158 resulted in the following two prime factors:

3388495837466721394368393204672181522815830368604993048084925840555281177

and

1165882340667125990314837655838327081813101
2258146392600439520994131344334162924536139.

So now all 8 prime factors of $2^{953} - 1$ have been found.

Links:

<http://www.loria.fr/~zimmerma/records/gnfs158>
<http://www.crypto-world.com/FactorRecords.html>
<http://www.crypto-world.com/announcements/c158.txt>

⁸⁶http://www.ercim.org/publication/Ercim_News/enw49/franke.html, 2002-01

⁸⁷Cunningham project: <http://www.cerias.purdue.edu/homes/ssw/cun/>

⁸⁸e.g. using CrypTool via menu **Indiv. Procedures \ RSA Demonstration \ Factorisation of a Number**.

⁸⁹<http://www.loria.fr/~zimmerma/ecmnet>

⁹⁰<http://www.swox.se/gmp/>

RSA-160

On January 18, 2002 researchers at the German University of Bonn⁹¹ factorised a 160-digit number into its both prime factors (these are build with each 80 decimal digits) using the GNFS method (General Number Field Sieve).

The computations for the factorization of RSA-160 took place at the German Information Security Agency (GISA) in Bonn⁹².

The 160-digit decimal number origins from the old challenge list of RSADSI. This number was retracted after RSA-155 (RSA512) had been factorized successfully, leaving RSA576, RSA640 etc. (using 64 bit steps upwards). The prime factors of RSA-160 were still unknown. So this record of the team of Prof. Franke provides the solution of the old challenge, for which RSADSI didn't award a price anymore.

The composite number called "RSA-160" is (the following 3 lines contain one number):

```
215274110271888970189601520131282542925777358884567598017049
767677813314521885913567301105977349105960249790711158521430
2079314665202840140619946994927570407753
```

The factorisation of RSA-160 resulted in the following two prime factors:

$$p = 45427892858481394071686190649738831$$
$$656137145778469793250959984709250004157335359$$

and

$$q = 47388090603832016196633832303788951$$
$$973268922921040957944741354648812028493909367$$

The calculations took place between December 2002 and April 2003.

Links:

<http://www.loria.fr/~zimmerma/records/rsa160>
<http://www.loria.fr/~zimmerma/records/factor.html>
<http://www.crypto-world.com/FactorWorld.html>

As you notice the factorised compound numbers built of 2 prime factors are much smaller than the especially structured numbers, for which primality tests are able to decide whether they are prime or not (see chapters 2.4, 2.5 and 2.6).

⁹¹<http://www.loria.fr/~zimmerma/records/rsa160>

⁹²Every year the GISA creates a paper to describe which crypto algorithms are feasible to generate digital signatures according to the German signature law – under participation of experts from economy and science. To review signature methods based on the factorisation problem the GISA also co-operates with researchers from the University of Bonn. Further information about crypto algorithms can be found on the web page of GISA: <http://www.bsi.bund.de/esig/basics/techbas/krypto/index.htm>.

Joanne K. Rowling⁹³:

It is our choices, that show what we truly are, far more than our abilities.

3.12 Further applications of number theory in cryptography

The results of modular arithmetic are used extensively in modern cryptography. Here we will provide a few examples from cryptography using small⁹⁴ numbers.

Enciphering a text entails applying a function (mathematical operation) to a character string (number) to generate a different number. Deciphering entails reversing this function, in other words using the distorted image that the function has created from the plaintext in order to restore the original image. For example, the sender could take the plaintext M of a confidential message and add a secret number, the key S , to obtain the cipher text C :

$$C = M + S.$$

The recipient can reconstruct the plaintext by reversing this operation, in other words by subtracting S :

$$M = C - S.$$

Adding S reliably makes the plaintext impossible to read. However, this encryption is rather weak, because all an interceptor needs to do to calculate the key is obtain a plaintext and the associated cipher text

$$S = C - M,$$

and can then read any subsequent messages encrypted using S .

The essential reason for this is that subtraction is just as simple an operation as addition.

3.12.1 One way functions

If the key is to be impossible to determine even with knowledge of both the plaintext and the cipher text, we need a function that is, on the one hand, relatively easy to calculate – we don't want to have problems encrypting messages. On the other hand, the inverse function should exist (otherwise information would be lost during encryption), but should be de facto in calculable.

What are possible candidates for such a **one way function**? We could take multiplication rather than addition, but even primary school children know that the inverse function, division, is only slightly more difficult than multiplication itself. We need to go one step higher in the hierarchy of calculation methods. It is still relatively simple to calculate the power of a number, but the corresponding two reverse functions – *taking roots* (find b in the equation $a = b^c$ when a and c

⁹³Joanne K. Rowling, "Harry Potter and the Chamber of Secrets", Bloomsbury, 1998, last chapter "Dobby's reward", p. 245, by Dumbledore.

⁹⁴In the RSA procedure, we call numbers "small" if the bit lengths are much less than 1024 bits (i.e. 308 decimal points). In practice, 1024 bits is currently the minimum length for a secure Certification Authority RSA modulus.

are known) and *calculating logarithms* (find c in the above equation when a and b are known) are so complicated that pupils normally do not learn them at school.

Although a certain structure can still be recognised for addition and multiplication, raising numbers to the power of another or calculating exponentials totally mixes up all the numbers. Knowing a few values of the function doesn't tell us much about the function as a whole (in contrast to addition and multiplication).

3.12.2 The Diffie-Hellman key exchange protocol

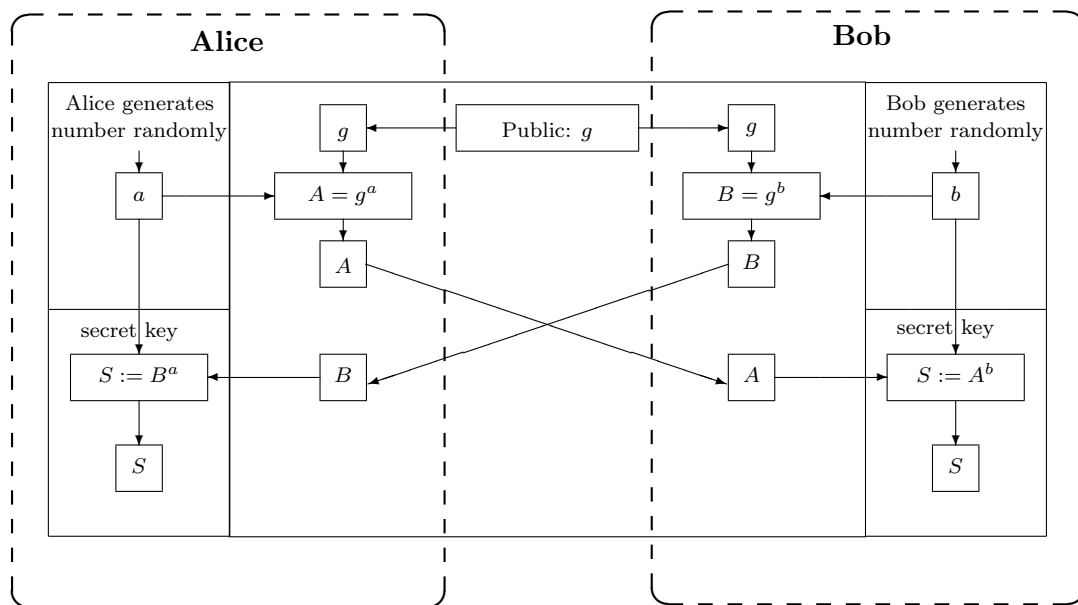
Whitfield Diffie, Martin E. Hellman and Ralph Merkle developed this DH key exchange protocol in Stanford in 1976⁹⁵.

Alice and Bob⁹⁶ use a one way function to obtain a key S , the session key, for subsequent correspondence. This is then a secret that is only known to the two of them. Alice selects a random number a and keeps it secret. She applies a one way function to a to calculate the number $A = g^a$ and sends it to Bob. He does the same, by selecting a secret random number b , calculating $B = g^b$ and sending it to Alice. The number g is random and can be publicly known. Alice applies the one way function together with her secret number a to B , while Bob does the same with his secret number b and the received number A .

The result S is the same in each case because the one way function is commutative: $(g^a)^b = (g^b)^a$. But even Bob cannot reconstruct Alice's secret number a from the data available to him, while Alice cannot determine Bob's secret number b . And a perpetrator who knows g and has intercepted both A and B cannot use this knowledge to determine a , b or S .

⁹⁵With CrypTool v1.3.04 this exchange protocol has been visualized: you can execute the single steps with concrete numbers using menu **Indiv. Procedures \ Diffie-Hellman Demonstration**.

⁹⁶Bob and Alice are the default names used for the two authorised participants in a protocol (see [Schneier1996, p. 23]).



Procedure:

Alice and Bob want to negotiate a secret session key S via a channel that may be intercepted.

1. They select a prime number p and a random number g and exchange this information openly.
2. Alice now selects a , a random number less than p and keeps it secret.
Similarly, Bob selects b , a random number less than p and keeps it secret.
3. Alice now calculates $A \equiv g^a \pmod{p}$.
Bob calculates $B \equiv g^b \pmod{p}$.
4. Alice sends the result A to Bob.
Bob sends the result B to Alice.
5. In order to now determine the session key to be used by both, they both separately raise the respective results they have received to the power of their secret random number modulo p . This means:
 - Alice calculates $S \equiv B^a \pmod{p}$ and
 - Bob calculates $S \equiv A^b \pmod{p}$.

Even if a spy intercepts g, p , and the interim results A and B , he cannot use these to determine the session key used due to the difficulty of calculating the discrete logarithm.

We will now use an example with (unrealistically) small numbers to illustrate this.

Example using numbers:

1. Alice and Bob select $g = 11, p = 347$.
2. Alice selects $a = 240$, Bob selects $b = 39$ and they keep a and b secret.
3. Alice calculates $A \equiv g^a \equiv 11^{240} \equiv 49 \pmod{347}$.
Bob calculates $B \equiv g^b \equiv 11^{39} \equiv 285 \pmod{347}$.
4. Alice sends Bob: $A \equiv 49$,
Bob sends Alice: $B \equiv 285$.
5. Alice calculates $B^a \equiv 285^{240} \equiv 268 \pmod{347}$,
Bob calculates $A^b \equiv 49^{39} \equiv 268 \pmod{347}$.

Alice and Bob can now communicate securely using their shared session key. Even if spies were to intercept everything transferred via the connection: $g = 11, p = 347, A = 49$ and $B = 285$, they would not be able to calculate the secret key.

Comment:

In this example using such small numbers, it would be possible, but with large numbers the discrete logarithm problem^{97,98} is extremely difficult to solve.

Here, we need to calculate:

For Alice: $11^x \equiv 49 \pmod{347}$, that means $\log_{11}(49) \pmod{347}$.

For Bob: $11^y \equiv 285 \pmod{347}$, that means $\log_{11}(285) \pmod{347}$.

3.13 The RSA procedure with actual numbers

Having described above [how the RSA procedure works](#), we will now work through the steps using actual, but small, numbers.

⁹⁷If you try to determine the discrete logarithm x that solves the equation $11^x \equiv 49 \pmod{347}$ with Mathematica by means of Solve, you obtain the em tdep message “The equations appear to involve the variables to be solved for in an essentially non-algebraic way”. Mathematica therefore claims not to know a direct algebraic procedure for solving the equation. Yet Mathematica is able to calculate this with the general function for the multiplicative order (here for Alice): `MultiplicativeOrder[11, 347, 49]` delivers the value 67.

The syntax with Pari-GP is: `znlog(Mod(49,347),Mod(11,347))`.

Such number-theory tasks can also be solved using other tools such as the LiDIA or BC package (see web links in appendix). The `dl` function in the **LC** user interface for LiDIA also delivers the value 67 for `dl(11,49,347)`.

⁹⁸Why have the functions delivered the value 67 rather than 240 for the `dl` problem for Alice? The discrete logarithm is the smallest natural exponent that solves the equation $11^x \equiv 49 \pmod{347}$. Both $x = 67$ and $x = 240$ (the number selected in the example) satisfy the equation and can therefore be used to calculate the session key: $285^{240} \equiv 285^{67} \equiv 268 \pmod{347}$. If Alice and Bob had selected a primitive root modulo p as base g , then for every remainder from the set $\{1, 2, \dots, p-1\}$ there is exactly one exponent from the set $\{0, 1, \dots, p-2\}$.

For info: there are 172 different primitive roots for modulo 347, 32 of which are prime (not necessary). Since the number 11 selected for g in the example is not a primitive root of 347, the remainders do not take all values from the set $\{1, 2, \dots, 346\}$. Thus, for a particular remainder there may be more than one exponent or even no exponent at all in the set $\{0, 1, \dots, 345\}$ that satisfies the equation.

`PrimeQ[347] = True; EulerPhi[347] = 346; GCD[11, 347] = 1; MultiplicativeOrder[11, 347] = 173`

The syntax with Pari-GP is: `isprime(347); eulerphi(347); gcd(11,347); znorder(Mod(11,347))`.

3.13.1 RSA with small prime numbers and with a number as message

Before applying the RSA procedure to a text, we will first demonstrate it directly using a single number as message⁹⁹.

1. Let the selected prime numbers be $p = 5$ and $q = 11$.
Thus, $n = 55$ and $J(n) = (p - 1) * (q - 1) = 40$.
2. $e = 7$ (should lie between 11 and 40 and must be relatively prime to 40).
3. $d = 23$ (since $23 * 7 \equiv 161 \equiv 1 \pmod{40}$),
→ Public key of the recipient: $(55, 7)$,
→ Private key of the recipient: $(55, 23)$.
4. Let the message be the number $M = 2$ (so no division into blocks is required).
5. Encryption: $C \equiv 2^7 \equiv 18 \pmod{55}$.
6. The cipher text is simply the number $C = 18$ (we therefore do not need to divide it into blocks).
7. Decryption: $M \equiv 18^{23} \equiv 18^{(1+2+4+16)} \equiv 18 * 49 * 36 * 26 \equiv 2 \pmod{55}$.

We will now apply the RSA procedure to a text, first using the upper case alphabet (26 characters), then using the entire ASCII character set as the basis for the messages.

i	$11^i \bmod 347$	
0	1	
1	11	
2	121	
3	290	
67	49	searched exponent
172	284	
173	1	= multiplicative order of 11 (mod 347)
174	11	
175	121	
176	290	
240	49	searched exponent

⁹⁹Using CrypTool v1.3 you can solve this with the **Indiv.Procedures \ RSA Demonstration \ RSA Cryptosystem**.

3.13.2 RSA with slightly larger primes and a text of upper case letters

We have the text “ATTACK AT DAWN” and the characters are coded in the following simple manner¹⁰⁰:

Table 7: capital letters alphabet

Character	Numerical value	Character	Numerical value
Blank	0	M	13
A	1	N	14
B	2	O	15
C	3	P	16
D	4	Q	17
E	5	R	18
F	6	S	19
G	7	T	20
H	8	U	21
I	9	V	22
J	10	W	23
K	11	X	24
L	12	Y	25
		Z	26

Key generation (steps 1 to 3):

1. $p = 47, q = 79$ ($n = 3,713$; $J(n) = (p - 1) * (q - 1) = 3,588$).
2. $e = 37$ (should lie between 79 and 3,588 and must be relatively prime to 3,588).
3. $d = 97$ (since $e * d = 1 \pmod{J(n)}$; $37 * 97 \equiv 3,589 \equiv 1 \pmod{3,588}$)¹⁰¹.

4. Encryption:

Text: A T T A C K A T D A W N
Number: 01 20 20 01 03 11 00 01 20 00 04 01 23 14

This 28-digit number is divided into 4-digit parts (because 2,626 is still smaller than $n = 3,713$):
0120 2001 0311 0001 2000 0401 2314

All 7 parts are encrypted using: $C \equiv M^{37} \pmod{3,713}$ ¹⁰²:
1404 2932 3536 0001 3284 2280 2235

5. Decryption:

Cipher text: 1404 2932 3536 0001 3284 2280 2235

This 28-digit number is divided into 4-digit parts.

¹⁰⁰Using CrypTool v1.3 you can solve this with the **Indiv.Procedures \ RSA Demonstration \ RSA Cryptosystem**. This is also described in the tutorial/scenario in CrypTool’s online help [Options, specify alphabet, number system, block length 2 and decimal representation].

¹⁰¹How to compute $d = 97$ using the *extended* gcd algorithm is shown in [appendix A of this chapter](#)

¹⁰²See [Appendix E of this chapter](#) for source code to do RSA encryption using Mathematica and Pari-GP. You can also encrypt the message with CrypTool v1.3 **Indiv. Procedures \ RSA Demonstration \ Factorisation of a Number**.

All 7 parts are decrypted using: $M \equiv C^{97} \pmod{3,713}$:

0120 2001 0311 0001 2000 0401 2314

The 2-digit numbers are transformed into capital letters and blanks.

Using the selected values it is easy for a cryptanalyst to derive the secret values from the public parameters $n = 3,713$ and $e = 37$ by revealing that $3,713 = 47 * 79$.

If n is a 768-bit number, there is, according to present knowledge, little chance of this.

3.13.3 RSA with even larger primes and a text made up of ASCII characters

In real life, the ASCII alphabet is used to code the individual characters of the message as 8-bit numbers.

The idea for this task¹⁰³ is taken from the example in [Eckert2003, p. 271].

Coded in decimal notation, the text “RSA works!” is as follows:

Text:	R	S	A		w	o	r	k	s	!
Number:	82	83	65	32	119	111	114	107	115	33

We will work through the example in 2 variants. The steps 1 to 3 are common for both.

Key generation (steps 1 to 3):

1. $p = 503, q = 509$ ($n = 256,027$; $J(n) = (p-1)(q-1) = 255,016 = 2^3 * 127 * 251$)¹⁰⁴.
2. $e = 256,027$ (should lie between 509 and 255,016 and must be relatively prime to 255,016)¹⁰⁵.
3. $d = 231,953$
(since $e \equiv d^{-1} \pmod{J(n)}$: $65,537 * 231,953 \equiv 15,201,503,761 \equiv 1 \pmod{67,000}$)¹⁰⁶.

Variant 1: All ASCII characters are en-/decrypted separately (no blocks are formed).

4. Encryption:

Text:	R	S	A		w	o	r	k	s	!
Number:	82	83	65	32	119	111	114	107	115	33

The letters are not combined¹⁰⁷!

¹⁰³Using CrypTool v1.3 you can solve this with the **Indiv.Procedures \ RSA Demonstration \ RSA Cryptosystem**.

¹⁰⁴See **Appendix E of this chapter** for the source code to factorise the number $J(n)$ using Mathematica and Pari-GP. Using CrypTool v1.3 you can solve this with the **Indiv.Procedures \ RSA Demonstration \ Factorisation of a Number**.

¹⁰⁵ e cannot, therefore, be 2, 127 or 251 ($65537 = 2^{16} + 1$).

In real life, $J(n)$ is not factorised but rather the Euclidean algorithm is used for the selected e to guarantee that $\gcd(d, J(n)) = 1$.

¹⁰⁶Other possible combinations of (e, d) include: (3, 170, 011), (5, 204, 013), (7, 36, 431).

¹⁰⁷For secure procedures we need large numbers that assume – as far as possible – all values up to $n-1$. If the possible value set for the numbers in the message is too small, even large prime numbers cannot make the procedure secure. An ASCII character is represented by 8 bits. If we want larger values we must combine several numbers. Two characters need 16 bits, whereby the maximum value that can be represented is 65536. The modulus n must then be greater than $2^{16} = 65536$. This is applied in variant 2. When the numbers are combined, the leading zeros are kept in binary notation (just as if we were to write all numbers with 3 digits in decimal notation above and were then to obtain the sequence 082 083, 065 032, 119 111, 114 107, 115 033).

Each character is encrypted using: $C = M^{65,537} \pmod{256,027}$ ¹⁰⁸:

```
212984 025546 104529 031692 248407
100412 054196 100184 058179 227433
```

5. Decryption:

Cipher text:

```
212984 025546 104529 031692 248407
100412 054196 100184 058179 227433
```

Each character is decrypted using: $M \equiv C^{231,953} \pmod{256,027}$:

```
82 83 65 32 119 111 114 107 115 33
```

Variant 2: The ASCII characters are en-/decrypted two at a time as blocks.

In variant 2 the block formation is done in two different sub-variants: (4./5. and 4'./5').

```
Text:   R   S   A           w   o   r   k   s   !
Number: 82  83  65  32  119 111 114 107 115 33
```

4. Encryption:

Blocks are formed¹⁰⁹ (each ASCII character is encoded into a 8 digit binary number below):

```
21075 16672 30575 29291 29473110
```

Each block is encrypted using: $C \equiv M^{65,537} \pmod{256,027}$ ¹¹¹:

```
158721 137346 37358 240130 112898
```

5. Decryption:

Cipher text:

```
158721 137346 37358 240130 112898
```

Each block is decrypted using: $M \equiv C^{231,953} \pmod{256,027}$:

```
21075 16672 30575 29291 29473
```

4'. Encryption:

Blocks are formed: (each ASCII character is encoded into a 3 digit decimal number below):

¹⁰⁸See [Appendix E of this chapter](#) for the source code for RSA exponentiation using Mathematica and Pari-GP.

¹⁰⁹

	binary representation	decimal representation
01010010, 82	01010010 01010011	=21075
01010011, 83		
01000001, 65	01000001 00100000	=16672
00100000, 32		
01110111, 119	01110111 01101111	=30575
01101111, 111		
01110010, 114	01110010 01101011	=29291
01101011, 107		
01110011, 115	01110011 00100001	=29473
00100001, 33:		

¹¹⁰Using CrypTool v1.3 you can solve this with the **Indiv.Procedures \ RSA Demonstration \ RSA Cryptosystem** with the following options: all 256 ASCII characters, b-adic, block length 2 and decimal representation.

¹¹¹See [Appendix E of this chapter](#) for the source code for RSA exponentiation using Mathematica and Pari-GP.

82083 65032 119111 114107 115033¹¹²

Each block is encrypted using: $C \equiv M^{65,537} \pmod{256,027}$ ¹¹³:

198967 051405 254571 115318 014251

5'. Decryption:

Cipher text:

198967 051405 254571 115318 014251

Each block is decrypted using: $M \equiv C^{2,473} \pmod{67,519}$:

82083 65032 119111 114107 115033

3.13.4 A small RSA cipher challenge (1)

The task is taken from [Stinson1995, Exercise 4.6]: The pure solution has been published by Prof. Stinson at <http://www.cacr.math.uwaterloo.ca/~dstinson/solns.html>.¹¹⁴

However, it is not the result that is important here but rather the individual steps of the solution, that is, the explanation of the cryptanalysis¹¹⁵:

Two samples of RSA cipher text are presented in Tables 4.1 and 4.2. Your task is to decrypt them. The public parameters of the system are

$n = 18,923$ and $e = 1,261$ (for Table 4.1) and

$n = 31,313$ and $e = 4,913$ (for Table 4.2).

This can be accomplished as follows. First, factor n (which is easy because it is so small). Then compute the exponent d from $J(n)$, and, finally, decrypt the cipher text. Use the square-and-multiply algorithm to exponentiate modulo n .

In order to translate the plaintext back into ordinary English text, you need to know how alphabetic characters are “encoded” as elements in \mathbb{Z}_n . Each element of \mathbb{Z}_n represents three alphabetic characters as in the following examples:

$$\text{DOG} \mapsto 3 * 26^2 + 14 * 26 + 6 = 2,398$$

$$\text{CAT} \mapsto 2 * 26^2 + 0 * 26 + 19 = 1,371$$

$$\text{ZZZ} \mapsto 25 * 26^2 + 25 * 26 + 25 = 17,575.$$

You will have to invert this process as the final step in your program.

The first plaintext was taken from “The Diary of Samuel Marchbanks”, by Robertson Davies, 1947, and the second was taken from “Lake Wobegon Days”, by Garrison Keillor, 1985.

¹¹²The RSA encryption works correctly with the modulus $n = 256,027$ because each ASCII block of two characters will be encoded into a number that is smaller or equal than the number 255,255.

¹¹³See [Appendix E of this chapter](#) for the source code for RSA exponentiation using Mathematica and Pari-GP.

¹¹⁴or <http://bibd.unl/~stinson/solns.html>.

¹¹⁵The method of solving the problem is outlined in the scenario of the online help to CrypTool and in the presentation on the website. If anyone sends us a well prepared exact method of solving the problem, we would be pleased to include it in the documentation.

TABLE 4.1¹¹⁶: RSA cipher text

12423	11524	7243	7459	14303	6127	10964	16399
9792	13629	14407	18817	18830	13556	3159	16647
5300	13951	81	8986	8007	13167	10022	17213
2264	961	17459	4101	2999	14569	17183	15827
12693	9553	18194	3830	2664	13998	12501	18873
12161	13071	16900	7233	8270	17086	9792	14266
13236	5300	13951	8850	12129	6091	18110	3332
15061	12347	7817	7946	11675	13924	13892	18031
2620	6276	8500	201	8850	11178	16477	10161
3533	13842	7537	12259	18110	44	2364	15570
3460	9886	8687	4481	11231	7547	11383	17910
12867	13203	5102	4742	5053	15407	2976	9330
12192	56	2471	15334	841	13995	17592	13297
2430	9741	11675	424	6686	738	13874	8168
7913	6246	14301	1144	9056	15967	7328	13203
796	195	9872	16979	15404	14130	9105	2001
9792	14251	1498	11296	1105	4502	16979	1105
56	4118	11302	5988	3363	15827	6928	4191
4277	10617	874	13211	11821	3090	18110	44
2364	15570	3460	9886	9988	3798	1158	9872
16979	15404	6127	9872	3652	14838	7437	2540
1367	2512	14407	5053	1521	297	10935	17137
2186	9433	13293	7555	13618	13000	6490	5310
18676	4782	11374	446	4165	11634	3846	14611
2364	6789	11634	4493	4063	4576	17955	7965
11748	14616	11453	17666	925	56	4118	18031
9522	14838	7437	3880	11476	8305	5102	2999
18628	14326	9175	9061	650	18110	8720	15404
2951	722	15334	841	15610	2443	11056	2186

¹¹⁶The numbers of this table can be worked with via Copy and Paste.

TABLE 4.2¹¹⁷: RSA cipher text

6340	8309	14010	8936	27358	25023	16481	25809
23614	7135	24996	30590	27570	26486	30388	9395
27584	14999	4517	12146	29421	26439	1606	17881
25774	7647	23901	7372	25774	18436	12056	13547
7908	8635	2149	1908	22076	7372	8686	1304
4082	11803	5314	107	7359	22470	7372	22827
15698	30317	4685	14696	30388	8671	29956	15705
1417	26905	25809	28347	26277	7897	20240	21519
12437	1108	27106	18743	24144	10685	25234	30155
23005	8267	9917	7994	9694	2149	10042	27705
15930	29748	8635	23645	11738	24591	20240	27212
27486	9741	2149	29329	2149	5501	14015	30155
18154	22319	27705	20321	23254	13624	3249	5443
2149	16975	16087	14600	27705	19386	7325	26277
19554	23614	7553	4734	8091	23973	14015	107
3183	17347	25234	4595	21498	6360	19837	8463
6000	31280	29413	2066	369	23204	8425	7792
25973	4477	30989					

3.13.5 A small RSA cipher challenge (2)

The following task is a corrected version from the excellent book written by Prof. Yan [Yan2000, Example 3.3.7, p. 318]. However, it is not the result that is important here but rather the individual steps of the solution, that is, the explanation of the cryptanalysis¹¹⁸.

There are three tasks with completely different degrees of difficulty here. In each case we know the cipher text and the public key (e, n) :

- (a) Known plaintext: find the secret key d using the additionally known original message.
- (b) Cipher text only: find d and the plaintext.
- (c) Calculate the RSA modulus, in other words factorisation (with no knowledge of the message).

¹¹⁷The numbers of this table are in the online-help “Example illustrating the RSA demonstration” of CrypTool.

¹¹⁸The method of solving the problem is outlined in the scenario of the online help to CrypTool and in the CrypTool presentation. If anyone sends us a well prepared exact method of solving the problem, we would be pleased to include it in the documentation.

$n = 63978486879527143858831415041$, $e = 17579$

Message¹¹⁹:

1401202118011200,
1421130205181900,
0118050013010405,
0002250007150400

Cipher:

45411667895024938209259253423,
16597091621432020076311552201,
46468979279750354732637631044,
32870167545903741339819671379

Comments:

The original message consisted of a sentence containing 31 characters (coded with the capital letters alphabet from section 3.13.2). Each group of 16 decimal numbers is then combined to form one number (the last number is filled with zeros). These numbers are raised to the power of e .

When you decrypt the message you must fill the calculated numbers with leading zeros in order to obtain plaintext.

This needs to be stressed because the type of padding is extremely important during implementation and standardisation for interoperable algorithms.

¹¹⁹The numbers of this table are in the online help “Example illustrating the RSA demonstration” of CrypTool.

References

- [Agrawal2002] M. Agrawal, N. Kayal, N. Saxena,
PRIMES in P, August 2002
<http://www.cse.iitk.ac.in/news/primalty.html>.
- [Bartholome1996] A. Bartholome, J. Rung, H. Kern,
Zahlentheorie für Einsteiger, Vieweg 1995, 2nd edition 1996.
- [Bauer1995] Friedrich L. Bauer,
Entzifferte Geheimnisse, Springer, 1995.
- [Bauer2000] Friedrich L. Bauer,
Decrypted Secrets, Springer 1997, 2nd edition 2000.
- [Bernstein2001] D. J. Bernstein,
Circuits for integer factorization: a proposal,
<http://cr.yp.to/papers/nfscircuit.ps>
<http://cr.yp.to/djb.html>.
- [Beutelspacher1996] Albrecht Beutelspacher,
Kryptologie, Vieweg 1987, 5th edition 1996.
- [Bourseau2002] F. Bourseau, D. Fox, C. Thiel,
Vorzüge und Grenzen des RSA-Verfahrens,
In: Datenschutz und Datensicherheit (DuD) 26/2002, pp 84-89 (see www.dud.de).
- [Brands2002] Gilbert Brands,
Verschlüsselungsalgorithmen – Angewandte Zahlentheorie rund um Sicherheitsprotokolle,
Vieweg, 2002.
- [Buchmann1999] Johannes Buchmann,
Einführung in die Kryptographie, Springer, 1999.
- [Buhler1993] J.P. Buhler, H.W. Lenstra, C. Pomerance,
Factoring integers with the number field sieve,
In: A.K. Lenstra, H.W. Lenstra (Hrsg.): The Development of the Number Field Sieve,
Lecture Notes in Mathematics, Vol. 1554, Springer, Heidelberg 1993, pp 50–94.
- [Eckert2003] Claudia Eckert,
IT-Sicherheit: Konzepte-Verfahren-Protokolle, Oldenbourg 2001, 2nd edition 2003.
- [Ertel2001] Wolfgang Ertel,
Angewandte Kryptographie, Fachbuchverlag Leipzig FV 2001.
- [GISA2002] GISA (German Information Security Agency),
Recommendation for key length selection, Bonn, Sep. 2002,
<http://www.bsi.bund.de/esig/basics/techbas/krypto/bund02v7.pdf>

A statement on these recommendations:
[http://www.secorvo.de/publikat/
stellungnahme-algorithmenempfehlung-020307.pdf](http://www.secorvo.de/publikat/stellungnahme-algorithmenempfehlung-020307.pdf)

- [Graham1994] Graham, Knuth, Patashnik,
Concrete Mathematics, a Foundation of Computer Science,
Addison Wesley 1989, 6th printing 1990.
- [Kippenhahn1997] Rudolph Kippenhahn,
Verschlüsselte Botschaften – Geheimschrift, Enigma und Chipkarte, Rowohlt, 1997.
- [Kippenhahn1999] Rudolph Kippenhahn,
Code Breaking – A History and Exploration, Constable, 1999.
- [Knuth1998] Donald E. Knuth,
The Art of Computer Programming, vol 2: Seminumerical Algorithms,
Addison-Wesley, 2nd edition 1998.
- [Lenstra1993] A. Lenstra, H. Lenstra:
The development of the Number Field Sieve,
Lecture Notes in Mathematics 1554, Springer, New York 1993
- [Lenstra2002] Arjen K. Lenstra, Adi Shamir, Jim Tomlinson, Eran Tromer,
Analysis of Bernstein's Factorization Circuit,
<http://www.cryptosavvy.com/mesh.pdf>
- [Menezes2001] Alfred J. Menezes, Paul C. van Oorschot, Scott A. Vanstone
Handbook of Applied Cryptography, CRC Press 1997, 5th printing 2001.
- [Pfleeger1997] Charles P. Pfleeger,
Security in Computing, Prentice-Hall, 2nd edition 1997.
- [Pomerance1984] C. Pomerance,
The quadratic sieve factoring algorithm,
In: G.R. Blakley, D. Chaum (Hrsg.): *Proceedings of Crypto '84*, LNCS 196, Springer
Berlin 1995, pp 169-182.
- [RSA Security 2002] RSA Security,
Has the RSA algorithm been compromised as a result of Bernstein's Paper?,
April 8th, 2002,
<http://www.rsasecurity.com/>.
- [Schneier1996] Bruce Schneier,
Applied Cryptography, Protocols, Algorithms, and Source Code in C,
Wiley and Sons, 2nd edition 1996.

- [Schwenk2002] Jörg Schwenk,
Sicherheit und Kryptographie im Internet, Vieweg 2002.
- [Sedgewick1990] Robert Sedgewick,
Algorithms in C, Addison-Wesley, 1990.
- [Shamir2003] Adi Shamir, Eran Tromer,
Factoring Large Numbers with the TWIRL Device, Januar 2003,
<http://www.wisdom.weizmann.ac.il/~tromer/>.
- [Silverman2000] Robert D. Silverman:
A Cost-Based Security Analysis of Symmetric and Asymmetric Key Lengths
In: RSA Laboratories Bulletin, No. 13, April 2000, p. 1-22
- [Stinson1995] Douglas R. Stinson,
Cryptography - Theory and Practice, CRC Press, 1995.
- [Weis2003] Rüdiger Weis, Stefan Lucks, Andreas Bogk,
Sicherheit von 1024 bit RSA-Schlüsseln gefährdet,
In: Datenschutz und Datensicherheit (DuD) 6/2003, pp 360-362 (see www.dud.de)
The article explains details about the TWIRL device.
- [Welschenbach2001] Welschenbach, Michael,
Kryptographie in C und C++, Springer 2001.
- [Wiles1994] Wiles, Andrew,
Modular elliptic curves and Fermat's Last Theorem,
In: Annals of Mathematics 141 (1995).
- [Wolfenstetter1998] Albrecht Beutelspacher, Jörg Schwenk, Klaus-Dieter Wolfenstetter,
Moderne Verfahren in der Kryptographie, Vieweg 1995, 2nd edition 1998.
- [Yan2000] Song Y. Yan,
Number Theory for Computing, Springer, 2000.

Web links

1. Ron Knott's Fibonacci page,
Here, everything revolves around Fibonacci numbers.
<http://www.mcs.surrey.ac.uk/personal/R.Knott/Fibonacci/fib.html>
2. CrypTool (Version 1.3.04, 2003),
Freeware to illustrate cryptography
<http://www.cryptool.de>,
<http://www.cryptool.org>,
<http://www.cryptool.com>
3. Mathematica,
Commercial mathematics package
<http://www.wolfram.com>
4. LiDIA,
Extensive library containing number-theory functions and the LC interpreter
<http://www.informatik.tu-darmstadt.de/TI/LiDIA>
5. BC,
Interpreter with number-theory functions
<http://www.maths.uq.edu.au/~krm/gnubc.html>
6. Pari-GP,
Excellent, fast, free interpreter with number theoretical functions
<http://www.parigp-home.de> and <http://www.parigp-home.com>
7. Only after I had completed this article, did I come across the website of Mr. Münchenbach, which interactively and didactically uses elementary number theory to provide a sophisticated description of the fundamental mathematical thought processes. It was created for a teaching project in the 11th grade of the technical grammar school (unfortunately only available in German):
<http://www.hydrargyrum.de/kryptographie>
8. Once again only after finishing this I happened upon the web site of Mr. Wagner, who is responsible for the development of the curriculum of computer science in one of the German federal states (Länder). Here you can get hold of a collection of texts and (Java-)programs (available only in German):
<http://www.hom.saar.de/~awa/kryptolo.htm>
9. GISA,
German Information Security Agency
<http://www.bsi.bund.de>
10. Factorisation records and challenges,
<http://www.crypto-world.com/>

<http://www.crypto-world.com/FactorWorld.html>, page by Scott Contini
<http://www.loria.fr/~zimmerma/records/factor.html>

<http://www.uni-bonn.de/Aktuelles/Pressemitteilungen/pm02/pm035-02.html>
http://www.ercim.org/publication/Ercim_News/enw49/franke.html, 2002-01
<http://www.loria.fr/~zimmerma/records/rsa160>

<http://www.rsasecurity.com/rsalabs/challenges/factoring/numbers.html>

11. The Cunningham Project,
<http://www.cerias.purdue.edu/homes/ssw/cun/>

Acknowledgments

I would like to take this opportunity to thank

- Mr. Henrik Koy for making many very useful suggestions, for the very constructive proof-reading this article and for helping with TeX.
Mr. Koy designed and developed in his leisure time the functions and the complex dialog box of the RSA cryptosystem, which enables you to execute the RSA samples of this article.
- Jörg Cornelius Schneider for his TeX support und for the many cases where he helped when facing programming or design problems.
- Dr. Illies for pointing me to Pari-GP .

Appendix A: the greatest common divisor (gcd) of whole numbers and the two Algorithms of Euclid

1. The greatest common divisor of two natural numbers a and b is an important value that can be calculated very quickly. Here we make use of the fact that if a number c divides the numbers a and b (i.e. there exists an a' and a b' such that $a = a' * c$ and $b = b' * c$), then c also divides the remainder r of a/b . In short notation we can write: If c divides a and b it follows that c divides $r = a - \lfloor a/b \rfloor * b$ ¹²⁰. As the latter statement is valid for each common divisor c of a and b it follows that:

$$\gcd(a, b) = \gcd(a - \lfloor a/b \rfloor * b, b).$$

Using this information, the algorithm for calculating the gcd of two numbers can be written as follows (in pseudo code):

```

INPUT: a, b != 0
1. if ( a < b ) then  x = a; a = b; b = x; // Swap a and b (a > b)
2. a = a - int(a/b) * b           // a is smaller than b, the
                                   // gcd(a, b) is unchanged
3. if ( a != 0 ) then goto 1.     // a falls after each step and
                                   // the algorithm ends when a==0.
OUTPUT "gcd(a,b) = " b          // b is the gcd of the original a and b

```

2. However, to other relationships can be derived from the gcd: For this, we need the set of equations for a and b :

$$\begin{aligned} a &= 1 * a + 0 * b \\ b &= 0 * a + 1 * b, \end{aligned}$$

or, in matrix notation:

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} * \begin{pmatrix} a \\ b \end{pmatrix}.$$

We summarise this information in the extended matrix:

$$\left(\begin{array}{cc|cc} a & & 1 & 0 \\ b & & 0 & 1 \end{array} \right)$$

If we apply the above gcd algorithm to this matrix, we obtain the extended gcd algorithm:

¹²⁰The Gauss bracket $\lfloor x \rfloor$ of a real number x is defined via: $\lfloor x \rfloor$ is the next integer less or equal x .

INPUT: $a, b \neq 0$

0. $x_{1,1} := 1, x_{1,2} := 0, x_{2,1} := 0, x_{2,2} := 1$
1. $\left(\begin{array}{c|cc} a & x_{1,1} & x_{1,2} \\ b & x_{2,1} & x_{2,2} \end{array} \right) := \left(\begin{array}{c|cc} 0 & 1 & \\ 1 & -\lfloor a/b \rfloor * b & \end{array} \right) * \left(\begin{array}{c|cc} a & x_{1,1} & x_{1,2} \\ b & x_{2,1} & x_{2,2} \end{array} \right).$
2. if ($b \neq 0$) then goto 1.

OUTPUT: “gcd(a, b) = $a * x + b * y$: ”, “gcd(a, b) = ” b , “ $x =$ ” $x_{2,1}$, “ $y =$ ” $x_{2,2}$

Since this algorithm only performs linear transformations, the same equations always apply

$$\begin{aligned} a &= x_{1,1} * a + x_{1,2} * b \\ b &= x_{2,1} * a + x_{2,2} * b, \end{aligned}$$

and we have the extended gcd equation at the end of the algorithm¹²¹:

$$\gcd(a, b) = a * x_{2,1} + b * x_{2,2}.$$

Example:

Using the extended gcd we can determine for $e = 37$ the multiplicative inverse number d to modulo 3588 (i.e. $37 * d \equiv 1 \pmod{3588}$):

0. $\left(\begin{array}{c|cc} 3588 & 1 & 0 \\ 37 & 0 & 1 \end{array} \right)$
1. $\left(\begin{array}{c|cc} 37 & 1 & 0 \\ 36 & 0 & -96 \end{array} \right) = \left(\begin{array}{c|cc} 0 & 1 & \\ 1 & -(\lfloor 3588/36 \rfloor = 96) * 37 & \end{array} \right) * \left(\begin{array}{c|cc} 3588 & 1 & 0 \\ 37 & 0 & 1 \end{array} \right).$
2. $\left(\begin{array}{c|cc} 36 & 1 & -96 \\ 1 & -1 & 97 \end{array} \right) = \left(\begin{array}{c|cc} 0 & 1 & \\ 1 & -(\lfloor 37/36 \rfloor = 1) * 36 & \end{array} \right) * \left(\begin{array}{c|cc} 37 & 1 & 0 \\ 36 & 0 & -96 \end{array} \right).$
3. $\left(\begin{array}{c|cc} 1 & -1 & 97 \\ 0 & 37 & -3588 \end{array} \right) = \left(\begin{array}{c|cc} 0 & 1 & \\ 1 & -(\lfloor 36/1 \rfloor = 36) * 1 & \end{array} \right) * \left(\begin{array}{c|cc} 36 & 1 & -96 \\ 1 & -1 & 97 \end{array} \right).$

OUTPUT:

gcd(37, 3588) = $a * x + b * y$: gcd(37, 3588) = 1, $x = -1$, $y = 97$.

Thus

- (a) 37 and 3588 are relatively prime (37 has an inverse modulo 3588).
- (b) $37 * 97 = (1 * 3588) + 1$ in other words $37 * 97 \equiv 1 \pmod{3588}$. and therefore the number 97 is the multiplicative inverse to 37 modulo 3588.

¹²¹By termination of the gcd algorithm, the program variables a and b contain the values $a = 0$ and $b = \gcd(a, b)$. Please keep in mind, that the program variables are different to the numbers a and b and that they are only relevant for the scope of the algorithm.

Appendix B: Forming closed sets

The property of closeness is always defined in relation to an operation in a set. The following shows how to construct the “closed set” G with respect to the operation $+$ (mod 8) for a given initial set G_0 :

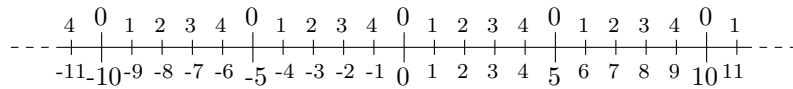
$$\begin{aligned}
 G_0 &= \{2, 3\} \text{ addition of the numbers in } G_0 \text{ determines further numbers :} \\
 &\quad 2 + 3 \equiv 5 \pmod{8} = 5 \\
 &\quad 2 + 2 \equiv 4 \pmod{8} = 4 \\
 &\quad 3 + 3 \equiv 6 \pmod{8} = 6 \\
 G_1 &= \{2, 3, 4, 5, 6\} \text{ addition of the numbers in } G_1 \text{ determines :} \\
 &\quad 3 + 4 \equiv 7 \pmod{8} = 7 \\
 &\quad 3 + 5 \equiv 8 \pmod{8} = 0 \\
 &\quad 3 + 6 \equiv 9 \pmod{8} = 1 \\
 G_2 &= \{0, 1, 2, 3, 4, 5, 6, 7\} \text{ addition of the numbers in } G_2 \text{ does not extend the set!} \\
 G_3 &= G_2 \text{ we say : } G_2 \text{ is closed for addition (mod 8).}
 \end{aligned}$$

End of forming a closed set.

Appendix C: Comments on modulo subtraction

Comment on subtraction modulo 5: $2 - 4 \equiv -2 \equiv 3 \pmod{5}$. It is therefore not true modulo 5 that $-2 = 2$! People often make the mistake of equating this. You can show this clearly if you place the permutation $(0, 1, 2, 3, 4)$ in \mathbb{Z}_5 , for example from -11 to $+11$, over the range of numbers in \mathbb{Z} .

range of numbers modulo 5



range of numbers in \mathbb{Z}

Appendix D: Base representation of numbers, estimation of length of digits

For a given number z one may ask how to represent such a number. In general we use representations like $z = 2374$ or $z = \sqrt{2}$. The second number consists of an infinite number of digits and therefore it can never be described precisely by the first representation. In this case the number must be rounded.

We represent numbers usually in the decimal system (base 10). Computers are working with the binary representation of numbers — only for the display numbers are represented in decimal or sometimes hexadecimal (base 16).

This appendix describes how to generate arbitrary base representations of a positive integers and how to determine the number of required digits via the logarithm function.

b -adic sum representation of positive integers

Given base b , each positive integer z can be represented as a b -adic sum

$$z = a_n b^n + a_{n-1} b^{n-1} + \cdots + a_1 b + a_0,$$

where $a_i \in \{0, 1, \dots, b-1\}$, $i = 0, 1, \dots, n$ are called *digits*. For this sum, it follows that for arbitrary digits a_0, a_1, \dots, a_n we have $b^{n+1} > a_n b^n + a_{n-1} b^{n-1} + \cdots + a_1 b + a_0$. On the other hand there exist a_0, a_1, \dots, a_n (namely $a_i = b-1$ for $i = 0, \dots, n$) following that $b^{n+1} - 1 \leq a_n b^n + a_{n-1} b^{n-1} + \cdots + a_1 b + a_0$ (using these inequalities it can be shown that each positive integer can be represented by a b -adic sum).

By writing the digits $a_n a_{n-1} \cdots a_1 a_0$ in a row directly after each other (without the b^i) the usual writing for numbers comes to hand.

Examples:

Base $b = 10$: $10278 = 1 \cdot 10^4 + 0 \cdot 10^3 + 2 \cdot 10^2 + 7 \cdot 10^1 + 8$ und

Base $b = 16$: $FE70A = 15 \cdot 16^4 + 14 \cdot 16^3 + 7 \cdot 16^2 + 0 \cdot 16^1 + 10$.

Number of digits to represent a positive integer

For a positive integer z the length of the b -adic representation can be determined via the following steps. Starting from the inequality $b^{n+1} > z \geq b^n$ we have — after applying the logarithm function on basis b ¹²² — $n+1 > \log_b z \geq n$. Therefore we have $n = \lfloor \log_b z \rfloor$ ¹²³. We call $l_b(z)$ the *number of required digits to represent the number z on the base b* . We have

$$l_b(z) := \lfloor \log_b z \rfloor + 1.$$

¹²²Applying the logarithm formula on base b and b' we have $\log_b z = \log_{b'} z / \log_{b'}(b)$. It is therefore easy using e.g. logarithm tables for the base $b' = 10$ to compute the logarithm of base $b = 2$.

¹²³The function $\lfloor x \rfloor$ determines the next integer smaller than x (in case $x \geq 0$ the digits after the decimal point are truncated).

Example 1: We compute for the decimal number $z = 234$ the hexadecimal representation (number base $b = 16$)

$$l_{16}(z) = \lfloor \log_{16}(z) \rfloor + 1 = \lfloor \ln(z)/\ln(16) \rfloor + 1 = \lfloor 1.96... \rfloor + 1 = 1 + 1 = 2.$$

Example 2: We compute for the decimal number $z = 234$ the binary representation (number base $b = 2$)

$$l_2(z) = \lfloor \log_2(z) \rfloor + 1 = \lfloor \ln(z)/\ln(2) \rfloor + 1 = \lfloor 7.87... \rfloor + 1 = 7 + 1 = 8.$$

Algorithm to compute the base representation

Given the number z one can compute the base b representation of z using the following algorithm

```

input:  $z, b$ 
 $n := 0, z' := z$ 
while  $z' > 0$  do
     $a_n := z' \bmod b,$ 
     $z' := \lfloor z'/b \rfloor$ 
     $n := n + 1$ 
end do
output:  $a_n a_{n-1} \cdots a_1 a_0$  in base  $b$  representation.

```

Example 1: The integer $z = 234$ on the number base 10 will be transformed into the hex representation via $a_0 = 234 \bmod 16 = 10 = A$, $234/16 = 15 = E$, $a_1 = 14 \bmod 16 = E$ and therefore we have EA .

Example 2: The binary number $z = 1000100101110101$ is computed into the decimal representation via the following steps:

```

1000100101110101 = 1001 (mod 1010)  $\implies a_0 = 9$ ,  $1000100101110101/1010 = 110110111110$ 
110110111110 = 1000 (mod 1010)  $\implies a_1 = 8$ ,  $110110111110/1010 = 101011111$ 
101011111 = 1 (mod 1010)  $\implies a_2 = 1$ ,  $10101111/1010 = 100011$ 
100011 = 101 (mod 1010)  $\implies a_3 = 5$ ,  $100011/1010 = 1$ 
11 = 11 (mod 1010)  $\implies a_4 = 3$ 
therefore  $z = 35189$ .

```

Appendix E: Examples using Mathematica and Pari-GP

This appendix gives you the source code to compute the tables and examples using Mathematica or the free software Pari-GP.

Multiplication table modulus m

The multiplication tables modulo $m = 17$ for $a = 5$ and $a = 6$ on page 57 can be computed in Mathematica with the following commands:

```
m = 17; iWidth = 18; iFactor1 = 5; iFactor2 = 6;
Print[ 'i ', Table[ i, {i, 1, iWidth} ] ];
Print[ iFactor1, '*i ', Table[ iFactor1*i, {i, 1, iWidth} ] ];
Print[ 'Remainder ', Table[ Mod[iFactor1*i, m], {i, 1, iWidth} ] ];
Print[ iFactor2, '*i ', Table[ iFactor2*i, {i, 1, iWidth} ] ];
Print[ 'Remainder ', Table[ Mod[iFactor2*i, m], {i, 1, iWidth} ] ];
```

Pari-GP computes the tables via:

```
m=17; iWidth=18; iFactor1=5; iFactor2=6;
matrix(1,iWidth, x,y, iFactor1*y) yields
[5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90]
matrix(1,iWidth, x,y, (iFactor1*y)%m ) yields
[5 10 15 3 8 13 1 6 11 16 4 9 14 2 7 12 0 5]
```

Note: Pari-GP generates when using the Mod function compound Mod objects, which are displayed as shown below:

```
matrix(1,iWidth, x,y, Mod(iFactor1*y, m))
[Mod(5, 17) Mod(10, 17) Mod(15, 17) Mod(3, 17) Mod(8, 17) Mod(13, 17) Mod(1, 17)
Mod(6, 17) Mod(11, 17) Mod(16, 17) Mod(4, 17) Mod(9, 17) Mod(14, 17) Mod(2, 17)
Mod(7, 17) Mod(12, 17) Mod(0, 17) Mod(5, 17)]
```

From a Mod object you can get back the components with the component or lift function:

```
component(Mod(5, 17),1) → 17
component(Mod(5, 17),2) → 5
component(Mod(17,5), 1) → 5
component(Mod(17,5), 2) → 2
lift(Mod(17,5)) → 2
```

The other multiplication table examples modulo 13 and modulo 12 on page 57 can be computed by replacing $m=17$ with $m=13$ and $m=12$ respectively.

Fast exponentiation

The fast exponentiation modulo m belongs to the built in functions of Mathematica and Pari-GP. Using those programs you can comprehend the idea of the square and multiply method. With Mathematica you can compute the exponentiations of the example on page 60 as follows:

```
Mod[{87^43, 87^2, 87^4, 87^8, 87^16, 87^32}, 103] = {85, 50, 28, 63, 55, 38}.
```

and in Pari-GP the syntax is:

```
Mod([87^43,87^2,87^4,87^8,87^16,87^32],103)
```

Multiplicative order and primitive roots

The order $ord_m(a)$ of a number a in the multiplicative group Z_m^* is the smallest number $i \geq 1$, for which $a^i \equiv 1 \pmod{m}$ holds. For the example on page 68 you can make Mathematica print all exponentiations $a^i \pmod{11}$ using the following syntax:

```
m=11; Table[ Mod[a^i, m], {a, 1, m-1}, {i, 1, m-1} ]
```

Equivalent Pari-GP syntax:

```
m=11; matrix(10,10, x,y, (x^y)%m )
```

The table on page 69 gives examples for the order modulo 45 $ord_{45}(a)$ and the Euler number $J(45)$. Mathematica can be used to create this table with the following program (please note that Print cannot be used inside of Do-loops and each Print outputs a newline).

```
m = 45;
Do[ Print[ Table[ Mod[a^i, m], {i, 1, 12} ],
'', '', MultiplicativeOrder[a, m, 1],
'', '', EulerPhi[m] ],
{a, 1, 12} ];
```

Here is the corresponding Pari-GP syntax:

```
m=45;
matrix(12,14, x,y,
    if( y<=12, (x^y)%m,
    if( y==13, if( gcd(x,m)==1, znorder(Mod(x,m)), "--"),
    eulerphi(m))))
```

$znorder(\text{Mod}(x,m))$ can only be calculated if x is relatively prime to m , which can be checked with $\text{gcd}(x,m)$.

Performance can be improved by using $\text{Mod}(x,m)^y$ instead of $(x^y)\%m$.

Loops are also supported by Pari-GP. When you remove the table formatting the result looks like this:

```
for( x=1,12,
  for(y=1,12, print(Mod(x^y,m)));
  if(gcd(x,m)==1, print(znorder(Mod(x,m))), print("--"));
  print(eulerphi(m)))
```

The third example on page 70 displays exponentiations $a^i \bmod 46$ as well as the order $\text{ord}_{46}(a)$. Mathematica can create this table with the following loop:

```
m = 46;
Do[ Print[ Table[ Mod[a^i, m], {i, 1, 23} ] ,
'', '', MultiplicativeOrder[a, m, 1]
{a, 1, 23} ] ];
```

In Pari-GP the syntax looks like this:

```
m=46;
matrix(23,24, x,y,
  if( y<=23, (x^y)%m,
  if( y==24, if( gcd(x,m)==1, znorder(Mod(x,m)), "--"))))
```

RSA examples

This section list the source code of the RSA examples in section 3.13 (“The RSA procedure with actual numbers”) using Mathematica and Pari-GP syntax.

Example on page 88.

The RSA exponentiation $M^{37} \bmod 3713$ on message $M = 120$ can be calculated in Mathematica like this: `PowerMod[120, 37, 3713]`.

Here is the corresponding Pari-GP syntax:

`Mod(120,3713)^37` or `Mod(120^37,3713)`.

Example on page 89.

The factorisation of $J(256, 027) = 255,016 = 2^3 * 127 * 251$ can be calculated with Mathematica like this: `FactorInteger[255016] = {{2,3}, {127,1}, {251,1}}`.

Pari-GP does the same with:

`factor(255016)`.

Example on page 89.

Mathematica can do RSA encryption with the command:

```
PowerMod[{82, 83, 65, 32, 119, 111, 114, 107, 115, 33}, 65537, 256027]}
```

Pari-GP needs the following syntax:

```
vecextract( [Mod(82,256027)^65537, Mod(83,256027)^65537, Mod(65,256027)^65537,  
Mod(32,256027)^65537, Mod(119,256027)^65537, ...])
```

Remarks on using Mod in Pari-GP:

$\text{Mod}(82, 256027)^{65537}$ is much faster than

- $\text{Mod}(82^{65537}, 256027)$ and
- $(82^{65537}) \% 256027$.

Example on page 90.

Mathematica can do RSA encryption with the following command:

```
PowerMod[{21075, 16672, 30575, 29291, 29473}, 65537, 256027]
```

The same calculation with Pari-GP:

```
vecextract( [Mod(21075,256027)^65537, Mod(16672,256027)^65537,  
Mod(30575,256027)^65537, Mod(29291,256027)^65537,  
Mod(29473,256027)^65537], 31)
```

Example on page 91.

RSA encryption using Mathematica:

```
PowerMod[{82083, 65032, 119111, 114107, 115033}, 65537, 256027]
```

RSA encryption with Pari-GP:

```
vecextract( [Mod(82083,256027)^65537, Mod(65032,256027)^65537,  
Mod(119111,256027)^65537, Mod(114107,256027)^65537,  
Mod(115033,256027)^65537], 31)
```

Appendix F: List of the formulated definitions and theorems

	Short description	Page
Definition 3.1	prime numbers	48
Definition 3.2	composite numbers	48
Theorem 3.1	factors of composite numbers	49
Theorem 3.2	1. fundamental theorem of number theory	49
Definition 3.3	divisibility	50
Definition 3.4	remainder class r modulo m	50
Definition 3.5	congruent	51
Theorem 3.3	congruence with difference	51
Theorem 3.4	multiplicative inverse	56
Theorem 3.5	exhaustive permutation	57
Theorem 3.6	power mod m	59
Definition 3.6	\mathbb{Z}_n	61
Definition 3.7	\mathbb{Z}_n^*	62
Theorem 3.7	multiplicative inverse in \mathbb{Z}_n^*	63
Definition 3.8	Euler function $J(n)$	64
Theorem 3.8	$J(p)$	64
Theorem 3.9	$J(p * q)$	64
Theorem 3.10	$J(p_1 * \cdots * p_k)$	64
Theorem 3.11	$J(p_1^{e_1} * \cdots * p_k^{e_k})$	64
Theorem 3.12	little Fermat	65
Theorem 3.13	Euler-Fermat theorem	65
Definition 3.9	multiplicative order $\text{ord}_m(a)$	68
Definition 3.10	primitive root of m	68
Theorem 3.14	exhausting of all possible values	70

4 The Mathematical Ideas behind Modern Cryptography

(Oyono R. / Esslinger B., Sep. 2000, Updates Nov. 2000, Feb. 2003)

4.1 One way functions with trapdoor and complexity classes

A **one way function** is a function that can be calculated efficiently, but whose inverse is extremely complicated and practically impossible to calculate.

To put it more precisely: A one way function is a mapping f from a set X to a set Y , such that $f(x)$ can be calculated easily for each element x of X , whereas for (almost) every y from Y it is practically impossible to find an inverse image x (i.e. an x where $f(x) = y$).

An everyday example of a one way function is a telephone book: the function to be performed is to assign a name to the corresponding telephone number. This can be done easily due to the fact that the names are sorted alphabetically. However, the inverse function - assigning a name to a given number - is obviously difficult if you only have a telephone book available.

One way functions play a decisive role in cryptography. Almost all cryptographic terms can be rephrased using the term one way function. Let's take for example public key encryption (asymmetric cryptography):

Each subscriber T to the system is assigned a private key d_T and what is known as a public key e_T . These keys must have the following property (public key property):

For an opponent who knows the public key e_T , it is practically impossible to determine the private key d_T .

In order to construct useful public key procedures, therefore, we look for a one way function that is "easy" to calculate in one direction, but is "difficult" (practically impossible) to calculate in the other direction, provided that a particular piece of additional information (trapdoor) is not available. This additional piece of information allows the inverse to be found efficiently. Such functions are called **trapdoor one way functions**. In the above case, d_T is the trapdoor information.

In this process, we describe a problem as "easy" if it can be solved in polynomial time as a function of the length of the input. If the length of the input is n bits, then the time for calculating the function is proportional to n^a , where a is a constant. We say that the complexity of such problems is $O(n^a)$ [Landau- or Big-O notation].

If you compare two functions 2^n and n^a , where a is a constant, then there always exists a value for n , from which for all further n applies: $n^a < 2^n$. The function n^a has a lower complexity. Sample: for $a = 5$ the following applies: from the length $n = 23$, 2^n is greater than n^5 ; for further n 2^n clearly increases more quickly [($2^{22} = 4,194,304$, $22^5 = 5,153,632$), ($2^{23} = 8,388,608$, $23^5 = 6,436,343$), ($2^{24} = 16,777,216$, $24^5 = 7,962,624$)].

The term "practically impossible" is slightly less precise. In general, we can say that a problem cannot be solved efficiently, if the time required to solve it increases more quickly than the polynomial time as a function of the size of the input. If, for example, the length of the input is n bits and the time required for calculating the function is proportional to 2^n , then the following

currently applies: the function practically cannot be calculated for $n > 80$.

In order to develop a public key procedure that can be implemented in practice, it is therefore necessary to discover a suitable trapdoor one way function.

In order to tidy things up among this confusing multitude of possible problems and their complexities, we group problems with similar complexities into classes.

The most important complexity classes are the classes **P** and **NP**:

- The class **P**: This class contains those problems that can be solved in a polynomial amount of time.
- The class **NP**: The definition of this class doesn't look at the time required to solve a problem, but rather at the time required to verify a given solution. The class **NP** consists of those problems for which a given solution can be verified in a polynomial amount of time. Hereby, the term **NP** "non-deterministic" means polynomial and is based on a calculation model, i.e. on a computer that only exists in theory and can "guess" correct solutions non-deterministically then verify them in polynomial time.

The class **P** is contained in the class **NP**. A well-known unsolved problem is the question whether or not $\mathbf{P} \neq \mathbf{NP}$ is true, i.e. whether or not **P** is a true subset. An important property of the class **NP** is that it also contains what are known as "**NP**-complete" problems. These are problems that represent the class **NP** as follows: If a "good" algorithm for such a problem exists, then "good" algorithms exist for all problems from **NP**. In particular: if **P** only contained one complete problem, i.e. if a polynomial solution algorithm existed for this problem, then **P** would be equal to **NP**. In this sense, the **NP**-complete problems are the most difficult problems in **NP**.

Many cryptographic protocols are formed in such a way that the "good" subscribers only have to solve problems from **P**, whereas a perpetrator is faced with problems from **NP**.

Unfortunately, we do not yet know whether one way functions actually exist. However, we can prove that one way functions exist if and only if $\mathbf{P} \neq \mathbf{NP}$ [Balcazar1988, S.63].

Mathematicians have again and again claimed to have proven the equivalence, e.g.

http://www.geocities.com/st_busygin/clipat.html),

but so far the claims have always turned out to be false.

A number of algorithms have been suggested for public key procedures. In many cases - although they at first appeared promising - it was discovered that they could be solved polynomially. The most famous failed applicant is the knapsack with trapdoor, suggested by Ralph Merkle [Merkle1978].

4.2 Knapsack problem as a basis for public key procedures

4.2.1 Knapsack problem

You are given n objects G_1, \dots, G_n with the weights g_1, \dots, g_n and the values w_1, \dots, w_n . The aim is to carry away as much as possible in terms of value while restricted to an upper weight limit g . You therefore need to find a subset of $\{G_1, \dots, G_n\}$, i.e. $\{G_{i_1}, \dots, G_{i_k}\}$, so that $w_{i_1} + \dots + w_{i_k}$ is maximised under the condition $g_{i_1} + \dots + g_{i_k} \leq g$.

Such questions are called **NP**-complete problems (not deterministically polynomial) that are difficult to calculate.

A special case of the knapsack problem is:

Given the natural numbers a_1, \dots, a_n and g , find $x_1, \dots, x_n \in \{0, 1\}$ where $g = \sum_{i=1}^n x_i a_i$ (i.e. where $g_i = a_i = w_i$ is selected). This problem is also called a **0-1 knapsack problem** and is identified with $K(a_1, \dots, a_n; g)$.

Two 0-1 knapsack problems $K(a_1, \dots, a_n; g)$ and $K(a'_1, \dots, a'_n; g')$ are called congruent if two co-prime numbers w and m exist in such a way that

1. $m > \max\{\sum_{i=1}^n a_i, \sum_{i=1}^n a'_i\}$,
2. $g \equiv wg' \pmod{m}$,
3. $a_i \equiv wa'_i \pmod{m}$ for all $i = 1, \dots, n$.

Comment: Congruent 0-1 knapsack problems have the same solutions. No quick algorithm is known for clarifying the question as to whether two 0-1 knapsack problems are congruent.

A 0-1 knapsack problem can be solved by testing the 2^n possibilities for x_1, \dots, x_n . The best method requires $O(2^{n/2})$ operations, which for $n = 100$ with $2^{100} \approx 1.27 \cdot 10^{30}$ and $2^{n/2} \approx 1.13 \cdot 10^{15}$ represents an insurmountable hurdle for computers. However, for special a_1, \dots, a_n the solution is quite easy to find, e.g. for $a_i = 2^{i-1}$. The binary representation of g immediately delivers x_1, \dots, x_n . In general, the a 0-1 knapsack problem can be solved easily if a permutation¹²⁴ π of $1, \dots, n$ exists with $a_{\pi(j)} > \sum_{i=1}^{j-1} a_{\pi(i)}$. If, in addition, π is the identity, i.e. $\pi(i) = i$ for $i = 1, 2, \dots, n$, then the sequence a_1, \dots, a_n is said to be super-increasing. The following algorithm solves the knapsack problem with a super-increasing sequence in the timeframe of $O(n)$.

¹²⁴A permutation π of the numbers $1, \dots, n$ is a change in the order in which these numbers are listed. For example, a permutation π of $(1, 2, 3)$ is $(3, 1, 2)$, i.e. $\pi(1) = 3$, $\pi(2) = 1$ and $\pi(3) = 2$.

```

for  $i = n$  to 1 do
  if  $T \geq a_i$  then
     $T := T - s_i$ 
     $x_i := 1$ 
  else
     $x_i := 0$ 
if  $T = 0$  then
   $X := (x_1, \dots, x_n)$  is the solution.
else
  No solution exists.

```

Algorithm 1. Solving knapsack problems with super-increasing weights

4.2.2 Merkle-Hellman knapsack encryption

In 1978, Merkle and Hellman [Merkle1978] specified a public key encryption procedure that is based on “defamiliarising” the easy 0-1 knapsack problem with a super-increasing sequence into a congruent one with a super-increasing sequence. It is a block ciphering that ciphers an n -bit plaintext each time it runs. More precisely:

Let (a_1, \dots, a_n) be super-increasing. Let m and w be two co-prime numbers with $m > \sum_{i=1}^n a_i$ and $1 \leq w \leq m-1$. Select \bar{w} with $w\bar{w} \equiv 1 \pmod{m}$ the modular inverse of w and set $b_i := wa_i \pmod{m}$, $0 \leq b_i < m$ for $i = 1, \dots, n$, and verify whether the sequence b_1, \dots, b_n is not super-increasing. A permutation $b_{\pi(1)}, \dots, b_{\pi(n)}$ of b_1, \dots, b_n is then published and the inverse permutation μ to π is defined secretly. A sender writes his/her message in blocks $(x_1^{(j)}, \dots, x_n^{(j)})$ of binary numbers n in length, calculates

$$g^{(j)} := \sum_{i=1}^n x_i^{(j)} b_{\pi(i)}$$

and sends $g^{(j)}$, $(j = 1, 2, \dots)$.

The owner of the key calculates

$$G^{(j)} := \bar{w} g^{(j)} \pmod{m}, \quad 0 \leq G^{(j)} < m$$

and obtains the $x_{\mu(i)}^{(j)} \in \{0, 1\}$ (and thus also the $x_i^{(j)}$) from

$$\begin{aligned} G^{(j)} &\equiv \bar{w} g^{(j)} = \sum_{i=1}^n x_i^{(j)} b_{\pi(i)} \bar{w} \equiv \sum_{i=1}^n x_i^{(j)} a_{\pi(i)} \pmod{m} \\ &= \sum_{i=1}^n x_{\mu(i)}^{(j)} a_{\pi(\mu(i))} = \sum_{i=1}^n x_{\mu(i)}^{(j)} a_i \pmod{m}, \end{aligned}$$

by solving the easier 0-1 knapsack problems $K(a_1, \dots, a_n; G^{(j)})$ with super-increasing sequence a_1, \dots, a_n .

Merkle-Hellman procedure (based on knapsack problems).

In 1982, Shamir [Shamir1982] specified an algorithm for breaking the system in polynomial time without solving the general knapsack problem. Len Adleman [Adleman1982] and Jeff Lagarias [Lagarias1983] specified an algorithm for breaking the twice iterated Merkle-Hellman knapsack encryption procedure in polynomial time. Ernst Brickell [Brickell1985] then specified an algorithm for breaking multiply iterated Merkle-Hellman knapsack encryption procedures in polynomial time. This made this procedure unsuitable as an encryption procedure. It therefore delivers a one way function whose trapdoor information (defamiliarisation of the 0-1 knapsack problem) could be discovered by an eavesdropper.

4.3 Decomposition into prime factors as a basis for public key procedures

4.3.1 The RSA procedure¹²⁵

As early as 1978, R. Rivest, A. Shamir, L. Adleman [RSA1978] introduced the most important asymmetric cryptography procedure to date.

Key generation:

Let p and q be two different prime numbers and $N = pq$. Let e be any prime number relative to $\phi(N)$, i.e. $\gcd(e, \phi(N)) = 1$. Using the Euclidean algorithm, we calculate the natural number $d < \phi(N)$, such that

$$ed \equiv 1 \pmod{\phi(N)}.$$

whereby ϕ is the **Euler phi Function**.

The output text is divided into blocks and encrypted, whereby each block has a binary value $x^{(j)} \leq N$.

Public key:

$$N, e.$$

Private key:

$$d.$$

Encryption:

$$y = e_T(x) = x^e \pmod{N}.$$

Decryption:

$$d_T(y) = y^d \pmod{N}.$$

RSA procedure (based on the factorisation problem).

Comment: The Euler phi function is defined as: $\phi(N)$ is the number of natural numbers that do not have a common factor with N $x \leq N$. Two natural numbers a and b are co-prime if $\gcd(a, b) = 1$.

For the Euler phi function: $\phi(1) = 1, \phi(2) = 1, \phi(3) = 2, \phi(4) = 2, \phi(6) = 2, \phi(10) = 4, \phi(15) = 8$. For example, $\phi(24) = 8$, because $|\{x < 24 : \gcd(x, 24) = 1\}| = |\{1, 5, 7, 11, 13, 17, 19, 23\}|$.

If p is a prime number, then $\phi(p) = p - 1$.

If we know the various prime factors p_1, \dots, p_k of N , then $\phi(N) = N \cdot (1 - \frac{1}{p_1}) \cdots (1 - \frac{1}{p_k})$ ¹²⁶.

¹²⁵Using CrypTool v1.3 you gain practical experience with the RSA procedure via the menu **Indiv.Procedures \ RSA Demonstration \ RSA Cryptosystem**.

¹²⁶Further formulas for the Euler phi function are in the article “Introduction to Elementary Number Theory with Examples”, chapter 3.8.1.

In the case of $N = pq$, $\phi(N) = pq(1 - 1/p)(1 - 1/q) = p(1 - 1/p)q(1 - 1/q) = (p - 1)(q - 1)$.

n	$\phi(n)$	The natural numbers that are co-prime to n and less than n .
1	1	1
2	1	1
3	2	1, 2
4	2	1, 3
5	4	1, 2, 3, 4
6	2	1, 5
7	6	1, 2, 3, 4, 5, 6
8	4	1, 3, 5, 7
9	6	1, 2, 4, 5, 7, 8
10	4	1, 3, 7, 9
15	8	1, 2, 4, 7, 8, 11, 13, 14

The function e_T is a one way function whose trapdoor information is the decomposition into primes of N .

At the moment, no algorithm is known that can factorise two prime numbers sufficiently quickly for extremely large values (e.g. for several hundred decimal places). The quickest algorithms known today [Stinson1995] factorise a compound whole number N in a time period proportional to $L(N) = e\sqrt{\ln(N) \ln(\ln(N))}$.

N	10^{50}	10^{100}	10^{150}	10^{200}	10^{250}	10^{300}
$L(N)$	$1.42 \cdot 10^{10}$	$2.34 \cdot 10^{15}$	$3.26 \cdot 10^{19}$	$1.20 \cdot 10^{23}$	$1.86 \cdot 10^{26}$	$1.53 \cdot 10^{29}$

As long as no better algorithm is found, this means that values of the order of magnitude 100 to 200 decimal places are currently safe. Estimates of the current computer technology indicate that a number with 100 decimal places could be factorised in approximately two weeks at justifiable cost. Using an expensive configuration (e.g. of around 10 million US dollars), a number with 150 decimal places could be factorised in about a year. A 200-digit number should remain impossible to factorise for a long time to come, unless there is a mathematical breakthrough. For example, it would take about 1000 years to decompose a 200-digit number into prime factors using the existing algorithms; this applies even if 10^{12} operations can be performed per second, which is beyond the performance of current technology and would cost billions of dollars in development costs. However, you can never be sure that there won't be a mathematical breakthrough tomorrow.

To this date, it has not been proved that the problem of breaking RSA is equivalent to the factorisation problem. Nevertheless, it is clear that the RSA procedure will no longer be safe if the factorisation problem is "solved".

4.3.2 Rabin public key procedure (1979)

In this case it has been shown that the procedure is equivalent to breaking the factorisation problem. Unfortunately, this procedure is susceptible to chosen-cipher text attacks.

Let p and q be two different prime numbers with $p, q \equiv 3 \pmod{4}$ and $n = pq$. Let $0 \leq B \leq n - 1$.	
<u>Public key:</u>	$e = (n, B).$
<u>Private key:</u>	$d = (p, q).$
<u>Encryption:</u>	$y = e_T(x) = x(x + B) \pmod{n}.$
<u>Decryption:</u>	$d_T(y) = \sqrt{y + B^2/4} - B/2 \pmod{n}.$

Rabin procedure (based on the factorisation problem).

Caution: Because $p, q \equiv 3 \pmod{4}$ the encryption is easy to calculate (if the key is known). This is not the case for $p \equiv 1 \pmod{4}$. In addition, the encryption function is not injective: There are precisely four different source codes that have $e_T(x)$ as inverse image: $x, -x - B, \omega(x + B/2) - B/2, -\omega(x + B/2) - B/2$, where ω is one of the four roots of unity. The source codes therefore must be redundant for the encryption to remain unique!

Backdoor information is the decomposition into prime numbers of $n = pq$.

4.4 The discrete logarithm as a basis for public key procedures

Discrete logarithms form the basis for a large number of algorithms for public-key procedures.

4.4.1 The discrete logarithm in \mathbb{Z}_p

Let p be a prime number and let $g \in \mathbb{Z}_p^* = \{1, \dots, p-1\}$. Then the discrete exponential function base g is defined as

$$e_g : k \longrightarrow y := g^k \pmod{p}, \quad 1 \leq k \leq p-1.$$

The inverse function is called a discrete logarithm function \log_g ; the following holds:

$$\log_g(g^k) = k.$$

The problem of the discrete logarithm (in \mathbb{Z}_p^*) is understood to be as follows:

Given p, g and y , determine k such that $y = g^k \pmod{p}$.

It is much more difficult to calculate the discrete logarithm than to evaluate the discrete exponential function (see chapter 3.9). There are several procedures for calculating the discrete logarithm [Stinson1995]:

Name	Complexity
Baby-Step-Giant-Step	$O(\sqrt{p})$
Silver-Pohlig-Hellman	polynomial in q , the greatest prime factor of $p - 1$.
Index-Calculus	$O(e^{(1+o(1))\sqrt{\ln(p) \ln(\ln(p))}})$

4.4.2 Diffie-Hellman key agreement¹²⁷

The mechanisms and algorithms of classical cryptography only take effect when the subscribers have already exchanged the secret key. In classical cryptography you cannot avoid exchanging secrets without encrypting them. Transmission safety here must be achieved using non-cryptographic methods. We say that we need a secret channel for exchanging secrets. This channel can be realised either physically or organisationally.

What is revolutionary about modern cryptography is, amongst other things, that you no longer need secret channels: You can agree secret keys using non-secret, i.e. public channels.

One protocol that solves this problem is that of Diffie and Hellman.

Two subscribers A and B want to agree on a joint secret key.
Let p be a prime number and g a natural number. These two numbers do not need to be secret.
The two subscribers then select a secret number a and b from which they calculate the values $\alpha = g^a \pmod{p}$ and $\beta = g^b \pmod{p}$. They then exchange the numbers α and β . To end with, the two subscribers calculate the received value to the power of their secret value to get $\beta^a \pmod{p}$ and $\alpha^b \pmod{p}$.
Thus

$$\beta^a \equiv (g^b)^a \equiv g^{ba} \equiv g^{ab} \equiv (g^a)^b \equiv \alpha^b \pmod{p}$$

Diffie-Hellman key agreement.

¹²⁷With CrypTool v1.3.04 this exchange protocol has been visualized: you can execute the single steps with concrete numbers using menu **Indiv. Procedures \ Diffie-Hellman Demonstration**.

The safety of the **Diffie-Hellman protocol** is closely connected to calculating the discrete logarithm mod p . It is even thought that these problems are equivalent.

4.4.3 ElGamal public key encryption procedure in \mathbb{Z}_p^*

By varying the Diffie-Hellman key agreement protocol slightly, you can obtain an asymmetric encryption algorithm. This observation was made by Taher ElGamal.

Let p be a prime number such that the discrete logarithm in \mathbb{Z}_p is difficult to compute. Let $\alpha \in \mathbb{Z}_p^*$ be a primitive element. Let $a \in \mathbb{N}$ and $\beta = \alpha^a \mod p$.

Public key:

$$p, \alpha, \beta.$$

Private key:

$$a.$$

Let $k \in \mathbb{Z}_{p-1}$ be a random number and $x \in \mathbb{Z}_p^*$ the plaintext.

Encryption:

$$e_T(x, k) = (y_1, y_2),$$

where

$$y_1 = \alpha^k \mod p$$

and

$$y_2 = x\beta^k \mod p.$$

Decryption:

$$d_T(y_1, y_2) = y_2(y_1^a)^{-1} \mod p$$

ElGamal procedure (based on the factorisation problem).

4.4.4 Generalised ElGamal public key encryption procedure

The discrete logarithm can be generalised in any number of finite groups (G, \circ) . The following provides several properties of G , that make the discrete logarithm problem difficult.

Calculating the discrete exponential function Let G be a group with the operation \circ and $g \in G$. The (discrete) exponential function base g is defined as

$$e_g : k \mapsto g^k, \quad \text{for all } k \in \mathbb{N}.$$

where

$$g^k := \underbrace{g \circ \dots \circ g}_{k \text{ times}}.$$

The exponential function is easy to calculate:

Lemma.

The power g^k can be calculated in at most $2 \log_2 k$ group operations.

Proof

Let $k = 2^n + k_{n-1}2^{n-1} + \dots + k_12 + k_0$ be the binary representation of k . Then $n \leq \log_2(k)$, because $2^n \leq k < 2^{n+1}$. k can be written in the form $k = 2k' + k_0$ with $k' = 2^{n-1} + k_{n-1}2^{n-2} + \dots + k_1$. Thus

$$g^k = g^{2k' + k_0} = (g^{k'})^2 g^{k_0}.$$

We therefore obtain g^k from $g^{k'}$ by squaring and then multiplying by g . The claim is thus proved by induction to n . \square

Problem of the discrete logarithm

Let G be a finite group with the operation \circ . Let $\alpha \in G$ and $\beta \in H = \{\alpha^i : i \geq 0\}$.
 We need to find a unique $a \in \mathbb{N}$ with $0 \leq a \leq |H| - 1$ and $\beta = \alpha^a$.
 We define a as $\log_\alpha(\beta)$.

Calculating the discrete logarithm A simple procedure for calculating the discrete logarithm of a group element, that is considerably more efficient than simply trying all possible values for k , is the Baby-Step-Giant-Step algorithm.

Theorem 4.1. *[Baby-Step-Giant-Step algorithm] Let G be a group and $g \in G$. Let n be the smallest natural number with $|G| \leq n^2$. Then the discrete logarithm of an element $h \in G$ can be calculated base g by generating two lists each containing n elements and comparing these lists. In order to calculate these lists, we need $2n$ group operations.*

Proof

First create the two lists

Giant-Step list: $\{1, g^n, g^{2n}, \dots, g^{n \cdot n}\},$

Baby-Step list: $\{hg^{-1}, hg^{-2}, \dots, hg^{-n}\}.$

If $g^{jn} = hg^{-i}$, i.e. $h = g^{i+jn}$, then the problem is solved. If the lists are disjoint, then h cannot be represented as g^{i+jn} , $i, j \leq n$. As all powers of g are thus recorded, the logarithm problem does not have a solution. \square

You can use the Baby-Step-Giant-Step algorithm to demonstrate that it is much more difficult to calculate the discrete logarithm than to calculate the discrete exponential function. If the numbers that occur have approximately 1000 bits in length, then you only need around 2000

multiplications to calculate g^k but around $2^{500} \approx 10^{150}$ operations to calculate the discrete logarithm using the Baby-Step-Giant-Step algorithm.

In addition to the Baby-Step-Giant-Step algorithm, there are also numerous other procedures for calculating the discrete logarithm [Stinson1995].

The theorem from Silver-Pohlig-Hellman In finite Abelian groups, the discrete logarithm problem can be reduced to groups of a lower order.

Theorem 4.2. [Silver-Pohlig-Hellman] Let G be a finite Abelian group with $|G| = p_1^{a_1} p_2^{a_2} \dots p_s^{a_s}$. The discrete logarithm in G can then be reduced to solving logarithm problems in groups of the order p_1, \dots, p_s .

If $|G|$ contains a “dominant” prime factor p , then the complexity of the logarithm problem is approximately

$$O(\sqrt{p}).$$

Therefore, if the logarithm problem is to be made difficult, the order of the group used G should have a large prime factor. In particular, if the discrete exponential function in the group \mathbb{Z}_p^* is to be a one way function, then $p - 1$ must be a large prime factor.

Let G be a finite group with operation \circ , and let $\alpha \in G$, so that the discrete logarithm in $H = \{\alpha^i : i \geq 0\}$ is difficult, Let a with $0 \leq a \leq H - 1$ and let $\beta = \alpha^a$.	
<u>Public key:</u>	$\alpha, \beta.$
<u>Private key:</u>	$a.$
Let $k \in \mathbb{Z}_{ H }$ be a random number and $x \in G$ be a plaintext.	
<u>Encryption:</u>	$e_T(x, k) = (y_1, y_2),$
where	$y_1 = \alpha^k$
and	$y_2 = x \circ \beta^k.$
<u>Decryption:</u>	$d_T(y_1, y_2) = y_2 \circ (y_1^a)^{-1}$

Generalised ElGamal procedure (based on the factorisation problem).

Elliptic curves provide useful groups for public key encryption procedures.

References

- [Adleman1982] Adleman L.:
On breaking the iterated Merkle-Hellman public key Cryptosystem.
Advances in Cryptology, Proceedings of Crypto 82, Plenum Press 1983, 303-308.
- [Balcazar1988] Balcazar J.L., Daaz J., Gabarr´ J.:
Structural Complexity I.
Springer Verlag, pp 63.
- [Brickell1985] Brickell E.F.:
Breaking Iterated Knapsacks.
Advances in Cryptology: Proc. CRYPTO84, Lecture Notes in Computer Science, vol. 196,
Springer-Verlag, New York, 1985, pp. 342-358.
- [Lagarias1983] Lagarias J.C.:
Knapsack public key Cryptosystems and diophantine Approximation.
Advances in Cryptology, Proceedings of Crypto 83, Plenum Press.
- [Merkle1978] Merkle R. and Hellman M.:
Hiding information and signatures in trapdoor knapsacks.
IEEE Trans. Information Theory, IT-24, 1978.
- [RSA1978] Rivest R.L., Shamir A. and Adleman L.:
A Method for Obtaining Digital Signatures and Public Key Cryptosystems.
Commun. ACM, vol 21, April 1978, pp. 120-126.
- [Shamir1982] Shamir A.:
A polynomial time algorithm for breaking the basic Merkle-Hellman Cryptosystem.
Symposium on Foundations of Computer Science (1982), 145-152.
- [Stinson1995] Stinson D.R.:
Cryptography.
CRC Press, Boca Raton, London, Tokyo, 1995.

Web links

1. http://www.geocities.com/st_busygin/clipat.html

5 Hash Functions and Digital Signatures

(Schneider J. / Esslinger B. / Filipovics B. / Koy H., June 2002, Update: Feb. 2003)

The aim of digital signatures is to guarantee the following two points:

- User authenticity:
It can be checked whether a message really does come from a particular person.
- Message integrity:
It can be checked whether the message has been changed (on route).

An asymmetric technique is used again (see encryption procedures). Participants who wish to generate a digital signature for a document must possess a pair of keys. They use their secret key to generate signatures and the recipient uses the sender's public key to verify whether the signature is correct. As before, it must be impossible to use the public key to derive the secret key¹²⁸.

In detail, a *Signature procedure* looks like this:

Senders use their message and secret key to calculate the digital signature for the message. Compared to hand-written signatures, digital signatures therefore have the advantage that they also depend on the document to be signed. Signatures from one and the same participant are different unless the signed documents are completely identical. Even inserting a blank in the text would lead to a different signature. The recipient of the message would therefore detect any injury to the message integrity as this would mean that the signature no longer matches the document and is shown to be incorrect when verified.

The document is sent to the recipient together with the signature. The recipient can then use the sender's public key, the document and the signature to establish whether or not the signature is correct. In practice, however, the procedure we have just described has a decisive disadvantage. The signature is approximately as long as the document itself. To prevent an unnecessary increase in data traffic, and also for reasons of performance, we use a cryptographic hash function¹²⁹.

¹²⁸With CrypTool v1.3 you can also generate and check digital signatures:
using the submenus of the main menu **Digital Signatures** or
using menu **Indiv. Procedures \ RSA Demonstration \ Signature Demonstration**.

¹²⁹Hash functions are implemented within CrypTool v1.3.04 at several places.
Using menus **Individual Procedures \ Hash** and **Individual Procedures \ Attack on the Hash Value of the Digital Signature** you have the possibilities

- to apply one of 6 hash functions to the content of the current window,
- to calculate the hash value of a file,
- to test, how changes to a text change the according hash value and
- to perform a simulation, how digital signatures could be attacked by a targeted search for hash value collisions.

5.1 Hash functions

A cryptographic *hash function*¹³⁰ maps a message of any length to a string of characters with a constant size, the hash value. It should be practically impossible, for a given number, to find a message that has precisely this number as hash value. Furthermore, it should be practically impossible to find two messages with the same hash value. In both cases the final signature procedure would have weaknesses.

So far, no formal proof has been found that perfectly secure cryptographic hash functions exist. However, there are several good candidates that have not yet shown any weak points in their structure in practice (e.g. SHA-1¹³¹ or RIPEMD-160¹³²).

The hash function procedure is as follows:

Rather than signing the actual document, the sender now first calculates the hash value of the message and signs this. The recipient also calculates the hash value of the message (the algorithm used must be known), then verifies whether the signature sent with the message is a correct signature of the hash value. If this is the case, the signature is verified to be correct. This means that the message is authentic, because we have assumed that knowledge of the public key does not enable you to derive the secret key. However, you would need this secret key to sign messages in another name.

Some digital signature schemes are based on asymmetric *encryption* procedures, the most prominent example being the RSA system, which can be used for signing by performing the private key operation on the hash value of the document to be signed.

Other digital signature schemes were developed exclusively for this purpose, as the DSA (Digital Signature Algorithm), and are not directly connected with a corresponding encryption scheme.

¹³⁰Secure hash algorithms compute a condensed representation of electronic data (message). When a message is input to a hash algorithm, the result is an output called a message digest. The message digests typically range in length from 128 to 512 bits, depending on the algorithm. Secure hash algorithms are typically used with other cryptographic algorithms, such as digital signature algorithms and keyed-hash message authentication codes, or in the generation of random numbers (bits).

¹³¹SHA-1 is a 160 bit hash function specified in FIPS 180-1, ANSI X9.30 Part 2 and [FIPS186].

SHA means Secure Hash Algorithm, and is widely used, e.g. with DSA, RSA or ECDSA.

The current standard [FIPS180-2] defines four secure hash algorithms – SHA-1, SHA-256, SHA-384, and SHA-512. For these hash algorithms there are also validation tests defined in the test suite FIPS 140-2.

The output length of the SHA algorithms was enhanced because of the possibility of birthday attacks: these make n-bit AES and a 2n-bit hash roughly equivalent:

- 128-bit AES – SHA-256
- 192-bit AES – SHA-384
- 256-bit AES – SHA-512.

With CrypTool v1.3.04 you can comprehend the birthday attack on digital signatures: using menu **Indiv. Procedures \ Attack on the hashvalue of the digital signature**.

¹³²RIPEMD-160, RIPEMD-128 and the optional extension RIPEMD-256 have object identifiers defined by the ISO-identified organization TeleTrusT, both as hash algorithm and in combination with RSA. RIPEMD-160 is also part of the ISO/IEC international standard ISO/IEC 10118-3:1998 on dedicated hash functions, together with RIPEMD-128 and SHA-1.

Further details:

<http://www.esat.kuleuven.ac.be/~bosselae/ripemd160.html>

<http://www.ietf.org/rfc/rfc2857.txt> (“The Use of HMAC-RIPEMD-160-96 within ESP and AH”).

Both, RSA and DSA signature are discussed in more detail in the following two sections. After that we go one step further and show how digital signatures can be used to create the digital equivalent of ID cards. This is called Public Key Certification.

5.2 RSA signatures

As mentioned in the comment at the end of [section 3.10.3](#) it is possible to perform the RSA private and public key operation in reverse order, i. e. raising M to the power of d and then to the power of $e \pmod{N}$ yields M again. Based on this simple fact, RSA can be used as a signature scheme.

The RSA signature S for a message M is created by performing the private key operation:

$$S \equiv M^d \pmod{N}$$

In order to verify, the corresponding public key operation is performed on the signature S and the result is compared with message M :

$$S^e \equiv (M^d)^e \equiv (M^e)^d \equiv M \pmod{N}$$

If the result matches the message M , then the signature is accepted by the verifier, otherwise the message has been tampered with, or was never signed by the holder of d .

As explained above, signatures are not performed on the message itself, but on a cryptographic hash value of the message. To prevent certain attacks on the signature procedure (alone or in combination with encryption) it is necessary to format the hash value before doing the exponentiation, as described in the PKCS#1 (Public Key Cryptography Standard #1 [\[PKCS1\]](#)). The fact that this standard had to be revised recently, after being in use for several years, can serve as an example of how difficult it is to get the details of cryptography right.

5.3 DSA signatures

In August of 1991, the U.S. National Institute of Standards and Technology (NIST) proposed a digital signature algorithm (DSA), which was subsequently adopted as a U.S. Federal Information Processing Standard (FIPS 186 [\[FIPS186\]](#)).

The algorithm is a variant of the ElGamal scheme. Its security is based on the Discrete Logarithm Problem. The DSA public and private key and its procedures for signature and verification are summarised below.

Public Key

p prime

q 160-bit prime factor of $p - 1$

$g = h^{(p-1)/q} \pmod{p}$, where $h < p - 1$ and $h^{(p-1)/q} > 1 \pmod{p}$

$y \equiv g^x \pmod{p}$

Remark: Parameters p, q and g can be shared among a group of users.

Private Key

$x < q$ (a 160-bit number)

Signing

m the message to be signed

k choose at random, less than q

$r = (g^k \bmod p) \bmod q$

$s = (k^{-1}(\text{SHA-1}(m) + xr)) \bmod q$

Remark:

- (s, r) is the signature.
- The security of the signature depends not only on the mathematical properties, but also on using a good random source for k .
- SHA-1 is a 160-bit hash function.

Verifying

$w = s^{-1} \bmod q$

$u_1 = (\text{SHA-1}(m)w) \bmod q$

$u_2 = (rw) \bmod q$

$v = (g^{u_1}y^{u_2} \bmod p) \bmod q$

Remark: If $v = r$, then the signature is verified.

While DSA was specifically designed, so that it can be exported from countries regulating export of encryption soft and hardware (like the U.S. at the time when it was specified), it has been noted [Schneier1996, p. 490], that the operations involved in DSA can be used to emulate RSA and ElGamal encryption.

5.4 Public key certification

The aim of public key certification is to guarantee the connection between a public key and a user and to make it traceable for external parties. In cases in which it is impossible to ensure that a public key really belongs to a particular person, many protocols are no longer secure, even if the individual cryptographic modules cannot be broken.

5.4.1 Impersonation attacks

Assume Charlie has two pairs of keys $(PK1, SK1)$ and $(PK2, SK2)$, where SK denotes the secret key and PK the public key. Further assume that he manages to palm off PK1 on Alice as Bob's public key and PK2 on Bob as Alice's public key (by falsifying a public key directory).

Then he can attack as follows:

- Alice wants to send a message to Bob. She encrypts it using PK1 because she thinks that this is Bob's public key. She then signs the message using her secret key and sends it.
- Charlie intercepts the message, removes the signature and decrypts the message using SK1. If he wants to, he can then change the message in any way he likes. He then encrypts the message again, but this time using Bob's genuine public key, which he has taken from a public key directory, signs the message using SK2 and forwards it to Bob.
- Bob verifies the signature using PK2 and will reach the conclusion that the signature is correct. He then decrypts the message using his secret key.

In this way Charlie can listen in on communication between Alice and Bob and change the exchanged messages without them noticing. The attack will also work if Charlie only has one pair of keys.

Another name for this type of attack is “man-in-the-middle attack”. Users are promised protection against this type of attack by publickey certification, which is intended to guarantee the authenticity of public keys. The most common certification method is the X.509 standard.

5.4.2 X.509 certificate

Each participant who wants to have an X.509 certificate ([X.509]) verifying that his public key belongs to a real person consults what is known as a certification authority (CA)¹³³. He proves his identity to this CA (for example by showing his ID). The CA then issues him an electronic document (certificate) which essentially contains the name of the certificate-holder and the name of the CA, the certificate-holder's public key and the validity period of the certificate. The CA then signs the certificate using its secret key.

Anyone can now use the CA's public key to verify whether a certificate is falsified. The CA therefore guarantees that a public key belongs to a particular user.

This procedure is only secure as long as it can be guaranteed that the CA's public key is correct. For this reason, each CA has its public key certified by another CA that is superior in the hierarchy. In the upper hierarchy level there is usually only one CA, which can of course then have its key certified by another CA. It must therefore transfer its key securely in another way. In the case of many software products that work with certificates (such as the Microsoft and Netscape Web browsers), the certificates of these root CAs are permanently embedded in the program right from the start and cannot be changed by users at a later stage. However, (public) CA keys, in particularly those of the root entity, can also be secured by means of making them available publicly.

¹³³Often called trust center, if the certificates are not only offered to a closed user group.

References

- [Schneier1996] Bruce Schneier,
Applied Cryptography, Protocols, Algorithms, and Source Code in C,
Wiley, 2nd edition, 1996.
- [PKCS1] RSA Laboratories,
PKCS #1 v2.1 Draft 3: RSA Cryptography Standard,
April 19, 2002.
- [X.509] ITU-T,
ITU-T Recommendation X.509 (1997 E): Information Technology – Open Systems Interconnection – The Directory: Authentication Framework,
June 1997.
- [X.509v3] ITU-T,
X.509 (1993) Amendment 1: Certificate Extensions, The Directory Authentication Framework,
International Telecommunication Union, Geneva, Switzerland, July 1995
(equivalent to amendment 1 to ISO/IEC 9594-8).
- [FIPS180-2] U.S. Department of Commerce/N.I.S.T. ,
Secure Hash Standard (SHS),
August 1, 2002.
(FIPS 180-2 will supersede FIPS 180-1 beginning February 1, 2003.)
- [FIPS186] U.S. Department of Commerce/N.I.S.T. ,
Entity authentication using public key cryptography,
February 18, 1997.
No more valid.
- [FIPS186-2] U.S. Department of Commerce/N.I.S.T. ,
Digital Signature Standard (DSS),
January 27, 2000. Change Note: October 5, 2001.
<http://csrc.nist.gov/publications/fips/fips186-2/fips186-2-change1.pdf>

6 Elliptic Curves

(Filipovics B. / Büger M. / Esslinger B. / Oyono R., April 2000, Updates: Dec. 2001, June 2002, Mar. 2003)

6.1 Elliptic curve cryptography – a high-performance substitute for RSA?

In many business sectors secure and efficient data transfer is essential. In particular, the RSA algorithm is used in many applications. Although the security of RSA is beyond doubt, the evolution in computing power has caused a growth in the necessary key length. Today, 1024-bit RSA keys are standard, but the GISA (German Information Security Agency) recommends the usage of 2048-bit keys from 2006 on (compare section 3.11). The fact that most chips on smart cards cannot process keys extending 1024 bit shows that there is a need for alternatives. Elliptic curve cryptography (ECC) can be such an alternative in the area of asymmetric cryptography.

The efficiency of a cryptographic algorithm depends on the key length and the calculation effort that is necessary to provide a prescribed level of security. The major advantage of ECC compared to RSA is that it requires much shorter key lengths. If we assume that the computing power increases by Moore's law (i. e. it doubles every 18 months)¹³⁴, then the evolution of the key lengths for secure communication will be as figure 1 [Lenstra1999] (source: Arjen Lenstra and Eric Verheul: <http://cryptosavvy.com/table.htm>).

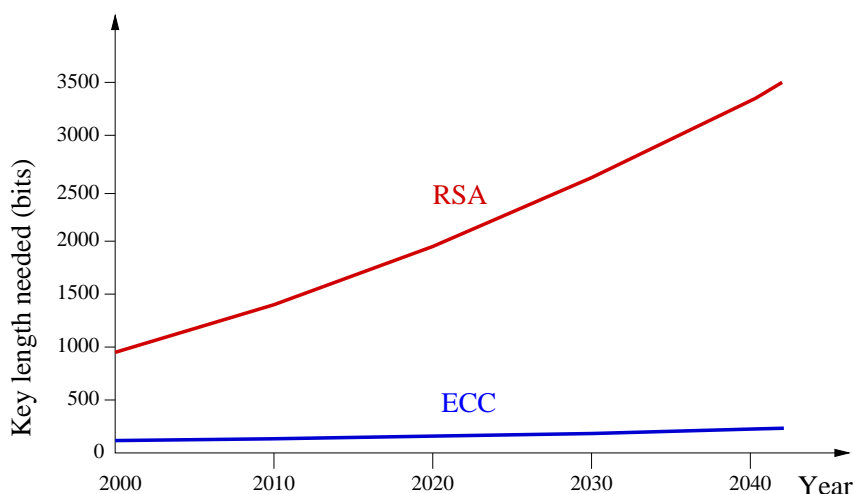


Figure 1: Prognosis of the key lengths to be regarded safe for RSA and Elliptic Curves

In addition, a digital signature can be processed 10-times faster with ECC than with RSA. However, verification of a given signature is still more efficient with RSA than with ECC. Refer to figure 2 (source: Dr. J. Merkle, Elliptic Curve Cryptography Workshop, 2001) for a comparison.

¹³⁴empirical knowledge by Gordon Moore, co-founder of Intel, 1965

The reason is that RSA public keys can be chosen relatively small as long as the secret key is long enough.

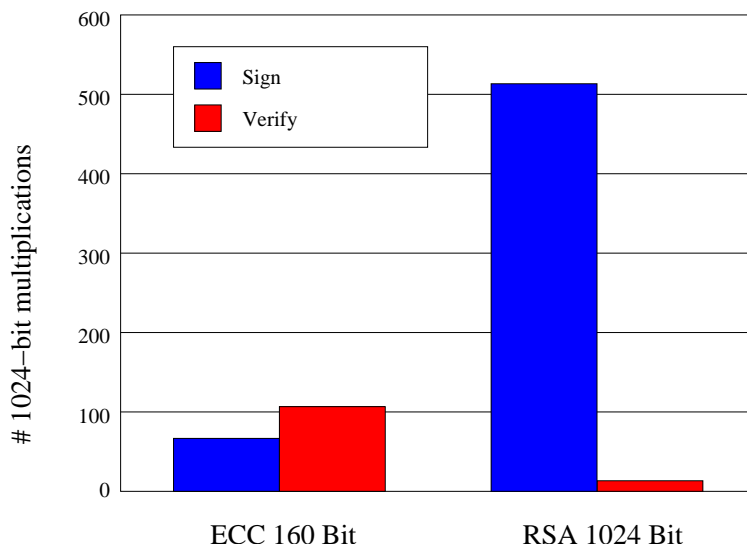


Figure 2: Comparison of signing and verification time for RSA and Elliptic Curves

Nevertheless, thin clients like smart cards usually have to store the (long) secret key and have to process a digital signature rather than verify one. Therefore, there is a clear advantage in using ECC in terms of efficiency.

Nowadays, the major problem with ECC-implementations is the lack of standardization. There is only one way to implement RSA, but there are many ways for ECC: One can work with different sets of numbers, different (elliptic) curves — described by parameters¹³⁵ — , and a variety of representations of the elements on the curve. Each choice has its advantages and disadvantages, and one can certainly construct the most efficient for each application. However, this causes problems in interoperability. But if all ECC-tools should be able to communicate with each other, they will have to support all different algorithms, which might put the advantage of efficient computation and the need of less storage capacity to the contrary.

Therefore, international standardization organizations like IEEE (P1363), ASC (ANSI X9.62, X9.63), ISO/IEC as well as major players like RSA labs or Certicom have recently started standardization initiatives. While the IEEE only describes the different implementations, the ASC has explicitly stated 10 elliptic curves and recommends their usage. The advantage of the ASC approach is that one needs only a single byte to indicate which curve is meant. However, it is

¹³⁵see chapter 6.4

not yet clear whether the ASC-curves will become a de facto standard.

Although we see no need to replace RSA in any application today¹³⁶, one should take the usage of ECC-based tools into consideration whenever a new system is set up — in particular, when the tool should be available beyond 2005¹³⁷.

6.2 Elliptic curves – history

Mathematicians have been researching elliptic curves for over 100 years. Over the course of time, many lengthy and mathematically complex results have been found and published which are connected to elliptic curves. A mathematician would say that elliptic curves (or the mathematics behind them) are widely understood. This research was originally purely mathematical. That is to say, elliptic curves were investigated, for example, in the mathematical areas of number theory and algebraic geometry, which are generally highly abstract. Even in the recent past, elliptic curves played an important role in pure mathematics. In 1993 and 1994, Andrew Wiles published mathematical works that triggered enthusiasm far beyond the specialist audience. In these works, he proved a conjecture put forward in the 1960's. To put it short, this conjecture was concerned with the connection between elliptic curves and what are called module forms. What is particularly interesting for most people is that the works of Wiles also proved the famous second theorem of Fermat. Mathematicians had spent centuries (Fermat lived from 1601 to 1665) trying to find a strict proof of this theorem. Understandably, therefore, Wiles' proof got a good response. Fermat formulated his theorem as follows (written in the border of a book):

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos ejusdem nominis fas est dividere: cujus rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.

With a free translation, using the denotation of modern mathematics, this means:

No positive whole numbers x, y and z greater than zero exist such that $x^n + y^n = z^n$ for $n > 2$. I have found an amazing proof of this fact, but there is too little space within the confines of this book to include it.

This is truly amazing: A statement that is relatively simple to understand (we are referring to Fermat's second theorem here) could only be proved after such a long period of time, although Fermat himself claimed to have found a proof. What's more, the proof found by Wiles is extremely extensive (all of Wiles publications connected with the proof made up a book in themselves). This should therefore make it obvious that elliptic curves are generally based on highly complex mathematics.

Anyway that's enough about the role of elliptic curves in pure mathematics. In 1985 Neal Koblitz and Victor Miller independently suggested using elliptic curves in cryptography. Elliptic curves have thus also found a concrete practical application. Another interesting area of application

¹³⁶Current information on the security of the RSA algorithm can be found in chapter 3.11.

¹³⁷Compare the recommendation of GISA: "Fitting Crypto Algorithms" from October 24th, 2002.

for elliptic curves is for factorising whole numbers (the RSA cryptographic system is based on the difficulty/complexity of finding prime factors of an extremely large number; compare section 3.11.). In this area, procedures based on elliptic curves have been investigated and used since 1987 (compare section 6.8).

There are also prime number tests based on elliptic curves.

Elliptic curves are used differently in the various areas. Encryption procedures based on elliptic curves are based on the difficulty of a problem known as elliptic curve discrete logarithm. The factorisation of whole numbers uses the fact that a large number of elliptic curves can be generated for a natural composite number n with several prime factors; however, these curves are not then groups for composite n . More information about this can be found under the chapter 6.8.

6.3 Elliptic curves – mathematical basics

This section provides information about *groups* and *fields*.

6.3.1 Groups

Because the term *group* is used differently in everyday language than in mathematics, we will, for reasons of completeness, begin by introducing the essential statement of the formal definition of a group:

- A group is a non-empty set G on which an operation “ \cdot ”. The set G is closed under this operation, which means that for any two elements a, b taken from G , performing the operation on them gives an element in G , i.e. $ab = a \cdot b$ lies in G .
- For all elements a, b and c in G : $(ab)c = a(bc)$ (associative law).
- There exists an element e in G that behaves neutrally with respect to the operation \cdot . That means that for all a in the set G : $ae = ea = a$.
- For each element a in G there exists a so-called inverse¹³⁸ element a^{-1} in G such that: $aa^{-1} = a^{-1}a = e$.

If also $ab = ba$ (commutative law) for all a, b in G , then we call the group an *Abelian* group.

Since we may define different operations on the same set, we distinguish them by giving them different names (e.g. $+$ addition or \cdot multiplication).

The simplest example of an (Abelian) group is the group of whole numbers under the standard operation of addition. The set of whole numbers is denoted as \mathbb{Z} . \mathbb{Z} has an infinite number of elements, because $\mathbb{Z} = \{\dots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}$. For example, the operation of $1 + 2$ lies in \mathbb{Z} , for $1 + 2 = 3$ and 3 lies in \mathbb{Z} . The neutral element in the group \mathbb{Z} is 0. The inverse element of 3 is -3 , for $3 + (-3) = 0$.

¹³⁸The inverse is uniquely determined because if $x, y \in G$ are each inverse to a , i.e. $ax = xa = e$ and $ay = ya = e$, then $x = xe = x(ay) = (xa)y = ey = y$.

For our purpose, so-called *finite* groups play an important role. This means that there exists a set \mathcal{M} with a fixed number of elements and an operation $+$ such that the above conditions are fulfilled. One example of this is any set \mathbb{Z}_n where $\mathbb{Z}_n = \{0, 1, 2, 3, \dots, n-1\}$, n is a positive whole number and the operation is addition mod n , i.e. a and b in \mathbb{Z}_n are subject to the operation $a + b \bmod n$.

Cyclic groups Cyclic groups¹³⁹ are those groups G' that possess an element g from which the group operation can be used to generate all other elements in the group. This means that for each element a in G' there exists a positive whole number i such that if g is subject to the operation i times (i.e. “ $g \cdot i$ ”), $g + g + \dots + g = a$ (additive group) or $g^i = g \cdot g \dots g = a$ (multiplicative group). The element g is the *generator* of the cyclic group — each element in G can be generated using g and the operation.

Group order Now to the order of an element of the group: Let a be in G . The smallest positive whole number r for which a subject to the operation with itself r times is the neutral element of the group G (i.e.: $r \cdot a = a + a + \dots + a = e$ respectively $a^r = e$), is called the *order* of a .

The order of the group is the number of elements in the set G .

6.3.2 Fields

In mathematics, one is often interested in sets on which at least two (group) operations are defined — frequently called addition and multiplication. Most prominent are so called fields.

A field is understood to be a set K with two operations (denoted as $+$ and \cdot) which fulfils the following conditions:

- The set K forms an Abelian group together with the operation $+$ (addition), where 0 is the neutral element of the operation $+$.
- The set $K \setminus \{0\}$ also forms an Abelian group together with the operation \cdot (multiplication).
- For all elements a, b and c in K , we have $c \cdot (a + b) = c \cdot a + c \cdot b$ and $(a + b) \cdot c = a \cdot c + b \cdot c$ (distributive law).

Fields may contain an infinite number of elements (e.g. the field of real numbers). They are called *infinite* fields. In contrast we call a field *finite*, if it contains only a finite number of elements (e.g. $\mathbb{Z}_p = \{0, 1, 2, 3, \dots, p-1\}$, where p is a prime. \mathbb{Z}_p with addition mod p and multiplication mod p).

¹³⁹Cyclic groups can be in general also endless like the additive group of the integer numbers. We consider here only finite cyclic groups.

Characteristic of a field Let K be a field and 1 be the neutral element of K with respect to the multiplicative operation “ \cdot ”. Then the characteristic of K is said to be the order of 1 with respect to the additive operation. This means that the characteristic of K is the smallest positive integer n such that

$$\underbrace{1 + 1 + \cdots + 1}_{n \text{ times}} = 0.$$

If there is no such n , i.e. if $1 + 1 + \cdots + 1 \neq 0$ no matter how many 1s we add, then we call K a field with characteristic 0.

Thus, fields with characteristic 0 are infinite since they contain the (pairwise distinct) elements 1, $1 + 1$, $1 + 1 + 1$, \dots . On the other hand, fields with finite characteristic may be finite or infinite.

If the characteristic is finite, it has to be prime. This fact can easily be proved: Assume $n = pq$, $p, q < n$, is the characteristic of a field K . By definition of n , the elements $\bar{p} = \underbrace{1 + 1 + \cdots + 1}_{p \text{ times}}$, $\bar{q} = \underbrace{1 + 1 + \cdots + 1}_{q \text{ times}}$ of K are not equal to 0. Thus, there exist inverse elements $\bar{p}^{-1}, \bar{q}^{-1}$ with

respect to multiplication. It follows that $(\bar{p}\bar{q})(\bar{p}^{-1}\bar{q}^{-1}) = 1$, which contradicts the fact that $\bar{p}\bar{q} = \bar{n} = \underbrace{1 + 1 + \cdots + 1}_{n \text{ times}} = 0$ and, hence, $\underbrace{(\bar{p}\bar{q})(\bar{p}^{-1}\bar{q}^{-1})}_{=0} = 0$.

Comment: The field of real numbers has the characteristic 0; the field \mathbb{Z}_p has the characteristic p . If p is not prime, \mathbb{Z}_p is not a field at all.

The most simple field is $\mathbb{Z}_2 = \{0, 1\}$. It contains only two elements, the neutral elements with respect to addition and multiplication. In particular, we have $0 + 0 = 0$, $0 + 1 = 1 + 0 = 1$, $1 + 1 = 0$, $1 \cdot 1 = 1$, $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0$.

Finite Fields As mentioned above, each finite field has a characteristic $p \neq 0$, where p is a prime. On the other hand, given a prime p there is a field which has exactly p elements, that is \mathbb{Z}_p .

However, the number of elements of a field need not be prime in general. For example, it is not hard to construct a field with 4 elements¹⁴⁰.

One can show that the order of any field is a prime power (i.e. the power of a prime number). On the other hand, we can construct a field with p^n elements for any given prime p and positive integer n . Since two fields that have the same number of elements can not be distinguished¹⁴¹,

¹⁴⁰The set $K = \{0, 1, a, b\}$ fitted with the operation defined in the tabular below is a field:

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

und

\cdot	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

¹⁴¹If K, K' are fields with $k = p^n$ elements, then there is a one-to-one map $\varphi : K \rightarrow K'$, that respects the arithmetic of the field. Such a map is called an isomorphism. Isomorphic fields mathematically behave in the same way so that it makes no sense to distinguish between them. For example, \mathbb{Z}_2 und $K' = \{ZERO, ONE\}$ with zero-

we say that there is **the field with p^n elements** and denote it by $GF(p^n)$. Here GF stands for *Galois Field* to commemorate the French Mathematician Galois.

The fields $GF(p)$ of prime order play a prominent role. They are called prime fields and often denoted by \mathbb{Z}_p ¹⁴².

6.4 Elliptic curves in cryptography

In cryptography elliptic curve are a useful tool. Such curves are described by some equation. A detailed analysis has shown that curves of the form¹⁴³

$$F(x_1, x_2, x_3) = -x_1^3 + x_2^2 x_3 + a_1 x_1 x_2 x_3 - a_2 x_1^2 x_3 + a_3 x_2 x_3^2 - a_4 x_1 x_3^2 - a_6 x_3^3 = 0, \quad (1)$$

are especially useful. The variables x_1, x_2, x_3 and parameters a_1, \dots, a_4, a_6 are elements of a given field K , which has certain properties that make it useful from the cryptographic point of view. The underlying field K might be the well known field of real numbers or some finite field (see last section). In order to obtain a curve that is useful for cryptography, the parameters have to be chosen in a way that the following conditions hold

$$\frac{\partial F}{\partial x_1} \neq 0, \quad \frac{\partial F}{\partial x_2} \neq 0, \quad \frac{\partial F}{\partial x_3} \neq 0.$$

We identify points on the curve that can be derived from each other by multiplying each component with some scalar. This makes sense since (x_1, x_2, x_3) solves (1) if and only if $\alpha(x_1, x_2, x_3)$ ($\alpha \neq 0$) does. Formally, this means that we consider classes of equivalent points instead of single points, where points are called equivalent if one is the scalar multiple of the other one.

If we put $x_3 = 0$ in the basic equation (1), then this equation collapses to $-x_1^3 = 0$, leading to $x_1 = 0$. Thus, the equivalence class which includes the element $(0, 1, 0)$ is the only one that contains a point with $x_3 = 0$. For all points on the curve that are not equivalent to $(0, 1, 0)$, we may apply the following transformation

$$K \times K \times (K \setminus \{0\}) \ni (x_1, x_2, x_3) \mapsto (x, y) := \left(\frac{x_1}{x_3}, \frac{x_2}{x_3} \right) \in K \times K,$$

which reduces the number of variables to two instead of three. We note that the basic equation (1) $F(x_1, x_2, x_3) = 0$ was chosen in a way that this transformation leads to the famous so-called Weierstrass-Equation¹⁴⁴ holds

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6. \quad (2)$$

element *ZERO* and one-element *ONE* are isomorphic. We note that mathematical objects are only defined by their mathematical properties.

¹⁴²For prime fields additive as well as multiplicative group are cyclic. Furthermore, each field $GF(p^n)$ contains a subfield that is isomorphic to the prime field \mathbb{Z}_p .

¹⁴³This curve is given by the zeros of a *polynomial* F of degree three in three variables. In general, expressions of the form $P = \sum_{i_1, \dots, i_n \in \mathbb{N}_0} a_{i_1 \dots i_n} x_1^{i_1} \dots x_n^{i_n}$ with coefficients $a_{i_1 \dots i_n} \in K$ are called polynomials in n variables x_1, \dots, x_n with underlying field K , if $\deg P := \max\{i_1 + \dots + i_n : a_{i_1 \dots i_n} \neq 0\}$ is finite, i.e. the sum has only finitely many non-zero terms (monomials). The sum of the exponents of each term of the sum is at most 3, at least one term of the sum has a single variable with 3 as value of the according exponent.

¹⁴⁴Karl Weierstrass, 31.10.1815–19.12.1897, German mathematician, famous for his rigorous formal approach to mathematics.

Since all but one point (i.e. equivalence class) of the elliptic curve can be described using equation (2), this equation is often called the elliptic equation, and its solutions written as

$$\mathbf{E} = \{(x, y) \in K \times K \mid y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6\} \cup \{\mathcal{O}\}.$$

Here, \mathcal{O} represents the point $(0, 1, 0)$ that is loosely speaking mapped to infinity by the transformation (division by x_3) that reduces the three variables to two.

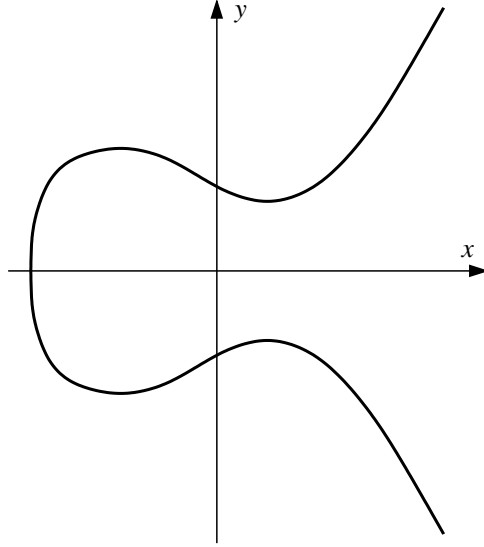


Figure 3: Example of an elliptic curve with the real numbers as underlying field.

In contrast to figure 3 only finite fields $K = GF(p^n)$ are used in elliptic curve cryptography. The reason is loosely speaking that in modern communication engineering data processing is always based on discrete data (simply because computers accept only discrete data).

For practical reasons, it turned out to be useful to take either $GF(p)$ with a large prime p or $GF(2^n)$ with a (large) positive integer n . Using $GF(p)$ has the advantage of providing a relatively simple arithmetic; on the other hand $GF(2^n)$ allows a binary representation of each element that supports the way computers work. Other fields like, for example, $GF(7^n)$ do not have any of these advantages and are, thus, not considered, although there is no mathematical reason why they should not.

A coordinate transformation can result in a simpler version¹⁴⁵ of the Weierstrass equation. Depending whether $p > 3$, different transformations are used, and we obtain

- in case of $GF(p)$, $p > 3$, the elliptic curve equation of the form

$$y^2 = x^3 + ax + b \tag{3}$$

with $4a^3 + 27b^2 \neq 0$

¹⁴⁵Such a coordinate transformation is combination of a rotation and a dilatation of the coordinate system without changing the elliptic curve itself.

- in case of $GF(2^n)$ the elliptic curve equation of the form

$$y^2 + xy = x^3 + ax^2 + b \quad (4)$$

with $b \neq 0$ ¹⁴⁶.

This conditions on the parameters a, b ensure that the elliptic equation can be used in the context of cryptography¹⁴⁷.

Let $|E|$ denote the number of elements of an elliptic curve E given an underlying field $GF(k)$ (for practical reasons either $k = p$ with p prim or $k = 2^n$). Then Hasse's theorem[Silverman1986] yields $||E| - k - 1| \leq 2 \cdot \sqrt{k}$. This Inequality is equivalent to $k + 1 - 2\sqrt{k} < |E| < k + 1 + 2\sqrt{k}$. In particular, this means that the number of elements of an elliptic curve is approximately k (for large k).

6.5 Operating on the elliptic curve

In order to work with elliptic curves in practice, we define an operation (often written in an additive way $+$) on the set of points on the curve. If we have a curve over the field $GF(p)$, we define the commutative operation $+$ by

1. $P + \mathcal{O} = \mathcal{O} + P = P$ for all $P \in E$,
2. for $P = (x, y)$ and $Q = (x, -y)$ we set $P + Q = \mathcal{O}$,
3. for $P_1 = (x_1, y_1), P_2 = (x_2, y_2) \in E$ with $P_1, P_2 \neq \mathcal{O}$ and $(x_2, y_2) \neq (x_1, -y_1)$ we set $P_3 := P_1 + P_2, P_3 = (x_3, y_3)$ defined by

$$x_3 := -x_1 - x_2 + \lambda^2, \quad y_3 := -y_1 + \lambda(x_1 - x_3)$$

with the auxiliary quotient

$$\lambda := \begin{cases} \frac{y_1 - y_2}{x_1 - x_2} & \text{if } P_1 \neq P_2, \\ \frac{3x_1^2 + a}{2y_1} & \text{if } P_1 = P_2. \end{cases}$$

In particular, we obtain $-P = (x, -y)$ for $P = (x, y) \in E$.

If we deal with a curve over the field $GF(2^n)$, we define the operation $+$ in an analogous way by

1. $P + \mathcal{O} = \mathcal{O} + P = P$ for all $P \in E$,
2. for $P = (x, y)$ and $Q = (x, x + y)$ we set $P + Q = \mathcal{O}$,

¹⁴⁶The form (3) is called the standard form of the Weierstrass-equation. If the characteristic of the field is 2 or 3, we obtain $4 = 0$ respectively $27 = 0$, which means that the condition on parameters a, b collapse. Loosely speaking, this is the reason why the transformation to the standard form does not work in these cases.

¹⁴⁷Formally we call such curves non singular.

3. for $P_1 = (x_1, x_2), P_2 = (x_2, y_2) \in E$ with $P_1, P_2 \neq \mathcal{O}$ and $(x_2, y_2) \neq (x_1, x_1 + y_1)$ we set $P_3 := P_1 + P_2, P_3 = (x_3, y_3)$ defined by

$$x_3 := -x_1 + x_2 + \lambda + \lambda^2 + a, \quad y_3 := y_1 + x_3 + \lambda(x_1 + x_3)$$

with auxiliary quotient

$$\lambda := \begin{cases} \frac{y_1 + y_2}{x_1 + x_2} & \text{if } P_1 \neq P_2, \\ x_1 + \frac{y_1}{x_1} & \text{if } P_1 = P_2. \end{cases}$$

In particular, we obtain $-P = (x, -y)$ for $P = (x, y) \in E$.

(Note that $-(-P) = (x, x + (x + y)) = (x, 2x + y) = (x, y)$, since the underlying field has characteristic 2.)¹⁴⁸

One can verify that $+$ defines a group operation on the set $E \cap \{\mathcal{O}\}$. In particular this means that the sum of two points is again a point on the elliptic curve. How this operation works is geometrically visualized in the following section.

¹⁴⁸An animation of the addition of points on elliptic curves can be found on the Certicom-Homepage http://www.certicom.com/resources/ecc_tutorial/ecc_tutorial.html

How to add points on an elliptic curve

The following figures show how points on an elliptic curve over the field of real numbers are summed up using affine coordinates. We note that the point infinity \mathcal{O} cannot be shown in the affine plane.

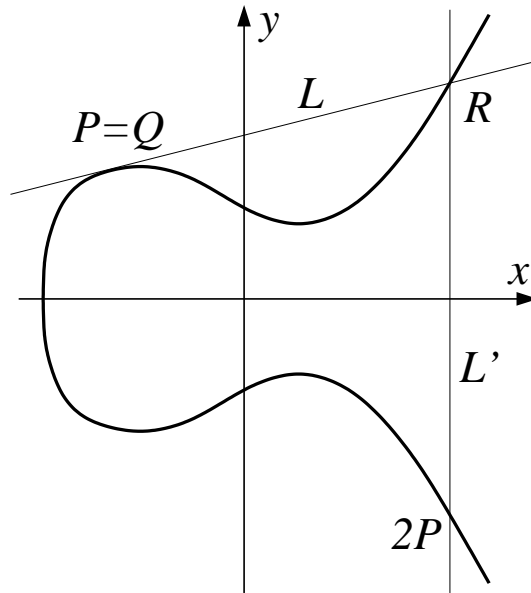


Figure 4: Doubling of a point

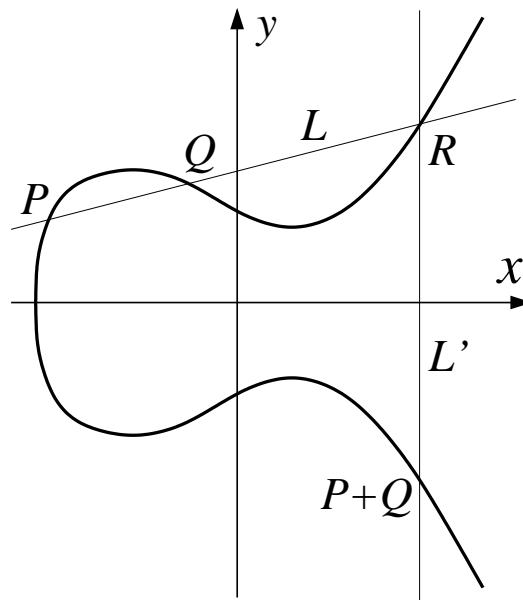


Figure 5: Summing up two different points over the real number field

6.6 Security of elliptic-curve-cryptography: The ECDLP

As mentioned above in section 6.4, we only consider elliptic curves over the finite¹⁴⁹ fields $GF(2^n)$ or $GF(p)$ (for a large prime p). This means that all parameters that describe the curve are taken from this underlying field. If E is an elliptic curve over such a field and P is a point on the curve E , then we can derive for all positive integers m

$$mP := \underbrace{P + P + \cdots + P}_{m \text{ times}}.$$

Looking on this operation from the cryptographic point of view, it turns out to be very interesting by the following reason: On the one hand one needs only $\log m$ operations to calculate mP — one simply has to calculate $P, 2P, 2^2P, 2^3P, \dots$, write m in a binary form and finally add all these multiples 2^kP of P with respect to the binary representation of m — on the other hand it seems to be very hard to find m given P and $Q = mP$ on E . Of course, we may simply calculate $P, 2P, 3P, 4P, 5P, \dots$ and compare each of them with Q . But this will take as much as m operations.

Yet there is no algorithm known that efficiently derives m given P and G . The best algorithms known so far need about \sqrt{q} operations where q is the (largest) prime factor of $p - 1$, in case the underlying field is $GF(p)$; here m should be between 1 and q liegen so that one needs at most $\log q$ operations to calculate mP . However, the quotient $\frac{\sqrt{q}}{\log q}$ tends to $+\infty$ very fast for large q .

If we choose the parameters sufficiently large (for example, let p be prime and at least 160 bits long), a computer will easily be able to calculate mP (in less than a second). The *inverse problem* however, to derive m from mP and P , can (still) not be solved in reasonable time.

This problem is known as the “over Elliptic Curve Discrete Logarithm Problem” (for short ECDLP).

In elliptic curve cryptography we formally look at points on the elliptic curve as elements of a group with point addition $+$ as operation. Furthermore, we use only elliptic curves that have a sufficiently large number of points. However, in special cases curves may be weak and not useful due to other reasons. For such special cases the ECDLP can be much easier to solve than in the general case. This means that one has to look carefully at the parameters when choosing an elliptic curve for cryptographic applications.

Not useful for cryptography are *a-normal* (that are curves over \mathbb{Z}_p , for which the set \mathbf{E} consists of exactly p elements) and *supersingular* curves (that are curves, for which the ECDLP can be reduced to the “normal” discrete logarithms in another, smaller finite field). This means that there are cryptographically useful and non-useful elliptic curves. Given the parameters a and b , it is possible to determine whether a curve is useful or not. In many publications one can find parameters that turned out to be useful for cryptography. The open (scientific) discussion guarantees that these results take into account latest research.

Given a secure curve, the time that is needed to solve the ECDLP is strongly correlated with parameter p in case $GF(p)$ respectively n in case of $GF(2^n)$. The larger these parameters become,

¹⁴⁹Discrete in contrast to continuous.

the more time an attacker needs to solve the ECDLP — at least with the best algorithms known so far. Experts recommend bit-lengths of 200 for p for secure curves. A comparison with RSA modulus length shows why elliptic curves are so interesting for applications. We note that the computation effort for signing and encryption is closely related to the bit-length of the parameters. In addition the initiation process, i.e. the generation of the private-public-key-pair, becomes more complicated the larger p is. Thus, one looks for the smallest parameters that still come along with the security required. It is remarkable that a length of 200 bits for p is sufficient to construct a *good* elliptic curve that is as secure as RSA with a 1024 bit RSA modulus (as far as we know today). For short, the reason for this advantage of ECC lies in the fact that the best algorithms known for solving the ECDLP need exponential time while the best algorithms for factorizing are sub-exponential (number field sieve, quadratic sieve or factorizing with elliptic curves). Hence, the parameters for a cryptosystem that is based on the problem of *factorizing large integers* have to be larger than the parameters for a system based on ECDLP.

6.7 Encryption and signing with elliptic curves

The *elliptic curve discrete logarithm problem* (ECDLP) is the basis for elliptic curve cryptography. Based on this problem, there are different signature schemes. In order to apply one of these, we need:

- An elliptic curve \mathbf{E} with an underlying field $GF(p^n)$.
- A prime $q \neq p$ and a point G on the elliptic curve \mathbf{E} with order q . This means that $qG = \mathcal{O}$ and $rG \neq \mathcal{O}$ for all $r \in \{1, 2, \dots, q-1\}$. Thus q is a factor of the group order (i.e. the number of elements) $\#\mathbf{E}$ of E . Since q is prime, G generates a cyclic sub-group of \mathbf{E} of order q .

The parameters mentioned are often called *Domain* parameter. They describe the elliptic curve \mathbf{E} and the cyclic sub-group of \mathbf{E} on which the signature scheme is based.

6.7.1 Encryption

Using elliptic curves one can construct a key exchange protocol based on the Diffie-Hellman protocol (see chapter 4.4.2). The key exchanged can be used for a subsequent symmetric encryption. We note that in contrast to RSA there is no pair of private and public key that can be used for encryption and decryption!

In the notation of elliptic curves, the Diffie-Hellman protocol reads as follows: First both partners (A und B) agree on a group G and an integer q . Then they choose $r_A, r_B \in \{1, 2, \dots, q-1\}$ at random, derive the points $R_A = r_A G$, $R_B = r_B G$ on the elliptic curve and exchange them (using an insecure channel). After that A easily obtains $R = r_A R_B$; B gets the same point ($R = r_A r_B G$) by calculating $r_B R_A = r_B r_A G = r_A r_B G = R$. We note that R_A, R_B are easy to derive as long as r_A respectively r_B are known G . However, the inverse operation, to get R_A respectively R_B from r_A respectively r_B is hard.

Using the best algorithms known so far, it is impossible for any attacker to obtain R without knowing either r_A or r_B — otherwise he would have to solve the ECDLP.

In order to prohibit a “Man-in-the-middle” attack, one may sign the values G, q, R_A, R_B as described in chapter 5.4.1.

6.7.2 Signing

Using the DSA signature scheme, one can proceed as follows: The signing party chooses a (non-trivial) number $s \in \mathbb{Z}_q$, which will be the private key, and publishes q, G and $R = sG$. We note that s cannot be obtained from G and R are not sufficient — a fact on which the security of the signature scheme is based.

Given the message m , which should be signed, one first constructs a digital finger print using a hash-algorithm h such that $h(m)$ has its values in $\{0, 1, 2, \dots, q-1\}$. Thus, $h(m)$ can be considered as an Element of \mathbb{Z}_q . Then the signing party chooses a random number $r \in \mathbb{Z}_q$ and derives $R = (r_1, r_2) = rG$. We note that the first component r_1 of R is an element of $GF(p^n)$. This component will then be projected onto \mathbb{Z}_q , i.e. in case of $n = 1$ it is interpreted as the remainder of an element of $\{0, 1, \dots, p-1\}$ divided by q . This projection of r_1 onto \mathbb{Z}_q is denoted by \bar{r}_1 . Then one determines $x \in \mathbb{Z}_q$ such that

$$rx - s\bar{r}_1 - h(m) = 0.$$

The triple (m, r_1, x) is then published as the digital signature of message m .

6.7.3 Signature verification

In order to verify a signature, one has to build $u_1 = h(m)/x$, $u_2 = \bar{r}_1/x$ (in \mathbb{Z}_q and derive

$$V = u_1G + u_2Q.$$

Since we have $Q = sG$, the point $V = (v_1, v_2)$ satisfies $v_1 = u_1 + u_2s$. We note that this operations take place in the field $GF(p^n)$. The projection of $GF(p^n)$ on \mathbb{Z}_q mentioned above should be chosen in such a way that $\bar{v}_1 = u_1 + u_2s$ is an element of \mathbb{Z}_q . Then it follows that

$$\bar{v}_1 = u_1 + u_2s = h(m)/x + \bar{r}_1s/x = (h(m) + \bar{r}_1s)/x = rx/x = r.$$

Since $R = rG$, we obtain $\bar{v}_1 = \bar{r}_1$, i.e. R and V coincide modulo the projection onto \mathbb{Z}_q .

6.8 Factorisation using elliptic curves

There are factorisation¹⁵⁰ algorithms based on elliptic curves¹⁵¹. More precisely, these procedures exploit the fact that elliptic curves can be defined over \mathbb{Z}_n (n composite number). Elliptic curves over \mathbb{Z}_n do not form a group, because not every point on such an elliptic curve has an inverse point. This is connected with the fact that - if n is a composite number - there exist elements in \mathbb{Z}_n that do not have an inverse with respect to multiplication mod n . In order to add two points on an elliptic curve over \mathbb{Z}_n , we can calculate in the same way as on elliptic curves over \mathbb{Z}_p . Addition of two points (on an elliptic curve over \mathbb{Z}_n), however, fails if and only if a factor of n has been found. The reason for this is that the procedure for adding points on elliptic curves gives elements in \mathbb{Z}_n and calculates the inverse elements for these (with respect to multiplication mod n) in \mathbb{Z}_n . The extended Euclidean algorithm is used here. If the addition of two points (that lie on an elliptic curve over \mathbb{Z}_n) gives an element in \mathbb{Z}_n that does not have an inverse element in \mathbb{Z}_n , then the extended Euclidean algorithm delivers a genuine factor of n .

Factorisation using elliptic curves thus principally works as follows: Random curves over \mathbb{Z}_n are selected, as well as random points (that lie on this curve) and add them; you thus obtain points that also lie on the curve or find a factor of n . Factorisation algorithms based on elliptic curves therefore work probabilistically. The opportunity of defining large number of elliptic curves over \mathbb{Z}_n allows you to increase the probability of finding two points which you can add to obtain a factor of n . These procedures are therefore highly suitable for parallelisation.

6.9 Implementing elliptic curves

CrypTool also offers elliptic curves for the digital signature function.

It implements the basic algorithms for group operations, for generating elliptic curves, for importing and exporting parameters for elliptic curves over finite fields with p (p prime) elements. The algorithms have been implemented in ANSI C and comply with draft no. 8 of the IEEE P1363 work group *Standard Specifications for Public Key Cryptography*

<http://grouper.ieee.org/groups/1363>.

The procedure implements the cryptographic primitives for generating and verifying signatures for the variations of Nyberg-Rueppel signatures and DSA signatures based on elliptic curves (in accordance with draft no. 8 of the IEEE P1363 work group). This was done in collaboration with the Secude GmbH — using the above library and the Secude SDK.

In case one uses the field $GF(2^n)$ instead of the prime field $GF(p)$, one has to make substantial changes in the implementation. The advantage of $GF(2^n)$ lies in the fact that calculations in $GF(2^n)$ can be implemented very efficiently using the binary representation. In particular,

¹⁵⁰Especially John M. Pollard was involved in the development of many different factorisation algorithms; also at factorisation with ECC he was one of the leading heads. As an employee of British Telekom he never published much. At the RSA data Security Conference in 1999 he was awarded for his “outstanding contributions in mathematics”: http://www.eff.org/Privacy/Crypto/misc/DESCracker/HTML/19990118_rsa_awards.html.

¹⁵¹In 1987 H.W. Lenstra published a factorization algorithm, based on elliptic curves (see [Lenstra1987]). The biggest compound number currently factorised with elliptic curves is the number $628^{59} - 1$, which has 55 decimal digits. It was found Oct. 6th, 2001 by M. Izumi (See [ECMNET](#)).

divisions are much easier to process compared to $GF(p)$ (this is particularly important in the signature scheme mentioned above where a division is needed for processing a signature as well as for the verification).

In order to achieve maximal gain in efficiency, one may choose a field that allows special basis like polynomial basis (useful for software implementations) or normal basis (best for hardware implementations). For special n (like, for example, $n = 163, 179, 181$) one may even combine both advantages. However, they are still non-standard.

Sometimes only the first component and one additional bit is used as representation of a point on the elliptic curve instead of the full two components. Since the first component together with the additional bit is sufficient to derive the full point, this representation minimizes the memory capacity needed. In particular, for normal basis this point compression can be implemented efficiently. In addition, the cryptographic protocols themselves become more effective. A disadvantage is, however, that *point compression* can be used for about half of all elliptic curves only and is protected under US patent (US Patent 6141420, Certicon), causing additional costs. In the general case $GF(p^n)$ (and also in case $n = 1$) often so called affine or projective co-ordinates are used. Depending on the application, these co-ordinates may result in a gain in efficiency as well.

A comprehensive description of all implementations and their advantages and disadvantages would go far beyond the scope of this paper. We only want to state that there is a variety of possible implementations for elliptic curve cryptography, much more than for RSA. Therefore, there are serious efforts to reduce this large to a small number of standard implementations. Some standardization committees even try to reduce the complexity by focussing on a small number of (prescribed) curves (ASC-approach).

Today it is still not clear whether these standardization initiatives will be successful or not. However, without agreed standards, ECC is not likely to become a real alternative for RSA. The committees might be forced to act fast if there was a break-through in factorization.

6.10 Elliptic curves in use

Today elliptic curve cryptography is already in use. A prominent example is the information network Bonn-Berlin¹⁵², used for the exchange of strictly confidential documents between different German federal governmental institutions in Berlin and Bonn. With the help of ECC a high security solution could be realized. Interoperability, however, played only a minor role.

Based on information from the head of the Austrian e-Government projects, Prof. Posch, a smartcard based on ECC will shortly be launched in Austria: A bank card that allows digital signing will be issued in Austria from 2004 on to all citizens.

Both examples show the typical range of application for elliptic curve cryptography: For high security solutions and for implementations on smartcards in which the key length is crucial (because of physical memory available).

¹⁵²The Informationsverbund Bonn-Berlin (IVBB) connects governmental institutions in the old and new German capital.

References

- [Cassels1991] J. W. S. Cassels,
Lectures on elliptic curves,
Cambridge University Press, 1991, 143 pages.
- [Koblitz1984] N. Koblitz,
Introduction to elliptic curves and modular forms,
Graduate Texts in Mathematics, Springer-Verlag, 1984.
- [Koblitz1998] N. Koblitz,
Algebraic aspects of Cryptography. With an appendix on Hyperelliptic curves by Alfred J. Menezes, Yi Hong Wu and Robert J. Zuccherato,
Springer-Verlag, 1998, 206 pages.
- [Menezes1993] A. J. Menezes,
Elliptic curve public key cryptosystems,
Kluwer Academic Publishers, 1993.
- [Lenstra1987] H.W. Lenstra,
Factoring integers with elliptic curves,
Annals of Mathematics 126, pp. 649-673, 1987.
- [Lenstra1999] Arjen K. Lenstra, Eric R. Verheul
Selecting Cryptographic Key Sizes (1999),
Journal of Cryptology: the journal of the International Association for Cryptologic Research
<http://www.cryptosavvy.com/cryptosizes.pdf>
- [Silverman1986] J. Silverman,
The Arithmetic of Elliptic Curves,
Springer-Verlag, 1986.
- [Silverman1992] J. Silverman,
The arithmetic of elliptic curves,
Graduate Texts in Mathematics, Springer-Verlag, 1992.
- [SilvermanTate1992] J. Silverman, J. Tate,
Rational points on elliptic curves,
Springer-Verlag, 1992.

Web links

1. Certicom Online Tutorial,
http://www.certicom.com/resources/ecc_tutorial/ecc_tutorial.html

2. Working group IEEE P1363
<http://grouper.ieee.org/groups/1363>
3. An informative web page about factorisation with elliptic curves.
<http://www.loria.fr/~zimmerma/records/ecmnet.html>
It contains literature related to the topic factorisation with elliptic curves as well as links to other web page.
4. Key length comparison by Arjen Lenstra and Eric Verheul
<http://cryptosavvy.com/table.htm>

A CrypTool Menus

This appendix contains the complete menu tree of CrypTool version 1.3.04. Please notice that not all menu items are always available, because the menu tree is built up dynamically: which menu items are available depends on the type of the currently active document window. The brute-force analysis for DES e. g. is only available, if the active window is opened in the hexadecimal view. On the other hand the menu item “Generate Random Numbers...” is always available. Within the figures 6 and 7 this relation is shown in the brackets after the menu items, e. g. “Find... [ATH]” means, that the function “Find...” is only available for ASC-, text- and hexadecimal windows.

Code letter	Type of document
M	(no document opened)
A	ASC
T	Text
H	Hexadecimal
P	Plot

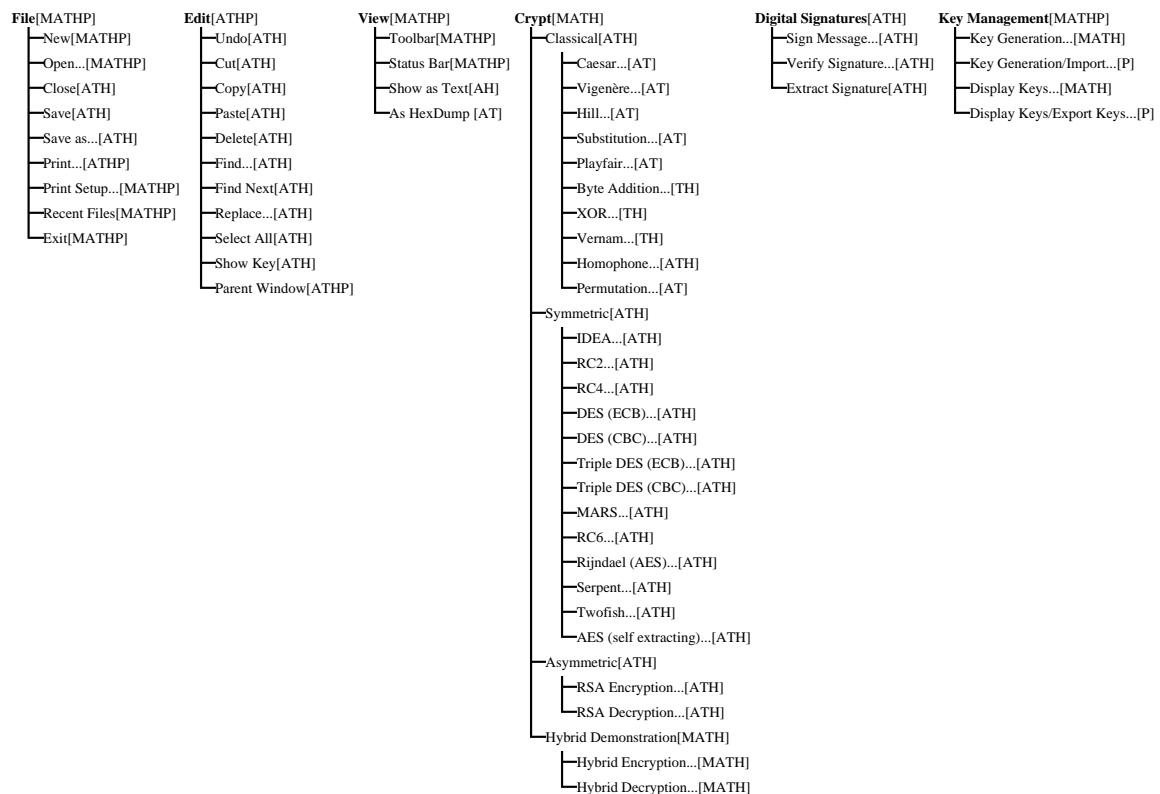


Figure 6: Detailed view of the CrypTool menu tree – part 1

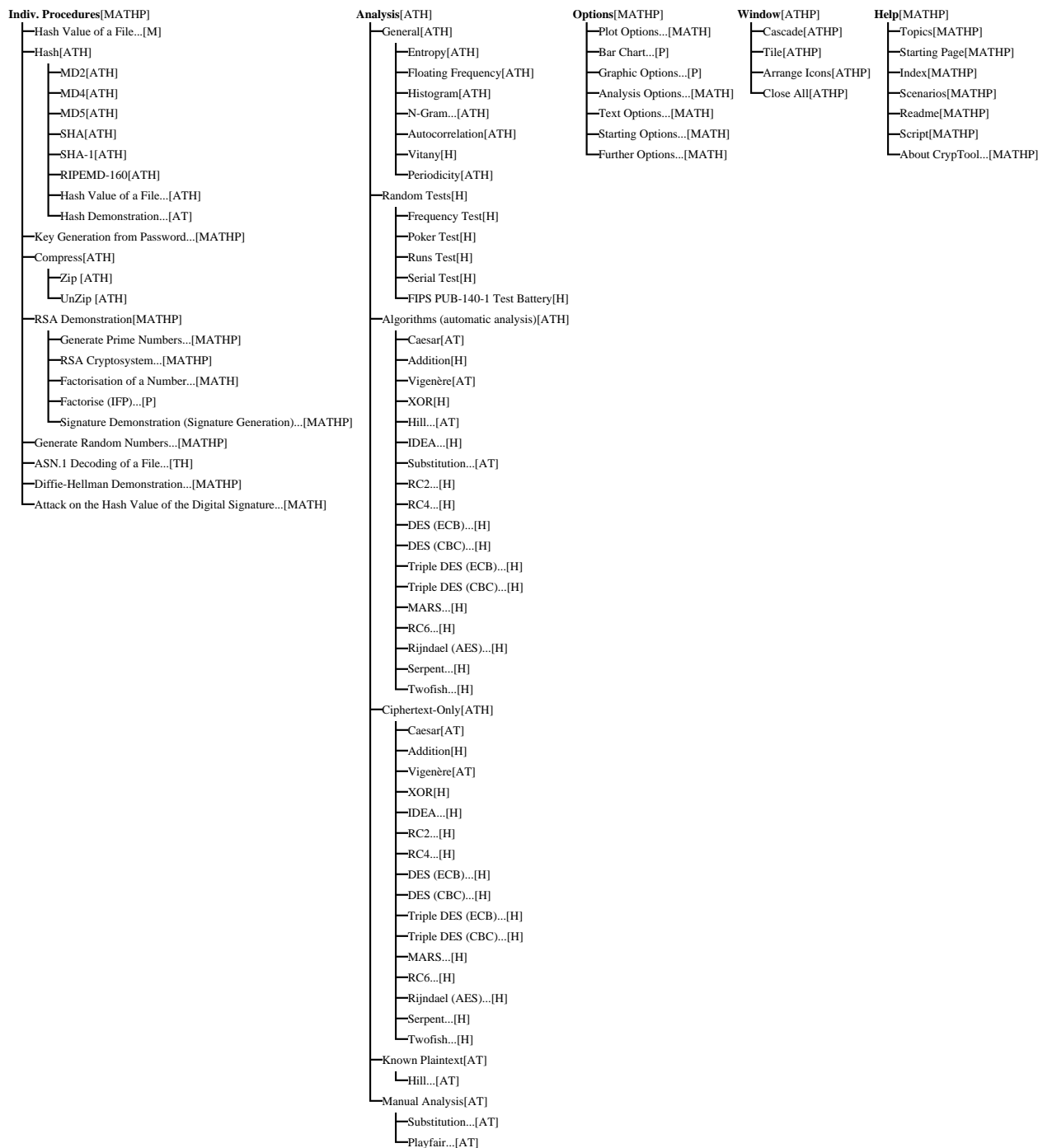


Figure 7: Detailed view of the CrypTool menu tree – part 2

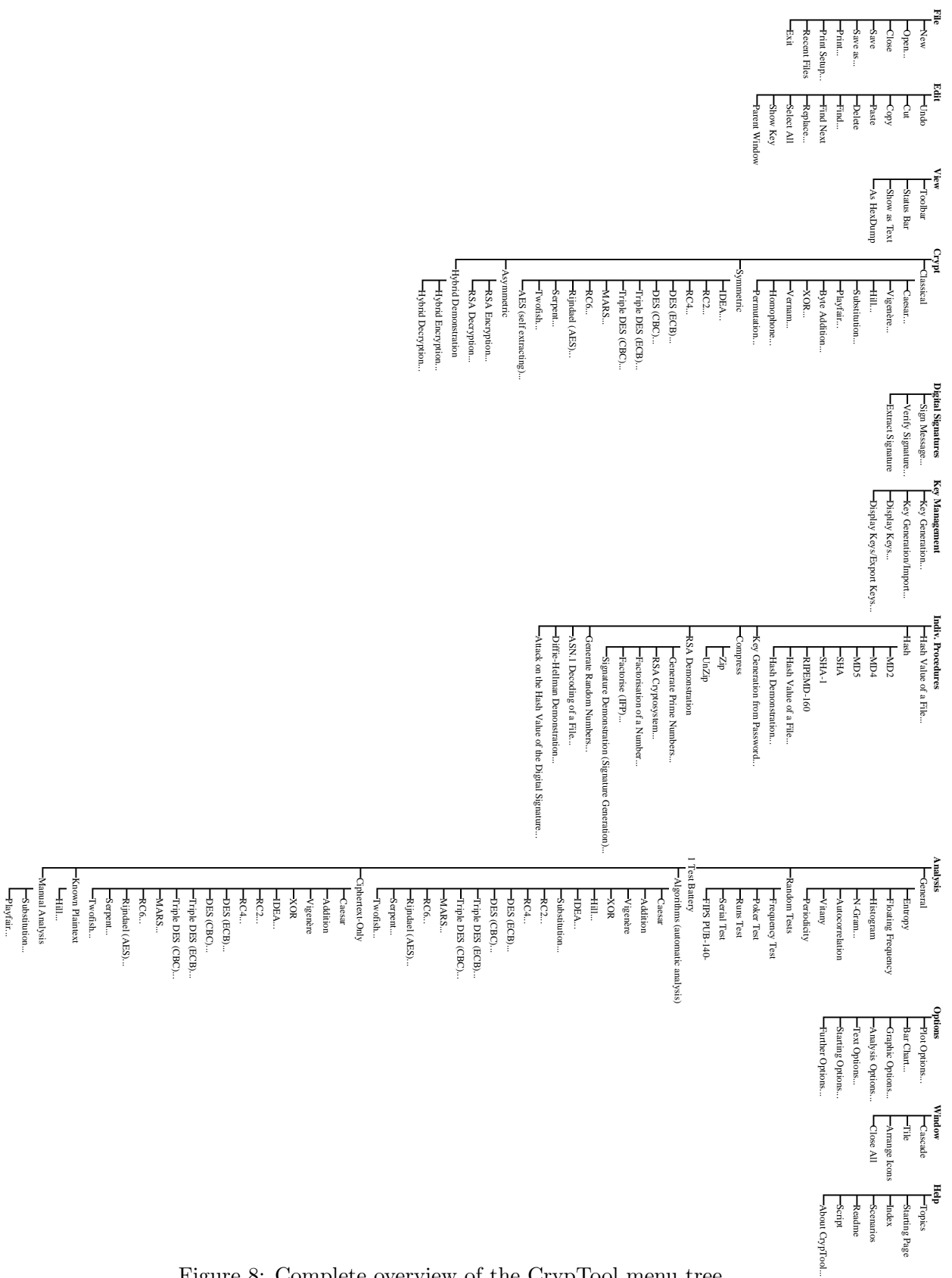


Figure 8: Complete overview of the CryptTool menu tree

B Authors of this script

This appendix lists the authors of this document. Please refer to the top of the individual chapters for their contribution.

Bernhard Esslinger, initiator of the CrypTool project, head IT security at Deutsche Bank and lecturer on IT security at the University of Siegen. E-mail: besslinger@web.de.

Matthias Büger, contribution to the chapter “Elliptic Curves”, research analyst at Deutsche Bank and responsible for Deutsche Bank’s crypto standard.

Bartol Filipovic, original author of the CrypTool elliptic curve implementation and the corresponding chapter in this script.

Henrik Koy, main developer and co-ordinator of CrypTool development since version 1.3; script reviewer and T_EX guru; project leader IT at Deutsche Bank.

Roger Oyono, implementer of the CrypTool factorization dialogue and original author of chapter “The Mathematical Ideas behind Modern Cryptography”.

Jörg Cornelius Schneider, design and support of CrypTool; crypto enthusiast and project leader IT at Deutsche Bank.

Index

- Aaronson 2003, 40
- Addition, 54, 62
- Adleman 1982, 122
- Adleman Leonard, 13, 114, 115
- Agrawal 2002, 95
- AKS, 36, 80
- Alice, 12
- Arthur Conan Doyle, 71
- Associative law, 52
- Attack
 - birthday, 124
 - brute-force, 7, 10–12, 147
 - chosen-cipher text, 117
 - cipher text only, 93
 - known plaintext, 93
 - man-in-the-middle, 127
- Authenticity, 13, 127
 - user, 123
- Baby-Step-Giant-Step, 120
- Balcazar 1988, 122
- Bartholome 1996, 40, 95
- Bauer 1995, 95
- Bauer 2000, 95
- BC, 98
- Berne Eric, 64
- Bernstein 2001, 95
- Beutelspacher 1996, 95
- Block length, 88–90
- Blum 1999, 40
- Bogk 2003, 97
- Bourseau 2002, 95
- Brands 2002, 95
- Brickell 1985, 122
- Brickell Ernst, 114
- Buchmann 1999, 95
- Buhler 1993, 95
- Bundschuh 1998, 40
- C158, 81
- Caldwell Chris, 42
- Capital letters alphabet, 88, 94
- Cascade cipher, 11
- Cassels 1991, 145
- Catalan Eugene, 32
- Certicom, 138, 145
- Certification
 - public key, 126
- Certification authority (CA), 127
- Closeness, 53, 61, 102
- Co-prime, 112, 115
- Cole Frank Nelson, 22
- Collision, 123, 124
- Commutative law, 52
- Complexity, 75, 110, 121, 132
- Congruence, 50, 51
- Congruent, 51
- Coppersmith 2002, 15
- Courtois 2002, 15
- Cryptanalysis, 10, 89, 91, 93
- Cryptography
 - modern, 17, 83, 110
 - public key, 17, 71, 111, 112
- CrypTool, 1, 7–10, 12–14, 22, 71, 72, 81, 84, 87–91, 93, 94, 98, 115, 118, 123, 124, 143
- Cunningham project, 27, 42, 81, 99
- Dedekind Julius, 45
- DES, 9
- Diffie Whitfield, 13, 84, 118
- Diffie-Hellman, 45, 84, 118, 119, 141
- Discrete logarithm, 85, 86, 117
- Distributive law, 52
- Divisibility, 50
- Divisible, 50
- Division modulo n , 52, 54
- Divisor, 50
- Domain parameter, 141
- DSA, 13, 142, 143
 - signature, 125
- ECDLP, 140, 141
- Eckert 2003, 72, 89, 95

ECMNET, 143
 EFF, 24
 ElGamal, 13
 public key, 119
 Elliptic curves, 129
 Elsnar 1999, 71
 Encryption, 9
 asymmetric, 12, 44, 110
 El Gamal public key, 119
 hybrid, 13
 Merkle-Hellman, 113
 public key, 110
 symmetric, 9
 Eratosthenes
 sieve, 24, 34
 Erdős Paul, 35
 Ertel 2001, 95
 Euclid, 19
 Euclid's proof by contradiction, 20
 Euclidean algorithm, 143
 extended, 57, 66, 100
 Euclidean number, 29
 Euler
 (phi) function, 56, 61, 64, 115
 Euler Leonhard, 64, 65
 Exponential function
 calculation, 119
 discrete, 118

 Factor, 50
 Factorisation, 22, 67, 132
 factoring records, 36, 80, 98, 143
 factorisation problem, 67, 77, 93, 115
 Ferguson 2001, 15
 Fermat
 last theorem, 97
 little theorem, 25, 56
 number, 24, 27
 generalized, 28
 prime number, 27
 Fermat Pierre, 24, 65
 Fibonacci, 45, 98
 Field, 132
 characteristic, 133
 finite, 134
 FIPS180-2, 128
 FIPS186, 128
 FIPS186-2, 128
 Fixpoint, 64, 66
 Fox 2002, 95

 Gödel Kurt, 35
 Gallot Yves, 27, 28
 Gauss bracket, 100
 Gauss Carl Friedrich, 27, 33, 44, 45, 49, 71
 Gcd, 44, 56, 61, 100, 115
 General Number Field Sieve (GNFS), 75–77,
 79, 81, 82
 GIMPS, 23, 28, 42
 GISA, 72, 77, 82, 98, 129
 GISA 2002, 95
 Goldbach Christian, 35
 Graham 1989, 40
 Graham 1994, 45, 96
 Grid computing, 76
 Group, 44, 61, 119, 132
 cyclic, 133

 Hash function, 123, 124
 Hash value, 124
 Hellman Martin, 13, 84, 113, 118
 Hybrid procedure, 13

 IDEA, 10
 Identity, 53
 Impersonation attack, 126
 Inverse
 additive, 53
 multiplicative, 53
 Invertibility, 63
 Isau 2003, 71
 ISO/IEC 9594-8, 128
 ITU-T, 128
 IVBB, 144

 Key
 private, 110
 public, 12, 110
 secret, 12

- Key agreement (key exchange)
 - Diffie-Hellman, 84, 118
- Key management, 13, 14
- Kippenhahn 1997, 96
- Kippenhahn 1999, 96
- Klee 1997, 40
- Knapsack, 112
 - Merkle-Hellman, 113
- Knott, Ron, 45, 98
- Knuth 1981, 40
- Knuth 1998, 96
- Koblitz 1984, 145
- Koblitz 1998, 145
- Koblitz Neal, 131
- Kronecker Leopold, 45

- Lagarias 1983, 122
- Lagarias Jeff, 114
- Legendre Adrien-Marie, 33, 71
- Lenstra 1987, 145
- Lenstra 1993, 96
- Lenstra 2002, 96
- Lenstra/Verheul 1999, 145
- LiDIA, 86, 98
- Literature, 71
- Logarithm, 60, 119
- Logarithm problem
 - discrete, 60, 85, 117, 120, 125, 132
- Long integer, 59
- Lorenz 1993, 40
- Lucas Edouard, 21, 24
- Lucks 2002, 15
- Lucks 2003, 97

- Mathematica, 57, 86, 98, 105
- Menezes 1993, 145
- Menezes 2001, 96
- Merkle 1978, 122
- Merkle Ralph, 113
- Mersenne
 - number, 21, 22
 - generalized, 27, 28
 - prime number, 21, 22, 26, 36, 42
 - M-37, 22
 - M-38, 22, 27
 - M-39, 23
 - M-40, 23
 - theorem, 21
- Mersenne Marin, 21, 24
- Message integrity, 123
- Miller Gary L., 26
- Miller Victor, 131
- Modulus, 50
- Moore's law, 76
- Movies, 71
- Multiplication, 54, 62
- Münchenbach Carsten, 98

- Nichols 1996, 15
- NIST, 125
- NSA, 10
- Number
 - Carmichael, 25, 29
 - Catalan, 32
 - composite, 18, 48
 - Fermat, 24, 27
 - Mersenne, 21
 - natural, 17, 45
 - prime, 17, 18
 - pseudo prime, 25, 29
 - strong pseudo prime, 25, 29
- Number theory
 - elementary, 44, 48
 - fundamental theorem, 18, 49, 61
 - introduction, 45
 - modern, 46

- One Time Pad, 9
- One way function, 61, 83, 110
 - with trapdoor , 110
- Open Source, 72
- Order
 - maximum, 68
 - multiplicative, 68

- P(n), 33
- Padberg 1996, 40
- Palladium, 79
- Pari-GP, 57, 86, 98, 99, 105

Patent, 72
 Performance, 17, 76, 107, 116, 123, 129
 Permutation, 57, 70, 112
 Pfleeger 1997, 96
 $PI(x)$, 33
 Pieper 1983, 40
 PKCS#1, 128
 PKI, 126
 Pohlig, 118
 Pollard John M., 143
 Polynomial, 31, 36, 76, 80, 110–112, 114, 135
 Pomerance 1984, 96
 Power, 59
 Primality testing, 36, 80, 82
 Prime factor, 49
 decomposition, 49, 61, 64, 115
 Prime number, 17, 48
 density, 32
 Fermat, 27
 formula, 26
 gigantic, 22
 Mersenne, 22, 26, 36
 number of, 71
 pseudo prime, 25, 29
 relative prime, 30, 115
 strong pseudo prime, 25, 29
 test, 22, 24, 132
 theorem, 33
 titanic, 22
 Problem of discrete logarithm, 141
 Proof by contradiction, 20, 21

 Rabin
 public key procedure, 117
 Rabin Michael O., 26, 117
 Raising to the power, 58
 Random, 13, 126
 RC5, 12
 Reducibility, 52
 Relatively prime, 30, 56, 57, 66
 Remainder class, 50
 Remainder set
 full, 63
 reduced, 63

 Richstein 1999, 35, 40
 Riemann Bernhard, 35
 RIPEMD-160, 124
 Rivest Ronald, 13, 115
 Robshaw 2002, 15
 Root, 60
 Rowling Joanne, 48, 83
 RSA, 13, 17, 45, 59, 65, 66, 71, 72, 86, 115
 cipher challenge, 91, 93
 modulus, 141
 RSA procedure, 71
 signature, 125
 RSA 1978, 122
 RSA Laboratories, 128
 RSA Security 2002, 96
 RSA-155, 80
 RSA-160, 82
 Runtime
 efficient, 110
 not polynomial NP, 112
 polynomial, 110

 Scheid 1994, 40
 Schmeh 2003, 15
 Schneier 1996, 15, 41, 96, 128
 Schnorr C.P., 13
 Schroeder 1999, 41
 Schwenk 1996, 41
 Schwenk 2002, 97
 Secude, 8, 13, 143
 Sedgewick 1990, 72, 97
 Seneca, 54
 Session key, 13
 SHA-1, 124, 126
 Shamir 1982, 122
 Shamir 2003, 97
 Shamir Adi, 13, 114, 115
 Short integer, 59
 Signature
 digital, 13, 17, 75, 123, 125
 Signature procedure, 123
 Silver, 118
 Silverman 1986, 145
 Silverman 1992, 145

Silverman 2000, 97
Siverman/Tate 1992, 145
Square and multiply, 60, 91
Stinson 1995, 91, 97, 116, 118, 121, 122
Structure, 53, 62, 64, 68

Tietze 1973, 41
Transitivity, 53
TWIRL device, 7, 79, 97

Weierstrass, Karl, 135–137
Weis 2003, 97
Welschenbach 2001, 97
Wiles Andrew, 46, 97, 131
Wobst 2002, 15
Wolfenstetter 1998, 97

X.509, 127, 128

Yan 2000, 93, 97
Yates Samuel, 22

 \mathbb{Z}_n , 61
 \mathbb{Z}_n^* , 62