1

Introduction

Joppe W. Bos and Martijn Stam

This introductory chapter provides a sketch of Arjen Klaas Lenstra's scientific career, followed by a preview of the technical chapters. We are indebted to Ronald Cramer and Herman te Riele for information about Arjen's longstanding connections to CWI, Monique Amhof for providing additional information about Arjen's time at EPFL and Peter van Emde Boas for his kind permission to use a photograph from his private collection.

1.1 Biographical Sketch

Arjen Klaas Lenstra was born on 2 March 1956 in Groningen, the Netherlands. He received a BA and an MA degree in Mathematics and Physics at the University of Amsterdam in 1975 and 1980, respectively. The picture in Figure 1.1 shows Arjen defending his master's thesis. Afterwards, he obtained a research position at CWI – the Dutch national research institute for mathematics and computer science – in Amsterdam and started conducting his PhD research. Actually, when Arjen started in 1980, the institute was still called Mathematisch Centrum, but in 1983 it would be renamed CWI, an abbreviation for 'Centrum voor Wiskunde en Informatica', using the Dutch words 'wiskunde' for mathematics and 'informatica' for computer science. This renaming reflected the emergence of computer science as a discipline in its own right, separate from mathematics. In retrospect, the new name CWI turned out to be indicative for Arjen's research career as well. Throughout, his work has provided exemplary proof of the impressive powers summoned by harnessing both mathematics and computer science. Since his PhD days, he has always maintained close links with CWI, for instance, as an officially appointed advisor from 2004 until 2017, but also by co-supervision of various PhD students from CWI over the years.

Figure 1.1 Arjen defending his master's thesis on 10 December 1980. The committee, seated behind the desk, was formed by Th. J. Dekker, Peter van Emde Boas and Hendrik W. Lenstra, Jr. Photo from the private collection of van Emde Boas.

Arjen's own PhD research at CWI was conducted under the external supervision of Peter van Emde Boas and, in 1984, he received his PhD degree at the University of Amsterdam for his work on 'polynomial time algorithms for the factorization of polynomials' [312, 362, 363, 364, 380], which was also the title of his PhD thesis [365]. One key ingredient of this research became one of his most widely known results: the Lenstra–Lenstra–Lovász (LLL) lattice basis reduction algorithm, a polynomial time lattice basis reduction algorithm [380]. The LLL algorithm found countless applications, as described in more detail in the LLL book [457] that was published to commemorate its 25th birthday. Accurately predicting the effectiveness and efficiency of lattice reduction algorithms continues to be a key tool in the cryptanalytic toolbox. Indeed, with the ongoing post-quantum cryptographic standardisation effort, LLL and similar algorithms play an essential role in determining the practical parameters for lattice-based cryptography. This is detailed further in Chapter 2.

From 1984 to 1989, Arjen was a visiting professor at the computer science department of the University of Chicago. During this time he retained a position as visiting researcher at CWI and he conducted multiple summer research visits to Digital Equipment Corporation (DEC) in Palo Alto, CA. These latter visits resulted in the famous distributed effort using factoring by electronic mail together with Mark S. Manasse [371]. Such computational cryptanalysis

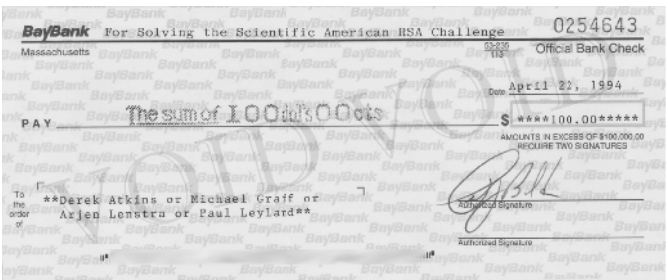

Figure 1.2 Reward for the factorisation of the *Scientific American* RSA challenge.

is essential to understand the practical security strength of the Rivest–Shamir–Adleman (RSA) cryptosystem [501], which is related to the presumed hardness of integer factorisation. It marked the start of Arjen's research in all computational aspects of integer factorisation, leading to his involvement in virtually all integer factorisation records. His records include the factorisation of the ninth Fermat number [382] using the first implementation of the number field sieve [381], solving the famous *Scientific American* RSA challenge [24] by using the quadratic sieve algorithm [478], the kilo-bit [22] special number field sieve factorisation [474], new elliptic-curve integer factorisation [388] records by using game consoles [90] and the factorisation of the RSA-768 challenge [327] (but see also [90, 117, 118, 143, 156, 165, 329, 371, 372, 384]). His work on the *Scientific American* RSA challenge was even featured on the front page of the *New York Times* on 12 October 1988 and, after some delay, he and his collaborators received their due reward for solving this challenge, as evidenced by Figure 1.2. His brother Hendrik would eventually refer to him as 'world champion in factoring'.

Arjen's involvement led to many practical optimisations of these algorithms [167, 224, 372, 392, 428] and most notably to the development of the number field sieve [381]. A historical overview of integer factorisation techniques, including the asymptotically best methods to which Arjen contributed, is provided in Chapter 3.

In order to run integer factorisation algorithms on a large variety of computer architectures, a portable and fast arbitrary-length integer arithmetic software library was needed. Arjen developed the software library FreeLIP (1988–1992), later maintained by Paul Leyland. FreeLIP was used in the early integer factorisation records and formed the early backbone of Shoup's Number Theory Library (NTL). More details about the relationship between FreeLIP, NTL,

and other high-performance, portable software libraries for doing number theory are outlined and explained in Chapter 9.

Recall that Arjen's PhD research on polynomial factoring included the LLL algorithm for lattice reduction. In 1996, Coppersmith [136, 137] showed how the LLL algorithm can be used to factor poorly generated RSA keys in polynomial time. Lattice-based integer factorisation methods are the main topic of Chapter 4. In this chapter, the details behind Coppersmith's method and how to use this for assessing the security of the RSA cryptosystem are explained.

Showing the practical impact of various asymptotic methods and determining, for given circumstances, which is superior is one of the main ingredients when recommending cryptographic key sizes in standards. Arjen pioneered the concrete extrapolation of known factoring and DLP methods to determine meaningful bit-level security estimates for common cryptosystems, including their long-term security. His contributions in this field are widely known and valued by academia, industry and government agencies alike [328, 368, 376, 379] and served as the foundation to determine the exact key sizes for different security levels for virtually all public-key cryptographic standards. More details about the history of selecting appropriate cryptographic keys and current recommendations are described in Chapter 11.

From 1989 to 1996, Arjen held various positions at Bell Communications Research, Morristown, NJ, in the Mathematics and Cryptology Research Group. From 1996 to 2002 he was Vice President of Emerging Technologies in the Corporate Technology Office of Citibank and from 2002 to 2004 Vice President, Information Security Services, still at Citibank. He joined Lucent Technologies' Bell Labs in 2004 where he stayed until the end of 2005.

In addition to his corporate positions at Citibank and later Bell Labs, from 2000 to the end of 2005 Arjen held a position as part-time professor at the Department of Mathematics and Computer Science at the Technische Universiteit Eindhoven in the Netherlands. When the appointment was still being negotiated, the Dean had to keep the Faculteitsraad ('Departmental Council') informed, although without mentioning any concrete names. At the time, the Dean happened to be his brother Jan Karel, who appeared to take some pleasure in referring to an excellent candidate for a part-time professorship as a reknowned Dutch cryptographer who – tongue in cheek – had turned banker.

At the beginning of this period, Arjen invented XTR, together with Eric Verheul [375, 377, 378]. In Eindhoven, Arjen remotely supervised a number of PhD students, including this book's second editor, whose topic was related to XTR [559, 560]. An overview of XTR and subsequent developments is provided in Chapter 10.

At the start of 2006, Arjen was appointed a professor at the École

Figure 1.3 Photo of Arjen in front of the PS3 cluster taken in 2014. Copyright @Tamedia Publications romandes/Sabine Papilloud.

Polytechnique Fédérale de Lausanne (EPFL) in Lausanne, Switzerland. At EPFL, it is customary for professors to name their group a 'laboratory' with associated acronym. Arjen proudly settled on LACAL, as it worked both in French and in English. Of course, for any abbreviation to work in both those languages, it helps if this is a palindrome and LACAL fitted the bill perfectly. In English it stands for LAboratory for Cryptologic ALgorithms, whereas in French it reads LAboratoire de Cryptologie ALgorithmique.

At LACAL, Arjen continued his focus on the design and analysis of algorithms used in cryptographic protocols or security assessments. Given Arjen's interest in computational cryptography, in hindsight it was no surprise that he started to investigate cheap and powerful alternatives to conventional computer clusters, culminating in the purchase of 215 first generation PlayStation 3 (PS3) game consoles in 2007. The purchase of this exotic computer cluster coincided with the start of the PhD of this book's first editor in Arjen's group at EPFL. Arjen and his cluster were featured in a local newspaper in 2014; a photo as used by the newspaper of Arjen posing in front of the cluster is shown in Figure 1.3.

The main motivation for using the PS3 game console for computational cryptography was the Cell processor: a powerful general-purpose processor, which on the first generation PS3s could be accessed using Sony's hypervisor. What made a PS3's Cell processor especially interesting were its eight Synergistic Processing Units (SPUs), although when running Linux only six of these SPUs can be used (as one is disabled and one is reserved by the hypervisor).

Each SPU runs independently from the others at 3.2GHz, using its own 256 kilobyte of fast local memory for instructions and data. It is cache-less and has 128 registers of 128 bits each, allowing Single Instruction, Multiple Data (SIMD) operations on 16 8-bit, 8 16-bit, or 4 32-bit integers. This opens up a myriad of algorithmic possibilities to enhance cryptographic as well as cryptanalytic applications. Indeed, a wide variety of projects were executed on the PS3 cluster.

A particular notorious one was a continuation of Arjen's previous work with Xiaoyun Wang and Benne de Weger related to the possibility of constructing meaningful hash collisions for public keys [369, 385]. This research effort, led by Marc Stevens, resulted in short chosen-prefix collisions for MD5 and, more impactful, rogue CA certificates [570, 573, 574]. The details are described in Chapter 7.

The PS3 cluster was also used by the first author for a PhD project to compute a 112-bit elliptic curve discrete logarithm, an effort roughly equivalent to 14 full 56-bit DES key searches and, at the time, a new ECDLP record [89, 91]. It is worth highlighting that members of Arjen's lab LACAL played an important role in the fall of discrete logarithms in finite fields of fixed characteristic over the past decade. Arjen's involvement in computing discrete logarithms [91, 330, 366] as well as recent progress is detailed in Chapter 5.

Other computational research sparked by the cluster include fast modular multiplication [85, 86, 90]. Chapter 8 provides a detailed description on fast modular reduction.

Working with the cluster was a fun and unique experience: after all, not every PhD student gets the opportunity to work with such challenging computer equipment. The bespoke installation of the cluster also led to some hilarious moments. At some point, after a global power glitch at EPFL, the PhD students (first author included) had to climb the racks in order to manually switch on every PS3 again, all much to Arjen's entertainment.

In an original twist on 'dual-use cryptography', Arjen had part of the cluster reconfigured as PlayLab: a room of 25 PS3s equipped with game controllers and headsets (see Figure 1.4). This room was used to host groups of children between 7 and 18 years old. There were various kinds of demos and presentations, ranging from explaining the Cell processor as an eight-headed dragon to a live demo on MD5 password cracking. The presentations in the PlayLab always ended with the opportunity for the children to play some games. Arjen himself is not in any sense a gamer, a fact that is emphasised in an interview in 2011 [607]:

Figure 1.4 PlayLab: a room of 25 PS3s equipped with game controllers and headsets for visits of groups of children aged 7 to 18.

My first exposure to computer games was in 1974, in the basement of the mathematics department of the university of Amsterdam, where about half a dozen terminals were remotely connected to Amsterdam's only computer (a CDC with 12-bit bytes, and computationally speaking probably less powerful than a current microwave oven). It may have been the dungeon-like environment or the unhealthy look of those who were playing, but it did not attract me and that never changed. But, 25 of our playstations are equipped with monitors and students can come and play one of the many games that we have.

Besides these computational results, Arjen has a keen interest in the security assessment of widely deployed security systems. One year he enthusiastically returned from the RSA conference with a sizeable stash of security dongles that, upon pressing a button, would reveal a supposedly random six-digit security code. He had been using this exact same device for some application already and, after an eery feeling of not-quite-randomness, he had started keeping track of the numbers, which indeed revealed a pattern. Now, the question was whether the problem was device specific or not, so upon spotting a large bowl of such devices at the RSA conference, he quietly set upon collecting a few more samples to confirm his suspicions. After a responsible disclosure to those hosting the stand with the bowl, the bowl promptly disappeared for the rest of the conference. Yet, enough devices had been secured to task a master's student with pressing the button sufficiently often to reverse-engineer and successfully cryptanalyse the device.

More seriously, the interest in deployed security system extended to implementation mistakes or misuse [374, 386]. An age-old adage in cryptanalysis is 'to look for plaintext'. In the context of RSA public keys, specifically the public moduli, this can be translated to look for primes (indeed!) or common factors. It turns out RSA moduli are often not as random as one might wish. Progress in this fascinating area is described in Chapter 6.

Arjen's expertise, broad interest and entertaining way of asking questions, make him a regular member of PhD committees. Memorable are the moments where he asks seemingly naive questions which, in the heat of a defense, can catch a candidate off guard. In one instance, a candidate working on lattice reduction to find the shortest vector was asked why one would not look for the longest vector in lattice instead. In another instance, a candidate working on weaknesses of RSA when either the public exponent e or the private exponent d is small, was asked what would happen if e and d are small simultaneously.

Among EPFL students, Arjen is well known for his clear, fun and interesting lectures. To illustrate examples, he often uses characters from popular television series, for example from the American television sitcom *The Big Bang Theory*. The students' appreciation is highlighted by him twice receiving the 'polysphère de faculté IC' (the teaching award handed out by the students from the computer science department), first in 2008 and then again in 2018. Even more impressively, in 2011 he received the overall, EPLF-wide best teaching award (the 'polysphère d'or').

The Big Bang Theory also featured at LACAL's movie lunches, where the team watched an episode of the series, or 30 to 45 minutes of a movie, every day while eating: one summer we consumed a substantial part of the best Spaghetti Westerns. Perhaps it is a tradition inspired by an earlier one, where Arjen would watch the latest movies with Professor Johannes Buchmann while visiting the flagship conference 'Crypto' in Santa Barbara, USA. One such movie was *Pulp Fiction*, an all-time favorite. Indeed, Arjen's preference for movies from the director Quentin Tarantino or the Coen brothers is no secret. His ideal table partner would be Jules Winnfield (a character in *Pulp Fiction* played by Samuel L. Jackson) 'to make philosophy lessons less unbearable' [177]. His adoration for Tarantino even made it into a scientific publication: see e.g. 'that's a bingo' [327, Section 2.5] from the movie *Inglourious Basterds*.

1.2 Outline

Computational cryptography has two sides: a potentially destructive side and a constructive one. The destructive side challenges the hardness assumptions underlying modern cryptographic systems by the development of suitable number-theoretic algorithms; the constructive side offers techniques and recommendations that help implement and deploy cryptosystems. Although the two sides are intricately intertwined, this book is divided into two parts accordingly. The first part is dedicated to cryptanalysis, whereas the second part focusses on the implementation aspects of computational cryptography.

In Chapter 2, 'Lattice Attacks on NTRU and LWE: A History of Refinements', Martin R. Albrecht and Léo Ducas provide an overview of the advances and techniques used in the field of lattice reduction algorithms. Four decades after its invention, the LLL algorithm still plays a significant role in cryptography, not least because it has become one of the main tools to assess the security of a new wave of lattice-based cryptosystems intended for the new post-quantum cryptographic standard. The runtime of the LLL algorithm was always well understood, but the quality of its output, i.e., how short its output vectors were, could be hard to predict, even heuristically. Yet, an important aspect in the evaluation of the new lattice schemes are accurate predictions of the hardness of the underlying lattice problems, which crucially relies on estimating the 'shortness' of the vectors that can be efficiently found using lattice reduction and enumeration. Albrecht and Ducas have been on the forefront of improving such estimators and build upon their expertise in their chapter.

In Chapter 3, 'History of Integer Factorisation', Samuel S. Wagstaff, Jr, gives a thorough overview of the hardness of one of the cornerstones of modern public-key cryptography. The history starts with the early effort by Eratosthenes and his sieve, eventually leading to the modern number field sieve, currently the asymptotically fastest general-purpose integer factorisation method known. Also included are 'special' integer factorisation methods like the elliptic curve method, where the run-time depends mainly on the size of the unknown prime divisor. Modern factorisation efforts often include a gradual escalation of different methods, so it is essential to be familiar with a wide range of methods and the essence of all relevant algorithms is explained clearly. Wagstaff's chapter is based on his far more extensive book on the same topic [611].

In Chapter 4, 'Lattice-Based Integer Factorisation: An Introduction to Coppersmith's Method', Alexander May investigates the use of LLL to factor integers as pioneered by Coppersmith. Conceptually, Coppersmith's method can be deceptively simple: given additional information about an integer to factor

(e.g., the knowledge that an RSA key pair (N, e) has a small corresponding private exponent d), derive a system of equations with a small root that reveals the factorisation and use LLL to find the small root. As a result, it becomes possible to explore exponentially sized search spaces, while preserving polynomial time using the famous LLL lattice reduction algorithm. Yet, exploiting Coppersmith's method in a cryptographic context optimally often involves a number of clever choices related to which system of equations to consider. At first, a tantalisingly annoying problem where the choice may appear obvious only in retrospect. May uses his extensive experience in improving the state of the art to explain the reasoning behind various applications in his chapter.

In Chapter 5, 'Computing Discrete Logarithms', Robert Granger and Antoine Joux discuss the question 'how hard is it to compute discrete logarithms in various groups?'. The key ideas and constructions behind the most efficient algorithms for solving the discrete logarithm problem are detailed, with a focus on the recent advances related to finite fields of extension degree >1. A highlight is the rapid development, in the period 2012–2014, of quasi-polynomial time algorithms to solve the DLP in finite fields of fixed charaterstic. Both Granger and Joux contributed significantly to this development, albeit on competing teams. For this book, they join forces and explain how different ideas eventually led to the fall of the fixed characteristic finite field discrete logarithm problem.

In Chapter 6, 'RSA, DH and DSA in the Wild', Nadia Heninger outlines the various cryptographic pitfalls one can – but really should not – make in practice. Often it is possible to bypass the 'hard' mathematical problem a cryptosystem is based upon, and instead take advantage of implementation, deployment or protocol mistakes to extract the private key. Often, the techniques used are excellent examples of the interplay of mathematics and computer science, requiring a combination of ingenuity to find the core idea and perseverance to exploit the weakness in practice. Heninger gives a wide-ranging overview of the multitude of cryptographic implementation vulnerabilities that have been found in the past decades and their impact in practice, including a fair number where she was personally involved in identifying the vulnerability. In her chapter, she wonders whether after several decades of implementation chaos and catastrophic vulnerabilities, we are doomed, but concludes that there is hope yet by bringing into practice the lessons learned.

In Chapter 7, 'A Survey of Chosen-Prefix Collision Attacks', Marc Stevens surveys the technical advances, impact and usage of collision attacks for the most widely used cryptographic hash functions. Cryptographic hash functions are the Swiss army knives within cryptography and are used in many applications including digital signature schemes, message authentication codes,