Gröbner Bases, Coding, and Cryptography

Bearbeitet von
Massimiliano Sala, Teo Mora, Ludovic Perret, Shojiro Sakata, Carlo Traverso

1. Auflage 2009. Buch. xvi, 430 S. Hardcover
ISBN 978 3 540 93805 7
Format (B x L): 15,5 x 23,5 cm
Gewicht: 1770 g

Weitere Fachgebiete > Mathematik > Operational Research > Symbolisches Rechnen
Zu Inhaltsverzeichnis

schnell und portofrei erhältlich bei

beck-shop.de
DIE FACHBUCHHANDELUNG

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.
1 In the Beginning

Last century saw a number of landmark scientific contributions, solving long-standing problems and opening the path to entirely new subjects. We are interested in three\(^1\) of these:

1. Claude Shannon’s (1948),
2. Claude Shannon’s (1949),
3. Bruno Buchberger’s (1965)

The title of Shannon’s (1948) paper says it all: “A mathematical theory of communication”. It was later reprinted as Shannon and Weaver (1949) with an even more ambitious title: “The Mathematical Theory of Communication”. Although people have exchanged information in speech and writing for centuries, nobody had ever treated the information exchange (or even information itself) in a rigorous mathematical way. In Shannon’s time there was a need for it, since the last century saw a dramatic increase in the amount and speed of information exchange, with the spreading of new media, like radio, television and telephone.

In Shannon (1948), communication theory is the study of some stationary stochastic processes. Random variables describe information sources and probability distributions describe channels, through which information is sent. Noisy channels are modelled and (error correcting) codes are introduced to permit information recovery after the transmission. In particular, the (probabilistic) foundation of Coding Theory was laid.

One year later, another astonishing paper by Shannon appeared: Shannon (1949). For centuries “secret codes” have been used to protect messages from unauthorized readers. Unsurprisingly, the lack of a rigorous model for communication prevented the study of a more specific model for secure communication. Cryptography had been largely regarded as an art, often mixed with esoteric and obscure references. A cipher was considered secure until an attacker could break it. Like a lighthouse in the dark, Shannon’s paper introduces basic definitions and results, which make

\(^1\)Here listed in chronological order.

---

M. Sala
Dept. of Mathematics, Univ. of Trento, Trento, Italy
e-mail: sala@science.unitn.it

M. Sala et al. (eds.), *Gröbner Bases, Coding, and Cryptography*, DOI 10.1007/978-3-540-93806-4_1, © Springer-Verlag Berlin Heidelberg 2009
cryptography into a science. Shannon views a cipher as a set of indexed functions from the plain-text space to the cipher-text space, where the index space is the key space. Building on his previous paper, he focuses on the probability distribution of (the use of) the keys and of the plain-texts, on the way they determine the cipher distribution and on how an attacker can use them. The paper is also full of invaluable (and prophetic) remarks, such as: “The problem of good cipher design is essentially one of finding difficult problems … How can we ever be sure that a system which is not ideal … will require a large amount of work to break with every method of analysis? … We may construct our cipher in such a way that breaking it is equivalent to … the solution of some problem known to be laborious.”

Among the mathematical problems known to be “laborious” (to use Shannon’s terminology), there is one which has always received a lot of interest: how to “solve” a system of polynomial equations. This reduces to a more general problem: how to represent in a “standard” way a (multivariable) polynomial ideal. Even a simple decision problem like ideal membership\(^2\) had no way to be solved and some even believed it was undecidable, after the word problem in group theory was proved so in Novikov (1955, 1958).

However, in 1965 Buchberger’s (1965, 2006) thesis he presented the appropriate framework for the study of polynomial ideals, with the introduction of Gröbner bases. There is no way to summarize in a few pages the surge in computational algebra research originated from Buchberger’s stunning contribution, with uncountable applications in Mathematics, Engineering, Physics and recently even Biology and other sciences. Fortunately, this book deals only with the applications of Gröbner bases to coding theory and cryptography, and in the next section we will hint at them within the book.

\section{Until Now}

A finite field \(\mathbb{F}\) may not look particularly interesting to mathematicians accustomed to infinite fields. After all, it contains only a finite number of elements. Also, all nonzero elements are exactly the powers of a primitive element, providing a rather dull group structure for its multiplicative elements. Nevertheless, it is a field, which means a lot\(^3\) from the point of view of its polynomial rings and their algebraic varieties. Moreover, it has a very peculiar property: all functions from \(\mathbb{F}^n\) to \(\mathbb{F}\) can be represented as polynomials in \(\mathbb{F}[x_1, \ldots, x_n]\). Here lies the heart of the interaction\(^4\) between Gröbner bases and coding theory/cryptography.

\(^2\)Determining whether a polynomial belongs to an ideal \(I\) given a finite basis for \(I\).

\(^3\)For example, the number of roots of \(p \in \mathbb{F}[x]\) is \(\deg(p)\) (counting multiplicities).

\(^4\)Some recent research has focused on special classes of rings, we will discuss it at the end of Sect. 2.3.
2.1 Classical Coding Theory

After Shannon (1948), coding theory has developed along two main directions: algebraic coding theory and probabilistic coding theory. The rationale behind the (apparently unnatural) introduction of algebra is that it is very difficult to predict (or even to estimate) the performance of codes constructed and decoded in a probabilistic way, while already the pioneeristic work by Hamming (1950) showed how easy it is to construct algebraic codes, with algebraic decoding, whose performance can be easily estimated by the computation of a parameter called the (Hamming) distance.

The main objects of study in algebraic coding theory are “codes”, that is, subsets of finite-dimensional vector spaces over $\mathbb{F}$. There has been extensive study into linear codes (subspaces) and much less into non-linear codes, due to implementation issues. A lot of research has been devoted to cyclic codes, that form a class of linear codes enjoying special algebraic properties, allowing both easier determination of their distance and low-complexity decoders. An introduction to linear and cyclic codes is provided in our chapter (Augot et al. 2009). The two introductory chapters (Mora 2009a, 2009b) lay down our commutative algebra notation, sketch Gröbner basis theory and describe its powerful results for 0-dimensional ideals. The first instance of applications we present is the chapter on the “Cooper philosophy” (Mora and Orsini 2009), where it is showed how to decode efficiently cyclic codes using Gröbner bases. We have a few short notes on linear and non-linear codes, where some Gröbner basis computation is needed:

- Lally (2009) gives a description of quasi-cyclic codes in term of Gröbner bases of polynomial modules,
- Gioretti (2009) introduces nth root codes and show how to compute their distance and weight distribution,
- Bulygin and Pellikaan (2009) explains how to decode a (general) linear code,
- Guerrini et al. (2009) explains how to find the distance of (systematic) non-linear codes (and of linear codes as a special case); a variation allows to classify all such codes with some given parameters,
- Kim (2009) presents a prize problem in coding theory about the existence of a code with special parameters (it could be solved by a variation to the methods in Guerrini et al. 2009),
- Borges-Quintana et al. (2009) provides a Gröbner basis description for binary linear codes, allowing their decoding and the calculation of their distance,
- Martinez-Moro and Ruano (2009) presents a new family of linear codes endowed with a natural Gröbner basis description.

---

5See our note Gluesing-Luerssen et al. (2009) for a hybrid approach.
6I.e., ideals having a finite number of solutions, as it is always the case in coding and cryptography.
7A class of linear codes which can be seen as a generalization of cyclic codes.
8A wide class of linear codes containing cyclic codes.
2.2 AG Codes

In the eighties (Goppa 1981) the so-called AG (short for “Algebraic Geometry”) codes were proposed. These are linear codes obtained as evaluation of function spaces on algebraic curves. Standard results in curve theory yield sharp estimates for their distance. Their geometric structure permits specific decoding algorithms. For problems related to these codes, a polynomial formulation is natural and hence Gröbner bases find a field fertile in applications. Our treatment (chapters) of AG codes is as follows:

- Leonard (2009a) introduces the AG codes, especially the one-point AG codes,\(^9\)
- Little (2009) explains their encoding (with Gröbner bases) and the relation with the curve automorphisms,
- Sakata (2009a) describes the Berlekamp–Massey–Sakata (BMS) algorithm, which can be specialized to decode AG codes,\(^10\) as explained in Sakata (2009b),
- Leonard (2009b) further explores their decoding.

Recently, it has been observed that the classical presentation of AG codes suffers from some limitations, such as the need for a lot\(^11\) of theoretical prerequisites in order to understand theory and the absence of explicit code descriptions.\(^12\) To overcome these difficulties, a new constructive approach has been proposed: the Order Domain codes. These codes and their relation to classical AG codes are discussed in our chapter (Geil 2009). Interestingly, Gröbner bases have turned out to be very convenient tools for their study.

2.3 Coding Miscellanea

Classical decoding algorithms for cyclic and AG codes can be reinterpreted in terms of Gröbner basis computation, as explained in our chapter (Guerrini and Rimoldi 2009), where also list-decoding algorithms are detailed. A list-decoding algorithm is an algorithm\(^13\) that decodes a received message into a list of possible codewords. A probabilistic algorithm is then used to choose the most likely among them. These algorithms are a compromise between algebraic decoding and probabilistic decoding, which is necessary in order to fully exploit the channel capacity without losing the advantage of the algebraic approach. Also the BMS algorithm can be adapted to a list-decoding algorithm (Sakata 2009b).

---

\(^9\)Which is their most important subclass, enjoying an easier description. See our note (Matthews 2009) for multi-point AG codes.

\(^10\)Historically, this was the first fast algorithm to decode such codes.

\(^11\)In comparison to the prerequisites for standard linear code theory.

\(^12\)Which would prevent actual use of these codes.

\(^13\)See also our notes (Augot and Stepanov 2009; Beelen and Brander 2009).
We report that recently also (linear and cyclic) codes over rings have been studied. For an introduction to this theory see our chapter (Greferath 2009). Also Gröbner basis theory can be adapted to special classes of rings. This is sketched in our chapter (Byrne and Mora 2009), where it is also explained how the Gröbner basis decoding techniques in Guerrini and Rimoldi (2009) are extended to codes over (special) rings.

### 2.4 Cryptography

After Shannon’s (1949) paper two main kinds of ciphers have been developed: block ciphers and stream ciphers. Block ciphers are closer to Shannon’s original idea of key-indexed transformations from the plain-text space to the cipher-text space, and can be viewed as maps from $\mathbb{F}^n$ to $\mathbb{F}^m$, for some $n, m \geq 1$. Stream ciphers assume the message to come in a (ideally) infinite stream (of field elements in $\mathbb{F}$) and they add\(^{14}\) element by element the message stream with a key stream produced by the cipher itself. Block ciphers and their relation to Gröbner bases are discussed in our chapter (Cid and Weinmann 2009), while stream ciphers and their relation to Gröbner bases are discussed in chapter (Armknecht and Ars 2009). It is interesting to note that Gröbner basis attacks on some stream ciphers have outmatched all classical attacks and so they are now widely used for assessing the security of keystream generators (Armknecht and Ars 2009). This is not the case for Gröbner basis attacks on block ciphers, yet.

The problem with the ciphers as designed by Shannon is that the two peers need to exchange the key before data transmission. This can be difficult since it requires the presence of a secure channel. In Diffie and Hellman (1976) they solved this problem with an ingenious key exchange protocol and their ideas were adapted to design a cipher based on two keys, a public $K_P$ and a secret $K_S$, such that only a key exchange of $K_P$ in a public channel is required (see e.g. Rivest et al. 1978; McEliece 1978). This branch of cryptography is nowadays called public key (or asymmetric) cryptography (PKC), while traditional cryptography is called symmetric cryptography. Although PKC cannot provide the same security level as symmetric cryptography without a larger computational cost, in many real situations (such as in the Internet) there is little choice. Among the PKC systems brought forward in the last 40 years, there are two families that rely on “laborious” problems in polynomial rings. They are deeply discussed in our chapters (Billet and Ding 2009) and (Levy-dit-Vehel et al. 2009). The ciphers in the latter family are called Polly Cracker systems. Although Gröbner bases are used to attack the systems discussed in both chapters, Gröbner bases are used to build the systems themselves in the Polly Cracker case (which then deserves a deeper analysis).

---

\(^{14}\)Or, rarely, perform more complicate transformations.
As mentioned at the beginning of the section, it is the polynomial nature of all functions from $\mathbb{F}^n$ to $\mathbb{F}$ that allows the use of Gröbner bases in coding and cryptography. A special case is the binary case, i.e. when $\mathbb{F} = \mathbb{F}_2$, since in most applications the encoding/enciphering is binary. Any function from $(\mathbb{F}_2)^n$ to $\mathbb{F}_2$ is called a Boolean function and any function from $(\mathbb{F}_2)^n$ to $(\mathbb{F}_2)^m$ is a vectorial Boolean functions. As expected, their properties are amply studied in connection with cryptography problems.

We present three notes dealing with three different aspects:

- Simonetti (2009) shows how to use Gröbner bases to compute the non-linearity of any Boolean function $f$, which is an important parameter in evaluating the security of using $f$ in building a cipher;
- Gligoroski et al. (2009b) sketches the use of (vectorial) Boolean functions in building hash functions;\(^\text{15}\)
- Gligoroski et al. (2009a) uses Gröbner bases to represent a special class of Boolean functions (quasigroups) which are used to construct a PKC system.

### 3 Final Comments

In the previous sections, I have tried to convey the general plan behind our book and its chapters (notes) division. This book is a collection of papers by many authors, some of them with a very different background.\(^\text{16}\) As such, it cannot be read as a text-book, but the accurate choice of the subjects should allow the reader to have a comprehensive view of the most common applications of Gröbner bases to coding and cryptography. It is especially important to read carefully the introductory chapters and understand their notation. Within every chapter and note, I have done my best to insert all inter-book cross-references that I felt adequate. Still, there are many parts of the theory we have not been able to cover and a lot of further interactions that we have not detailed.

It is my belief (shared by the Board) that this book can be an excellent guide to the subject, both for the researcher wishing to go deeper into some unfamiliar part of the theory and for the student approaching this area.

### References


\(^{15}\)These are cryptographic methods utilized to guarantee the authentication of a pair message/sender.

\(^{16}\)There is also a note (Matsumoto 2009) which does not apparently fit with the rest of the book’s material, but which I felt it should be included because it hints at possible new developments.


J. B. Little, *Automorphisms and encoding of AG and order domain codes*, this volume, 2009, pp. 107–120.


G. L. Matthews, *Viewing multipoint codes as subcodes of one-point codes*, this volume, 2009, pp. 399–402.


T. Mora, *The FGLM problem and Moeller’s algorithm on zero-dimensional ideals*, this volume, 2009a, pp. 27–45.


Gröbner Bases, Coding, and Cryptography
(Eds.) M. Sala; T. Mora; L. Perret; S. Sakata; C. Traverso
2009, XVI, 430 p., Hardcover
ISBN: 978-3-540-93805-7