

Duale Hochschule Baden-Württemberg Mannheim

Studienarbeit

Konstruktion eines kryptographischen Verfahrens basierend auf Reed-Solomon-Codes und Diskussion möglicher Angriffsmethoden

Studiengang Informatik

Studienrichtung Cyber Security

Verfasser(in):	Roman Wetenkamp
Matrikelnummer:	5533869
Kurs:	TINF20CS1
Studiengangsleiter:	Prof. Dr. Konstantin Bayreuther
Wissenschaftliche(r) Betreuer(in):	Prof. Dr. Reinhold Hübl
Bearbeitungszeitraum:	18.10.2022 – 18.04.2023
Version:	4. Februar 2023

Ehrenwörtliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Titel "*Konstruktion eines kryptographischen Verfahrens basierend auf Reed-Solomon-Codes und Diskussion möglicher Angriffsmethoden*" selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich versichere zudem, dass die eingereichte elektronische Fassung mit der gedruckten Fassung übereinstimmt.

Ort, Datum

Roman Wetenkamp

Inhaltsverzeichnis

Abbildungsverzeichnis	iii
Quelltextverzeichnis	iv
Abkürzungsverzeichnis	v
1 Einleitung	1
1.1 Thematische Übersicht	2
2 Codierungstheorie	3
2.1 Übermittlung von Informationen	3
2.2 Problemstellung und Zielsetzung	5
2.3 Kanalcodierung	5
2.3.1 Grundbegriffe	6
2.3.2 Noisy Channels Coding Theorem	9
2.4 Finitfeldarithmetik	10
2.4.1 Gruppen	10
2.4.2 Körper	12
2.4.3 Polynome	15
3 Lineare fehlerkorrigierende Codes für kryptographische Zwecke	17
3.1 Zyklische Codes	19
3.2 Definition der Klasse von Reed-Solomon-Codes	21
3.2.1 Verallgemeinerte Reed-Solomon-Codes	22
3.2.2 Duale Codes	23
3.2.3 Kontroll- und Generatormatrizen	24
3.3 Kodierung, Fehlerkorrektur und Dekodierung	26
3.3.1 Kodierung	26
3.3.2 Dekodierung	28
Anhang	
A Anhang	32
A.1 Implementierung der GRS-Kodierung	32
A.2 Transformieren der Generatormatrix in die systematische Form	34
Literaturverzeichnis	38

Abbildungsverzeichnis

1.1	Fortlaufend ergänzte Übersicht über relevante und abgegrenzte Teilgebiete der Informationstheorie	2
2.1	Gegenstand der Codierungstheorie (nach [7, S. 1])	5
2.2	Vereinfachte Darstellung eines Information Transmission System (ITS) (nach [4, S. 3])	6
2.3	Visualisierung der Kugelinterpretation (nach [9])	9

Quelltextverzeichnis

A.1 SageMath-Implementierung des Kodieralgorithmus für GRS-Codes	32
--	----

Abkürzungsverzeichnis

BSC	Binary Symmetric Channel
EDV	Elektronische Datenverarbeitung
ITS	Information Transmission System
QR	Quick Response
RSC	Reed-Solomon-Codes

1 Einleitung

„The lesson here is that it is insufficient to protect ourselves with laws;
we need to protect ourselves with mathematics.“
– BRUCE SCHNEIER in [1]

In einer Welt, in der so viele Daten wie nie zuvor übertragen werden, digitale Kriegsführung und *Nation-state-attacks* nicht mehr bloß Gegenstand dystopischer Science-Fiction-Literatur, sondern Alltag sind, steigt die Relevanz kryptographischer Verfahren, die es ermöglichen, die Vertraulichkeit und Integrität schützenswerter Daten selbst unter der Annahme, dass Angreifenden nahezu unbegrenzte Ressourcen zur Verfügung stehen, sicherzustellen.

Das Forschungsgebiet der *Post-Quanten-Kryptographie* [vgl. 2] hat die Entwicklung kryptographischer Systeme zum Gegenstand, die selbst mit den durch Quantentechnologie anzunehmenden Rechenleistungssteigerungen nicht gebrochen werden können. Ein aussichtsreicher Kandidat dafür ist das *McEliece*-Kryptosystem, das auf linearen, fehlerkorrigierenden Codes basiert. Jene Verbindung der Codierungstheorie und Kryptographie ist Gegenstand dieser Arbeit: Basierend auf dem McEliece-Kryptosystem soll ein aufbauender Ansatz von HARALD NIEDERREITER betrachtet werden, der im Vergleich zum McEliece-Kryptosystem nicht auf *Goppa*-, sondern auf *Reed-Solomon*-Codes basiert und dadurch zwar bessere Rechenzeiten erreicht, jedoch vermutlich auch an Sicherheit einbüßt.

Diese Arbeit stellt zunächst die theoretischen Hintergründe der Codierungstheorie für kryptographische Zwecke dar, bevor basierend darauf die Arbeiten von MC ELIECE und NIEDERREITER analysiert und für die Entwicklung eines eigenen Kryptosystems genutzt werden. Auf jenes Verfahren werden abschließend Methoden der Kryptoanalyse unter Einbeziehung der Arbeit von SIDELNIKOV und SHESTAKOV angewandt, um Aussagen über die Sicherheit des Verfahrens treffen zu können.

1.1 Thematische Übersicht

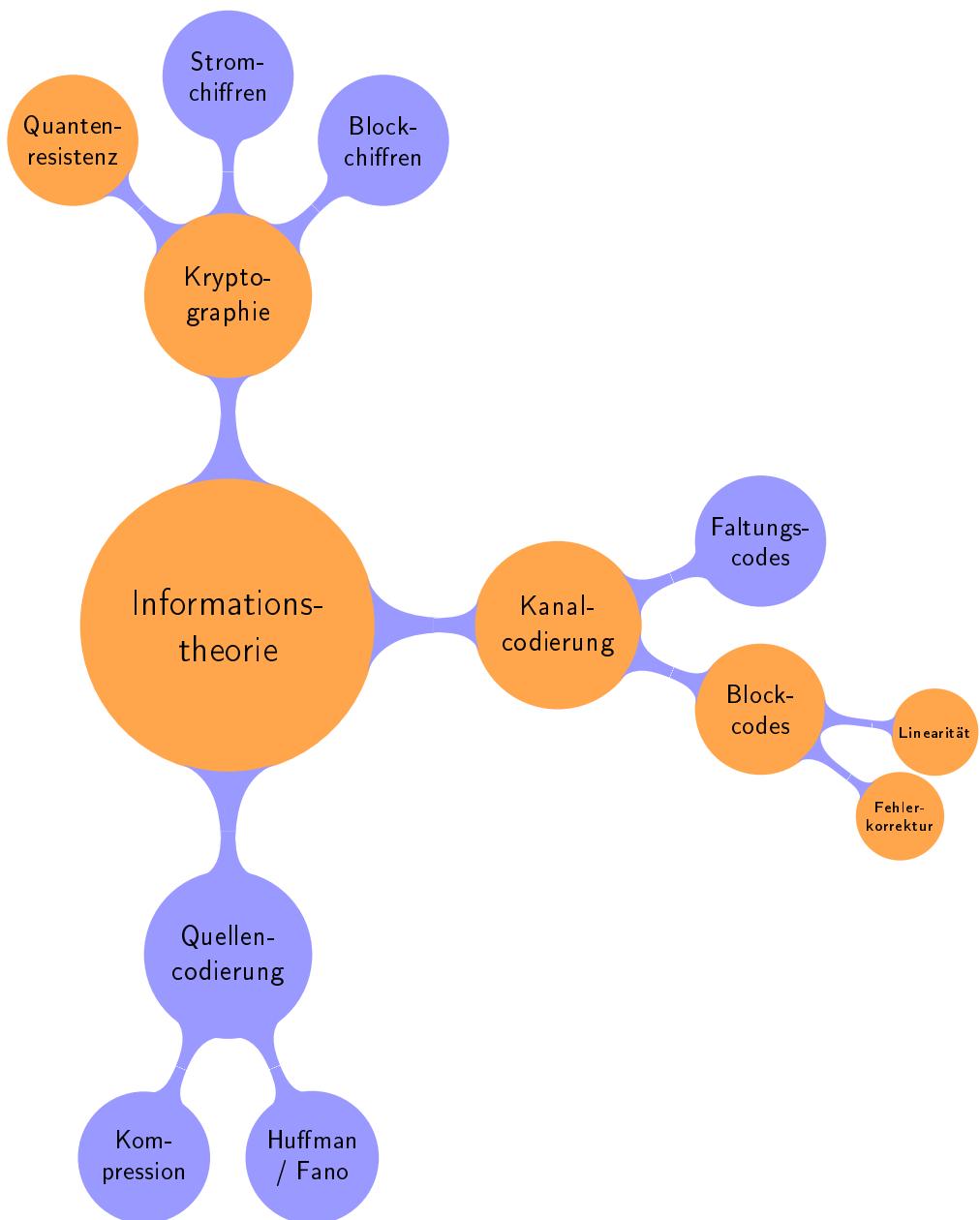


Abbildung 1.1: Fortlaufend ergänzte Übersicht über relevante und abgegrenzte Teilgebiete der Informationstheorie

2 Codierungstheorie

Da im Rahmen dieser Arbeit ein kryptographisches Verfahren entwickelt wird, das auf Elementen der Codierungstheorie basiert, wird diese nun zunächst in Definitionen und Hintergründen dargestellt.

2.1 Übermittlung von Informationen

Sowohl in der Kryptographie, als auch in der Codierungstheorie fungieren **Nachrichten**, die über mit bestimmten Eigenschaften behafte **Kanäle** übertragen werden, als die Objekte der Anschauung.

Definition 1

Eine **Nachricht** m sei definiert als eine endliche Folge von Zeichen $a_i \in \Sigma$, wobei Σ eine endliche Menge von Zeichen (genannt **Alphabet**) bezeichnet.

$$m = \langle a_1, a_2, \dots, a_{n-1}, a_n \rangle \quad \forall i = 1, \dots, n : a_i \in \Sigma$$

Ein typisches Alphabet sind die Zeichen der *ASCII*-Kodierung, mit denen nahezu alle Worte und Sätze der natürlichen englischen Sprache gebildet werden können [vgl. 3]. Dieses Alphabet besteht nun nicht aus Zeichen der natürlichen Sprache, sondern aus 7 Bit langen Zahlenwerten, was die Anwendung von Codes oder kryptographischen Verfahren ermöglicht. Im Rahmen dieser Arbeit wird angenommen, dass Zeichen stets in einem Zahlenformat repräsentiert werden.

Die Definition einer informationstheoretischen Nachricht impliziert eine Autorenschaft, folglich muss jeder Nachricht eine Partei (ein natürliche Person, ein System oder ein Dienst) zugeordnet werden können, die im Folgenden als **Sender**¹ der Nachricht bezeichnet wird. Wird diese Nachricht nun über einen Kanal an eine andere Partei übertragen, so nennen wir

¹Da sich die Anwendung der modernen Kryptographie sehr überwiegend mit dem Austausch von verschlüsselten Nachrichten zwischen Systemen und nicht unmittelbar zwischen natürlichen Personen befasst, wird hier die männliche Form verwendet (Sender = Dienst/System).

diese den **Empfänger**. Entgegen der in der Kryptographie üblichen *Alice-Bob-Notation* wird in dieser Arbeit diese Terminologie beibehalten, um an den codierungstheoretischen Hintergrund anzuknüpfen.

Ein **Kanal** bezeichne ein Medium zur Datenübertragung wie beispielsweise einen elektrischen Leiter, einen Lichtwellenleiter oder die Luft für eine drahtlose Verbindung. Es gibt Kanäle, die Informationen **digital** übertragen, also als diskrete Binärwerte, und im Gegensatz dazu Kanäle, die fortlaufend und stetig Signale übertragen [vgl. 4, S. 1]. **Rauschen** bezeichne nicht-deterministische Daten, die Nachrichten bei einer Übertragung über einen Kanal unbeabsichtigt hinzugefügt werden und so die Nachricht verändern, sie folglich mit **Fehlern** versetzen [vgl. 5, S. 1].

Definition 2

Ein **Kanal** K mit der Menge der Eingabewerte I und der Menge der Ausgabewerte J sei definiert als Matrix

$$K_{ij} = \Pr(j \mid i) \quad i \in I, j \in J$$

wobei $\Pr(j \mid i)$ definiert sei als die Wahrscheinlichkeit für die Ausgabe j unter der Bedingung i .

Sofern mindestens ein Koeffizient der Matrix $i \neq j$ einen Wert > 0 aufweist, ist der Kanal **rauschend** [vgl. 6, S. 73f.].

Beispiel 1

Ein **Binary Symmetric Channel (BSC)** Γ mit einer Bitfehlerwahrscheinlichkeit e ist definiert durch

$$\Gamma = \begin{pmatrix} \Gamma_{00} & \Gamma_{01} \\ \Gamma_{10} & \Gamma_{11} \end{pmatrix} = \begin{pmatrix} 1-e & e \\ e & 1-e \end{pmatrix}$$

Dass ein Kanal eine Nachricht ohne Rauschen überträgt, ist zwar ein erstrebenswerter Zustand, praktisch jedoch aufgrund der Physik nicht zu erreichen. Jede physische Datenübertragung verläuft nicht fehlerfrei, weshalb die Beziehung $m = m'$ daher nur in einem theoretischen Idealfall gilt. Diese Feststellung liefert die Begründung für die Beschäftigung mit der Codierungstheorie.

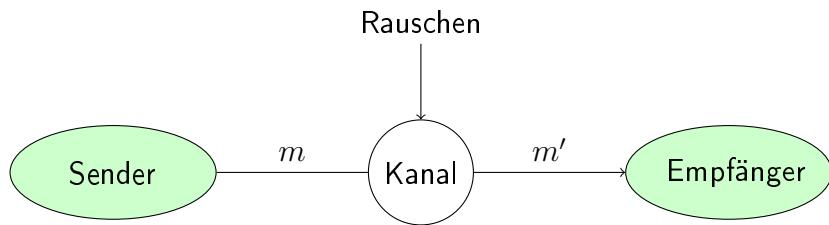


Abbildung 2.1: Gegenstand der Codierungstheorie (nach [7, S. 1])

2.2 Problemstellung und Zielsetzung

Da die Datenübertragung über eine Vielzahl von Kanälen eben nicht fehlerfrei verläuft, liegt es nahe, die Daten so zu übertragen, dass fehlende Bits aus dem Rest der Nachricht erschlossen werden können, wie es zum Beispiel bei natürlicher Sprache der Fall ist. Unsere Worte enthalten häufig Buchstaben, die nicht zwingend erforderlich sind, um das gemeinte Wort zu erkennen [vgl. 5, S. 3].

Überträgt man diese Erkenntnis auf Nachrichten einer beliebigen Sprache, so lassen sich auch im allgemeinen Fall durch das Hinzufügen von redundanten Informationen Nachrichten erzeugen, deren Informationsgehalt sich auch nach der Übermittlung nicht verringert hat. Ein spezieller Typ dieser Verfahren wird als **fehlerkorrigierende Codes** bezeichnet [vgl. 5, S. 3]. Diese Codes sind dem Gebiet der **Kanalcodierung** zuzurechnen, die das Ziel hat, die Qualität der Übertragung auf verlustbehafteten Kanälen sicherzustellen. Sie grenzt sich ab von der **Quellencodierung**, die Verfahren bündelt, welche die Transformation der zu versendenden Daten zum Ziel hat, beispielsweise zur Kompression und Redundanzverringerung [vgl. 8, S. 1]. Der Fokus dieser Arbeit liegt dabei auf der Kanalcodierung, da die betrachteten kryptographischen Verfahren auf ihr basieren.

2.3 Kanalcodierung

Die Relevanz der Frage, wie möglichst viele Übertragungsfehler in einer Datenübertragung vermieden oder korrigiert werden können, stieg mit der Verbreitung von EDV-Systemen und nicht zuletzt dem Internet rasant an [vgl. 4, S. 209]. Populäre Beiträge der Grundlagenforschung wie jener von HAMMING sind daher auch trotz ihres einige Dekaden umspannenden Alters Fundament der folgenden Definitionen.

2.3.1 Grundbegriffe

Um Codier- und Decodervorgänge beschreiben zu können, ist eine Erweiterung und Präzisierung der Abbildung 2.1 erforderlich. Die Zuordnung eines Codewortes c zu einer Nachricht m wird als **Codierung** (der Nachricht) bezeichnet.

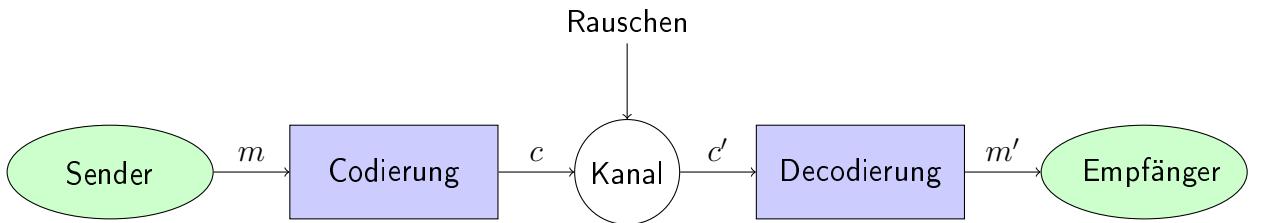


Abbildung 2.2: Vereinfachte Darstellung eines ITS (nach [4, S. 3])

Definition 3

Sei A ein Alphabet und $n \in \mathbb{N}$.

Dann sei A^n die Menge aller n -Tupel der Form $A^n = \{\langle a_1, a_2, \dots, a_n \rangle \mid a_i \in A\}$.

Ein **Blockcode** C der Länge n über dem Alphabet A ist definiert als $C \subseteq A^n$, wobei $|C| > 0$ gelten muss.

Ist $m = |C|$ und B mit $|B| < m$ die Menge zu codierender Informationseinheiten über einem Alphabet A' , so ist jede injektive Abbildung $f: B \rightarrow C$ eine **Codierfunktion** [vgl. 8, S. 10].

Neben Blockcodes können auch sogenannte **Faltungscodes** zur Fehlerkorrektur verwendet werden. Hierbei werden die Informationen nicht unabhängig voneinander blockweise codiert, sondern über Schieberegister in Abhängigkeit zueinander, wodurch sich eine bessere Performanz im Gegensatz zu Blockcodes ergibt [vgl. 10, S. 752]. Der Ansatz, Daten blockweise oder als Datenstrom über Schieberegister zu codieren, weist Parallelen zur kryptographischen Unterscheidung zwischen Block- und Stromchiffren auf. Der Fokus dieser Arbeit wird aufgrund der kryptographischen Bedeutung auf Blockcodes liegen.

Als Maß der Qualität einer Übertragung beziehungsweise der Verfälschung einer Nachricht durch Rauschen, aber auch für die Unterscheidbarkeit von Codeworten, bietet sich die **Hamming-Distanz** an.

Definition 4

Seien $a = \langle a_1, a_2, \dots, a_n \rangle$, $b = \langle b_1, b_2, \dots, b_n \rangle \in A^n$ und A ein Alphabet, so ist die **Hamming-Distanz** d eine Metrik auf A^n . Sie ist definiert als die Anzahl der abweichenden Stellen in zwei Codeworten a und b :

$$d(a, b) = |\{i \mid 1 \leq i \leq n, a_i \neq b_i\}|$$

Des Weiteren sei die **Minimaldistanz** eines Blockcodes C mit $|C| > 1$ definiert als

$$d(C) = \min\{d(x, y) \mid x, y \in C, x \neq y\}$$

[vgl. 8, S. 11] [vgl. 11, S. 105] [vgl. 9, S. 155].

Beispiel 2

Seien $A = \{0, 1\}$ und $n = 5$. Dann ist $A^n = \{\langle 0, 0, 0, 0, 0 \rangle, \langle 0, 0, 0, 0, 1 \rangle, \dots, \langle 1, 1, 1, 1, 1 \rangle\}$. Dann gilt für einen Blockcode C :

- $C = A^n \rightarrow d(C) = 1$
- $C = \{\langle 0, 0, 0, 0, 1 \rangle, \langle 0, 0, 0, 1, 0 \rangle, \dots, \langle 1, 0, 0, 0, 0 \rangle\} \rightarrow d(C) = 2$
- $C = \{\langle 0, 1, 0, 1, 0 \rangle, \langle 1, 0, 1, 0, 1 \rangle\} \rightarrow d(C) = 5$

Wenn die Werte einer Codierungsfunktion im Bildbereich weit verstreut sind, der Code also eine große Minimaldistanz aufweist, werden die einzelnen Codeworte unterscheidbar. Folglich ergibt sich aus der Minimaldistanz ein relevantes Kriterium für die Eigenschaft eines Codes, **fehlererkennend** oder sogar **fehlerkorrigierend** zu sein. Sie beeinflusst, wie viele Fehler erkannt beziehungsweise korrigiert werden können [vgl. 9, S. 155].

HAMMING nutzt für seine Argumentation in [9] eine geometrische Betrachtung mithilfe von Kugeln:

Definition 5

Eine **abgeschlossene Kugel** $B(a, r)$ um einen Punkt a mit Radius r in einem (metrischen) Raum X sei definiert durch

$$B(a, r) = \{x \in X \mid d(a, x) \leq r\}$$

Als Abstandsmaß d kann hier beispielsweise die eingeführte *Hamming*-Distanz verwendet werden.

Definition 6

Sei B eine Kugel mit Radius $r \geq 2$ und Mittelpunkt $x \in A^n$. Dann liegen alle $a_i \in A^n$ mit $d(a_i, x) = r$ auf der Kugeloberfläche von B . Sei $C \subseteq A^n$ nun ein Blockcode mit $d(C) = r$ und $x \in C$. Dann liegen alle $c \in (C \setminus \{x\})$ auf der Kugeloberfläche oder außerhalb der Kugel B .

Bemerkung 1

Ein Punkt $a_j \neq x$ mit $d(a_j, x) < r$ (unterscheidet sich in in weniger als r Stellen von x) liegt nicht auf der Kugeloberfläche und kann folglich kein fehlerfreies Codewort sein. Ein Code C mit Minimaldistanz 2 ist damit **fehlererkennend** für maximal ein falsch übertragenes Zeichen (notiert: „1-fehlererkennend“).

Ferner ist ein Code C' mit Minimaldistanz 3 **fehlerkorrigierend**: Sei a_r ein gültiges Codewort mit einer Minimaldistanz von 3 zu allen anderen Codeworten in C' . Für ein in einer Stelle abweichendes Wort a_f gilt folglich $d(a_r, a_f) = 1$, aber $\forall a_i \in C' \setminus \{a_r, a_f\} : d(a_i, a_f) > 1$. Damit lässt sich a_f eindeutig a_r zuordnen, wodurch sich der Fehler korrigieren lässt und C' **1-fehlerkorrigierend** ist [vgl. 9, S. 155f.].

Die Eigenschaft eines Codes, fehlererkennend oder fehlerkorrigierend zu sein, ist skalierbar:

Definition 7

Seien $c, c' \in C; c \neq c'$ Codeworte eines Blockcodes $C \in A^n$, $\epsilon \in \mathbb{N}$, $B_e(c) = \{x \in A^n \mid d(x, c) \leq \epsilon\}$ und $B_e(c')$ analog $B_e(c') = \{x \in A^n \mid d(x, c') \leq \epsilon\}$.

Dann ist C **ϵ -fehlerkorrigierend**, wenn $\forall c, c' \in C : B_e(c) \cap B_e(c') = \emptyset$ gilt [vgl. 8, S. 12f.].

Bemerkung 2

Ein Blockcode C ist ϵ -fehlerkorrigierend, wenn $d(C) \geq 2 \cdot \epsilon + 1$ gilt [vgl. 8, S. 12f.].

Definition 8

Sei $c \in C$ Codewort eines Blockcodes $C \in A^n$ und $B_t(c) = \{x \in A^n \mid d(x, c) \leq t\}$ die zu

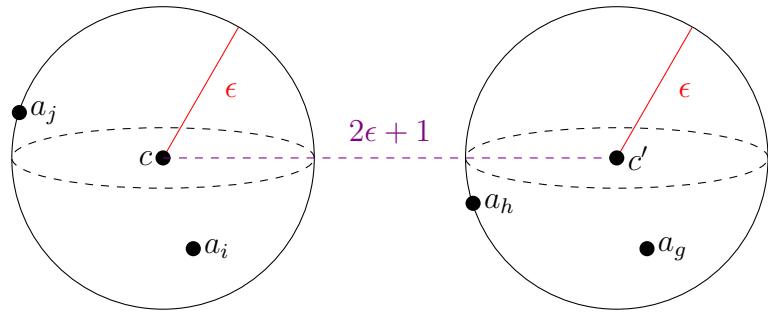


Abbildung 2.3: Visualisierung der Kugelinterpretation (nach [9])

c gehörige Kugel mit allen Elementen aus A^n , die höchstens t entfernt sind.

Dann ist C **t -fehlererkennend**, wenn $\forall c \in C : B_t(c) \cap (C \setminus \{c\}) = \emptyset$, die Kugel um c also keine anderen Codeworte enthält [vgl. 8, S. 13].

Bemerkung 3

Ein Blockcode C ist t -fehlererkennend, wenn $d(C) \geq t + 1$ gilt [vgl. 8, S. 13].

Beispiel 3

Sei $A = \{0, 1\}$ und $A^3 = \{\langle 0, 0, 0 \rangle, \langle 0, 0, 1 \rangle, \dots, \langle 1, 1, 1 \rangle\}$. Ferner sei $C \in A^3$ ein Blockcode für Nachrichten der Form $m = \langle m_1, m_2 \rangle$ mit der Codierfunktion $f: M \rightarrow C, \langle m_1, m_2 \rangle \mapsto \langle m_1, m_2, ((m_1 + m_2) \bmod 2) \rangle$.

Für die Nachricht $x = \langle 0, 1 \rangle$ ergibt sich also $f(x) = \langle 0, 1, 1 \rangle$.

Dann ist $C = \{\langle 0, 0, 0 \rangle, \langle 0, 1, 1 \rangle, \langle 1, 0, 1 \rangle, \langle 1, 1, 0 \rangle\}$. Daraus folgt $d(C) = 2$. Damit ist C 1-fehlererkennend, da $2 = 1 + 1$, aber nicht fehlerkorrigierend, da $2 \not\geq 2 \cdot 1 + 1$ gilt.

Welches Potenzial von fehlerkorrigierenden Codes für das zugrundeliegende Problem ausgeht, verdeutlicht das folgende Theorem:

2.3.2 Noisy Channels Coding Theorem

SHANNON konnte 1948 zeigen, dass es für einen binären, symmetrischen und rauschenden Kanal (BSC) Codes gibt, bei der die auftretende Fehlerwahrscheinlichkeit beliebig klein wird, solange sich die Datenrate der Übertragung unterhalb der Kapazität eines Kanals befindet.

Theorem 1

Sei C die **Kapazität** eines binären, rauschenden Kanals, n die Länge uniformer Codeworte, $P(E)$ die Fehlerwahrscheinlichkeit und D die Datenübertragungsrate, wobei $D < C$. Dann gilt

$$P(E) \leq 2^{-n \cdot e(D)}$$

wobei $e(D)$ eine vollständig von Kanalparametern abhängige positive Funktion, genannt **Fehlerexponent**, ist.

Folglich ist es möglich, Datenübertragungsvorgänge unter der Wahl einer geeigneten Datenübertragungsrate und einer Codierung nahezu fehlerfrei umzusetzen, völlig unabhängig von der Länge der zu übertragenden Nachricht. Es folgt auch, dass sich die Fehleranzahl nicht linear zur Fehleranfälligkeit (angegeben durch die Kapazität) des Kanals verhält [vgl. 4, S. 209ff.].

2.4 Finitfeldarithmetik

Viele der in der Praxis häufig eingesetzten Codierfunktionen basieren auf der Arithmetik in endlichen Körpern und *Galois*-Feldern, die im Folgenden eingeführt werden.

Ausgangspunkt der Definition von endlichen Körpern ist nach LIDL UND NIEDERREITER die algebraische Struktur der **Gruppe**.

2.4.1 Gruppen

Definition 9

Sei G eine Menge und \circ eine binäre Relation $\circ : G \times G \rightarrow G$ mit folgenden Eigenschaften:

1. **Assoziativität:**

$$\forall a, b, c \in G : a \circ (b \circ c) \Leftrightarrow (a \circ b) \circ c$$

2. **Existenz eines neutralen Elements:**

$$\exists e \in G, \forall a \in G : a \circ e \Leftrightarrow e \circ a \Leftrightarrow a$$

3. **Existenz inverser Elemente:**

$$\forall a \in G, \exists a^{-1} \in G : a \circ a^{-1} \Leftrightarrow a^{-1} \circ a \Leftrightarrow e$$

Dann ist (G, \circ) eine **Gruppe**.

Gilt für eine Gruppe zudem $\forall a, b \in G : a \circ b \Leftrightarrow b \circ a$, so ist sie **kommutativ**, auch bezeichnet als **abelsch** [vgl. 12, S. 2].

Beispiel 4

Gegeben seien die Menge der ganzen Zahlen \mathbb{Z} und eine binäre Relation.

- $(\mathbb{Z}, +)$, wobei $+$ die Addition bezeichne, ist eine **abelsche Gruppe**, da die folgenden Eigenschaften erfüllt sind:

1. Assoziativität:

Seien $a, b, c \in \mathbb{Z}$. Dann gilt mit dem Distributivgesetz der Addition $a + (b + c) \Leftrightarrow a + b + c \Leftrightarrow (a + b) + c$

2. Existenz eines neutralen Elements:

$$\forall x \in \mathbb{Z} : x + 0 \Leftrightarrow 0 + x \Leftrightarrow x$$

3. Existenz inverser Elemente:

$$\forall x \in \mathbb{Z}, \exists -x \in \mathbb{Z} : x + (-x) = 0$$

4. Kommutativität:

$$\forall a, b \in \mathbb{Z} : a + b \Leftrightarrow b + a$$

- (\mathbb{Z}, \cdot) , wobei \cdot die Multiplikation bezeichne, ist keine Gruppe, da die inversen Elemente zu Elementen von \mathbb{Z} nicht Teil von \mathbb{Z} sind.

$$\forall x \in \mathbb{Z} \setminus \{1, -1\} : x \cdot \frac{1}{x} = 1, \text{ aber } \frac{1}{x} \notin \mathbb{Z}$$

Definition 10

Eine multiplikative Gruppe ist **zyklisch**, sofern

$$\exists a \in G, \forall b \in G, \exists j \in \mathbb{Z} : b = a^j$$

Dann ist a ein **Generator** der multiplikativen Gruppe² G [vgl. 12, S. 3f.]. Die Gruppe ist durch $G = \langle a \rangle$ hinreichend beschrieben.

²Wird zu einer Gruppe keine binäre Relation angegeben, handelt es sich implizit um eine multiplikative Verknüpfung.

Kryptographische Bedeutung Zu Gruppen existieren einige mathematische Probleme, die sich für kryptographische Verfahren anbieten. MYASNIKOV, USHAKOV UND SHPILRAIN stellen in [13] kryptographische Protokolle dar, die beispielsweise auf dem *Konjugationssuchproblem* nicht-kommutativer Gruppen basieren:

Für zwei Elemente $u, v \in G$ einer Gruppe G soll mindestens ein $x \in G$ gefunden werden, sodass $u^x = v$ gilt. Ist kein anderer Algorithmus für dieses Problem für die Gruppe G bekannt, so lässt sich $x \rightarrow u^x$ als Einwegfunktion auffassen und so ein kryptographisches Protokoll konstruieren [vgl. 13, S. 37f.].

Abhängig davon, ob eine Gruppe abelsch ist oder nicht, lassen sich unterschiedliche Such- oder Entscheidungsprobleme zur Konstruktion kryptographischer Verfahren verwenden, die in [13] dargestellt, aber in dieser Arbeit nicht weiter behandelt werden.

2.4.2 Körper

Definition 11

Ein **Ring** $(R, +, \circ)$ ist eine nicht-leere Menge R mit den zwei auf ihr definierten binären Relationen, notiert durch $+$ und \circ , für die gilt:

- R ist abelsche Gruppe bezüglich der Operation $+$
- Die Operation \circ ist assoziativ, folglich gilt

$$\forall a, b, c \in R : (a \circ b) \circ c \Leftrightarrow a \circ (b \circ c)$$

- Das Distributivgesetz gilt, also ist

$$\forall a, b, c \in R : (a \circ (b + c)) \Leftrightarrow (a \circ b + a \circ c) \wedge ((b + c) \circ a) \Leftrightarrow (b \circ a + c \circ a)$$

Ein Ring ist **kommutativ**, wenn die Operation \circ kommutativ ist. Ferner ist ein Ring ein **Divisionsring**, falls $(R \setminus \{0\}, \circ)$ eine Gruppe bildet [vgl. 12, S. 13][vgl. 14, S. 1f.].

Ein Ring beinhaltet folglich eine Gruppe, jedoch auch eine Operation, die nicht notwendigerweise die Anforderungen an eine Gruppe erfüllt. Da beispielsweise der Menge \mathbb{Z} inverse Elemente bezüglich der Multiplikation fehlen, so ist $(\mathbb{Z}, +, \circ)$ dennoch ein Ring, wenn auch kein Divisionsring.

Definition 12

Ein kommutativer Divisionsring wird als **Feld** oder **Körper** bezeichnet [vgl. 12, S. 13].

Um nun endliche Körper und *Galois-Körper* definieren zu können, ist das Konzept der **Restklasse** unabdingbar, das nun zunächst wieder auf Ringen definiert wird.

Definition 13

Sei R ein Ring. Dann ist J ein **Unterring** von R , falls J abgeschlossen ist unter beiden Operationen $+$ und \circ [vgl. 12, S. 13].

Bemerkung 4

Abgeschlossenheit bezeichnet hier, dass das Ergebnis der Operation auf zwei beliebigen Elementen des Rings ein Element ergibt, das wiederum Element des Rings ist.

Definition 14

Ferner ist J ein **Ideal** von R , falls J ein Unterring von R ist und

$$\forall a \in J, r \in R : ((ar \in J) \wedge (ra \in J))$$

gilt [vgl. 12, S. 13].

Gilt $I = (a)$ für ein $a \in I$, so ist I ein von a erzeugtes **Hauptideal**.

Bemerkung 5

Ein Ideal ist nun nicht mehr nur bezüglich eigenen Elementen und beiden Operationen abgeschlossen, sondern in Bezug auf die Multiplikation auch bezüglich Elementen des übergeordneten Rings.

Definition 15

Sei R eine Restklasse, $a \in R$ und J ein Ideal von R . Die **Restklasse** von $a \bmod J$ sei definiert als Menge aller Elemente der Form $a + c \in R, \forall c \in J$ [vgl. 12, S. 13][vgl. 14, S. 10]:

$$[a] = a + J = \{a + c \in R : c \in J\}$$

Bemerkung 6

Zwei Elemente $a, b \in R$ sind *kongruent* modulo J , wenn sie Elemente der gleichen Restklasse bezüglich J sind, geschrieben $a \equiv b \bmod J$.

Auch auf Restklassen lassen sich die Operationen eines Rings anwenden:

$$(a + J) + (b + J) = (a + b) + J$$

$$(a + J) \circ (b + J) = ab + J$$

Bemerkung 7

Durch diese beiden Operationen auf Restklassen entsteht ein Ring, der als **Restklassenring** bezeichnet und mit R/J notiert wird [vgl. 12, S. 13f.] [vgl. 14, S. 10].

Beispiel 5

Ein klassisches Beispiel für einen auf diese Weise entstandenen Restklassenring ist $\mathbb{Z}/(p)$, $p \in \mathbb{P}$, wobei \mathbb{P} die Menge aller Primzahlen bezeichnet. Hier ist (p) ein durch p generiertes Hauptideal. Dieser Ring ist ein Körper, wie in [12] bewiesen wird und aufgrund der endlichen Anzahl an Elementen **finit**. Die Elemente sind beschrieben durch

$$\{[0] = 0 + (n), [1] = 1 + (n), \dots\}$$

Der Restklassenring $\mathbb{Z}/(3)$ enthält die Elemente $[0]$, $[1]$ und $[2]$.

Nun ist die Definition der *Galois-Körper* unmittelbar möglich:

Definition 16

Sei $p \in \mathbb{P}$. Dann sei \mathbb{F}_p der Körper, der durch den Restklassenring $\mathbb{Z}/(p) = \mathbb{Z}/p\mathbb{Z}$ beschrieben wird und folglich die Struktur $\mathbb{F}_p = \{[0], [1], \dots, [p - 1]\}$ aufweist. \mathbb{F}_p wird dann als **Galois-Körper** der Ordnung p bezeichnet [vgl. 12, S. 15][vgl. 15, S. 75].

Beispiel 6

Beispiele für endliche Körper sind:

- $\mathbb{F}_2 = \{[0], [1]\}$
- $\mathbb{F}_3 = \{[0], [1], [2]\}$, beispielsweise $\{0, 1, -1\}$

Neben *Galois-Körpern*, die auf diese Weise erzeugt wurden, bilden auch alle Primzahlpotenzen p^f , $p \in \mathbb{P}$ und $f \in \mathbb{N}$ jeweils einen endlichen Körper.

Theorem 2

Für jede Primzahlpotenz $p^f, p \in \mathbb{P}$ und $f \in \mathbb{N}$ gibt es (bis auf Isomorphie) genau einen endlichen Körper K mit $|K| = p^f$.

Die durch Primzahlpotenzen entstandenen endlichen Körper basieren nun nicht bloß auf den Restklassen des Unterkörpers \mathbb{F}_p , sondern werden mit **Polynomen** erweitert.

2.4.3 Polynome

Durch die Verkettung der beiden Operationen eines Körpers \mathbb{K} lassen sich **Polynome** erzeugen. Die Menge aller Polynome über einem Körper \mathbb{K} wird definiert als

$$\mathbb{K}[x] = \{p(x) = a_n x^n + \dots + a_1 x + a_0 \mid a_i \in \mathbb{K}\}$$

[vgl. 16, S. 118f.] und bildet einen kommutativen Ring, aufgrund der fehlenden Abgeschlossenheit der Division jedoch keinen Körper. Wird ein Polynom durch ein anderes dividiert, so kann dabei ein Rest entstehen, was die Anwendung des Kongruenzbegriffs nahelegt. Die Polynomdivision aller Polynome aus einem Körper \mathbb{K} mit einem Polynom $m(x)$ erzeugt einen **Restklassenring** $\mathbb{K}[x]/(m(x))$ mit der bekannten Addition und Multiplikation auf Polynomen.

Die Frage, ob jener Ring auch ein (endlicher) Körper ist, ergibt sich daraus, ob zu jedem Polynom ein multiplikatives Inverses Teil des Rings ist. Dafür ist folgende Eigenschaft eines Polynoms relevant:

Definition 17

Ein Polynom $p(x) \in \mathbb{K}$ vom Grad (also dem höchsten Exponenten des Polynoms) $\deg(p) > 1$ wird als **irreduzibel** bezeichnet, falls es kein Polynom $q(x) \in \mathbb{K}$ mit $0 < \deg(q) < \deg(p)$ gibt, das $p(x)$ teilt [vgl. 16, S. 129].

Wird ein Polynomring nun durch reduzibles Polynom dividiert, also durch ein Polynom, dass sich selbst weiter dividieren lässt, existiert zu jedem seiner Teiler kein multiplikatives Inverses mehr, weshalb dadurch folglich kein Körper beschrieben wird.

Theorem 3

Ein Ring $\mathbb{K}[x]/(m(x))$ ist genau dann ein Körper, wenn $m(x)$ irreduzibel über \mathbb{K} ist [vgl. 16, S. 130].

Ein Galois-Körper \mathbb{F}_p kann auf diese Weise zu einem Körper \mathbb{F}_{p^k} erweitert werden, in dem aus dem Polynomring $\mathbb{F}_p[x]$ ein irreduzibles Polynom $m(x)$ vom Grad k ausgewählt wird. Dann ist

$$\mathbb{F}_{p^k} \Leftrightarrow \mathbb{F}_p[x]/(m(x))$$

Beispiel 7

Typische Beispiele für endliche Körper sind:

- Über dem Körper \mathbb{F}_2 ist x^2+x+1 ein irreduzibles Polynom, das zur Körpererweiterung genutzt werden kann und nun den endlichen Körper $\mathbb{F}_{2^2} = \{[0], [1], [x], [x + 1]\}$ erzeugt.
- Für das AES-Kryptosystem und eine Vielzahl weiterer kryptographischer und codierungstheoretischer Anwendungen wird der Körper \mathbb{F}_{2^8} mit dem definierenden Polynom $x^8 + x^4 + x^3 + x + 1$ verwendet. Er enthält 256 Elemente [vgl. 16, S. 131].

3 Lineare fehlerkorrigierende Codes für kryptographische Zwecke

Mithilfe der vorgestellten Polynome soll nun eine Klasse von Blockcodes betrachtet werden, zu der auch *Reed-Solomon-Codes* und *Goppa-Codes*, die beiden Codes aus den Verfahren von NIEDERREITER und MC ELIECE, zählen. Diese Klasse wird als **lineare Codes** bezeichnet.

Für die Definition linearer Codes ist eine Wiederholung der algebraischen Struktur des **Vektorraums** notwendig. In [15] ist folgende Definition zu finden:

Definition 18

Sei K ein Körper. Eine Menge V mit einer additiven Verknüpfung

$$+: V \times V \rightarrow V; \quad \langle v, w \rangle \mapsto v + w$$

und einer Skalarmultiplikation

$$\circ: K \times V \rightarrow V; \quad \langle \lambda, v \rangle \mapsto \lambda \circ v$$

heißt **K -Vektorraum**, sofern gilt:

- $(V, +)$ ist **abelsche Gruppe**.
- Die Skalarmultiplikation ist sowohl bezüglich des Skalars als auch der Komponente distributiv und assoziativ, ferner ist das Element $1 \in K$ das neutrale Element für diese Skalarmultiplikation. So muss $\forall \lambda, \mu \in K; v, w \in V$ gelten:

$$(\lambda + \mu) \circ v = \lambda \circ v + \mu \circ v$$

$$\lambda \circ (v + w) = \lambda \circ v + \lambda \circ w$$

$$\lambda \circ (\mu \circ v) = (\lambda \circ \mu) \circ v$$

$$1 \circ v = v$$

[vgl. 15, S. 95]

Vektorräume können mit einer beliebigen Dimension konstruiert werden und beinhalten folglich Vektoren dieser Dimension, deren Komponenten addiert und mit Skalaren multipliziert werden können. Für die Definition linearer Codes ist es erforderlich, Vektorräume \mathbb{F}_q^n auf Primzahlkörpern \mathbb{F}_q zu definieren.

Theorem 4

Endliche Körper \mathbb{F}_q bilden \mathbb{F}_q^n -Vektorräume.

Beweis. Es bietet sich an, induktiv vorzugehen:

- **Annahme:** \mathbb{F}_q^n bildet einen Vektorraum $V(q, n)$.
- **Induktionsanfang:** Zu zeigen ist, dass \mathbb{F}_q^1 die Forderungen eines Vektorraums erfüllt.
 - Die Forderung danach, dass $(\mathbb{F}_q, +)$ eine abelsche Gruppe bildet, folgt unmittelbar aus der Körperdefinition.
 - Die Forderungen nach Distributivität und Assoziativität der Skalarmultiplikation bzw. Addition folgen direkt aus der Ringeigenschaft, da $\lambda, \mu, v \in \mathbb{F}_q^1$.
 - Die Existenz eines neutralen Elements bezüglich der Skalarmultiplikation ergibt sich daraus, dass gemäß Definition $q \geq 2 > 1 \Rightarrow 1 \in \mathbb{F}_q$ gilt.
- **Induktionsschritt** $1 \rightarrow n$: Sei die komponentenweise Addition induktiv definiert als

$$(v_1, v_2, \dots, v_n) + (w_1, w_2, \dots, w_n) = (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n)$$

und die Skalarmultiplikation als

$$k \circ (v_1, v_2, \dots, v_n) = (k \circ v_1, k \circ v_2, \dots, k \circ v_n)$$

. Alle obigen Forderungen folgen damit induktiv und \mathbb{F}_q^n ist ein \mathbb{F}_q -Vektorraum.

□

Definition 19

Ein **Untervektorraum** W eines Vektorraums V ist eine Teilmenge $W \subset V$, für die gilt:

- $W \neq \emptyset$
- $v, w \in W : v + w \in W$

- $v \in W, \lambda \in K : \quad \lambda \circ v \in W$

[vgl. 15, S. 96]

Definition 20

Sei \mathbb{F}_q ein endlicher Körper, $|K| = q$. Dann ist ein Blockcode C ein **linearer Code** mit dem Alphabet \mathbb{F}_q und Länge n , falls die Menge C ein **Untervektorraum** von \mathbb{F}_q^n ist [vgl. 8, S. 29].

Folglich beschreibt ein linearer Code Tupel der Länge n mit Komponenten aus \mathbb{F}_q . Ferner stellt die Untervektorraum-Eigenschaft sicher, dass die Addition von Codeworten und die Multiplikation eines Codewortes mit einem Skalar aus dem übergeordneten endlichen Körper \mathbb{F}_q weiterhin valide Codeworte erzeugt. Dies ist eine elementare Eigenschaft linearer Codes.

Eine Subklasse der linearen Codes sind die **zyklischen Codes**.

3.1 Zyklische Codes

Definition 21

Ein k -dimensionaler Untervektorraum C von \mathbb{F}_q^n ist dann ein **zyklischer Code**, wenn

$$\forall \langle a_0, a_1, \dots, a_{n-1} \rangle \in C : \quad \langle a_{n-1}, a_0, a_1, \dots, a_{n-2} \rangle \in C$$

gilt [vgl. 5, S. 42].

Ein linearer Code ist folglich dann zyklisch, falls zu jedem Codewort auch das Codewort Teil des Codes ist, das durch eine zyklische Verschiebung entsteht. Ein solcher Code ist beispielsweise der *gewöhnliche Reed-Colomon-Code*, der neben der verallgemeinerten, nicht zyklischen Version Gegenstand des nächsten Abschnitts ist.

Theorem 5

Zyklische Codes der Länge n sind Ideale des Rings $R = \mathbb{F}_q[x]/(x^n - 1)$.

Das Polynom $x^n - 1$ fungiert hierbei als Modul, sodass der Ring R folglich die Restklassen bezüglich dieser modularen Reduktion von \mathbb{F}_q enthält.

Bemerkung 8

Sei C ein linearer zyklischer Code der Länge n über einem Körper F_q und $a = a(x) = a_0 + a_1 \cdot x + \dots + a_{n-1} \cdot x^{n-1} \in C$. Dann ist

$$\begin{aligned} x \cdot a &= x \cdot (a_0 + a_1 \cdot x + \dots + a_{n-1} \cdot x^{n-1}) \\ &= a_0 \cdot x + a_1 \cdot x^2 + \dots + a_{n-1} \cdot x^n \\ &= a_{n-1} + a_0 \cdot x + \dots + a_{n-2} \cdot x^{n-1} \end{aligned}$$

genau das zyklisch verschobene Codewort $a' \in C$.

Dies folgt aus $x^n \equiv 1 \pmod{x^n - 1}$ [vgl. 6, S. 149f.].

Da durch die Linearität des Codes C lässt sich ein $a \in C$ nicht nur beliebig oft zyklisch verschieben, also mit x multiplizieren, sondern auch lineare Kombinationen (also Summen) gültiger Codeworte ergeben wiederum ein gültiges Codewort in C . Somit ergibt auch das Produkt eines Codewortes mit einem Polynom, also

$$a(x) \cdot c(x) \in C, c(x) \in C, a \in \mathbb{F}_q[x]/(x^n - 1)$$

ein gültiges Codewort [vgl. 17, S. 121].

Da Ideale eines Polynomrings erzeugende Polynome besitzen, ist die Angabe eines **Generatorpolynoms** $g(x)$ möglich, für das gilt:

- $\deg g(x) = n - k$
- $g(x) \mid (x^n - 1)$

Theorem 6

Sei C ein linearer zyklischer Code über einem endlichen Körper \mathbb{F}_q . Dann gibt es ein Polynom $g(x)$, für das $C = \langle g(x) \rangle$ gilt.

Ein Codewort beziehungsweise ein Codewortpolynom $c(x)$ eines linearen, zyklischen Codes mit Generatorpolynom $g(x)$ kann nun als Produkt von Nachrichtenpolynom $m(x)$ und Generatorpolynom aufgefasst werden [vgl. 17, S. 121].

$$c(x) = m(x) \cdot g(x)$$

Bei vielen aktuellen Anwendungen von *Reed-Solomon-Codes* werden Generatorpolynome angegeben und genutzt [vgl. 18, S. 6].

3.2 Definition der Klasse von Reed-Solomon-Codes

Die grundlegende Idee von linearen (zyklischen) Codes über Polynomen besteht darin, Fehler über Interpolation zu korrigieren. Bildlich gesehen soll also ein Codewort konstruiert werden, dessen einzelnen Komponenten entlang einer interpolierbaren Abbildung aufgereiht sind. Sollte nun ein Übertragungsfehler die Lage eines Punktes (also einer Komponente) verändern, so ist die korrekte Lage durch die anderen Punkte jedoch hinreichend bestimmt und kann korrigiert werden.

Idee Eine Nachricht der Länge k in Form eines Elementes eines endlichen Körpers \mathbb{F}_q^k kann als Polynom aufgefasst werden, indem die entsprechenden Komponenten der Nachricht die Koeffizienten des Polynoms bilden. Werden nun n paarweise verschiedene Stützstellen aus dem dem Code zugrundeliegenden Körper \mathbb{F}_q^n ausgewählt, so ergibt die Auswertung des Polynoms an den jeweiligen Stützstellen die Komponenten eines betrachteten Codewortes.

Durch bereits k Komponenten des Codewortes ist das Polynom hinreichend bestimmt, sodass $n - k$ Punkte, die während der Übertragung ausgelöscht wurden, und $(n - k)/2$ verfälscht übertragene Punkte als Fehler erkannt und korrigiert werden können.

Definition 22

Sei $u = \langle u_1, u_2, \dots, u_n \rangle \in \mathbb{F}_q^n$, wobei $\forall i, j \in \{1, \dots, n\} : u_i \neq u_j$. Dann ist $a = \langle a_1, a_2, \dots, a_n \rangle \in \mathbb{F}_q^n$ ein Codewort eines **Reed-Solomon-Codes** C mit Länge n , Dimension k und Minimaldistanz d , falls es ein Polynom $f(x)$ mit Grad $\deg f < k$ gibt, sodass $f(u_i) = a_i$ ist. Die Menge aller dieser Codeworte a ist C :

$$C = \{a \mid a_i = f(u_i), \deg f(x) < k, f \in \mathbb{F}_q^n[x]\}$$

[vgl. 19, S. 9]

Beispiel 8

Sei $m = \langle 0, 1, 2, 1 \rangle \in \mathbb{F}_7^4$. Dann ist $m(X) = 0 + 1 \circ X + 2 \circ X^2 + 1 \circ X^3 \in \mathbb{F}_7[X]$ die polynomiale Darstellung dieser Nachricht m .

Betrachtet werde nun ein **Reed-Solomon-Code** C mit Länge $n = 7$ und Dimension $k = 4$:

$$C = \{a \mid a_i = f(u_i), \deg f(x) < 4, f \in \mathbb{F}_7[x]\}$$

- Gewählt sind folglich n paarweise verschiedene Stützstellen, deren Wahlfreiheit in \mathbb{F}_7 lediglich in der Anordnung besteht. Möglich ist beispielsweise $u = \langle 0, 1, 2, 3, 4, 5, 6 \rangle \in \mathbb{F}_7^7$.

• Auswertung:	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: center;">Stützstelle u_i</th><th style="text-align: center;">0</th><th style="text-align: center;">1</th><th style="text-align: center;">2</th><th style="text-align: center;">3</th><th style="text-align: center;">4</th><th style="text-align: center;">5</th><th style="text-align: center;">6</th></tr> </thead> <tbody> <tr> <td style="text-align: center;">Auswertung $m(u_i)$</td><td style="text-align: center;">0</td><td style="text-align: center;">4</td><td style="text-align: center;">4</td><td style="text-align: center;">6</td><td style="text-align: center;">2</td><td style="text-align: center;">5</td><td style="text-align: center;">0</td></tr> </tbody> </table>	Stützstelle u_i	0	1	2	3	4	5	6	Auswertung $m(u_i)$	0	4	4	6	2	5	0
Stützstelle u_i	0	1	2	3	4	5	6										
Auswertung $m(u_i)$	0	4	4	6	2	5	0										

- Folglich ergibt sich das Codewort $c = \langle 0, 4, 4, 6, 2, 5, 0 \rangle \in \mathbb{F}_7^7$.

Bemerkung 9

Das Generatorpolynom zu einem *Reed-Solomon-Code* ist definiert durch

$$g(x) = \prod_{j=1}^{2t} (x - \alpha^j)$$

wobei t die Anzahl der durch diesen Code korrigierbaren Fehler und α ein primitives Element des zugrundeliegenden Körpers repräsentiert, es gilt $n - k = 2t$ [vgl. 20, S. 17].

3.2.1 Verallgemeinerte Reed-Solomon-Codes

Reed-Solomon-Codes lassen sich verallgemeinern, in dem ein zusätzlicher Vektor $v = \langle v_1, \dots, v_n \rangle$ eingebracht wird.

Definition 23

Sei $u = \langle u_1, u_2, \dots, u_n \rangle \in \mathbb{F}_q^n$, wobei $\forall i, j \in \{1, \dots, n\}: u_i \neq u_j$. Ferner sei $v = \langle v_1, v_2, \dots, v_n \rangle$, $\forall j \in \{1, \dots, n\}: v_j \in \mathbb{F}_q \setminus \{0\}$. Der **verallgemeinerte Reed-Solomon-Code** $GRS_k(u, v)$ besteht nun aus der Menge der Codeworte

$$c = \langle v_1 f(u_1), v_2 f(u_2), \dots, v_n f(u_n) \rangle$$

für Polynome $f \in \mathbb{F}_q[x]/(x^n - 1)$ mit $\text{Grad } \deg f < k$ [vgl. 19, S. 13].

Verallgemeinerte Reed-Solomon-Codes sind im Allgemeinen nicht zyklisch, folglich entfällt die Möglichkeit, ein Generatorpolynom anzugeben. Die Codierung basiert entsprechend auch nicht mehr auf einer Multiplikation des Informationstupels mit dem Generatorpolynom, sondern auf einer Multiplikation mit einer Generatormatrix, die im Folgenden betrachtet wird. Verallgemeinerte Reed-Solomon-Codes finden Verwendung im Ansatz von Niederreiter, da aufgrund der zusätzlich eingebrachten **Gewichte** die Anzahl der möglichen Codes signifikant steigt und die Codeklasse somit an kryptographischer Bedeutung gewinnt.

Für den weiteren Verlauf der Arbeit ist die Betrachtung **dualer Codes** notwendig. Duale Codes können zu linearen Codes gebildet werden:

3.2.2 Duale Codes

Definition 24

Sei C ein linearer $[n, k, d]_q$ -Code mit Länge n , Dimension k und Minimaldistanz d über einem Körper \mathbb{F}_q . Ferner bezeichne die Operation \cdot das Skalarprodukt über dem Untervektorraum \mathbb{F}_q^n , das definiert ist durch

$$\langle a_1, a_2, \dots, a_n \rangle \cdot \langle b_1, b_2, \dots, b_n \rangle = \sum_{i=1}^n a_i b_i$$

Nun ist der zu C **duale Code** C^\perp definiert als

$$C^\perp = \{v \in \mathbb{F}_p^n \mid v \cdot c = 0, \forall c \in C\}$$

[vgl. 8, S. 35f.]

Diese Definition mag an die aus der linearen Algebra im Reellen bekannte Definition von Orthogonalität („Rechtwinkligkeit“) erinnern, jedoch lässt sich dies auf die hier betrachteten Körper nicht übertragen, unter Anderem, da es auch **selbstduale** Codes gibt [vgl. 8, S. 36].

Bemerkung 10

Der zu einem $[n, k, d]_q$ -Code C duale Code C^\perp ist ein $[n, n - k, d]_q$ -Code [vgl. 8, S. 36].

Wird der duale Code zu einem dualen Code eines linearen Codes C gebildet, so ist dieser wieder C , also gilt:

$$C^{\perp\perp} = C$$

Theorem 7

Der duale Code zu einem Reed-Solomon-Code ist ebenfalls ein Reed-Solomon-Code und der duale Code zu einem verallgemeinerten Reed-Solomon-Code ist ebenfalls ein verallgemeinerter Reed-Solomon-Code [vgl. 19, S. 14].

So gilt $GRS_k(u, v)^\perp = GRS_{n-k}(u, v')$, wobei für die Komponenten des Vektors v' gilt:

$$v_i \cdot v'_i \prod_{j=1, j \neq i}^n (u_j - u_i) = 1$$

[vgl. 19, S. 14]

3.2.3 Kontroll- und Generatormatrizen

Da lineare Codes wie gezeigt Vektorräume bilden, ist die Angabe von **Basen** möglich. Durch die Definition der Dualität von Codes zueinander, können **Generator-** und **Kontrollmatrizen** angegeben werden.

Definition 25

Sei C ein linearer $[n, k, d]_q$ -Code und C^\perp der duale Code zu C .

Ferner sei $(\langle h_{11}, h_{12}, \dots, h_{1n} \rangle, \langle h_{21}, \dots, h_{2n} \rangle, \dots, \langle h_{(n-k)1}, \dots, h_{(n-k)n} \rangle)$ eine Basis zu C^\perp .

Dann ist die Matrix

$$H = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \dots & \dots & \dots & \dots \\ h_{(n-k)1} & h_{(n-k)2} & \dots & h_{(n-k)n} \end{pmatrix}$$

die **Generatormatrix** von C^\perp und **Kontrollmatrix** oder **Paritätsprüfmatrix** von C [vgl. 8, S. 37]

Der Zusammenhang zwischen Generator- und Kontrollmatrix ergibt sich daraus, dass über die Generatormatrix des dualen Codes C^\perp zu C überprüft werden kann, ob für ein $x \in \mathbb{F}_q$

auch $x \in C$ gilt, indem die Skalarprodukte $x \cdot h_i, \forall i \in \{1, \dots, n-k\}$ gebildet werden. Sollte für mindestens eines dieser Produkte $x \cdot h_i \neq 0$ gelten, so ist $x \notin C$ [vgl. 8, S. 37]. So ließe sich beispielsweise eine Identitätsabbildung $1_C(x)$ auf C bilden.

Generatormatrizen des verallgemeinerten Reed-Solomon-Codes

Zur Kodierung eines Informationstupels mittels des verallgemeinerten Reed-Solomon-Codes wird die Generatormatrix des Codes benötigt.

Bemerkung 11

Gegeben sei $GRS_k(u, v)$ und der dazu duale Code $GRS_{n-k}(u, y)$. Die Paritätsprüfmatrix zu $GRS_k(u, v)$ ist dann genau die Generatormatrix zu $GRS_{n-k}(u, y)$ und definiert durch

$$H = (y_i u_i^j)_{0 \leq j \leq n-k-1, 1 \leq i \leq n} = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_1 u_1 & y_2 u_2 & \cdots & y_n u_n \\ \vdots & \vdots & & \vdots \\ y_1 u_1^{n-k-1} & y_2 u_2^{n-k-1} & \cdots & y_n u_n^{n-k-1} \end{pmatrix}$$

Entsprechend ist die Generatormatrix zu $GRS_k(u, v)$ definiert durch

$$G = (v_i u_i^j)_{0 \leq j \leq k-1, 1 \leq i \leq n} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \\ v_1 u_1 & v_2 u_2 & \cdots & v_n u_n \\ \vdots & \vdots & & \vdots \\ v_1 u_1^{k-1} & v_2 u_2^{k-1} & \cdots & v_n u_n^{k-1} \end{pmatrix}$$

[vgl. 17, S. 277] und [vgl. 21, S. 304].

Diese Matrizen sind für die Kodierung und Dekodierung – insbesondere auch für die beabsichtigte kryptographische Verwendung – von besonderer Bedeutung, da mithilfe ihrer Codes definiert werden können und die Frage, ob ein Wort Codewort dieses Codes ist, nun entscheidbar wird. Generatormatrizen lassen sich durch elementare Zeilen- und Spaltenoperationen verändern, sodass zwar weiterhin gültige Codeworte eines Codes C gebildet werden, jedoch wird die Projektion des Informationstupels auf ein Codewort dadurch beeinflusst, sodass womöglich andere Codeworte entstehen. Diese Eigenschaft wird genutzt, um beispielsweise eine systematische Codierung zu ermöglichen.

3.3 Kodierung, Fehlerkorrektur und Dekodierung

In diesem Abschnitt wird mit der algorithmischen Betrachtung der Kodierung und Dekodierung samt der Fehlerkorrektur verallgemeinerter Reed-Solomon-Codes die Grundlage dafür gelegt, ein kryptographisches Verfahren auf Basis fehlerkorrigierender Codes zu konstruieren, wie es McELIECE und NIEDERREITER in ihren Ansätzen tun.

Sämtliche Listings in diesem Abschnitt zeigen Implementierungen im auf Python basierenden Mathematiksoftwaresystem *SageMath* [22]. Für die Ausführung dieser Skripte ist ein gegebenenfalls auch virtualisiertes Linux-Betriebssystem erforderlich.

3.3.1 Kodierung

Bei der Codierung linearer Codes im Allgemeinen werden zwei Ansätze unterschieden:

- **nicht-systematische Codierung** – Hier ergeben sich die Codeworte direkt aus der Multiplikation von Informationstupel und einer beliebigen Generatormatrix des Codes.
- **systematische Codierung** – Hier wird eine Generatormatrix der Form $[I_k \mid C]$ verwendet, wobei I die $k \times k$ -Identitätsmatrix und C die $k \times (n-k)$ -Codeworterzeugungsmatrix ist. Eine solche Generatormatrix kann durch elementare Zeilen- und Spaltenoperationen erzeugt werden [vgl. 17, S. 84f.].

Nicht-Systematische Kodierung

Sei $h = \langle h_0, \dots, h_{k-1} \rangle \in \mathbb{F}_q^k$ ein Informationstupel und $GRS_k(a, v)$ mit $a = \langle a_0, \dots, a_{n-1} \rangle$, $\forall i, j \in \{0, \dots, n-1\}: a_i, a_j \in \mathbb{F}_q \wedge a_i \neq a_j$ und $v = \langle v_0, \dots, v_{n-1} \rangle$, $\forall i \in \{0, \dots, n-1\}: v_i \in \mathbb{F}_q \setminus \{0\}$ ein verallgemeinerter Reed-Solomon-Code.

Dann ist eine Generatormatrix zu $GRS_k(a, v)$ definiert als

$$G = \begin{pmatrix} v_0 & v_1 & \cdots & v_{n-1} \\ v_0 a_0 & v_1 a_1 & \cdots & v_{n-1} a_{n-1} \\ \vdots & \vdots & & \vdots \\ v_0 a_0^{k-1} & v_1 a_1^{k-1} & \cdots & v_{n-1} a_{n-1}^{k-1} \end{pmatrix}$$

Die nicht-systematische Codierung ist nun gegeben durch die Zeilenvektor-Matrix-Multiplikation von h und G [vgl. 23, S. 37f]:

$$\begin{aligned} c = h \cdot G &= \left(\begin{array}{ccc} h_0 & \dots & h_{k-1} \end{array} \right) \cdot \begin{pmatrix} v_0 & v_1 & \dots & v_{n-1} \\ v_0 a_0 & v_1 a_1 & \dots & v_{n-1} a_{n-1} \\ \vdots & \vdots & & \vdots \\ v_0 a_0^{k-1} & v_1 a_1^{k-1} & \dots & v_{n-1} a_{n-1}^{k-1} \end{pmatrix} \\ &= \left(\begin{array}{cccc} h_0 v_0 + h_1 v_0 a_0 + \dots + h_{k-1} v_0 a_0^{k-1} & \dots & h_0 v_{n-1} + h_1 v_{n-1} a_{n-1} + \dots + h_{k-1} v_{n-1} a_{n-1}^{k-1} \end{array} \right) \\ &= \langle c_0, \dots, c_{n-1} \rangle \end{aligned}$$

Systematische Codierung

Wie bereits beschrieben, ist es für eine systematische Codierung lediglich erforderlich, die Generatormatrix durch Zeilen-/Spaltenoperationen in eine systematische Form zu bringen. Dies kann durch den GAUSS'SCHEN Eliminationsalgorithmus mit Pivotierung erreicht werden [vgl. 17, S. 87f].

Nun liegt die Generatormatrix in folgender Form vor:

$$G_S = \left(\begin{array}{cccc|cccccc} 1 & 0 & \dots & 0 & p_{0,0} & p_{0,1} & \dots & p_{0,n-k-1} \\ 0 & 1 & \dots & 0 & p_{1,0} & p_{1,1} & \dots & p_{1,n-k-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & p_{k-1,0} & p_{k-1,1} & \dots & p_{k-1,n-k-1} \end{array} \right)$$

Die systematische Codierung bewirkt nun, dass das Informationstupel unverändert die ersten k Stellen des Codewortes ergibt. Die restlichen Stellen sind $n - k$ Paritätsprüfbits, die über den rechten Teil der Generatormatrix erzeugt werden:

$$\begin{aligned} c = h \cdot G_S &= \left(\begin{array}{ccc} h_0 & \dots & h_{k-1} \end{array} \right) \cdot \left(\begin{array}{cccc|cccccc} 1 & 0 & \dots & 0 & p_{0,0} & p_{0,1} & \dots & p_{0,n-k-1} \\ 0 & 1 & \dots & 0 & p_{1,0} & p_{1,1} & \dots & p_{1,n-k-1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & p_{k-1,0} & p_{k-1,1} & \dots & p_{k-1,n-k-1} \end{array} \right) \\ &= \left(\begin{array}{cccc|cc} h_0 & h_1 & \dots & h_{k-1} & \sum_{i=0}^{k-1} h_0 p_{i,0} & \dots & \sum_{i=0}^{k-1} h_{k-1} p_{i,n-k-1} \end{array} \right) \\ &= \langle h_0, h_1, \dots, h_{k-1}, c_k, c_{k+1}, \dots, c_{n-1} \rangle \end{aligned}$$

Implementierung

Im Anhang A.1 ist die Implementierung des Kodievorgangs eines verallgemeinerten Reed-Solomon-Codes über endlichen Körpern in *SageMath* dargestellt. Das Skript generiert aus den Stützstellen und Gewichten eine Generatormatrix für die nicht-systematische Kodierung. Die Generatormatrix für die systematische Kodierung wird dadurch erzeugt, dass die $k \times k$ -Untermatrix der ersten k Spalten invertiert und mit G multipliziert wird. Dadurch wird die systematische Form erreicht. Referenzen und alternative Algorithmen sind dem Anhang A.2 zu entnehmen.

3.3.2 Dekodierung

Nachdem (nicht-)systematisch erzeugte Codeworte nun über einen unzuverlässigen Kanal übertragen wurden, werden Fehlerkorrektur- und Dekoderalgorithmen angewandt, um die übertragenen Informationen zu erhalten.

Im Folgenden wird der **syndrombasierte** Dekoderalgorithmus für eine maximale Fehleranzahl von $t = \lfloor \frac{n-k}{2} \rfloor$ dargestellt. Der alternative Ansatz des **listen-basierten** Dekodierens, der die Korrektur von einer größeren Anzahl an Fehlern ermöglicht, wird ohne vertiefte Darstellung erwähnt.

Syndrombasierter Dekoderalgorithmus

Gegeben sei ein verallgemeinerter Reed-Solomon-Code $GRS_k(a, v)$ mit Auswertungsstellen $a = \langle a_0, \dots, a_{n-1} \rangle$, $\forall i, j \in \{0, \dots, n-1\}: a_i, a_j \in \mathbb{F}_q \wedge a_i \neq a_j$ und Gewichten $v = \langle v_0, \dots, v_{n-1} \rangle$, $\forall i \in \{0, \dots, n-1\}: v_i \in \mathbb{F}_q \setminus \{0\}$. Ferner sei α primitives Element des zugrundeliegenden Körpers \mathbb{F}_q .

Nach einem Datenübertragungsvorgang wurde ein Tupel $r = \langle r_0, r_1, \dots, r_{n-1} \rangle$ empfangen, das sich aus dem korrekten **Codeworttupel** $c = \langle c_0, c_1, \dots, c_{n-1} \rangle$ und einem **Fehlertupel** $e = \langle e_0, e_1, \dots, e_{n-1} \rangle$ zusammensetzt, wobei die Anzahl der Komponenten von e , die nicht null sind, nicht größer als $t = \lfloor \frac{n-k}{2} \rfloor$ sein darf, andernfalls ist es nicht möglich, mit diesem Dekoderalgorithmus das korrekte Codewort aus r zu ermitteln. Im Folgenden bezeichne k_1, \dots, k_ϵ die **Fehlerstellen** im empfangenen Wort, wobei $\epsilon < t$ gelten muss. α^{k_j} bezeichne die **Fehlstellennummer** und e_{k_j} die **Fehlergröße** an der Stelle k_j .

Basis des Dekodierens bilden nun $n - k - 1$ Syndrome, die sich aus der Multiplikation der Paritätsprüfmatrix H mit dem empfangenen Wort r ergeben:

Definition 26

Das l -te **Syndrom** eines Wortes r sei definiert als

$$S_l = \sum_{j=0}^{n-1} r_j v_j \alpha_j^l$$

[vgl. 24, S. 184]

Aus der Definition der Paritätsprüfmatrix folgt, dass r genau dann ein gültiges Codewort des betrachteten $GRS_k(a, v)$ -Codes ist, falls alle Syndrome gleich null sind. Da zweifelsfrei für jedes Codewort-Syndrom $\sum_{j=0}^{n-1} c_j v_j \alpha_j^l = 0$ gilt, sind die Fehlerwortsyndrome äquivalent zu den Syndromen obiger Definition. Ferner gilt sogar $S_l = \sum_{j=1}^e e_{k_j} v_j \alpha_j^l$, da die Fehlergröße e_j an nicht fehlerhaften Stellen gleich 0 sein muss. Die Betrachtung dieser Syndrome gibt Aufschluss darüber, an welchen Stellen Fehler während der Übertragung aufgetreten sind.

Durch die Syndrome ergibt sich folgendes Gleichungssystem:

$$\begin{aligned} S_1 &= e_0 v_0 \alpha_0 + e_1 v_1 \alpha_1 + \dots + e_{n-1} v_{n-1} \alpha_{n-1} \\ S_2 &= e_0 v_0 \alpha_0^2 + e_1 v_1 \alpha_1^2 + \dots + e_{n-1} v_{n-1} \alpha_{n-1}^2 \\ &\vdots \\ S_{n-k-1} &= e_0 v_0 \alpha_0^{n-k-1} + e_1 v_1 \alpha_1^{n-k-1} + \dots + e_{n-1} v_{n-1} \alpha_{n-1}^{n-k-1} \end{aligned}$$

Da dieses Gleichungssystem jedoch angesichts der Exponenten nicht linear ist, lässt sich ohne großen Rechenaufwand keine triviale Lösung bestimmen, die jedoch für die Berechnung der Fehlerpositionen und Fehlergrößen notwendig wäre [vgl. 17, S. 248] [vgl. 23, S. 180f] [vgl. 25, S. 148].

Stattdessen wird ein **Fehlerlokalisationspolynom** definiert:

$$\sigma(x) = \prod_{j=1}^e (1 - x \cdot \alpha_{k_j}) = 1 + \sum_{i=1}^e \sigma_i x^i$$

Jede inverse Fehlerstellenummer $\alpha_{k_j}^{-1}$ ist Wurzel dieses Polynoms [vgl. 23, S. 180f] [vgl. 24, S. 185]. Ziel ist es nun, die Unbekannten σ_j zu bestimmen, um damit die Fehlerstellenummern ermitteln zu können.

Gemäß obiger Definition folgt

$$\sigma(\alpha_{k_j}^{-1}) = 1 + \sigma_1 \alpha_{k_j}^{-1} + \sigma_2 \alpha_{k_j}^{-2} + \dots + \sigma_\epsilon \alpha_{k_j}^{-\epsilon} = 0$$

Um nun ein lineares Gleichungssystem zu erhalten, dessen Lösung die Fehlerstellenummern angibt, ist es erforderlich, das durch die Syndrome erzeugte Gleichungssystem mit dem Fehlerlokalisationspolynom in Verbindung zu setzen [vgl. 23, S. 181f.][vgl. 25, S. 148f.]. Zunächst wird obige Gleichung dafür mit $e_j v_j \alpha_j^{l+\epsilon}$ multipliziert:

$$e_j v_j \alpha_j^{l+\epsilon} + \sigma_1 e_j v_j \alpha_j^{l+\epsilon-1} + \dots + \sigma_\epsilon e_j v_j \alpha_j^l = 0$$

Um so nun zu den Syndromen zu gelangen, ist die Summenbildung für $j = 1, \dots, e$ erforderlich:

$$\sum_{j=1}^{\epsilon} e_j v_j \alpha_j^{l+\epsilon} + \sigma_1 \sum_{j=1}^{\epsilon} e_j v_j \alpha_j^{l+\epsilon-1} + \dots + \sigma_\epsilon \sum_{j=1}^{\epsilon} e_j v_j \alpha_j^l = 0$$

Gemäß der Definition der Syndrome ergibt sich

$$S_{l+e} + \sigma_1 S_{l+e-1} + \dots + \sigma_\epsilon S_l = 0$$

und damit ein lineares Gleichungssystem

$$\begin{aligned} S_1 \sigma_\epsilon + S_2 \sigma_{\epsilon-1} + \dots + S_\epsilon \sigma_1 &= 0 \\ S_2 \sigma_\epsilon + S_3 \sigma_{\epsilon-1} + \dots + S_{\epsilon+1} \sigma_1 &= 0 \\ &\vdots \\ S_\epsilon \sigma_\epsilon + S_{\epsilon+1} \sigma_{\epsilon-1} + \dots + S_{2\epsilon-1} \sigma_1 &= 0 \end{aligned}$$

Dieses Gleichungssystem wird lösbar, sobald die korrekte Fehleranzahl ϵ bestimmt wurde. Dafür wird die in [23] beschriebene Eigenschaft ausgenutzt, dass die größte nicht-singuläre

durch Syndrome beschriebene Matrix der Form

$$M_\mu = \begin{pmatrix} S_1 & S_2 & \cdots & S_\mu \\ S_2 & S_3 & \cdots & S_{\mu+1} \\ \vdots & & & \\ S_\mu & S_{\mu+1} & \cdots & S_{2\mu-1} \end{pmatrix}$$

die Anzahl der Fehler angibt. Zunächst wird geprüft, ob die Matrix $\mu \leftarrow t: M_t$ singulär ist. Falls ja, so wird μ iterativ um 1 reduziert, bis M_μ nicht-singulär ist [vgl. 23, S. 181f.].

Um nun die Fehlerstellennummern α^{k_j} bestimmen zu können, müssen die Wurzeln des Fehlerlokalisationspolynoms durch Berechnen von $\sigma(\alpha^i)$ für alle $i \in \{0, \dots, n-1\}$ ermittelt und anschließend invertiert werden.

Der finale Schritt des hier betrachteten Algorithmus besteht darin, die Fehlergrößen aus den Syndromen zu ermitteln. Dafür wird das Gleichungssystem

$$\begin{aligned} S_1 &= e_{k_1} v_{k_1} \alpha^{k_1} + e_{k_2} v_{k_2} \alpha^{k_2} + \dots + e_{k_\epsilon} v_{k_\epsilon} \alpha^{k_\epsilon} \\ S_2 &= e_{k_1} v_{k_1} \alpha^{2k_1} + e_{k_2} v_{k_2} \alpha^{2k_2} + \dots + e_{k_\epsilon} v_{k_\epsilon} \alpha^{2k_\epsilon} \\ &\vdots \\ S_\epsilon &= e_{k_1} v_{k_1} \alpha^{\epsilon k_1} + e_{k_2} v_{k_2} \alpha^{\epsilon k_2} + \dots + e_{k_\epsilon} v_{k_\epsilon} \alpha^{\epsilon k_\epsilon} \end{aligned}$$

für die Fehlergrößen e_{k_j} gelöst. Daraus ergibt sich das Fehlerstellentupel $e = \langle e_0, e_1, \dots, e_{n-1} \rangle$, das nun vom empfangenen Wort abgezogen wird, um das korrekte Codewort zu erhalten [vgl. 23, S. 182f.] [vgl. 25, S. 150]:

$$\langle c_0, c_1, \dots, c_{n-1} \rangle = \langle r_0, r_1, \dots, r_{n-1} \rangle - \langle e_0, e_1, \dots, e_{n-1} \rangle$$

Der hier vorgestellte Algorithmus folgt den in [23] beschriebenen Schritten und wird als *Peterson-Gorenstein-Zierler*-Algorithmus bezeichnet.

Implementierung

Weitere Algorithmen und Ansätze

A Anhang

A.1 Implementierung der GRS-Kodierung

```
1 #!/usr/bin/env sage
2 print("### GENERALIZED REED-SOLOMON ENCODING")
3 print("### This sage script allows you to perform a (non-
      systematic
4    Generalized Reed-Solomon encoding over a finite field F_q.")
5 print("### (C) Roman Wetenkamp, 2023.\n")
6
7 # Basic Code properties
8 q = input("Please type in q, e.g. 2^2:\t\t")
9 n = int(input("Please type in length n:\t\t"))
10 k = int(input("Please type in dimension k:\t\t"))
11
12 # Generate Galois field
13 F_q = None
14 if len(q.split("^")) < 2:
15     F_q.<a> = GF(int(q))
16 elif len(q.split("^")) == 2:
17     F_q.<a> = GF(pow(int(q.split("^")[0]), int(q.split("^")[1])))
18 else:
19     print(f"ERROR: Sorry, unable to work with this q={q}!")
20     exit(1)
21
22 # GRS Code properties
23 a = input(f"\nPlease type in evaluation tuple\na=(a_0, a_1,
      ..., a_{n-1}):\t\t")
24 a = [F_q(x) for x in a[1:-1].split(",")]
25
26 if len(a) != n:
27     print(f"ERROR: Tuple length {len(a)} differs from code length
      {n}!")
28     exit(1)
```

```

29
30     for element in a:
31         if a.count(element) > 1:
32             print("ERROR:\tElements of a must be pairwise distinct!")
33             exit(1)
34     else:
35         pass
36
37 v = input(f"Please type in weight tuple\nv=(v_0, v_1, ..., v_{n-1}):\t\t")
38 v = [F_q(x) for x in v[1:-1].split(",")]
39
40 if len(v) != n:
41     print(f"ERROR:\tTupel length {len(v)} differs from code length {n}!")
42     exit(1)
43
44 for element in v:
45     if v == 0:
46         print("ERROR:\tElements of v must not be zero!")
47         exit(1)
48     else:
49         pass
50
51 # Generate Generator Matrix G of GRS_k(a, v)
52 rows = []
53 i = 0
54 while i < k:
55     row = []
56     j = 0
57     while j < n:
58         row.append(v[j] * (a[j]**i))
59         j += 1
60     rows.append(row)
61     i += 1
62
63 G = matrix(F_q, rows)

```

```

64
65 print(f"\nGenerator\u2022matrix:\u2022\n{G}")
66
67 # Generate systematic Generator Matrix G_S of GRS_k(a, v)
68 A = G.submatrix(0,0,k,k)
69 G_S = A.inverse() * G
70 print(f"\nSystematic\u2022form\u2022matrix:\u2022\n{G_S}")
71
72 # Perform Encoding
73 h = input(f"\nPlease\u2022type\u2022in\u2022message\u2022tupel\u2022\nh\u2022=\u2022(h_0,\u2022\dots,\u2022h_{k-1}): \u2022\t\t\t")
74 h = [F_q(x) for x in h[1:-1].split(",")]
75
76 if len(h) != k:
77     print(f"ERROR:\tTupel\u2022length\u2022{len(h)}\u2022differs\u2022from\u2022input\u2022word\u2022length\u2022{k}!")
78     exit(1)
79
80 h_v = matrix(F_q, h)
81
82 C = h_v * G
83 C_S = h_v * G_S
84
85 print(f"\nCodeword\u2022non-systematic\u2022approach:\t\t{C.rows()[0]}")
86 print(f"Codeword\u2022systematic\u2022approach:\t\t\t{C_S.rows()[0]}")

```

Quelltext A.1: SageMath-Implementierung des Kodieralgorithmus für GRS-Codes

A.2 Transformieren der Generatormatrix in die systematische Form

Die folgende Ausarbeitung stellt die Ergebnisse einer Recherche zu Algorithmen dar, die es ermöglichen, aus nicht-systematischen Generatormatrizen systematische zu konstruieren. Sie entstand als Nebenprodukt zu dieser Studienarbeit und ist hier in Gänze dargestellt, da sie weder publiziert noch zitierfähig ist.

On transforming Generator Matrices of GRS_k Codes for systematic Encoding

A Literature Review

Roman Wetenkamp*



January 29, 2023

Abstract

Linear codes like GRS_k are frequently used for error-correction and even cryptographic purposes. Since the encoding is often required to be *systematic*, the goal of this paper is to find algorithms that transform *non-systematic* generator matrices into *systematic* ones. The conducted literature review shows that most frequently the inverse of the $k \times k$ -submatrix is computed and multiplied to the generator matrix to achieve the *standard form*.

Linear Codes can be represented by a generator matrix and a parity-check matrix which is the generator matrix of the dual code.

An example for linear codes is the class of *Generalized Reed Solomon* codes:

Definition 2 ([2]). Let $a = \langle a_0, a_1, \dots, a_{n-1} \rangle$ be a tupel of pairwise distinct components from \mathbb{F}_q and $v = \langle v_0, v_1, \dots, v_{n-1} \rangle$ be a tupel of components from $\mathbb{F}_q \setminus \{0\}$. The **generalized Reed-Solomon code** $GRS_k(a, v)$ is now defined by all codewords

$$c = \langle v_0 f(a_0), v_1 f(a_1), \dots, v_{n-1} f(a_{n-1}) \rangle$$

for all polynomials $f \in \mathbb{F}_q[x]/(x^n - 1)$ with $\deg f < k$.

The generator matrix for this class of codes as in [2] is given by

$$G_{GRS} = \begin{pmatrix} v_0 & v_1 & \cdots & v_{n-1} \\ v_0 a_0 & v_1 a_1 & \cdots & v_{n-1} a_{n-1} \\ \vdots & \vdots & & \vdots \\ v_0 a_0^{k-1} & v_1 a_1^{k-1} & \cdots & v_{n-1} a_{n-1}^{k-1} \end{pmatrix}$$

Definition 3 ([1]). A generator matrix G is called to be **systematic** if it has the form $G = [I_k \mid P_{n-k}]$ with I_k being the $k \times k$ -identity matrix and P_{n-k} being the $k \times (n-k)$ -matrix for generating parity-check bits.

Since the matrix G_{GRS} is in *non-systematic* form, an algorithm for transforming this matrix into a *systematic* form is required.

1 Introduction

The encoding of information tupels with linear codes is done by matrix multiplication. Since several generator matrices for the same linear code exist, approaches like *systematic encoding* make use of this in order to produce codewords with certain properties. This article deals with the question how existing generator matrices of a linear code can be transformed into *systematic* ones algorithmically.

2 Problem

Definition 1 ([1]). A code C with length n over a finite field \mathbb{F}_q with $q \in \mathbb{P}$ or $q = p^m, p \in \mathbb{P} \wedge m \in \mathbb{N}$ is said to be **linear** if C forms a linear subspace of \mathbb{F}_q^n .

*✉ s200376@student.dhbw-mannheim.de

3 Method

A literature review is conducted.

- *Question of Research:* Which algorithms generate *systematic* generator matrices for linear codes like Generalized Reed Solomon?
- *Derived search terms:* "Generalized Reed Solomon" \wedge ("Systematic Encoding" \vee "Generator Matrix"); "Generator Matrix" \wedge ("Standard Form" \vee "Systematic Form");
- *Databases:* IEEE Xplore (1); Google Scholar (2); JSTOR (3); KIT KV (4);
- *Scope:* 1+3: Metadata; 2+4: Full-texts
- *Criteria for Selection:* Implementability; Efficiency; Applicability;

Since a variety of different shaped and constructed generator matrices is found among linear codes in general, it seems ineffective to search for terms like "Linear codes *systematic encoding*" containing no information on specific code classes. To find general algorithms that work for several types of linear codes, the second search term was introduced.

4 Results

Using the first search term "Generalized Reed Solomon *systematic encoding*", the following results were gathered:

- BRAUCHLE points out that matrices of Generalized Reed Solomon codes are *Vandermonde* matrices. He proposes a new algorithm for the recovery of erased symbols [3]. The search term "Vandermonde matrix *systematic conversion*" will be added to this papers scope. Since [3] references [4], it will be added to the review.
- In [4] an algorithm for the *systematic encoding* of Reed-Solomon codes with arbitrary parity positions (neither pre- nor postponed to the information bits) is proposed. BRAUCHLE AND KOETTER show that for a given parity-check *Vandermonde*

matrix $H =$

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha^0 & \alpha^1 & \cdots & \alpha^{n-1} \\ \vdots & \vdots & & \vdots \\ \alpha^{0 \cdot (2t-1)} & \alpha^{1 \cdot (2t-1)} & \cdots & \alpha^{(n-1) \cdot (2t-1)} \end{pmatrix}$$

a *systematic* form can be derived by multiplying the original matrix with the inverse of the $2t \times 2t$ -submatrix B_{2t} :

$$H = [A \mid B_{2t}] \Rightarrow H_{sys} = B_{2t}^{-1} H = [P \mid I_{2t}]$$

- MATTOUSSI, ROCA, AND SAYADI compare in [5] the construction of *systematic* generator matrices using *Vandermonde* and *Hankel* matrices. Their *Hankel*-based approach has no need for matrix inversions and is therefore considered to be faster and simpler [5].

The second search term "Generalized Reed Solomon Generator Matrix" delivers one additional result:

- ROTH AND SEROUSSI proved in [6] that a $GRS_k(\alpha, v)$ code with length n has a *systematic* generator matrix $G = [I \mid A]$ where A is a $k \times (n - k)$ – Generalized Cauchy Matrix. Their proof is constructive, hence a *systematic* generator matrix can be derived from a given *non-systematic* generator matrix using the formulas in [6].

The term "Generator Matrix Standard Form" gives related procedures:

- NAKKIRAN, RASHMI, AND RAMCHANDRAN describe in [7] the "systematic remapping" operation in order to generate *systematic* codewords from *non-systematic* generator matrices: Let G be a $k \times n$ -generator matrix. G_k denotes the $k \times k$ -submatrix consisting of the first k columns of G . Since encoding is done by $c = G \cdot m$ for an information tuple m of length k , the *remapping step* and *systematic* encoding is defined by

$$\bar{m} = G_k^{-1} \cdot m \Rightarrow c_{sys} = G \cdot \bar{m}$$

Several papers described here reference [8]. HILL proposes an algorithm for transforming a generator matrix G to *standard form* using elementar operations like row/column permutation, row multiplications with scalars or adding multiples of a row to another.

Algorithm 1 Algorithm for generating a standard form matrix after [8]

Require: $G = (g_{ij})_{1 \leq i \leq k, 1 \leq j \leq n}$

```

for  $j \in \{1, \dots, k\}$  do
    if  $g_{jj} = 0$  then
        if  $i \in \{j+1, \dots, k\}: g_{ji} \neq 0$  then
             $\text{row}(g_{jj}) \leftarrow \text{row}(g_{ij})$ 
        else
             $\text{col}(g_{jj}) \leftarrow \text{col}(g_{jh} \neq 0)$ 
        end if
    end if
     $\text{row}(g_{jj}) \leftarrow \text{row}(g_{jj}) \cdot g_{jj}^{-1}$ 
    for  $i \in \{1, 2, \dots, k\}: i \neq j$  do
         $\text{row}(g_{i0}) \leftarrow \text{row}(g_{i1}) - g_{ij} \cdot \text{row}(g_{1j})$ 
    end for
end for

```

5 Conclusion

The research has shown that the default approach for *systematic* encoding of linear codes is given by *Vandermonde* matrices. For a given *non-systematic* generator matrix, the $k \times k$ -submatrix is inverted and multiplied with G to achieve the *systematic* form $G_{sys} = [I_k \mid P]$. Since matrix inversion can be an exhaustive task for large matrices, alternative approaches using *Hankel* or *Cauchy* matrices could be taken into consideration. Even row and column operations could be used algorithmically to achieve a *systematic* generator matrix, but it should be pointed out that this algorithm does not seem to be more efficient than matrix inversion since matrix inversion is performance engineered in many programming libraries.

Specific solutions for GRS_k codes were not discovered during the conducted literature review. Since this literature review was not done systematically and a comparison of results did not take place, further investigation is required to find and verify the most efficient algorithm for the underlying question.

References

- [1] W. C. Huffman and V. Pless, “Basic concepts of linear codes,” in *Fundamentals of Error-Correcting Codes*, Cambridge, UK: Cambridge University Press, 2010, ch. 1, pp. 1–52, ISBN: 9780521131704. [Online]. Available: <https://archive.org/details/fundamentalsoferror0000huff/> (visited on 01/25/2023).
- [2] F. MacWilliams and N. Sloane, “10 reed-solomon and justesen codes,” in *The Theory of Error-Correcting Codes*, ser. North-Holland Mathematical Library, vol. 16, Elsevier, 1977, pp. 294–316. DOI: [https://doi.org/10.1016/S0924-6509\(08\)70535-X](https://doi.org/10.1016/S0924-6509(08)70535-X).
- [3] J. Brauchle, “On efficient recovery of erased symbols in generalized reed-solomon codes,” in *2011 IEEE International Conference on Communications (ICC)*, 2011, pp. 1–5. DOI: [10.1109/icc.2011.5962475](https://doi.org/10.1109/icc.2011.5962475).
- [4] J. Brauchle and R. Koetter, “A systematic reed-solomon encoder with arbitrary parity positions,” in *GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference*, 2009, pp. 1–4. DOI: [10.1109/GLOCOM.2009.5426304](https://doi.org/10.1109/GLOCOM.2009.5426304).
- [5] F. Matoussi, V. Roca, and B. Sayadi, “Complexity comparison of the use of vandermonde versus hankel matrices to build systematic mds reed-solomon codes,” in *2012 IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2012, pp. 344–348. DOI: [10.1109/SPAWC.2012.6292924](https://doi.org/10.1109/SPAWC.2012.6292924).
- [6] R. M. Roth and G. Seroussi, “On generator matrices of mds codes (corresp.),” *IEEE Transactions on Information Theory*, vol. 31, no. 6, pp. 826–830, 1985. DOI: [10.1109/TIT.1985.1057113](https://doi.org/10.1109/TIT.1985.1057113).
- [7] P. Nakkiran, K. V. Rashmi, and K. Ramchandran, “Optimal systematic distributed storage codes with fast encoding,” in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 430–434. DOI: [10.1109/ISIT.2016.7541335](https://doi.org/10.1109/ISIT.2016.7541335).
- [8] R. Hill, “Introduction to linear codes,” in *A First Course in Coding Theory*, ser. Oxford Applied Mathematics and Computing Science Series, Oxford University Press, 1986, ISBN: 0-19-853804-9.

Literaturverzeichnis

- [1] B. Schneier, *Applied Cryptography: Protocols, Algorithms and Source Code in C*, 20th anniversary edition. Indianapolis: John Wiley und Sons, 2015, ISBN: 978-1-119-09672-6. Adresse: <https://www.schneier.com/books/applied-cryptography-2preface/> (besucht am 06. 11. 2022).
- [2] D. J. Bernstein, J. Buchmann und E. Dahmen, Hrsg., *Post-Quantum Cryptography*. Heidelberg: Springer, 2009, ISBN: 978-3-540-88701-0.
- [3] V. Cerf, „ASCII format for Network Interchange,“ RFC Editor, RFC 20, Okt. 1969, S. 1–56. DOI: 10.17487/RFC0020.
- [4] M. Borda, *Fundamentals in Information Theory and Coding*, 1. Aufl. Berlin: Springer, 2011. DOI: 10.1007/978-3-642-20347-3.
- [5] J. H. van Lint, *Coding Theory*, Lecture Notes in Mathematics, A. Dold und B. Eckmann, Hrsg. Berlin: Springer, 1973, ISBN: 3-540-06363-3.
- [6] N. L. Biggs, *Codes: An Introduction to Information Communication and Cryptography* (Springer Undergraduate Mathematics Series). London: Springer, 2008, DHBW-Bestand. DOI: 10.1007/978-1-84800-273-9.
- [7] W. Willems, *Codierungstheorie und Kryptographie*. Basel: Birkhäuser, 2008. DOI: 10.1007/978-3-7643-8612-2.
- [8] O. Manz, *Fehlerkorrigierende Codes*. Wiesbaden: Springer Vieweg, 2017. DOI: <https://doi.org/10.1007/978-3-658-14652-8>.
- [9] R. W. Hamming, „Error detecting and error correcting codes,“ *The Bell system technical journal*, Jg. 29, Nr. 2, S. 147–160, 1950. DOI: 10.1002/j.1538-7305.1950.tb00463.x.
- [10] A. Viterbi, „Convolutional Codes and Their Performance in Communication Systems,“ *IEEE Transactions on Communication Technology*, Jg. 19, Nr. 5, S. 751–772, 1971. DOI: 10.1109/TCOM.1971.1090700.
- [11] S. Roman, *Coding and Information Theory* (Graduate Texts in Mathematics). New York: Springer, 1992, Bd. 134, ISBN: 3-540-97812-7.

- [12] R. Lidl und H. Niederreiter, „Algebraic Foundations,“ in *Finite Fields* (Encyclopedia of Mathematics and its Applications), 2. Aufl., Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1996, S. 1–46, UB-Bestand. DOI: 10.1017/CBO9780511525926.003.
- [13] A. Myasnikov, A. Ushakov und V. Shpilrain, *Group-based Cryptography* (Advanced Courses in Mathematics - CRM Barcelona). Basel: Birkhäuser, 2008, DHBW-Bestand. DOI: 10.1007/978-3-7643-8827-0.
- [14] J. M. Howie, *Fields and Galois Theory* (Springer Undergraduate Mathematics Series). London: Springer, 2006, UB-Bestand. DOI: 10.1007/978-1-84628-181-5.
- [15] G. Fischer und B. Springborn, *Lineare Algebra, Eine Einführung für Studienanfänger* (Grundkurs Mathematik), 19. Aufl. Berlin: Springer, 2020. DOI: 10.1007/978-3-662-61645-1.
- [16] G. Teschl und S. Teschl, *Mathematik für Informatiker, Band 1: Diskrete Mathematik und Lineare Algebra* (eXamen.press), 4. Aufl. Heidelberg: Springer, 2013. DOI: 10.1007/978-3-642-37972-7.
- [17] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms*. New York: Wiley-Interscience, 2005, ISBN: 978-0-471-64800-0.
- [18] S. B. Wicker und V. K. Bhargava, „An Introduction to Reed-Solomon Codes,“ in *Reed-Solomon Codes and Their Applications*. 1994, S. 1–16. DOI: 10.1109/9780470546345.ch1.
- [19] C. Mahr, *Zur Codierung von RS-, BCH- und Goppa-Codes* (Fortschritt-Berichte 95). Düsseldorf: VDI-Verlag, 1988, Bd. 10, UB-Bestand (Magazin).
- [20] N. Zivic, *Coding and Cryptography, Synergy for a Robust Communication*. München: Oldenbourg Wissenschaftsverlag, 2013, DHBW-Bestand. DOI: 10.1524/9783486781267.
- [21] „10 Reed-Solomon and Justesen codes,“ in *The Theory of Error-Correcting Codes*, Ser. North-Holland Mathematical Library, F. MacWilliams und N. Sloane, Hrsg., Bd. 16, Elsevier, 1977, S. 294–316. DOI: [https://doi.org/10.1016/S0924-6509\(08\)70535-X](https://doi.org/10.1016/S0924-6509(08)70535-X). Adresse: <https://www.sciencedirect.com/science/article/pii/S092465090870535X>.
- [22] „SageMath - Open-Source Mathematical Software System,“ SageMath Open Source Collective. (2022), Adresse: <https://www.sagemath.org/> (besucht am 17.01.2023).
- [23] W. C. Huffman und V. Pless, *Fundamentals of Error-Correcting Codes*. Cambridge, UK: Cambridge University Press, 2010, ISBN: 9780521131704. Adresse: <https://archive.org/details/fundamentalsof0000huff/> (besucht am 25.01.2023).

- [24] R. Roth, *Introduction to Coding Theory*. Cambridge University Press, 2006. DOI: 10.1017/CBO9780511808968.
- [25] D. G. Hoffman, D. A. Leonard, C. C. Lindner, K. T. Phelps, C. Rodger und J. R. Wall, *Coding Theory: The Essentials*. New York: Marcel Dekker, Inc., 1991, KIT-Bestand, ISBN: 0-8247-8611-4.