Vorlesung Kryptographie

WiSe '23/'24, hhu
K. Halupczok

K1: Einführung, Kryptologie, Enigma

Stichworte: Begriffsklärungen, Ver- und Entschlüsseln, (a-)symmetrische Verschlüsselung, Einwegfunktionen, Anhang: die Enigma

---

1.1. **Einleitung:** Die Vorlesung "Kryptographie" behandelt Grundlagen der Kryptologie, algorithmischen Zahlentheorie, elliptische-Kurven-Arithmetik und ihre kryptographische Eignung, kryptologische Bedeutung des Quantencomputers.

Kryptologie

1.2. **Def.:** Die **Kryptologie** besteht aus den folgenden beiden Gebieten:

Kryptographie: Studium mathematischer Techniken zur Verschlüsselung von Informationen oder geheimen Nachrichten und dem Schutz von Daten.

Kryptoanalyse: Beschreibung der Rückgewinnung von Informationen aus verschlüsselten Texten, der Entschlüsselung.

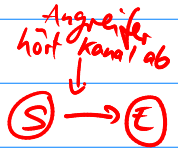
Oft meint man mit "Kryptographie" die Kryptologie (wird auch in dieser Vorlesung so sein).

1.3. **Wozu Kryptologie?**

Früher wurde die Kryptologie vor allem im militärischen oder diplomatischen Sektor verwendet, heutzutage steht in unserer vernetzten Welt vor allem auch der praktische Nutzen im Alltag im Vordergrund: im Internet einkaufen, online-Banking, persönliche Daten geheimhalten bzw. Datenschutz, Nachrichten und Dokumente digital unterschreiben etc.

Im Prinzip geht es um die Geheimhaltung von Informationen, die ausgetauscht werden sollen, vor dem Zugriff Dritter / Unbefugter.

Das Internet liefert schnelle Informationswege über "öffentliche" Kanäle, die leicht abgehört werden können, so dass die Verschlüsselung schützenswerter Daten unumgänglich wird. Auch die Möglichkeit zur Signierung/Identifikation wird nötig, weil sehr leicht Absenderangaben gefälscht werden können.
Eventuell nicht abhörsichere Kanäle können außer dem Internet aber auch Briefe, Radio, Boten, Rauchzeichen, Photonensender, etc. sein.



1.4. Bei der symmetrischen Verschlüsselung von Daten gibt es einen Sender S und einen Empfänger E, die sich beide im Vorfeld auf einen gemeinsamen Schlüssel geeinigt haben, der zum Ver- und Entschlüsseln dient. Beim Caesar-Code z.B. ist dies die Vereinbarung, jeden Buchstaben durch den 3. nachfolgenden im Alphabet zu ersetzen, also $A \mapsto D$, $B \mapsto E$, $C \mapsto F$, $D \mapsto G$, usw., die Entschlüsselung ist klar.
(Sogenannte "ROT3"-Verschlüsselung, da das Alphabet um 3 Buchstaben rotiert wird.)

1.5. Derartige monoalphabetische Chiffrierungen, bei der jeder Buchstabe des Alphabets stets durch denselben Geheimtextbuchstaben chiffriert wird, sind durch Häufigkeitsanalysen durch einen Angreifer, der die verschlüsselten Nachrichten abhört, sehr leicht zu entschlüsseln. Übrigens gibt es auch heutzutage pdf-Verschlüsselungsprogramme, die so arbeiten: mit einer ROT13-Verschlüsselung, die zur Ver- und Entschlüsselung denselben Programmcode verwendet, da die Verschlüsselungsabbildung involutorisch ist, vgl. Def. 1.16. Eine ROT-Verschlüsselung ist quasi so sicher als würde man gar keine Verschlüsselung anwenden: Durch Häufigkeitanalysen der Buchstaben im verschlüsselten Text z.B. lässt sich diese Art von Verschlüsselung leicht bestimmen, sogar dann, wenn irgendeine Permutation des Alphabets vorgenommen wurde.

1.6. In dieser Vorlesung behandeln wir die heutzutage gängigen modernen Methoden, die als sicher gelten. Worauf diese starke Sicherheit beruht, hat mathematische Gründe, die wir besprechen möchten. Vor allem interessiert uns, wie und welche Mathematik in die Kryptologie kommt.

1.7. Die Anwendungen erfordern die Lösung folgender Probleme bei symmetrischen Verschlüsselungsverfahren:

- Schlüsselaustausch über öffentliche Kanäle ("öffentliche Schlüssel")
- Verschlüsselung ohne vorherigen Schlüsselaustausch (mit "geheimen Schlüsseln", die nicht versendet werden)
- digitale Signierung / Authentifizierung

1.8. Dies können asymmetrische Verfahren leisten (auch "Public-key-Kryptographie" genannt) und gehen u.a. zurück auf Ideen von Diffie und Hellman aus den 70er Jahren:
Jeder Nutzer eines Kommunikationskanals hat einen privaten Schlüssel, den er geheim hält und niemand sonst kennt, sowie einen öffentlichen Schlüssel, den jeder einsehen kann. Eine Nachricht wird dann unter Ansetzen einer Funktion $x \mapsto f(x)$ verschlüsselt, die zwar leicht zu berechnen, aber praktisch nur mit Kenntnis des privaten Schlüssels des rechtmäßigen Empfängers entschlüsselt werden kann. Der Sender der Nachricht wird dafür den öffentlichen Schlüssel des Empfängers zur Verschlüsselung benutzen.
Eine derartige Funktion heißt Einwegfunktion / Hashfunktion.

- Beim RSA-Verfahren ist diese Fkt. die Multiplikation zweier Primzahlen $(p,q) \mapsto p \cdot q$.
- Beim ECC-Verfahren ist dies die Fkt. $x \mapsto m \cdot x$ in einer abelschen Gruppe, nämlich die Gruppe auf einer elliptischen Kurve.

1.9. In einem ersten Teil der Vorlesung stellen wir gängige Verfahren dar, die leicht mit dem Zahlring $\mathbb{Z}$ und Strukturen darin realisiert werden können. Die zentralen Fragestellungen der algorithmischen Zahlentheorie, die sich daraus ergeben, sind das Primzahltestproblem und das Faktorisierungsproblem, denen wir uns widmen werden. In einem zweiten Teil studieren wir die Eigenschaften elliptischer Kurven als interessante geometrische und arithmetische Objekte, die sich in der Praxis der Kryptographie als nützlich erwiesen haben. Wir besprechen dann auch die Sicherheit und Implementierung dieser Verfahren und vergleichen sie miteinander.

1.10. Für die Kryptographie sind elliptische Kurven (über einem Körper k) interessant, weil sich eine Verknüpfung auf ihrer Punktemenge definieren lässt, mit der diese zu einer Gruppe wird. Dabei gerade auch endliche Körper k zuzulassen, macht diese Verknüpfung auf Rechenmaschinen realisierbar; da nur wenig Rechenoperationen dafür nötig sind (s. K21), ist dies technisch vorteilhaft. Die Sicherheit der darauf beruhenden "ECC" (elliptic curve cryptography) beruht darauf, dass das Problem des diskreten Logarithmus auf einer elliptischen Kurve E, nämlich die Umkehrung der Fkt. $P \mapsto m \cdot P$ für $m \in \mathbb{N}$ fest, nach heutigem Wissensstand rechnerisch i.a. extrem schwer realisierbar ist (Bem.: $m \cdot P := \underbrace{P + \ldots + P}_{m\text{-mal}}$, wobei "+" die Gruppenverknüpfung auf der elliptischen Kurve bezeichnet.), d.h. dass diese Funktion eine Einwegfunktion im Sinne von 1.8 ist.

Anhang / Zur Einführung: Die Enigma

↗ mit rudimentären elektr. Schaltkreisen

1.11. Die Enigma ist eine mechanische Verschlüsselungsmaschine, die im 2. Weltkrieg zur Verschlüsselung des Nachrichtenverkehrs des deutschen Militärs und anderer Dienste (SS, Polizei, Reichspost, Reichsbahn...) verwendet wurde. Die Entwicklung mechanischer Verschlüsselungsmaschinen begann nach dem 1. Weltkrieg und wurde fortlaufend verbessert. Trotzdem gelang es den Alliierten im 2. Weltkrieg mit hohem Aufwand, die deutschen Funksprüche, die mit der Enigma verschlüsselt waren, zu entziffern und komplett abzuhören (Codename "ULTRA"), wovon die Deutschen nichts ahnten. Dies hatte strategische Bedeutungen für den Verlauf des 2. Weltkriegs; Historiker sind sich darin einig, dass dadurch der 2. Weltkrieg erheblich verkürzt wurde (ca. 2-4 Jahre) und so vielen Menschen auf allen Seiten des Krieges das Leben gerettet werden konnte.
Als Erfinder zählt der deutsche Elektroingenieur A. Scherbius, der 1918 das Patent zur Enigma anmeldete. Ab 1934 wurden die Maschinen in großen Stil hergestellt, vermutlich wurden ca. 100.000 Stück eingesetzt. Zur damaligen Zeit galt die Enigma als "unknackbar", d.h. kryptographisch sicher.

1.12. Aufbau und Prinzip

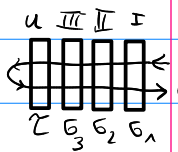
Idee: die Maschine verwendet drei bewegliche Walzen und eine feste unbewegliche Walze, die als "Umkehrwalze" dient. Durch die Walzen sind Verdrahtungen gelegt, durch die Strom fließen kann. Jede Verdrahtung verbindet einen Buchstaben des Alphabets auf einer Seite mit einem der anderen Seite der Walze. Diese werden bei der Verschlüsselung eins weiterbewegt, die Drahtenden der Walzen haben bei Stillstand Kontakt.

Unser Wissen zu Permutationen aus Grundvorlesungen:

1.13. **Def.:** Bijektive Abbildungen $\sigma: \{1,\dots,n\} \to \{1,\dots,n\}$ heißen Permutationen, man bezeichnet mit $S_n := \{\sigma;\ \sigma: \{1,\dots,n\} \to \{1,\dots,n\}\}$ die Menge der Permutationen von n Elementen.
Mit der Verknüpfung $\circ: S_n \times S_n \to S_n$
$(\sigma, \tau) \mapsto \sigma\circ\tau$, wo $\sigma\circ\tau(i) := \sigma(\tau(i))$,
wird $(S_n, \circ)$ zu einer (nichtabelschen) Gruppe,
der symmetrischen Gruppe von n Elementen.

Konvention ist, das Verknüpfungszeichen $\circ$ in Produkten darin wegzulassen, d.h. z.B. $\sigma\circ\tau\circ\sigma^{-1} = \sigma\tau\sigma^{-1}$ zu schreiben; wie beim Malpunkt, den man weglässt. Wir schreiben $(S_n, \circ)$ also multiplikativ. Man "liest" ein Produkt $\sigma\tau\eta$ von "rechts nach links": erst wird η angewendet, dann τ, dann σ. Man beachte die Reihenfolge beim Invertieren, z.B. ist $(\sigma\tau\eta)^{-1} = \eta^{-1}\tau^{-1}\sigma^{-1}$.

1.14. **Bem.:** Wir identifizieren $\mathbb{Z}_{26} = \{0,1,\dots,25\}$ mit dem Alphabet $\{A,B,\dots,Z\}$.
Die Vertauschung von Buchstaben entspricht einer Permutation aus S_{26}.
Eine Walze der Enigma ist ca. 10 cm im Durchmesser groß und hat auf beiden Seiten das Alphabet $A,\dots,Z$ stehen, wobei die Buchstaben links und rechts jeweils fest miteinander verdrahtet sind. Dieser inneren Verdrahtung entspricht also einer festen Permutation $\sigma \in S_{26}$.



Es werden drei austauschbare Walzen (mit Permutationen $\sigma_1, \sigma_2, \sigma_3 \in S_{26}$) verwendet. Strom kann von rechts nach links durch diese Drähte fließen, kehrt an einer Umkehrwalze ($\tau \in S_{26}$) um, und wieder zurück von links nach rechts.
Die Maschine hat ein Tastenfeld mit Buchstaben A–Z. Durch Druck einer Taste bewegen sich die Walzen – ähnlich wie bei einem Kilometerzähler – um einen Schritt weiter, lässt Strom durch den Walzenweg des gedrückten Buchstabens fließen und schließlich eine Anzeigelampe aufleuchten, die den kodierten (=verschlüsselten) Buchstaben anzeigt.

1.15. Eine Walze mit Verdrahtung $\sigma \in S_{26}$ erzeugt so beim i-ten Tastendruck die Permutation $\sigma(k(i)) = \eta^{-k(i)} \sigma \eta^{k(i)}$, wobei $\eta : \mathbb{Z}_{26} \to \mathbb{Z}_{26}$ die Verschiebung um Eins, $\eta(i) = i+1$ bezeichnet, und $k(i) \equiv i \bmod 26$ (der Rest von $i \bmod 26$) die Walzenstellung. Ein Eingangstext mit den Buchstaben $p_1, p_2, p_3, \dots$ wird von einer Walze mit $E(p_i) = \eta^{-k(i)} \sigma \eta^{k(i)}(p_i)$ verschlüsselt zu $E(p_1), E(p_2), E(p_3), \dots$, wobei $k(i) \equiv i \bmod 26$ mit $k(i) \in \mathbb{Z}_{26}$ die aktuelle Walzenstellung bezeichnet. Es werden drei solcher Walzen verwendet, sowie eine Umkehrwalze mit einer Permutation $\tau \in S_n$, die eine Involution darstellt: sie besitzt nur ein Alphabet, und je zwei verschiedene Buchstaben davon sind verdrahtet.

1.16. Def.: Ein $\tau \in S_n$ heißt Involution, falls $\tau^2 = \tau \circ \tau = id$ ist, d.h. es werden immer zwei Buchstaben vertauscht: $\tau(\tau(i)) = i$ für $i,j \in \{1,\dots,n\}$, $i \neq j$, oder ein Buchstabe wird auf sich selbst abgebildet: $\tau(\tau(i)) = i$ für $i \in \{1,\dots,n\}$. z.B. ist $\tau(1)=2$, $\tau(2)=1$, $\tau(3)=4$, $\tau(4)=3$ eine Involution in S_4. Es gibt eine elegante Kurzschreibweise dafür: $\tau = (1\,2)(3\,4)$, um die Vertauschungen (= Transpositionen) $1 \leftrightarrow 2$, $3 \leftrightarrow 4$ zu beschreiben. Entsprechend würde $\tau = (a_1\,a_2)(a_3\,a_4)\cdots(a_{25}\,a_{26}) \in S_{26}$ die fixpunktfreie Involution bezeichnen, welche die 13 Vertauschungen $a_1 \leftrightarrow a_2$ usw. vornimmt.

1.17. Nun werden drei Walzen verwendet, welche 26^3 Walzenstellungen möglich machen und die sich wie bei einem Kilometerzähler bewegen: bei jedem Tastendruck dreht sich Walze 1 um einen Schritt weiter, nach 26 Schritten von Walze 1 dreht sich Walze 2 um einen Schritt weiter, nach 26 Schritten von Walze 2 dreht sich Walze 3 um einen Schritt weiter. Sagen wir, für jede Walze 1,2,3 erhalten wir jeweils die Walzenstellung $k_1(i), k_2(i), k_3(i) \in \mathbb{Z}_{26}$ nach dem i-ten Schritt, bei der der Buchstabe p_i verschlüsselt werden soll.

1.18. Wir setzen nun $k := k_3(i) + k_2(i)\cdot 26 + k_1(i)\cdot 26^2 \equiv i \bmod 26^3$ mit $0 \le k_j(i) \le 25$, d.h. $k_3(i), k_2(i), k_1(i)$ ist die 26-adische Darstellung der Schrittzahl $i = 1, 2, 3, \ldots$ Damit wird die Stellung der 3 Walzen beschrieben.

1.19. Setze $\sigma[i] := \eta^{-k_3(i)}\sigma_3\eta^{k_3(i)}\,\eta^{-k_2(i)}\sigma_2\eta^{k_2(i)}\,\eta^{-k_1(i)}\sigma_1\eta^{k_1(i)}$,

diese Permutation bewirkt die Verschlüsselung von rechts nach links durch die Walzen 1, 2 und 3, der Umkehrweg ist dann

$$\sigma[i]^{-1} = \eta^{-k_1(i)}\sigma_1^{-1}\eta^{k_1(i)}\,\eta^{-k_2(i)}\sigma_2^{-1}\eta^{k_2(i)}\,\eta^{-k_3(i)}\sigma_3^{-1}\eta^{k_3(i)}.$$

Der Buchstabe p_i wird somit zu $E(p_i) = \sigma[i]^{-1}\tau\,\sigma[i](p_i)$ verschlüsselt.

Schlüssel:
- Auswahl der Walzen (3 aus 5 oder 4 möglichen)
- Anfangsstellung der Walzen (z.B. Anzeige MCK)
- zusätzliches Steckbrett (= stationäre Walze mit frei wählbarer Verdrahtung $\rightsquigarrow$ "Eingangspermutation")
- durch Verdrehen eines Ringes konnte auch i zu $i+I$ ersetzt werden für ein festes I

Probleme eines Kryptoanalytikers, der die verschlüsselten Texte empfängt (hier der Alliierten, die verschlüsselte Funksprüche empfangen):
Bestimmen der inneren Verdrahtung bzw. der verwendeten Schlüssel, falls zufällig ein Klartext (oder Teile davon) samt Verschlüsselung vorliegt – eine sogenannte "chosen plaintext – Attacke".
Zuerst wurde eine Enigma von polnischen Mathematikern untersucht und ihre Verdrahtung bestimmt, war den Alliierten also bekannt.
$\rightsquigarrow \sigma_1, \sigma_2, \sigma_3, \tau$ ($\rightsquigarrow$ Marian Rejewski).

1.20. Es mussten dann die verwendeten Schlüssel bestimmt werden. Allerdings wurden täglich wechselnde Walzeneinstellungen verwendet, die durch eine geheime Schlüsseltabelle vorgegeben war ("Tagesschlüssel"); der Wechsel war immer genau um Mitternacht. Zum Zweck der Schlüsselermittlung bauten polnische Mathematiker Maschinen, genannt Bomba, welche z.T. erfolgreich waren. Während des Krieges wurde in Bletchley Park nahe London ein Angriff auf die Enigma-Funksprüche gestartet mit enormen Einsatz – über 10.000 Frauen und Männer wurden beschäftigt. Von Alan Turing wurde dabei eine spezielle elektromagnetische Maschine ("Turing-Bombe" genannt) zur Schlüsselsuche eingesetzt. Dabei spielte die Suche von wahrscheinlichen Wörtern im Text eine große Rolle, wie z.B. OBERKOMMANDODERWEHRMACHT, womit der mögliche Schlüsselraum erheblich eingeschränkt wurde.

↝ Filmtipp: "The Imitation Game"

1.21. Die Verschlüsselung mit $E(p_i) = \sigma[i]^{-1} \circ \tau \circ \sigma[i](p_i)$ hatte den Vorteil, involutorisch zu sein, d.h.

Entschlüsseln = Verschlüsseln, da

$$E(E(p_i)) = \sigma[i]^{-1} \circ \tau \underbrace{\sigma[i] \circ \sigma[i]^{-1}}_{=id} \circ \tau \circ \sigma[i](p_i) = id(p_i) = p_i \text{ ist,}$$

was praktisch ist,

Für die kryptographische Sicherheit war diese Eigenschaft jedoch nachteilig, ebenso die Tatsache, dass die Verschlüsselung fixpunktfrei ist (d.h. $E(p_i) \neq p_i$, weil sonst $\sigma[i]^{-1} \circ \tau \circ \sigma[i](p_i) = p_i$

$\Leftrightarrow \tau \circ \sigma[i](p_i) = \sigma[i](p_i) \Leftrightarrow \tau(b) = b$ für $b = \sigma[i](p_i)$ wäre;

die Umkehrung erfüllt aber laut Konstruktion eine fixpunktfreie Permutation τ, nämlich eine Involution wie oben erläutert.)

Von den Deutschen wurde diese Eigenschaft fälschlicherweise für einen Vorteil gehalten, was aber nicht stimmt:
Von den $26! \approx 4 \cdot 10^{26}$ Permutationen aus S_{26} gibt es nur höchstens $25 \cdot 23 \cdot 21 \cdot \ldots \cdot 5 \cdot 3 \cdot 1 \approx 8 \cdot 10^{12}$ viele, die von einer Enigma realisiert werden können. Der Anteil beträgt also nur $\approx \frac{8 \cdot 10^{12}}{4 \cdot 10^{26}} = 2 \cdot 10^{-14} \approx 0.\underbrace{0\ldots02}_{14 \text{ Stellen}}$.

Herleitung:
- Stirlingsche Formel: $n! \sim \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n$, also
$$26! \approx \sqrt{2\pi \cdot 26} \cdot \left(\frac{26}{e}\right)^{26} \overset{TR}{=} 4.03 \cdot 10^{26}$$
- #Möglichkeiten für Involution $(a_1\ a_2)(a_3\ a_4)\cdots(a_{25}\ a_{26})$:
  Setze a_1 auf A, für a_2 gibt es dann noch 25 Möglichkeiten.
  Setze a_3 auf B oder C (den nächsten noch nicht vergebenen Buchstaben),
  für a_4 gibt es dann noch 23 Möglichkeiten, ...
  Symbolisch: $(A\ \underbrace{\ \ }_{25})\ (\underbrace{B \text{ od. } C}_{2}\ \underbrace{\ \ }_{23})\ (\underbrace{C \text{ od. } D/E}_{2}\ \underbrace{\ \ }_{21}) \cdots (\underbrace{\text{letzte freie}}_{1}\ \underbrace{\ \ }_{1})$
  nächster Buchstabe
  Dies macht insg. $25 \cdot 23 \cdot 21 \cdots 5 \cdot 3 \cdot 1$ viele Möglichkeiten, eine Involution zu definieren. Dies ist also deren Anzahl.

Haben weiter
$$25 \cdot 23 \cdot 21 \cdots 5 \cdot 3 \cdot 1 = \frac{25!}{2 \cdot 4 \cdot 6 \cdot \ldots \cdot 24} = \frac{25!}{2^{12} \cdot 12!} \overset{\text{TR/Stirling}}{\approx} \frac{1.55 \cdot 10^{25}}{4096 \cdot 4.79 \cdot 10^{8}} = 7.9 \cdot 10^{12}$$

Aufgabe: Gegeben ist eine "Mini-Enigma" mit dem Alphabet $\Sigma = \{A, E, H, N\}$ mit zwei Walzen und einer Umkehrwalze.
Bekannt ist, dass der Klartext AHNEHANNAHNAHEANNE verschlüsselt wurde zu HAENAHEEHAEHANHEEN
- Mit welcher inneren Verdrahtung arbeitet die Enigma?
- Welchen Geheimtext wird (bei gleicher Anfangsstellung) dem Klartext NAEHEANNA zugeordnet?