Elliptische Kurven in der Charakteristik p > 3 und die Implementierung der Arithmetik in der Programmiersprache Python

Studienarbeit T3_3101

Hochschule:	Duale Hochschule Baden-Württemberg Mannheim
Kurs:	TINF20IT2
Name:	Vorname Nachname
Matrikelnummer:	XXXXXX
E-Mail:	sXXXXXX@student.dhbw-mannheim.de

Studiengangsleiter:	Prof. Dr. Nathan Sudermann-Merx
Betreuer:	Prof. Dr. Reinhold Hübl
Bearbeitungszeitraum:	18.10.2022 - XX.XX.2023

Unterschrift des Betreuers: _____

Selbstständigkeitserklärung

Hiermit erkläre ich durch meine Unterschrift, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und keine anderen Hilfsmittel als die angegebenen verwendet habe.

Insbesondere versichere ich, dass ich alle wörtlichen und sinngemäßen Übernahmen aus anderen Werken – dazu gehören auch Internetquellen – als solche kenntlich gemacht habe.

Ort, Datum	Unterschrift Student

Zusammenfassung

Hier Text des Abstract in Deutsch.

Abstract

Hier Text des Abstract in Englisch.

Inhaltsverzeichnis

Abkürzungsverzeichnis

Abbildungsverzeichnis

1. Grundlagen

Diese Studienarbeit befasst sich mit dem komplexen Thema der Elliptischen Kurven in der Kryptographie. Die Kryptographie ist ein mathematisches Thema, bei welchem es zu Anfang der Legung einer Grundlage für das Verständnis der Inhalte dieser Studienarbeit bedarf. In diesem Kapitel werden sowohl die mathematischen als auch die kryptographischen Grundlagen zum Verständnis der Inhalte dieser Studienarbeit gelegt.

1.1. Primzahlen

In der Zahlentheorie, einem Teilbereich der Mathematik, werden viele unterschiedliche Eigenschaften von Zahlen untersucht. Durch die Untersuchung erhofft man sich neue Erkenntnisse für Wissenschaft und Technik. Die Primzahlen als mathematisches Forschungsgebiet sind hierbei ein Teilbereich der Zahlentheorie. Im Folgenden werden Primzahlen definiert und deren Eigenschaften erläutert. Anschließend wird untersucht, wie Primzahlen berechnet werden können. Am Ende wird erläutert, welche Rolle Primzahlen in der Kryptologie und modernen Kryptosystemen innehaben.

1.1.1. Definition und Eigenschaften

Es gibt viele unterschiedliche Zahlenmengen. Beispielsweise gibt es die Menge der reellen Zahlen $\mathbb{R}$. Diese beinhalten als Teilmenge die rationalen und die irrationalen Zahlen. Die natürlichen Zahlen $\mathbb{N}$ bilden hierbei alle positiven ganzen Zahlen ab. Dabei gibt es $\mathbb{N}^+$ exklusive der Zahl 0 als Teilmenge mit

$$\mathbb{N}^+ = \{1, 2, 3, 4, 5, ...\}$$

und $\mathbb{N}_0$ inklusive der Zahl 0 als Teilmenge mit

$$\mathbb{N}_0 = \{1, 2, 3, 4, 5, ...\}.$$

Die Primzahlen $\mathbb{P}$ sind hierbei etwas ganz besonderes. Sie unterscheiden sich von anderen Zahlen. Sie sind eine Teilmenge der natürlichen Zahlen und die Kardinalität ihrer Elemente ist unendlich respektive die Anzahl der Primzahlen ist unendlich. Die Unendlichkeit der Primzahlen konnte schon mit mehreren mathematischen Sätzen bewiesen werden, unter anderem dem Satz von Euklid. Auf die unendlichkeit der

Primzahlen sowie deren bestimmung wird später in 1.1.2 eingegangen.

Doch wie genau sind Primzahlen definiert? Dafür muss erst geklärt werden, was zusammengesetzte Zahlen sind. Dadurch können die Primzahlen klarer von anderen natürlichen Zahlen abgegrenzt werden. Eine natürliche Zahl mit $n \geq 2$ ist eine zusammengesetzte Zahl, falls es zwei natürliche Zahlen m und k mit den Eigenschaften:

$$m, k \geq 2 \text{ oder } m, k \neq n, \text{ für die gilt: } m \cdot k = n.$$

Zusammengesetzte natürliche Zahlen können also immer als Produkt zweier natürlicher Zahlen ≥ 2 beschrieben werden. Primzahlen bilden hierzu das Gegenstück. Eine Primzahl p ist eine natürliche Zahl mit $p \geq 1$, wobei p nur durch 1 und sich selbst teilbar sein darf. Durch diese Eigenschaft sind Primzahlen nicht zusammengesetzt. Sie können nicht als Produkt von zusammengesetzten natürlichen Zahlen gebildet werden. Man nehme als Beispiel die Primzahl 7. Sie lässt sich nicht als Produkt von natürlichen Zahlen darstellen. Als Gegenbeispiel nimmt man die zusammengesetzte natürliche Zahl 28. Sie kann durch Multiplikation aus den Zahlen 2 und 14 gebildet werden:

$$2 \cdot 14 = 28.$$

Eine weitere Eigenschaft von Primzahlen ist, dass sie das Grundgerüst zur Bildung von Zahlen sind, da man aus ihnen alle natürlichen Zahlen bilden kann. Eine zusammengesetzte natürliche Zahl n mit $n \geq 2$ kann wie bereits beschrieben immer als Produkt von mindestens zwei weiteren natürlichen Zahlen dargestellt werden. Die einzelnen Faktoren dieses Produktes heißen Primfaktoren. Die Zerlegung einer zusammengesetzten natürlichen Zahl in ihre Primfaktoren nennt man Primfaktorzerlegung. Dadurch ist die Zahl als Produkt von mehreren Primzahlen dargestellt. Nehmen wir als Beispiel die Zahl 28. Im vorigen Absatz stellten wir diese zusammengesetzte natürliche Zahl durch die Multiplikation von 2 und 14 dar. Die Zahl 2 ist eine Primzahl. Die Zahl 14 ist noch nicht in ihre Primzahlfaktoren zerlegt. Sie lässt sich als folgendes Produkt darstellen:

$$2 \cdot 7 = 14.$$

Da 7 auch eine Primzahl ist, wurden alle Primfaktoren gefunden. Die Zahl 28 lässt sich in ihrer Primfaktorzerlegung also wie folgt darstellen:

$$2 \cdot 2 \cdot 7 = 28.$$

Die Mehrfachheit von Primzahlen lässt sich auch als Potenz schreiben. Somit wird

daraus

$$2^2 \cdot 7 = 28.$$

Der Vorteil durch die Potenzen zeigt sich besonders bei großen Zahlen, da diese oft eine große Anzahl an Primfaktoren haben können. Nimmt man als Beispiel die Zahl 5281250000. Diese setzt sich mit ihren Primfaktoren wie folgt zusammen:

$$2 \cdot 2 \cdot 2 \cdot 2 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 13 \cdot 13 = 5281250000.$$

Man erkennt rasch, dass sich die Primfaktoren mit der Potenzschreibweise zusammenfassen lassen und man so die Primfaktorzerlegung wie folgt darstellen kann:

$$2^4 \cdot 5^9 \cdot 13^2 = 5281250000.$$

Die Vorteile der Potenzschreibweise liegt hier auf der Hand, da man erheblich Zeit beim Aufschreiben und Platz auf dem Papier spart.

1.1.2. Bestimmung von Primzahlen

Nachdem die grundlegenen Eigenschaften der Primzahlen angeführt wurden, muss auf die Bestimmung von Primzahlen eingegangen werden. Paulo Ribenboim geht in seinem Buch „*Die Welt der Primzahlen: Geheimnisse und Rekorde*“der Frage auf den Grund, ob primzahldefinierende Funktionen existieren. An einer Stelle des Buches geht er auf diese möglichen Funktionen und ihre Eigenschaften ein [**Ribenboim.2011**]. Solch eine Funktion müsse laut ihm eine der folgenden drei Eigenschaften aufweisen, damit man sie zur Bestimmung von Primzahlen nutzen könne:

(a) $f(n) = p_n$ (die n-te Primzahl) für alle n ≥ 1;

(b) $f(n)$ ist immer prim und wenn n $\neq$ m, dann gilt: $f(n) \neq f(m)$;

(c) der positive Wertebereich der Funktion ist identisch mit der Menge der Primzahlen

Ribenboin erklärt, dass die Bedingung, um (a) zu erfüllen schärfer sei als (b) und als (c). Die bisher erzielten Resultate zur Findung einer Formel zur Bestimmung von Primzahlen seien außerdem eher enttäuschend. Doch wenn die Funktionen zur Bestimmung von Primzahlen bisher enttäuschend waren, wie wurden diese bisher bestimmt?

Eine der simpelsten und sicher auch eine der ältesten Methoden ist das „Sieb des Eratosthenes“. Der Übersetzer Kai Brodersen beschreibt in seiner Übersetzung aus dem Jahre 2021 eines Buches aus dem Griechischen von Nikomachos von Gerasa, wie dieser sehr simpel die Funktionsweise des Siebes erläuterte [**Nikomachos+2021+7+7**]. Die Richtigkeit dieses Verfahrens wurde von Nikomachos im frühen 2. Jh. n. Chr. belegt. Bei dem Verfahren schreibt man alle natürlichen Zahlen von 2 bis zu einer gewählten Zahl n in eine Liste. Um die Primzahlen zu erhalten, siebt man jetzt die zusammengesetzten natürlichen Zahlen aus, indem man Vielfache streicht. Man beginnt bei der kleinsten Zahl, der 2. Man schreitet in der Liste fort und streicht alle Vielfachen der 2 bis zur höchsten gewählten Zahl n durch. Anschließend beginnt man mit der nächstgrößeren Zahl, welche nicht durchgestrichen ist respektive ausgesiebt wurde und streicht von dieser ebenfalls alle Vielfachen bis zur höchsten Zahl n durch. Den simplen Algorithmus führt man nun solange fort, bis man keine Vielfachen mehr streichen kann. Die übriggebliebenen Zahlen sind die Reihe der Primzahlen bis n. Die Darstellung in einer Tabelle ist heutzutage geläufig, da dies übersichtlicher ist. In der folgenden Tabelle wurde der Algorithmus des Siebes des Eratosthenes von 2 bis 100 angewandt:

1	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	~~49~~	~~50~~
~~51~~	~~52~~	53	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	59	~~60~~
61	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	67	~~68~~	~~69~~	~~70~~
71	~~72~~	73	~~74~~	~~75~~	~~76~~	~~77~~	~~78~~	79	~~80~~
~~81~~	~~82~~	83	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	89	~~90~~
~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	97	~~98~~	~~99~~	~~100~~

Das Sieb des Eratosthenes ist eine Möglichkeit, Primzahlen genau zu bestimmen. Problematisch wird es jedoch bei großen Zahlen. Für jede Zahl müssen je alle anderen Zahlen durchgegangen werden und es muss eine Teilbarkeitsprüfung durchgeführt werden. Umso größer die Zahlen werden, umso rechenaufwendiger wird die Anwendung des Sieb des Eratosthenes, um Primzahlen zu finden. Dieses ist also kein optimaler Ansatz, große Primzahlen zu bestimmen. Neben dem Sieb des Eratosthenes gibt es viele weitere Methoden, Primzahlen zu bestimmen.

Die Besprechung aller dieser Verfahren und Algorithmen soll nicht Thema dieser Arbeit sein, jedoch soll noch eine gängige Methode erläutert werden. Über Probabilistische Verfahrne kö
WEITERSCHREIBEN

1.1.3. Rolle der Primzahlen in der Kryptografie

Gemäß dem Buch „*Kryptographie: Grundlagen, Algorithmen, Protokolle*“von Dietmar Wätjen aus dem Jahr 2018 beschreibt den Begriff der Kryptografie [**Watjen.2018**]. Diese ist die Wissenschaft vom geheimen Schreiben. Kernziel ist es dabei, einen unverschlüsselten Text, genannt Klartext in einen Chiffretext überführt. Dieser Vorgang heißt *chiffrieren*. Der Vorgang, bei welchem der Chiffretext wieder in den Klartext überführt wird, heißt *dechiffrieren*. Für beide Vorgänge werden Schlüssel notwendig. Die Abbildung 1.1 stellt dies übersichtlich dar:

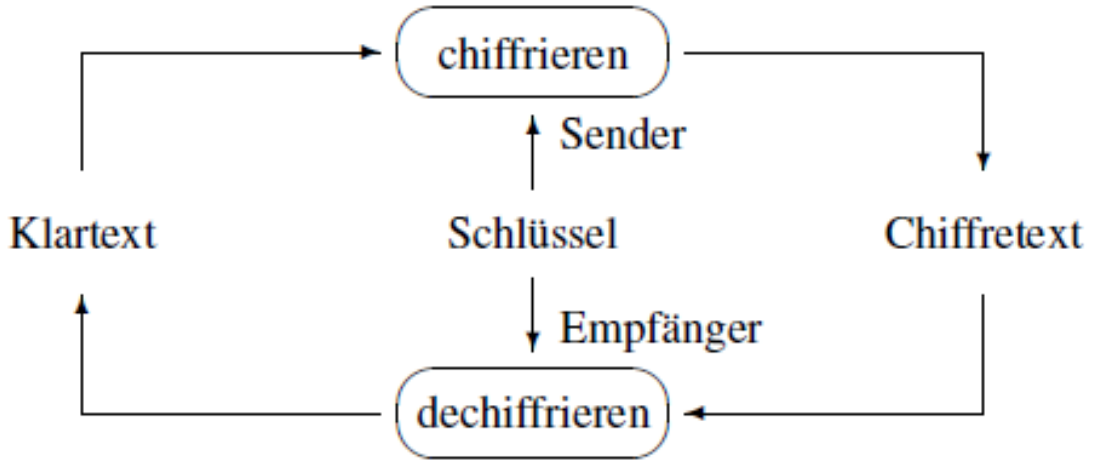

Abbildung 1.1.: Kryptografische Verschlüsselung
Quelle: [**Watjen.2018**]

Wätjen beshreibt Kryptografische System, auch Kryptosystem genannt, als ein System aus fünf Komponenten:

1. Klartextraum M
2. Chiffretextraum C
3. Schlüsselraum K
4. Familie von Chiffriertransformationen $E_K : M \to C$ mit $K \in K$

5. Familie von Dehiffriertransformationen $D_K : C \to M$ mit $K \in K$

Laut Watjen sind M, C und K höchstens, abzählbare Mengen. Eine Chiffriertransformation E_K wird durch einen Schlüssel K und einen Chiffrieralgorithmus E definiert, welcher für jede Familie gleich ist. Eine Dechiffriertransformation D_K wird ebenfalls durch einen Schhlüssel K durchgeführt. Desweiteren sollen die Kryptografischen Systeme nach Watjen die folgenden drei Eigenschaften aufweisen:

(1) Klartextraum M

(2) Chiffretextraum C

(3) Schlüsselraum K

Laut
Durch die Kryptografie können dadurch Übertragungen von sensiblen Informationen zum einen sicherer als auch privater ablaufen, da ein abgefangenes Chiffrat nicht direkt lesbar ist.
Primzahlen sind aufgrund ihrer hervorragenden Eigenschaften für die Kryptologie nützlich. Viele Verfahren, welche den Klartext in einen Chiffretext überführen, benötigen für ihren Algorithmus Primzahlen. Ein gutes Beispiel ist der Diffie-Hellmann-Schlüsselaustausch
hier diffie-hellmann

1.2. Algebraische Strukturen

Definiert durch die Zahlentheorie und als zentraler Untersuchungsgegenstand des mathematischen Teilgebietes der universellen Algebra, liefern algebraische Strukturen die Basis zur Realisierung komplexer symmetrischer und asymmetrischer Kryptosysteme, weshalb wir im folgenden Kapitel die Eigenschaften relevanter algebraischer Strukturen näher betrachten wollen. Darüber hinaus möchten wir Ihnen auch einige Werkzeuge zum Rechnen in der jeweiligen algebraischen Struktur an die Hand geben, welche zur späteren Realisierung von Kryptosystemen benötigt werden.

Unter einer sehr allgemeinen Betrachtung ist eine mathematische Struktur eine Liste nichtleerer Mengen, genannt Trägermengen, mit Elementen aus den Trägermengen, genannt Konstanten, und mengentheoretischer Konstruktionen über den Trägermengen. Diese sind konkret Funktionen über den Trägermengen. Im Weiteren beschränken wir uns auf den Fall einer einzigen Trägermenge, wodurch die Strukturen als homogen bezeichnet werden können.

Definition: Homogene algebraische Struktur Eine homogene algebraische Struktur ist ein Tupel $(M, c_1, \ldots, c_m, f_1, \ldots, f_n)$ mit $m, n \in \mathbb{N}$ und $n \geq 1$. Dabei ist M eine nichtleere Menge, genannt **Trägermenge**, alle c_i sind Elemente aus M, genannt die **Konstanten**, und alle f_i sind s_i-stellige Funktionen $f_i : M \to M$ im Fall $s = 1$ und $f_i : M^{s_i} \to M$ im Fall $s_i > 1$, genannt die (inneren) **Operationen**. Die lineare Liste$(0, \ldots, 0, s_1, \ldots, s_n)$ mit m Nullen heißt **Typ** oder die **Signatur**.

Laut dieser Definition muss eine homogen algebraische Struktur nicht unbedingt Konstanten enthalten, jedoch mindestens eine Operation. Das Paar $(\mathbb{N}, +)$ bildet beispielsweise eine homogene algebraische Struktur des Typs (2). Das 5-Tupel $(\mathbb{N}, 0, 1, +, \cdot)$ bildet ebenfalls eine homogen algebraische Struktur des Typs (0,0,2,2).
Algebraische Strukturen unterscheiden sich grundsätzlich durch ihren Typ. Wirklich charakterisiert werden sie aber erst durch die jeweils geltenden Axiome, d.h. bestimmte Eigenschaften, welche für die Konstanten und Operationen gefordert werden. Durch die Hinzunahme immer weiterer Axiome, entsteht eine Hierarchie immer feinerer Strukturen, an deren Anfang der Monoid steht.

1.2.1. Monoid

Definition: Monoid Eine algebraische Struktur $(M, e, \cdot)$ des Typs (0,2) heißt ein Monoid, falls für alle $x, y, z \in M$ die folgenden Monoid Axiome gelten:

- (Ass) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
- (Neu) $e \cdot x = x = x \cdot e$

Gilt zusätzlich noch für alle $x, y \in M$ die Gleichung $x \cdot y = y \cdot x$, so heiß $(M, e, \cdot)$ ein **kommutatives Monoid**.

Die erste und die letzte Gleichung bilden das Assoziativ- und Kommutativgesetz ab. Durch die mittlere Gleichung wird ein neutrales Element e bezüglich der Operation gefordert, wobei sowohl die **Linksneutralität** als auch die **Rechtsneutralität** spezifiziert wird.

Einfache Beispiele für Monoide sind $(\mathbb{N}, 0, +)$, $(\mathbb{N}, 1, \cdot)$ und $(\mathbb{Z}, 0, +)$. Die Potenzierung in solchen Monoiden ist folgendermaßen definiert.

Definition: Potenzierung In einem Monoid $(M, e, \cdot)$ definiert man die n-te **Potenz** x^n von $x \in M$ durch $x^0 := e$ und $x^{x+1} = x \cdot x^n$ für alle $n \in \mathbb{N}$.

Daraus ergibt sich für den Monoid $(\mathbb{N}, 1, \cdot)$ die aus $\mathbb{R}$ gewohnte Potenzierung. Nach welcher für ein $x \in \mathbb{N}$ die Potenzierung $x^n = x_1 \cdot x_2 \cdot \cdots \cdot x_n$ ergibt. Betrachtet man jedoch den Monoid $(\mathbb{N}, 0, +)$, so ergibt analog dazu für ein $x \in \mathbb{N}$ die Potenzierung $x^n = x_1 + x_2 + \cdots + x_n = x \cdot n$, was also einer Multiplikation von x mit n entspricht.

1.2.2. Gruppe

Definition: Gruppe Eine algebraische Struktur $(G, e, \cdot, inv)$ des Typs $(0, 2, 1)$ heißt **Gruppe**, falls für alle $x, y, z \in G$ die folgenden Axiome gelten:

- (Ass) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
- (Neu) $e \cdot x = x$
- (Inv) $inv(x) \cdot x = e$

Gilt wiederum die Gleichung $x \cdot y = y \cdot x$ für alle $x, y \in G$, so heißt $(G, e, \cdot, inv)$ eine **kommutative Gruppe** oder Abelsche Gruppe.
In jeder Gruppe $(G, e, \cdot, inv)$ gelten für alle $x \in G$ folgende Formeln:

- $x \cdot x = x \Rightarrow x = e$
- $x \cdot e = x$
- $x \cdot inv(x) = e$
- $(\forall z \in G : x \cdot z = z) \Rightarrow x = e$
- $x \cdot y = e \Rightarrow x = inv(y)$
- $inv(x \cdot y) = inv(x) \cdot inv(y)$
- $inv(inv(x)) = x$
- $inv(e) = e$

1.2.3. Ring

Ein **Ring** ist eine algebraische Struktur $(R, 0, 1, +, \cdot, -)$ des Typs $(0, 0, 2, 2, 1)$ mit den folgenden Eigenschaften:

1. Es ist $(R, 0, +, -)$ eine kommutative Gruppe
2. Es ist $(R, 1, \cdot)$ ein Monoid.

3. Für alle $x, y, z \in R$ gelten die Distributivgesetze $x(y+z) = xy$ und $(y+z)x = yx + zx$

Ist $(R, 1, \cdot)$ ein kommutatives Monoid, so nennt man $(R, 0, 1, +, \cdot, -)$ einen kommutativen Ring.

1.2.4. Körper

1.3. Allgemeines zur Verschlüsselung

XXX

1.3.1. Symmetrische und Asymmetrische Verschlüsselung

x

1.3.2. Diffie-Hellmann

x

1.4. Ziel der Arbeit

XXX

1.5. Geplante Vorgehensweise

XXX

2. Elliptische Kurven

Als Basis für Asymentrische Kryptosysteme können elliptische Kurven dazu genutzt werden, die verschlüsselungstechnische Effektivität mathematischer Probleme, wie das des diskreten Logarithmus, zu erhöhen. Bei der Kryptographie unter Verwendung elliptsicher Kurven bei deutlich kürzerer Schlüssellänge ein gleichwertigers Ergebnis erzielt werden. Dieser Effekt wird durch die spezielle Arithmetik auf elliptischen Kurven erzielt, deren mathematische Grundlage, konkrete Eingenschaften und Funktionsweise im folgenden Kapitel erörtert werden soll.

Elliptische Kurven können über beliebigen Körpern definiert werden. Für die Kryptographie interessant sind elliptische Kurven über Primkörpern.
Um das weitere Verständnis zu verbessern, wollen wir erst eine uns schon bekannte Kurve ansehen. In Abbildung XY ist das Polynom $x^2 + y^2 = r^2$ über $\mathbb{R}$ dargestellt. Wie zu sehen ist, handelt es sich hierbei um die Kreisgleichung. Der zu sehende Kreis ist nichts anderes als die Menge aller Punkte, welche die Kreisgleichung erfüllen. Ein Beispiel für eine solchen ist der Punkt $(r, 0)$. Wenn x den Wert r hat, muss y folglich den Wert 0 haben. Ein Gegenbeispiel ist der Punkt $(r, r/2)$. Dieser erfüllt die Kreisgleichung nicht. Die Kreisgleichung kann verallgemeinert werden, indem

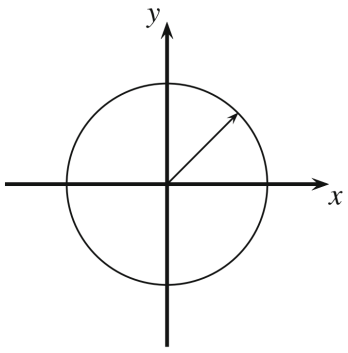

Abbildung 2.1.: r

den Termen x^2 und y^2 Koeffizienten voran gesetzt werden. Eine solche Gleichung, $ax^2 + by^2 = c$ erzeugt über $\mathbb{R}$ eine Ellipse, wie in Abbildung XY zu sehen.

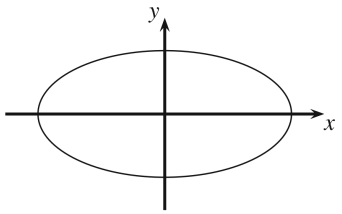

Abbildung 2.2.: r

Eine elliptische Kurve ist nun eine spezielle Polynomgleichung, der Form $y^2 = x^3+ax+b$, unter der Bedingung $4a^3+27b^3 \neq 0$. Eine solche Gleichung über $\mathbb{R}$ ist in Abbildung XY dargestellt. Damit elliptische Kurven sinnvoll in der Kryptologie eingesetzt werden

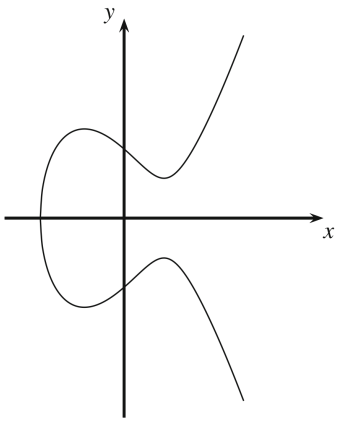

Abbildung 2.3.: r

können, muss die Polynomgleichung über einem Primkörper betrachtet werden. Das heißt einfach gesprochen, alle Berechnungen werden modulo p durchgeführt.

Definition: Elliptische Kurven über Primkörpern Die *elliptische Kurve* über $\mathbb{F}_{\text{|}}$, ist die Menge aller Punkte (x, y) mit $x, y \in \mathbb{F}_{\text{|}}$, welche die folgende Gleichung erfüllen:

$$y^2 \equiv x^3 + ax + b \bmod p, \text{ wobei } a, b \in \mathbb{F}_{\text{|}}$$

und die Bedingung

$$4a^3 + 27b^3 \neq 0$$

gelten müssen. Zu der elliptischen Kurve gehört des Weiteren auch der imaginäre *Punkt im Unendlichen* $\mathcal{O}$.

Durch die Bedingung XY werden sog. Singularitäten ausgeschlossen. Andernfalls gäbe es Punkte, deren Tangente nicht wohldefiniert ist, was für das Rechnen auf elliptischen Kurven jedoch erforderlich ist.

Nachdem elliptische Kurven nun definiert wurden, stellt sich die Frage, wie diese in der Kryptographie eingesetzt werden können. Wenn wir uns an das in Kapitel XY zurückerinnern, wird für die Konstruktion eines **DLP**s eine zyklische Gruppe benötigt. Eine eben solche findet sich in der Punktmenge der elliptischen Kurve wieder. Offen bleibt wie die Gruppenoperation definiert ist. Diese muss die in Kapitel XY geforderten Gruppengesetze erfüllen.

Als Symbol für die Gruppenoperation wird das Additionszeichen + verwendet. Durch die Gruppenoperation muss aus zwei Punkten $P = (x_1, y_2)$ und $Q = (x_2, y_2)$ der Kurve ein dritter Punkt R auf der Kurve berechnet werden.

$$P + Q = R$$

$$(x_1, y_1) + (x_2, y_2) = (x_3, y_3)$$

Am verständlichsten lässt sich diese Operation grafisch zeigen.
Elliptische Kurven über endlichen Körpern können grafisch nicht sinnvoll dargestellt werden. Ihre Form und Arithmetik lassen sich jedoch gut veranschaulichen wenn man sie auf $\mathbb{R}$ abbildet. Im Folgenden betrachten wir eine Elliptische Kurve, dargestellt in einem kartesischen Koordinatensystem, um die Gruppeneigenschaften bezüglich der Punktaddition zu zeigen. Hierbei sind nun zwei Fälle zu unterscheiden.

Punktaddition $P + Q$: Falls $P \neq Q$ erfolgt die geometrische Konstruktion, indem zunächst eine Grade durch die beiden Punkte gelegt wird. Aufgrund der Kurveneigenschaften hat diese immer einen dritten Schnittpunkt mit der Kurve. Dieser wird an der x-Achse gespiegelt um den gesuchten Punkt R zu erhalten. Abbildung XY zeigt die beschriebene Konstruktion.

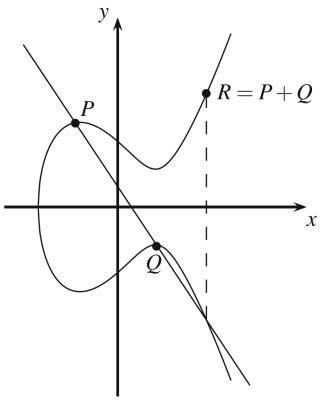

Abbildung 2.4.: r

Punktverdopplung $P + P$: Falls P und Q identisch sind erfolgt die geometrische Konstruktion, indem eine Tangente an den Punkt P angelegt wird. Diese liefert wieder einen weiteren Schnittpunkt mit der Kurve, welcher an der x-Achse gespiegelt wird um den Punkt R zu erhalten. Anstatt $R = P + Q$ schreibt man in diesem Fall $R = P + P = 2P$ Abbildung XY zeigt die beschriebene Konstruktion.

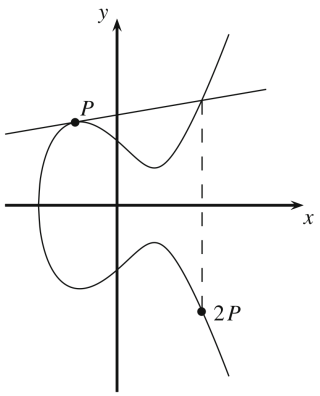

Abbildung 2.5.: r

Nach dieser grafischen Veranschaulichung sollte es leichter fallen die folgenden Formeln für die Punktaddition bzw. Punktverdopplung nachvollziehen zu können. Die Gruppenoperation existiert in jedem Körper, weshalb die Berechnung von R, wie grade gezeigt über den reellen Zahlen $\mathbb{R}$, als auch über einem Primkörper $\mathbb{F}_{\text{|}}$ durchgeführt werden kann.

Die Formeln für die Punktaddition und - verdopplung auf elliptischen Kurven können anhand der grade gezeigten Veranschaulichung hergeleitet werden. Das Vorgehen hierbei ist prinzipiell recht simpel.

Gegeben ist die Gleichung der elliptischen Kurve $y^2 = x^3 + ax + b$ und die Punkte $P = (x_1, y_1)$ und $Q = (x_2, y_2)$. Zunächst ist die Geradengleichung der Sekante durch P und Q zu ermitteln. Eine Grade im Allgemeinen hat die Form

$$g : y = sx + m.$$

Der Parameter s ist dabei die Steigung der Geraden und m ist der Schnittpunkt mit der y-Achse. Die Steigung s lässt sich wie gewohnt durch Anlegen des Steigungsdreiecks berechnen, also mit der Formel

$$s = \frac{y_2 - y_1}{x_2 - x_1}.$$

Zur Bestimmung des Schnittpunkts mit der y-Achse kann nun einer der beiden Punkte P oder Q in die Geradengleichung $y = \frac{y_2-y_1}{x_2-x_1} * x + m$ eingesetzt werden. Wenn wir $P = (x_1, y_1)$ einsetzen erhalten wir die folgende Gleichung

$$y_1 = \frac{y_2 - y_1}{x_2 - x_1} * x_1 + m,$$

welche nach m aufgelöst folgendermaßen aussieht:

$$m = y_1 - \frac{y_2 - y_1}{x_2 - x_1} * x_1$$

Durch Einsetzten aller Parameter in die obige Geradengleichung ergibt sich

$$y = \frac{y_2 - y_1}{x_2 - x_1} * x + y_1 - \frac{y_2 - y_1}{x_2 - x_1} * x_1$$

für die gesuchte Gerade g durch die Punkte P und Q. Um den dritten Schnittpunkt dieser Geraden g mit der elliptischen Kurve E zu ermitteln, sind beide Kurven gleichzusetzen. Da es für das weitere Vorgehen keine Rolle spielt und es der Übersichtlichkeit dient, werden im Folgenden wieder die Parameter s und m statt eben gezeigten Konkretisierungen verwendet. Es ergibt sich die Gleichung

$$(sx + m)^2 = x^3 + ax + b.$$

Im Normalfall ist das allgemeine Lösen eines solchen kubischen Polynoms nicht trivial. Wir haben hier jedoch den Vorteil, dass zwei der drei Schnittpunkte von g mit E schon bekannt sind. Im Grunde sind wir auf der Suche nach den Nullstellen des kubischen Polynoms, weshalb wir zur Verringerung des Funktionsgrades die Polynomdivision anwenden können, wobei die schon bekannten Nullstellen eben die x-Koordinaten der schon bekannten Schnittpunkte sind. Vorher wollen wird die Gleichung durch Umformung auf eine Seite bringen:

$$0 = x^3 - s^2x^2 - ax - 2smx - m^2 + b$$

Die Nullstellen sind $x_1 = x_1$ und $x_2 = x_2$. Daraus ergibt sich die folgende Polynomdivision:

Formel: Punktaddition und -verdopplung auf elliptischen Kurven:

$$x_3 = s^2 - x_1 - x_2$$

$$y_3 = s(x_1 - x_2) - y_1$$

, wobei

$$s = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & \text{, falls } P \neq Q \text{ (Punktaddition)} \\ \frac{3x_1^2 + a}{2y_1} & \text{, falls } P = Q \text{ (Punktverdopplung)} \end{cases}$$

Zur Erfüllung der Gruppeneigenschaften wird außerdem ein neutrales Element $\mathcal{O}$ benötigt. Alle Punkte P der elliptischen Kurve müssen die Eigenschaft $P + \mathcal{O} = P$. Da kein Punkt der elliptischen Kurve diese Eigenschaft erfüllen kann, wird der imaginäre *unendlich ferne Punkt* als neutrales Element $\mathcal{O}$ definiert. Dieser Punkt liefert den dritten *Schnittpunkt* mit der Kurve im Falle, dass ein Punkt P und der bezüglich der x-Achse gegenüberliegende Punkt $-P$ addiert werden. Abbildung XY zeigt den Fall grafisch.
Die Existenz eines neutralen Elements ermöglicht die Definition eines Inversen $-P$ für jeden Punkt P auf der Kurve, für welches $P + (-P) = \mathcal{O}$. Wie Abbildung XY entnommen werden kann, ist für den Punkt $P = (x_p, y_p)$ das Inverse also $-P = (y_p, -y_p)$ zu definieren. In einem Primkörper berechnet sich die negative y-Koordinate durch $-y_p = p - y_p$.

2.1. Punktbestimmung

Die Arithmetik für Elliptische Kurven wurde bereits besprochen. Nun wird erarbeitet, wie man die Punkte einer elliptischen Kurve bestimmt. Doch was ist ein Punkt einer elliptischen Kurve? Im Foglenden wird erklärt, was ein Punkt auf einer elliptischen Kurve ist und wie diese berehcnet werden können. Die Erklärungen werden anschließend anhand einige Beispiele näher erläutert. Anschließend wird Python-Code präsentiert, welcher die Punktbestimmung für eine elliptische Kurve mit p $>$ 3 durchführt und die Punkte anschließend in der Konsole ausgibt.

2.1.1. Rechnerische Grundlagen

Viele Mathematiker suchten auch eine Formel, mit denen sich die Anzahl der Punkte einer elliptischen Kurve schätzen lässt, ohne dass man diese vorher aus- oder berechnen muss. Joseph H. Silverman beweist in seinem Buch einen mathematischen Satz aus der Zahlentheorie, welcher Aussagen über die Anzahl der rationalen Punkte auf einer elliptischen Kurve trifft [**silverman**]. Bei dem mathematischen Satz handelt es sich um die Hasse–Weil–Schranke. Diese wird im allgemeinen dafür benutzt, um die Anzahl der Lösungen der Gleichung und der Bedingungen aber auch für die Einschränkung des Lösungsraums. Die Hasse-Weil-Schranke wird angelehnt an

[**reinholdhuebl**] wie folgt beschrieben. Sei k = $\mathbb{F}_p$ ein endlicher Körper und $\overline{E}$ eine elliptische Kurve über k, dann gilt

$$p + 1 - 2 * \sqrt{p} \leq |\overline{E}| \leq p + 1 + 2 * \sqrt{p}$$

Was ist jedoch die Hauptaussage der Hasse-Weil-Schranke? Sie besagt, dass sich bei großen p die Anzahl der Elemente der elliptischen Kurve in der Größenoprdnung von p bewegen. Diese Schranke bildet hierbei eine obere und untere Grenze. Die Anzahl der Punkte bewegt sich also innerhalb dieser Schranke. Dies ist für elliptische Kurven mit kleinem p uninteressant, jedoch wird diese Schranke für elliptische Kurven mit großem gewählten p, was in der Kryptographie gängig ist, relevant.

Doch wie lassen sich die Punkte konkret berechnen? Um diese Frage zu beantworten, müssen wir noch einmal die Grundlagen für elliptische Kurven aufgreifen. Wie Prof. Dr. Reinhold Hübl in seinem Manuskript der Kryptologie [**reinholdhuebl**] erläutert, ist eine elliptische Kurve mit den Charakteristiken char(k) = 0 oder char(k) = 3 eine Kurve mit der Form

$$F(X, Y) = Y^2 - X^3 - aX - b$$

mit $a, b \in k$, für die $4a^3 + 27b^2 \neq 0$ mod(p) gilt.

Stellt man die Gleichung der Funktion nach y^2 um, dann erhält man folgendes Polynom:

$$y^2 = x^3 + \text{ax} + \text{b mod(p)},$$

wobei $a, b \in \mathbb{Z}_p$. Diese Gleichung ist der Schlüssel für die im Voraus aufgeworfene Frage, was ein Punkt einer elliptischen Kurve ist. Diese sind nämlich alle Punkte (x, y), welche die Gleichung lösen und gleichzeitig die Bedingung $4a^3 + 27b^2 \neq 0$ mod(p) erfüllen.

Um die Punte auf einer Kurve zu berechnen, prüft man zunächst, ob es sich bei der betrachteten Kurve um eine elliptische Kurve mit p > 3 handelt. Dafür arbeitet man mit der Formel $4a^3 + 27b^2 \neq 0$. Man setzt die Parameter a und b der vermeintlichen elliptischen Kurve ein. Wenn die linke Seite $\neq 0$ ist, dann handelt es sich bei der besagten Kurve tatsächlich um eine elliptische Kurve mit p > 3. Rechnen wir dies nun einmal schematisch durch. Gegeben ist eine Funktion F mit

$$F(X, Y) = Y^2 - X^3 + 3X - 3 \in \mathbb{F}_{13}$$

Wie man der Funktion entnehmen kann ist a = - 3 = 10 und b = 3 in $\mathbb{F}_{13}$. Diese setzen wir nun in die Formel ein.

$$4 * 10^3 + 27 * 3^2 = 6 \neq 0 \text{ mod}(13)$$

Da $6 \neq 0$ ist definiert die Funktion eine elliptische Kurve in $\mathbb{F}_{13}$. Die Abbildung **??** zeigt die Zeichnung der Funktion dieser elliptischen Kurve im Koordinatensystem kartesischen Koordinatensystem:

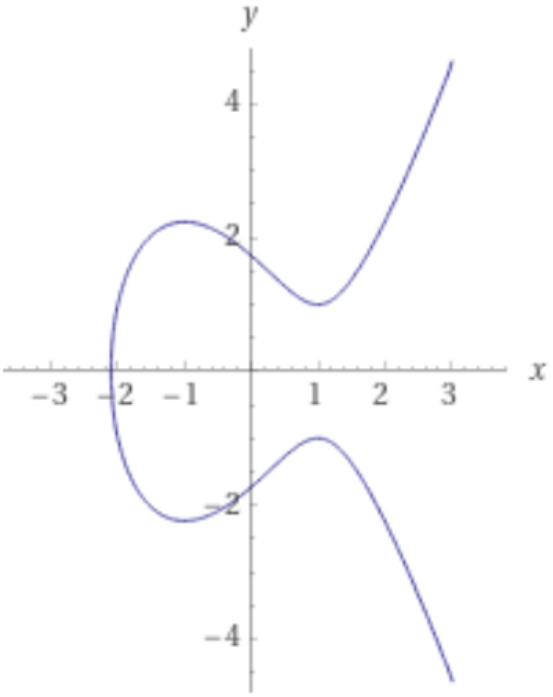

Abbildung 2.6.: Zeichnung der elliptischen Kurve
Quelle: Wolframalpha

Nachdem man allgemein geprüft hat, ob es sich bei der besagten Funktion um eine elliptische Kurve ahndelt, geht es jetzt um die Findung der Lösungen der umgestellten Gleichung, um alle Punkte zu finden.

2.1.2. Speziallfall: 4 ist Teiler von p + 1

U

2.1.3. Implementierung in Python

U