Lukuteoriaan perustuvia salausmenetelmiä

Rasmus Rehn

Matematiikan pro gradu -tutkielma

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
Kevät 2019

Tiivistelmä: Rasmus Rehn, *Lukuteoriaan perustuvia salausmenetelmiä* (engl. *Number theoretic cryptography*), matematiikan pro gradu -tutkielma, 49 sivua, Jyväskylän Yliopisto, Matematiikan ja tilastotieteen laitos, kevät 2019.

Tämän tutkielman tarkoitus on tutustuttaa lukija salakirjoituksen maailmaan lukuteorian näkökulmasta. Tutkielma sisältää salausmenetelmiin tarvittavat matemaattiset pohjatiedot, Diffie–Hellmanin salausmenetelmän ja RSA-salausmenetelmän, sekä molemmille salausmenetelmille soveltuvat purkumenetelmät. Monet esitetyistä tuloksista perustuvat algebraan, mutta tähän tutkielmaan tulokset on pyritty muuttaaman lukuteorian piiriin siten, että lukija ei välttämättä edes tiedä lukevansa algebraan perustuvia tuloksia.

Ensimmäisessä luvussa tutustutaan jaollisuuteen, lukuteoriaan ja kokonaislukujen tekijöihinjakoon. Tarkoituksena on esittää riittävän kattava pohja salausmenetelmiä varten, jotta kaikki niissä käytetty matematiikka olisi hyvin perusteltua. Toisessa luvussa käsitellään lisää jaollisuustuloksia modulaariaritmetiikan näkökulmasta. Kolmannessa luvussa käsitellään primitiivisiä juuria, jotka ovat tärkeässä roolissa Diffie–Hellmanin salausmenetelmän turvallisuuden perustelemisessa. Ensin osoitetaan, miten primitiivisten juurten moninkertojen avulla saadaan esitettyä salausmenetelmässä käytössä oleva lukujoukko. Tämän jälkeen osoitetaan, että salausmenetelmässä käytetyssä lukujoukossa on aina olemassa primitiivinen juuri.

Salausmenelmät esitetään neljännessä ja viidennessä luvussa siten, että niiden toimivuus perustellaan aiemmin esiteltyjen tuloksien avulla ja niistä näytetään havainnollistavat esimerkit. Neljännessä luvussa käsitellään Diffie–Hellmanin salausmenetelmää, jonka tarkoituksena on pystyä luomaan informaation lähettäjälle ja vastaanottajalle yhteinen salausavain, jota muut ulkopuoliset eivät pääse lukemaan. Sen jälkeen tutustutaan Shankin algoritmiin, jonka avulla ulkopuolinen henkilö voi yrittää selvittää tätä salausavainta. Viidennessä luvussa käsitellään RSA-menetelmää, jonka tarkoituksena on pystyä lähettämään viesti kahden henkilön välillä siten, että ulkopuolinen ei pysty lukemaan sen sisältöä. Lopuksi esitellään Pollardin $p-1$ -menetelmä, jolla ulkopuolinen henkilö voi yrittää avata salattua viestiä. Käyttämällä tutkielmassa esitettyjä salausmenetelmiä yhdessä, voidaan tietojen vaihtaminen toteuttaa turvallisesti esimerkiksi internet-sivustoilla.

Liitteissä on laskettu Maxima-ohjelmalla esimerkit kaikista käytetyistä salaus- ja purkumenetelmistä. Liitteet on pyritty tekemään mahdollisimman helppolukuisiksi Maximaa käyttämättömille henkilöille siten, että he voisivat toteuttaa itsenäisesti samat esimerkit.

Sisältö

Merkintöjä

Tässä lukijan tueksi muistilista tutkielmassa käytettävistä merkinnöistä:

- $\mathbb{N}$ Luonnollisten lukujen joukko, eli $\{1, 2, 3, \dots\}$
- $\mathbb{Z}$ Kokonaislukujen joukko, eli $\{\dots, -2, -1, 0, 1, 2, \dots\}$
- $a \mid b$ Luku b on jaollinen luvulla a
- $a \nmid b$ Luku b ei ole jaollinen luvulla a
- $a \equiv b \pmod{p}$ Luku a on kongruentti luvun b kanssa modulo p
- $a \not\equiv b \pmod{p}$ Luku a ei ole kongruentti luvun b kanssa modulo p
- $a \pmod{b}$ Luvun a jakojäännös luvulla b jaettaessa
- $\mathbb{Z}_n$ Kokonaislukujen joukko, jossa jokainen alkio on $(\text{Mod } n)$
- $\mathbb{Z}_p^*$, jossa p on alkuluku: Kokonaislukujen joukko $(\text{Mod } p)$, jossa alkioiden kesken operoidaan kertolaskuilla
- $\text{syt}(a, b)$ Lukujen a ja b suurin yhteinen tekijä

Termistöä

Alla olevassa listassa on lueteltu tässä tutkielmassa käytettävää termistöä. Termistöön kannattaa tutustua ennen lukemista, jotta tekstiä on helpompi ymmärtää.

- *Kryptografia, kryptologia ja salakirjoitus*
  - Termeillä tarkoitetaan oppia turvallisesta, salatusta viestinnästä kahden tai useamman kohteen välillä. Viestintä on turvallista silloin, kun ulkopuolinen henkilö ei pysty ymmärtämään tai selvittämään viestin sisältöä. Termi juontaa juurensa kreikan kielen sanoista *kryptos* eli ”näkymätön, salainen, piilossa”, *graphein* eli ”kirjoitus” ja *logia* eli ”oppi, tutkimus”.
- *Kolmas osapuoli*
  - Ulkopuolinen henkilö, joka ei saa ymmärtää viestin sisältöä.
- *Salauksen purkaminen*

  - Tutkielmassa salauksen purkamisella tarkoitetaan salatun tekstin muuttamista takaisin luettavaan, alkuperäiseen muotoon.
- *Algoritmi*
  - Tarkka ohje, jonka avulla jokin tehtävä tai tapahtuma voidaan suorittaa.
- *Salausavain*
  - Salausavain kertoo, miten viestin sisältöä muutetaan. Salauksen purkaminen tapahtuu käyttämällä salausavainta käänteisesti. Salausavaimet jaetaan *salaisiin-* ja *julkisiin* avaimiin.
- *Salainen avain*
  - Viestin salaaminen ja avaaminen tehdään saman avaimen avulla. Salaaminen ja purkaminen tapahtuvat nopeasti menetelmissä, joissa käytetään salaisia avaimia.
- *Julkinen avain*
  - Julkinen avain on nimensä mukaisesti julkinen. Avain voidaan jakaa kaikille, mutta vain vastaanottajalla on oma, salainen avain, jonka avulla hän saa salauksen purettua.
- *Symmetrinen salaus*
  - Lähettäjä ja vastaanottaja salaavat ja purkavat sisällön samalla salausavaimella, kuten *salaisella avaimella*.
- *Epäsymmetrinen salaus*
  - Salaamiseen käytetään eri avainta kuin salauksen purkamiseen, kuten *julkisella avaimella*.
- *Raaka voima*
  - Ongelman ratkaisemiseksi kokeillaan kaikki mahdolliset vastausvaihtoehdot. Tämä voi olla luonnollisesti todella hidasta.
- *Raa'an voiman menetelmä*
  - Menetelmä, jossa ongelman ratkaiseminen perustuu *raakaan voimaan*.

Johdanto

Kryptografialla eli salakirjoituksella tarkoitetaan sananmukaisesti oppia viestien salaamisesta. Tiedon salaamisen perustana on tiedon pysyminen luottamuksellisena kahden tai useamman osapuolen välillä. Ihmissuhteet, sodat, politiikka ja ihmisten väliset salaisuudet ovat olleet aikojen alusta asti asioita, joissa tietyn tiedon pysyminen salaisena on ollut välttämätöntä. Näin ollen on syntynyt tarve lähettää tietoa salaisena siten, että kukaan ulkopuolinen ei pääse tätä tietoa lukemaan. Tutustutaan ensin esimerkkiin, joka avaa salakirjoituksen ideaa hieman paremmin.

Salataan seuraava teksti: "*Kissalla ei ole turkkia*". Toteutetaan tämä siten, että siirretään jokaista kirjainta aakkosissa kolme pykälää eteenpäin ja aakkosten loppuessa siirrytään aakkosten alkuun seuraavan taulukon mukaisesti:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c

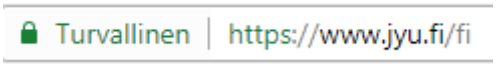
Näin jokainen aakkosten kirjain saadaan esitettyä jonain toisena kirjaimena. Kirjoitetaan viesti tätä salausta käyttäen:

k	i	s	s	a	l	l	a	e	i	o	l	e	t	u	r	k	k	i	a
n	l	v	v	d	o	o	d	h	l	r	o	h	w	x	u	n	n	l	d

Viesti "*Kissalla ei ole turkkia*" saatiin muotoon "*nlvvdood hl roh wxunnld*". Ulkopuolisen silmiin viesti näyttää järjettömältä, joskin näin yksinkertaisen salauksen purkaminen kävisi monelta ihmiseltä hyvinkin nopeasti hetken miettimisen jälkeen. Jos tilannetta vielä hieman vaikeutetaan, voidaan edellinen lause esittää neljän kirjaimen pätkissä: "*nlvv dood hlro hwxu nnld*". Suoraan viestiä katsomalla on mahdotonta sanoa, mitä alkuperäisessä viestissä lukee. Tässä esimerkissä käytetty salausmenetelmä on nimeltään Caesarin salakirjoitus. Caesarin salakirjoitus pohjautuu Julius Caesarin käskyihin, joita hän lähetti salattuna kirjeenvaihdossaan yllä esitetyllä tavalla. Caesarin salaus on yksi vanhimpia tunnettuja salakirjoitusmenetelmiä [**1**, ss. 1–4]. Caesarin salakirjoituksen purkamiseen ei tarvita monimutkaista matematiikkaa, mutta aikojen saatossa ja teknologian kehittyessä tarve turvallisemmille ja monimutkaisemmille salausmenetelmille on kasvanut.

Nyky-yhteiskunnan toimivuus on turvattu moneltakin eri näkökannalta kryptografian avulla. Rahaliikenne, politiikka, sähköposti, internet-sivustot ja yleisesti viestintä on tehtävä huolellisen salauksen kautta. Tietojen on säilyttävä luottamuksellisena ja eheinä, jotta niitä ei käytettäisi vääriin tarkoituksiin. Meillä kaikilla on henkilökohtaisia tietoja ja viestintää, joiden päätyminen vääriin käsiin voisi tuottaa harmia.

Kun yhdistät tietokoneella turvalliselle verkkosivustolle, näet osoiterivin vasemmalla puolella merkinnän HTTPS:// ja pienen lukon kuvan. (ks. kuva alla)



Termillä *HTTPS* tarkoitetaan sitä, että HTTP-protokolla, joka vastaa palvelinten ja selainten välisestä tietoliikenteestä, on suojattu SSL/TLS -salausprotokollan avulla. Protokollat ovat ohjenuoria sille, miten tietoa siirretään. Lukon kuvasta painamalla käyttäjä saa lisää tietoa siitä, miten tieto on salattu. Kuvassa olevan Jyväskylän yliopiston verkkosivuston lukkoa painamalla avautuvat seuraavanlaiset tiedot:

Yhteys-osion viimeisellä merkinnällä "*ECDHE_RSA*" tarkoitetaan sitä, että yhteyden salaamisessa on käytetty elliptisten käyrien Diffie–Hellmanin menetelmää ja RSA-menetelmää. Tässä tutkielmassa tutustutaan perinteiseen Diffie–Hellmanin menetelmään ja RSA-menetelmään. Päivittäisessä käytössä olevat verkkosivustot käyttävät siis muun muassa tässä tutkielmassa esitettyjä menetelmiä tietojen salauksessa. Tutkielmassa esitetyissä esimerkeissä lasketaan pienillä luvuilla havainnollistamisen vuoksi, mutta todellisuudessa näin pienillä luvuilla ei saavutettaisi tietoturvaa. Salausmenetelmien toimivuus perustuu raskaisiin ja pitkiin laskutoimituksiin, joita nykyisten tietokoneiden laskentatehoilla ei pystytä laskemaan. Menetelmien taustalla on lukuteoriaan pohjautuvaa matematiikkaa.
Tutkielmassa esiintyvät salausmenetelmät perustuvat niin kutsuttuihin *yksisuuntaisiin funktioihin* (engl. *trapdoor function, one-way function*). Yksisuuntaisella funktiolla tarkoitetaan sellaista funktiota, joka on nopea ja helppo laskea yhteen suuntaan millä tahansa arvoilla,

mutta vaikea tai lähes mahdoton peruuttaa takaisin lähtötilanteeseen. Peruuttamisella tarkoitetaan sellaista tilannetta, jossa funktion tuottama lopputulos tiedetään ja funktion toimintalogiikka tunnetaan, mutta laskutoimituksessa käytettyjä lukuarvoja ei tunneta. Joissakin funktioissa on tämän lisäksi niin kutsuttu takaovi (engl. *trapdoor*), jonka avulla yksisuuntainen funktio saadaankin laskettua takaisin alkuperäiseen tilanteeseen. Tutkielmassa esiintyvät salausmenetelmät perustuvat näihin havaintoihin. Täytyy kuitenkin muistaa, että yksisuuntaiset funktiot ovat aina oletettavasti turvallisia. Matematiikka ja teknologia kehittyvät jatkuvasti ja tulevaisuudessa tämänkaltaiset funktiot eivät välttämättä tuota enää riittävää tietoturvaa. Tulevaisuudessa voi olla mahdollista, että joku keksii sellaisen menetelmän, jonka avulla nämä salausmenetelmät voidaan purkaa helpoillakin toimenpiteillä. Tämän lisäksi muun muassa kvanttitietokoneiden kehittyessä laskentateho voi kasvaa niin suureksi, että tutkielmassa esitellyt salausmenetelmät voidaan purkaa nopeasti ja tämän seurauksena ne eivät olisi enää turvallisia.

Salausmenetelmiä lähestytään tässä tutkielmassa lukuteorian näkökulmasta. Monet käytetyistä tuloksista juontavat juurensa algebrasta, mutta tulokset on muutettu toimiviksi siten, että tutkielma pysyy lukuteorian piirissä. Suosittelen tutkielman lukijaa etenemään järjestyksessä. Lukija tutustutetaan termistöön, jota tutkielmaa lukiessa tarvitaan. Ensin käydään kattavasti läpi matematiikan pohjatiedot, joita esitellyissä salausmenetelmissä tarvitaan. Näiden jälkeen päästään viimein salausmenetelmien pariin. Salausmenetelmiä käsitellään niiden toimivuuden ja niihin kohdistuvien hyökkäysmenetelmien kannalta. Lopun liitteistä löytyy esimerkkejä tutkielman menetelmistä Maxima-laskentaohjelmistolla toteutettuna. Liitteiden esimerkit on toteutettu siten, että lukijan olisi mahdollisimman helppo seurata niiden välivaiheita.

LUKU 1

Matematiikkaa kryptografian taustalla

Tässä luvussa tutustutaan välttämättömiin matemaattisiin pohjatietoihin, joita tarvitaan tulevien salausmenetelmien ymmärtämiseksi ja perustelemiseksi. Tulevat tulokset rakentuvat aiemmin esitettyjen tulosten päälle, joten lukijaa suositellaan etenemään järjestyksessä pystyäkseen seuraamaan rakennettuja tuloksia. Tämän luvun tulokset ovat lukuteoriaa.

1.1. Jaollisuudesta

Käsitellään ensin jaollisuutta. Jakolaskujen laskeminen kokonaisluvuilla ja erityisesti teoria jakojäännösten taustalla luovat välttämättömän pohjan tässä tutkielmassa esitellyille salausmenetelmille. Tämän vuoksi on hyvä käydä ensin läpi perusteet jaollisuudesta. Tämä luku pohjautuu viitteeseen [**1**, ss. 10–34].

MÄÄRITELMÄ 1.1. Olkoot a ja b kokonaislukuja. Luku b on jaollinen luvulla a, jos $b = aq$ jollain $q \in \mathbb{Z}$. Merkitään tätä $a \mid b$. Jos b ei ole jaollinen luvulla a, niin merkitään $a \nmid b$.

LAUSE 1.2. *Olkoot a, b ja c kokonaislukuja. Tällöin*
(a) Jos $a \mid b$ ja $b \mid c$, niin $a \mid c$.
(b) Jos $a \mid b$ ja $a \mid c$, niin $a \mid (mb + nc)$ kaikilla $m, n \in \mathbb{Z}$.
(c) Jos $a \mid b$, mutta $a \nmid c$, niin tällöin $a \nmid (b + c)$.
(d) Jos $a \mid b$ ja $b \mid a$, niin $a = b$ tai $a = -b$.

TODISTUS. Olkoot a, b ja c kokonaislukuja.

(a) Koska $a \mid b$ ja $b \mid c$, niin on olemassa kokonaisluvut q ja p siten, että $b = qa$ ja $c = pb$. Tällöin

$$\begin{aligned} c &= pb \\ &= p(qa) \\ &= a(pq). \end{aligned}$$

Koska pq on kokonaisluku, niin Määritelmän 1.1 nojalla $a \mid c$.

(b) Olkoot m ja n kokonaislukuja. Koska $a \mid b$ ja $a \mid c$, niin on olemassa kokonaisluvut q ja p siten, että $b = qa$ ja $c = pa$. Edelleen

$$\begin{aligned} mb + nc &= m(qa) + n(pa) \\ &= (mq)a + (np)a \\ &= a(mq + np). \end{aligned}$$

Koska $mq + np$ on kokonaisluku, niin Määritelmän 1.1 nojalla $a \mid (mb + nc)$.

(c) Tehdään antiteesi: $a \mid (b + c)$. Koska oletuksen nojalla $a \mid b$, niin (b)-kohdan nojalla $a \mid ((b + c) - b)$. Mutta tällöin $a \mid c$, joka on ristiriita oletuksen $a \nmid c$ kanssa. Toisin sanottuna $a \nmid (b + c)$.

(d) Koska $a \mid b$ ja $b \mid a$, niin on olemassa kokonaisluvut q ja p siten, että $b = qa$ ja $a = pb$. Tällöin $a = pb = p(qa) = (pq)a$. Näin ollen on oltava $p = q = 1$ tai $p = q = -1$. Tästä seuraa, että $a = b$ tai $a = -b$.

□

LAUSE 1.3 (Jakoyhtälö). *Olkoot a ja b kokonaislukuja siten, että $b \neq 0$. Tällöin on olemassa yksikäsitteiset kokonaisluvut q ja r siten, että $a = qb + r$ ja $0 \leq r < |b|$.*

TODISTUS. Todistetaan ensin jakoyhtälön yksikäsitteisyys ja sen jälkeen sen ratkaisun olemassaolo. Olkoon $a = qb + r$, jossa $0 \leq r < |b|$ ja olkoon toisaalta $a = \hat{q}b + \hat{r}$, jossa $0 \leq \hat{r} < |b|$. Koska $-|b| < -\hat{r} \leq 0$, niin tällöin

$$\begin{aligned} -|b| < r - \hat{r} < |b|, \quad &\text{eli} \\ |r - \hat{r}| < |b|. & \end{aligned}$$

Koska $a = qb + r$ ja toisaalta $a = \hat{q}b + \hat{r}$, niin

$$\begin{aligned} qb + r &= \hat{q}b + \hat{r} \\ \Leftrightarrow r - \hat{r} &= \hat{q}b - qb \\ \Leftrightarrow r - \hat{r} &= (\hat{q} - q)b \end{aligned}$$

Joten $|b|$ jakaa kokonaisluvun $|r - \hat{r}|$. Jos nyt oletetaan, että $|r - \hat{r}| \neq 0$, niin $|b| \leq |r - \hat{r}|$. Tämä johtaisi ristiriitaan ylemmän johtopäätöksen $|r - \hat{r}| < |b|$ kanssa, joten on oltava $r = \hat{r}$, eli $(\hat{q} - q)b = 0$. Koska $b \neq 0$, niin myös $q = \hat{q}$. Jakoyhtälön ratkaisu on siis yksikäsitteinen.

Todistetaan vielä jakoyhtälön ratkaisun olemassaolo. Olkoon $b > 0$ ja $A = \{a - nb : n \in \mathbb{Z} \text{ ja } a - nb \geq 0\}$. Joukko A on alhaalta rajoitettu ja epätyhjä, joten siinä on pienin alkio r, joka voidaan kirjoittaa muodossa $r = a - qb \geq 0$ jollakin $q \in \mathbb{Z}$. Tällöin $r < b$, koska jos olisi

$r \geq b$, niin olisi olemassa luku $t = a - (q+1)b \geq 0$, mutta r on jo joukon A pienin alkio. Saadaan siis $a = qb + r$ ja $0 \leq r < |b|$.

Jos taas $b < 0$, niin soveltamalla yllä olevaa tapausta lukuun $-b > 0$ saadaan $a = \hat{q}(-b)+r$ ja $0 \leq r < -b$. Näin ollen $a = qb + r$, missä $q = -\hat{q}$ ja $0 \leq r < |b|$. □

HUOMAUTUS 1.4. Jakoyhtälössä 1.3 esiintyvää lukua r kutsutaan *jakojäännökseksi*. Tässä tutkielmassa jakojäännöstä merkitään a (Mod b). Tämä kannattaa laittaa tässä vaiheessa muistiin, sillä jatkossa on tärkeää ymmärtää tämän merkinnän tarkoitus jakojäännöksenä.

ESIMERKKI 1.5. Koska $31 = 10 \cdot 3 + 1$, niin 31 (Mod 3)$= 1$.

Seuraavaksi katsotaan, mitä tarkoitetaan kahden kokonaisluvun suurimmalla yhteisellä tekijällä. Suurinta yhteistä tekijää tullaan käyttämään monessa tuloksessa tämän tutkielman aikana.

MÄÄRITELMÄ 1.6. Olkoot a ja b kokonaislukuja. Lukujen $a \neq 0$ ja $b \neq 0$ *suurin yhteinen tekijä* $\text{syt}(a, b)$ on suurin kokonaisluku, joka jakaa molemmat luvut a ja b.

ESIMERKKI 1.7. Lukujen 15 ja 5 suurin yhteinen tekijä on $\text{syt}(15, 5) = 5$.

LEMMA 1.8 (Bezoutin lemma). *Olkoon $a \neq 0$ ja $b \neq 0$ kokonaislukuja. Tällöin on olemassa kokonaisluvut x ja y siten, että*

$$\text{syt}(a, b) = ax + by.$$

TODISTUS. Olkoon $A = \{ax + by : \ x, y \in \mathbb{Z} \text{ ja } ax + by > 0\}$. Joukko A on alhaalta rajoitettu ja epätyhjä, sillä jos valitaan $x = a$ ja $y = b$, niin $ax + by = a^2 + b^2 > 0$. Olkoon c joukon A pienin alkio. Osoitetaan, että a ja b ovat jaollisia luvulla c.

Tehdään antiteesi. Oletetaan, että a ei ole jaollinen luvulla c. Tällöin Jakoyhtälön 1.3 nojalla $a = qc + r$, jossa $0 < r < c$. Edelleen saadaan, että $qc = a - r$. Koska $c = ax + by$, niin

$$\begin{aligned} & qc = a - r \\ \Leftrightarrow \quad & q(ax + by) = a - r \\ \Leftrightarrow \quad & qax + qby - a = -r \\ \Leftrightarrow \quad & r = a(1 - qx) + b(-qy) \end{aligned}$$

Jakojäännös $r > 0$ saatiin siis esitettyä muodossa $\hat{a}x + \hat{b}y$, joten $r \in A$. Jakoyhtälön 1.3 nojalla $r < c$. Luvun c piti olla joukon A pienin alkio, joten päädytään ristiriitaan. Luku a on siis jaollinen luvulla c.

Vastaavasti voidaan osoittaa, että luku b on jaollinen luvulla c. Tästä seuraa, että myös $ax + by$ on jaollinen luvulla c siten, että $1 \leq c \leq \text{syt}(a, b)$. Toisaalta $c = ax + by$ ja siten $\text{syt}(a, b) \mid c$, joten tästä seuraa, että $\text{syt}(a, b) \leq c$. Nämä tiedot yhdistämällä huomataan, että

$\text{syt}(a,b) \leq c \leq \text{syt}(a,b)$. Näin ollen $c = ax + by$ on lukujen a ja b suurin yhteinen tekijä ja $\text{syt}(a,b) = ax + by$.

□

HUOMAUTUS 1.9. Olkoon a, b ja x, y kokonaislukuja. Muodossa $ax + by$ esitettyä lukua kutsutaan lukujen a ja b *lineaarikombinaatioksi*.

1.2. Alkuluvut ja lukujen tekijöihinjako

Alkulukuja käytetään jokaisessa tässä tutkielmassa esitetyssä salausmenetelmässä. Alkuluvuilla on monia mielenkiintoisia ominaisuuksia, jotka osaltaan mahdollistavat salausmenetelmien toimivuuden. Esitellään seuraavaksi, mitä alkuluvut ovat ja tutustutaan joihinkin niiden perusominaisuuksiin. Luvun tulokset pohjautuvat viitteisiin [**1**] ja [**4**].

MÄÄRITELMÄ 1.10. Olkoon p luonnollinen luku siten, että $p \geq 2$. Luku p on *alkuluku*, jos se on jaollinen vain itsellään ja luvulla 1.

MÄÄRITELMÄ 1.11. Kokonaisluvut $a \neq 0$ ja $b \neq 0$ ovat *suhteellisia alkulukuja* keskenään, jos $\text{syt}(a,b) = 1$.

ESIMERKKI 1.12. Positiiviset, lukua 15 pienemmät ja luvun 15 kanssa suhteelliset alkuluvut ovat $1, 2, 4, 7, 8, 11, 13$ ja 14.

LEMMA 1.13 (Eukleideen lemma). *Olkoon a, b ja c kokonaislukuja. Jos $a \mid bc$ ja a ja b ovat keskenään suhteellisia alkulukuja, niin $a \mid c$.*

TODISTUS. Olkoon a ja b suhteellisia alkulukuja. Tällöin Määritelmän 1.11 nojalla $\text{syt}(a,b) = 1$ ja edelleen Bezoutin Lemman 1.8 nojalla $ax + by = 1$ joillekin $x, y \in \mathbb{Z}$. Kerrotaan yhtälön molemmat puolet c:llä, jolloin

$$cax + cby = c.$$

Koska $a \mid cb$ ja $a \mid ca$, niin Lauseen 1.2 nojalla $a \mid (cax + cby)$, eli $a \mid c$. □

Suuria kokonaislukuja käsiteltäessä on luontevaa jakaa ne pienempiin tekijöihin, jotta laskeminen nopeutuisi ja helpottuisi. Kokonaislukujen tekijöihinjako on erityisen tärkeässä roolissa RSA-menetelmässä, jota käsitellään tämän tutkielman Luvussa 5. Tutustutaan seuraavaksi *aritmetiikan peruslauseeseen*, jonka mukaan jokainen kokonaisluku voidaan esittää alkulukujen tulona.

LAUSE 1.14 (Aritmetiikan peruslause). *Jokainen luonnollinen luku $a > 1$ voidaan esittää yksikäsitteisesti alkulukujen $p_1, p_2, \ldots, p_n$ tulona, kun $p_1 \leq p_2 \leq \cdots \leq p_n$.*

TODISTUS. Osoitetaan ensin tuloksen yksikäsitteisyys. Olkoon $a = p_1 p_2 \cdots p_n$ jollain $n \geq 1$ siten, että $p_1 \leq p_2 \leq \cdots \leq p_n$. Olkoon toisaalta $a = q_1 q_2 \cdots q_m$ jollain $m \geq 1$ siten, että $q_1 \leq q_2 \leq \cdots \leq q_m$. Näytetään induktion avulla, että tällöin $n = m$ ja $p_i = q_i$ kaikille $1 \leq i \leq n$.

Jos $n = 1$, niin $a = p_1$ on alkuluku. Tällöin $p_1 = a = q_1 q_2 \cdots q_m$. Koska p_1 on alkuluku, niin on oltava vastaavasti $m = 1$ ja $p_1 = q_1$. Oletetaan, että tulos on totta kaikille $1 \leq n \leq k$. Näytetään, että tulos pätee, kun $n = k + 1$. Olkoon $a = p_1 p_2 \cdots p_k p_{k+1}$, jossa $p_1 \leq p_2 \leq \cdots \leq p_k \leq p_{k+1}$ ja $a = q_1 q_2 \cdots q_m$, jossa $q_1 \leq q_2 \leq \cdots \leq q_m$. Koska p_{k+1} on luvun a tekijä, niin $p_{k+1} \mid a$ ja vastaavasti $p_{k+1} \mid q_1 \cdots q_m$. Eukleideen Lemman 1.13 nojalla $p_{k+1} \mid q_i$ jollain $1 \leq i \leq m$. Tästä seuraa, että $p_{k+1} = q_i$, koska muuten p_{k+1} olisi alkuluvun q_i tekijä, joka ei tietenkään ole mahdollista alkuluvulle. Jaetaan luku $p_{k+1} = q_i$ yhtälön

$$p_1 p_2 \cdots p_k p_{k+1} = q_1 q_2 \cdots q_m$$

oikealta ja vasemmalta puolelta. Tällöin induktio-oletuksen nojalla $n = m$. Koska lukuja p_i ja q_j on yhtä monta, niin luvut uudelleen järjestelemällä saadaan $p_i = q_j$ kaikilla $1 \leq i \leq n$.

Todistetaan vielä vahvan induktion avulla, että jokainen kokonaisluku $a > 1$ on alkulukujen tulo tai alkuluku. Induktion ensimmäinen askel on selvä, sillä luku 2 on alkuluku. Induktio-oletuksena on, että kaikki kokonaisluvut $2, 3, \ldots, n$ voidaan esittää alkulukujen tulona tai alkulukuna. Induktio-askeleessa tulee osoittaa, että luku $n + 1$ voidaan esittää vastaavasti. Oletetaan ensin, että $n + 1$ on alkuluku. Tämä tilanne on selvä ja väite pätee. Jos $n + 1$ ei ole alkuluku, se voidaan esittää lukujen a ja b tulona siten, että $n + 1 = ab$ ja $2 \leq a \leq b < n + 1$. Koska induktio-oletuksen mukaan luvut $2, 3, \ldots, n$ olivat joko alkulukuja tai niiden tuloja, niin myös luku $n + 1 = ab$ on alkulukujen tulo. Vahvan induktioperiaatteen nojalla jokainen kokonaisluku $a > 1$ on siis alkulukujen tulo tai alkuluku. □

1.3. Tekijöihinjaon algoritmeista

Tässä luvussa tustutaan jakoyhtälöön pohjautuviin Eukleideen algoritmeihin. Eukleideen algoritmin perusversiolla pystytään laskeman kahden kokonaisluvun suurin yhteinen tekijä. Perusversiosta johdetulla laajennetulla Eukleideen algoritmilla on merkittävä rooli RSA-menetelmän käyttämisessä käytännön sovelluksissa, joten on tärkeää ymmärtää sen toiminta ennen RSA-menetelmään tutustumista.

ALGORITMI 1.15 (Eukleideen Algoritmi). Esitetään Eukleideen algoritmi ensin sanallisesti. Olkoot annettuna kaksi luonnollista lukua.

(1) Jaa kahdesta annetusta luvusta suurempi pienemmällä luvulla ja ota jakojäännös talteen.

(2) Jaa jakaja edellisen jakolaskun talteen otetulla jakojäännöksellä ja ota uusi jakojäännös talteen.
(3) Toista kohtaa (2), kunnes jakojäännös on 0.
(4) Alkuperäisten kokonaislukujen suurin yhteinen tekijä on viimeinen nollasta poikkeava jakojäännös.

Esimerkki 1.16. Määritetään syt(621, 420) Eukleideen algoritmin avulla:

$$621 = 1 \cdot 420 + 201$$
$$420 = 2 \cdot 201 + 18$$
$$201 = 11 \cdot 18 + \boxed{3}$$
$$18 = 6 \cdot 3 + \underline{0}$$

Joten syt(621, 420) = 3.

Esitetään sama tulos vielä yleimmin:

Lause 1.17 (Eukleideen Algoritmi). *Olkoon a ja b luonnollisia lukuja siten, että $a \geq b$. Tällöin syt(a, b) saadaan laskettua seuraavan algoritmin mukaisesti:*

(1) Olkoon $r_0 = a$ ja $r_1 = b$.
(2) Asetetaan $i = 1$.
(3) Jaetaan r_{i-1} luvulla r_i, jolloin

$$r_{i-1} = r_i \cdot q_i + r_{i+1} \quad jossa\ 0 \leq r_{i+1} < r_i.$$

(4) Jos jakojäännös $r_{i+1} = 0$, niin tällöin $r_i = syt(a, b)$ ja lopetetaan algoritmi.
(5) Jos jakojäännös $r_{i+1} > 0$, niin asetetaan $i = i + 1$ ja toistetaan kohdasta 3.

Todistus. Katso liite [**1**, s. 13]. □

Esitellään seuraavaksi laajennettu versio Eukleideen algoritmista 1.17. Sen ajatuksena on etsiä kahden kokonaisluvun suurin yhteinen tekijä näiden samojen lukujen lineaarikombinaationa. Laajennettu versio Eukleideen algoritmista perustuu yllä esitetyn Eukleideen algoritmin peruuttamiseen lopusta alkuun. Laajennettua algoritmia tullaan käyttämään luvun 5 RSA-menetelmässä. Seuraava lause muistuttaa hieman Bezoutin Lemmaa 1.8 ja yhdessä todistuksen kanssa se antaa tavan esittää kahden kokonaisluvun suurin yhteinen tekijä samojen lukujen lineaarikombinaationa.

Lause 1.18 (Laajennettu Eukleideen algoritmi). *Olkoon a ja b luonnollisia lukuja. Tällöin yhtälöllä*

$$\text{syt}(a, b) = ax + by$$

on aina olemassa ratkaisu joillain kokonaisluvuilla x ja y.

TODISTUS. Olkoot a ja b luonnollisia lukuja. Merkitään jakojäännöksiä $r_i \in \mathbb{N}$, $i = 1, 2, 3 \ldots$ ja osamääriä $q_i \in \mathbb{N}$, $i = 1, 2, 3 \ldots$, kuten Lauseessa 1.17. Tarkastellaan ensin havainnollistavaa taulukkoa [**1**, s. 13] Eukleideen algoritmista:

$a = b \cdot q_1 + r_2$	kun $0 \leq r_2 < b$,
$b = r_2 \cdot q_2 + r_3$	kun $0 \leq r_3 < r_2$,
$r_2 = r_3 \cdot q_3 + r_4$	kun $0 \leq r_4 < r_3$,
$r_3 = r_4 \cdot q_4 + r_5$	kun $0 \leq r_5 < r_4$
$\vdots$	$\vdots$
$r_{t-2} = r_{t-1} \cdot q_{t-1} + r_t$	kun $0 \leq r_t < r_{t-1}$,
$r_{t-1} = r_t \cdot q_t$	
Lopuksi $r_t = \text{syt}(a, b)$.	

TAULUKKO 1. Eukleideen algoritmi

Todistetaan induktion avulla, että jokainen jakojäännös r_i voidaan esittää lukujen a ja b lineaarikombinaationa. Voidaan olettaa, että $a > b$. Jos $r_0 = a$ ja $r_1 = b$, niin tällöin $a = bq_1 + r_2$ ja edelleen $r_2 = a + bq_1$. Oletetaan, että jakojäännökset r_{k+1} ja r_k voidaan esittää lukujen a ja b lineaarikombinaationa. Tällöin Eukleideen algoritmista saadaan $r_k = q_{k+1}r_{k+1} + r_{k+2}$. Koska r_k ja r_{k+1} olivat lukujen a ja b lineaarikombinaatioita, niin luvun $r_{k+2} = -q_{k+1}r_{k+1} + r_k$ on myös oltava lukujen a ja b lineaarikombinaatio. Induktioperiaatteen nojalla jokainen jakojäännös r_i voidaan siis esittää lukujen a ja b lineaarikombinaationa, jolloin myös $\text{syt}(a, b) = r_t = ax + by$ joillain kokonaisluvuilla x ja y. □

ESIMERKKI 1.19. Etsitään kokonaisluvut x ja y samoilla luvuilla kuin Esimerkissä 1.16. Olkoon $a = 621$ ja $b = 420$. Esimerkistä 1.16 saatiin selvitettyä $\text{syt}(621, 420) = 3$. Etsitään vielä luvut $x, y \in \mathbb{Z}$ yhtälöstä

$$3 = 621x + 420y$$

Lähdetään peruuttamaan Eukleideen algoritmia Esimerkistä 1.16 toiseen suuntaan.

$$3 = 201 - 11 \cdot 18$$

Korvataan luku 18 sijoittamalla sen tilalle Esimerkin 1.16 aiemmasta välivaiheesta:

$$\begin{aligned} 3 &= 201 - 11 \cdot (420 - 2 \cdot 201) \\ \Leftrightarrow 3 &= 201 - 11 \cdot 420 + 22 \cdot 201 \\ \Leftrightarrow 3 &= 23 \cdot 201 - 11 \cdot 420 \end{aligned}$$

Korvataan jälleen luku 201 sijoittamalla sen tilalle Esimerkin 1.16 aiemmasta välivaiheesta:

$$\begin{aligned} & 3 = 23 \cdot (621 - 420) - 11 \cdot 420 \\ \Leftrightarrow\ & 3 = 23 \cdot 621 - 23 \cdot 420 - 11 \cdot 420 \\ \Leftrightarrow\ & 3 = 621 \cdot 23 + 420 \cdot (-34) \end{aligned}$$

Saatiin siis $x = 23$ ja $y = -34$. Näin voimme esittää luvun $\text{syt}(621, 420)$ lukujen 621 ja 420 lineaarikombinaationa muodossa $\text{syt}(621, 420) = 621 \cdot 23 + 420 \cdot (-34)$.

LUKU 2

Modulaariaritmetiikkaa

Modulaariaritmetiikka on toinen tapa tutkia jaollisuutta ja jakojäännöksiä. Tässä luvussa esitellään tutkielmassa esitettäviin salausmenetelmiin tarvittavia pohjatietoja perusteluiden ja esimerkkien kera. Aloitetaan kongruenssin määritelmällä ja tutustutaan sen ominaisuuksiin. Seuraavat tulokset perustuvat osittain viitteisiin [**1**] ja [**4**].

MÄÄRITELMÄ 2.1 (Kongruenssi). Olkoon $n > 0$ kokonaisluku. Kokonaisluku a on *kongruentti* kokonaisluvun b kanssa modulo n, merkitään $a \equiv b \pmod{n}$, jos $n \mid (a - b)$.

HUOMAUTUS 2.2. Merkintää (Mod n) käytetään jakojäännöksen yhteydessä, kun taas kongruenssin yhteydessä käytetty merkintä (mod n) tarkoittaa koko kongruenssin jakavaa lukua. Tässä tutkielmassa käytetään molempia merkintöjä hieman tilanteesta riippuen.

ESIMERKKI 2.3. $17 \equiv 2 \pmod{5}$, sillä $17 = 3 \cdot 5 + 2$ joten $5 \mid (17 - 2)$.
Yllä olevan huomautuksen tapauksessa merkittäisiin $17 \text{ (Mod } 5) = 2$.

LAUSE 2.4. *Kongruenssi on ekvivalenssirelaatio kokonaislukujen joukossa, eli jos a, b, c ja n ovat kokonaislukuja siten, että $n > 0$, niin tällöin*

(1) $a \equiv a \pmod{n}$.
(2) Jos $a \equiv b \pmod{n}$, *niin* $b \equiv a \pmod{n}$.
(3) Jos $a \equiv b \pmod{n}$ *ja* $b \equiv c \pmod{n}$, *niin* $a \equiv c \pmod{n}$.

TODISTUS. Olkoot a, b ja c kokonaislukuja ja n luonnollinen luku.

(1) Koska $(a - a) = 0$ ja $n \mid 0$, niin $a \equiv a \pmod{n}$.

(2) Olkoon $a \equiv b \pmod{n}$. Koska $a = kn + b$ jollain $k \in \mathbb{Z}$, niin $b = (-k)n + a$. Tästä seuraa, että $b - a = (-k)n$, ja siten $b \equiv a \pmod{n}$.

(3) Olkoon $a \equiv b \ (mod\ n)$ ja $b \equiv c \pmod{n}$. Tällöin on olemassa $k_1, k_2 \in \mathbb{Z}$ siten, että

$$\begin{aligned} a &= k_1 n + b \\ &= k_1 n + k_2 n + c \\ &= (k_1 + k_2)n + c. \end{aligned}$$

Koska $(k_1 + k_2) \in \mathbb{Z}$, niin $a \equiv c \pmod{n}$.

□

Kolmessa seuraavassa lauseessa esitetään ja todistetaan kongruenssin laskusääntöjä. Esitetään ensin, kuinka kongruenssi on yhteensopiva kerto- ja yhteenlaskun suhteen.

LAUSE 2.5. *Olkoot a, b, c, d ja n kokonaislukuja siten, että $n > 0$.*

(1) Olkoot $a \equiv b \pmod{n}$ ja $c \equiv d \pmod{n}$. Tällöin kaikilla $x, y \in \mathbb{Z}$ pätee, että $ax+cy \equiv bx + dy \pmod{n}$.

(2) Olkoot $a \equiv b \pmod{n}$ ja $c \equiv d \pmod{n}$. Tällöin $ac \equiv bd \pmod{n}$.

(3) Olkoot $a \equiv b \pmod{n}$. Tällöin $a^m \equiv b^m \pmod{n}$ kaikilla $m \in \mathbb{N}$.

TODISTUS. Todistetaan Lauseen kohdat 1)–3):

(1) Koska $n \mid (a - b)$ ja $n \mid (c - d)$, niin Lauseen 1.2 nojalla $n \mid [x(a - b) + y(c - d)]$. Toisaalta

$$\begin{aligned} x(a-b) + y(c-d) &= xa - xb + yc - yd \\ &= ax + cy - bx - dy \\ &= ax + cy - (bx + dy). \end{aligned}$$

Joten $ax + cy \equiv bx + dy \pmod{n}$.

(2) Koska $n \mid (a - b)$ ja $n \mid (c - d)$, niin on olemassa kokonaisluvut x ja y siten, että $xn = a - b$ ja $yn = c - d$. Tällöin $a = xn + b$ ja $c = yn + d$. Kerrotaan luvut a ja c keskenään, jolloin

$$\begin{aligned} & ac = (xn + b)(yn + d) \\ \Leftrightarrow \quad & ac = xyn^2 + xnd + byn + bd \\ \Leftrightarrow \quad & ac - bd = xyn^2 + xnd + byn \\ \Leftrightarrow \quad & ac - bd = n(xyn + xd + by). \end{aligned}$$

Nyt $n \mid n(xyn + xd + by)$, joten $n \mid (ac - bd)$. Saatiin siis näytettyä, että $ac \equiv bd \pmod{n}$.

(3) Osoitetaan väite induktion avulla. Kun $m = 1$, niin $a \equiv b \pmod{n}$. Oletetaan, että väite pätee, kun $m = k$, eli $a^k \equiv b^k \pmod{n}$. Lauseen 2.5 kohdan (2) nojalla $aa^k \equiv bb^k \pmod{n}$, eli $a^{k+1} \equiv b^{k+1} \pmod{n}$. Väite siis pätee, kun $m = k + 1$. Tällöin induktioperiaatteen nojalla alkuperäinen väite pätee.

□

LAUSE 2.6. *Olkoot a ja b kokonaislukuja ja n luonnollinen luku. Tällöin*

$$ab \ (\mathrm{Mod}\ n) \equiv (a \ (\mathrm{Mod}\ n))(b \ (\mathrm{Mod}\ n)) \pmod{n}$$

TODISTUS. Jakoyhtälön 1.3 nojalla luvut a ja b voidaan kirjoittaa muodossa $a = nq_1 + r_1$ ja $b = nq_2 + r_2$. Merkitään jakojäännöksiä kuten Huomautuksessa 1.4, jolloin $a \pmod{n} = r_1$ ja $b \pmod{n} = r_2$. Tällöin kongruenssin vasemman puolen termille saadaan

$$ab \pmod{n} \equiv r_1 r_2 \pmod{n}.$$

Toisaalta $r_1 = a \pmod{n}$ ja $r_2 = b \pmod{n}$, joten

$$ab \pmod{n} \equiv (a \pmod{n})(b \pmod{n}) \pmod{n}.$$

□

LAUSE 2.7. *Olkoot a kokonaisluku sekä m ja n luonnollisia lukuja. Tällöin*

$$a^m \pmod{n} \equiv (a \pmod{n})^m \pmod{n}$$

TODISTUS. Todistetaan tulos induktion avulla. Väite pätee, kun $m = 1$, sillä $a \pmod{n} \equiv (a \pmod{n}) \pmod{n}$. Oletetaan, että väite pätee kun $m = k - 1$ ja osoitetaan edelleen, että se pätee kun $m = k$. Kerrotaan kongruenssin

$$a^{k-1} \pmod{n} \equiv (a \pmod{n})^{k-1} \pmod{n}$$

molemmat puolet termillä $a \pmod{n}$. Tällöin kongruenssin vasemman puolen termille saadaan Lauseen 2.6 nojalla

$$\begin{aligned}(a^{k-1} \pmod{n}) \cdot (a \pmod{n}) &\equiv (a^{k-1} \cdot a) \pmod{n} \\ &\equiv a^k \pmod{n} \pmod{n}.\end{aligned}$$

Lisäksi kongruenssin oikealle puolelle saadaan

$$(a \pmod{n})^{k-1} \cdot (a \pmod{n}) \equiv (a \pmod{n})^k \pmod{n}.$$

Yhdistämällä vasemmalle ja oikealle puolelle tehdyt toimitukset, saadaan

$$a^k \pmod{n} \equiv (a \pmod{n})^k \pmod{n}.$$

Väite siis pätee kun $m = k$, joten induktioperiaatteen nojalle lause pätee. □

Määritellään seuraavaksi jäännösluokat ja niiden muodostama joukko $\mathbb{Z}_n$. Jäännösluokat tulevat olemaan tärkeässä roolissa tutkielman lopuissa tuloksissa ja esitettävissä salausmenetelmissä.

MÄÄRITELMÄ 2.8. Olkoot a ja n kokonaislukuja siten, että $n > 0$. Kokonaislukujen, jotka ovat kongruentit luvun a kanssa modulo n, muodostamaa joukkoa kutsutaan luvun a *jäännösluokaksi modulo n*. Tätä merkitään

$$[a]_n = \{b \in \mathbb{Z} : b \equiv a \pmod{n}\}.$$

Jäännösluokkien joukkoa merkitään $\mathbb{Z}_n = \{[0]_n, [1]_n, [2]_n, \ldots [n-1]_n\}, n \in \mathbb{N}$.

HUOMAUTUS 2.9. Tässä tutkielmassa hyödynnetään erityisesti joukkoa $\mathbb{Z}_p$, jossa p on alkuluku.

HUOMAUTUS 2.10. Joukon $\mathbb{Z}_n$ alkioita $[a]_n$ merkitään monesti myös ilman hakasulkeita lukemisen helpottamiseksi.

ESIMERKKI 2.11. Olkoon jäännösluokkien joukko $\mathbb{Z}_4$. Tällöin sen alkioille on esimerkiksi $[0]_4 = [4]_4 = [8]_4$ ja $[1]_4 = [5]_4 = [9]_4$.

Määritellään seuraavaksi kokonaisluvun *kertaluku* modulo n. Lukujen kertalukuja tarvitaan erityisesti luvussa 3 ja luvun 4 Shankin algoritmissa.

MÄÄRITELMÄ 2.12. Olkoot a ja n suhteellisia alkulukuja. Tällöin luvun a *kertaluku* modulo n on pienin luonnollinen luku m, jolle $a^m \equiv 1 \pmod{n}$.

ESIMERKKI 2.13. Luvut $a = 2$ ja $m = 7$ ovat suhteellisia alkulukuja keskenään. Tällöin

$$2^1 = 2 \equiv 2 \pmod{7},$$
$$2^2 = 4 \equiv 4 \pmod{7},$$
$$2^3 = 8 \equiv 1 \pmod{7},$$

Joten Määritelmän 2.12 nojalla luvun 2 kertaluku modulo 7 on 3.

Esitetään seuraavaksi kaksi lausetta joiden avulla todistetaan, että kertaluku on aina hyvin määritelty. Ensimmäisessä tuloksessa perustellaan, milloin jakolasku on sallittua kongruenssissa ja toisessa luvussa näytetään, että Määritelmän 2.12 kongruenssilla on aina olemassa ratkaisu.

LAUSE 2.14. *Jos c ja n ovat suhteellisia alkulukuja, joille $ac \equiv bc \pmod{n}$ jollain $a, b \in \mathbb{Z}$, niin tällöin $a \equiv b \pmod{n}$.*

TODISTUS. Kongruenssin määritelmän nojalla $n \mid (ac - bc)$, joten $n \mid (a - b)c$. Koska c ja n ovat suhteellisia alkulukuja, niin $\text{syt}(c, n) = 1$. Tällöin Eukleideen lemman 1.13 nojalla $n|(a - b)$, eli $a \equiv b \pmod{n}$. □

LAUSE 2.15. *Olkoot a ja n suhteellisia alkulukuja. Tällöin kongruenssilla*

$$a^m \equiv 1 \pmod{n}$$

on olemassa ainakin yksi ratkaisu $m \in \mathbb{N}$.

TODISTUS. Tarkastellaan lukuja $a, a^2, a^3, \ldots$ Joukossa $\mathbb{Z}_n$ on n kappaletta alkioita, joten on olemassa kaksi indeksiä $i > j$ siten, että

$$a^i \equiv a^j \pmod{n}.$$

Muussa tapauksessa kaikki joukon $\mathbb{Z}_n$ alkiot $[a^i]_n$ olisivat eri alkioita. Tämä on ristiriita.

Kongruenssin $a^i \equiv a^j \pmod{n}$ vasemman puolen termi voidaan kirjoittaa muodossa $a^i = a^{i-j} \cdot a^j$ ja oikean puolen termi muodossa $a^j = a^j \cdot 1$. Tällöin kongruenssi saadaan muotoon

$$a^{i-j} \cdot a^j \equiv a^j \cdot 1 \pmod{n}.$$

Koska $\text{syt}(a, n) = 1$, niin myös $\text{syt}(a^j, n) = 1$. Tällöin Lauseen 2.14 nojalla

$$a^{i-j} \equiv 1 \pmod{n}.$$

Löydettiin siis kokonaisluku $m = i - j$ siten, että väite pätee. □

Seuraava tulos liittyy kongruenssin modulaariin käänteislukuun. Tulosta tarvitaan erityisesti RSA-menetelmän salausavaimen luomisessa. Tuloksen todistamiseen tarvitaan ensin yksi aputulos.

LEMMA 2.16. *Olkoon $a \neq 0$ ja $n \neq 0$ kokonaislukuja. Tällöin $ab \equiv 1 \pmod{n}$ jollekin kokonaisluvulle b, jos ja vain jos $\text{syt}(a, n) = 1$.*

TODISTUS. Olkoon ensin $\text{syt}(a, n) = 1$. Tällöin Bezoutin Lemman 1.8 nojalla $ax + ny = 1$ jollain $x, y \in \mathbb{Z}$ ja tästä edelleen $ax - 1 = -ny$. Tämän nojalla $ax - 1$ on jaollinen luvulla $-ny$ ja edelleen luvulla n. Tällöin kongruenssin Määritelmän 2.1 nojalla $ax \equiv 1 \pmod{n}$. Kun valitaan $b = x$, niin saadaan

$$ab \equiv 1 \pmod{n}.$$

Olkoon sitten $ab \equiv 1 \pmod{n}$ jollakin $b \in \mathbb{Z}$. Tällöin $ab - 1 = vn$ jollain $v \in \mathbb{Z}$. Tästä saadaan edelleen, että $ab - vn = 1$. Koska $ab - cn$ on jaollinen luvulla $\text{syt}(a, n)$ ja $ab - cn = 1$, niin on oltava $\text{syt}(a, n) = 1$. Näin ollen lause pätee. □

LAUSE 2.17 (Modulaari käänteisluku). *Olkoon p alkuluku. Tällöin alkiolle $[a]_p \in \mathbb{Z}_p$, $a \neq 0$, on olemassa yksikäsitteinen alkio $[b]_p \in \mathbb{Z}_p$, $b \neq 0$ siten, että*

$$ab \equiv 1 \pmod{p}$$

Lukua b kutsutaan luvun a modulaariksi käänteisluvuksi.

TODISTUS. Olemassaolo seuraa suoraan, kun valitaan $n = p$ Lemmassa 2.16. Koska p on alkuluku, niin $\text{syt}(a, p) = 1$. Osoitetaan vielä luvun b yksikäsitteisyys. Olkoon c alkio siten,

että $ac \equiv 1 \pmod{p}$. Tällöin $ab \equiv ac \equiv 1 \pmod{p}$. Siten Lauseen 2.4 nojalla

$$\begin{aligned} b &\equiv 1 \cdot b \\ &\equiv acb \\ &\equiv abc \\ &\equiv 1 \cdot c \\ &\equiv c \pmod{p}. \end{aligned}$$

Joten $b \equiv c \pmod{p}$. □

Esimerkki 2.18. Modulaareja käänteislukuja pystytään laskemaan kätevästi Laajennetun Eukleideen algoritmin avulla. Olkoon $7x \equiv 1 \pmod{40}$, jossa $x \in \mathbb{Z}$. Kun $y \in \mathbb{Z}$, niin kongruenssi voidaan kirjoittaa muodossa

$$7x + 40y = 1.$$

Ratkaistaan luku x käyttämällä Laajennettua Eukleideen algoritmia 1.18:

$$\begin{aligned} 40 &= 5 \cdot 7 + 5 \\ 7 &= 1 \cdot 5 + 2 \\ 5 &= 2 \cdot 2 + 1. \end{aligned}$$

Peruutetaan algoritmia sijoittelemalla:

$$\begin{aligned} & 5 - 2 \cdot 2 = 1 \\ \Leftrightarrow\ & 5 - 2 \cdot (7 - 5) = 1 \\ \Leftrightarrow\ & 3 \cdot 5 - 2 \cdot 7 = 1 \\ \Leftrightarrow\ & 3 \cdot (40 - 5 \cdot 7) - 2 \cdot 7 = 1 \\ \Leftrightarrow\ & 3 \cdot 40 - 17 \cdot 7 = 1. \end{aligned}$$

Joten saatiin $x = -17$ ja $y = 3$. Koska $[-17]_{40} = [23]_{40}$, niin tällöin luvun 7 käänteisluku modulo 40 on 23.

LUKU 3

Primitiiviset juuret joukossa $\mathbb{Z}_p^*$

3.1. Primitiivinen juuri

Määritellään ensin joukko, jossa tulemme operoimaan.

MÄÄRITELMÄ 3.1. Joukko $\mathbb{Z}_p^*$, jossa p on alkuluku, koostuu alkioista $1, 2, ..., p-1$ siten, että perusoperaationa joukon alkioiden välillä käytetään kertolaskua.

Tutkitaan seuraavaksi tällaisen joukon ominaisuuksia. Tulevassa luvussa osoitetaan, että joukossa $\mathbb{Z}_p^*$ on aina olemassa niin sanottu *primitiivinen juuri*, jonka moninkertojen avulla voidaan esittää kaikki joukon $\mathbb{Z}_p^*$ alkiot. Tuloksen avulla voidaan varmistaa, että Diskreetin logaritmin ongelmalla 3.18 on aina olemassa ratkaisu joukossa $\mathbb{Z}_p^*$. Luvun tulokset pohjautuvat viitteisiin [**1**], [**2**] ja [**3**]. Tulokset on rakennettu lukuteorian avulla, vaikkakin monet lauseista juontavat juurensa algebrasta.

MÄÄRITELMÄ 3.2. Luku $g \in \mathbb{Z}_p^*$ on joukon $\mathbb{Z}_p^*$ *primitiivinen juuri*, jos sen kertaluku modulo p on $p-1$.

ESIMERKKI 3.3. Luku 3 on joukon $\mathbb{Z}_7^*$ primitiivinen juuri, koska $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$. Lisäksi luvun 3 moninkerroilla saadaan esitettyä kaikki $\mathbb{Z}_7^*$ alkiot seuraavasti:

$$3^1 \equiv 3 \pmod 7$$
$$3^2 \equiv 2 \pmod 7$$
$$3^3 \equiv 6 \pmod 7$$
$$3^4 \equiv 4 \pmod 7$$
$$3^5 \equiv 5 \pmod 7$$
$$3^6 \equiv 1 \pmod 7.$$

HUOMAUTUS 3.4. Yllä olevasta esimerkistä voidaan huomata eräs primitiivisten juurten tärkeä ominaisuus. Jos nimittäin g on primitiivinen juuri, niin sen moninkertojen $g, g^2, \ldots, g^{p-1}$ avulla voidaan esittää kaikki joukon $\mathbb{Z}_p^*$ alkiot. Sanotaan, että tällöin primitiivinen juuri *virittää* joukon $\mathbb{Z}_p^*$.

Osoitetaan Huomautuksen 3.4 väite todeksi yleisessä tapauksessa. Tätä varten tarvitaan ensin yksi aputulos, joka lienee tutumpi algebran puolelta.

LEMMA 3.5. *Olkoon $a \in \mathbb{Z}_p^*$ siten, että alkion a kertaluku on n. Tällöin $a^k \equiv 1 \pmod{p}$ jos ja vain jos $n \mid k$.*

TODISTUS. Oletetaan ensin, että $n \mid k$. Tällöin Määritelmän 1.1 nojalla $k = nt$ jollain $t \in \mathbb{Z}$. Tällöin $a^k = a^{nt}$ ja koska alkion a kertaluku modulo p on n, niin $a^n \equiv 1 \pmod{p}$. Lauseen 2.5 nojalla

$$\begin{aligned} a^k &\equiv (a^n)^t \\ &\equiv (1)^t \\ &\equiv 1 \pmod{p}. \end{aligned}$$

Oletetaan sitten, että $a^k \equiv 1 \pmod{p}$. Jakoyhtälön 1.3 nojalla luku k voidaan esittää muodossa $k = nt + r$, missä $0 \leq r < n$. Nyt koska $a^k \equiv 1 \pmod{p}$, niin

$$\begin{aligned} 1 &\equiv a^k \\ &\equiv a^{nt} \cdot a^r \\ &\equiv (a^n)^t \cdot a^r \pmod{p}. \end{aligned}$$

Alkion a kertaluku modulo p on n, joten$(a^n)^t \equiv (1)^t \pmod{p}$ ja $a^r \equiv a^r \pmod{p}$. Siispä Lauseen 2.5 nojalla

$$\begin{aligned} (a^n)^t \cdot a^r &\equiv (1)^t \cdot a^r \\ &\equiv a^r \pmod{p}. \end{aligned}$$

Joten $a^r \equiv 1 \pmod{p}$. Tällöin on oltava $r = 0$, sillä muuten alkion a kertaluku modulo p olisi $r < n$, mutta $n > 0$ on luvun a kertalukuna pienin tällainen luku. Tällöin $k = nt$, eli $n \mid k$. □

LAUSE 3.6. *Olkoon $g \in \mathbb{Z}_p^*$ joukon $\mathbb{Z}_p^*$ primitiivinen juuri. Tällöin luvun g moninkerrat (Mod p) virittävät joukon $\mathbb{Z}_p^*$.*

TODISTUS. Joukossa $\mathbb{Z}_p^*$ on $p-1$ erillistä alkiota $1, 2, \ldots p-1$. Riittää osoittaa, että alkiot $g \pmod{p}, g^2 \pmod{p}, g^3 \pmod{p}, \ldots, g^{p-1} \pmod{p}$ ovat erillisiä, jolloin nämä g:n moninkerrat (Mod p) palauttavat itse asiassa alkiot $1, 2, \ldots, p-1$. Lisäksi alkiot $g \pmod{p}, g^2 \pmod{p}$, $g^3 \pmod{p}, \ldots, g^{p-1} \pmod{p}$ ovat nollasta poikkeavia, koska $g \nmid p$.

Primitiivisen juuren g kertaluku on $p-1$. Olkoon $1 \leq l \leq k < p-1$. Jos $g^k \equiv g^l \pmod{p}$, niin Lauseen 2.14 nojalla $g^{k-l} \equiv 1 \pmod{p}$. Tällöin Lemman 3.5 nojalla $(p-1) \mid (k-l)$. Mutta koska $0 \leq k - l \leq p - 2 < p - 1$, niin on oltava $k - l = 0$. Eli $g^k \equiv g^l \pmod{p}$ vain silloin, kun $k = l$. Näin ollen kaikki alkiot $g \pmod{p}, g^2 \pmod{p}, g^3 \pmod{p}, \ldots, g^{p-1} \pmod{p}$ ovat erillisiä.

Koska joukoissa $\mathbb{Z}_p^*$ ja $\{g \pmod p, g^2 \pmod p, g^3 \pmod p, \ldots, g^{p-1} \pmod p\} \subset \mathbb{Z}_p^*$ on molemmissa $p-1$ kappaletta erillisiä alkioita, niin nämä joukot ovat samat. Näin ollen primitiivisen juuren g moninkerrat (Mod p) virittävät joukon $\mathbb{Z}_p^*$. □

3.2. Primitiivisten juurten olemassaolo

Osoitetaan seuraavaksi muutama aputulos, joiden avulla todistetaan primitiivisten juurten olemassaolo joukossa $\mathbb{Z}_p^*$, kun p on alkuluku. Aloitetaan nämä aputulokset ranskalaisen matemaatikon Pierre de Fermatin pienellä lauseella. Olemassaolon todistus ja osa seuraavista aputuloksista perustuvat viitteisiin [**1**] ja [**2**, ss. 43–46].

LAUSE 3.7 (Fermatin pieni lause). *Olkoon p alkuluku ja a kokonaisluku. Tällöin*

$$a^{p-1} \equiv \begin{cases} 1 \pmod p & \text{jos } p \nmid a, \\ 0 \pmod p & \text{jos } p \mid a. \end{cases}$$

TODISTUS. Oletetaan ensin, että $p \mid a$. Tällöin $p \mid a^k$ kun k on kokonaisluku. Jos nyt valitaan $k = p-1$, niin $p \mid a^{p-1}$ ja tällöin $a^{p-1} \equiv 0 \pmod p$.

Oletetaan, että $p \nmid a$. Tutkitaan lukuja

$$a \pmod p, 2a \pmod p, 3a \pmod p, \ldots, (p-1)a \pmod p$$

Osoitetaan, että tämän listan luvut ovat erillisiä ja nollasta poikkeavia. Valitaan listasta kaksi indeksiä $1 \leq k, j \leq p-1$ siten, että $k \neq j$. Tehdään vastaoletus, että $ka \equiv ja \pmod p$. Tällöin

$$\begin{aligned} & ka \equiv ja \pmod p \\ \Leftrightarrow\ & (k-j)a \equiv 0 \pmod p. \end{aligned}$$

Koska p on alkuluku, niin joko $p \mid a$ tai $p \mid (k-j)$. Oletuksen nojalla $p \nmid a$, joten on oltava $p \mid (k-j)$. Luvut j ja k ovat välillä $[1, p-1]$, joten $j-k$ on tällöin välillä $[-(p-2), p-2]$. Nolla on ainoa luku tällä välillä, joka on jaollinen luvulla p. Tällöin on siis $k-j=0$ ja näin ollen $ka = ja$, joka on ristiriita oletuksen kanssa. Näin ollen yllä olevan listan luvut ovat erillisiä. Vastaava päättely osoittaa myös, että listan luvut modulo p ovat nollasta poikkeavia. Lukujen 1 ja $p-1$ välissä on muutenkin $p-1$ eri kokonaislukua, joten listassa on oltava numerot $1, 2, 3, 4, \ldots, (p-1)$ vain eri järjestyksessä.

Kerrotaan listan alkiot keskenään. Tällöin

$$a \cdot 2a \cdot 3a \cdot (p-1)a \equiv 1 \cdot 2 \cdot 3 \cdots (p-1) \pmod p.$$

Järjestellään lukuja hieman uudelleen, jolloin

$$a^{p-1} \cdot (p-1)! \equiv (p-1)! \pmod p.$$

Edelleen kongruenssin Määritelmän 2.1 nojalla

$$a^{p-1} \cdot (p-1)! - (p-1)! = (p-1)! \cdot (a^{p-1} - 1)$$

on jaollinen luvulla p. Koska p on alkuluku, niin $(p-1)!$ ei ole jaollinen luvulla p. Tällöin Eukleideen Lemman 1.13 nojalla p jakaa luvun $a^{p-1}-1$. Uudelleen kongruenssin Määritelmää 2.1 käyttämällä saadaan

$$a^{p-1} \equiv 1 \pmod{p}.$$

Joten päästiin haluttuun lopputulokseen. □

Primitiivisen juuren olemassaolon todistamiseksi tullaan tarvitsemaan tietoa polynomeista ja funktioista joukossa $\mathbb{Z}_p^*$. Tutustutaan hieman polynomien ja funktioiden ominaisuuksiin ja välttämättömiin pohjatietoihin, joita tarvitaan olemassaolon todistuksessa.

MÄÄRITELMÄ 3.8. Kokonaisluku x on funktion $f : \mathbb{Z} \to \mathbb{Z}$ juuri joukossa $\mathbb{Z}_p$, jos

$$f(x) \equiv 0 \pmod{p}.$$

MÄÄRITELMÄ 3.9. Kokonaislukukertoimisen polynomin

$$a_0 + a_1 x + \cdots + a_n x^n,$$

jossa $a_0, \ldots, a_n \in \mathbb{Z}$, *aste* on suurimman potenssin omaavan ja nollasta poikkeavan monomin $a_n x^n$ potenssi n. Tällaisen polynomin astetta merkitään $\deg(f) = n$.

ESIMERKKI 3.10. Olkoon $f(x) = x^2 + 1$. Tällöin $\deg(f) = 2$. Funktion f juuret joukossa $\mathbb{Z}_5$ ovat on $x = 1$ ja $x = 4$, sillä tällöin $f(x) = x^2 - 1 \equiv 0 \pmod{5}$.

LEMMA 3.11. *Olkoot $f, g : \mathbb{Z} \to \mathbb{Z}$ funktioita ja p alkuluku. Funktion $h(x) = f(x)g(x)$ juuret joukossa $\mathbb{Z}_p$ ovat joko funktion $f(x)$ tai $g(x)$ juuria.*

TODISTUS. Olkoon x funktion h juuri joukossa $\mathbb{Z}_p^*$, eli $h(x) \equiv 0 \pmod{p}$. Koska $h(x) = f(x)g(x)$, niin $f(x)g(x) \equiv 0 \pmod{p}$. Koska p on alkuluku, niin joko $p \mid f(x)$ tai $p \mid g(x)$. Jos $p \mid f(x)$, niin kongruenssin Määritelmän 2.1 nojalla $f(x) \equiv 0 \pmod{p}$. Jos taas $p \mid g(x)$, niin $g(x) \equiv 0 \pmod{p}$. Toisin sanottuna funktion $h(x)$ juuret ovat joko funktion $f(x)$ tai $g(x)$ juuria joukossa $\mathbb{Z}_p$. □

LEMMA 3.12. *Olkoon f asteen n polynomi kuten Määritelmässä 3.9, jolle syt$(a_n, p) = 1$. Tällöin funktiolla f on korkeintaan n juurta joukossa $\mathbb{Z}_p^*$.*

Todistus. Todistetaan väite induktion avulla. Kun $\deg(f) = 0$, niin tapaus on selvä. Olkoon $f(x) = a_n x^n + \cdots + a_1 x + a_0$, jolle $\text{syt}(a_n, p) = 1$. Jos $f(z) = 0$ siten, että $a_n \neq 0$, niin

$$\begin{aligned} f(x) &= f(x) - f(z) \\ &= a_n x^n + \cdots + a_1 x + a_0 - a_n z^n - \cdots - a_1 z - a_0 \\ &= a_n(x^n - z^n) + \cdots + a_1(x - z) \end{aligned}$$

Koska $(x^n - z^n) = (x - z)(x^{n-1} + x^{n-2}z + \cdots + xz^{n-2} + z^{n-1})$, niin termi $(x - z)$ voidaan ottaa yhteiseksi tekijäksi. Olkoon lisäksi $g(x) = (a_n(x^{n-1} + x^{n-2}z + \cdots + xz^{n-2} + z^{n-1}) + \cdots + a_2(x + z) + a_1)$, jossa $\text{syt}(a_n, p) = 1$. Tällöin edellistä jatkaen saadaan

$$\begin{aligned} f(x) &= a_n(x^n - z^n) + \cdots + a_1(x - z) \\ &= (x - z)(a_n(x^{n-1} + x^{n-2}z + \cdots + xz^{n-2} + z^{n-1}) + \cdots + a_2(x + z) + a_1) \\ &= (x - z)g(x), \end{aligned}$$

Tällöin Määritelmän 3.9 nojalla $\deg(g) = n - 1$. Oletetaan, että $f(b) \equiv 0 \pmod{p}$ siten, että $b \neq z \pmod{p}$. Tällöin ylemmän laskun nojalla $(b - z)g(b) \equiv 0 \pmod{p}$, joten koska $b \neq z \pmod{p}$, niin Lemman 3.11 nojalla $g(b) \equiv 0 \pmod{p}$. Induktio-oletuksen ja oletuksen $\text{syt}(a_n, p) = 1$ nojalla funktiolla g on korkeintaan $n-1$ juurta, joten on korkeintaan $n-1$ mahdollisuutta luvulle b. Näin ollen induktioperiaatteesta seuraa, että funktiolla f on korkeintaan n juurta joukossa $\mathbb{Z}_p^*$. □

Esimerkki 3.13. Kongruenssin juurien kanssa on oltava tarkkana, sillä ilman lisäoletusta $\text{syt}(a_n, p) = 1$ päädytään erikoiseen tilanteeseen. Olkoon esimerkiksi $g(x) = 5x$. Tällöin $\text{syt}(5, 5) = 5$ ja funktion g juuria ovat kaikki joukon $\mathbb{Z}_5$ alkiot, sillä $5x \equiv 0 \pmod{5}$ kaikilla $x \in \mathbb{Z}_5$.

Lemma 3.14. *Olkoon p alkuluku ja $d \neq 0$ kokonaisluku siten, että $d \mid (p - 1)$. Tällöin funktiolla f, jolle $f(x) = x^d - 1$, on d kappaletta juuria joukossa $\mathbb{Z}_p$.*

Todistus. Olkoon $a = \frac{p-1}{d}$. Tällöin $ad = p - 1$ ja edelleen

$$\begin{aligned} x^{p-1} - 1 &= (x^d)^a - 1 \\ &= (x^d - 1)(x^{d(a-1)} + x^{d(a-2)} + \cdots + x^d + 1) \\ &= (x^d - 1)g(x), \end{aligned}$$

jossa g on astetta $d(a-1) = da - d = p - 1 - d$ oleva polynomi. Fermatin pienen Lauseen 3.7 nojalla polynomilla $x^{p-1} - 1$ on $(p-1)$ juurta joukossa $\mathbb{Z}_p$. Lauseen 3.12 nojalla funktiolla g on korkeintaan $p-1-d$ juurta ja vastaavasti polynomilla $x^d - 1$ on korkeintaan d juurta. Lemman 3.11 nojalla polynomin $(x^d - 1)g(x)$ juuri on joko polynomin $x^d - 1$ tai $g(x)$ juuri. Funktiolla g

on siis oltava täsmälleen $p-1-d$ juurta ja tällöin polynomilla x^d-1 on täsmälleen d juurta, kuten väitettiinkin. □

LEMMA 3.15. *Olkoon $a, b \in \mathbb{Z}_p^*$ siten, että luvun a kertaluku modulo p on r ja luvun b kertaluku modulo p on s. Jos* $\text{syt}(r,s)=1$*, niin tällöin luvun ab kertaluku modulo p on rs.*

TODISTUS. Todistus mukailee lähdettä [**2**, s. 45]. Oletuksen nojalla $a^r \equiv 1 \pmod{p}$. Tällöin Lauseen 2.5 nojalla myös

$$\begin{aligned}(a^r)^s &\equiv (1)^s \\ &\equiv 1 \pmod{p}.\end{aligned}$$

Vastaavasti voidaan päätellä, että $(b^s)^r \equiv 1 \pmod{p}$. Tällöin Lauseen 2.5 nojalla

$$\begin{aligned}a^{rs}b^{rs} &\equiv (ab)^{rs} \\ &\equiv 1 \pmod{p}\end{aligned}$$

Lemman 3.5 nojalla luvun ab kertaluku modulo p on luvun rs tekijä. Merkitään tätä tekijää luvulla r_1s_1, missä $r_1 \mid r$ ja $s_1 \mid s$. Tällöin Lauseen 2.5 nojalla

$$\begin{aligned}a^{r_1s_1}b^{r_1s_1} &\equiv (ab)^{r_1s_1} \\ &\equiv 1 \pmod{p}\end{aligned}$$

Korotetaan molemmat puolet potenssiin $r_2 = \frac{r}{r_1}$. Tällöin

$$a^{r_1r_2s_1}b^{r_1r_2s_1} \equiv 1 \pmod{p}.$$

Koska $a^{r_1r_2s_1} \equiv 1 \pmod{p}$, niin Lauseen 2.5 nojalla myös $b^{r_1r_2s_1} \equiv 1 \pmod{p}$. Tällöin $s \mid r_1r_2s_1$. Koska

$$\text{syt}(s, r_1r_2) = \text{syt}(s, r) = 1,$$

niin Eukleideen Lemman 1.13 nojalla $s \mid s_1$. Koska myös $s_1 \mid s$ ja $s, s_1 > 0$, on oltava $s = s_1$. Vastaavasti osoitetaan, että $r = r_1$, joten luvun ab kertaluku modulo p on rs. □

LEMMA 3.16. *Olkoot p ja q alkulukuja siten, että $q \mid p-1$. Jos $p-1 = q^nk$, jossa $q \nmid k$, niin tällöin on olemassa kokonaisluku a, jonka kertaluku modulo p on q^n.*

TODISTUS. Tehdään vastaoletus: ei ole olemassa tällaista alkiota $a \in \mathbb{Z}_p^*$, jonka kertaluku modulo p on q^n. Olkoon $a \in \mathbb{Z}_p^*$. Koska $q^nk = p-1$ ja $p \nmid a$, niin Fermatin pienen Lauseen 3.7 nojalla

$$\begin{aligned}a^{p-1} &\equiv (a^{q^n})^k \\ &\equiv (1)^k \\ &\equiv 1 \pmod{p}.\end{aligned}$$

Lemman 3.5 nojalla luvun a^k kertaluku modulo p jakaa luvun q^n. Olkoon d alkion $a^k \in \mathbb{Z}_p^*$ kertaluku modulo p. Vastaoletuksen nojalla luku d ei voi olla q^n. Tästä seuraa, että d on jokin luvuista $1, \ldots, q^{n-1}$. Koska $d \mid q^n$, niin on olemassa $m \in \mathbb{N}$ siten, että $d = q^m$. Tällöin

$$\begin{aligned}(a^k)^{q^{n-1}} &\equiv ((a^k)^{q^m})^{q^{(n-1)-m}} \\ &\equiv 1^{q^{(n-1)-m}} \\ &\equiv 1 \pmod{p}\end{aligned}$$

Joten $(a^k)^{q^{n-1}} - 1 \equiv 0 \pmod{p}$ kaikille $a \in \mathbb{Z}_p^*$. Koska $q^{n-1}k < p-1$, niin seuraa ristiriita, sillä polynomilla jonka aste on pienempi kuin $p-1$ ei Lemman 3.14 nojalla voi olla $p-1$ kappaletta juuria joukossa $\mathbb{Z}_p^*$. □

Viimein pääsemme todistamaan primitiivisten juurten olemassaolon joukossa $\mathbb{Z}_p^*$.

LAUSE 3.17. *Olkoon p alkuluku. Tällöin joukossa $\mathbb{Z}_p^*$ on olemassa primitiivinen juuri.*

TODISTUS. Tapaus $p = 2$ on selvä, joten voidaan olettaa, että $p > 2$. Tällöin erityisesti $p-1$ ei ole alkuluku. Aritmetiikan peruslausetta 1.14 hieman soveltamalla $p-1$ voidaan esittää muodossa

$$p - 1 = p_1^{e_1} \cdot p_2^{e_2} \cdots p_n^{e_k},$$

jossa $p_1 < p_2 < \cdots < p_n$. Lemman 3.16 nojalla on olemassa alkio a_i, jonka kertaluku modulo p on $p_i^{e_i}$. Lisäksi, kun käytetään Lausetta 3.15 induktiivisesti, luvun $g = a_1 a_2 \ldots a_r$ kertaluku modulo p on $p_1^{e_1} p_2^{e_2} \cdots p_n^{e^k} = p - 1$. Tällöin luku g on Määritelmän 3.2 nojalla joukon $\mathbb{Z}_p^*$ primitiivinen juuri. □

3.3. Diskreetti logaritmi

Seuraavassa luvussa käsiteltävä Diffie–Hellmanin algoritmi pohjautuu *diskreetin logaritmin ongelmaan* [**1**, s. 62]. Tutustutaan tähän seuraavaksi tarkemmin.

MÄÄRITELMÄ 3.18. Olkoon p alkuluku, $g \in \mathbb{Z}_p^*$ sen primitiivinen juuri ja $h \in \mathbb{Z}_p^*$. Kongruenssin

$$g^x \equiv h \pmod{p}$$

toteuttavaa eksponenttia $x \in \mathbb{Z}$ kutsutaan luvun h diskreetiksi logaritmiksi kannalla g ja sitä merkitään $x = \log_g(h)$. Esitys on yksikäsitteinen jos vaaditaan, että $0 \leq x < p$.

LAUSE 3.19. *Olkoon p alkuluku ja g joukon $\mathbb{Z}_p^*$ primitiivinen juuri, sekä $h \in \mathbb{Z}_p^*$. Tällöin kongruenssilla*

$$g^x \equiv h \pmod{p}$$

on aina olemassa yksikäsitteinen ratkaisu x, jolle $1 \leq x \leq p-1$.

TODISTUS. Lauseen 3.17 nojalla joukossa $\mathbb{Z}_p^*$ on aina olemassa primitiivinen juuri ja lisäksi Lauseen 3.6 nojalla sen moninkertojen avulla pystytään esittämään kaikki joukon $\mathbb{Z}_p^*$ alkiot, kun $1 \leq x \leq p-1$. Lisäksi nämä moninkerrat olivat erillisiä, joten tällöin myös kongruenssilla $g^x \equiv h \pmod{p}$ on aina olemassa yksikäsitteinen ratkaisu. □

Määritelmässä 3.18 esiintyvän eksponentin x etsimistä kutsutaan *diskreetin logaritmin ongelmaksi*. Diskreetin logaritmin ongelma on hankala ratkaista. Tällä hetkellä ei ole olemassa yleistä ja tehokasta menetelmää, jolla pystyttäisiin laskemaan diskreettejä logaritmeja. Ainoat tunnetut menetelmät diskreetin logaritmin laskemiseen perustuvat niin kutsuttuihin raa'an voiman menetelmiin, joissa lähdetään järjestäen kokeilemaan eksponentteja, kunnes ratkaisu löydetään. Kun alkuluvuksi p valitaan suuri luku, niin raa'alla voimalla laskettuna laskuun kulunut aika on käytännössä niin suuri, että nykyisillä tietokoneilla ei pystytä ratkaisemaan ongelmaa. Tämä on perusteena luvun 4 Diffie–Hellmanin menetelmän toimivuudelle ja sille, että sitä on vaikea purkaa. Havainnollistetaan tätä vielä esimerkillä:

ESIMERKKI 3.20. Tarkastellaan kongruenssia $2^x \equiv 3 \pmod{29}$. Luku 29 on alkuluku ja 2 on sen primitiivinen juuri, joten Lauseen 3.17 nojalla kongruenssilla on ratkaisu. Kokeilemalla huomataan, että jos $x = 5$, niin $2^5 = 32$ ja tällöin kongruenssi pätee.

Tarkastellaan seuraavaksi yhtälöä $6^x \equiv 857 \pmod{2791}$. Tässä 2791 on alkuluku ja 6 on sen primitiivinen juuri. Huomataan, että kokeiluja täytyy tehdä jo melko paljon enemmän oikean eksponentin x löytämiseksi. Tämän kongruenssin ratkaisee muun muassa $x = 4046$, sillä $6^{4046} \equiv 857 \pmod{2791}$. Esimerkissä 4.9 ja liitteessä 2 ratkaistaan tämä sama kongruenssi käyttäen *Shankin algoritmia*.

LUKU 4

Diffie–Hellmanin menetelmä

4.1. Taustatietoa

Diffie–Hellmanin avaimenvaihdoksi kutsuttu menetelmä juontaa juurensa 1970-luvulle. Whitfield Diffy ja Martin Hellman julkaisivat vuonna 1976 kryptografialle tärkeää suuntaa antavan artikkelin "*New Directions in Cryptography*" [**5**]. Mielenkiintoinen taustaseikka on se, että brittiläisen tiedustelupalvelu GCHQ:n (*Government Communications Headquarters*) matemaatikko Malcolm J. Williamson oli työskennellyt samaisen menetelmän kanssa jo muutamaa vuotta aiemmin [**6**]. Kyseinen projekti oli ollut salainen, joten siitä ei oltu julkaistu tietoa ennen kuin Diffie ja Hellman sattuivat keksimään saman menetelmän muutamaa vuotta myöhemmin ja täten myös julkaisemaan sen. Menetelmään tarvittavat taustatiedot löytyvät luvuista 2 ja 3. Tutustutaan ensin algoritmiin ja sitten perustellaan sen toimivuus.

4.2. Algoritmi

Algoritmi 4.1 (Diffie–Hellmanin algoritmi). Diffie–Hellmanin menetelmä toimii seuraavanlaisen algoritmin mukaisesti henkilöiden A ja B välillä:

(1) Valitaan suuri alkuluku p ja sen primitiivinen juuri g joukossa $\mathbb{Z}_p^*$. Luvut p ja g jaetaan julkisesti molemmille osapuolille A ja B.
(2) A valitsee salaisen kokonaisluvun $a \in \mathbb{Z}$, jota hän ei paljasta muille. Vastaavasti B valitsee oman salaisen kokonaisluvun $b \in \mathbb{Z}$.
(3) A laskee laskutoimituksen $\hat{A} \equiv g^a \pmod{p}$.
(4) B laskee laskutoimituksen $\hat{B} \equiv g^b \pmod{p}$.
(5) A lähettää tuloksen $\hat{A}$ B:lle ja B lähettää tuloksen $\hat{B}$ A:lle.
(6) A laskee laskutoimituksen $s \equiv \hat{B}^a \pmod{p}$ ja B laskee laskutoimituksen $s \equiv \hat{A}^b \pmod{p}$.

Merkitsimme syystäkin molempien A:n ja B:n laskutoimitusten lopputulosta symbolilla s. A:lla ja B:llä on nimittäin algoritmin päätteeksi sama salausavain s, vaikka he eivät tiedä toistensa valitsemia salaisia lukuja. Todistetaan vielä yllä olevan algoritmin toimivuus yleisessä tapauksessa:

Lause 4.2. *Olkoot p ja g sekä a ja b kuten algoritmissa 4.1. Tällöin*

$$(g^a \ (\mathrm{Mod}\ p))^b \equiv (g^b \ (\mathrm{Mod}\ p))^a \pmod{p}.$$

TODISTUS. Lauseen 2.7 nojalla $(g^a \ (\text{Mod } p))^b \equiv g^{ab} \ (\text{Mod } p) \ (\text{mod } p)$. Koska $g^{ab} = g^{ba}$, niin käyttämällä Lausetta 2.7 uudelleen, saadaan

$$g^{ba} \ (\text{Mod } p) \equiv (g^b \ (\text{Mod } p))^a \ (\text{mod } p).$$

Joten $(g^a \ (\text{Mod } p))^b \equiv (g^b \ (\text{Mod } p))^a \ (\text{mod } p)$. □

HUOMAUTUS 4.3. Tässä tutkielmassa valitaan diskreetin logaritmin ongelmaan ja Diffie–Hellmanin algoritmiin joukon $\mathbb{Z}_p^*$ primitiivinen juuri $g \in \mathbb{Z}_p^*$. Diffie–Hellmanin algoritmi toimii myös käyttämättä primitiivistä juurta, mutta tällöin saavutettaisiin heikompi turvallisuus. Tämä johtuu siitä, että primitiivisen juuren g moninkerrat g^i, $1 \leq i \leq p-1$, virittävät koko joukon $\mathbb{Z}_p^*$. Jos luvun g paikalle valittaisiin jokin joukon $\mathbb{Z}_p^*$ alkio, joka ei ole primitiivinen juuri, niin tällöin tämän alkion potensseilla ei voisi esittää kaikkia joukon $\mathbb{Z}_p^*$ alkioita. Tämä pienentäisi mahdollisten ratkaisujen lukumäärää, jolloin kokeilemalla eri ratkaisuja, kokeilijalla olisi suurempi todennäköisyys arvata oikein. Lisäksi primitiivisen juuren valitseminen varmistaa sen, että diskreetin logaritmin ongelmalla on aina olemassa ratkaisu.

Tehdään algoritmista vielä esimerkki kahdella konkreettisella luvulla. Valitaan tähän esimerkkiin pienet luvut, jotta laskeminen olisi helpompaa. Todellisuudessa näin pienillä luvuilla laskettaessa ei saavutettaisi minkään tasoista tietoturvaa.

ESIMERKKI 4.4. Valitaan $p = 71$ ja $g = 33$. Tällöin p on alkuluku ja g on alkuluvun p primitiivinen juuri joukossa $\mathbb{Z}_{71}^*$, sillä luvun g moninkertojen avulla voidaan esittää kaikki joukon $\mathbb{Z}_{71}^*$ alkiot. Olkoot esimerkkihenkilöt Aava ja Brutus. Aava valitsee salaiseksi luvukseen $a = 3$ ja Brutus valitsee salaiseksi luvukseen $b = 4$. Tällöin Aava saa lähetettäväksi luvukseen

$$33^3 \ (\text{Mod } 71) = 35937 \ (\text{Mod } 71) = 11$$

ja vastaavasti Brutus saa lähetettäväksi luvukseen

$$33^4 \ (\text{Mod } 71) = 1185921 \ (\text{Mod } 71) = 8.$$

Seuraavaksi Aava ja Brutus vaihtavat keskenään lukuja. Aava tietää nyt siis luvun 8 ja Brutus luvun 11. Seuraavaksi molemmat laskevat ylemmän laskutoimituksen vastaavasti, mutta uusilla luvuillansa:

$$\text{Aava: } 8^3 \ (\text{Mod } 71) = 512 \ (\text{Mod } 71) = 15$$

$$\text{Brutus: } 11^4 \ (\text{Mod } 71) = 14641 \ (\text{Mod } 71) = 15.$$

Molemmat pääsivät samaan lopputulokseen $s = 15$. Aava ja Brutus eivät kuitenkaan missään vaiheessa vaihtaneet keskenään salaisia lukujaan ja kummallakaan ei siis ole mitään tietoa toistensa salaisista luvuista. Silti he molemmat jakavat nyt saman salaisuuden, eli luvun 15.

HUOMAUTUS 4.5. Liitteessä 1 on toinen esimerkki Maxima-laskentaohjelmistolla lasketuna.

Mikä yllä olevan Esimerkin 4.4 tiedonvaihdosta oli siis salaista? Pohditaan vielä seuraavanlaista tilannetta: Rasmus on ulkopuolinen henkilö, joka kuulee kun Aava ja Armas huutavat lukuja toisilleen. Tällöin Rasmus kuulee siis heti aluksi luvut $p = 71$ ja $g = 33$. Tämän lisäksi Rasmus kuulee laskutoimitusten jälkeen vaihdetut luvut 8 ja 11. Tässä tuleekin eteen pieni ongelma. Rasmus tietää seuraavat laskutoimitukset:

$$33^a \equiv 11 \pmod{71} \qquad \text{ja}$$
$$33^b \equiv 8 \pmod{71}.$$

Rasmuksen olisi selvitettävä molempien salainen eksponentti laskeakseen toimituksen

$$33^{ab} \pmod{p}.$$

Rasmuksen olisi siis ratkaistava diskreetin logaritmin ongelma 3.18. Riittävän suurella alkuluvulla p tämän ratkaiseminen olisi niin työlästä, että Rasmus ei saisi selville salaista lukua s. Salauksen purkamisen kannalta olisi mielekästä, jos työmäärää saataisiin hieman vähennettyä. Tämä onnistuu, mutta tällöinkään salausta ei saada purettua kovin nopeasti. Tutustutaan seuraavaksi tämänkaltaiseen algoritmiin.

4.3. Shankin algoritmi

Seuraava algoritmi pohjautuu viitteeseen [**1**, ss. 80–81]. Shankin algoritmilla pystytään nopeuttamaan diskreetin logaritmin ongelman ratkaisua. Algoritmi perustuu niin sanottuun *yhteentörmäykseen*, jossa kahdesta lukulistasta etsitään vastaavuuksia. Operoidaan edelleen joukossa $\mathbb{Z}_p^*$, jossa p on alkuluku.

HUOMAUTUS 4.6. Shankin algoritmissa käytetään niin kutsuttua *lattiafunktiota* $\lfloor x \rfloor$. Merkintä tarkoittaa sitä, että reaaliluku x pyöristetään lähimpään pienempään kokonaislukuun.

ESIMERKKI 4.7. Tutkitaan lattiafunktion toimintaa esimerkkien avulla:

$$\lfloor \sqrt{7} \rfloor = 2, \lfloor \sqrt{8} \rfloor = 2, \lfloor 3 \rfloor = 3$$

LAUSE 4.8 (Shankin algoritmi). *Olkoon g joukon $\mathbb{Z}_p^*$ primitiivinen juuri, jonka kertaluku modulo p on $p-1$. Tällöin seuraava algoritmi ratkaisee diskreetin logaritmin ongelman $g^x \equiv h \pmod{p}$, missä $h \not\equiv 0 \pmod{p}$:*

(1) Olkoon $n = 1 + \lfloor \sqrt{p-1} \rfloor$. Tällöin erityisesti $n > \sqrt{p-1}$.

(2) Tehdään kaksi listaa:

$$L_1 = (g \ (Mod\ p), g^2 \ (Mod\ p), g^3 \ (Mod\ p), \ldots, g^n \ (Mod\ p))$$
$$L_2 = (h \ (Mod\ p), h \cdot u \ (Mod\ p), h \cdot u^2 \ (Mod\ p), \ldots, h \cdot u^n \ (Mod\ p)),$$

jossa u on luvun g^n modulaari käänteisluku, eli $u \cdot g^n \equiv 1 \pmod{p}$.

(3) Etsitään listojen luvuista kaksi samaa lukua $g^i \ (Mod\ p) = h \cdot u^j \ (Mod\ p)$, jossa $1 \leq i \leq n$ ja $0 \leq j \leq n$.

(4) Tällöin $x = i + jn$ on kongruenssin $g^x \equiv h \pmod{p}$ ratkaisu.

TODISTUS. Todistetaan ensin, että algoritmin listoissa L_1 ja L_2 on aina olemassa kaksi samaa lukua. Lauseen 3.19 nojalla kongruenssilla $g^x \equiv h \pmod{p}$ on olemassa ratkaisu x siten, että $1 \leq x \leq p-1$ ja lisäksi Lauseen 3.6 nojalla luvut g^x ovat eri lukuja, kun $x = 1, \ldots, p-1$. Siten voidaan valita $x = nq + r$ siten, että $0 \leq r < n$ ja $0 \leq nq \leq p-1$. Erityisesti $0 \leq q < n$, sillä muuten olisi $nq \geq n^2 > p-1$.

Kongruenssille $g^x \equiv h \pmod{p}$ saadaan nyt

$$\begin{aligned} & g^x \equiv h \pmod{p} \\ \Leftrightarrow \quad & g^{nq+r} \equiv h \pmod{p} \\ \Leftrightarrow \quad & g^{nq} \cdot g^r \equiv h \pmod{p}. \end{aligned}$$

Koska $u \cdot g^n \equiv 1 \pmod{p}$, niin

$$\begin{aligned} & g^{nq} \cdot g^r \equiv h \cdot 1 \pmod{p} \\ \Leftrightarrow \quad & g^{nq} \cdot g^r \equiv h \cdot (1)^q \pmod{p} \\ \Leftrightarrow \quad & g^{nq} \cdot g^r \equiv h \cdot (u \cdot g^n)^q \pmod{p} \\ \Leftrightarrow \quad & g^{nq} \cdot g^r \equiv h \cdot u^q \cdot g^{nq} \pmod{p} \end{aligned}$$

Koska p on alkuluku, niin $\text{syt}(g^{nq}, p) = 1$. Tällöin Lauseen 2.14 nojalla termi g^{nq} voidaan jakaa pois kongruenssin molemmilta puolilta. Tällöin

$$g^r \equiv h \cdot u^q \pmod{p}$$

Nyt yllä olevan kongruenssin vasen puoli on muotoa g^r ja oikea puoli muotoa $h \cdot u^q$. Koska nämä ovat nyt samat modulo p ja ne olivat eri listoissa, niin listoissa L_1 ja L_2 on aina olemassa kaksi samaa lukua.

Osoitetaan vielä, että ratkaisu on muotoa $x = i + jn$. Olkoot kaksi tällaista indeksiä $1 \leq i \leq n$ ja $0 \leq j \leq n$ siten, että $g^i \equiv h \cdot u^j \pmod{p}$. Tällöin edelleen ylempiä laskuja peruuttamalla toiseen suuntaan $g^{i+jn} \equiv h \pmod{p}$, joten $x = i + jn$ on diskreetin logaritmin ongelman ratkaisu. □

ESIMERKKI 4.9. Ratkaistaan Esimerkin 3.20 diskreetin logaritmin ongelma käyttäen Shankin algoritmia. Laskut löytyvät liitteistä Maxima-laskentaohjelmistolla laskettuna. Olkoon kongruenssi

$$6^x \equiv 857 \pmod{2791}.$$

Tällöin Lauseessa 4.8 on $g = 6$, $h = 857$ ja $p = 2791$. Koska g on primitivinen juuri ja p on alkuluku, niin luvun g kertaluku on $2791 - 1 = 2790$.

(1) Olkoon $n = 1 + \lfloor\sqrt{2790}\rfloor = 1 + 52 = 53$.

(2) Koska $6^{53} \cdot 2036 \equiv 1 \pmod{2791}$, niin $u = 2036$. Luodaan listat L_1 ja L_2 laskemalla luvut $6^i \pmod{2791}$ ja $857 \cdot 2036^j \pmod{2791}$, kun $1 \leq i \leq n$ ja $0 \leq j \leq n$.

(3) Liitteessä 2 Maxima -laskentaohjelmistolla löytyi $g^i \pmod{p} = h \cdot u^j \pmod{p}$, kun $i = 37$ ja $j = 23$. Tällöin siis

$$6^{37} \equiv 2446 \pmod{2791}$$

ja

$$857 \cdot 2036^{23} \equiv 2446 \pmod{2791}.$$

(4) Shankin algoritmin nojalla $x = 37 + 23 \cdot 53 = 1256$ ratkaisee diskreetin logaritmin ongelman. Muitakin ratkaisuja on olemassa ja ne ovatkin itse asiassa $x = 1256 + 2791 \cdot m$, kun $m \in \mathbb{Z}$. Tämä johtuu siitä, että luvun 2791 moninkerrat voidaan aina supistaa pois ratkaisusta vastaavasti kuten algoritmin todistuksessa tehtiin.

Shankin algoritmi ratkaisi ongelman hieman helpommin kuin kokeilemalla kaikki eksponentit. Täytyy huomioida, että ratkaisun $x = 1256$ saamiseksi listojen L_1 ja L_2 alkioita verrattiin toisiinsa $23 \cdot 37 = 851$ kertaa. Algoritmi perustuu kokeilemiseen listoja läpi käydessä, joten sekin on raa'an voiman menetelmä joka ei ole vielä tarpeeksi tehokas tuottamaan ratkaisua nopeasti ja tehokkaasti.

LUKU 5

RSA-menetelmästä

Diffie–Hellmanin menetelmällä pystyttiin luomaan molemmille yhteinen salausavain, mutta tiedonsiirto tai viestien lähettäminen sen avulla ei onnistunut. RSA-menetelmällä tätä vastoin pystytään lähettämään viestejä osapuolten välillä turvallisesti. RSA-menetelmän synty ajoittuu samalle aikakaudelle Diffie–Hellmanin menetelmän kanssa. Algoritmi julkaistiin vuonna 1978 Ron Rivestin, Adi Shamirin ja Leonard Adlemanin toimesta [**8**]. Nimi *RSA* tulee edellä mainittujen henkilöiden sukunimien alkukirjaimista. Tavoitteena oli luoda tapa siirtää tietoa siten, että salaukseen käytettävä avain on julkinen, mutta purkamiseen tarvittava salasana on henkilökohtainen. Tiedot vastaanottava osapuoli voi julkaista salaamiseen käytettävän salausavaimen kaikille, mutta hän pitää henkilökohtaisen avaimen itsellään. Näin vastaanottajalle voidaan lähettää tietoa, jota vain vastaanottaja voi lukea. RSA-menetelmä onkin esimerkki epäsymmetrisestä salausmenetelmästä.

5.1. RSA-menetelmä

Tämä luku pohjautuu viitteeseen [**1**]. RSA-menetelmä on yksisuuntainen funktio. RSA-menetelmän toimivuus perustuu siihen, että ensin lasketaan kahden suuren alkuluvun tulo, jonka avulla lähetettävä viesti salataan. Jos haluaa palata salauksen alkutilanteeseen, tämä tulo on jaettava tekijöihin. Tekijöihin jakaminen on vaikeaa, joten tämä ei käytännössä onnistu tämänhetkisillä menetelmillä ilman lisätietoja käytetyistä alkuluvuista.

Algoritmi 5.1 (RSA). RSA-menetelmä toimii seuraavan algoritmin mukaisesti henkilöiden A ja B välillä:

(1) A valitsee kaksi suurta alkulukua p ja q ja kokonaisluvun e siten, että $(p-1)(q-1)$ ja e ovat suhteellisia alkulukuja keskenään.
(2) A ratkaisee luvun d kongruenssista $de \equiv 1 \pmod{(p-1)(q-1)}$.
(3) A julkaisee luvut $N = pq$ ja e.
(4) B valitsee viestin m, jossa $1 \leq m \leq N-1$.
(5) B lähettää salatun viestin $c \equiv m^e \pmod{N}$ A:lle.
(6) A avaa salatun viestin laskemalla kongruenssin $m' \equiv c^d \pmod{N}$.

Huomautus 5.2. Algoritmin kohdassa (2) tällainen luku d on aina olemassa. Tulos on todistettu Lemmassa 2.16. Lisäksi jos luku e valitaan siten, että se on suuri alkuluku, niin luvut e ja $(p-1)(q-1)$ ovat aina suhteellisia alkulukuja.

Toteutetaan algoritmista esimerkki pienillä luvuilla, jotta lukijan on helpompi seurata menetelmän toimintaa. Salausmenetelmä ei ole turvallinen näin pienillä luvuilla.

Esimerkki 5.3. Aava haluaa vastaanottaa viestejä Brutukselta siten, että Rasmus ei pääse lukemaan viestejä. Aava ja Brutus sopivat, että he käyttävät viestin salaamisessa RSA-menetelmää. Seurataan algoritmia askel kerrallaan:

(1) Aava valitsee alkuluvut $p = 601$ ja $q = 307$. Seuraavaksi Aava etsii luvun e siten, että $(p-1)(q-1)$ ja e ovat suhteellisia alkulukuja keskenään. Lasketaan tulo

$$\begin{aligned}(p-1)(q-1) &= 600 \cdot 306 \\ &= 183600.\end{aligned}$$

Aava valitsee luvuksi $e = 11$. Koska $e = 11$ on alkuluku, niin Määritelmän 1.11 nojalla e ja $(p-q)(q-1)$ ovat suhteellisia alkulukuja keskenään.

(2) Aava ratkaisee luvun d kongruenssista $de \equiv 1 \pmod{(p-1)(q-1)}$. Kun $n \in \mathbb{N}$, niin kongruenssi voidaan kirjoittaa muodossa

$$11d + 183600n = 1.$$

Ratkaistaan luku d käyttäen Laajennettua Eukleideen algoritmia 1.18. Aluksi

$$\begin{aligned}183600 &= 11 \cdot 16690 + 10 \\ 11 &= 1 \cdot 10 + 1.\end{aligned}$$

Peruutetaan algoritmia sijoittelemalla, jotta saadaan selvitettyä luvut d ja n:

$$\begin{aligned}11 - 1 \cdot 10 &= 1 \\ 11 - 1 \cdot (183600 - 11 \cdot 16690) &= 1 \\ 11 + 11 \cdot 16690 - 183600 &= 1 \\ 11 \cdot 16691 + (-1) \cdot 183600 &= 1.\end{aligned}$$

Joten saatiin $d = 16691$.

(3) Aava laskee vielä luvun $N = pq = 601 \cdot 307 = 184507$. Nyt Aava julkaisee luvut $e = 11$ ja $N = 184507$. Luvut $p = 601$ ja $q = 307$ pysyvät Aavan omana tietona.
Nyt Aava ja Brutus ovat valmiita vaihtamaan tietoja keskenään. Aava on unohtanut piin likiarvon, mutta hän ei halua, että Rasmus saa tietää unohduksesta. Brutus auttaa Aavaa ja lähettää tälle muutaman numeron piistä ilman pilkkua, jotta Rasmus ei osaisi epäillä mitään. Brutus haluaa lähettää luvun $m = 31415$ Aavalle.

(4) Brutus lähettää Aavalle salatun viestin $c \equiv 31415^{11} \pmod{184507}$. Tällöin $c = 150191$.

(5) Aava avaa viestin laskemalla $m' \equiv 150191^{16691} \pmod{184507}$. Tällöin $m' = m = 31415$.

HUOMAUTUS 5.4. Esimerkin 5.3 laskut ovat liitteessä 3 Maximalla toteutettuna.

5.2. Miksi RSA toimii?

Pohditaan seuraavaksi, minkä vuoksi RSA-menetelmä toimii siten, että viestejä saadaan välitettyä salassa. Näytetään, miksi salausavain valitaan tällä tavalla ja mitä hyötyjä siitä saadaan. Pohditaan myös hieman salauksen purkamista ulkopuolisen henkilön näkökulmasta. Todistetaan ensin lause, joka takaa viestin salausavaimen toimivuuden viestiä salatessa ja purkaessa.

LAUSE 5.5. *Olkoon luvut p, q, d, e kuten RSA-menetelmän algoritmissa. Tällöin*

$$(m^e)^d \equiv m \ (mod\ N).$$

TODISTUS. Salausavain d luotiin kongruenssin $de \equiv 1 \pmod{(p-1)(q-1)}$ avulla. Lausekkeesta seuraa, että $(p-1)$ ja $(q-1)$ jakavat luvun $de - 1$. Tällöin olemassa kokonaisluvut x ja y siten, että

$$\begin{aligned} de - 1 &= x(p-1) \\ &= y(q-1). \end{aligned}$$

Todistetaan kaksi eri tapausta, jotka yhdistämällä lause saadaan todistettua.

(1) Näytetään, että kongruenssi pätee modulo p, eli $m^{de} \equiv m \pmod{p}$. Jos $m \equiv 0 \pmod{p}$, niin m on jaollinen luvulla p, ja edelleen m^{de} on jaollinen luvulla p. Tällöin myös $m^{de} \equiv 0 \pmod{p}$, joten Lauseen 2.4 nojalla $m^{de} \equiv m \pmod{p}$. Jos taas $m \not\equiv 0 \pmod{p}$, niin on olemassa kokonaisluku x siten, että $ed - 1 = x(p-1)$. Tällöin

$$\begin{aligned} m^{de} &= m \cdot m^{ed-1} \\ &= m \cdot m^{x(p-1)} \\ &= m \cdot (m^{p-1})^x. \end{aligned}$$

Fermatin pienen Lauseen 3.7 nojalla saadaan

$$\begin{aligned} m \cdot (m^{p-1})^x &\equiv m \cdot 1^x \\ &\equiv m \pmod{p}. \end{aligned}$$

(2) Kongruenssi pätee vastaavasti modulo q, eli $m^{ed} \equiv m \pmod{q}$. Todistus tehdään vastaavasti kuin luvulle p.

Tällöin $m^{ed} - m$ on jaollinen luvulla p ja jaollinen luvulla q. Tästä saadaan edelleen, että $p \mid q \cdot \frac{m^{ed}-m}{q}$. Koska $\text{syt}(p, q) = 1$, niin Eukleideen Lemman 1.13 nojalla $p \mid \frac{m^{ed}-m}{q}$. Tästä seuraa, että $pq \mid m^{ed} - m$. Koska $N = pq$, niin edelleen Kongruenssin Määritelmän 2.1 nojalla $(m^e)^d \equiv m \pmod{N}$. □

Miten tämä Lause 5.5 sitten perustelee sen, että viestit pysyvät salassa ja molemmilla henkilöillä on tietojen vaihtamisen jälkeen samat viestit? RSA-algoritmin askelissa (5) ja (6) todettiin, että henkilö B lähettää henkilölle A viestin $c \equiv m^e \pmod{N}$, Tällöin Lauseen 2.5 kohdan (3) nojalla $c^d \equiv (m^e)^d \pmod{N}$. Lauseen 5.5 nojalla $(m^e)^d \equiv m \pmod{N}$. Joten kun henkilö A ratkaisee kongruenssin

$$\begin{aligned} m' &\equiv c^d \\ &\equiv (m^e)^d \\ &\equiv m \pmod{N}, \end{aligned}$$

hän saa laskettua alkuperäisen henkilön B lähettämän viestin itselleen, kunhan $m < N$. Vain henkilö A tiesi luvun d, joten kukaan muu ei viestiä pysty avaamaan.

RSA-menetelmän turvallisuus perustuu kahteen ongelmaan. Ensimmäinen on suurten kokonaislukujen tekijöihinjako salausavainta luodessa ja toinen on yllä olevan Lauseen 5.5 kongruenssin ratkaiseminen. Molemmat ongelmat ovat vaikeita ratkaista ja toistaiseksi ei ole olemassa tunnettua algoritmia näiden ongelmien ratkaisemiseksi nopeasti.

Pohditaan seuraavaksi kokonaislukujen tekijöihinjaon ongelmaa. RSA-menetelmän luku $N = pq$ on vaikea jakaa tekijöihin. Koska luvut p ja q ovat alkulukuja, niin Aritmetiikan peruslauseen 1.14 nojalla luku N voidaan esittää vain näiden kahden alkuluvun tulona yksikäsitteisesti. Otetaan esimerkiksi luku $N = 993982939$. Luku N on nyt kahden alkuluvun tulo, jolle $2^{29} < N < 2^{30}$. Luku N on lukujen $p = 9491$ ja $q = 104729$ tulo. Luku p on järjestykseltään 1176. alkuluku ja 104729 on järjestykseltään 10000. alkuluku. Toisin sanottuna raa'alla voimalla laskettaessa pitäisi laskea korkeintaan $1176 \cdot 10000 = 11760000$ alkulukujen tuloa, ennen kuin osutaan oikeaan tulokseen. Kun lukua N kasvatetaan siten, että se on esimerkiksi kokoluokkaa 2^{252}, tarvitaan mahdollisia laskutoimituksia vielä huomattavasti paljon enemmän. On kuitenkin olemassa menetelmiä, joiden avulla tekijöihinjakoa saadaan hieman helpotettua joissakin tapauksissa.

5.3. Pollardin menetelmä

RSA-menetelmässä julkaistiin luku $N = pq$ siten, että p ja q ovat alkulukuja. Emme kuitenkaan tiedä, mitkä nämä kaksi alkulukua ovat, ennen kuin jaamme luvun N tekijöihin. Koska tiedämme, että luvun N tekijöitä on kaksi kappaletta, riittää saada selville edes toinen

niistä. Kun toinen tekijä on saatu selville, toinen voidaan laskea jakamalla luku N tällä tekijällä. Tutustutaan seuraavaksi menetelmään, josta on apua joidenkin kokonaislukujen tekijöihinjaon kanssa. Menetelmä perustuu kokeilemiseen, joten se on luokiteltavissa raa'an voiman menetelmäksi. Kehitellään ensin pohja menetelmälle tässä tutkielmassa esitettyjen tietojen avulla.

Valitaan satunnaisesti kokonaisluku a, jolle $1 < a < N$. Tällöin voimme selvittää luvun $\text{syt}(a, N)$, jonka on oltava joko 1 tai jokin luvun N alkulukutekijä. Joten jos $\text{syt}(a, n) \neq 1$, niin tällöin suurin yhteinen tekijä on luvun N tekijä. Suurin osa kokonaisluvuista eivät kuitenkaan ole luvun N tekijöitä, joten suurella osalla luvuista $a < N$ on $\text{syt}(a, N) = 1$. Miten sitten löytää sellainen luku, jolla olisi yhteinen tekijä luvun N kanssa? Esitetään ja todistetaan ensin lause, joka hieman auttaa tällaisen luvun etsinnässä.

LAUSE 5.6. *Olkoon $N = pq$, jossa p ja q ovat alkulukuja. Olkoon luku L jokin sellainen kokonaisluku, että $(p-1) \mid L$ ja a kokonaisluku, jolle $1 < a < N$. Tällöin syt$(a^L - 1, N)$ on joko p tai N.*

TODISTUS. Koska $(p-1) \mid L$, niin Määritelmän 1.1 nojalla on olemassa kokonaisluku k siten, että $L = k(p-1)$. Fermatin pieni lause 3.7 toteaa seuraavan:

$$a^{p-1} \equiv \begin{cases} 1 \pmod{p} & \text{jos } p \nmid a, \\ 0 \pmod{p} & \text{jos } p \mid a. \end{cases}$$

Voidaan olettaa, että $p \nmid a$, sillä jos $p \mid a$, niin a on luvun N tekijä. Fermatin pientä lausetta 3.7 hyödyntämällä voidaan päätellä seuraavasti:

$$\begin{aligned} a^L &\equiv a^{(p-1)k} \\ &\equiv (1)^k \\ &\equiv 1 \pmod{p} \end{aligned}$$

Tällöin kongruenssin Määritelmän 2.1 nojalla $p \mid a^L - 1$. Tämä tarkoittaa sitä, että syt$(a^L - 1, N)$ on joko p tai N. □

Lauseen 5.6 perusteella voimme saada toisen luvun N tekijöistä selville, jos löydämme sellaisen kokonaisluvun L, että $(p-1) \mid L$. Jos syt$(a^L - 1, N)$ ei ole N tai 1, olemme löytäneet luvun p, eli toisen luvun N alkulukutekijöistä ja pystymme selvittämään luvun $q = \frac{N}{p}$.

Keksimme siis tavan, jolla voidaan saada selvitettyä toinen luvun N tekijöistä. Herääkin vielä kysymys siitä, mistä tällainen luku L sitten löydetään? Suoraa tapaa tähän ei ole, joten lukua L tulee etsiä kokeilemalla. Lukujen kertomat sisältävät paljon tekijöitä, joten on

kannattavaa kokeilla Lauseen 5.6 pohdintaa luvuilla $L = 1!, 2!, 3!, \ldots$, kunnes jollakin näistä luvuista $\text{syt}(a^L - 1, N)$ on joko p tai q. Tästä johtuvaa menetelmää kutsutaan Pollardin $p-1$ -menetelmäksi. Esitetään se seuraavaksi algoritmina:

ALGORITMI 5.7 (Pollardin $p-1$ -menetelmä). Olkoon N RSA-menetelmässä julkaistu kahden suuren alkuluvun tulo. Pollardin $p-1$ -menetelmä toimii seuraavan algoritmin mukaisesti:

(1) Valitse $a = 2$.
(2) Jos $\text{syt}(a, p) \neq 1$ ja $\text{syt}(a, p) < N$, niin löysit luvun N tekijän.
(3) Olkoon $j = 2$.
(4) Laske $a_j \equiv a^{j!} \pmod{N}$.
(5) Laske $\text{syt}(a_j - 1, N)$.
(6) Jos $1 < \text{syt}(a_j - 1, N) < N$, niin löysit luvun N tekijän.
(7) Jos $\text{syt}(a_j - 1, N) = N$, niin kasvata lukua $a \to a + 1$ ja toista kohdasta 2.
(8) Kasvata lukua $j \to j + 1$. Toista kohdasta 4.

HUOMAUTUS 5.8. Jos esimerkiksi $j = 50$ ja $a = 2$, niin luku $a^{j!} = 2^{50!}$ olisi valtava. Lukua $a^{j!}$ ei onneksi tarvitse laskea, sillä riittää laskea $a^{j!} - 1$ (Mod N). Tämä on huomattavasti helpompaa, sillä pääsemme laskemaan lukua N pienemmillä luvuilla. Lisäksi, koska algoritmin aiemmassa välivaiheessa on laskettu $a^{(j-1)!}$, niin seuraavassa välivaiheessa riittää, kun korotetaan tämä potenssiin j, jolloin $(a^{(j-1)!})^j = a^{j!}$. Jokaisella algoritmin kierroksella ei siis tarvitse laskea kertomaa uudelleen.
Algoritmi myös päättyy aina, joskin se voi olla joskus todella hidasta. Vaikka kaikilla tähän mennessä tehdyillä luvun a valinnoilla olisi päädytty tilanteeseen $\text{syt}(a^L - 1, N) = N$, niin koska lukua a kasvatetaan näissä tilanteissa, niin jossain vaiheessa a on toinen luvun N tekijöistä.

ESIMERKKI 5.9. Valitaan RSA-menetelmän luvuksi $N = 8083$. RSA-menetelmän toiminnasta tiedetään, että luku N on alkulukujen p ja q tulo. Käytetään Pollardin $p-1$ -menetelmää näiden alkulukujen selvittämiseksi. Seurataan algoritmia askel kerrallaan.

(1) Valitaan $a = 2$.
(2) Olkoon $j = 2$.
(3) Lasketaan $2^{j!} \equiv a_j \pmod{8083}$.

(4) Lasketaan $\text{syt}((a_j - 1), 8083)$ kunnes $1 < \text{syt}((a_j - 1), 8083) < N$:

$$\begin{aligned} 2^{2!} &\equiv 4 (\text{mod } 8083), \quad \text{syt}(3, 8083) = 1 \\ 2^{3!} &\equiv 64 (\text{mod } 8083), \quad \text{syt}(63, 8083) = 1 \\ 2^{4!} &\equiv 4991 (\text{mod } 8083), \quad \text{syt}(4990, 8083) = 1 \\ &\vdots \\ 2^{16!} &\equiv 2379 (\text{mod } 8083), \quad \text{syt}(2378, 8083) = 1 \\ 2^{17!} &\equiv 138 (\text{mod } 8083), \quad \text{syt}(137, 8083) = 137 \end{aligned}$$

(5) Löydettiin luku $\text{syt}(137, 8083) = 137$ siten, että $1 < 137 < N$.
Huomataan, että $\frac{8083}{137} = 59$. Toisin sanottuna löydettiin alkuluvut $p = 137$ ja $q = 59$ siten, että $N = 137 \cdot 59 = 8083$.

Liitteet

Alla olevista liitteistä löytyy esimerkkejä tutkielmassa esitetyistä salausmenetelmistä Maxima -laskentaohjelmistolla toteutettuna. Maxima on yksi käytössä olevista sähköisten ylioppilaskirjoitusten laskentaohjelmistoista. Maxima on komentorivipohjainen ohjelmisto ja sen avulla pystyy laskemaan sekä symbolisesti että numeerisesti. Käyttöliittymä ja syntaksi ovat yksinkertaiset, joten ohjelma on melko helposti lähestyttävissä myös aloittelijalle. Liitteet ovat suoraa syötettä ohjelman käyttöliittymästä. Esimerkit on keksitty ja toteutettu itse. Esimerkit on pidetty yksinkertaisina ja kirjoitettu siten, että lukija pystyy seuraamaan välivaiheita. Menetelmät voisi toki kirjoittaa suoraan toimiviksi funktioksi, mutta tällöin esimerkkien seurattavuus kärsii.

Liite 1: Diffie-Hellman

Valitse ensin, kuinka suuren alkuluvun p haluat käyttöösi.

(% i1) n:2^200;

(n) 1606938044258990275541962092341162602522202993782792835301376

Valitaan alkuluku $p > n$ siten, että g on sen primitiivinen juuri modulo p.

(% i2) p:next_prime(n);

(p) 1606938044258990275541962092341162602522202993782792835301611

(% i3) g:zn_primroot(p);

(g) 2

Jaetaan luvut p ja g henkilöille A ja B. Henkilö A valitsee kokonaisluvun a ja B valitsee kokonaisluvun b:

(% i4) a:241;

(a) 241

(% i5) b:179;

(b) 179

Henkilöt A ja B laskevat laskutoimituksensa:

(% i6) A: mod(g^a, p);

(A) 1606938044258990275541962092341162602522202993266022370246891

(% i7) B:mod(g^b, p);

(B) 766247770432944429179173513575154591809369561091801088

Henkilöt A ja B vaihtavat keskenään saamansa tulokset, ja laskevat vielä algoritmin viimeisen laskutoimituksen:

(% i8) s1:mod(B^a, p);

(s1) 450513153024804253805182499363750225286339450204403389438784

(% i9) s2:mod(A^b, p);

(s2) 450513153024804253805182499363750225286339450204403389438784

Saatiin $s1 = s2$, eli molemmilla on nyt sama salausavain.

Liite 2: Shankin algoritmi

Valitaan ensin luvut g, p ja h diskreetin logaritmin ongelmasta.

(% i3) g:6;h:857;p:2791;

(g) 6

(h) 857

(p) 2791

Lasketaan luvun 11 kertaluku modulo 2791. Määritetään tätä varten funktio, jolla kertalukuja voidaan laskea:

(% i4) kertaluku(b):=~ for i:1 thru (p-1) do if(mod(bˆi, p)~=1) then (return (i));

(% o4) kertaluku(b) := for i thru $p-1$ do if mod $\left(b^i, p\right) = 1$ then return(i)

(% i5) f:kertaluku(g);

(f) 2790

Lasketaan kertaluvun avulla algoritmiin n:

(% i6) n:floor(sqrt(f)) + 1;

(n) 53

Lasketaan luvun 6ˆ53 modulo 2791 modulaari käänteisluku:

(% i7) u:inv_mod(gˆ(n), p);

(u) 2036

Etsitään listoista gˆi ja h*uˆj kaksi samaa lukua:

(% i8) for i:1 thru 53 do for j:1 thru 53 do if(mod(gˆi, p) =~mod(h*(uˆj) , p)) then display(i, j);

$\{37, 23\}$

Liite 3: RSA-menetelmä

1) Valitaan luvut p, q ja e.

(% i3) p:next_prime(600); q: next_prime(300); e:11;

(p) 601

(q) 307

(e) 11

2) Ratkaistaan luku d kongruenssista de $\equiv$ 1 (mod (p-1)(q-1)). Tämä onnistuu Laajennetun Eukleideen algoritmin avulla tai laskemalla suoraan invmod funktiolla laskutoimitus d $\equiv e^{-1}$(mod (p-1)(q-1)).

(% i4) mod1:(p-1)*(q-1);

(mod1) 183600

(% i6) d: inv_mod(e, mod1);

(d) 16691

3) Lasketaan N=pq ja juulkaistaan luvut N ja e.

(% i7) N:p*q;

(N) 184507

4) Valitaan lähetettävä viesti m. Huom! 1 < m $\leq$ N.

(% i8) m: 31415;

(m) 31415

5) Lähetetään viesti c $\equiv m^e$ (mod N).

(% i9) c: mod(m^e, N);

(c) 150191

6) Vastaanottaja avaa viestin laskemalla m'$\equiv c^d$ (mod N)

(% i10) avattuviesti: mod(c^d, N);

(avattuviesti) 31415

Liite 4: Pollardin p-1 menetelmä

Ladataan tarvittavia paketteja ja alustetaan suurin yhteinen tekijä:

(% i1) load(gcdex);

(% i2) syt(a,b):=part(igcdex(a,b), 3);

(% o2) $$\text{syt}\left(a,b\right) := \text{part}\left(\text{igcdex}\left(a,b\right),3\right)$$

Syötä RSA-menetelmässä käytetty luku N:

(% i3) N: 2194889357;

(N) 2194889357

Valitaan luku a:

(% i4) a:2;

(a) 2

(% i5) for j:2 thru N do (a: mod(a^j, N), p: syt((a-1), N), if(p>1 and p<N) then (display(j), return(p)));

$$j = 499$$

(% o5) 20959

(% i6) q:N/p; p;

(q) 104723

(% o6) 20959

Nyt $N = pq$.

Lähdeluettelo

[1] Jeffrey Hoffstein, Jill Pipher ja Joseph H. Silverman: *An Introduction to Mathematical Cryptography.* Springer, 2010.

[2] William Stein: *Elementary Number Theory, A Computational Approach.* Springer, 2007. *https://wstein.org/edu/2007/spring/ent/ent.pdf*

[3] Kenneth H. Rosen: *Elemenrary Number Theory and Its Applications.* Addison-Wesley, 1986.

[4] Heli Tuominen: *Lukuteorian alkeet.* Jyväskylän Yliopisto, 2011.

[5] Whitefield Diffie, Martin E. Hellman: *New Directions in Cryptography.* IEEE Transactions on Information Theory, Vol. IT-22, No.6, 1976: *https://ee.stanford.edu/∼hellman/publications/24.pdf*

[6] GCHQ News Articles: *Malcolm John Williamson 1950-2015.* GCHQ News Articles, 2016: *https://www.gchq.gov.uk/news-article/malcolm-john-williamson-1950-2015*

[7] Kai Rajala: *Algebra 1: Ryhmät.* Jyväskylän Yliopisto, 2017.

[8] Ron Rivest, Adi Shamir ja Leonard Adleman: *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems.* Communications of the ACM, Vol. 21, 1978.