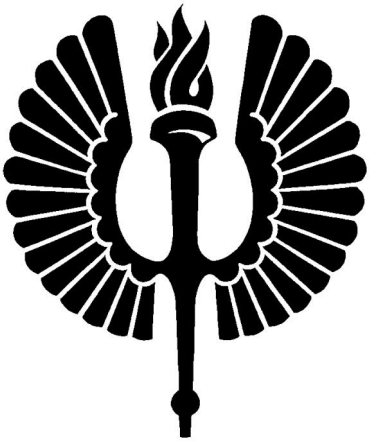
Johdatus julkisen avaimen kryptosysteemeihin

Joni Söderblom

Pro gradu -tutkielma
Turun yliopisto

Maaliskuu 2019



Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

TURUN YLIOPISTO
Matematiikan ja tilastotieteen laitos

SÖDERBLOM, JONI: Johdatus julkisen avaimen kryptosysteemeihin
Pro gradu -tutkielma, 32 s.
Matematiikka
Maaliskuu 2019

Tämän tutkielman tarkoituksena on toimia johdatuksena julkisen avaimen kryptosysteemeihin sekä niiden sovelluksiin. Lukijalta odotetaan vähintään lukion pitkän matematiikan oppimäärän hallitsemista, mutta tästä huolimatta tutkielma pyrkii olemaan mahdollisimman itsenäinen kokonaisuus, jonka vuoksi pohjatietojen esittely alkaa jaollisuuden ja alkulukujen kaltaisista yksinkertaisista määritelmistä, joiden pohjalta edetään hiljalleen erilaisiin kryptografian kannalta olennaisiin tuloksiin.

Toisessa luvussa esitellään tärkeimpiä lukuteoriaan ja ryhmien teoriaan liittyviä pohjatietoja, joita myöhemmin käytetään kryptosysteemien määritelmissä ja soveltamisessa. Keskeisimpiä aiheita ovat esimerkiksi suurin yhteinen tekijä, kiinalainen jäännöslause ja diskreetin logaritmin ongelma, minkä lisäksi luku sisältää useita pienempiä lauseita, joita käytetään ensisijaisesti aputuloksina tärkeämpiä tuloksia todistaessa.

Kolmannessa luvussa siirrytään itse kryptosysteemien käsittelyyn, joka alkaa kryptografian ja kryptosysteemien esittelystä yleisellä tasolla, minkä jälkeen tarkasteluun otetaan yksi kerrallaan RSA-, Rabin- ja ElGamal-kryptosysteemit. Jokaisen kryptosysteemin kohdalla käydään yksityiskohtaisesti läpi systeemin rakenne ja toiminta, todistetaan salauksen ja sen purkamisen oikeellisuus sekä lasketaan pieniä lukuarvoja käyttäen yksi laskuesimerkki.

Tutkielman lopuksi neljännessä luvussa otetaan vielä tarkasteluun julkisen avaimen kryptografian todennäköisesti tärkein sovellus eli digitaaliset allekirjoitukset. Lukija johdatetaan digitaalisten allekirjoitusmenetelmien toimintaan samalla tavalla kuin vastaaviin salaussysteemeihinkin, eli menetelmän toiminta käydään ensin yksityiskohtaisesti läpi, minkä jälkeen lasketaan pieniä lukuja käyttäen yksi käytännön laskuesimerkki.

Asiasanat: kryptografia, RSA, ElGamal, digitaaliset allekirjoitukset.

Sisältö

1 Johdanto **1**

2 Pohjatietoja lukuteoriasta ja algebrasta **2**
2.1 Alkuluvut ja jaollisuus 2
2.2 Suurin yhteinen tekijä ja Eukleideen algoritmi 4
2.3 Kongruenssit ja kiinalainen jäännöslause 8
2.4 Joukot $\mathbb{Z}_m$ ja $\mathbb{Z}_m^*$ 14

3 Julkisen avaimen kryptosysteemeistä **17**
3.1 Kryptografiasta yleisesti 17
3.2 RSA-kryptosysteemi 19
3.3 Rabin-kryptosysteemi 23
3.4 ElGamal-kryptosysteemi 26

4 Digitaaliset allekirjoitukset **28**
4.1 RSA-allekirjoitus 29
4.2 ElGamal-allekirjoitus 30

Lähteet **32**

1 Johdanto

Nykyaikaisen digitaalisen viestinnän turvallisuus perustuu hyvin pitkälti matemaattiselle kryptografialle, jossa hyödynnetään erilaisia tuloksia esimerkiksi lukuteorian, algebran ja todennäköisyyslaskennan aloilta. Tässä tutkielmassa lukijalle esitellään tärkeimpiä julkisen avaimen kryptografiassa tarvittavia esitietoja ja niiden pohjalta muutamia keskeisiä salausmenetelmiä sekä niiden sovelluksia.

Kryptografia itsessään ei ole mitenkään uusi aihealue, vaan viestien salausta on suoritettu monenlaisilla matemaattisilla menetelmillä jo ennen ajanlaskun alkua. Suurimman osan historiastaan kryptografian piirissä ajateltiin, että jos kaksi henkilöä haluavat lähettää salattuja viestejä toisilleen, heidän olisi ensin sovittava jonkinlainen yhteinen salainen avain, jota käytettäisiin sekä viestien salaamiseen että salauksen purkamiseen. Kryptografia kuitenkin mullistui pysyvästi 1970-luvulla julkisen avaimen kryptosysteemien myötä.

Aikaisemmista kryptosysteemeistä poiketen julkisen avaimen kryptosysteemeissä on etuna se, että viestintään osallistuvien henkilöiden ei tarvitse sopia yhteistä salaista avainta, vaan ainoastaan viestien vastaanottajalla on oma salainen avaimensa salausten purkamiseen. Viestien salaamiseen käytettävä avain sen sijaan on julkista tietoa, jonka voi ilmoittaa myös kolmansille osapuolille vaarantamatta viestinnän turvallisuutta. Käytännössä tämä tarkoittaa sitä, että kahden henkilön ei tarvitse edes tavata toisiaan, jotta he voisivat lähettää salattuja viestejä toisilleen. Tällaisen viestinnän mahdollistavien systeemien toimintaa tarkastelemme yksityiskohtaisemmin luvussa 3.

Julkisen avaimen kryptosysteemeillä on erilaisia käyttökelpoisia sovelluksia, joista kenties tärkeimpiä ovat digitaaliset allekirjoitukset. Fyysisten allekirjoitusten tavoin digitaalisten allekirjoitusten tarkoituksena on antaa keino todentaa, että vastaanotettu viesti on saapunut oikealta lähettäjältä. Tällaisia allekirjoituksia käytetään hyödyksi esimerkiksi tietokoneiden parissa silloin, kun tietokoneelle asennetulle ohjelmalle julkaistaan uusi päivitys. Jotta tietokone pystyisi varmistamaan päivityksen saapuvan luotettavalta taholta, se käyttää ohjelman mukana tullutta julkista todennusavainta, jonka avulla se voi hyväksyä oikean päivityksen mukana tulevan digitaalisen allekirjoituksen ja hylätä väärät[1], sallien täten ainoastaan ohjelman kehittäjien tekemät päivitykset. Digitaalisten allekirjoitusmenetelmien toimintaa esitellään tutkielman lopussa luvussa 4.

[1]Tämä on yksinkertaistettu selitys, mutta antaa karkean kuvan digitaalisten allekirjoitusten käytöstä.

2 Pohjatietoja lukuteoriasta ja algebrasta

Tässä luvussa käsittelemme niitä tärkeimpiä pohjatietoja, joita tarvitsemme myöhemmin luvuissa 3 ja 4 kryptosysteemien ja digitaalisten allekirjoitusten parissa. Keskeisimmistä asioista kuten Eukleideen algoritmista ja kiinalaisesta jäännöslauseesta on puhtaan teorian lisäksi esimerkkilaskuja, joiden lisäksi luku sisältää useita pienempiä tuloksia, joiden pääasiallisena tarkoituksena on toimia apuna tärkeimpien tulosten todistamisessa. Luku perustuu lähteisiin [3] ja [6].

2.1 Alkuluvut ja jaollisuus

Määritelmä 2.1. Olkoot a ja b kokonaislukuja ja $a \neq 0$. Jos on olemassa kolmas kokonaisluku c, joka toteuttaa ehdon $b = ac$, sanomme, että a *jakaa* luvun b ja kirjoitamme $a|b$. Tällöin lukua a myös kutsutaan luvun b *jakajaksi*.

Määritelmä 2.2. Kokonaislukua $p > 1$ kutsutaan *alkuluvuksi*, jos sen ainoat positiiviset jakajat ovat p itse ja 1. Jos p ei ole alkuluku, sitä kutsutaan *yhdistetyksi luvuksi*.

Lause 2.3. Jokainen kokonaisluku $n \geq 2$ on joko alkuluku tai se voidaan kirjoittaa alkulukujen tulona.

Todistus. Todistetaan lauseen väite induktiolla. Väite pitää selvästi paikkansa kun $n = 2$, sillä 2 on alkuluku. Oletetaan seuraavaksi, että väite pitää paikkansa kaikille kokonaisluvuille n, kun $2 \leq n \leq k$ ja k on kokonaisluku, joka on suuruudeltaan vähintään 2. Nyt on osoitettava, että tämän oletuksen perusteella väite pätee myös silloin, kun $n = k + 1$.

Jos $k + 1$ on alkuluku, väite pitää paikkansa eikä mitään todistettavaa ole. Jos $k + 1$ taas on yhdistetty luku, niin pystymme löytämään luvut r ja s ehdoilla $2 \leq r \leq k$ ja $2 \leq s \leq k$, jotka toteuttavat yhtälön $k + 1 = rs$. Koska r ja s ovat molemmat välillä $[2, k]$, niin alkuoletuksen perusteella ne ovat joko alkulukuja tai alkulukujen tuloja, eli $k + 1$ on voidaan kirjoittaa vähintään kahden alkuluvun tulona. Kummassakin tapauksessa väite siis pätee kun $n = k + 1$, eli induktion perusteella se pätee kaikille kokonaisluvuille $n \geq 2$. □

Lause 2.4 (Hyvän järjestyksen periaate)**.** Jokaisella epätyhjällä positiivisten kokonaislukujen osajoukolla on pienin alkio.

Todistus. Todistetaan väite vastaoletuksella. Olkoon B positiivisten kokonaislukujen epätyhjä osajoukko, jolla ei ole pienintä alkiota. Olkoon lisäksi C niiden positiivisten kokonaislukujen joukko, jotka eivät kuulu joukkoon B. Selvästi nähdään, että jos luvut $1, 2, \ldots, n \in C$, niin tällöin myös $n + 1 \in C$,

sillä muuten se olisi joukon B pienin alkio. Luku 1 on pienin positiivinen kokonaisluku, joten se kuuluu joukkoon C. Induktiolla seuraa, että kaikki muutkin kokonaisluvut kuuluvat joukkoon C, jolloin joukko B on tyhjä. Tämä on ristiriita vastaoletuksen kanssa, joten alkuperäinen väite pitää paikkansa. □

Lause 2.5 (Jakoalgoritmi)**.** Olkoot a ja b mielivaltaisia kokonaislukuja ja $b > 0$. Tällöin on olemassa yksikäsitteiset kokonaisluvut q (*osamäärä*) ja r (*jakojäännös*), $0 \leq r < b$, jotka toteuttavat ehdon $a = bq + r$.

Todistus. Olkoon C kaikkien muotoa $a - sb$ olevien ei-negatiivisten kokonaislukujen joukko. Jos $a \geq 0$, niin $a - 0b$ on joukon C alkio. Jos taas $a < 0$, niin $a - ab \geq 0$ on joukon C alkio, sillä $b \geq 1$. Kummassakin tapauksessa siis C on epätyhjä, eli lauseen 2.4 perusteella joukolla C on pienin alkio. Olkoon q se luvun s arvo, jonka avulla saadaan joukon C pienin alkio, ja merkitään $a - bq = r$. Nyt siis $0 \leq r$, ja lisäksi $r - b < 0$, koska

$$r - b = a - bq - b = a - (q + 1)b,$$

eli $r - b$ on muotoa $a - sb$ ja pienempi kuin r, joka on pienin ei-negatiivinen joukon C alkio. Edellisen perusteella siis seuraa, että $0 \leq r < b$.

Seuraavaksi on vielä osoitettava lukujen q ja r olevan yksikäsitteisiä. Oletetaan, että $a = bq' + r'$ ja $0 \leq r' < b$. Nyt riittää osoittaa, että $r = r'$ ja $q = q'$. Jos $q' < q$, niin $q' + 1 \leq q$, koska luvut q ja q' ovat molemmat kokonaislukuja. Tällöin

$$r = a - bq \leq a - b(q' + 1) = a - bq' - b = r' - b < 0,$$

mikä on ristiriita, koska r on määritelty ei-negatiiviseksi luvuksi. Vastaavanlainen ristiriita saavutetaan myös silloin, kun $q' > q$, eli siis $q = q'$. Yhtälöstä $bq + r = a = bq' + r'$ seuraa nyt, että $r = r'$, eli sekä q että r ovat yksikäsitteisiä. □

Lause 2.6. Jos $a|b$ ja $b|c$, niin $a|c$.

Todistus. Jos $a|b$ ja $b|c$, niin määritelmän 2.1 perusteella on olemassa kokonaisluvut r ja s, jotka toteuttavat yhtälöt $b = ar$ ja $c = bs$. Tällöin siis $c = a(rs)$, eli $a|c$. □

Lause 2.7. Jos $a|b$ ja $a|c$, niin $a|(bx + cy)$ kaikilla lukujen x ja y arvoilla.

Todistus. Jos $a|b$ ja $a|c$, niin määritelmän 2.1 perusteella on olemassa luvut u ja v, jotka toteuttavat yhtälöt $b = au$ ja $c = av$. Tällöin

$$bx + cy = aux + avy = a(ux + vy),$$

eli $a|(bx + cy)$ kaikilla lukujen x ja y arvoilla. □

Lause 2.8. Jos $a|b$ ja $b \neq 0$, niin $|a| \leq |b|$.

Todistus. Jos $a|b$ ja $b \neq 0$, niin määritelmän 2.1 perusteella on olemassa nollasta poikkeava luku c, joka toteuttaa yhtälön $b = ac$. Tällöin

$$|b| = |a| \cdot |c| \geq |a|,$$

koska $|c| \geq 1$. □

2.2 Suurin yhteinen tekijä ja Eukleideen algoritmi

Määritelmä 2.9. Olkoot a ja b kokonaislukuja ja d suurin luku, joka jakaa ne molemmat. Tällöin sanomme, että d on lukujen a ja b *suurin yhteinen tekijä*, ja merkitsemme $d = \text{syt}(a, b)$.

Huomautus 2.10. Määritelmän 2.9 perusteella on helppoa todeta, että $\text{syt}(a, b)$ on aina positiivinen luku.

Määritelmä 2.11. Olkoon $\text{syt}(a, b) = 1$. Tällöin sanomme, että a ja b ovat *suhteellisia alkulukuja.*

Lause 2.12. Jos ainakin toinen luvuista a ja b poikkeaa nollasta ja $d = \text{syt}(a, b)$, niin d on pienin positiivinen kokonaisluku, joka on muotoa $ax + by$.

Todistus. Olkoon C kaikkien niiden positiivisten kokonaislukujen joukko, jotka ovat muotoa $ax + by$. Alkuoletuksen perusteella ainakin toinen luvuista a ja b eroaa nollasta, eli joukko C ei ole tyhjä. Tällöin lauseen 2.4 perusteella joukossa C on pienin alkio. Olkoon $e = ax_0 + by_0$ tämä pienin alkio. Nyt on siis osoitettava, että $d = e$.

Jakoalgoritmin perusteella on olemassa yksikäsitteiset kokonaisluvut q ja r, jotka toteuttavat ehdot $0 \leq r < e$ ja $a = eq + r$. Nyt siis

$$r = a - eq = a - (ax_0 + by_0)q = a(1 - qx_0) + b(-qy_0),$$

eli r on muotoa $ax + by$. Jos r eroaisi nollasta, se kuuluisi joukkoon C, jolloin päätyisimme ristiriitaan sen oletuksen kanssa, että e on joukon C pienin alkio. Tästä siis seuraa, että $r = 0$ ja $e|a$. Vastaavanlaisella päättelyllä voimme osoittaa, että $e|b$. Luku e siis jakaa sekä luvun a että luvun b, jolloin määritelmän 2.9 perusteella $e \leq d$. Toisaalta koska $d|a$ ja $d|b$, niin lauseen 2.7 perusteella $d|e$. Tällöin lauseesta 2.8 seuraa, että $d \leq e$, eli $d = e$. □

Lause 2.13. Suurin yhteinen tekijä $\text{syt}(a, b) = 1$ jos ja vain jos on olemassa jotkin kokonaisluvut x ja y, jotka toteuttavat yhtälön $1 = ax + by$.

Todistus. Oletetaan ensin, että $\text{syt}(a, b) = 1$. Tällöin lauseesta 2.12 seuraa suoraan, että luku 1 voidaan esittää muodossa $ax + by$.

Oletetaan sitten, että $1 = ax + by$. Suurimman yhteisen tekijän määritelmän perusteella $\text{syt}(a, b)|(ax + by) = 1$, josta seuraa automaattisesti, että $\text{syt}(a, b) = 1$, koska luvun 1 ainoa positiivinen jakaja on 1 itse, ja kahden luvun suurin yhteinen tekijä on aina positiivinen. Näin ollen lause on siis saatu todistettua. □

Seuraus 2.14. Olkoot $d = \text{syt}(a, b)$, $a = Ad$ ja $b = Bd$. Tällöin $\text{syt}(A, B) = 1$.

Todistus. Koska $d = \text{syt}(a, b)$, niin lauseen 2.12 perusteella $d = ax + by$ joillakin lukujen x ja y arvoilla. Jakamalla yhtälön molemmat puolet luvulla d saamme muutettua sen muotoon

$$1 = \frac{a}{d}x + \frac{b}{d}y = Ax + By,$$

jolloin lauseen 2.13 perusteella $\text{syt}(A, B) = 1$. □

Lause 2.15. Jos $a|bc$ ja $\text{syt}(a, b) = 1$, niin $a|c$.

Todistus. Koska $\text{syt}(a, b) = 1$, niin lauseen 2.12 perusteella on olemassa luvut x ja y, jotka toteuttavat yhtälön $1 = ax + by$. Kertomalla yhtälön molemmat puolet luvulla c saamme $c = acx + bcy$. Koska $a|a$ ja $a|bc$, niin lauseen 2.7 perusteella $a|(acx + bcy)$, eli $a|c$. □

Lause 2.16. Olkoot p alkuluku ja $p|bc$. Tällöin $p|b$ tai $p|c$.

Todistus. Jos $p|b$ niin mitään todistettavaa ei ole. Jos taas $p \nmid b$, niin $\text{syt}(p, b) = 1$, koska luvun p ainoat positiiviset jakajat ovat p itse ja 1. Tällöin lauseen 2.15 perusteella $p|c$. □

Lause 2.17. Olkoot p alkuluku ja $p|a_1a_2 \cdots a_n$. Tällöin $p|a_i$ jollakin luvun i arvolla $1, 2, \ldots, n$.

Todistus. Merkitään $a_1a_2 \cdots a_{n-1} = b$. Tällöin $p|ba_n$, eli lauseen 2.16 perusteella $p|a_n$, jolloin p siis jakaa jonkin luvuista a_i, tai $p|b$, jolloin $p|a_1a_2 \cdots a_{n-1}$. Jälkimmäisessä tapauksessa voimme käyttää samaa menetelmää ja merkitä $a_1a_2 \cdots a_{n-2} = c$, jolloin $p|ca_{n-1}$ eli $p|a_{n-1}$ tai $p|c$. Jatkamalla tätä päättelyketjua päädymme lopulta siihen tulokseen, että luvun p on jaettava jokin luvuista $a_1, a_2, \ldots, a_n$. □

Lause 2.18. Olkoon $\text{syt}(a, b_i) = 1$ kaikilla luvun $i = 1, 2, \ldots, n$ arvoilla. Tällöin $\text{syt}(a, b_1b_2 \cdots b_n) = 1$.

Todistus. Todistetaan lause vastaoletuksella. Oletetaan ensin, että suurin yhteinen tekijä $\text{syt}(a, b_1 b_2 \cdots b_n) = d > 1$. Tällöin lauseen 2.3 perusteella on olemassa jokin alkuluku p, joka jakaa luvun d. Luvun d määritelmän perusteella $d|a$ ja $d|b_1 b_2 \cdots b_n$, jolloin lauseesta 2.6 seuraa, että $p|a$ ja $p|b_1 b_2 \cdots b_n$. Tällöin lauseen 2.17 perusteella $p|b_i$ jollakin luvun $i = 1, 2, \ldots n$ arvolla. Nyt siis $p|a$ ja $p|b_i$, mikä on ristiriita alkuoletuksen $\text{syt}(a, b_i) = 1$ kanssa. Tämän perusteella saadaan siis, että $d = 1$. □

Lause 2.19. Jos $a|c$, $b|c$ ja $\text{syt}(a, b) = 1$, niin $ab|c$.

Todistus. Koska $a|c$ ja $b|c$, on määritelmän 2.1 perusteella olemassa kokonaisluvut r ja s, jotka toteuttavat yhtälön $ar = c = bs$. On selvää, että $b|bs$, joten $b|ar$. Tällöin alkuoletuksen ja lauseen 2.15 perusteella $b|r$, eli $r = bt$ jollakin luvun t arvolla ja $c = ar = abt$. Siis $ab|c$, eli väite on todistettu. □

Lause 2.20. Olkoot $m_1, m_2, \ldots, m_n$ pareittain suhteellisia alkulukuja, eli $\text{syt}(m_i, m_j) = 1$ kaikilla $i \neq j$, ja lisäksi $m_i|a$ kaikilla luvun $i = 1, 2, \ldots n$ arvoilla. Tällöin $m|a$, kun $m = m_1 m_2 \cdots m_n$.

Todistus. Lause pitää selvästi paikkansa, kun $n = 1$, koska tällöin $m = m_1$. Oletetaan seuraavaksi, että lause pätee kun $n = k$, ja tarkastellaan kokonaislukuja $m_1, m_2, \ldots, m_{k+1}$, missä $\text{syt}(m_i, m_j) = 1$ kaikilla $i \neq j$. Merkitään lisäksi $m' = m_1 m_2 \cdots m_k$. Nyt lauseen 2.18 perusteella $\text{syt}(m', m_{k+1}) = 1$, ja induktio-oletuksen perusteella $m'|a$. Tällöin lauseesta 2.19 seuraa, että $m' m_{k+1}|a$, ja koska $m' m_{k+1} = m_1 m_2 \cdots m_{k+1}$, niin lause pätee induktion perusteella kaikilla luvun $n \geq 1$ arvoilla. □

Lause 2.21. Olkoot $f_1, f_2, \ldots, f_n$ kaikki lukujen a ja b yhteiset jakajat. Tällöin $d = \text{syt}(a, b)$ jos ja vain jos $d > 0$, $d|a$, $d|b$ ja $f_i|d$ kaikilla luvun $i = 1, 2, \ldots, n$ arvoilla.

Todistus. Koska kyse on ekvivalenssista, on implikaatio todistettava molempiin suuntiin. Oletetaan ensin, että $d = \text{syt}(a, b)$. Tällöin määritelmän 2.9 perusteella $d|a$ ja $d|b$, ja lisäksi lauseen 2.12 perusteella on olemassa kokonaisluvut x ja y, jotka toteuttavat yhtälön $d = ax + by > 0$. Tällöin myös lauseen 2.7 perusteella $f_i|d$ kaikilla luvun i arvoilla, koska $f_i|a$ ja $f_i|b$.

Oletetaan seuraavaksi, että $d > 0$, $d|a$, $d|b$ ja $f_i|d$ kaikilla luvun i arvoilla. Tällöin d on lukujen a ja b yhteinen jakaja, ja lauseen 2.8 perusteella $|f_i| \leq d$, eli määritelmän 2.9 perusteella $d = \text{syt}(a, b)$. □

Lause 2.22 (Eukleideen algoritmi)**.** Olkoot a ja b kokonaislukuja ja $a > b > 0$. Tällöin voimme jakoalgoritmin perusteella löytää osamäärän q_1 ja jakojäännöksen r_1, jotka toteuttavat ehdot $a = bq_1 + r_1$ ja $0 \leq r_1 < b$.

Jos $r_1 = 0$, niin $b|a$ eli $\text{syt}(a, b) = b$. Jos taas $r_1 \neq 0$, algoritmia jatketaan jakamalla b luvulla r_1, jolloin löydetään uudet luvut q_2 ja r_2 ja saadaan yhtälö $b = r_1 q_2 + r_2$. Jos $r_2 \neq 0$, prosessia jatketaan jälleen eteenpäin siihen asti, kunnes jakojäännökseksi saadaan 0. Algoritmin etenemisestä voidaan kirjoittaa seuraavat yhtälöt:

$$\begin{aligned} a &= bq_1 + r_1, \\ b &= r_1 q_2 + r_2, \\ r_1 &= r_2 q_3 + r_3, \\ &\dots \\ r_{k-3} &= r_{k-2} q_{k-1} + r_{k-1}, \\ r_{k-2} &= r_{k-1} q_k + r_k \text{ ja} \\ r_{k-1} &= r_k q_{k+1}. \end{aligned} \tag{1}$$

Yhtälöryhmässä (1) esiintyvä viimeinen nollaa suurempi jakojäännös r_k on lukujen a ja b suurin yhteinen tekijä.

Todistus. On selvää, että $r_k|r_k$, ja lisäksi yhtälöryhmän (1) perusteella $r_k|r_{k-1}$. Tällöin lauseen 2.7 ja yhtälöryhmän (1) toiseksi viimeisen yhtälön perusteella $r_k|r_{k-2}$. Nyt siis $r_k|r_{k-1}$ ja $r_k|r_{k-2}$, joten vastaavanlaisen päättelyn perusteella $r_k|r_{k-3}$. Jatkamalla tätä päättelyketjua loppuun asti saamme osoitettua, että $r_k|a$ ja $r_k|b$.

Oletetaan seuraavaksi, että $f_1, f_2, \dots, f_n$ ovat kaikki lukujen a ja b yhteiset jakajat. Tällöin lauseen 2.7 ja yhtälöryhmän (1) ensimmäisen yhtälön perusteella $f_i|r_1$ kaikilla luvun $i = 1, 2, \dots, n$ arvoilla. Nyt siis $f_i|b$ ja $f_i|r_1$, jolloin vastaavanlaisen päättelyn perusteella $f_i|r_2$. Samaa päättelyketjua jatkamalla saamme osoitettua, että $f_i|r_k$ kaikilla luvun i arvoilla, jolloin lauseen 2.21 ehdot täyttyvät. Tämän perusteella $r_k = \text{syt}(a, b)$. □

Esimerkki 2.23. Etsitään lukujen 1994 ja 2019 suurin yhteinen tekijä Eukleideen algoritmilla.

$$\begin{aligned} 2019 &= 1994 \cdot 1 + 25 \\ 1994 &= 25 \cdot 79 + 19 \\ 25 &= 19 \cdot 1 + 6 \\ 19 &= 6 \cdot 3 + 1 \\ 6 &= 1 \cdot 6 \end{aligned} \tag{2}$$

Lauseen 2.22 perusteella lukujen suurin yhteinen tekijä on viimeinen nollaa suurempi jakojäännös, eli $\text{syt}(1994, 2019) = 1$. Saimme siis myös selville, että luvut 1994 ja 2019 ovat suhteellisia alkulukuja.

Lisäksi tiedämme lauseen 2.12 perusteella, että on olemassa luvut x ja y, jotka toteuttavat yhtälön $1 = 1994x + 2019y$. Näiden lukujen löytämiseen voimme käyttää apuna yhtälöryhmän (2) yhtälöitä eliminoimalla niistä kaikki jakojäännökset aloittaen toiseksi alimmasta yhtälöstä.

$$\begin{aligned}
1 &= 19 - 3 \cdot 6 \\
&= 19 - 3 \cdot (25 - 19) \\
&= 4 \cdot 19 - 3 \cdot 25 \\
&= 4 \cdot (1994 - 79 \cdot 25) - 3 \cdot 25 \\
&= 4 \cdot 1994 - 319 \cdot 25 \\
&= 4 \cdot 1994 - 319 \cdot (2019 - 1994) \\
&= 1994 \cdot 323 + 2019 \cdot (-319)
\end{aligned}$$

Siis $1 = 1994 \cdot 323 + 2019 \cdot (-319)$.

2.3 Kongruenssit ja kiinalainen jäännöslause

Määritelmä 2.24. Olkoon $m > 0$ ja a ja b kokonaislukuja. Jos $m|(a - b)$, sanomme, että a ja b ovat *kongruentteja modulo* m ja kirjoitamme $a \equiv b \pmod{m}$. Jos taas $m \nmid (a - b)$, sanomme lukujen olevan *epäkongruentteja modulo* m ja kirjoitamme $a \not\equiv b \pmod{m}$. Luvusta m käytetään lisäksi nimitystä *moduli*.

Kongruenssin määritelmästä on helppoa todeta seuraavat väitteet paikkansapitäviksi.

(i) Kaikilla luvun a arvoilla pätee $a \equiv a \pmod{m}$.
(ii) Jos $a \equiv b \pmod{m}$, niin $b \equiv a \pmod{m}$.
(iii) Jos $a \equiv b \pmod{m}$ ja $b \equiv c \pmod{m}$, niin $a \equiv c \pmod{m}$.

Lause 2.25. Olkoon $ac \equiv bc \pmod{m}$ ja $d = \text{syt}(c, m)$. Tällöin $a \equiv b \pmod{\frac{m}{d}}$

Todistus. Määritelmän perusteella $m|(ac - bc)$, jolloin siis on olemassa jokin kokonaisluku q, joka toteuttaa yhtälön $mq = ac - bc = (a - b)c$. Merkitään $c = Cd$ ja $m = Md$. Kun edellisen rivin yhtälö jaetaan puolittain luvulla d, muuttuu se muotoon $Mq = (a - b)C$. Nähdään, että $M|(a - b)C$, ja lisäksi seurauksen 2.14 perusteella $\text{syt}(C, M) = 1$, jolloin lauseesta 2.15 seuraa, että $M|(a - b)$. Siis $a \equiv b \pmod{M}$, missä $M = \frac{m}{d}$. □

Lause 2.26. Olkoon $a \equiv b \pmod{m_i}$ kaikilla luvun $i = 1, 2, \ldots, r$ arvoilla, ja olkoot lisäksi luvut $m_1, m_2, \ldots, m_r$ kaikki pareittain suhteellisia alkulukuja. Tällöin $a \equiv b \pmod{m}$, missä $m = m_1 m_2 \cdots m_r$.

Todistus. Oletuksesta $a \equiv b \pmod{m_i}$ seuraa, että $m_i|(a-b)$ kaikilla luvun i arvoilla. Koska luvut $m_1, m_2, \ldots, m_r$ ovat lisäksi pareittain suhteellisia alkulukuja, niin lauseen 2.20 perusteella $m|(a-b)$ eli $a \equiv b \pmod{m}$. □

Lause 2.27. Olkoon a kokonaisluku. Tällöin kongruenssin $ab \equiv 1 \pmod{m}$ toteuttava luku b on olemassa jos ja vain jos $\text{syt}(a, m) = 1$. Jos tällainen luku b on olemassa, sitä kutsutaan luvun a *käänteisalkioksi* modulo m.

Todistus. Oletetaan ensin, että $\text{syt}(a, m) = 1$. Tällöin lauseen 2.13 perusteella voimme löytää kokonaisluvut x ja y, jotka toteuttavat yhtälön $ax+my = 1$. Yhtälö voidaan kirjoittaa muotoon $ax-1 = -my$, josta nähdään, että $ax-1$ on jaollinen luvulla m, eli $ax \equiv 1 \pmod{m}$, jolloin luvuksi b voidaan valita x.

Oletetaan sitten, että luvulla a on käänteisalkio b modulo m, eli $ab \equiv 1 \pmod{m}$. Nyt kongruenssin määritelmän perusteella on olemassa jokin kokonaisluku c, joka toteuttaa yhtälön $ab-1 = cm$. Tästä seuraa, että $\text{syt}(a, m)$ jakaa luvun $ab - cm = 1$, jolloin $\text{syt}(a, m) = 1$. □

Käänteisalkion laskemisessa modulo m voidaan hyödyntää seuraavanlaista variaatiota Eukleideen algoritmista, jossa yhtälöitä muokataan matriisimuodossa. Jos tehtävänä on laskea luvun a käänteisalkio modulo m, kirjoitamme luvut osana matriisia muodossa

$$\left(\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & m \end{array}\right). \tag{3}$$

Oletetaan seuraavaksi, että $m > a$. Päinvastainen tapaus etenee muuten samalla tavalla, mutta luvut a ja m vaihtavat paikkoja. Jakoalgoritmin perusteella tiedämme, että jotkin kokonaisluvut q_1 ja r_1 toteuttavat yhtälön $m = aq_1+r_1$. Vähennetään seuraavaksi matriisin (3) alemmasta rivistä ylempi rivi luvulla q_1 kerrottuna, siis

$$\left(\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & m \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 0 & a \\ -q_1 & 1 & r_1 \end{array}\right).$$

Tämän jälkeen rivien roolit vaihtuvat, eli matriisin ylemmästä rivistä vähennetään alempi luvulla q_2 kerrottuna, kun $a = r_1q_2+r_2$. Tätä rivien vuorottaista kertomista ja vähentelyä jatketaan siihen asti, kunnes toinen pystyviivan oikealla puolella olevista luvuista saavuttaa arvon 1. Tällöin kyseisen rivin ensimmäinen luku on alkuperäisen luvun a käänteisluku modulo m.

Esimerkki 2.28 (Käänteisluku Eukleideen algoritmilla)**.** Havainnollistetaan edellä esitettyä sanallista selitystä laskuesimerkillä ja lasketaan luvun 13 käänteisalkio modulo 36.

$$\left(\begin{array}{cc|c} 1 & 0 & 13 \\ 0 & 1 & 36 \end{array}\right) \xrightarrow{q_1=2} \left(\begin{array}{cc|c} 1 & 0 & 13 \\ -2 & 1 & 10 \end{array}\right)$$
$$\xrightarrow{q_2=1} \left(\begin{array}{cc|c} 3 & -1 & 3 \\ -2 & 1 & 10 \end{array}\right)$$
$$\xrightarrow{q_3=3} \left(\begin{array}{cc|c} 3 & -1 & 3 \\ -11 & 4 & 1 \end{array}\right)$$

Algoritmi on nyt saatu päätökseen, koska toinen pystyviivan oikealla puolella olevista luvuista on 1, ja rivin ensimmäinen luku -11 on etsitty käänteisluku. Siis

$$13^{-1} = -11 \equiv 25 \pmod{36}.$$

Tuloksen voi vielä helposti todeta oikeaksi seuraavilla laskuilla:

$$13 \cdot 25 = 325 = 9 \cdot 36 + 1.$$

Määritelmä 2.29. Jos $f(a) \equiv 0 \pmod{m}$, niin lukua a kutsutaan *kongruenssiyhtälön* $f(x) \equiv 0 \pmod{m}$ *ratkaisuksi*. Kongruenssiyhtälöä kutsutaan *lineaariseksi kongruenssiksi*, jos $f(x)$ on ensimmäisen asteen polynomi.

Lause 2.30. Lineaarinen kongruenssi $ax \equiv b \pmod{m}$ on ratkaistavissa jos ja vain jos $d|b$, missä $d = \text{syt}(a, m)$. Jos ratkaisuja on olemassa, niin niistä täsmälleen d kappaletta on keskenään epäkongruentteja.

Todistus. Oletetaan ensin, että $d|b$, jolloin $b = kd$ jollakin kokonaisluvulla k. Koska $d = \text{syt}(a, m)$, niin lauseen 2.12 perusteella on olemassa jotkin kokonaisluvut r ja s, jotka toteuttavat yhtälön $d = ar + ms$. Kertomalla yhtälön molemmat puolet luvulla k voimme kirjoittaa $b = kd = akr + mks$, jolloin $akr - b = -mks$. Näemme siis, että $m|(akr - b)$ eli $akr \equiv b \pmod{m}$ ja kr on alussa esitetyn lineaarisen kongruenssin ratkaisu.

Oletetaan seuraavaksi, että x_0 on alussa esitetyn kongruenssin ratkaisu. Tällöin siis $ax_0 \equiv b \pmod{m}$, eli $ax_0 - b = qm$ jollakin luvun q arvolla. Luvun d määritelmän perusteella $d|a$ ja $d|m$, jolloin $d|b$, koska $b = ax_0 - qm$.

Tarkastellaan lopuksi vielä ratkaisujen lukumäärää. Jos x_0 on alussa esitetyn kongruenssin ratkaisu, niin tällöin myös $x_0 + \dfrac{km}{d}$ on ratkaisu jokaisella luvun k arvolla, koska

$$a\Big(x_0 + \frac{km}{d}\Big) = ax_0 + km \cdot \frac{a}{d} \equiv ax_0 \equiv b \pmod{m}.$$

Lisäksi jos x_0 ja x_1 ovat molemmat esitetyn kongruenssin ratkaisuja, niin $ax_0 \equiv b \equiv ax_1 \pmod{m}$. Tällöin lauseen 2.25 perusteella $x_0 \equiv x_1 \left(\mathrm{mod}\ \frac{m}{d}\right)$, eli $x_1 = x_0 + \frac{km}{d}$ jollakin luvun k arvolla.

Edellisten perusteella siis jos x_0 on jokin kongruenssin ratkaisu, niin $x_0 + \frac{km}{d}$ on ratkaisu jokaisella luvun k arvolla ja kongruenssin kaikki ratkaisut ovat tätä muotoa, jolloin jono

$$\ldots, x_0 - \frac{m}{d}, x_0, x_0 + \frac{m}{d}, x_0 + \frac{2m}{d}, \ldots \tag{4}$$

sisältää kaikki tutkittavan kongruenssin ratkaisut. On helppoa todeta, että ratkaisut

$$x_0, x_0 + \frac{m}{d}, \ldots, x_0 + \frac{(d-1)m}{d} \tag{5}$$

ovat kaikki keskenään epäkongruentteja modulo m, sillä minkään kahden erotus ei ole luvun m monikerta. Lisäksi jokainen listan (4) luvuista on kongruentti jonkun listan (5) luvun kanssa, joita on yhteensä d kappaletta. Tämän osoittamiseksi hyödynnämme jälleen jakoalgoritmia, jonka perusteella jokainen k voidaan kirjoittaa lukujen q ja r avulla muodossa $k = qd+r$, missä $0 \leq r < d$. Tällöin

$$\begin{aligned} x_0 + \frac{km}{d} &= x_0 + \frac{(qd+r)m}{d} \\ &= x_0 + qm + \frac{rm}{d} \\ &\equiv x_0 + \frac{rm}{d} (\mathrm{mod}\ m), \end{aligned}$$

ja $x_0 + \frac{rm}{d}$ sisältyy listaan (5), koska $r < d$. □

Lause 2.31 (Kiinalainen jäännöslause)**.** Olkoon $\mathrm{syt}(m_i, m_j) = 1$ kun $i \neq j$. Tällöin kongruenssiryhmällä

$$\begin{cases} x \equiv c_1 & (\mathrm{mod}\ m_1) \\ x \equiv c_2 & (\mathrm{mod}\ m_2) \\ \qquad\vdots & \\ x \equiv c_r & (\mathrm{mod}\ m_r) \end{cases}$$

on ratkaisu, ja kyseinen ratkaisu on yksikäsitteinen modulo m, kun $m = m_1 m_2 \cdots m_r$.

Todistus. Osoitetaan ensin ratkaisun olemassaolo. Merkitään $M_i = \frac{m}{m_i}$. Tällöin $m_j | M_i$ kaikilla $i \neq j$. Koska $\text{syt}(m_i, m_j) = 1$, niin lauseen 2.18 perusteella myös $\text{syt}(m_i, M_i) = 1$. Tästä seuraa lauseen 2.30 mukaan, että kongruenssilla $M_i y \equiv 1 \pmod{m_i}$ on aina olemassa ratkaisu b_i, eli $M_i b_i \equiv 1 \pmod{m_i}$ jokaisella luvun i arvolla.

Merkitään seuraavaksi

$$x_0 = \sum_{i=1}^{r} c_i M_i b_i \tag{6}$$

ja osoitetaan tämän olevan ratkaisu alkuperäiselle kongruenssiryhmälle. Selvästi nyt $c_i M_i b_i \equiv c_i \pmod{m_i}$, koska $M_i b_i \equiv 1 \pmod{m_i}$. Lisäksi $M_j \equiv 0 \pmod{m_i}$ kun $i \neq j$, joten $x_0 \equiv c_i \pmod{m_i}$ ja täten x_0 on kongruenssiryhmän ratkaisu.

Osoitetaan vielä ratkaisun yksikäsitteisyys. Jos x_1 ja x_0 ovat molemmat ratkaisuja kongruenssiryhmälle, niin

$$x_1 \equiv c_i \equiv x_0 \pmod{m_i}$$

jokaisella luvun i arvolla. Koska lisäksi $\text{syt}(m_i, m_j) = 1$ kaikilla $i \neq j$, niin lauseen 2.26 perusteella $x_1 \equiv x_0 \pmod{m}$, eli ratkaisu on yksikäsitteinen modulo m. □

Esimerkki 2.32. Ratkaistaan kongruenssiryhmä

$$\begin{cases} x & \equiv 1 \pmod{13} \\ x & \equiv 3 \pmod{4} \\ x & \equiv 8 \pmod{9} \end{cases}$$

kiinalaisen jäännöslauseen avulla. Nyt $m_1 = 13$, $m_2 = 4$ ja $m_3 = 9$, eli $m = 13 \cdot 4 \cdot 9 = 468$. Lisäksi $M_1 = \frac{468}{13} = 36$, $M_2 = \frac{468}{4} = 117$ ja $M_3 = \frac{468}{9} = 52$. Kongruenssin määritelmän perusteella todetaan, että $36 \equiv 10 \pmod{13}$, $117 \equiv 1 \pmod{4}$ ja $52 \equiv 7 \pmod{9}$, joten seuraavaksi on ratkaistava kongruensseista

$$\begin{aligned} 10b_1 &\equiv 1 \pmod{13}, \\ 1b_2 &\equiv 1 \pmod{4} \text{ ja} \\ 7b_3 &\equiv 1 \pmod{9} \end{aligned}$$

luvut b_1, b_2 ja b_3. Keskimmäinen kongruenssi on jo valmiiksi ratkaistuna, eli $b_2 = 1$. Luvut b_1 ja b_3 saavat molemmat arvokseen 4, koska $10 \cdot 4 = 40 =$

$3 \cdot 13+1$, ja $7 \cdot 4 = 28 = 3 \cdot 9+1$. Nyt kun kaikki luvut M_i ja b_i ovat selvitettynä, voidaan alkuperäisen kongruenssiryhmän ratkaisu laskea kaavalla (6).

$$\begin{aligned} x_0 &= \sum_{i=1}^{r} c_i M_i b_i \\ &= 1 \cdot 36 \cdot 4 + 3 \cdot 117 \cdot 1 + 8 \cdot 52 \cdot 4 \\ &= 2159 \\ &\equiv 287 \pmod{468} \text{ (pienin positiivinen ratkaisu)} \end{aligned}$$

Vastauksen voi vielä tarkistaa oikeaksi seuraavien yhtälöiden avulla:

$$\begin{aligned} 287 &= 22 \cdot \mathbf{13} + \mathbf{1}, \\ 287 &= 71 \cdot \mathbf{4} + \mathbf{3} \text{ ja} \\ 287 &= 31 \cdot \mathbf{9} + \mathbf{8}. \end{aligned}$$

Tarkastellaan tämän osion lopuksi vielä sitä, kuinka pystymme laskemaan korkeita potensseja modulo m. Tällaisia laskuja joutuu usein suorittamaan julkisen avaimen kryptosysteemien yhteydessä, ja siksi siihen tarkoitukseen esitelläänkin seuraavaksi esimerkin avulla eräs algoritmi, joka variaatioineen tunnetaan englanninkielisessä lähdemateriaalissa muun muassa nimillä *square-and-multiply* ja *exponentiation by squaring*.

Esimerkki 2.33 (Potenssiinkorotusalgoritmi)**.** Lasketaan $7^{966} \pmod{35}$. Laskun 7^{966} näppäileminen tavalliseen laskimeen antaa todennäköisesti virheilmoituksen, sillä vastaus on liian suuri laskimen käsiteltäväksi, jonka vuoksi laskun suorittamiseen on syytä hyödyntää algoritmia, jossa potenssiinkorotus tapahtuu vaiheittain, eikä tulos pääse missään vaiheessa kasvamaan liian suureksi laskimen käsiteltäväksi.

Aloitetaan laskemalla arvoja luvuille $7^{2^n} \pmod{35}$, missä $n = 0, 1, 2, \ldots$, ja jatketaan tätä niin kauan kun eksponentit 2^n pysyvät pienempinä kuin alkuperäinen eksponentti 966.

$$\begin{aligned} 7^1 &\equiv 7 \pmod{35} \\ 7^2 &= 49 \equiv 14 \pmod{35} \\ 7^4 &= (7^2)^2 \equiv 14^2 \equiv 196 \equiv 21 \pmod{35} \\ 7^8 &= (7^4)^2 \equiv 21^2 \equiv 441 \equiv 21 \pmod{35} \end{aligned}$$

Tässä kohtaa huomaamme, että $21^2 \equiv 21 \pmod{35}$, jonka vuoksi voimme suoraan todeta loppujen tarvittavien potenssien kaikkien olevan 21 (mod 35),

sillä seuraava lukuarvo saadaan aina edellisestä korottamalla se potenssiin 2. Alla olevat kongruenssit eivät siis tässä tapauksessa vaadi mitään laskemista, mutta yleisesti ne pitäisi laskea kuten edellä.

$$\begin{aligned} 7^{16} &\equiv 21 \pmod{35} \\ 7^{32} &\equiv 21 \pmod{35} \\ 7^{64} &\equiv 21 \pmod{35} \\ 7^{128} &\equiv 21 \pmod{35} \\ 7^{256} &\equiv 21 \pmod{35} \\ 7^{512} &\equiv 21 \pmod{35} \end{aligned}$$

Nyt voimme lopettaa potenssien listaamisen, sillä seuraava eksponentti $2^{10} = 1024$ on suurempi kuin 966. Seuraavaksi esitetään alkuperäinen potenssi 966 binäärimuodossa.

$$966_{10} = 1111000110_2$$

Luvun binääriesityksestä voimme todeta, että

$$\begin{aligned} 966 &= 2^9 + 2^8 + 2^7 + 2^6 + 2^2 + 2^1 \\ &= 512 + 256 + 128 + 64 + 4 + 2. \end{aligned}$$

Käyttämällä kyseistä binääriesitystä, potenssien laskusääntöjä sekä edellä laskettuja kongruensseja voimme kirjoittaa alkuperäisen laskutoimituksen sellaisessa muodossa, että sen laskeminen onnistuu tavallisilla laskimillakin.

$$\begin{aligned} 7^{966} &= 7^{512+256+128+64+4+2} \\ &= 7^{512} \cdot 7^{256} \cdot 7^{128} \cdot 7^{64} \cdot 7^4 \cdot 7^2 \\ &\equiv 21 \cdot 21 \cdot 21 \cdot 21 \cdot 21 \cdot 14 \pmod{35} \\ &\equiv 57177414 \pmod{35} \\ &\equiv 14 \pmod{35} \end{aligned}$$

Siis $7^{966} \equiv 14 \pmod{35}$.

2.4 Joukot $\mathbb{Z}_m$ ja $\mathbb{Z}_m^*$

Edellä käsittelemämme teorian perusteella tiedämme, että mielivaltainen kokonaisluku a voidaan kirjoittaa muodossa $a = mq + r$, missä $0 \leq r < m$, jolloin kongruenssin määritelmän perusteella $a \equiv r \pmod{m}$. Tästä seuraa, että kun puhumme kokonaisluvuista modulo m, meille riittää käyttää ainoastaan lukuja $0, 1, 2, \ldots, m-1$. Tämän pohjalta on luontevaa ottaa käyttöön seuraava määritelmä.

Määritelmä 2.34. Kirjoitamme

$$\mathbb{Z}_m = \{0, 1, 2, \ldots, m-1\}$$

ja käytämme kyseisestä joukosta nimeä *pienimmät ei-negatiiviset jäännökset modulo m*. Yleisesti ottaen kun laskutoimituksia suoritetaan modulo m, lopulliset vastaukset esitetään joukon $\mathbb{Z}_m$ alkioina, ellei ole jotain erityistä syytä toimia toisin.

Määritelmä 2.35. Lauseen 2.27 perusteella tiedämme, että luvulla a on käänteisalkio modulo m jos ja vain jos $\text{syt}(a, m) = 1$. Otetaan käyttöön merkintä $\mathbb{Z}_m^*$, jolla merkitsemme tällaisten lukujen joukkoa modulo m. Siis

$$\begin{aligned}\mathbb{Z}_m^* &= \{a \in \mathbb{Z}_m \mid \text{syt}(a, m) = 1\} \\ &= \{a \in \mathbb{Z}_m \mid \text{luvulla } a \text{ on käänteisalkio modulo } m\}.\end{aligned}$$

Erityishuomiona voidaan mainita, että kun modulina on jokin alkuluku p, kaikki luvuista $1, 2, \ldots, p-1$ ovat suhteellisia alkulukuja modulin p kanssa, eli

$$\mathbb{Z}_p^* = \{1, 2, \ldots, p-1\}.$$

Lause 2.36 (Fermat'n pieni lause)**.** Olkoon a jokin kokonaisluku, joka ei ole jaollinen alkuluvulla p. Tällöin

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

Todistus. Tarkastellaan lukuja $a, 2a, 3a, \ldots, (p-1)a$ modulo p. Luvut ovat kaikki keskenään erisuuria, mikä voidaan osoittaa seuraavalla tavalla. Oletetaan, että kaksi kyseisestä joukosta mielivaltaisesti valittua lukua $ka \pmod{p}$ ja $la \pmod{p}$ ovat samoja, eli $ka \equiv la \pmod{p}$. Nyt siis $(k-l)a \equiv 0 \pmod{p}$, jolloin p jakaa tulon $(k-l)a$. Lauseen 2.16 perusteella p jakaa toisen luvuista $(k-l)$ ja a, mutta alkuoletuksen perusteella tiedämme, että $p \nmid a$, joten $p|(k-l)$. Molemmat luvuista k ja l ovat välillä $[1, p-1]$, eli niiden erotus $k-l$ on lukujen $-(p-2)$ ja $p-2$ välissä. Tällä välillä on kuitenkin vain yksi luku, joka on jaollinen luvulla p, ja tämä luku on nolla. Siis $j - k = 0$, eli $j = k$.

Edellisen tuloksen lisäksi lukuja on yhteensä $p-1$ kappaletta, joten niiden on oltava samat kuin lukujen $1, 2, 3, \ldots, p-1$ jossakin mielivaltaisessa järjestyksessä. Tämän perusteella voimme kirjoittaa

$$a \cdot 2a \cdot 3a \cdots (p-1)a \equiv 1 \cdot 2 \cdot 3 \cdots (p-1) \pmod{p}. \tag{7}$$

Seuraavaksi voimme käyttää kertoman määritelmää ja potenssien laskusääntöjä, joiden perusteella kongruenssi (7) muuttuu muotoon

$$a^{p-1} \cdot (p-1)! \equiv (p-1)! \pmod{p}. \tag{8}$$

Lopuksi voimme kertoa kongruenssin (8) puolittain luvun $(p-1)!$ käänteisalkiolla, mistä seuraa

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

eli lause on saatu todistettua. □

Seuraus 2.37. Jos $b \equiv c \pmod{p-1}$, niin $a^b \equiv a^c \pmod{p}$.

Todistus. Oletuksesta $b \equiv c \pmod{p-1}$ seuraa, että $(p-1)|(b-c)$, eli $b = c + (p-1)d$, missä d on jokin kokonaisluku. Nyt voimme muokata lauseketta a^b seuraavalla tavalla:

$$\begin{aligned} a^b \pmod{p} &\equiv a^{c+(p-1)d} \pmod{p} \\ &\equiv a^c \cdot (a^{p-1})^d \pmod{p} \\ &\equiv a^c \cdot 1^d \pmod{p} \textbf{ (Fermat'n lause)} \\ &\equiv a^c \pmod{p}. \end{aligned}$$

□

Lause 2.38. Olkoon p alkuluku. Tällöin joukossa $\mathbb{Z}_p^*$ on olemassa jokin alkio g, jonka potensseina saadaan kaikki joukon muut alkiot. Siis

$$\mathbb{Z}_p^* = \{1, g, g^2, g^3, \ldots, g^{p-2}\}.$$

Alkiota jolla on tällainen ominaisuus sanotaan joukon $\mathbb{Z}_p^*$ *generoijaksi* tai *primitiiviseksi juureksi.*

Lauseen todistusta ei tässä yhteydessä esitetä sen pituuden vuoksi, mutta se löytyy sähköisenä esimerkiksi lähteestä [7]. Myöhemmin ElGamal-kryptosysteemistä puhuttaessa on kuitenkin hyvä tietää tällaisten primitiivisten juurten olemassaolo.

Määritelmä 2.39 (Diskreetin logaritmin ongelma)**.** Olkoon g joukon $\mathbb{Z}_p^*$ generoija ja h jokin kyseisen joukon nollasta poikkeava alkio. Tällöin käytämme nimitystä *diskreetin logaritmin ongelma* ongelmasta, jossa tehtävänä on löytää kongruenssin

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

toteuttava luku x. Kyseinen x on luvun h g-kantainen *diskreetti logaritmi*, ja siitä käytetään merkintää $\log_g(h)$.

3 Julkisen avaimen kryptosysteemeistä

Tässä luvussa siirrymme viimeinkin käsittelemään itse kryptografiaa. Tarkasteltavaksi on valittu RSA-, Rabin- ja ElGamal-kryptosysteemit, joita ennen lukijalle esitellään yleisemmällä tasolla kryptografian tarkoitusta sekä olennaisimpia käytettäviä termejä ja käsitteitä. Luku perustuu lähteisiin [5] ja [2].

3.1 Kryptografiasta yleisesti

Kryptografian perimmäinen tarkoitus on mahdollistaa kahden henkilön – olkoot heidän nimensä englanninkielistä lähdemateriaalia kunnioittaen Alice ja Bob – keskustelu jonkin epäluotettavan viestintäkanavan kautta siten, että kukaan ulkopuolinen taho ei saa selville, mistä he keskustelevat. Epäluotettava tarkoittaa tässä yhteydessä sitä, että kolmannet osapuolet pystyvät vapaasti seuraamaan ja lukemaan kaikkea kanavan kautta kulkevaa viestintää.

Aluksi on syytä ottaa käyttöön muutamia termejä. Alicen Bobille lähettämästä viestistä käytetään nimitystä *selkoteksti* (engl. *plaintext*). Ennen viestinsä lähettämistä Alice *salaa* (*encrypt*) selkotekstin käyttämällä jotain ennalta valitsemaansa salausavainta, jolloin hän saa tuloksena *salatekstin* (*ciphertext*, usein myös *cryptotext*, "kryptoteksti"). Alice lähettää salatekstin epäluotettavan viestintäkanavan kautta Bobille, eikä kukaan kolmas osapuoli pysty kryptotekstin perusteella päättelemään, minkä viestin selkoteksti sisältää. Bob sen sijaan pystyy viestin laillisena vastaanottajana käyttämään hänen tiedossaan olevaa purkamisavainta ja *purkamaan* viestin salauksen (*decrypt*, "dekryptata"), jolloin hän pystyy lukemaan alkuperäisen selkotekstin.

Tämän lähtöasetelman pohjalta voidaan ottaa käyttöön seuraava määritelmä.

Määritelmä 3.1. *Kryptosysteemi* on viisikko $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$, joka toteuttaa seuraavat ehdot:

(i) $\mathcal{P}$ on *selkotekstiavaruus* eli kaikkien mahdollisten selkotekstien joukko,

(ii) $\mathcal{C}$ on *salatekstiavaruus* eli kaikkien mahdollisten salatekstien joukko,

(iii) $\mathcal{K}$ on *avainavaruus* eli kaikkien mahdollisten avainten joukko ja

(iv) jokaista avainta $K \in \mathcal{K}$ kohti on olemassa salausfunktio $e_K \in \mathcal{E}$ ja sitä vastaava purkamisfunktio $d_K \in \mathcal{D}$, jotka toteuttavat ehdon $d_K(e_K(x)) = x$ jokaiselle selkotekstille $x \in \mathcal{P}$.

Jotta viesteille voidaan suorittaa salaukseen vaadittavat laskuoperaatiot, on kirjaimet muutettava ensin numeroiksi. Yksinkertaisin tapa tähän on se, että merkitsemme $A = 0, B = 1, C = 2, \ldots$ ja käymme näin läpi kaikki

aakkosten kirjaimet. Halutessaan myös välilyönnille voi antaa oman numeroarvon, mutta jatkossa salaamme kaikki selkotekstit ilman välilyöntejä. Kirjaimet siis muuttuvat numeroiksi alla olevan taulukon mukaisesti.

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

Taulukko 1: Englannin aakkosten numeroarvot.

Kryptosysteemit voidaan karkeasti jakaa kahteen eri pääluokkaan, joista ensimmäisen muodostavat "perinteiset"eli symmetriset tai salaisen avaimen (*private-key*) kryptosysteemit ja toisen epäsymmetriset eli *julkisen avaimen* (*public-key*) kryptosysteemit. Symmetrisissä kryptosysteemeissä joko käytetään samaa avainta sekä viestin salaamiseen että purkamiseen, tai purkamisavaimen päättely on triviaalia, jos salausavain on tiedossa. Julkisen avaimen kryptosysteemeissä sen sijaan purkamisavainta ei pysty päättelemään salausavaimesta, jonka vuoksi salausavain voi olla täysin julkista tietoa, mistä menetelmää käyttävät kryptosysteemit ovat saaneet nimensäkin.

Julkisen avaimen kryptosysteemeissä on selkeä etu symmetrisiin systeemeihin nähden. Jos Alice ja Bob haluavat käyttää symmetristä salausta viestinnässään, heidän on ennen viestien lähettämistä sovittava etukäteen käytettävästä avaimesta. Tämä ei kuitenkaan ole aina mahdollista, sillä Alicella ja Bobilla ei välttämättä ole mahdollisuutta tavata toisiaan henkilökohtaisesti, ja salaisten avainten sopiminen epäluotettavan viestintäkanavan kautta tekee koko salauksesta hyödytöntä. Julkisen avaimen salausta käyttämällä tämä ei kuitenkaan ole ongelma, sillä Alice voi vapaasti kertoa Bobille ja kenelle tahansa keskustelua seuraavalle kolmannelle osapuolelle julkisen salausavaimen, jolla salatut viestit ovat ainoastaan Alicen purettavissa. Bob voi vastaavalla tavalla julkistaa oman salausavaimensa, jolloin he voivat turvallisesti lähettää viestejä toisilleen ilman, että heidän täytyy etukäteen sopia mitään.

Toimintaperiaatteen ajatusta voi selkeyttää vaikkapa seuraavalla laatikko ja lukko -vertauksella. Alice vie julkiselle paikalle kaikkien saataville lukkoja, joihin vain hänellä itsellään on avain. Viestien lähettäjät voivat laittaa viestinsä laatikkoon, jonka lukitsevat Alicen antamalla lukolla ja lähettävät sitten hänelle. Vaikka kaikki tietävät, minkälaisella lukolla laatikot on lukittu, kukaan ei silti pysty avaamaan niitä, sillä heillä ei ole lukkoon sopivaa avainta.

Matematiikassa tällaiseen salaukseen käytetään niin sanottuja *yhdensuun-*

taisia funktioita (*one-way function*). Yhdensuuntaisille funktioille on ominaista se, että laskutoimituksen suorittaminen funktion lähtöjoukosta maalijoukkoon on helppoa, mutta käänteinen operaatio on liian työläs laskettavaksi missään käyttökelpoisessa ajassa. Tällaiset funktiot eivät kuitenkaan yksinään vielä riitä, vaan lisäksi on oltava olemassa jonkinlainen *takaportti* (*trapdoor*) eli ylimääräinen salassapidettävä tieto, joka mahdollistaa käänteisen operaation laskemisen helposti viestin lailliselle vastaanottajalle.

Seuraavissa osioissa tarkastelemme muutamia yleisimpiä julkisen avaimen kryptosysteemejä.

3.2 RSA-kryptosysteemi

RSA (nimetty kehittäjiensä Rivestin, Shamirin ja Adlemanin mukaan) on kenties tunnetuin ja käytetyin julkisen avaimen kryptosysteemi. RSA:n toiminta perustuu siihen, että kun p ja q ovat satunnaisesti valittuja suuria alkulukuja, luvun $N = pq$ tekijöihinjako ei ole käytännössä mahdollista nykyisillä laskenta-algoritmeilla kenenkään sellaisen henkilön toimesta, jonka tiedossa on ainoastaan luku N.

Asetelma on seuraavanlainen.

1. Alice valitsee ensin kaksi suurta alkulukua p ja q ja laskee niiden tulon $N = pq$. Seuraavaksi hän valitsee jonkin luvun e, joka toteuttaa ehdon $\text{syt}(e, (p-1)(q-1)) = 1$. Alice julkistaa parin (e, N), joka on systeemin julkinen salausavain, ja pitää itsellään parin (p, q), joka on hänen salainen avaimensa.

2. Bob esittää haluamansa viestin m kokonaislukuna ja laskee $c = m^e \pmod{N}$, jonka hän sitten lähettää Alicelle.

3. Alice laskee ensin luvun $d = e^{-1} \pmod{(p-1)(q-1)}$ ja purkaa sen jälkeen Bobin viestin salauksen laskemalla $c^d \equiv m \pmod{N}$. Luvun d tehokas laskeminen vaatii tekijöiden p ja q tietämisen, jonka vuoksi salauksen purkaminen ei onnistu vastaavalla tavalla kenenkään ulkopuolisen tahon toimesta.

Seuraavat kaksi lausetta osoittavat, että RSA:n salauksen purkaminen todella toimii niin kuin sen kuuluukin.

Lause 3.2 (Eulerin lause)**.** Olkoot p ja q alkulukuja ja $d = \text{syt}(p-1, q-1)$. Tällöin

$$a^{(p-1)(q-1)/d} \equiv 1 \pmod{pq},$$

kun $\text{syt}(a, pq) = 1$. Jos p ja q ovat parittomia, niin

$$a^{(p-1)(q-1)/2} \equiv 1 \pmod{pq},$$

kun $\text{syt}(a, pq) = 1$.

Todistus. Oletusten perusteella tiedämme, että $p \nmid a$ ja $d|(q-1)$, joten

$$a^{(p-1)(q-1)/d} = \left(a^{(p-1)}\right)^{(q-1)/d}.$$

Nyt Fermat'n pienen lauseen perusteella

$$\begin{aligned}\left(a^{(p-1)}\right)^{(q-1)/d} &\equiv 1^{(q-1)/d} \pmod{p} \\ &\equiv 1 \pmod{p}.\end{aligned}$$

Täsmälleen samoilla laskuilla voimme osoittaa, että $a^{(p-1)(q-1)/d} \equiv 1 \pmod{q}$. Nyt siis $a^{(p-1)(q-1)/d} - 1$ on jaollinen molemmilla luvuista p ja q, eli se on myös jaollinen niiden tulolla pq. □

Lause 3.3. Olkoot p ja q alkulukuja ja $e \geq 1$ jokin luku, joka toteuttaa ehdon $\text{syt}(e, (p-1)(q-1)) = 1$. Tällöin

$$(c^e)^d \equiv c \pmod{pq}, \tag{9}$$

kun $ed \equiv 1 \pmod{(p-1)(q-1)}$.

Todistus. Todistettavana on kaksi eri tapausta. Oletetaan ensin, että $\text{syt}(c, pq) = 1$. Kongruenssin $ed \equiv 1 \pmod{(p-1)(q-1)}$ perusteella on olemassa jokin kokonaisluku k, joka toteuttaa yhtälön

$$de = 1 + k(p-1)(q-1).$$

Nyt voimme muokata kongruenssin (9) vasemman puolen muotoon

$$\begin{aligned}(c^e)^d &\equiv c^{ed} \pmod{pq} \\ &\equiv c^{1+k(p-1)(q-1)} \pmod{pq} \\ &\equiv c \cdot (c^{(p-1)(q-1)})^k \pmod{pq}.\end{aligned}$$

Eulerin lauseen perusteella kongruenssin oikea puoli muuttuu muotoon $c \cdot 1^k \pmod{pq} \equiv c \pmod{pq}$, eli $(c^e)^d \equiv c \pmod{pq}$ kuten haluttiinkin.

Seuraavaksi on vielä tutkittava tapausta, missä $\text{syt}(c, pq) > 1$, jolloin Eulerin lausetta ei voida käyttää kuten edellä. Oletuksesta kuitenkin seuraa,

että $\text{syt}(c, pq)$ on joko p, q tai pq. Viimeisenä mainittu tapaus on triviaali, joten tarkastelemme seuraavaksi ainoastaan kahta ensin mainittua tilannetta. Näiden kahden tapauksen todistukset ovat identtiset, joten valitaan $\text{syt}(c, pq) = p$. Tästä seuraa, että $\text{syt}(c, q) = 1$.

Kiinalaisen jäännöslauseen perusteella tiedämme, että jos todistamme molemmat yhtälöistä

$$\begin{aligned} (c^e)^d &\equiv c \pmod{p} \text{ ja} \\ (c^e)^d &\equiv c \pmod{q}, \end{aligned} \tag{10}$$

niin tällöin myös $(c^e)^d \equiv c \pmod{pq}$.

Oletuksesta $\text{syt}(c, pq) = p$ seuraa, että jokin kokonaisluku j toteuttaa yhtälön $c = jp$, jolloin $(c^e)^d = ((jp)^e)^d$. Molemmat termeistä c ja $(c^e)^d$ ovat siis jaollisia luvulla p, jolloin kongruenssiparin (10) ensimmäinen kongruenssi muuttuu muotoon $0 \equiv 0 \pmod{p}$, mikä selvästi pitää paikkansa.

Jälkimmäisen kongruenssin kohdalla toimitaan hyvin samalla tavalla kuin aikaisemmassa tapauksessa, missä oli voimassa $\text{syt}(c, pq) = 1$. Kongruenssin vasen puoli $(c^e)^d$ saadaan muokattua muotoon

$$c \cdot (c^{(q-1)})^{k(p-1)} \pmod{q},$$

joka edelleen Fermat'n pienen lauseen perusteella muuttuu muotoon

$$c \cdot 1^{k(p-1)} \equiv c \pmod{q},$$

eli molemmat kongruensseista (10) ovat voimassa. □

Havainnollistetaan seuraavaksi RSA:n toimintaa laskuesimerkillä. Kyseisessä esimerkissä on syytä ottaa huomioon, että käytetyt luvut ovat aivan liian pieniä oikean salaamisen näkökulmasta[2], ja tarkoituksena onkin vain selkeyttää systeemin käyttöä.

Esimerkki 3.4 (RSA)**.** Alice valitsee ensin alkuluvut $p = 67$ ja $q = 73$, jolloin $N = 67 \cdot 73 = 4891$ ja $(p-1)(q-1) = 66 \cdot 72 = 4752$. Seuraavaksi hän valitsee luvun $e = 61$, joka toteuttaa ehdon $\text{syt}(61, 4752) = 1$. Tämän jälkeen Alice julkistaa parin $(61, 4891)$.

Bob haluaa lähettää Alicelle viestin *CTR CONFIRMED*. Hän pilkkoo viestinsä kahden kirjaimen osiin, jotta $m < N$, ja sen jälkeen suorittaa tarvittavat laskut käyttämällä luvussa 2 esiteltyä potenssiinkorotusalgoritmia.

[2]Tämänhetkinen suositus käytettävän avaimen koolle on 2048 bittiä [1], ja sama suositus pätee myös muille tässä tutkielmassa käsiteltäville kryptosysteemeille.

C	T	R	C	O	N	F	I	R	M	E	D
02	19	17	02	14	13	05	08	17	12	04	03

Taulukko 2: Bobin viesti muutettuna numeroiksi.

m	219	1702	1413	508	1712	403
m^2	3942	1332	1041	3732	1235	1006
m^4	657	3682	2770	3147	4124	4490
m^8	1241	4163	3812	4225	1369	4289
m^{16}	4307	1756	183	3366	908	470
m^{32}	3577	2206	4143	2400	2776	805
$m^{61} = m^{32}m^{16}m^8m^4m$	3285	914	84	2552	2583	3954

Taulukko 3: Bobin suorittamat laskut modulo 4891.

Lopuksi Bob lähettää Alicelle salatekstin

$$c = (3285, 914, 84, 2552, 2583, 3954).$$

Vastaanotettuaan Bobin lähettämän viestin Alice laskee ensin $d = 61^{-1} \pmod{4752} = 3973$ ja sen avulla

$$\begin{aligned} m &= (3285^{3973},\ 914^{3973},\ 84^{3973},\ 2552^{3973},\ 2583^{3973},\ 3954^{3973}) \\ &\equiv (219, 1702, 1413, 508, 1712, 403) \pmod{4891} \end{aligned}$$

käyttämällä samaa potenssiinkorotusalgoritmia, jolloin hän pystyy lukemaan alkuperäisen selkotekstin.

Huomautus 3.5. Potenssiinkorotuksessa modulo m voidaan käyttää myös apuna kiinalaista jäännöslausetta. Olkoot p ja q alkulukuja ja lisäksi $N = pq$, $a \in \mathbb{Z}_N$ ja $d < N$. Tehtävänä on laskea $a^d \pmod{N}$.

Otetaan käyttöön merkinnät $a_p \equiv a \pmod{p}$, $a_q \equiv a \pmod{q}$, $d_p \equiv d \pmod{p-1}$ ja $d_q \equiv d \pmod{q-1}$. Lasketaan ensin $a_p^{d_p} \pmod{p}$ ja $a_q^{d_q} \pmod{q}$. Nyt voimme kiinalaisen jäännöslauseen avulla ratkaista kongruenssiparin

$$\begin{cases} x &\equiv a_p^{d_p} \pmod{p} \\ x &\equiv a_q^{d_q} \pmod{q}, \end{cases}$$

jonka yksikäsitteinen ratkaisu on kaivattu $a^d \pmod{N}$. Huomioitavaa on se, että algoritmin käyttö vaatii luvun N tekijöiden olevan tiedossa, jonka vuoksi esimerkiksi RSA:n tapauksessa sen käyttö onnistuu ainoastaan viestin lailliselta vastaanottajalta.

3.3 Rabin-kryptosysteemi

Rabin-kryptosysteemi perustuu RSA:n tavoin luvun $N = pq$ tekijöihinjaon vaikeuteen ja on muutenkin hyvin samankaltainen pienillä muutoksilla. Tässä kryptosysteemissä viestin salaaminen tapahtuu korottamalla viesti potenssiin 2, jolloin salauksen purkaminen luonnollisesti tehdään laskemalla neliöjuuri salatekstistä. Kuten reaalilukujen kohdalla yleisesti, kokonaisluvun a neliöjuurella modulo m tarkoitetaan sellaista lukua b, joka toteuttaa kongruenssin $b^2 \equiv a \pmod{m}$. Huomionarvoista kuitenkin on, että modulista m riippuen luvulla a voi olla useita edellisen ehdon täyttäviä lukuja b.

Tarkastellaan seuraavaksi yksityiskohtaisemmin, kuinka viestien salaaminen ja purkaminen toimii Rabin-kryptosysteemissä.

1. Alice valitsee kaksi suurta kokonaislukua p ja q ja laskee niiden tulon $N = pq$. Neliöjuurten laskemisen helpottamiseksi valitaan usein sellaiset luvut, jotka toteuttavat kongruenssin $p \equiv q \equiv 3 \pmod{4}$, jolloin $\sqrt{c} \equiv c^{\frac{p+1}{4}} \pmod{p}$ (tämä osoitetaan paikkansapitäväksi hieman edempänä). Alice julkistaa luvun N, joka on systeemin julkinen salausavain.

2. Bob muuttaa haluamansa viestin m numeroksi ja laskee $c \equiv m^2 \pmod{N}$, jonka hän lähettää Alicelle.

3. Alice laskee edellä esitetyllä laskukaavalla neliöjuuret modulo p ja q ja yhdistää ne käyttämällä kiinalaista jäännöslausetta, jolloin hän saa selville neliöjuuret modulo N ja täten Bobin kirjoittaman alkuperäisen selkotekstin.

Rabin-kryptosysteemin etu RSA:han nähden on se, että sen murtamisen on todistettu olevan yhtä haastavaa kuin luvun N tekijöihinjaon, johon ei siis toistaiseksi ole olemassa mitään käyttökelpoisessa ajassa toimivia algoritmeja. RSA:n kohdalla tällaista todistusta ei ole, mutta väitteen puolesta on runsas määrä numeerista näyttöä.

Valitettavasti Rabin-kryptosysteemissä on myös eräs selkeä haittapuoli RSA:han nähden, joka on kenties syy siihen, miksi se ei ole koskaan yleistynyt samalla tavalla. Salauksen purkamisessa viestin vastaanottaja laskee neliöjuuret modulo p ja q, joita on kaksi kappaletta kummankin modulin kohdalla, jolloin neliöjuuria modulo $N = pq$ on yhteensä neljä kappaletta. Tämä tarkoittaa sitä, että jopa viestin lailliselle vastaanottajalle mahdollisia selkotekstejä on neljä eri kappaletta, joista oikea on osattava valita jollakin keinolla. Oikean selkotekstin poimiminen on todennäköisesti helppoa, jos kyseessä on jokin merkityksellinen luonnollisen kielen viesti, jonka vastaanottaja pystyy helposti tunnistamaan, mutta mahdollisten selkotekstien määrä

muodostuu ongelmaksi viestin ollessa numerosarja, jokin mielivaltainen kirjainyhdistelmä tai vaikkapa sellaista kieltä, mitä vastaanottaja ei ymmärrä. Tämän ongelman pystyy kiertämään lisäämällä viestiin jollakin tavalla redundanssia, mutta se luonnollisesti vaatii ylimääräistä työtä sekä lähettäjältä että vastaanottajalta.

Osoitetaan vielä ennen laskuesimerkkiä, että systeemin esittelyssä mainittu laskuvaaka neliöjuurten laskemiseksi toimii, eli $(c^{\frac{p+1}{4}})^2 = c \pmod{p}$.

$$\begin{aligned}
(c^{\frac{p+1}{4}})^2 &\equiv c^{\frac{p+1}{2}} \pmod{p} \\
&\equiv c \cdot c^{\frac{1}{2}(p-1)} \pmod{p} \\
&\equiv c \cdot (m^2)^{\frac{1}{2}(p-1)} \pmod{p} \qquad (11) \\
&\equiv c \cdot m^{p-1} \pmod{p} \\
&\equiv c \cdot 1 \pmod{p} \qquad (12) \\
&\equiv c \pmod{p}
\end{aligned}$$

Yllä olevissa kongruensseissa kohdassa (11) on käytetty luvun c määritelmää ja kohta (12) seuraa Fermat'n pienestä lauseesta.

Esimerkki 3.6 (Rabin)**.** Alice valitsee salaiseksi avaimekseen alkuluvut $p = 307$ ja $q = 311$, jotka laskujen helpottamiseksi toteuttavat kongruenssin $307 \equiv 311 \equiv 3 \pmod{4}$. Valituista alkuluvuista saadaan laskettua julkinen salausavain $N = pq = 307 \cdot 311 = 95477$, jonka Alice julkistaa. Lisäksi Alice ilmoittaa, että oikean selkotekstin tunnistamiseksi lähetettävän viestin binääriesityksessä on toistettava viimeiset kuusi bittiä.

Bob haluaa lähettää Alicelle viestin $m = 1050$. Hän muuttaa ensin viestin binäärimuotoon, josta hän muodostaa luvun x toistamalla viimeiset kuusi bittiä Alicen ohjeiden mukaisesti.

$$\begin{aligned}
1050_{10} &= 10000\mathbf{011010}_2, \\
x &= 10000011010\mathbf{011010}_2 \\
&= 67226_{10}.
\end{aligned}$$

Lopuksi Bob laskee

$$\begin{aligned}
c &= x^2 \pmod{N} \\
&= 67226^2 \pmod{95477} \\
&\equiv 26758 \pmod{95477}
\end{aligned}$$

ja lähettää Alicelle salatekstin $c = 26758$.

Vastaanotettuaan Bobin viestin Alice laskee luvun 26758 neliöjuuret modulo $p = 307$ ja $q = 311$ aikaisemmin esiteltyä laskukaavaa käyttämällä. Käytetään näistä neliöjuurista merkintöjä m_p ja m_q.

$$\begin{aligned} m_p &= 26758^{\frac{307+1}{4}} \pmod{307} \equiv \mathbf{7} \pmod{307}, \\ -m_p &= -7 + 307 \equiv \mathbf{300} \pmod{307}, \\ m_q &= 26758^{\frac{311+1}{4}} \pmod{311} \equiv \mathbf{50} \pmod{311}, \\ -m_q &= -50 + 311 \equiv \mathbf{261} \pmod{311}. \end{aligned}$$

Näistä neljästä neliöjuuresta Alice muodostaa seuraavat neljä kongruenssiparia, jotka hän ratkaisee käyttämällä kiinalaista jäännöslausetta.

$$\begin{cases} m \equiv 7 \pmod{307} \\ m \equiv 50 \pmod{311} \end{cases} \tag{13}$$

$$\begin{cases} m \equiv 300 \pmod{307} \\ m \equiv 50 \pmod{311} \end{cases} \tag{14}$$

$$\begin{cases} m \equiv 7 \pmod{307} \\ m \equiv 261 \pmod{311} \end{cases} \tag{15}$$

$$\begin{cases} m \equiv 300 \pmod{307} \\ m \equiv 261 \pmod{311} \end{cases} \tag{16}$$

Ratkaistuaan jokaisen kongruenssiparin Alice muuttaa vastaukset binäärimuotoon ja tarkastelee, mikä niistä toteuttaa vaaditun ehdon lopussa toistetuista biteistä.

Kongruenssiparin (13) ratkaisu : $20576_{10} = 101000001100000_2$
Kongruenssiparin (14) ratkaisu : $67226_{10} = 10000\mathbf{011010011010}_2$
Kongruenssiparin (15) ratkaisu : $28251_{10} = 110111001011011_2$
Kongruenssiparin (16) ratkaisu : $74901_{10} = 10010010010010101_2$

Tässä vaiheessa Alice toteaa, että kongruenssiparin (14) ratkaisu on luvuista ainut, jonka binääriesityksessä kuusi viimeistä bittiä ovat samat kuin sitä edeltävät kuusi bittiä, joten sen on oltava oikea viesti. Alice poistaa ylimääräiset lisätyt bitit ja muuttaa luvun takaisin kymmenjärjestelmään, jolloin hän saa luettua alkuperäisen selkotekstin $m = 10000011010_2 = 1050_{10}$.

3.4 ElGamal-kryptosysteemi

ElGamal-kryptosysteemi on RSA:n ohella todennäköisesti tunnetuin julkisen avaimen kryptosysteemi, ja se perustuu määritelmässä 2.39 esiteltyyn diskreetin logaritmin ongelmaan. Kolmansien osapuolien olisi salauksen murtamiseksi ratkaistava kyseinen ongelma julkisten parametrien perusteella, mikä ei nykytiedon perusteella onnistu realistisessa ajassa, jos käytettävät luvut ovat riittävän suuria. Seuraavaksi tarkastelemme vaihe kerrallaan, kuinka viestin salaus ja salauksen purkaminen toimii ElGamal-kryptosysteemissä.

1. Ensiksi on valittava jokin suuri alkuluku p ja joukon $\mathbb{Z}_p^*$ primitiivinen juuri g. Alice voi valita nämä luvut itse tai ne voivat olla ennalta valittuja jonkin luotetun tahon toimesta. Luvut p ja g ovat julkisia ja siis kaikkien osapuolten tiedossa.

2. Alice valitsee salaiseksi avaimekseen jonkin kokonaisluvun a väliltä $[1, p-1]$ ja laskee $A = g^a \pmod{p}$. Kyseinen A on systeemin julkinen salausavain, jonka Alice julkistaa.

3. Bob valitsee ensin jonkin satunnaisen luvun k, joka on hänen salainen kertakäyttöinen avaimensa. Muutettuaan viestinsä m numeroksi Bob laskee lukujen A ja k avulla $c_1 = g^k \pmod{p}$ ja $c_2 = mA^k \pmod{p}$ ja lähettää Alicelle salatekstin (c_1, c_2).

4. Vastaanotettuaan Bobin lähettämän viestin Alice purkaa salauksen laskemalla $(c_1^a)^{-1} \cdot c_2 \pmod{p}$, joka on alkuperäinen Bobin kirjoittama selkoteksti.

Osoitetaan vielä, että salauksen purkaminen toimii niin kuin sen kuuluukin muokkaamalla Alicen laskemaa lauseketta $(c_1^a)^{-1} \cdot c_2 \pmod{p}$. Kryptosysteemin rakenteesta seuraa suoraan, että

$$\begin{aligned}(c_1^a)^{-1} \cdot c_2 &\equiv ((g^k)^a)^{-1} \cdot (mA^k) \pmod{p} \\ &\equiv (g^{ak})^{-1} \cdot (m(g^a)^k) \pmod{p} \\ &\equiv m \pmod{p},\end{aligned}$$

koska toiseksi viimeisellä rivillä luvut $(g^{ak})^{-1}$ ja g^{ak} kumoavat toisensa.

Havainnollistetaan seuraavaksi ElGamal-salauksen toimintaa esimerkillä. Jälleen kerran tässä yhteydessä on syytä muistuttaa, että esimerkissä käytetyt luvut ovat aivan liian pieniä vakavasti otettavaan viestien salaukseen.

Esimerkki 3.7 (ElGamal)**.** Valitaan ensin julkisiksi parametreiksi alkuluku $p = 2609$ ja primitiivinen juuri $g = 3$. Primitiivisten juurten löytämiseen

ei valitettavasti ole minkäänlaista valmista laskukaavaa, joten ainut keino on kokeilla potentiaalisia lukuja ja karsia vääriä valintoja pois yrityksen ja erehdyksen kautta.

Alice valitsee salaiseksi avaimekseen luvun $a = 949$ ja laskee arvon $A = 3^{949} \equiv 467 \pmod{2609}$. Tämän jälkeen Alice julkistaa luvun $A = 467$, joka on systeemin julkinen salausavain.

Bob haluaa lähettää Alicelle viestin *SOME SAY KOSM*. Hän muuttaa viestinsä numeroiksi ja pilkkoo sen kahden kirjaimen osiin[3], jotta $m < p$, ja saa näin ollen kuusi pienempää viestiä alla olevan taulukon mukaisesti.

S	O	M	E	S	A	Y	K	O	S	M	X
18	14	12	04	18	00	24	10	14	18	12	23

Taulukko 4: Bobin viesti muutettuna numeroiksi.

Parhaimman salauksen saavuttamiseksi Bob valitsee jokaista kuutta osaa varten oman kertakäyttöisen avaimensa käyttämällä satunnaislukugeneraattoria, joka antaa seuraavat arvot:

$$\begin{aligned} k_1 &= 1102, \quad k_2 = 886, \quad k_3 = 1048, \\ k_4 &= 2223, \quad k_5 = 1464, \quad k_6 = 603. \end{aligned}$$

Seuraavaksi Bob laskee lukujen $c_{1,i} \equiv g^{k_i} \pmod{p}$ ja $c_{2,i} \equiv m_i A^{k_i} \pmod{p}$ lukuarvot taulukon 5 mukaisesti.

$$\begin{aligned}
m_1 &= 1814, \quad c_{1,1} = 3^{1102} \equiv 561, \quad c_{2,1} = 1814 \cdot 467^{1102} \equiv 9 \\
m_2 &= 1204, \quad c_{1,2} = 3^{886} \equiv 1337, \quad c_{2,2} = 1204 \cdot 467^{886} \equiv 1512 \\
m_3 &= 1800, \quad c_{1,3} = 3^{1048} \equiv 1952, \quad c_{2,3} = 1800 \cdot 467^{1048} \equiv 545 \\
m_4 &= 2410, \quad c_{1,4} = 3^{2223} \equiv 955, \quad c_{2,4} = 2410 \cdot 467^{2223} \equiv 1984 \\
m_5 &= 1418, \quad c_{1,5} = 3^{1464} \equiv 2250, \quad c_{2,5} = 1418 \cdot 467^{1464} \equiv 2338 \\
m_6 &= 1223, \quad c_{1,6} = 3^{603} \equiv 1920, \quad c_{2,6} = 1223 \cdot 467^{603} \equiv 2028
\end{aligned}$$

Taulukko 5: Bobin suorittamat laskut modulo 2609.

Lopuksi Bob lähettää Alicelle salatekstin

$$c = ((561, 9), (1337, 1512), (1952, 545), (955, 1984), (2250, 2338), (1920, 2028)).$$

Vastaanotettuaan Bobin viestin Alice laskee lukuarvot $c_{1,i}^a$ ja sen jälkeen kyseisten lukujen käänteisalkiot modulo 2609.

[3]Alice ja Bob ovat voineet tässä tapauksessa esimerkiksi sopia etukäteen, että kirjain X viestin lopussa tarkoittaa tyhjää merkkiä.

$$
\begin{array}{lll}
c_{1,1} = 561, & c_{1,1}^a = 561^{949} \equiv 2481, & 2481^{-1} \equiv 1651 \\
c_{1,2} = 1337, & c_{1,2}^a = 1337^{949} \equiv 608, & 608^{-1} \equiv 339 \\
c_{1,3} = 1952, & c_{1,3}^a = 1952^{949} \equiv 138, & 138^{-1} \equiv 1607 \\
c_{1,4} = 955, & c_{1,4}^a = 955^{949} \equiv 698, & 698^{-1} \equiv 998 \\
c_{1,5} = 2250, & c_{1,5}^a = 2250^{949} \equiv 193, & 193^{-1} \equiv 1487 \\
c_{1,6} = 1920, & c_{1,6}^a = 1920^{949} \equiv 2280, & 2280^{-1} \equiv 1134
\end{array}
$$

Taulukko 6: Alicen suorittamat laskut modulo 2609.

Lopuksi Alice purkaa salaukset laskemalla $(c_{1,i}^a)^{-1} \cdot c_{2,i} \pmod{2609}$ kaikilla luvun i arvoilla, jolloin saadaan alkuperäinen viesti

$$
\begin{aligned}
m &= (1651 \cdot 9,\ 339 \cdot 1512,\ 1607 \cdot 545,\ 998 \cdot 1984,\ 1487 \cdot 2338,\ 1134 \cdot 2028) \\
&\equiv (1814, 1204, 1800, 2410, 1418, 1223) \pmod{2609}.
\end{aligned}
$$

4 Digitaaliset allekirjoitukset

Julkisen avaimen kryptosysteemien kenties tärkein sovellus on niiden käyttö *digitaalisiin allekirjoituksiin*. Digitaalisten allekirjoitusten tarkoitus on mahdollistaa sähköisten dokumenttien allekirjoittaminen siten, että kuka tahansa pystyy myöhemmin todentamaan kyseisen allekirjoituksen ja siten varmistumaan viestin lähettäjästä. Jotta tällainen allekirjoitus olisi käyttökelpoinen, sen tulisi täyttää seuraavat kolme ehtoa:

1. allekirjoituksia ei voi väärentää (tai väärentäminen olisi vähintäänkin liian työlästä missään käyttökelpoisessa ajassa),
2. allekirjoitetun viestin sisältöä ei voi muokata jälkeenpäin ja
3. viestin lähettäjä ei myöhemmin voi kieltää allekirjoittaneensa viestiä.

Digitaaliset allekirjoitukset toimivat hyvin samankaltaisella periaatteella kuin viestien salaaminen julkisen avaimen kryptografiassa. Viestin allekirjoittaja käyttää jonkinlaista salaista *allekirjoitusavainta* (*signing key*) muodostaakseen allekirjoituksen, jonka hän sitten lähettää valitsemansa dokumentin mukana. Viestin vastaanottaja tai kuka tahansa muu henkilö voi käyttää julkista *todentamisavainta* (*verification key*) tarkistaakseen, täyttääkö allekirjoitus tietyn käytetystä menetelmästä riippuvan ehdon, jonka toteutuessa hän hyväksyy allekirjoituksen aitouden. Allekirjoittaminen perustuu viestin salaamisen tapaan yhdensuuntaisiin funktioihin, eli allekirjoitusten väärentäminen vaatisi kolmansilta osapuolilta esimerkiksi diskreetin logaritmin ongelman ratkaisemista.

Tarkastellaan seuraavaksi lyhyesti kahta erilaista allekirjoitusmenetelmää, jotka perustuvat RSA- ja ElGamal-kryptosysteemeihin. Tämä luku pohjautuu lähteisiin [2] ja [4].

4.1 RSA-allekirjoitus

RSA-allekirjoituksen voi ajatella olevan ikään kuin käänteinen RSA-salaus. Viestin allekirjoittamiseksi Alice "purkaa" valitsemansa sähköisen dokumentin salauksen omalla salaisella avaimellaan, ja allekirjoituksen todentamiseksi Bob käyttää julkista todentamisavainta varmistaakseen, että laskutoimituksen vastauksena saadaan lähetetyn dokumentin numeroarvo. Yksityiskohtaisemmin menetelemä toimii seuraavalla tavalla.

1. Alice valitsee tuttuun tapaan kaksi suurta alkulukua p ja q ja laskee niiden tulon N, jonka hän julkistaa valitsemansa todentamisavaimen v kanssa, joka toteuttaa ehdon $\text{syt}(v, (p-1)(q-1)) = 1$. Koska Alice tietää luvun N tekijät, hän pystyy helposti löytämään luvun s, joka toteuttaa kongruenssin $sv \equiv 1 \pmod{(p-1)(q-1)}$. Tämä s on hänen salainen allekirjoitusavaimensa.

2. Olkoon Alicen dokumentti kokonaisluku D, joka toteuttaa ehdon $1 < D < N$. Allekirjoittaakseen tämän dokumentin Alice laskee luvun $S_A \equiv D^s \pmod{N}$ ja lähettää luvut D ja S_A Bobille.

3. Bob todentaa allekirjoituksen laskemalla $S_A^v \pmod{N}$ ja toteamalla, että tämän arvo on D. Lauseen 3.3 perusteella tiedämme, että

$$S_A^v \equiv D^{sv} \equiv D \pmod{N},$$

   eli todentaminen toimii kuten sen kuuluukin.

Lasketaan seuraavaksi vielä pieniä lukuja käyttämällä lyhyt laskuesimerkki salauksen toiminnasta.

Esimerkki 4.1 (RSA-allekirjoitus)**.** Olkoon lähetettävä dokumentti $D = 1487$. Alice valitsee alkuluvuikseen luvut $p = 59$ ja $q = 61$, jolloin $N = 59 \cdot 61 = 3599$. Lisäksi hän valitsee todentamisavaimeksi luvun $v = 1559$, jonka perusteella hän ratkaisee allekirjoitusavaimen $s = 2759$ kongruenssista $1559 \cdot s \equiv 1 \pmod{3480}$. Lopuksi Alice laskee

$$\begin{aligned} S_A &\equiv 1487^{2759} \pmod{3599} \\ &\equiv 2387 \pmod{3599} \end{aligned}$$

ja lähettää Bobille luvut $D = 1487$ ja $S_A = 2387$.

Vastaanotettuaan dokumentin ja allekirjoituksen Bob käyttää julkisia lukuarvoja $N = 3599$ ja $v = 1559$ laskeakseen

$$\begin{aligned} S_A^v \pmod N &\equiv 2387^{1559} \pmod{3599} \\ &\equiv 1487 \pmod{3599} \end{aligned}$$

ja hyväksyy allekirjoituksen, koska $S_A^v \equiv D \pmod N$.

4.2 ElGamal-allekirjoitus

RSA-allekirjoituksen toiminnasta poiketen ElGamal-allekirjoitus ei ole vain "käänteinen"ElGamal-salaus, mutta lähtöasetelmassa sekä muodostetussa allekirjoituksessa on silti selvät yhtäläisyydet taustalla olevaan kryptosysteemiin. Katsotaan seuraavaksi kohta kohdalta, kuinka ElGamal-allekirjoitukset ja niiden todentaminen toimivat.

1. Alice tai jokin luotettava kolmas osapuoli valitsee suuren alkuluvun p sekä primitiivisen juuren g modulo p, jotka ovat julkisia parametrejä. Alice valitsee salaiseksi allekirjoitusavaimekseen jonkin luvun s, ja laskee sen avulla todentamisavaimen $v \equiv g^s \pmod p$, joka julkistetaan lukujen p ja g kanssa.

2. Olkoon Alicen dokumentti kokonaisluku D, joka toteuttaa ehdon $1 < D < p$. Alice valitsee satunnaisesti kertakäyttöisen kokonaislukuavaimen k väliltä $1 < k < p$, laskee sen käänteisluvun $k^{-1} \pmod{p-1}$ ja lopuksi muodostaa allekirjoituksensa laskemalla lukuarvot S_1 ja S_2 seuraavien kaavojen avulla:

   $$\begin{aligned} &S_1 \equiv g^k \pmod p \text{ ja} \\ &S_2 \equiv (D - sS_1)k^{-1} \pmod{p-1}. \end{aligned}$$

   Pari (S_1, S_2) on Alicen allekirjoitus, jonka hän lähettää dokumentin D mukana.

3. Vastaanotettuaan Alicen lähettämät luvut Bob laskee $v^{S_1} S_1^{S_2} \pmod p$ sekä $g^D \pmod p$ ja hyväksyy allekirjoituksen, jos kyseiset lukuarvot ovat yhtäsuuret.

Katsotaan vielä ennen laskuesimerkkiä, miksi ElGamal-allekirjoituksen todentaminen toimii. Käyttämällä potenssien laskusääntöjä sekä lukujen v ja S_1 määritelmiä voimme kirjoittaa seuraavat kongruenssit:

$$\begin{aligned} v^{S_1} \cdot S_1^{S_2} &\equiv g^{sS_1} \cdot g^{kS_2} \pmod p \\ &\equiv g^{sS_1 + kS_2} \pmod p. \end{aligned}$$

Seurauksen 2.37 perusteella suorittaessamme laskutoimituksia modulo p voimme vaihtaa eksponentteja modulo $p-1$. Nyt luvun S_2 määritelmän perusteella voimme korvata sen lausekkeella $(D - sS_1)k^{-1}$, jolloin kongruenssien muokkaamista voidaan jatkaa seuraavalla tavalla:

$$\begin{aligned} g^{sS_1+kS_2} &\equiv g^{sS_1+k(D-sS_1)k^{-1}} \pmod{p} \\ &\equiv g^{sS_1+(D-sS_1)} \pmod{p} \\ &\equiv g^D \pmod{p}, \end{aligned}$$

eli todentaminen toimii kuten väitettiinkin.

Esimerkki 4.2 (ElGamal-allekirjoitus)**.** Olkoon Alicen sähköinen dokumentti lukuarvo $D = 1673$, ja valittakoon luvuiksi p ja g ElGamal-salauksen esimerkissä käytetyt $p = 2609$ sekä $g = 3$. Alice valitsee salaiseksi avaimekseen luvun $s = 458$ ja laskee sen avulla

$$\begin{aligned} v &\equiv 3^{458} \pmod{2609} \\ &\equiv 2141 \pmod{2609}, \end{aligned}$$

joka on systeemin julkinen todentamisavain.

Seuraavaksi Alice arpoo satunnaisen kokonaislukuavaimen $k = 329$ ja laskee $329^{-1} \pmod{2608} = 761$. Näiden lukujen avulla hän muodostaa allekirjoitukseen tarvittavat arvot S_1 ja S_2 seuraavien laskujen mukaisesti:

$$\begin{aligned} S_1 &\equiv 3^{329} \pmod{2609} \\ &\equiv 106 \pmod{2609} \text{ ja} \\ S_2 &\equiv (1673 - 458 \cdot 106) \cdot 761 \pmod{2608} \\ &\equiv 349 \pmod{2608}. \end{aligned}$$

Lopuksi Alice lähettää Bobille dokumentin $D = 1673$ ja allekirjoituksen $(S_1, S_2) = (106, 484)$.

Vastaanotettuaan Alicen viestin Bob suorittaa laskut

$$\begin{aligned} v^{S_1} S_1^{S_2} &\equiv 2141^{106} \cdot 106^{349} \pmod{2609} \\ &\equiv 102 \pmod{2609} \end{aligned}$$

sekä

$$\begin{aligned} g^D &\equiv 3^{1673} \pmod{2609} \\ &\equiv 102 \pmod{2609} \end{aligned}$$

ja hyväksyy allekirjoituksen, koska $v^{S_1} S_1^{S_2} \equiv g^D \pmod{p}$.

Lähteet

[1] E. Barker, Q. Dang: *NIST SP 800-57 Part 3 Rev. 1, Recommendation for Key Management, Part 3: Application-Specific Key Management Guidance.* `https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-57Pt3r1.pdf` (Luettu 19.3.2019)

[2] J. Hoffstein, J. Pipher, J. H. Silverman: *An Introduction to Mathematical Cryptography.* Springer, New York, 2008.

[3] C. T. Long: *Elementary Introduction to Number Theory, Third Edition.* Waveland Press, Inc., Illinois, 1995.

[4] A. Renvall: *Cryptography I.* Matematiikan ja tilastotieteen laitos, Turun yliopisto, 2012.

[5] A. Renvall, T. Meskanen: *Kryptografia II.* Matematiikan ja tilastotieteen laitos, Turun yliopisto, 2008.

[6] D. R. Stinson: *Cryptography: Theory and Practice.* CRC Press, Inc., Florida, 1995.

[7] A. Witno: *The Primitive Root Theorem.* `https://www.philadelphia.edu.jo/math/witno/notes/won5.pdf` (Luettu 29.12.2018)