Elliptkõverate krüptograafia

Jan Willemson

Cybernetica AS

1. Paar sõna krüptograafia ajaloost. Vajadus informatsiooni vahetada ning seejuures osaliselt varjata on peaaegu sama vana kui inimtsivilisatsioon. Esimesed märgid tahtlikult üldtuntud kirjaviisist kõrvalekaldumisest pärinevad juba Vana-Egiptusest ca 1900 eKr (tõsi, tolle kirjutaja eesmärgiks olid tõenäoliselt pigem kunstilis-religioossed taotlused). Viite salakirjutamisele leiame ka Piiblist, kus Jeremija raamatus seisab Paabeli asemel kahes kohas hoopis nimi Seesak, mis on Paabelist saadud tähtede väljavahetamise teel.

Ei läinud kaua, kui võimukandjad ning väejuhid mõistsid salakirjakunsti potentsiaali oma strateegiliste eesmärkide saavutamiseks. Järgnes sajanditepikkune võidujooks krüptograafide (salakirjutajate) ning krüptoanalüütikute (salakirjamurdjate) vahel, kus võitis see, kes suutis vastasest kavalamaid nõkse leiutada. Huvitatud lugeja leiab hea ülevaate krüptograafia ajaloost väliseesti majandus- ja keeleteadlase Nicolai Liventhali raamatust [8].

Šifrite loomise ja murdmise ajalugu näitas, et mida keerukam on krüptosüsteemi ehitus, seda raskem teda vastase eest varjata on. Seepärast sõnastas Hollandi krüptograaf Auguste Kerckhoffs[1] 19. sajandi lõpul printsiibi, mille tänapäevase tõlgenduse järgi peaks krüptogrammi turvalisus tuginema ainult võimalikult kompaktsele (ja sellisena paremini kaitstavale) saladusele, niinimetatud *võtmele* [7].

Kui suur võimalike võtmete hulk (nn *võtmeruum*) olema peab, sõltub vastase võimekusest. Selles osas tegi krüptoanalüüs suure sammu edasi II maailmasõja päevil koos elektronarvutite leiutamisega (kusjuures nende väljaarendamise üheks oluliseks eesmärgiks oligi just vastaste šifrite murdmine). Kulus veel

[1] Auguste Kerckhoffsi ei tohiks segi ajada saksa füüsiku Gustav Kirchhoffiga, kes sõnastas printsiibid elektrilaengu ja energia jäävuse kohta vooluahelates.

mõnikümmend aastat, enne kui arvutid masskasutusse jõudsid, aga umbkaudu 1960. – 70. aastateks oli selge, et salakirjutamisel ei saa enam loota sellele, et vastane ei suuda mõnda miljonit võimalikku võtmekombinatsiooni läbi vaadata. Otsinguruumi oli vaja märkimisväärselt suurendada ja selleks läks vaja süsteemsemat lähenemist.

Teine probleem, mis ajalooliste krüptosüsteemide kasutamisel järjest rohkem peavalu valmistas, oli võtmesümmeetria. See tähendab, et nii salakirja koostamiseks kui ka tema lugemiseks oli vaja sama võtit. Nõnda toimetades tuleb aga lahendada võtme levitamise probleem – kuidas tagada, et salajase sõnumi saajal on selle mõistmiseks vajalik võti olemas? Võtme edastamiseks füüsiliselt kokku saada pole alati võimalik. Teisalt aga ei saa turvalise kanali loomiseks alguses võtit üle ebaturvalise kanali ka saata, sest ründaja võib võtme pealt kuulata ning sellega hiljem kogu krüpteeritud kommunikatsiooni avada.

2. Krüptograafia rühmades. Esmapilgul tundub, et tegemist on muna ja kana probleemiga – turvalise sõnumikanali loomiseks läheb vaja turvalist võtmevahetuskanalit. Õnneks ei ole asi nii hull. Paradoksi lahendus peitub sõnas *turvaline*, mis on praegu kasutusel kahes erinevas tähenduses. Sõnumikanali turvalisuse all peame hetkel silmas eeskätt sõnumi *salajasust*, aga osutub, et võtmevahetuskanali puhul pole salajasust otseselt vaja. Piisab, kui see kanal on *autentne*, st suhtlevad osapooled suudavad kuidagi veenduda, et suhtluspartner on tõepoolest see, kelle ta väidab end olevat. Üle interneti toimivate sideseansside puhul pole ka autentsuse tagamise ülesanne tingimata lihtsasti lahenduv, aga see on teise artikli teema. Palun lugejal mind praegu uskuda, et teise (võibolla isegi täiesti võõra) osapoole autentimine on põhimõtteliselt võimalik.

Põhimõttelise lahenduse, kuidas üle potentsiaalselt pealt kuulatava kanali salajast võtit vahetada, pakkusid 1976. aastal välja Martin Hellman ja Whitfield Diffie [5]. Nende idee oli kasutada algarvulist järku lõpliku korpuse multiplikatiivset rühma $\mathbb{Z}_p^*$. Valime selle (teadupärast tsüklilise) rühma moodustaja g ning

edastame avalikult võtme kokku leppimisest huvitatud osapooltele, keda krüptograafia-alase kirjanduse traditsioonis nimetatakse Alice ja Bob.

Järgmiseks valivad Alice ja Bob endile kumbki oma juhusliku salajase naturaalarvu vahemikust $[1, p-1]$; olgu need arvud vastavalt a ja b. Alice arvutab rühmas $\mathbb{Z}_p^*$ väärtuse g^a ning saadab selle Bobile. Bob omakorda saadab Alice'ile tagasi g^b. Nüüd saab Alice arvutada väärtuse $(g^b)^a$ ning Bob $(g^a)^b$. Kuivõrd

$$(g^b)^a = (g^a)^b \,, \tag{1}$$

on Alice'il ja Bobil nüüd rühmas $\mathbb{Z}_p^*$ üks ühine element, mis pole ise kordagi üle võtmevahetuskanali liikunud; sellest saabki nende jagatud salajane võti.

Küll aga on avalikul kanalil ründajale näha väärtused g, g^a ja g^b. Kogu võtmevahetusprotokolli turvalisus sõltub sellest, kui lihtne või keeruline on neist tuletada salajane g^{ab}. Seda probleemi tuntakse *Diffie-Hellmani ülesande* nime all. Parim teadaolev üldine meetod tema lahendamiseks on leida kõigepealt väärtuste g ja g^a järgi astendaja a (sisuliselt siis logaritm) ning seejärel arvutada $(g^b)^a$. Diskreetsetes rühmades logaritmi leidmise ülesannet tuntaksegi *diskreetse logaritmi leidmise ülesandena*.

Astendamisoperatsioon on jäägiklassikorpuses teostatav võrdlemisi efektiivselt (selleks kulub ülimalt $2\log_2(p)$ korrutamist), aga kuidas on lugu pöördtehte, logaritmimisega? Meie kogemus reaalarvudega viitaks justkui sellele, et reaksarenduse abil ei saa ka logaritm ülearu keerukas olla. Siinkohal tuleb aga mängu jäägiklassistruktuuri diskreetne loomus. $\mathbb{Z}_p$ pole pidev ega isegi mitte tihe ja tänu sellele ei saa koonduvusest rääkida ning reaksarendused ei tööta.

Loomulikult on matemaatikud välja arendanud teisi lähenemisi. Üldistes rühmades saab diskreetse logaritmi arvutamiseks kasutada näiteks Shanksi *baby-step-giant-step* algoritmi ja Pollardi ρ-meetodit. Mõlema oodatav tööaeg n-elemendilise rühma korral on suurusjärku $\sqrt{n}$ rühmatehet [9].

Praktikas kasutatakse Diffie-Hellmani võtmevahetuse jaoks algarvu p väärtusi, mis on paar tuhat bitti pikad, st näiteks $p \approx 2^{2048}$ või $p \approx 2^{3072}$. Kuivõrd

$$\sqrt{2^{2048}} = 2^{1024}\,,$$

jäävad üldotsarbelised diskreetse logaritmi arvutamise algoritmid niisuguste väärtuste puhul jänni (mida me ju tegelikult taotlemegi).

Samas on jäägiklassikorpustes palju muud teadaolevat struktuuri kui abstraktse multiplikatiivse rühma oma. Neid struktuurseid iseärasusi saab ära kasutada efektiivsemate algoritmide loomisel.

Kõige paremaid tulemusi on andnud erinevad sõelumistehnikad (*sieving*), mille abil saab diskreetse logaritmi arvutamise (heuristilise) ajalise keerukuse tuua alla kuni suuruseni

$$L_p\left[\frac{1}{3}, \left(\frac{64}{9}\right)^{1/3}\right], \quad \text{kus} \quad L_p[\alpha, c] = e^{c(\ln(p))^{\alpha}(\ln(\ln(p))^{1-\alpha}} \tag{2}$$

(vt nt [12]). Kui $p \approx 2^{2048}$, siis

$$L_p\left[\frac{1}{3}, \left(\frac{64}{9}\right)^{1/3}\right] \approx 2^{116{,}9}\,,$$

mis on oluliselt vähem kui 2^{1024}, mille andsid üldised diskreetse logaritmi leidmise meetodid, kuid tänaste arvutite jõudluspiiridest jääb see suurusjärk siiski veel väljapoole.

Veelgi olulisem kui erinevus diskreetse logaritmi arvutamise konkreetses ressursivajaduses on nende meetodite *asümptootiline keerukus*. Vastavalt keerukusteoorias kasutusele võetud tavale väljendatakse algoritmi täitmiseks vajalikku aega (*ajalist keerukust*) funktsioonina sisendi pikkusest. Seejuures peetakse funktsiooni kasvumustrit pikas perspektiivis (st sisendite piiramatu suurenemise korral) olulisemaks võimalikest konstantsetest kordajatest funktsiooni avaldise ees. Lugeja võib ise veenduda, et iga (positiivne) ruutfunktsioon kasvab ükskord mööda lineaarfunktsioonist,

kuupfunktsioon omakorda ruutfunktsioonist, 1-st suurema alusega eksponentfunktsioon igast polünoomist jne.

Seetõttu eelistatakse võimalusel näiteks polünomiaalse ajahinnanguga algoritme eksponentsiaalsetele.

Keerukuse õigeks hindamiseks tuleb tähelepanelik olla ka sisendi pikkuse määramisel. Rühmas $\mathbb{Z}_p^*$ on $p-1$ elementi, kuid neid elemente esitatakse reeglina mingi positisoonilise arvusüsteemi abil, mistõttu elemendi pikkus on rühma võimsuse suhtes *logaritmiline*. Logaritmi alus sõltub valitud arvusüsteemi alusest. Samas erinevad erinevatel alustel võetud logaritmid üksteisest konstantse kordaja võrra, mistõttu algoritmi asümptootilise keerukusklassi määramisel pole logaritmi aluse valik eriti oluline. Kasutame siis järgnevas näiteks naturaallogaritmi.

Tuletame meelde, et üldistes rühmades rakendatavad diskreetse logaritmi arvutamise algoritmid vajavad töötamiseks umbes $\sqrt{n}$ rühmatehet, kus n on rühma järk. Rühma $\mathbb{Z}_p^*$ korral saame

$$\sqrt{|\mathbb{Z}_p^*|} = \sqrt{p-1} \approx \sqrt{p} = \sqrt{e^{\ln(p)}} = e^{\frac{1}{2}\cdot\ln(p)} ,$$

mis on eksponentsiaalne sisendi pikkuse $\ln(p)$ suhtes.[2]

Vaatleme nüüd valemis (2) toodud avaldist. Kui $\alpha = 1$, siis

$$L_p[\alpha, c] = e^{c\cdot\ln(p)} ,$$

mis on samuti eksponentsiaalne $\ln(p)$ suhtes. Kui aga $\alpha = 0$, saame

$$L_p[\alpha, c] = e^{c\cdot\ln(\ln(p)} = \ln(p)^c ,$$

mis on $\ln(p)$ suhtes polünomiaalne.

Muutes parameetri α suurust pidevalt 0-st 1-ni, liigume polünomiaalsest hinnangust järjest eksponentsiaalsema poole. Kuna sõelumistehnikate puhul on suudetud saavutada algoritme parameetriga $\alpha = \frac{1}{3}$, on tegemist keerukusfunktsioonidega, mis

[2]Kuna iga rühmatehe ise nõuab ka arvutusaega, on niimoodi diskreetse logaritmi arvutamine isegi veidi keerukam, kuid üldist eksponentsiaalset hinnangut see ainult tõstab.

jäävad eksponentsiaalsele märgatavalt alla. See on ka põhjus, miks sõelumistehnikad annavad jäägiklassikorpuse multiplikatiivse rühma korral oluliselt madalama konkreetse keerukuse.

Samas paneme tähele, et näiteks Diffie-Hellmani võtmevahetusprotokoll on defineeritud suvalises rühmas. Seega on õigustatud küsimus, kas protokolli aluseks oleva rühma saaks valida nii, et sõelumistehnikad ei töötaks. Selgub, et saab küll ja kõige levinumaks alternatiiviks jäägiklassistruktuuridele on elliptkõverate rühmad.

3. Elliptkõverad. Elliptkõverate, nagu suurema osa muu matemaatikagi, ajalugu algab juba Vana-Kreekast, kus Apollonios uuris koonuselõikeid ning kirjutas nende kohta 8-osalise uurimuse. Kahe aastatuhande vältel sisaldas see suuremat osa inimkonna teadmusest antud valdkonnas, kuid oli ka küsimusi, millele Apollonios vastust leida ei suutnud. Üks neist oli ellipsi kaare täpse pikkuse probleem. Matemaatiline aparatuur sellele küsimusele vastamiseks arendati välja alles 17. – 19. sajandil.

Selle töö käigus jõuti pika ringiga polünoomvõrrandite

$$y^2 = p(x)\ , \tag{3}$$

kus p on kuuppolünoom, uurimiseni. Läbi mitmete terminoloogiliste uperpallide hakati võrrandiga (3) määratud kõveraid nimetama *elliptkõverateks*. Kõige segadussejavamalt näeme, et ellips ise *ei ole* elliptkõver. Kõigi nende ajalooliste keerdkäikude kirjapanemine viiks meid siinkohal peateemast liigselt kõrvale, kuid huvitatud lugeja leiab hea ülevaate Adrian Rice'i ja Ezra Browni artiklist [10].

Nagu paljud teisedki võimsad matemaatilised tööriistad, kerkivad ka elliptkõverad esile mitmes sõltumatus kohas. Sageli võib neid kohata diofantiliste võrrandite lahendamisel (mis viib meid jälle otsaga Vana-Kreekasse![3]).

[3]Oma peateoses *Arithmetica* kirjeldab Diophantos mitut lahendusvõtet, milles hilisemad uurijad on tundnud ära elliptkõverate meetodi rakendamise, kuigi Diophantosel polnud tollal muidugi vastavat matemaatilist aparatuuri ega terminoloogiat. Huviline lugeja leiab rohkem informatsiooni Isabella

Vaatleme siinkohal lühidalt nn *kongurentsete arvude* ülesannet (pikema käsitluse annab Keith Conrad artiklis [4]).

Definitsioon 1. Positiivset täisarvu n nimetame *kongruentseks*, kui leidub ratsionaalarvuliste küljepikkustega täisnurkne kolmnurk, mille pindala on n.

Pythagorase kolmnurgast küljepikkustega 3, 4 ja 5 näeme, et 6 on kongruentne arv, aga kongruentsed on näiteks ka 5 (tänu kolmnurgale küljepikkustega $\frac{20}{3}$, $\frac{3}{2}$ ja $\frac{41}{6}$) ning 7 (tänu kolmnurgale küljepikkustega $\frac{35}{12}$, $\frac{24}{5}$ ja $\frac{337}{60}$). Teisalt on võimalik tõestada, et näiteks arv 1 ei ole kongruentne [4].

Niisiis otsime ratsionaalseid lahendeid a, b, c võrrandisüsteemile

$$\begin{cases} a^2 + b^2 &= c^2 \\ ab/2 &= n \end{cases}, \tag{4}$$

kus n on fikseeritud täisarv, mille kongruentsust me soovime kindlaks teha. Süsteemi (4) võrrandeid saab kolmemõõtmelises ruumis interpreteerida pindade võrranditena ning nende lõikumisel tekib joon, millel me otsime ratsionaalsete koordinaatidega punkte. Sobiva muutujavahetusega saab selle joone võrrandi teisendada kujule $y^2 = x^3 - n^2x$. Täpsemalt, kehtib järgmine

Teoreem 1. *Positiivse täisarvu n korral eksisteerib üksühene vastavus hulkade*

$$\{(a, b, c) : a^2 + b^2 = c^2, ab/2 = n\} \text{ ja } \{(x, y) : y^2 = x^3 - n^2x, y \neq 0\}$$

vahel. Sobiva vastavuse annavad teineteise pöördkujutused

$$(a, b, c) \mapsto \left(\frac{nb}{c-a}, \frac{2n^2}{c-a}\right), \quad (x, y) \mapsto \left(\frac{x^2 - n^2}{y}, \frac{2nx}{y}, \frac{x^2 + n^2}{y}\right).$$

Bašmakova suurepärasest raamatust [3].

Teoreemi 1 tõestus on lugejale paras näpuharjutus.

Kuidas tulla selle peale, et otsitav kõver võiks olla just kujul $y^2 = x^3 - n^2x$? Ka sellel on oma põhjus. Vaatleme süsteemi (4) abil järgmisi teisendusi:

$$\left(\frac{a+b}{2}\right)^2 = \frac{a^2+2ab+b^2}{4} = \frac{c^2+4n}{4} = \left(\frac{c}{2}\right)^2 + n\,,$$
$$\left(\frac{a-b}{2}\right)^2 = \frac{a^2-2ab+b^2}{4} = \frac{c^2-4n}{4} = \left(\frac{c}{2}\right)^2 - n\,.$$

Näeme, et ratsionaalruudud

$$\left(\frac{a-b}{2}\right)^2,\ \left(\frac{c}{2}\right)^2 \text{ ja } \left(\frac{a+b}{2}\right)^2$$

moodustavad aritmeetilise jada vahega n. Tähistades $x = \left(\frac{c}{2}\right)^2$, peab mingi ratsionaalarvu y ruut olema ka nende korrutis

$$(x-n)x(x+n) = x^3 - n^2x.$$

Paneme tähele, et teoreemis 1 antud vastavus säilitab lahendite ratsionaalsuse. Seega piisab kongruentseid arve andvate kolmnurkade leidmiseks otsida ratsionaalseid punkte elliptkõveratel kujuga $y^2 = x^3 - n^2x$. Selgub aga, et sellise kõvera abil saab teha veel midagi täiesti maagilist – moodustada kahest teadaolevast sama pindalaga täisnurksest kolmnurgast kolmanda. Uurime vastavat konstruktsiooni lähemalt.

Nagu teada, esituvad kõik primitiivsed Pythagorase kolmikud kujul $(k^2 - l^2, 2kl, k^2 + l^2)$, kus $k > l > 0$, $\text{SÜT}(k,l) = 1$ ning k ja l on erineva paarsusega. Katsetades väikeste k ja l väärtustega leiame, et $(k,l) = (6,1)$ ja $(k,l) = (5,2)$ annavad meile vastavalt Pythagorase kolmikud $(35,12,37)$ ja $(21,20,29)$ ning mõlema moodustuva kolmnurga pindala on 210. Teoreemi 1 abil saame kõveral $y^2 = x^3 - 210^2x$ kaks neile vastavat ratsionaalset punkti $(1260, 44100)$ ja $(525, 11025)$.

Neid punkte läbib sirge võrrandiga $y = 45x - 12600$. Osutub, et see sirge lõikab vaadeldavat elliptkõverat veel kolmandaski

ratsionaalses punktis. Asendades y sirge võrrandist elliptkõvera võrrandisse saame

$$\begin{aligned}(45x - 12600)^2 &= x^3 - 210^2 x\,,\\ x^3 - 2025x^2 \pm \ldots &= 0\,.\end{aligned}$$

Viète'i valemitest teame, et viimase võrrandi lahendite summa on 2025, seega kolmas lahend on

$$x_3 = 2025 - 1260 - 525 = 240$$

ja vastav y-koordinaat on

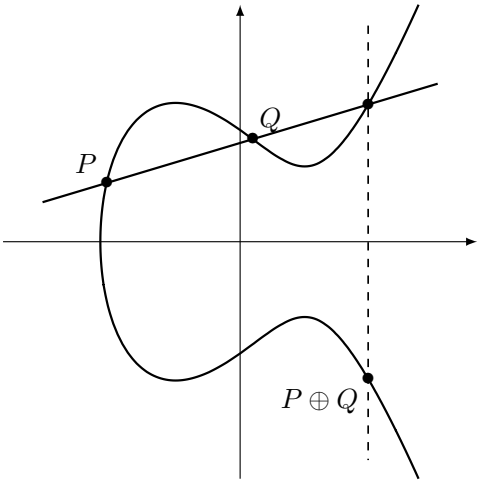
$$y_3 = 45 \cdot 240 - 12600 = -1800\,.$$

Punkt $(240, -1800)$ annab teoreemi 1 abil negatiivsed a, b ja c väärtused, mis ei sobi kolmnurga külgedeks. Paneme aga tähele, et leitud punktiga x-telje suhtes sümmeetriline punkt $(240, 1800)$ asub samuti samal elliptkõveral võrrandiga $y^2 = x^3 - 210^2 x$. See punkt annab teoreemi 1 abil aga uue kolmnurga $(\frac{15}{2}, 56, \frac{113}{2})$. Lugeja saab lihtsasti kontrollida, et tegemist on täisnurkse kolmnurgaga, mille pindala on 210 ja mis pole kongruentne kummagagi algsetest kolmnurkadest.

Ülalkirjeldatud võte, mille puhul läbi elliptkõvera kahe punkti tõmmatakse sirge ning peegeldatakse tema kolmandat lõikepunkti kõveraga x-telje suhtes, on elliptkõverate teoorias fundamentaalse tähtsusega. Sisuliselt defineerime me niimoodi binaarse tehte, mille abil etteantud kõvera kahest punktist P ja Q moodustatakse kolmas (tähistame teda $P \oplus Q$). Seda tehet illustreerib joonis 1.1.

Mis juhtub siis, kui $P = Q$? Sel juhul vaatleme lõikaja asemel lihtsalt kõvera puutujat punktis P ning leiame puutuja ning kõvera teise lõikepunkti peegelduse x-teljest (vt joonist 1.2).[4]

[4]Elliptkõvera lõikaja võimaldas meil kongruentsete arvude ülesannet uurides moodustada kahe teadaoleva ratsionaalse pindalaga kolmnurga järgi kolmanda. Jooniselt 1.2 näeme, et tegelikult piisab uue punkti/kolmnurga leidmiseks ka ühest teadaolevast. Huvitatud lugeja võib võtta näiteks $(3, 4, 5)$-kolmnurga

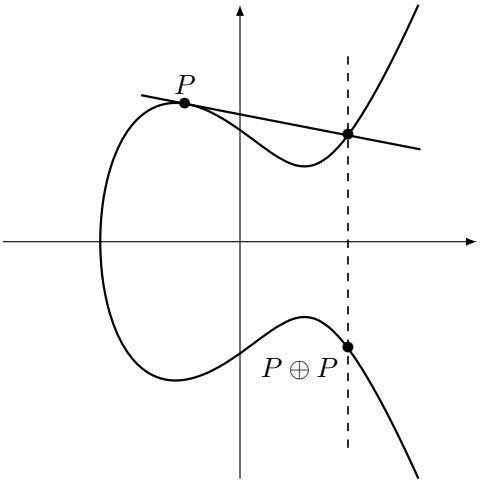


Joonis 1.1: Lõikaja abil määratud tehe elliptkõvera punktidel

Tehte $\oplus$ kommutatiivsus on ilmne, aga üllataval kombel osutub ta ka assotsiatiivseks. Assotsiatiivsuse tõestus ise on tehniline ja keerukas, mistõttu me teda siinkohal ei esita. Huvitatud lugeja leiab elementaarmatemaatika vahenditega esitatava tõestuse artiklist [6] ning rohkem tehte $\oplus$ sisulist tausta avava, aga see-eest ka pikemat süvenemist nõudva tõestuse Lawrence Washingtoni klassikalisest monograafiast [13].

Kas tehte $\oplus$ suhtes leidub ka neutraalne element? Osutub et vaikimisi mitte, kuid selle saab lisada. Võtame kõveral mingi punkti P ning mõtleme, kus peaks asuma punkt $\mathcal{O}$ nii, et $P \oplus \mathcal{O} = P$? Vastavalt tehte $\oplus$ ülaltoodud definitsioonile peaks punkte P ja $\mathcal{O}$ ühendav sirge lõikama kõverat punktiga P x-telje suhtes sümmeetrilises punktis, st see sirge peaks olema vertikaalne. Kuna elliptkõvera vertikaalsel lõikajal pole peale P ja temaga sümmeetrilise

pindalaga 6, leida teoreemi 1 abil talle vastava punkti, arvutada sellest punktist kõverale $y^2 = x^3 - 36x$ tõmmatud puutuja, leida selle puutuja teise ühise punkti kõveraga ning arvutada teoreemi 1 saadud punktile vastava uue kolmnurga külgede pikkused.



Joonis 1.2: Puutuja abil määratud tehe elliptkõvera punktidel

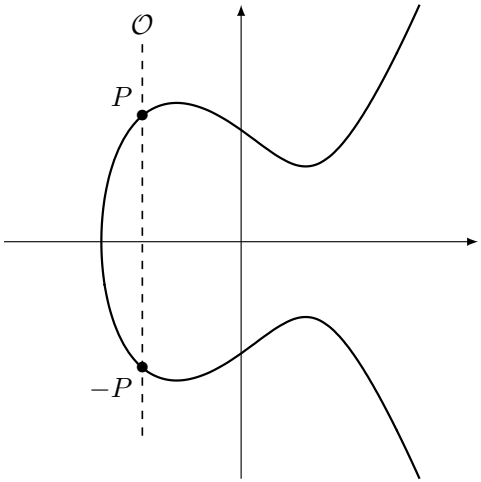
punkti kõveraga kolmandat lõikepunkti, defineeritakse sellesse rolli uus formaalne *lõpmatuspunkt*, mis käitub nii, nagu ta asuks vertikaalsihis lõpmata kaugel. Nüüd on lihtne näha, et punkti P vastandelemendi rolli tehte $\oplus$ suhtes täidab temaga x-telje suhtes sümmeetriline punkt. Neid kontseptsioone illustreerib joonis 1.3.

Kokkuvõttes võib öelda, et kehtib

Teoreem 2. *Elliptkõvera punktide hulk on tehte $\oplus$ suhtes kommutatiivne rühm neutraalse elemendiga lõpmatuspunktis.*

Tehet $\oplus$ nimetatakse elliptkõvera punktide *liitmiseks* ning tähistatakse sageli ka lihtsalt $+$.

4. Elliptkõverate krüptograafia. Tuleme nüüd tagasi oma eesmärgi juurde leida efektiivsemaid rühmi krüptograafiliste operatsioonide jaoks. Meenutame, et rühma kasutatavuseks Diffie-Hellmani võtmevahetuse alusena peab diskreetse logaritmi ülesanne temas raske olema. Milline üldse näeb välja diskreetse logaritmi ülesanne elliptkõvera punktide rühmas?



Joonis 1.3: Elliptkõvera punktide neutraalne element ja vastandelemendid

Klassikalise diskreetse logaritmi ülesande sõnastasime rühmas $\mathbb{Z}_n^*$, mille operatsioone tähistatakse traditsiooni kohaselt multiplikatiivses notatsioonis. Elliptkõvera puntkide puhul räägime aga liitmisest, st me töötame aditiivse notatsiooniga. Sellest johtuvalt pole ka diskreetse logaritmi leidmine sel juhul mitte astendaja, vaid skalaarse kordaja leidmise ülesanne.

Tähistame positiivse täisarvu k ja elliptkõvera punkti P korral

$$[k]P = \underbrace{P \oplus P \oplus \ldots \oplus P}_{k} .$$

Tähistust $[k]P$ saab loomulikul moel üldistada ka mittepositiivsetele täisarvudele:

$$[0]P = \mathcal{O},\ [-1]P = -P,\ [-2]P = (-P) \oplus (-P),\ \ldots\ .$$

Nüüd saame sõnastada *diskreetse logaritmi ülesande aditiivse versiooni* elliptkõverate jaoks:

On antud elliptkõvera punktid P ja Q. Leia täisarv k, mille korral kehtib võrdus

$$[k]P = Q\ .$$

Osutub, et kui aluseks olev elliptkõver hästi valida, pole vastava diskreetse logaritmi ülesande lahendamiseks teada efektiivsemaid algoritme kui need, mis töötavad üldistes rühmades. Jääb veel lahendada küsimus, milline on "hea" kõver.

Eelpool toodud näited kasutasid elliptkõveraid, mis olid defineeritud üle ratsionaal- või realaarvude. Arvutis pole nende korpuste tehted oma lõpmatuse tõttu kahjuks korrektselt teostatavad.[5] Seetõttu vaadeldakse krüptograafias elliptkõveraid üle lõplike korpuste, enamasti $\mathbb{Z}_p$ või $\mathbb{Z}_{2^m}$.[6]

Ka kõverat määrava võrrandi valikul on omajagu vabadust ning ausalt öeldes vaidlevad ka teadlased ikka veel omavahel, milline valik on parim või mida „parim" elliptkõverate kontekstis üldse tähendab. Siinkohal nendesse vaidlustesse laskumata märgime, et näiteks Eesti ID-kaardil kasutatakse elliptkõverat koodnimega P-384, mis on defineeritud üle korpuse $\mathbb{Z}_p$ (kus algarv $p = 2^{384} - 2^{128} - 2^{96} + 2^{32} - 1$) võrrandiga

$$y^2 = x^3 - 3x + b,$$

[5] „Aga arvutis saab ju reaal- ja ratsionaalarve kasutada!" hüüatab lugeja siinkohal. Tõsi, kuid sellel kasutamisel on piirid. Nii näiteks on suurim reaalarv, mida IEEE 754 standardi topelttäpsusega andmetüüp toetab, umbkaudu 1.7977×10^{308}. Sellest piirist suurmate arvude asemel annab arvuti ületäitumisvea. See omakorda tähendab, et reaalarvude teostus pole meie igapäevatarkvaras rangelt võttes tavaliste tehete suhtes kinnine. Teisalt on elliptkõvera rühmastruktuuri saamiseks alloleva algebralise struktuuri kinnisus vajalik.

[6] Kuna need korpused pole pidevad, tuleb eelpooltoodud geomeetrilise intuitsiooni abil puutujate ja lõikajate keeles defineeritud liitmisoperatsioonid veidi teisiti määrata. Lõikajaga tegelikult suurt probleemi ei olegi – läbi kahe punkti tõmmatud sirge võrrand töötab üle iga korpuse. Puutuja korral tuleb kõigepealt aga kõvera võrrandile formaalne tuletis arvutada ning selle abil „puutujasirge" leida. Tulemusena saadakse punktide koordinaatide kaudu määratud liitmisvalemid, mille huvitatud lugeja leiab näiteks artiklist [6] või [13].

kus

$$b = 27580193559959705877849011840389048093056905856361\\ 56852142870730198868924130986086513626076488374510\\ 7765439761230575.$$

Osutub, et saadav rühm on tsükliline[7] ning temas krüptograafiliste arvutuste tegemiseks on kasulik leppida kokku üks moodustaja. Selle moodustajapunkti G koordinaadid on paika pandud lausa ülemaailmselt ning lugeja leiab nad USA standardiorganisatsiooni NIST poolt välja töötatud standardist FIPS 186-4 [2].

Kui raske diskreetse logaritmi arvutamine saadavas rühmas on? Parimad teadaolevad algoritmid selleks töötavad ajas $\sqrt{n}$, kus n on rühma järk. Osutub, et üle algarvulise korpuse $\mathbb{Z}_p$ defineeritud elliptkõveral on ligikaudu p punkti (selle väite täpse sõnastuse annab Hasse teoreem, vt [13]). Seega saame otsitavaks keerukushinnanguks

$$\sqrt{n} \approx \sqrt{p} \approx \sqrt{2^{384}} = 2^{192}$$

operatsiooni, mis jääb kaugelt väljapoole tänapäeva arvutite jõudluspiire.

Kuidas nendest tükkidest nüüd krüptograafiline protokoll kokku panna? Näiteks alapunktis 2 multiplikatiivses notatsioonis esitatud Diffie-Hellmani võtmevahetuse saab otse aditiivseks "tõlkida". Ühise avaliku lähteväärtusena kasutavad Alice ja Bob kõvera baaspunkti G. Lisaks valivad mõlemad omale salajase täisarvu vahemikust $[1, \mathrm{ord}(G) - 1]$ (kus $\mathrm{ord}(G)$ on baaspunkti ja seega ka kogu rühma järk); olgu need vastavalt a ja b. Alice saadab Bobile punkti $[a]G$ ning Bob Alice'ile punkti $[b]G$. Alice saab nüüd arvutada $a([b]G) = [ab]G$ ja Bob arvutab samamoodi $b(a[G]) =$

[7] „Tohoh, tsükliline?" imestab lugeja. „Tsüklilised rühmad on ju kõige lihtsamat sorti rühmad, kuidas nendes saab üldse mingi arvutus keeruline olla?" Tõsi, kuid siinkohal on trikk selles, et diskreetne logaritm ise on tavaline täisarv ja sellisena rühmast väljapoole jääv matemaatiline objekt, mille leidmine osutub tõepoolest keeruliseks ülesandeks.

$[ba]G$. Kuna

$$[ab]G = [ba]G\,, \tag{5}$$

on nad edukalt ühise saladuse kokku leppinud. Tähelepanelik lugeja märkab kindlasti võrduste (1) ja (5) sarnasust.

Kuidas elliptkõverate krüptograafia ID-kaardi peal töötab? Olemuslikult on kõik väga lihtne. ID-kaart funktsioneerib salajase võtme turvalise hoidlana, kusjuures salajaseks võtmeks on juhuslikult valitud täisarv $k \in [1, \operatorname{ord}(G) - 1]$. Talle vastav avalik võti on $[k]G$. Paneme tähele, et tänu diskreetse logaritmi ülesande keerukusele pole avaliku võtme järgi salajase leidmine praktiliselt võimalik.

Avalike võtmete autentne levitamine on omaette keeruline teema, millesse me siinkohal ei lasku, aga näiteks Eestis on see lahendatud ühe suure, kõigi ID-kaartide (ja ka mobiil-ID kiipide) avalike võtmete taristu abil (inglise keeles kohtame selle jaoks sageli terminit *Public Key Infrastructure* ehk PKI).

Kui Alice tahab saata Bobile krüpteeritud sõnumit, hangib ta kõigepealt Bobi isikukoodi alusel PKI taristu kaudu tema avaliku võtme $[k_b]G$ ning valib ühekordse juhusliku täisarvu j_a. Sõnumi krüpteerimise võtmena kasutab Alice kõverapunkti[8]

$$j_a([k_b]G) = [j_a \cdot k_b]G.$$

Koos krüpteeritud sõnumiga paneb Alice kaasa ka punkti $[j_a]G$. Dekrüpteerides kasutab Bob oma ID-kaardil olevat salajast võtit k_b ja arvutab selle abil $k_b([j_a]G) = [k_b \cdot j_a]G$, mis ongi vajalikuks dekrüpteerimisvõtmeks.

Lugeja ei ole eksinud, kui ta tunneb siinkohal jälle ära Diffie-Hellmani võtmevahetuse. Ainsaks (kuigi mitmes mõttes oluliseks) erinevuseks on, et Bob ei moodusta igaks sõnumivahetuseks uut salajast võtit, vaid kasutab seda, mis tema ID-kaardi peal juba leidub.

[8] Siinkohal on elu tegelikkuses natuke keerulisem. Kõverapunkti koordinaadid ei sobi tavaliselt otse teiste krüptoalgoritmide võtmeteks, seega tuleb neid enne veel natuke töödelda, aga siinkohal palun lugejal mind lihtsalt uskuda, kui ma ütlen, et see on võimalik lihtsate ja standardsete meetoditega.

Salajaste sõnumite vahetamisest palju sagedasem ID-kaardi kasutamise stsenaarium on tegelikult hoopis allkirjastamine. Allkirju moodustatakse elliptkõverate digitaalsignatuurialgoritmi (*Elliptic Curve Digital Signature Algorithm*, ECDSA) abil, kuid selle detailide esitamine nõuaks pikemat süvenemist kui siinkohal mõistlik. Huvitatud lugeja leiab algoritmi täpse kirjelduse standardist SEC1 [1].

5. Mis edasi? Kas elliptkõverate krüptograafia kaitseb meid kõigi infoturbeprobleemide eest nüüd ja igavesti? Kahjuks mitte. Kõik šifrid muutuvad ajas ainult nõrgemateks ja seda mitmel põhjusel. Esiteks areneb arvutustehnika ikka veel hoogsalt ja ühel päeval võib 2^{192} operatsiooni realistlikuks muutuda (kuigi see ei paista lähiajal veel tõenäoline). Teiseks arenevad ka algoritmid ning mõni krüptograaf võib avastada teadaolevatega võrreldes palju efektiivsema meetodi elliptkõveratel diskreetsete logaritmide arvutamiseks.

Hetkel tundub kõige realistlikum oht elliptkõverate krüptograafiale aga hoopis uut tüüpi arvutusmasina, kvantarvuti, valmimine. Sellel saaks käivitada Peter Shori poolt 1994. aastal välja pakutud algoritmi, mis elementaarosakeste kvantefekte ära kasutades suudaks diskreetseid logaritme leida tunduvalt efektiivsemalt [11].

Väikesemõõtmelisi kvantarvuteid on uurimislaborites ehitatud juba kümmekond aastat, kuid siiani pole ükski neist piisavalt võimas, et praktikas kasutatavaid elliptkõverate krüptosüsteeme murda. Millal niisugune valmis saab, on raske öelda, aga ühel hetkel see tõenäoliselt juhtub. Hiljemalt selleks hetkeks peavad meil olema valmis uued krüptograafiastandardid. Töö nende kallal juba käib, aga võidujooksu šifriloojate ja -murdjate vahel ei lõpeta arvatavasti seegi. Nii jätkub krüptograafidele tööd ja leiba arvatavasti veel pikaks ajaks.

Kirjandus

[1] SEC 1: Elliptic Curve Cryptography, 2009. Certicom Research, Standards for Efficient Cryptography, `https://www.secg.org/sec`

1-v2.pdf.

[2] Digital Signature Standard (DSS), 2013. National Institute of Standards and Technology, Federal Information Processing Standard FIPS 186-4, https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.186-4.pdf.

[3] I. G. Bashmakova, *Diohantus and Diophantine equations*. The Dolciani Mathematical Expositions, 20. Mathematical Association of America, Washington, DC, 1997.

[4] K. Conrad, The congruent number problem. *The Harvard College Mathematics Review*, 2 (2008), 58–74. https://kconrad.math.uconn.edu/blurbs/ugradnumthy/congnumber.pdf.

[5] W. Diffie, M. E. Hellman, New directions in cryptography. *IEEE Trans. Information Theory*, 22, No 6 (1976), 644–654.

[6] S. Friedl, An elementary proof of the group law for elliptic curves. *Groups Complex. Gryptol.* 9, No 2 (2017), 117–123.

[7] A. Kerckhoffs, La cryptography militaire. *Le Journal des sciences militaires*, IX (1883), 5–38.

[8] N. Liventhal, *Krüptoloogia ja salaluure*. Perioodika, 1994.

[9] A. J. Menezes, P. C. van Oorschot, S. A. Vanstone, *Handbook of applied cryptography.* CRC Press Series on Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, 1997.

[10] A. Rice, E. Brown, Why ellipses are not elliptic curves. *Math. Mag.* 85, No 3 (2012), 163–176. https://www.maa.org/sites/default/files/pdf/upload_library/2/Rice-2013.pdf.

[11] P. W. Schor, Algorithms for quatum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, Nov. 1994.

[12] E. Thomé, Algorithms for discrete logarithms in finite fields and elliptic curves, 2015. https://www.math.u-bordeaux.fr/~aenge/ecc2015/documents/thome.pdf.

[13] L. C. Washington, *Elliptic curves. Number theory and cryptography.* Second edition. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2008.

