STŘEDOŠKOLSKÁ ODBORNÁ ČINNOST

Obor č. 1: Matematika a statistika

Isogenie v kryptografii

Zdeněk Pezlar
Jihomoravský kraj

Brno 2020

STŘEDOŠKOLSKÁ ODBORNÁ ČINNOST

Obor č. 1: Matematika a statistika

Isogenie v kryptografii

Isogeny Based Cryptography

Autor: Zdeněk Pezlar
Škola: Gymnázium Brno, třída Kapitána Jaroše, p. o.
Kraj: Jihomoravský
Konzultant: Bc. Vojtěch Suchánek

Prohlášení

Prohlašuji, že jsem svou práci SOČ vypracoval samostatně a použil jsem pouze prameny a literaturu uvedené v seznamu bibliografických záznamů. Prohlašuji, že tištěná verze a elektronická verze soutěžní práce SOČ jsou shodné. Nemám závažný důvod proti zpřístupňování této práce v souladu se zákonem č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) v platném znění.

V Brně dne: ..... Podpis: .....

jihomoravský kraj

Poděkování

++Tato práce byla vypracována za finanční podpory JMK.

Abstrakt

abstrakt

Klíčová slova

isogenie;klíčové slovo.

Abstract

abstrakt

Key words

isogenie;klíčové slovo.

Obsah

Úvod

celkem úvod

Použitá značení

$a \mid b$	a dělí b
$\mathcal{D}(a,b)$	největší společný dělitel a, b
$a \sim b$	a je asociované s b
$\overline{a+b\sqrt{m}}$	konjugát $a+b\sqrt{m}$, neboli $a-b\sqrt{m}$
$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$	množina přirozených, celých, racionálních, reálných, komplexních čísel
$\mathbb{Z}_d$	okruh zbytků modulo d
$R[x]$	okruh polynomů s koeficienty nad okruhem R
$K(a_1, \ldots, a_n)$	nejmenší podtěleso L, které obsahuje těleso K i prvky $a_1, \ldots, a_n \in L$
$[K : L]$	stupeň rozšíření tělesa K nad L, t.j. dimenze vektorového prostoru $K : L$
$\mathcal{O}_K$	okruh celých algebraických čísel tělesa K
$Cl(\mathcal{O}_K)$	grupa tříd ideálů tělesa K
h_K	řád grupy tříd ideálů tělesa K
$\mathcal{U}(\mathcal{O}_K)$	grupa jednotek tělesa K
(a)	hlavní ideál generovaný prvkem a
$\frac{\mathcal{I}}{m}$	lomený ideál $\frac{\mathcal{I}}{m}$
$\left(\frac{a}{m}\right)$	hlavní lomený ideál $\frac{(a)}{m}$
$N(a)$	norma prvku a
$N((a))$	norma ideálu generovaného a
$\mathcal{I} \mid \mathcal{J}$	ideál $\mathcal{I}$ dělí ideál $\mathcal{J}$
P_α	minimální polynom α nad K
G/H	faktorgrupa G podle H

Kapitola 1

Eliptické křivky

1.1 Základy

V naší první kapitole se budeme věnovat isogeniím eliptických křivek a práci s nimi. Budeme budovat teorii a intuici potřebnou k smysluplné diskuzi protokolu SIDH. Pro porozumění textu je třeba ovládat základy [čeho zjistím, až to napíšu]. Budeme postupovat vesměs dle ??, nicméně další vhodný úvodní materiál se nachází na ??. Ne vždy budeme uvádět důkazy tvrzení, neboť jsou mnohdy příliš pokročilé či technické, v takových případech se odkážeme na relevantní literaturu.

Po celou dobu budeme pracovat nad projektivním prostorem nad uzávěrem tělesa K, což je množina bodů v $\overline{K}^n$, kde dva body považujeme za ekvivalentní, pokud leží v přímce s počátkem, můžeme proto místo jednotlivých bodů pracovat s přímkami skrz počátek. Chtěli bychom, aby se každé dvě $n-1$ rozměrné roviny protínaly, a s tím máme problém pouze pokud protínáme dvě rovnoběžné. V každém směru si tak můžeme definovat projektivní prostor stupně $n-1$ v nekonečnu, kde se protínají rovnoběžné roviny.

Definice 1.1.1. *Projektivní prostor $\mathbb{P}^n(\overline{K})$ definujeme jako množinu tříd nenulových vektorů $(a_0, \ldots, a_n) \in \overline{K}^{n+1}$ s ekvivalentní relací $(a_0, \ldots, a_n) \sim (b_0, \ldots, b_n)$, pokud existuje nenulové λ, že $(a_0, \ldots, a_n) = \lambda(b_0, \ldots, b_n)$. Tyto třídy ekvivalence budeme značit $(a_0 : \cdots : a_n)$.*

Pokud je jedno z a_i nulové, získáme $n-1$ rozměrný prostor v nekonečnu.

Projektivní prostor $\mathbb{P}^2(\mathbb{R})$ je známý jako projektivní rovina. Každé dvě přímky se protínají v jednom bodě, přičemž rovnoběžné přímky se protínají v bodě v nekonečnu v daném směru.

Poznámka 1.1.2. *Je zajímavé uvážit spojitost projektivních prostorů s barycentrickými souřadnicemi, kde je každý bod vyjádřen jako vážený průměr vrcholů referenčního simplexu. Tyto souřadnice jsou též homogenní a každé dvě přímky se protínají, byť některé v ne-*

konečnu, takové body mají součet vah roven 0. *Můžeme tedy o barycentrických souřadnicích přemýšlet jako o projektivním prostoru s jiným základem.*

Přípomeňme si pak definici eliptické křivky. Často se definuje jako nesingulární projektivní křivka genu 1, pro naše účely si definici zúžíme.

Definice 1.1.3. *Mějme K těleso charakteristiky různé od* 2 *a* 3. *Pro $a, b \in K$, že $4a^2+27b^3 \neq 0$, definujeme v $\mathbb{P}^2(\overline{K})$ eliptickou křivku jako množinu bodů $(X : Y : Z) \in \overline{K}^3$ splňující:*

$$Y^2Z = X^3 + aXZ^2 + bZ^3.$$

Definice vylučující tělesa s charakteristikou 2 a 3 nám umožňuje zapsat křivku ve výše uvedené jednoduché formě. Avšak čtenář, jenž je již obeznámen s eliptickými křivkami, může protestovat, že eliptická křivka je množina bodů $x, y \in K$ splňující:

$$y^2 = x^3 + ax + b.$$

Pokud máme v rovnici eliptické křivky $Z = 0$, pak i $X = 0$ a máme jediný bod $(0 : 1 : 0)$. Jinak můžeme celou rovnici podělit Z^3 a přejít na proměnné $x := \frac{X}{Z}, y := \frac{Y}{Z}$ a získat nám známou formu, kterou budeme dále označovat jako *afinní*, často se v literatuře uvádí jako *Weierstrassova*. Pak naše křivka je množina bodů $(x, y) \in K^2$ splňujících $y^2 = x^3 + ax + b$ spolu s bodem v nekonečnu $\mathcal{O} = (0 : 1 : 0)$.

Definice 1.1.4. *Množinu všech bodů E nad K (společně s $\mathcal{O}$) budeme značit $E(K)$ a počet jejích prvků budeme značit $\#E(K)$.*

Podívejme se nyní na eliptickou křivku E geometricky. Je zjevné, že má graf symetrický podle osy x, definujme proto k $P \in E$ bod $-P \in E$ jako obraz P podle osy x. Pokud bychom na bodech naší křivky definovali součet, chtěli bychom, aby součet P a $-P$ byl $\mathcal{O}$.

OBrázky

Pokud řekneme, že tečna k E ji protíná ve dvou stejných bodech, pak každá přímka protíná E v právě třech bodech včetně multiplicity. Speciálně tečna v bodě s $y = 0$ tento bod protíná dvakrát a ten třetí je bod v nekonečnu E. Pak si sčítání $+$ na E můžeme definovat tak, že součet každých tří bodů v přímce je $\mathcal{O}$. Pokud tak přímka procházející $P, Q \in E$ protíná E potřetí v R, pak definujeme $P+Q = -R$. Pro takto definovaný součet můžeme pro $P, Q \in E$ odvodit několik důležitých vlastností:

(i) $P + Q = Q + P$,

(ii) $(P + Q) + R = P + (Q + R)$,

(iii) $P + \mathcal{O} = P$,

(iv) $P + (-P) = \mathcal{O}$.

Při takto definovaném součtu můžeme s body na E pracovat jako s abelovskou grupou se sčítáním $+$ a neutrálním prvkem $\mathcal{O}$. Samozřejmě součet dvou bodů dokážeme za pomocí analytické geometrie přímo spočíst:

Věta 1.1.5. *Mějme afinní body $P = (x_1, y_1), Q = (x_2, y_2)$ s $P \neq -P$. Pak $P+Q = (x_3, y_3)$ je daný:*

$$x_3 = \lambda^2 - x_1 - x_2,$$
$$y_3 = -\lambda x_3 - y_1 + \lambda x_1,$$

kde:

$$\lambda = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1}, & \text{pokud } x_1 \neq x - 2, \\ \frac{3x_1^2 + a}{2y_1}, & \text{pokud } x_1 = x - 2. \end{cases}$$

Důkaz s prominutím neuvádím. Pro zkrácení zápisu si budeme definovat skálární násobky bodů následovně:

Definice 1.1.6. *Mějme bod $P \in E$. Pak pro n přirozené definujeme jeho n-násobek:*

$$[n]_E P = \underbrace{P + \cdots + P}_{n},$$

přičemž pro $n < 0$ definujeme $[n]_E P = [-n]_E(-P)$ a $[0]_E P = \mathcal{O}$.

Pokud bude z kontextu jasné, nad kterou eliptickou křivkou pracujeme, budeme značit násobení skalárem pouze $[n]P$.

Příklad 1.1.7.

tu příklad, jak v Q tak v Fq, hezky graficky

Všimneme si, že pro P s $y = 0$ je $[2]P = \mathcal{O}$. Na příkladu (1.1.7) též ale vidíme, že trojnásobek bodu ??? dává též $\mathcal{O}$. Obecně by nás mohlo zajímat, které body pošle násobení n do nekonečna.

Definice 1.1.8. *Buď n celé číslo. O množině všech $P \in E$, že $[n]P = \mathcal{O}$, řekneme, že tvoří n-torzi E a tuto množinu budeme značit $E[n]$.*

Torzní podgrupy nám pomáhají hlouběji studovat eliptické křivky v mnohých směrech. Například pro eliptickou křivku E nad konečným tělesem $\mathbb{F}_q$, je konečná grupa $E(\mathbb{F}_q)$???????????

Věta 1.1.9. *Nechť je E eliptická křivka nad K a m nenulové číslo. Pak:*

- *Pokud* $\operatorname{char} K \nmid m$, *tak* $E[m] \cong \mathbb{Z}_m \times \mathbb{Z}_m$,

- *Pokud označíme* $c = \operatorname{char} K > 0$*:*

$$E[c^i] \cong \begin{cases} \{\mathcal{O}\}, & \text{pro každé nezáporné } i, \\ \mathbb{Z}_{c^i}, & \text{pro každé nezáporné } i. \end{cases}$$

Navíc pokud $m \neq \operatorname{char} K$ je prvočíslo, tak . Nicméně vidíme, že existuje rodina křivek, která má pouze triviální c-torzi.

Definice 1.1.10. *Pokud máme* $E[c] \cong \{\mathcal{O}\}$*, nazveme* E *supersingulární. Jinak* E *budeme říkat obyčejná.*

Rozdělení křivek na obyčejné a supersingulární bude vhodné v mnoha ohledech, jak při diskuzi vlastností křivek, tak z kryptografického hlediska.

1.2 Zobrazení mezi eliptickými křivkami

Násobení bodů v E skalárem nám dává homomorfismus $E \longrightarrow E$. Tvoří proto endomorfismus z E daný lomenou funkcí nad K. My se nyní podíváme na zobrazení mezi jednotlivými eliptickými křivkami, konkrétně homomorfismy grup $E_1(K) \longrightarrow E_2(K)$.

Uvažme zobrazení $(x, y) \mapsto (u^2x, u^3y)$, které převádí křivky:

$$E_1 : y^2 = x^3 + u^4ax + u^6b \mapsto E_2 : y^2 = x^3 + ax + b$$

pro libovolné $u \in \overline{K}$. To je lineární zobrazení mezi $E_1(K)$ a $E_2(K)$, které zachovává přímky a tedy i součet bodů na našich křivkách, definuje proto homomorfismus z $E_1(K)$ do $E_2(K)$. Navíc je zobrazení zjevně invertibilní, tudíž dokonce mezi $E_1(K)$ a $E_2(K)$ dává isomorfismus nad $\overline{K}$.

Věta 1.2.1. *(Sato-Tate) Dvě křivky* E_1, E_2 *nad* K *jsou nad* K *isomorfní právě pokud* $\#E_1(K) = \#E_2(K)$*.*

Speciálně dvě křivky, které mají nad K pouze body v nekonečnu, jsou nad K isomorfní. Ne vždy máme nutně isomorfismus nad K, ale nad jeho rozšířením. Aby byl náš isomorfismus nad $\overline{K}$ definovaný, musí být díky předpisu $(x, y) \mapsto (u^2x, u^3y)$ nutně nad rozšířením K stupně dělícího 6.

Definice 1.2.2. *Buďte* E, E' *křivky isomorfní nad rozšířením* K*, ale ne nad* K*. Pak řekneme, že* E' *je twistem* E *nad* K*.*

Zobrazení $(x, y) \mapsto \left(\frac{x}{d}, \frac{y}{\sqrt{d^3}}\right)$ pro $\sqrt{d} \notin K, d \in K$ nám dává isomorfismus z $E : y^2 = x^3 + ax + b$ na:

$$E : y^2 = x^3 + ax + b \simeq E_d : y^2 = x^3 + d^2ax + d^3b \Leftrightarrow dy^2 = x^3 + ax + b,$$

avšak ne nad K, ale nad jeho kvadratickým rozšířením $K(\sqrt{d})$. E_d nazveme *kvadratickým twistem* E.

Pro křivky s $a = 0$, resp. $b = 0$, můžeme analogicky najít *kubický* a *sextický*, resp. *kvartický* twist:

$$y^2 = x^3 + b \mapsto y^2 = x^3 + d^2b,$$
$$y^2 = x^3 + b \mapsto y^2 = x^3 + db,$$
$$y^2 = x^3 + ax \mapsto y^2 = x^3 + dax,$$

dané po řadě $(x, y) \mapsto \left(\frac{x}{\sqrt[3]{d^2}}, \frac{y}{d}\right)$ a $(x, y) \mapsto \left(\frac{x}{\sqrt[3]{d}}, \frac{y}{\sqrt{d}}\right)$, resp. $(x, y) \mapsto \left(\frac{x}{\sqrt{d}}, \frac{y}{\sqrt[4]{d^3}}\right)$. Vidíme, že poslední dvě zmíněné křivky jsou navíc kvadratickými twisty po řadě kubického a kvadratického twistu E.

Chtěli bychom říci, kdy mezi dvěma eliptickými křivkami existuje isomorfismus, tedy najít nějaký invariant, který isomorfní křivky sdílí. Takovou funkci splňuje právě j-invariant.

Definice 1.2.3. *Pro eliptickou křivku $E : y^2 = x^3 + ax + b$ definujeme její j-invariant jako:*

$$j(E) = 1728\frac{4a^3}{4a^3 + 27b^2}.$$

Poznamenejme, že ten je vždy nad K definovaný, neboť eliptické křivky mají nenulový diskriminant.

Věta 1.2.4. *Dvě křivky definované nad K jsou isomorfní nad $\overline{K}$ právě pokud mají stejný j-invariant.*

Mějme následujících pět křivek nad $\mathbb{Z}_{101}$:

$$E_1 : y^2 = x^3 + x + 1,$$
$$E_2 : y^2 = x^3 + 5x + 23,$$
$$E_3 : y^2 = x^3 + x - 1,$$
$$E_4 : y^2 = x^3 + 2,$$
$$E_5 : y^2 = x^3 + 2x.$$

Spočtěme si pak jejich j-invarianty nad $\mathbb{Z}_{101}$:

$$j(E_1) = 1728\frac{4}{31},$$
$$j(E_2) = 1728\frac{4 \cdot 5^3}{4 \cdot 5^3 + 27 \cdot 23^2} = 1728\frac{4 \cdot 24}{4 \cdot 24 + 27 \cdot 24} = 1728\frac{4}{31},$$
$$j(E_3) = 1728\frac{4}{31},$$
$$j(E_4) = 1728,$$
$$j(E_5) = 0.$$

Vidíme, že j-invarianty E_1 a E_2 se rovnají, přičemž v $\mathbb{Z}_{101}$ se oba rovnají $1728 \cdot 4 \cdot 88$, nutně mezi nimi existuje isomorfismus. Snadno ověříme, že zobrazení:

$$(x, y) \mapsto (3^2 x, 3^3 y) = (9x, 27y)$$

převádí:

$$\begin{aligned} y^2 = x^3 + x + 1 &\mapsto 27^2 y^2 = 9^3 x^3 + 9x + 1, \\ &\Leftrightarrow 22y^2 = 22x^3 + 9x + 1, \\ &\Leftrightarrow 22y^2 = 22x^3 + 110x + 506, \\ &\Leftrightarrow y^2 = x^3 + 5x + 23. \end{aligned}$$

Inverzní isomorfismus $E_2 \longrightarrow E_1$ je pak daný $(x, y) \mapsto (34^2 x, 34^3 y) = (45x, 15y)$, neboť multiplikativní inverz 3 v $\mathbb{Z}_{101}$ je 34.

Křivka E_3 má ale též stejný j-invariant jako E_1 a E_2, nad $\mathbb{Z}_{101}$ mezi nimi a E_3 přesto isomorfismus neexistuje. E_3 je kvadratickým twistem E_1 nad $\mathbb{Z}_{101^2} = \mathbb{Z}_{101}[i]$, jakožto zobrazení $(x, y) \mapsto \left(\frac{x}{i^2}, \frac{y}{i^3}\right) = (-x, iy)$ převádí:

$$\begin{aligned} y^2 = x^3 + x + 1 &\mapsto -y^2 = -x^3 - x + 1, \\ \Leftrightarrow y^2 = x^3 + x - 1. & \end{aligned}$$

Dvě speciální hodnoty j-invariantu jsou 0 a 1728, kterých nabývají křivky, které mají po řadě lineární, resp. konstantní člen roven 0. Právě křivky s j-invariantem 0 mají kubický (a sextický) twist a ty s j-invariantem 1728 zase kvartický.

Mohli bychom se nicméně zajímat, proč se v j-invariantu násobí číslem 1728. Důvodem jsou tělesa charakteristik 2 a 3, j-invariant je totiž definován pro libovolnou nesingulární projektivní křivku genu 1, tj.:

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6,$$

konkrétně jako:

$$\frac{(b_2^2 - 24b_4)^2}{\Delta},$$

kde Δ je diskriminant naší křivky a $b_2 = a_1^2 + 4a_2, b_4 = 2a_4 + a_1 a_3$. Pro tělesa s $\operatorname{char} K \neq 2, 3$ můžeme definovat zobrazení, která převádí naši křivku na nám známý afinní tvar, detaily důkazu jsou k nalezení v [5, kap. 3]. Obraz rovnice j-invariantu je právě takový, jak ho zde definujeme, násoben konstantou 1728.

Jak násobení bodů E skalárem, tak twistování, jsou homomorfismy bodů křivek nad tělesem K, resp. jeho rozšířením. Spadají tak pod rodinu zobrazení eliptických křivek zvaných *isogenie*, o kterých se budeme dále bavit.

1.3 Isogenie

Definice 1.3.1. *Ať $E_1, E_2 \in \overline{K}$ jsou eliptické křivky. Pak surjektivní morfismus $E_1 \longrightarrow E_2$ daný racionální funkcí nad K, který posílá bod v nekonečnu E_1 na bod v nekonečnu E_2, nazveme isogenií. Pokud mezi E_1, E_2 existuje isogenie, nazveme je isogenní.*

Definice 1.3.2. *Pod stupňem isogenie ϕ budeme rozumět jejímu stupni jako racionální funkci a budeme značit* $\deg \phi$. *Obraz křivky E v ϕ budeme značit $\phi(E)$.*

Všimneme si, že jak násobení bodů skalárem, tak naše isomorfismy zmíněné na konci předchozí kapitoly jsou isogenie. Zobrazení:

$$y^2 = x^3 + ax + b \mapsto ?$$

dané $(x, y) \mapsto (,)$ je isogenií též. Pojďme se nyní pobavit o několika základních vlastností isogenií.

teext

Věta 1.3.3. *Buď $\phi : E \longrightarrow E'$ isogenie stupně n. Pak existuje jediná isogenie $\hat{\phi} : E' \longrightarrow E$, která pro každou jinou isogenii $\psi : E' \mapsto E, \xi :$ splňuje:*

(i) $\phi \circ \hat{\phi} = [n]_{E'}$,

(ii) $\hat{\phi} \circ \phi = [n]_E$,

(iii) $\widehat{\phi \circ \psi} = \hat{\psi} \circ \hat{\phi}$,

(iv) $\widehat{\phi + \psi} = \hat{\phi} + \hat{\psi}$,

(v) $\hat{\hat{\phi}} = \phi$.

Isogenii $\hat{\phi}$ budeme označnovat jako isogenii duální k ϕ.

Stejně jako jsme se zabývali torsní podgrupou našich křivek, nebude překvapením, že bude pro studium isogenií důležité, které body zobrazí do nekonečna.

Definice 1.3.4. *Jádrem naší isogenie ϕ, značíme* $\ker \phi$, *rozumíme jejímu jádru jakožto homomorfismu grup, počet jeho prvků značme* $\# \ker \phi$.

Definice 1.3.5. *Mějme $E, E_1 \in \overline{K}$ a $\phi : E \longrightarrow E_1$ isogenii stupně k. Pokud je $\# \ker \phi = k$, pak o ϕ řekneme, že je separabilní. Jinak řekneme, že ϕ je neseparabilní. V případě, že je* $\deg \phi$ *roven mocnině* $\operatorname{char} K$, *mluvíme o ϕ jako o čistě neseparabilní.*

Jak tomu bylo v případě násobení skalárem, $\ker \phi$ tvoří podgrupu $E(K)$. Ukáže se, že separabilní isogenie E se dělí na třídy isomorfismů jednoznačně určené svým jádrem.

Věta 1.3.6. *Každá separabilní isogenie ϕ z E je, až na isomorfismus, jednoznačně určena svým jádrem.*

Pokud je tak $G = \ker \phi$ grupa tvořená jádrem ϕ, můžeme značit E/G cílovou křivku ϕ. Separabilní isogenie z $E \longrightarrow E_1$ je daná lomenou funkcí nad K a známe-li její jádro, dokážeme ji explicitně spočíst, přičemž libovolná podgrupa $E(K)$ je jádrem isogenie. Vzorce udávající přesný tvar separabilní isogenie z $E \longrightarrow E'$ s daným jádrem se nazývají *Véluovy* po Jeanu Véluovy, který je první publikoval (?).

Věta 1.3.7. *Vélu Formulas.*

Věta 1.3.8. *Buďte ϕ, ψ dvě isogenie z E. Pak:*

$$\phi(\psi(E)) \cong \psi(\phi(E)).$$

Pro separabilní isogenie tak můžeme říci:

$$(E/\langle A\rangle)/\langle B\rangle \cong (E/\langle B\rangle)/\langle A\rangle.$$

Kapitola 2

Užití v kryptografii

Přes Caesarovu šifru až po šifrování za pomocí Enigmy v období druhé světové války, po většinu lidské historie se využívaly kryptografické systémy založené na faktu, že obě komunikující partie si po domluvě vyberou způsob maskování zprávy a ten pro ostatní zůstává skrytý. Příkladem je právě o kolik písmen v Caesarově šifře transponujeme. Tento způsob nutně závisí na faktu, že se obě strany před výměnou mají možnost přes bezpečný kanál na tomto způsobu domluvit. S přibývajícím počtem účastníků a frekvencí komunikace, na příklad našeho každodenního interagování na internetu, je bohužel třeba vyšší počet klíčů a přibývá risk kompromitace.

Kvůli takovým obavám přišli Whitfield Diffie a Martin Hellman v roce 1976 s revolučním nápadem: asymetrickou kryptografií, kde každý z účastníků má svůj vlastní *privátní klíč*, který s nikým nesdílí. Všechny strany, i potenciální útočník, znají několik informací, které jsou známé jako *veřejné parametry*. Obě komunikující strany za pomocí veřejných informací tajně transformují svůj privátní klíč a výsledek, který budeme nazývat *veřejným klíčem*, publikují. Oba účastnící vezmou veřejný klíč toho druhého a provedou s ním ty samé tajné kroky závisící na jejich privátním klíči. Podstatou takové výměny je, že na jejím konci získají obě původní strany netriviální informaci, tedy informaci takovou, že žádná třetí strana ji nedokáže snadno uhodnout, za pomocí níž poté mohou společnou komunikaci šifrovat a nikdo jiný již jejich zprávy neuvidí. Předpokládá se, že pouze ze znalosti veřejného klíče je pro každou další partii těžké replikovat klíč privátní a že pole možných sdílených informací je obrovské. Vyhneme se tak přímočarým řešením hrubou silou.

Pojďme se podívat na protokol, který Diffie a Hellman navrhli. Budeme o něm dále mluvit jako o *Diffie-Hellmanově výměně*. Je založena na problému *diskrétního logaritmu* prvku $a \in \mathbb{Z}_p^*$. Tento problém po nás ze znalosti primitivního kořene g modulo p žádá najít k, že $g^k = a$ v $\mathbb{Z}_p$. Obecně můžeme $\mathbb{Z}_p$ nahradit cyklickou grupou G a mít g její generátor. Protokol požaduje, aby nebyl diskrétní logaritmus spočitatelný efektivně, tj. v polynomiálním čase vzhledem k velikosti grupy, jinak může útočník jednoduše privátní klíče obou stran spočíst, ale mocnění bylo . Umocnit číslo dokážeme v logaritmickém čase, a v konečné grupě nám stačí umocnit pouze na exponent modulo řádu grupy.

Veřejné parametry: Grupa G řádu p, kde p je prvočíslo, s generátorem g.

Alfréd		Blažena
Vstup: $a \in G^*$		**Vstup:** $b \in G^*$
	$\xrightarrow{g^a}$	
	$\xleftarrow{g^b}$	
$G_{AB} := \left(g^b\right)^a = g^{ab}$		$G_{BA} := \left(g^a\right)^b = g^{ba}$
Výstup: G_{AB}		**Výstup:** G_{BA}

Algoritmus 1: Diffie-Hellmanova výměna

Díky předpokladu, že G je cyklická, je i abelovská, tedy $G_{AB} = g^{ab} = g^{ba} = G_{BA}$. Na konci protokolu tak mají obě strany shodné tajemství g^{ab}.

Řád G se prakticky bere prvočíslo $q = 2p + 1$, že p je prvočíslo, p nazveme tzv. *Sophie-Germainovým prvočíslem* a q zase *bezpečným prvočíslem*. V takovém případě má G podgrupu prvočíselného řádu p, což je z kryptografického hlediska žádané, je tuto grupu totiž obtížnější spočíst. Navíc bezpečná prvočísla skýtají i výhody pro inicializování výměny, pro taková prvočísla dokážeme totiž snadno nalézt primitivní kořen. Konkrétně, g je primitivní kořen modulo $2p + 1$, tedy má řád $q - 1 = 2p$ modulo q, právě pokud $g^p \equiv -1 \pmod{q}$. Stačí nám pak najít $g^p \pmod{q}$, což nám mohou usnadnit nástroje jako Eulerovo kritérium, díky kterému je postačující mít g kvadratický nezbytek modulo q.

Veřejné klíče g^a, g^b, jsou nicméně, jak jejich název napovídá, veřejné, a má k nim přístup libovolná jiná osoba. Dejme tomu, že Eva, která má přístup pouze k veřejně dostupným informacím G, g, g^a, g^b, by chtěla též znát sdílené tajemství. Jeden způsob, jak by mohla tajnou informací získat, je pokud by spočítala diskrétní logaritmus $\log_g(g^a) = a$, nicméně předpokládáme, že to je obtížné. Na klasických počítačích jsou nejlepší známé útoky na problémy, jako diskrétní logaritmus a faktorizace čísla, na čemž jsou založené protokoly jako RSA, subexponenciální, nicméně na počítačích kvantových jsou už od poloviny 90. let známé algoritmy polynomiální. V čem však takto podstatné zrychlení spočívá?

2.1 Kvantové počítače

If computers that you build are quantum,
Then spies of all factions will want 'em.
Our codes will all fail,
And they'll read our email,
Till we've crypto that's quantum, and daunt 'em.

Jennifer a Peter Shorovi

Při diskuzi moderní kryptografie se často zmiňuje, že kvantové počítače dokáží problémy, jako faktorizaci čísla či diskrétní logaritmus, vyřešit v polynomiálním čase, přičemž nejrychlejší známé algoritmy pro klasické počítače pracují v čase subexponenciálním. V čem ale takto podstatné zrychlení spočívá?

Ve světě kvantových obvodů místo s klasickými bity pracuje s *qubity*. V n bitovém systémů máme 2^n různých stavů, které v n qubitovém systému tvoří generátory našeho prostoru. Podstatou je, že před pozorováním nemá daný qubit jednu z těchto hodnot, ale jejich (komplexní) superpozici. Generátory systému s jedním qubitem jsou stavy $|0\rangle = \begin{bmatrix}1\\0\end{bmatrix}, |1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$, systém je tedy:

$$\alpha|0\rangle + \beta|1\rangle,$$

kde α, β jsou komplexní čísla $|\alpha|^2 + |\beta|^2 = 1$. Zápis $|\psi\rangle$ je tzv. *ket* notace, kde ψ je vektor.

V dvojqubitovém systému máme čtyři báze a stav takového systému je:

$$\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle,$$

kde $\alpha, \beta, \gamma, \delta$ jsou komplexní čísla s $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. Qubity jsou značně nestabilní, musí být uchovány v izolované soustavě, nejčastěji v neutrinu, přičemž jakékoli narušení, i pouhé pozorování hodnoty qubitu, ho kolapsuje na jednu hodnotu, kterou už pak zůstane. Při pozorování má qubit pravděpodobnost ukázat stav právě takovou, kolik je druhá mocnina absolutní hodnoty příslušného koeficientu, proto ona normalizační podmínka. Pokud bychom pozorovali náš jedno-qubitový systém, s pravděpodobností $|\alpha|^2$ získáme výstup 0, s pravděpodobností $|\beta|^2$ získáme 1.

Můžeme ale též náš qubit vyjádřit ve vektorovém zápisu:

$$|\psi\rangle = \begin{bmatrix}\alpha\\\beta\\\gamma\\\delta\end{bmatrix},$$

což samozřejmě zobecníme pro systémy více qubitů. Tento vektor je díky naší podmínce jednotkový. V klasických obvodech máme brány, které jsou lineární zobrazení našich stavů, příklady takových bran jsou OR a NOT. V kvantových obvodech bereme jako brány právě unitární matice a jejich operaci násobení, neboť ty zachovávají normu vektoru, jejich výsledky jsou proto opět qubity.

Nedá moc práce ukázat, že všechny operace proveditelné na klasickém obvodu jsou replikovatelné kvantovými branami, model kvantového počítače, jakožto obvodu, je tak alespoň stejně silný jako počítač klasický.

Jedním z důvodů, proč se věří, že s veřejně dostupnými kvantovými počítači přijde nová éra výpočetní techniky je, že existují procesy, o kterých se dodnes neví, zda jsou v polynomiálním čase proveditelné na počítači klasickém, a jejichž kvantové algoritmy již byly nalezené. Příkladem je Fourierova transformace, která ?? Na klasickém počítače ??, nicméně Quantum Fourier Transfort (QFT)

Dále násobení matic, pro které nejlepší známé klasické algoritmy běží v ???? čase, čase je pro

proč to Shor rozbíjí. –

2.2 SIDH

Nyní, když jsme již trochu obeznámeni s kvantovými algoritmy, pojďme se vrátit k eliptickým křivkám. Zjevnou adaptací Diffie-Hellmanova protokolu je protokol, který nese název ECDH (Elliptic Curve Diffie-Hellman):

Veřejné parametry: Prvočíslo p a eliptická křivka E nad $\mathbb{Z}_p$ s generátorem $G \in E(\mathbb{Z}_p)$.

Alfréd		Blažena
Vstup: $a \leqslant \#E(\mathbb{Z}_p) - 1$		**Vstup:** $b \leqslant \#E(\mathbb{Z}_p) - 1$
	$\xrightarrow{[a]G}$	
	$\xleftarrow{[b]G}$	
$G_{AB} := [a]([b]G) = [a][b]G$		$G_{BA} := [b]([a]G) = [b][a]G$
Výstup: G_{AB}		**Výstup:** G_{BA}

Algoritmus 2: Protokol ECDH

Tento protokol je založen na předpokladu, že diskrétní logaritmus na eliptických křivkách, tedy ze znalosti P a $[n]P$ spočíst n, je těžký problém. Není znám žádný algoritmus, který by nezískal společné tajemství výpočtem privátních klíčů obou stran.

??

Pojďme se nyní znovu podívat na větu (1.3.8). Vidíme, že křivky $\phi(\psi(E)), \psi(\phi(E))$ sdílí j-invariant, neboť jsou isomorfní, což by v potenciálním protokolu založeném na isogeniích mohlo být sdílené tajemství obou stran. Pokud tak mají obě strany danou křivku E nad $\mathbb{Z}_p$, vyberou si tajné isogenie ϕ_A, resp. ϕ_B, pošlou druhé straně $\phi_A(E)$, resp. $\phi_B(E)$ a obě strany již snadno spočtou své tajemství. Takové myšlenky měli De Feo a Jao v [?], nicméně než se dostaneme přímo k jejich navrhovaném protokolu SIDH, musíme diskutovat několik důležitých detailů, které se vyhýbají známým útoků, případně usnadňují výměnu.

Jak napovídá název protokolu, požadujeme supersingularitu E. Pak totiž z věty ? je $\#E(\mathbb{Z}_{p^2}) = p+1$ a $E(\mathbb{Z}_{p^2}) \cong \mathbb{Z}_{p+1} \times \mathbb{Z}_{p+1}$. Pro prvočíslo $p = \ell_A^{e_A}\ell_B^{e_B} - 1$, kde ℓ_A, ℓ_B jsou prvočísla, proto existují dva body G_1, G_2 řádu $\ell_A^{e_A}\ell_B^{e_B}$, které generují $E(\mathbb{Z}_p^2)$. Speciálně dvojice $\langle P_A, Q_A\rangle := \langle [\ell_B^{e_B}]G_1, [\ell_B^{e_B}]G_2\rangle$, resp. $\langle P_B, Q_B\rangle := \langle [\ell_A^{e_A}]G_1, [\ell_A^{e_A}]G_2\rangle$, generují po řadě $\ell_A^{e_A}$, $\ell_B^{e_B}$ torzi.

Uvažme bod $P \in E[\ell_A^{e_A}]$ řádu ℓ_A^t a separabilní isogenii $\phi : E \mapsto E/\langle P\rangle$. Pokud bychom chtěli $E/\langle P\rangle$ spočíst, stačilo by spočítat celou $\langle P\rangle$ a za pomocí Véluových formulí spočíst výslednou křivku v exponenciálním čase $O(\ell_A^t)$, což zjevně není optimální.

Veřejné parametry: Grupa G řádu p, kde p je prvočíslo, s generátorem g.

Alfréd		Blažena
Vstup: $a \in G$		**Vstup:** $b \in G$
	$\xrightarrow{g^a}$	
	$\xleftarrow{g^b}$	
$G_{AB} := \left(g^b\right)^a = g^{ab}$		$G_{BA} := \left(g^a\right)^b = g^{ba}$
Výstup: G_{AB}		**Výstup:** G_{BA}

Algoritmus 1: Diffie-Hellmanova výměna

Závěr

zu ende

Literatura

[1] Marcus, Daniel A.: *Number fields*. New York: Springer-Verlag, 1977.

[2] Matushak, Andy a Nielsen, Michael. *Quantum computing for the very curious*. San Francisco, 2019. Dostupné z: `https://quantum.country/qcvc`.

[3] Raclavský, Marek. *Racionální body na eliptických křivkách*. Diplomová práce. Praha, 2014.

[4] Rosický, Jiří: *Algebra*. Brno: Masarykova univerzita, 2002.

[5] Silverman, Joseph H.: *The Arithmetic of Elliptic Curves*. New York: Springer-Verlag, 1992.

[6] Shor, Peter W.: *Polynomial-Time Algorithms for Prime Factorizationand Discrete Logarithms on a Quantum Computer*. New York: Springer-Verlag, 1994. Dostupné z: `https://arxiv.org/abs/quant-ph/9508027`.

[7] Washington, Lawrence C.: *Elliptic Curves: Number theory and cryptography*. Maryland, 2008.