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 ??. Na těchto adresách se dají najít důkazy všech uvedených tvrzení, u příliš technických či obtížných tvrzení důkaz uvádět nebudeme.

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, 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 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 nekonečnu, 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.

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 $(X : Y : Z) \in K^3$ splňující:*

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

Eliptické křivky pro naše účely nedefinujeme přes tělesa s $\operatorname{char} K = 2, 3$. Obecně se eliptické křivky definují přes všechna tělesa, nicméně pro tělesa s charakteristikou různou od $2, 3$, můžeme eliptické křivky napsat ve výše uvedené jednoduché formě.

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 *Weierstrassovu*. Pak naše křivka je množina bodů $(x, y) \in K^2$ splňujících $y^2 = x^3 + ax + b$ spolu s bodem v nekonečnu $\mathcal{O} = (0 : 1 : 0)$. Množinu těchto bodů (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 právě $\mathcal{O}$.

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 $y = 0$ tento bod protíná dvakrát a ten třetí je právě 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) $P + (-P) = \mathcal{O}$.

Poznamenejme, že asociativita není příliš jednoduchá dokázat, ??. Při takto definovaném sčítání 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 1.1.4. – *ta věta* –

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

Definice 1.1.5. *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 $[0]_E P = \mathcal{O}$.

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

Pro q mocninu prvočísla obecně platí $\mathbb{F}_q \cong \mathbb{Z}_q$, s okruhy $\mathbb{F}_q$ tak můžeme pracovat identicky jako při práci v $\mathbb{Z}_q$.

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

Definice 1.1.6. *Buď n celé číslo. O množině všech $P \in E$, že $[n]_E P = \mathcal{O}$, řekneme, že tvoří n-torzi E a tuto množinu budeme značit $E[n]$.*

Věta 1.1.7. *Nechť je E eliptická křivka nad K a m nenulové číslo. Pak:*

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

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

Definice 1.1.8. *Pokud máme $E[c] \cong \{\mathcal{O}\}$, nazveme E jako supersingulární. Jinak E budeme říkat ordinární.*

??

1.2 Zobrazení mezi eliptickými křivkami

V předchozí kapiole jsme se seznámili s násobení bodů čísel n, což je morfismus $E \mapsto E$. Množina všech endomorfismů z E Se sčítáním endomorfismů $\phi, \psi : E \mapsto E, P \in E$:

$$(\phi + \psi)(P) := \phi(P) + \psi(P)$$

a skládáním $\psi \circ \phi = \phi(\psi)$, spolu s nulovou mapou $0 : P \mapsto \mathcal{O}$, tvoří endomorfismy E okruh, který budeme značit $End(E)$. ??

(Tady bych diskutova End(E), Frobenius trochu)

Pojďme se místo endomorfismů dívat obecně na morfismy mezi eliptickými křivkami. Vidíme, že mapa $(x, y) \mapsto (u^2x', u^3y') := (x', y')$ nám dává isomorfismus přes K mezi křivkami:

$$y^2 = x^3 + ax + b \mapsto y^2 = x^3 + u^4x + u^6b \Leftrightarrow (y')^2 = (x')^3 + ax' + b$$

pro libovolné $u \in \overline{K}$. Takovéto mapy jsou zřejmě jediné isomorfismy přes K dané lineárním přenásobením souřadnic. Uvažme nicméně zobrazení převádějící křivky:

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

dané $(x, y) \mapsto (x, \sqrt{d}y)$. Nemáme nutně isomorfismus přes K, ale přes jeho rozšíření $K(\sqrt{d})$.

Definice 1.2.1. *Pro eliptickou křivku $E : y^2 = x^3+ax+b$ nad K definujeme její kvadratický twist jako křivku $E^d : dy^2 = x^3 + ax + b$ nad $K(\sqrt{d})$ pro $\sqrt{d} \notin K, d \in K$.*

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

Definice 1.2.2. *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 v K definovaný, neboť eliptické křivky mají nenulový diskriminant.

Věta 1.2.3. *Dvě křivky definované nad K jsou isomorfní nad $\overline{K}$ právě pokud mají stejný j-invariant.*

tu příklady !!!!, možná něco o tom, kdy je to nula etc –¿ proč je tam 1728

1.3 Isogenie

Pojďme se nyní z endomorfismů naší křivky E přesunot na zobrazení převádící na sebe i různé eliptické křivky. Speciálně se pojďme podívat na zobrazení daná lomenou funkcí nad K, která na sebe tyto křivky převádí.

Definice 1.3.1. *Ať $E_1, E_2 \in \overline{K}$ jsou eliptické křivky. Pak surjektivní morfismus $E_1 \mapsto E_2$ 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 rozumnět jejímu stupni jako racionální mapě a budeme značit* $\deg \phi$.

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

– Nějak chci zmínit duální isogenii, protože to chci zmínit a nevím, kam jinam to dát –

Věta 1.3.3. *Buď $\phi : E \mapsto E'$ isogenie stupně n. Pak existuje jediná isogenie $\hat{\phi} : E' \mapsto E$, která 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.

Do nekonečna E_2 se ne nutně z ϕ pošle pouze bod v nekončnu E_1, ale i kořeny jmenovatele jejího zlomku, které náleží do $\overline{K}$. O množině bodů, které se zobrazí do nekonečna, budeme hovořit jako o *jádře* naší isogenie ϕ a budeme jej značit $\ker \phi$, počet jeho prvků jako $\# \ker \phi$. Poznamenejme, že pokud $P \in \ker \phi$, pak je tam i $-P$, obecně $\ker \phi$ tvoří podgrupu $E(K)$. Navíc si za chvíli ukážeme, že existuje třída isogenii taková, že každá podrgupa $E(K)$ jednoznačně určuje isogenii z E.

Definice 1.3.4. *Mějme $E, E_1 \in \overline{K}$ a $\phi : E \mapsto E_1$ 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.*

Nás budou dále zajímat hlavně isogenie separabilní. ??

Věta 1.3.5. *Každá separabilní isogenie ϕ z E je jednoznačně určena svým jádrem.*

Pokud je tak $G = \ker \phi$ grupa tvořená jádrem ϕ, můžeme značit E/G cílovou křivku ϕ, jakožto faktorgrupa E/G. Separabilní isogenie z $E \mapsto E_1$ je daná racionální mapou a známe-li její jádro, dokážeme spočíst explicitně naši isogenii. Vzorce udávající přesný tvar separabilní isogenie z $E \mapsto E_1$ s daným jádrem se nazívají *Véluovy* po člověku, který je publikoval.

Věta 1.3.6. *Vélu Formulas.*

Pak už jen někde zmínit $(E/\langle A\rangle)/\langle B\rangle \cong (E/\langle B\rangle)/\langle A\rangle$

Kapitola 2

SIDH

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 příbývá 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 právě protokol, který Diffie a Hellman navrhli. Budeme o něm dále mluvit jako o *Diffie-Hellmanově výměně*. Je založena na problému *diskrétního logaritmu* prvku $a \in \mathbb{Z}_p \setminus \{0\}$. 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, t.j. v polynomiálním čase vzhledem k velikosti grupy, jinak může útočník jednoduše privátní klíče obou stran spočíst.

-¿ tu nějak hezky typeset algoritmus

Závěr

zu ende

Literatura

[1] BENEŠ, Petr: *Zákony Reciprocity.* Diplomová práce. Brno: Masarykova univerzita, 2010.

[2] DECHENNE, Spencer. The Ramanujan-Nagell Theorem: Understanding the Proof. Dostupné z: http://buzzard.ups.edu/courses/2013spring/projects/spencer-ant-ups-434-2013.pdf

[3] DOLEŽÁLEK, Matěj. Pellova rovnice a kvadratické okruhy. In: PraSe, Organizátoři. PraSe Sborníček 2019 [online]. Sklené, 2019. Dostupné z: `https://prase.cz/soustredeni/sbornik.php?sous=47`.

[4] HRNČIAR, Maroš: *Řešení diofantických rovnic rozkladem v číselných tělesech.* Diplomová práce. Praha: 2015.

[5] HUDEC, Pavel. Odmocniny z jedničky. In: *i*KS, Organizátoři. *i*KS Sborníček 2019 [online]. Kunžak, 2019. Dostupné z: `http://iksko.org/files/sbornik8.pdf`.

[6] KUŘIL, Martin: *Základy teorie grup.*

[7] LENSTRA JR, Hendrik W.: *Solving the Pell Equation.* 2002.

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

[9] PERUTKA, Tomáš: *Vyjadřování prvočísel kvadratickými formami.* Středoškolská odborná činnost. Brno: Masarykova univerzita, 2017.

[10] PERUTKA, Tomáš: *Užití dekompoziční grupy k důkazu zákona kvadratické reciprocity.* Středoškolská odborná činnost. Brno: Masarykova univerzita, 2018.

[11] PUPÍK, Petr: *Užití grupy tříd ideálů při řešení některých diofantických rovnic.* Diplomová práce. Brno: Masarykova univerzita, 2009.

[12] RACLAVSKÝ, Marek. *Racionální body na eliptických křivkách.* Diplomová práce. Praha, 2014.

[13] ROSICKÝ, Jiří: *Algebra.* Brno: Masarykova univerzita, 2002.

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

[15] *i*KS - mezinárodní korespondenční seminář [online]. Dostupné z: `iksko.org`.