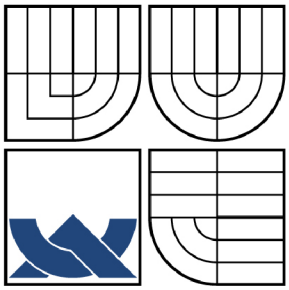
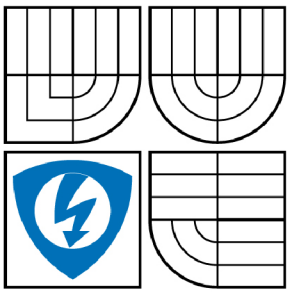


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

DEPARTMENT OF TELECOMMUNICATIONS

SOFTWAROVÁ PODPORA VÝUKY KRYPTOSYSTÉMŮ ZALOŽENÝCH NA ELIPTICKÝCH KŘIVKÁCH

SOFTWARE SUPPORT OF EDUCATION IN CRYPTOGRAPHY AREA BASED ON ELLIPTIC CURVES

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

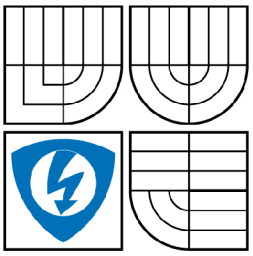
AUTOR PRÁCE — Bc. JAKUB SZTURC

AUTHOR

VEDOUCÍ PRÁCE — doc. Ing. KAREL BURDA, CSc.

SUPERVISOR

BRNO 2009



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky a komunikačních technologií

Ústav telekomunikací

Diplomová práce

magisterský navazující studijní obor
Telekomunikační a informační technika

Student:	Bc. Jakub Szturc	***ID:***	83196
Ročník:	2	***Akademický rok:***	2008/2009

NÁZEV TÉMATU:

Softwarová podpora výuky kryptosystémů založených na eliptických křivkách

POKYNY PRO VYPRACOVÁNÍ:

Prostudujte a popište problematiku kryptosystémů založených na eliptických křivkách (ECC). Na tomto základě navrhněte a prakticky zrealizujte software pro podporu výuky uvedené problematiky. Software musí být spustitelné na běžném webovém prohlížeči. Kromě popisné části a praktických příkladů musí obsahovat zadání příkladů k samostatnému řešení, výsledky jejich řešení a možnost interaktivních výpočtů a zobrazení. Minimálním obsahem práce je kryptosystém ECC pro šifrování i podepisování, vlastnosti eliptických křivek a algoritmus sčítání bodů.

DOPORUČENÁ LITERATURA:

[1] Stallings, W.: Cryptography and Network Security. Prentice Hall, Upper Saddle River 2006.
[2] Schneier, B.: Applied Cryptography. John Wiley, New York 1996.

Termín zadání:	9.2.2009	***Termín odevzdání:***	26.5.2009

Vedoucí práce: doc. Ing. Karel Burda, CSc.

prof. Ing. Kamil Vrba, CSc.
Předseda oborové rady

UPOZORNĚNÍ:

Abstrakt

Diplomová práce se zaměřením na kryptografii založenou na eliptických křivkách se skládá ze čtyř hlavních částí. První část poskytuje přehled o základních kryptografických a matematických pojmech. Klíčovým bodem této práce je druhá část, ve které jsou podrobně popsány mechanismy sčítání dvou bodu na eliptické křivce a přičtení bodu k sobě samému nad různými tělesy. Na tomto mechanismu je založena prakticky celá problematika. Ve třetí části jsou uvedeny nejznámější algoritmy a protokoly určené k výměně klíčů, šifrování a digitálnímu podpisu.

Cílem této práce je navržení softwaru pro podporu výuky. Tento materiál je vytvořen jako webová prezentace, ve které jsou popsány teoretické základy a hlavní vlastnosti kryptosystémů založených na eliptických křivkách. Celá problematika je podpořena praktickou ukázkou výpočtů příkladů, jsou zde i příklady pro samostatnou práci. Jako doplnění jsou připraveny java aplety, které umožňují interaktivní možnost vyzkoušení si základních parametrů křivek, nebo ověření výpočtů.

Klíčová slova

Kryptografie, asymetrický, eliptický, křivka, těleso, bod, ECDSA

Abstract

The master‘s thesis is focusing on cryptography based on elliptical curves consists of four main parts. The first part provides an overview of the basic cryptographic and mathematical concepts. A key element of this work is the second part which are described in detail the mechanisms of counting two points on elliptic curve and counting point to themselves over the various fields. On this mechanism is based almost the entire issue. In the third section provides the best-known algorithms and protocols for key exchange, encryption and digital signature.

The goal of this paper is to devise software to support teaching. This material is created as a web presentation, which described the theoretical foundations and the main characteristics of cryptosystems based on elliptical curves. The whole issue is supported by practical examples of calculations examples, there are also examples for independent work. Additionally, java applets are prepared that allow an interactive opportunity to try the basic parameters of curves, or verify the calculations.

Key words

Cryptography, asymmetrical, elliptic, curve, field, point, ECDSA

SZTURC, J. *Softwarová podpora výuky kryptosystémů založených na eliptických křivkách.* Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009. 65 s. Vedoucí diplomové práce doc. Ing. Karel Burda, CSc.

Prohlášení o původnosti práce

Prohlašuji, že svojí diplomovou práci na téma Softwarová podpora výuky kryptografie pomocí eliptických křivek jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením tohoto projektu jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č.140/1961 Sb.

V Brně dne 19. 5. 2009 .....

podpis autora

Poděkování

Děkuji vedoucímu diplomové práce doc. Ing. Karlu Burdovi, CSc., za velmi užitečnou metodickou pomoc, poskytnutí užitečných informací v oblasti kryptografie a v neposlední řadě také cenných rad při zpracování diplomové práce.

V Brně dne 19.5.2009 .....

OBSAH

SEZNAM OBRÁZKŮ

SEZNAM TABULEK

1 ÚVOD

Samotná problematika eliptického šifrování je dost rozsáhlá a také složitá, protože nejde o klasické metody šifrování, jako jsou RSA, IDEA, ElGamal nebo DES. Je to postup, který se značně liší od všech jiných metod svojí *složitou strukturou* a *těžkou pochopitelností.*

Diplomová práce se zabývá teoretickým vysvětlením, co se skrývá pod pojmem „kryptosystémy založené na bázi eliptických křivek“. Druhá a třetí kapitola jsou věnovány vysvětlení některých základních kryptografických pojmů, přímo spojených s danou problematikou. Je zde uveden i matematický background, kde jsou přiblíženy matematické pojmy, ze kterých vychází kryptografie založená na eliptických křivkách a *jejichž zvládnutí je nezbytné* k další práci s nimi. K těmto patří zvládnutí operací modulo *n*, výpočet inverzního prvku nebo porozumění konečným tělesům a grupám.

Ve čtvrté kapitole jsou podrobně popsány mechanismy *sčítání bodů* na eliptické křivce a tzv. *zdvojení bodu*, ze kterých vychází následná problematika. Složitost eliptického diskrétního logaritmu je založena právě na násobení bodu konstantou. Popsány jsou i základní vlastnosti eliptických kryptosystémů. Pro porovnání s jinými asymetrickými algoritmy jsou uvedeny použité délky klíčů při stejné bezpečnosti.

Jak je uvedeno v zadání, další kapitola se zabývá kryptosystémy ECC pro *šifrování* a *podepisování*. U každého systému je uvedeno několik možností a zástupců daného problému a jejich algoritmy. Mezi nejznámější patří algoritmy pro ustanovení společného klíče podle Difiie-Helmmana. U podpisu je to např. ECDSA.

Poslední část je věnována praktické stránce diplomové práce. Je zde popsán zrealizovaný software určený k podpoře výuky. Tento software je navržený tak, aby byl spustitelný z libovolného webového prohlížeče. Jsou zde shrnuty nejdůležitější vazby a teoretické informace, ke zvládnutí dané problematiky. Teorie je podpořena vypočtenými příklady. Studenti mají možnost si nabyté vědomosti procvičit na připravených příkladech a ověřit správnost výsledků.

2 ZÁKLADNÍ KRYPTOGRAFICKÉ POJMY

2.1 Obecné pojmy

Kryptologie - nauka o metodách utajování. Název z řeckého slova „kryptós“ = skrytý. Kryptologie je souhrnný název pro kryptografii, steganografii a kryptoanalýzu.

Kryptografie - technika a věda o utajování zpráv. Zabývá se šifrovacími technologiemi, algoritmy, mechanizmy, stavebními bloky a transformacemi používanými pro utajování zpráv. Kryptografie šifrované zprávy vytváří.

Steganografie - věda zabývající se „skrýváním zpráv“. Základem je utajený algoritmus na veřejném klíči, což může znamenat komunikaci tam, kde se zdá, že žádná komunikace není. (V praxi – text uložený v obrázku na místech, kde se vyskytuje náhodný šum.)

Kryptoanalýza - věda zabývající se získáváním obsahu zašifrovaných zpráv, aniž by bylo použito tajného klíče. Tzv. prolamování kódu. Kryptoanalýza šifrované zprávy luští a tím testuje odolnost kryptografických systémů.

Autentizace - je proces ověření identity. Např. opatření dokumentu digitálním podpisem, který zaručuje identitu uživatele, nebo ověření uživatele před přihlášením do systému elektronického bankovnictví.

Šifrovací klíč (k) - říká šifrovacímu algoritmu jak má data (de)šifrovat, podobá se počítačovým heslům, avšak neporovnává se zadaná hodnota s očekávanou, nýbrž se přímo používá a vždy tedy dostaneme nějaký výsledek, jehož správnost závisí právě na zadaném klíči. Existují 3 druhy klíčů. **Tajný klíč** (symetrická šifra→ $k_E=k_D=TK$), **veřejný** a **soukromý** klíč (asymetrická šifra → $SK \neq VK$).

Délka klíče ovlivňuje kromě jiného časovou náročnost při útoku hrubou silou - což je kryptoanalytická metoda, kdy postupně zkoušíme všechny možné hodnoty, kterých klíč může nabývat.

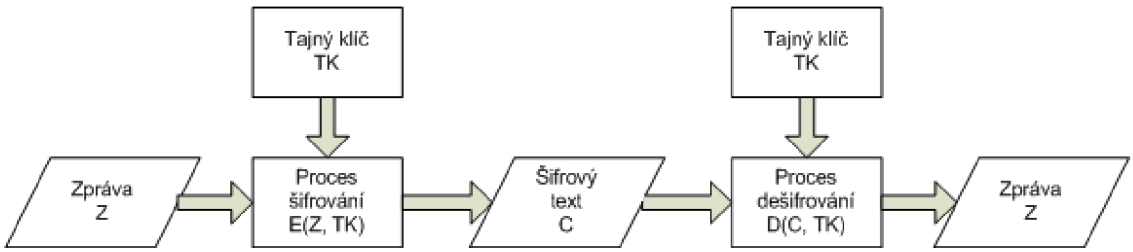
Hashovací funkce - pro pozdější vysvětlení digitálního podpisu se ještě musíme zastavit u pojmu **hash** (otisk), hashovací funkce [18]. Obecně řečeno se jedná o matematickou funkci, kterou lze v jednom (přímém) směru spočítat velice snadno, zatímco v opačném směru (inverzním zobrazení) se výpočty dají provést jen s velkými obtížemi (resp. jsou takřka nemožné). Výsledkem hashovací funkce je zpravidla 128 nebo 160 bitů dlouhá sekvence znaků, která jednoznačně charakterizuje vstupní blok dat

(dokument, soubor, e-mail apod.). Pokud by se v původních datech změnil byť jen jeden jediný bit, jejich hash se výrazně změní.

Digitální certifikát se skládá ze dvou základních komponent: veřejného klíče (VK) a osobních dat jeho vlastníka. Celistvost certifikátu je zaručena digitálním podpisem, který může vytvořit sám jeho vlastník pomocí soukromého klíče z tohoto klíčového páru (self-signed certificate) nebo certifikační autorita svým soukromým klíčem (SK). Z hlediska důvěryhodnosti má logicky větší váhu digitální certifikát vydaný certifikační autoritou, která ručí za ověření identity jeho vlastníka. Kromě již zmiňovaných obsahují certifikáty také další údaje, jako je doba vypršení platnosti certifikátu (expirace), jméno vydávající certifikační autority, evidenční číslo certifikátu, certifikační cesta (posloupnost certifikátů certifikační autority, které jsou potřebné k ověření pravosti certifikátu), případně ještě další doplňující údaje.

2.2 Symetrické šifrování

U symetrického šifrování se pro zašifrování i pro dešifrování dat používá jeden šifrovací klíč – tajný klíč *TK* ($k_E = k_D$), viz obr. **2.1**. Platí, že šifrovací klíč a dešifrovací klíč jsou od sebe odvoditelné v reálném čase. Stejný klíč musí mít k dispozici všichni, kdo se šifrovanými daty pracují [16]. Logicky tedy vyplývá potřeba zajistit jeho bezpečné předání určeným osobám. Ve chvíli, kdy dojde k jeho prozrazení, byť jen jedinou zúčastněnou osobou, jsou všechny jím zašifrované informace prozrazeny [18].



Obr. 2.1: Princip symetrického šifrování

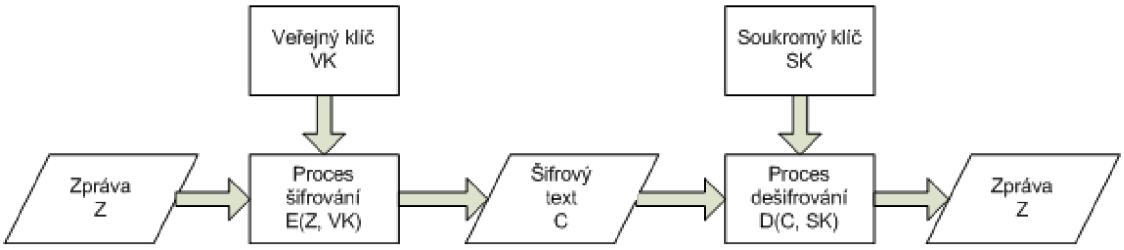
Mezi nejznámější symetrické šifrovací algoritmy patří DES, 3DES, IDEA, BlowFish a CAST. DES byl vyvinut v laboratořích IBM již v průběhu sedmdesátých let. V roce 1977 se dokonce stal vládní normou USA pro šifrování. Používá klíč, který má délku 56 bitů. Vzhledem k tomu, že se výpočetní technika postupem času výrazně zdokonalila, dnes již DES z hlediska bezpečnosti nedostačuje. Podařilo se jej dokonce prolomit pomocí tzv. „hrubého útoku" (vyzkoušením všech možných kombinací klíče). Proto byla posléze vyvinuta jeho zesílená varianta - 3DES (Triple-DES). Jeho princip je

velice jednoduchý. Data jsou jednoduše přešifrována dvakrát nebo třikrát. TripleDES proto pracuje s dvojnásobným (112 bitů) nebo trojnásobným klíčem (168 bitů). Oproti DES je relativně pomalý, ale mnohem bezpečnější. Jedním z důvodů použití tohoto řešení byla i kompatibilita se staršími systémy používajícími DES. Algoritmus IDEA je poměrně perspektivní. Je vyzbrojen klíčem o délce 128 bitů a vzhledem k DESu je znatelně rychlejší. Další algoritmus BlowFish je zajímavý tím, že může používat proměnnou délku šifrovacího klíče od 32 do 448 bitů. Obvykle se však používá s klíčem 128bitovým. Není zatížen patentovými právy, je rychlý a bezpečný. Velice se mu podobá CAST (tvůrci jsou CarlisleAdams a StaffordTaverns), který umí taktéž používat proměnnou délku klíče. A v neposlední řadě patří mezi zástupce symetrických šifer AES. Délka klíče může být 128, 192 nebo 256 bitů. Metoda šifruje data postupně v blocích s pevnou délkou 128 bitů. Šifra se vyznačuje vysokou rychlostí šifrování.

2.3 Asymetrické šifrování

Podstatou asymetrického šifrovacího systému jsou dva různé, ale „vzájemně kompatibilní“ šifrovací klíče – jeden veřejný (*VK* - pro zašifrování) a druhý soukromý (*SK* - pro dešifrování). Princip vidíme na obrázku **2.2.** Platí, že veřejný a soukromý klíč nejsou odvoditelné v reálném čase. Veřejný klíč je určen k volnému šíření a je distribuován všem osobám, se kterými komunikujeme. Oproti tomu soukromý klíč musí zůstat přísně utajen u jeho vlastníka, který by jej měl chránit jako oko v hlavě. To, co bylo zašifrováno veřejným klíčem, lze dešifrovat pouze soukromým a naopak. Jeden jediný klíč nelze použít k zašifrování i opětovnému dešifrování. Důvodem této vlastnosti asymetrických algoritmů jsou použité matematické funkce, jejichž reverzní výpočet je prakticky neproveditelný.

Asymetrické šifrovací algoritmy jsou v porovnání se symetrickými obecně výrazně pomalejší. Asymetrické kryptosystémy (např. RSA, ECPKC, LUC), kryptografické protokoly i metody digitálních podpisů používají komplikované operace s dlouhými čísly, které by standardnímu PC trvaly příliš dlouho. Proto se často šifruje klasickými symetrickými systémy (např. DES, IDEA, WinCros) a asymetrickými systémy se šifrují pouze relativně krátké použité symetrické klíče.



Obr. 2.2: Princip asymetrického šifrování

V praxi se nejčastěji používá algoritmus RSA a nově algoritmy na bázi eliptických křivek (Elliptic Curve Cryptography, dále jen ECC). Autory RSA jsou Rivest-Shamir-Adleman. Stupeň jeho bezpečnosti je odvislá od použité délky klíče. Pro vytváření elektronického podpisu se standardně používají klíče s minimální délkou 1024 bitů. Samotný algoritmus vznikl v roce 1977 a do podzimu roku 2000 byl chráněn patentem. Jeho prolomení závisí na schopnosti útočníkova systému řešit úlohy faktorizace velkých čísel. Ve srovnání s DES je RSA samozřejmě podstatně pomalejší [18]. Při softwarových realizacích se uvádí, že je to přibližně 100 krát, při hardwarových realizacích dokonce 1000 až 10000krát. RSA je součástí řady používaných norem.

Je tedy daleko výhodnější data zašifrovat symetrickým algoritmem a náhodným klíčem, který je vygenerován jako jedinečný pouze pro danou relaci, ten posléze zašifrovat pomocí asymetrického algoritmu a přibalit k zašifrovaným datům. Celý „balíček“ je pak odeslán příjemci, který nejprve dešifruje symetrický klíč a teprve s jeho pomocí samotná data. Tímto způsobem funguje drtivá většina softwaru používajícího asymetrickou kryptografii.

2.4 Digitální podpis (DP)

Digitální podpis lze přirovnat spíše k efektu, kterým je identifikace a autentizace např. autora určitého dokumentu. Digitální podpis je založen na metodách asymetrického šifrování.

Digitální podpis má hned několik zásadních výhod:

- jeho základní vlastností je nepopiratelnost,
- je prakticky nemožné jej zfalšovat,
- lze jednoduše ověřit jeho autenticitu,
- jeho použitím je zaručena neporušenost zprávy (resp. zjištění jejího porušení),

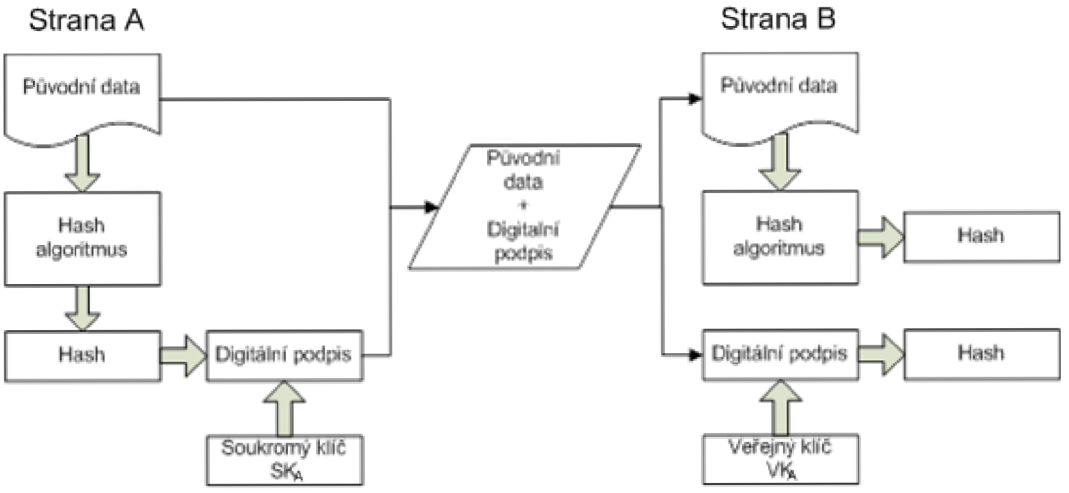
- v kombinaci se šifrováním je zpráva chráněna před vyzrazením obsahu. Navíc může obsahovat časovou značku a být tak jednoznačně určen v čase.

Sestavení DP

Pokud chceme vytvořit digitální podpis určitého souboru, musíme nejprve určit jeho hash. Jakýkoliv soubor nebo e-mailovou zprávu lze v podstatě chápat jako „obyčejný“ soubor čísel, na který aplikujeme hash algoritmus, viz obr. **2.3**. Na jeho výstupu získáme číslo o dané délce jednoznačně reprezentující vstupní data - otisk souboru. Hash je následně zašifrován pomocí soukromého klíče podepisující osoby a digitální podpis je na světě. Poněkud zjednodušeně řečeno je hash zašifrovaný soukromým klíčem digitálním podpisem [18]. Ten se pak může přidat k podepisovaným datům nebo může být transformován do podoby samostatného souboru (extra signature). Zpravidla se k němu ještě přidává i digitální certifikát podepsané osoby, který může posloužit adresátovi k ověření podpisu.

Ověření DP

Ověření podpisu probíhá analogicky k jeho vytvoření. Příjemce jednak znovu vypočte hash z původního souboru a jednak jej dešifruje pomocí certifikátu odesílatele z podpisu. Certifikát buď již má ověřující osoba k dispozici, nebo si jej může třeba stáhnout z webu certifikační autority, která jej vydala. Samotné ověření pak spočívá v porovnání obou hashů. Pokud jsou stejné, je zřejmé, že podepsanou osobou je skutečně ta, která to o sobě tvrdí (vlastník certifikátu, který jediný má k dispozici privátní klíč). Pokud se hash liší, může to znamenat pokus o padělání podpisu, pozměnění souboru cestou nebo cokoliv jiného, co ve svém důsledku vedlo k porušení nebo pozměnění dat či podpisu samotného.



Obr. 2.3: Princip digitálního podpisu

3 MATEMATICKÝ ZÁKLAD

Moderní kryptografie je postavena na algebraických strukturách. Každá algebraická struktura je systém s určitou množinou prvků, kde jsou definovány jisté operace s těmito prvky. Příkladem všeobecně známé algebraické struktury je množina reálných čísel spolu s operací sčítání a násobení - tzv. těleso reálných čísel.

V kryptografii mají uplatnění pouze struktury s konečnými množinami prvků. V současné praxi se v kryptografii nejvíce využívají tzv. komutativní grupy a tělesa.

3.1 Operace modulo n

Mohli bychom také říci zbytek po dělení nebo také, že modulo je početní operace související s operací celočíselného dělení. Např. 7 / 3 = 2 se zbytkem 1. Také můžeme říci, že 7 modulo 3 = 1, zkráceně 7 mod 3 = 1. Je-li zbytek po dělení a / n nula, říkáme, že a je dělitelné n [1]. Obecně lze tedy tuto operaci zapsat jako:

$$x \bmod n = y, \tag{3.1}$$

kde x je celé číslo, n je přirozené číslo a $y \in \{0,1,2 ..., n - 1\}$.

Pro výslednou hodnotu y platí:

$$y = x - \left\lfloor \frac{x}{n} \right\rfloor .n, \tag{3.2}$$

kde $\lfloor z \rfloor$ je celé číslo, které je ve směru k $(-\infty)$ nejbližší k číslu z. Libovolný celočíselný vstup se tak převede na jedno z n čísel konečné množiny $\{0,1,2 ..., n - 1\}$.

3.2 Konečná komutativní grupa

Označíme-li operaci jako sčítání (+), mluvíme o aditivní grupě a píšeme $(G, +)$, píšeme-li ji jako násobení $(\cdot)$, hovoříme o multiplikativní grupě a píšeme $(G, \cdot)$. Na výběru značky pro grupovou operaci nezáleží, jde vždy o tutéž operaci, jen zapisovanou jinak. Konečná komutativní grupa je tedy konečná množina prvků G spolu s operací sčítání, kdy pro všechny prvky $a, b, c \in G$ platí:

1. Když $a + b = c$, tak $c \in G$. (Algebraická struktura je vůči sčítání tzv. uzavřená.)
2. $(a + b) + c = a + (b + c)$. (Asociativní zákon.)
3. $a + b = b + a$. (Komutativní zákon.)

4. Existuje nulový prvek $0 \in G$, kdy pro každé $a \in G$ platí: $a + 0 = 0 + a = a$.
5. Pro každé $a \in G$ existuje opačný prvek $-a$, kdy platí: $a + (-a) = -a + a = 0$.

Příkladem konečné komutativní grupy jsou právě eliptické křivky. Prvky množiny jsou body eliptické křivky P_i a pro tyto body je definováno sčítání $P_k = P_i + P_j$. Tím je možné definovat násobení bodu konstantou c: $P_k = c \cdot P_i = P_i + P_i \ ... \ + P_i$. Pro tuto operaci platí, že ze znalosti bodu P_k a P_i je prakticky nemožné určit hodnotu c.

3.3 Konečné těleso

Konečné těleso je stejně jako u grupy konečná množina prvků F spolu s operací sčítání a operací násobení, kdy pro všechny prvky $a, b, c \in F$ platí:

1. když $a + b = c$, respektive $a \cdot b = c$, tak $c \in F$. (Algebraická struktura je vůči sčítání, respektive násobení uzavřená [1].)
2. $(a + b) + c = a + (b + c)$, respektive $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. (Asociativní zákon.)
3. $a + b = b + a$, respektive $a \cdot b = b \cdot a$. (Komutativní zákon.)
4. $a \cdot (b + c) = a \cdot b + a \cdot c$. (Distributivní zákon.)
5. existují neutrální prvky 0 a $1 \in F$, kdy pro každé $a \in F$ platí: $a + 0 = a$, respektive $a \cdot 1 = a$,
6. pro každé $a \in F$ existuje opačný prvek $-a$, kdy platí: $a + (-a) = 0$,
7. pro každé $a \neq 0$ z F existuje inverzní prvek a^{-1}, kdy platí: $a \cdot a^{-1} = 1$.

Příkladem konečného tělesa jsou prvočíselná a Galoisova tělesa.

3.3.1 Prvočíselné těleso

Příkladem konečného tělesa, které má velký význam v oblasti šifrování je těleso označené jako Fp. Konečné těleso Fp je algebraická struktura, která je tvořená zbytky po dělení celých kladných čísel prvočíslem p. Množinu Fp pak tvoří prvky $Fp = \{0, 1, 2, ..., p - 1\}$ spolu s operací sčítání modulo p a násobení modulo p tvoří prvočíselné těleso Fp [1]. Vlastností tělesa Fp je skutečnost, že operace sčítání a násobení modulo p prováděné s čísly z množiny Fp mají za výsledek opět číslo z této množiny.

- V prvočíselném tělese existuje pro každé číslo $a \in Fp$ opačné číslo $-a$, kdy platí: $a + (-a) = 0$.
- Pomocí opačných čísel můžeme definovat operaci odečítání: $a + (-b) = a - b$.

- V prvočíselném tělese existuje pro každé číslo $a \neq 0$ inverzní číslo a^{-1}, kdy platí: $a \cdot a^{-1} = 1$.
- Pomocí inverzních čísel můžeme definovat operaci dělení: $a \cdot b^{-1} = a / b$.

Dále platí:

- jestliže $a, b \in Fp$, pak výsledkem $a + b$ bude číslo $c \in Fp$, pro které platí $a + b \equiv c \bmod p$,
- jestliže $a, b \in Fp$, pak výsledkem $a \cdot b$ bude číslo $c \in Fp$, $0 \leq c \leq p - 1$, pro které platí $a \cdot b \equiv c \bmod p$,
- jestliže $a \in Fp$, pak multiplikativní inverzní prvek k prvku a je a^{-1}: $a \cdot a^{-1} \equiv 1 \bmod p$.

3.3.2 Binární konečné těleso

Konečné pole F_2^m se nazývá binární konečné pole. Můžeme si ho představit jako vektor délky m nad F_2. Potom existuje m elementů $\{\alpha_0, \alpha_1, ..., \alpha_{m-1}\}$ v F_2^m a každé $\alpha \in F_2^m$ může být jedinečně zapsání jako:

$$a = \sum_{i=0}^{m-1} a_i \alpha_i \,, \tag{3.3}$$

kde $a_i \in \{0, 1\}$.

Množina prvků $\{\alpha_0, \alpha_1, \ldots, \alpha_{m-1}\}$ se nazývá báze F_2^m nad F_2. Potom můžeme reprezentovat a jako binární vektor $(a_0, a_1, ..., a_{m-1})$. Všeobecně se používají dvě báze F_2^m nad F_2 a to polynomiální a normální [17].

Polynomiální báze

Nechť je:

$$f(x) = x^m + \sum_{i=0}^{m-1} f_i x^i \,, \tag{3.4}$$

kde $f_i \in \{0, 1\}$, pro $i = 0, 1, \ldots, m - 1$, je ireducibilní polynom stupně m nad F_2. Pro každý ireducibilní polynom existuje reprezentace polynomiální báze. V této reprezentaci, každý element F_2^m koresponduje s binárním polynomem stupně menšího než m. Proto pro $a \in F_2^m$ existuje m čísel $a_i \in \{0, 1\}$. Jako např.

$$a = a_{m-1} x^{m-1} + ... + a_1 x + a_0 \,.$$

Prvek pole $a_i \in F_2^m$ je obyčejně označovaný bitovým řetězcem $(a_{m-1}, \ldots, a_1, a_0)$ délky m. Následující operace jsou definované na elementech F_2^m, kdy je použitá polynomiální reprezentace s ireducibilním polynomem $F(x)$. Předpokládejme, že $a = (a_{m-1}, \ldots, a_1, a_0)$ a $b = (b_{m-1}, \ldots, b_1, b_0)$.

- Sčítání: $a + b = c = (c_{m-1}, \ldots, c_1, c_0)$, kde $c_i = (a_i + b_i) \bmod 2$. Potom sčítání odpovídá bitové operaci XOR.
- Násobení: $a \cdot b = c = (c_{m-1}, \ldots, c_1, c_0)$, kde

$$c(x) = \sum_{i=0}^{m-1} c_i x^i \tag{3.5}$$

je zbytek po dělení polynomu polynomem $f(x)$

$$(\sum_{i=0}^{m-1} a_i x^i)(\sum_{i=0}^{m-1} b_i x^i) .$$

3.4 Eliptický diskrétní logaritmus

Základem každého kryptosystému je obtížná matematická úloha, kterou je výpočetně nemožné řešit. Problém diskrétního logaritmu je základ pro bezpečnost mnoha šifrovacích systémů včetně eliptických kryptosystémů. Dá se říci, že ECC spoléhá na obtížnost eliptického diskrétního logaritmu [6].

Problém diskrétního logaritmu v multiplikativní grupě $(G, \cdot)$ je definován takto: jsou dány prvky r a q grupy a prvočíslo p a má se nalézt takové číslo k, aby byla splněna rovnice $r = qk \bmod p$.

Pokud grupa eliptické křivky bude popsána multiplikativní notací, pak problém diskrétního logaritmu bude formulován následovně: jsou dány body grupy P a Q, a je třeba určit takové číslo, aby platilo $Pk = Q$; číslo k je diskrétní logaritmus Q při základu P a platí: $0 \leq k \leq p - 1$.

Pokud je grupa eliptické křivky popsána aditivní notací, pak problém diskrétního logaritmu eliptické křivky je formulován takto: Jsou dány body grupy P a Q, a je třeba nalézt takové číslo k, aby platilo $Pk = Q$.

Příklad:

V grupě eliptické křivky definované rovnicí $y^2 = x^3 + x + 1$ nad F(23) se má určit diskrétní logaritmus k při základu P = (13, 16) pro Q = (5, 4). Jeden (naivní) způsob

stanovení k spočívá v počítání násobku P tak dlouho, až se nalezne vhodné Q. Několik prvních násobků P vypadá takto: P(13, 16), 2P(5, 19), 3P(17, 20), 4P(17, 3), 5P(5, 4). Protože 5P = (5, 4) = Q, tak diskrétním logaritmem Q při základu P je k = 5. V reálných aplikacích musí být *k* hodně velké, aby jeho hodnotu nebylo možné snadno získat.

V reálných kryptosystémech je *Q* veřejný klíč a *k* soukromý klíč. Křivka E a bod P jsou zveřejněny. Velikost *k,* tedy musí zabránit výpočtu diskrétního logaritmu postupným násobením.

3.5 Výpočet inverzního prvku

Multiplikativní inverze čísla *x* na tělese *Fp* (*p* je prvočíslo) je takové číslo x^{-1}, pro které platí, že $x \cdot x^{-1} \equiv 1$. Inverze může existovat i pro čísla na F_m (m je přirozené číslo), ale není to zaručeno. Pro zjištění hodnoty inverzního prvku se využívá dvou postupů. Prvním je hádání, případně zkoušení všech ostatních čísel. Tento postup je efektivní pro malá *p*, když je hádajícím člověk (výsledek je často vidět). Druhým postupem je rozšířený Euklidův algoritmus [1].

Euklidův algoritmus (v základní podobě) se používá pro výpočet největšího společného dělitele. Největší společný dělitel $k = \gcd(p, x)$ je největší celé číslo, které bezezbytku dělí *p* i *x*.

Aby existoval inverzní prvek x^{-1}: $x \cdot x^{-1} \bmod p = 1$, musí platit:

- $\gcd(p, x) = k = 1$, tzn. *p* a *x* musí být tzv. nesoudělná.

Postup hledání inverzního prvku x^{-1}:

1. vstupní proměnné jsou *x* a *p*,
2. provedeme Euklidův algoritmus k určení $\gcd(p, x) = 1$,
3. pomocí proměnných z tohoto algoritmu vyjádříme $\gcd(p, x)$ jako lineární kombinaci *p* a *x*: $\gcd(p, x) = 1 = a \cdot p + b \cdot x$,
4. inverzní prvek $x^{-1} = b \bmod n$.

Použití např. pro výpočet SK z VK u asymetrického kryptosystému RSA nebo při dílčích výpočtech sčítání bodu na eliptické křivce.

4 ELIPTICKÉ KŘIVKY

Při zamýšlení se nad pojmem „eliptická křivka“ určitě každého napadne, co má elipsa společného se šifrováním. Samotná elipsa mnoho ne, ale od této není daleko k eliptickým integrálům, funkcím a křivkám [8].

Matematici hledali jiné a nové cesty, nové algoritmy pro asymetrické kryptosystémy. Kryptografie na bázi eliptických křivek je moderní směr, který v řadě ukazatelů přináší lepší výsledky než nejrozšířenější používané kryptosystémy. Užití eliptických křivek pro návrh asymetrických kryptosystémů poprvé navrhli v r. 1985 nezávisle na sobě Victor Miller a Neal Koblitz. Jedná se vlastně o analogii již existujících systému s veřejným klíčem, kdy je modulární aritmetika nahrazena aritmetikou budovanou na základě operací s body na eliptické křivce. U asymetrických kryptosystémů definovaných nad eliptickou křivkou se hierarchicky volí dva typy algebraických struktur: konečné těleso a eliptická křivka reprezentující grupu bodů, nad níž je vlastní asymetrický algoritmus definován. Volba obou těchto algebraických struktur významně ovlivňuje bezpečnost a efektivitu kryptosystému. Požadavky kladené na tyto dvě struktury spolu vzájemně souvisí.

Obecné vysvětlení

Nyní nahradíme slovo *křivka* pro kryptografii příznačnějším termínem grupa (viz předchozí) a místo sčítání použijeme termín grupová operace [4]. Máme tedy definovanou grupu E. Protože jsme pracovali v rovině, byla souřadnicemi bodů na křivce (x, y) reálná čísla.

Proč ale vůbec potřebujeme nějaké těleso a nezůstaneme u reálných čísel? Výpočty nad množinou reálných čísel jsou pomalé a díky zaokrouhlovacím chybám i nepřesné. Kryptografické aplikace vyžadují rychlé a přesné výpočty, proto se v praxi používají právě grupy na eliptických křivkách nad konečnými tělesy typu Fp a F_2^m. Konečný počet prvků grupy je nutnou vlastností grup používaných v kryptografii. Proto si těleso reálných čísel nahradíme jiným tělesem (F), vhodným pro počítačové zpracování. Pak souřadnice (x, y) bodů na eliptické křivce E budou prvky tohoto tělesa F, a nikoli reálná čísla. Řečeno odborně, dostali jsme tak eliptickou křivku nad tělesem F.

Začněme s definicí eliptické křivky

Nechť F je těleso. Například F může být konečné těleso Fp, kde p je velké prvočíslo, pole R reálných čísel, pole Q racionálních čísel nebo C – komplexních čísel. Eliptická křivka nad tělesem F je definována pomocí Weierstrassovy rovnice:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \tag{4.1}$$

kde $a_1, a_3, a_4, a_6 \in F$.

Eliptická křivka E nad F je značena $E(F)$. Počet bodů na E (mohutnost) je značen $\#E(F)$, nebo jen $\#E$. Pro různá tělesa může být Weierstrassova rovnice transformována do různých tvarů [10].

4.1.1 Eliptická křivka nad množinou reálných čísel

Eliptická křivka nad množinou reálných čísel R je definována jako množina bodů $P = (x, y)$, kde x a y jsou reálná čísla splňující rovnici (4.2) společně s bodem v nekonečnu O. Koeficienty a, b jsou prvky určující eliptickou křivku a musí splňovat podmínku, viz rov 4.3 [14].

$$y^2 = x^3 + ax + b, \tag{4.2}$$

$$4a^3 + 27b^2 \neq 0. \tag{4.3}$$

Takto definované eliptické křivky nejsou vhodné do kryptografické praxe z důvodu nepřesností při zaokrouhlování. Naopak jsou však vhodné pro snazší pochopení dané problematiky.

4.1.2 Eliptická křivka nad tělesem *Fp*

Eliptická křivka E nad tělesem Fp je definována jako bod v nekonečnu O společně s množinou bodů $P = (x, y)$, kde x a y jsou z tělesa Fp a splňují rovnici (4.4). Koeficienty a, b v rovnici jsou také prvky tělesa Fp a musí splňovat podmínku, viz rov. 4.5.

$$y^2 \bmod p = (x^3 + ax + b) \bmod p, \tag{4.4}$$

$$4a^3 + 27b^2 (\bmod p) \neq 0. \tag{4.5}$$

Tato podmínka zaručuje, že takto definovaná množina bodů tvoří grupu (jinak koeficienty a a b můžeme volit libovolně – budou to později veřejné parametry příslušného kryptosystému) [15].

Křivky nad tělesem typu *Fp* jsou spíše vhodné pro softwarovou realizaci. Používané hodnoty p jsou:

- $p = 2^{192} - 2^{64} - 1$
- $p = 2^{224} - 2^{96} - 1$
- $p = 2^{256} - 2^{224} + 2^{192} + 2^{96} - 1$
- $p = 2^{384} - 2^{128} - 2^{96} - 2^{32} - 1$
- $p = 2^{521} - 1$

4.1.3 Eliptická křivka nad tělesem $F(_2^m)$

Prvky tělesa F_2^m jsou m-bitové posloupnosti. Aritmetické operace v tělese F_2^m lze definovat buď v reprezentaci polynomiální, nebo v reprezentaci optimální normální báze (ONB). Polynomiální báze není tak efektivní jako ONB, ale snazší pro pochopení principů. Hlavní rozdíl spočívá v definici násobení mezi prvky pole. Protože prvky F_2^m jsou bitové posloupnosti, tak počítačová realizace aritmetických operací je velmi efektivní.

Eliptická křivka s podložním polem (tělesem) F_2^m je určena výběrem prvků *a* a *b* z F_2^m. Jedinou omezující podmínkou je, že musí platit nerovnost, viz rov. 4.6. Důsledkem charakteristiky pole F_2^m s číslem 2 je úprava rovnice pro binární reprezentaci, viz rov. 4.7.

$$b \neq 0, \tag{4.6}$$

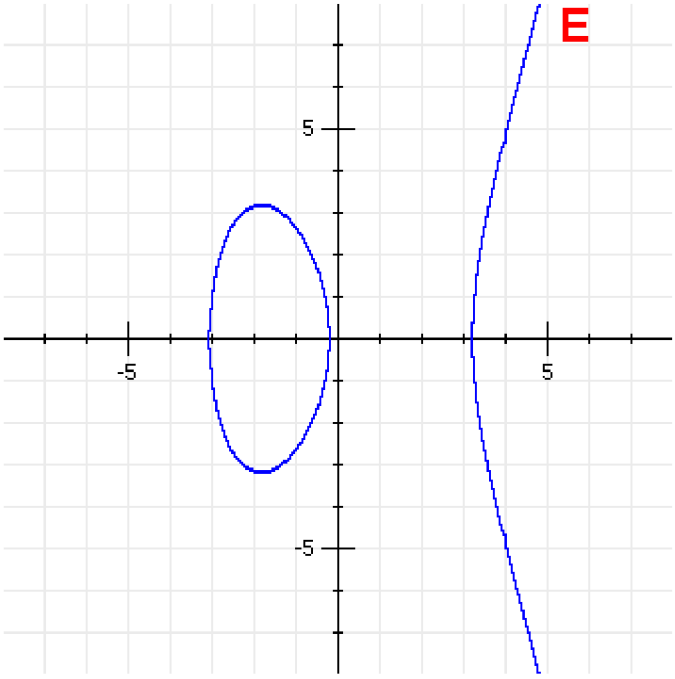
$$y^2 + xy = x^3 + ax^2 + b. \tag{4.7}$$

Eliptická křivka tedy zahrnuje všechny body (x, y), které vyhovují rovnici eliptické křivky nad tělesem F_2^m (x a y jsou prvky z F_2^m). Grupa eliptické křivky nad tělesem F_2^m zahrnuje body odpovídající eliptické křivky společně s bodem v nekonečnu O [9][17]. Takováto eliptická křivka má velký počet bodů, avšak počet konečný. Vlastnosti spojené s binární povahou pole vedou k tomu, že početní operace mohou být velmi efektivně implementovány hardwarově. Používané hodnoty *m* pak jsou:

- m = 163
- m = 233
- m = 283
- m = 409
- m = 571

4.2 Geometrický pohled na součet bodů na EC

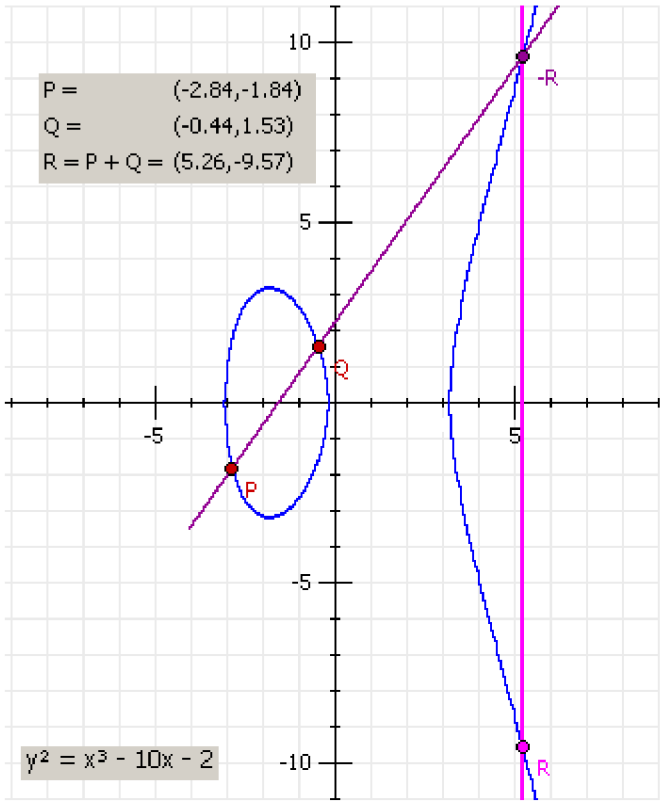
Ke snadnějšímu pochopení operací sčítání bodů na eliptické křivce nám pomůže grafické znázornění.



Obr. 4.1: Ukázka tvaru eliptické křivky

Reálná eliptická křivka, viz obr. **4.1**, je dána obecnou rovnicí (4.2). Musíme si zvyknout nato, že v algebře se dají například sčítat dva body na rovinné křivce. EC na obrázku 4.1, je dána rovnicí $y^2 = x^3 - 10x - 2$. Je to množina bodů (x, y) v rovině, jejichž souřadnice uvedenému vztahu vyhovují.

Křivka je složena ze dvou oddělených částí. Nyní vezmeme dva různé body P a Q, jak je vidět na obrázku **4.2**, které leží na křivce, a definujeme součet těchto bodů. Výsledkem bude další bod ležící na křivce E [8].



Obr. 4.2: Sčítání bodů *P* + *Q* = *R* křivky E

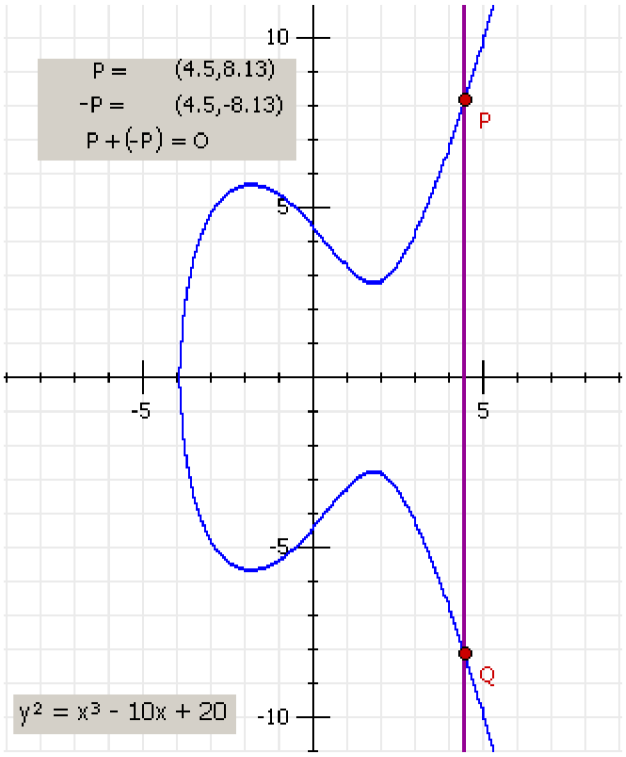
Postup jak tohoto dosáhnout je následující: body P (x_P, y_P) a Q (x_Q, y_Q) propojíme přímkou, která protne křivku v dalším bodě, jejž označíme $-R$ a výsledkem sčítání je bod R, symetrický k $-R$ podle osy x. Body symetrické podle osy x nazýváme opačné [5].

Mohou nastat tyto další případy vzájemného vztahu bodů P a Q.

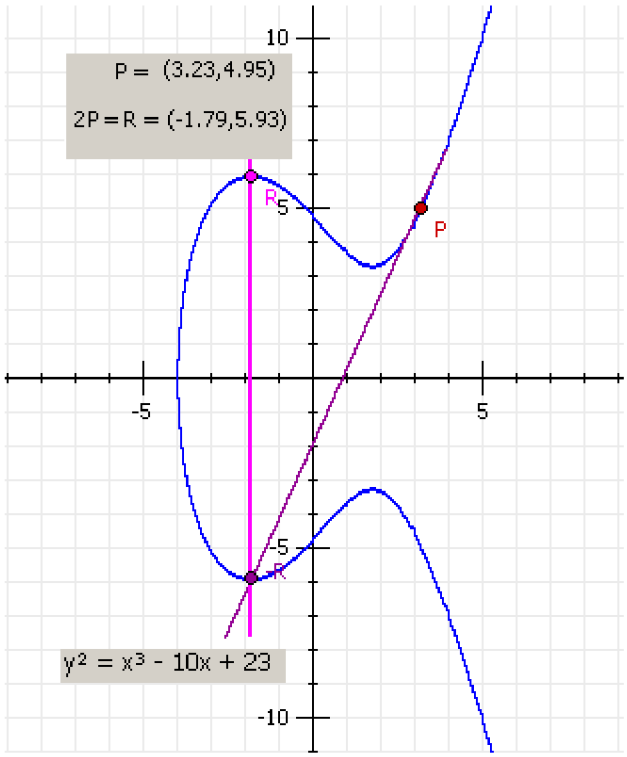
1. V případě $P = Q$ přechází jejich spojnice v tečnu ke křivce E.
2. Sčítáním bodů opačných $P = -Q$ bychom měli dostat 0 - „nulový bod“. Spojnice takových bodů ovšem křivku E už v žádném bodě neprotne, teoreticky v nekonečnu. Matematici bod O, kvůli jednotné funkci, v nekonečnu ke křivce E přidali a sčítání dodefinovali i pro body opačné $P + (-P) = O$, viz obr. **4.3**. Pro nulový bod (oficiálně bod v nekonečnu) křivky E musíme dodefinovat operaci pro sčítání. Proto pro každý bod P na křivce definujeme $P + O = -P$ a také $O + O = O$, kde $O = -O$. Tím je definováno sčítání pro všechny dvojice bodů na křivce E i bod O [5].

3. Pro přičtení bodu k sobě samému $P + P$, je vedena tečna ke křivce E z bodu P. Jestliže $y_P \neq 0$, protne tečna křivku v bodě $-R$. Výsledkem sčítání $P + P = 2P$ je bod R, symetrický k $-R$ podle osy x [5], viz obr. **4.4**.

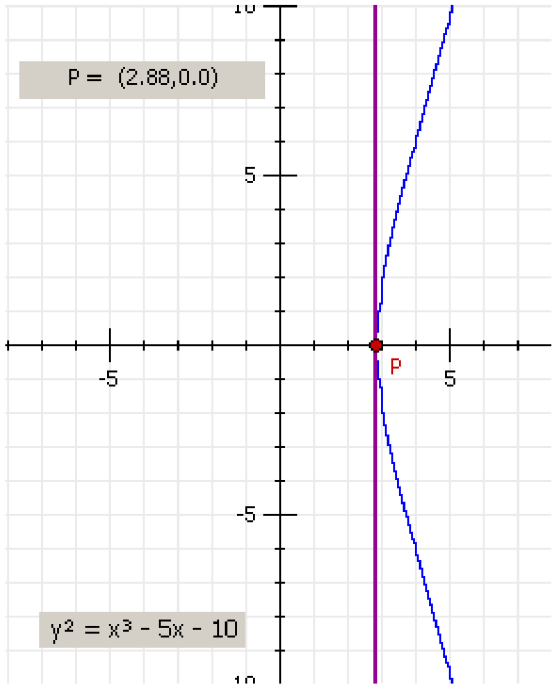
4. Když pro bod P platí $y_P = 0$, pak tečna v bodě P je vždy kolmá k ose x a neprotne EC v žádném dalším bodě, viz obr. **4.5**. Pro takový bod P podle definice platí 2P = 0. Kdybychom chtěli nalézt 3P, stačí určit P + 2P, resp. P + 0. Ale P + 0 = P, takže 3P = P. Dále 4P = 0, 5P = P, 6P = 0 atd.



Obr. 4.3: Sčítání opačných bodů P + (−P) = O



Obr. 4.4: Přičítání bodu k sobě samému P + P = 2P, pokud je $y_P \neq 0$



Obr. 4.5: Přičítání bodu k sobě samému P + P = 2P = O, pokud je yP = 0

4.3 Algebraický pohled na sčítání bodů EC nad R

I když geometrický popis výborně ilustruje aritmetiku EC, tak není vhodný k provádění aritmetických operací. Pro účinnou početní realizaci geometrické aritmetiky existují vhodné algebraické výrazy.

Součet bodů P a Q

Opačný bod k bodu P = (x_P, y_P) je bod –P = (x_P, $-y_P$). Nechť P a Q jsou takové, že $P \neq Q$, $P + Q = R$, kde:

s je směrnice přímky, která body P a Q spojuje je rovna:

$$s = \frac{\left(y_Q - y_P\right)}{\left(x_Q - x_P\right)}, \tag{4.8}$$

a souřadnice bodu $R = (x_R, y_R)$ lze poté odvodit z rovnice křivky jako:

$$x_R = \left(s^2 - x_P - x_Q\right), \tag{4.9}$$

$$y_R = s\left(x_P - x_R\right) - y_P. \tag{4.10}$$

V případě $P = Q$ přechází jejich spojnice v tečnu ke křivce E a její směrnice je rovna:

$$s = \frac{\left(3x_P^2 + \mathrm{a}\right)}{2y_P}. \tag{4.11}$$

4.4 Algebraický pohled na sčítání bodů EC nad *Fp*

Eliptickou křivku nad tělesem *Fp* je možné definovat jako množinu bodů (x, y) vyhovujících rovnici (4.4).

Součet bodů P a Q

Opačný bod k bodu $P = (x_P, y_P)$ je bod $-P = (x_P, y_P \bmod p)$ [3]. Nechť P a Q jsou takové, že $P \neq Q$, $P + Q = R$, kde:

$$s = \frac{y_P - y_Q}{x_P - x_Q} \bmod p, \tag{4.12}$$

$$x_R = s^2 - x_P - x_Q \bmod p, \tag{4.13}$$

$$y_R = -y_P + s(x_P - x_R) \bmod p, \tag{4.14}$$

kde s je směrnice přímky protínající body P a Q a x_R a y_R jsou souřadnice výsledného bodu R.

Zdvojení bodu P

Když $y_P \neq 0$, tak P + P = 2P = R a platí:

$$s = \frac{3x_P^2 + a}{2y_P} \bmod p\,, \tag{4.15}$$

$$x_R = s^2 - 2x_P \bmod p\,, \tag{4.16}$$

$$y_R = -y_P + s(x_P - x_R) \bmod p\,, \tag{4.17}$$

kde a je jeden z parametrů určujících eliptickou křivku.

4.5 Algebraický pohled na sčítání bodů EC nad $F(_2^m)$

Grupy nad F_2^m mají konečný počet bodů a jejich aritmetika nemá žádnou zaokrouhlovací chybu. Tato skutečnost a binární povaha pole umožňuje velmi účinnou počítačovou realizaci aritmetiky F_2^m. Pro aritmetiku nad F_2^m se používají následující algebraická pravidla.

Součet různých bodů P a Q

Opačný bod k bodu P = (x_p, y_p) je bod –P = (x_p, x_p + y_p) [17]. Jestliže P a Q jsou dva různé body, pro které platí P ≠ −Q, pak P + Q = R a platí:

$$s = \frac{y_P - y_Q}{x_P + x_Q}, \tag{4.18}$$

$$x_R = s^2 + s + x_P + x_Q + a\,, \tag{4.19}$$

$$y_R = s(x_P + x_R) + x_R + y_P\,. \tag{4.20}$$

Zdvojení bodu P

Jestliže $x_p = 0$, pak 2P = O. Za předpokladu $x_p \neq 0$ platí 2P = R.

$$s = \frac{x_P + y_P}{x_P}, \tag{4.21}$$

$$x_R = s^2 + s + a\,, \tag{4.22}$$

$$y_R = x_P^2 + (s+1).x_R \,. \tag{4.23}$$

4.6 Shrnutí teoretických poznatků v příkladu

Parametry eliptické křivky jsou:

a = 1, b = 1, p = 23,

zapsáno jako $E_{23}(1,1)$

Aby bylo možné na eliptické křivce nad *Fp* vytvořit grupu, je nutné aby člen $x^3 + ax + b$ byl nerozložitelný nebo, což je ekvivalentní, aby $(4a^3 + 27b^2) \bmod p \neq 0$. V našem případě:

$$4a^3 + 27b^2 (\bmod p) \neq 0$$

$$4.1^3 + 27.1^2 (\bmod 23) \neq 0$$

$$31 \bmod 23 \neq 0$$

$$8 \neq 0$$

Protože je splněna tato podmínka, existuje křivka E: $y^2 = x^3 + x + 1$ nad $F(_{23})$. Její body jsou vypsány v tabulce **4.1** a graficky znázorněny na obrázku **4.6**. Graf není spojitý, protože křivku nad *Fp* tvoří konečný počet bodů [9]. Můžeme si například ověřit, že bod (13, 16) patří této křivce - platí totiž:

$$y^2 \bmod p = (x^3 + ax + b) \bmod p$$

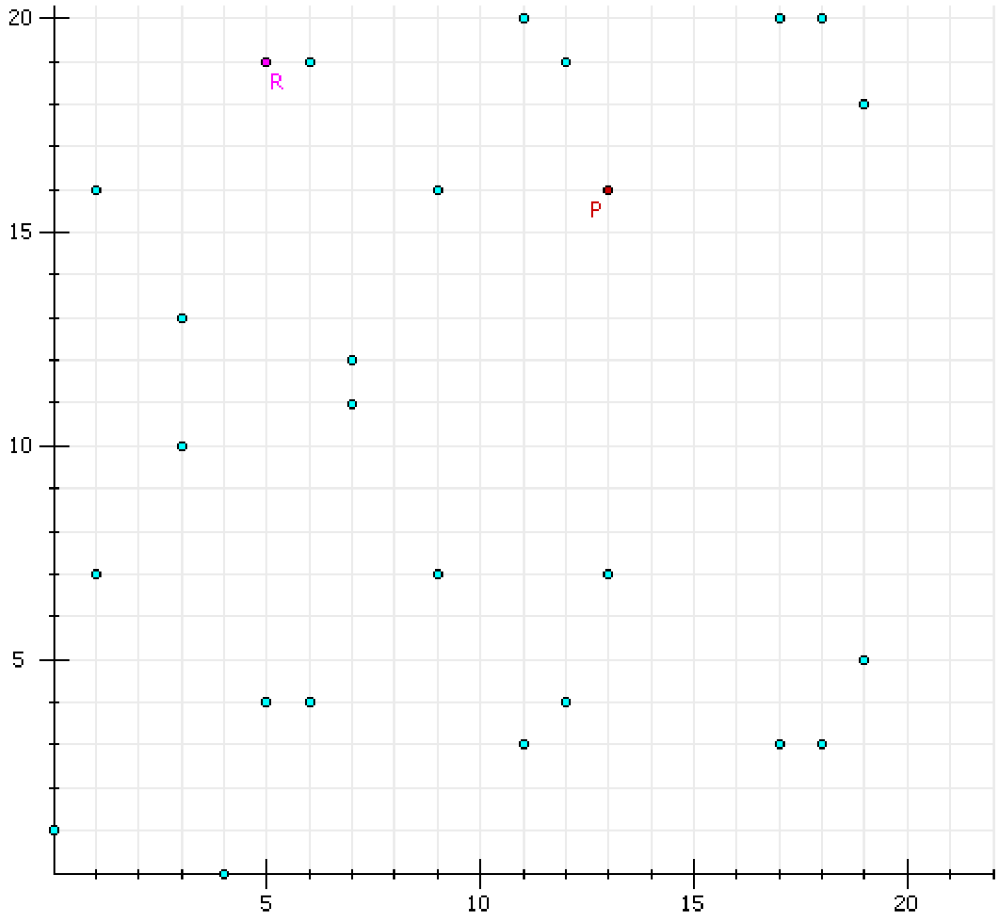
$$16^2 \bmod 23 = (13^3 + 1.13 + 13) \bmod 23$$

$$256 \bmod 23 = 2211 \bmod 23$$

$$3 = 3$$

Tab. 4.1: Body eliptické křivky $E_{23}(1,1)$

O	(4,0)	(9,7)	(13,16)
(0,1)	(5,4)	(9,16)	(17,3)
(0,22)	(5,19)	(11,3)	(17,20)
(1,7)	(6,4)	(11,20)	(18,3)
(1,16)	(6,19)	(12,4)	(18,20)
(3,10)	(7,11)	(12,19)	(19,5)
(3,13)	(7,12)	(13,7)	(19,18)



Obr. 4.6: Body křivky E: $y^2 = x^3 + x + 1$ nad $F_{(23)}$

Součet dvou různých bodu P a Q

Součet provedeme např. pro body P = (13, 16) a Q = (5,19). Budeme postupovat podle výše uvedených vzorců.

Výpočet směrnice přímky:

$$s = \frac{y_P - y_Q}{x_P - x_Q} \bmod p$$

$$= \frac{16-19}{13-5} \bmod 23$$

$$= \frac{-3}{8} \bmod 23$$

$$* = -3.\frac{1}{8} \bmod 23$$

Nyní hledáme inverzní prvek x^{-1}: $x \cdot x^{-1} \bmod p = 1$. V našem jednoduchém příkladu lze použít metodu hádání. Hledáme číslo X, vyhovující rovnici:

$$X.8 \bmod 23 = 1,$$

v našem případě je X = 3, neboť platí:

$$3.8 \bmod 23 = 24 \bmod 23 = 1$$

Z toho vyplývá výsledná směrnice:

$$\underline{* s = -3.3 \bmod 23 = 14}$$

Výpočet kořenů výsledného bodu R:

$$x_R = (s^2 - x_P - x_Q) \bmod p$$

$$x_r = (14^2 - 13 - 5) \bmod 23$$

$$x_r = 178 \bmod 23$$

$$\underline{x_R = 17}$$

$$y_R = -y_P + s(x_P - x_R) \bmod p$$

$$y_r = -¨16 + 14(13 - 17) \bmod 23$$

$$y_r = 72 \bmod 23$$

$$\underline{y_R = 20}$$

Souřadnice výsledného bodu jsou R = (17, 20). Jak je vidět v tabulce **4.1**, výsledný bod jakožto součet dvou bodů patřících eliptické křivce, leží opět na této křivce.

Zdvojení bodu P:

Nyní podle definice sečteme body $P + P$, kde P = (13, 16). Máme $R = P + P = (x_R, y_R)$, kde $s = (3.13^2 + 1)/(2.16) = 508/32 = 508.18 = 9144 = 13$**.** Dále dostáváme $x_R = 13^2 - 13 - 13 = 143 = 5$, $y_R = 13*(13 - 5) - 16 = 88 = 19$, tj. R = (5, 19), viz obr. **4.6**. Výsledek krátce označíme $2P$, což je symbolický zápis pro $P + P$.

Eliptická křivka nad tělesem F(23) na obrázku 4.6 nám naši známou reálnou křivku už geometricky příliš nepřipomíná. Určitě je však zřejmá symetrie kolem pomyslné osy $y = p*1/2 = 11{,}5$. To je zákonité, protože na křivce leží vždy jak bod (x, y), tak bod k němu opačný, tj. $(x, -y)$, což je $(x, p-y)$, tedy právě body symetrické podle uvedené osy $y = p*1/2$.

Převedení křivky nad prvočíselné těleso zkrátka boří jednoduché vztahy, což je vlastně celý záměr kryptografie eliptických křivek, neboť to, co mohlo být na reálné křivce řešitelné jednoduše, bude na diskrétní křivce složité.

4.7 Vlastnosti EC

Primární výhodou kryptosystémů na bázi eliptických křivek je jejich velká kryptografická bezpečnost vzhledem k dané velikosti klíče. Význačně kratší délka klíčů (např. oproti RSA) vede ke kratším certifikátům i menším parametrům systému a tedy i k větší výpočetní efektivnosti algoritmů [14].

Druhá výhoda je v tom, že fakticky všechna již známá použití v systémech na bázi diskrétního logaritmu (kryptografické protokoly, ElGamalův podpis atd.) lze převést do systémů na bázi eliptických křivek. Dá se říci, že existuje nekonečné množství konečných komutativních grup.

EC se používají k ustavení klíče nebo k podpisu. K šifrování zpráv se příliš nepoužívají, protože pro některé x-ové souřadnice body eliptické křivky neexistují [12].

4.8 Bezpečnost EC

V tabulce 4.2 vidíme porovnání bezpečnosti symetrických systémů, u nichž se předpokládá útok hrubou silou a bezpečnosti eliptických křivek, kde se uvažuje složitost řešení problému diskrétního logaritmu pomocí Pollardovy ρ-metody. Tuto tabulku zpracoval NIST jako doporučení pro federální použití v USA. V návrhu dokumentu „Příručka klíčového hospodářství" z 3. 7. 2002 pak NIST zpřesňuje délky u ECC tak, že uvádí logičtější požadavek na „velikost řádu generujícího bodu".

Pro podporu vývoje kryptografie nad eliptickými křivkami společnost CERTICOM sponzoruje prolomení eliptických křivek s různými délkami klíčů [12]. Do této soutěže se může přihlásit každý, kdo má k dispozici značně velké množství přebytečného výpočetního výkonu. Pro představu k prolomení křivky se 109-bitovými parametry nad prvočíselným tělesem (v tabulce **4.3** uvedená jako ECCp-109) potřeboval Chris Monico a jeho tým matematiků z Notre Dame v roce 2002 - 10 000 počítačů, které byly v provozu nepřetržitě 549 dní. V tabulce 4.3 je přehled křivek, které již byly prolomeny a které ještě odolávají a příslušná odměna, která čeká na jejich přemožitele [5].

Bezpečnost eliptických kryptosystémů souvisí s možnostmi řešení úlohy diskrétního logaritmu (DL). Zde existuje několik algoritmů, jejichž výpočetní složitost

lze popsat ve tvaru druhé odmocniny z *N*, kde *N* je počet bodů příslušné eliptické křivky. Jsou to zejména Pollardova ρ -metoda a Pollardova λ-metoda. Složitost z nich nejefektivnější Pollardovy ρ-metody je daná výrazem $(\pi.N/4)^{1/2}$ [10]

Tab. 4.2: Doporučené délky klíče podle NIST

Symetr. šifra délka klíče (bity)	RSA – velikost modulu (bity)	ECC nad *Fp* - velikost p (bity)	ECC nad F_2^m - číslo m (bity)	Řád generujícího bodu ECC (bity)
80	1024	192	163	160
112	2048	224	233	224
128	3072	256	283	256
192	7680	384	409	384
256	15360	521	571	512

Obecně pro řešení úlohy eliptického diskrétního logaritmu není znám žádný algoritmus mající subexponenciální výpočetní složitost jako je tomu pro řešení úlohy faktorizace (RSA) nebo řešení úlohy klasického diskrétního logaritmu. Jsou však určité speciální případy (speciální typy eliptických křivek), kde takovéto postupy existují. Např. v roce 1991 pánové Alfred Menezes, Tatsuaki Okamoto, and Scott Vanstone přišli se subexponenciálním algoritmem pro tzv. supersingulární eliptické křivky (MOV útok) a v roce 1997 byl nalezen algoritmus s lineární výpočetní složitostí pro eliptické křivky. Podobné útoky jsou stále předmětem úvah odborníků, jsou však obvykle orientovány na speciální případy eliptických křivek [3].

Z tohoto důvodu je také v současnosti doporučováno používat eliptické křivky s náhodně generovanými parametry, kde pravděpodobnost existence podobných útoků je minimální. Zejména atraktivním postupem je možnost generovat parametry eliptické křivky tzv. prokazatelně náhodně. Konstruktér eliptické křivky se takto může vyhnout potenciálním obviněním ze strany uživatele, že vložil do hodnot parametrů určitá zadní vrátka, která mu umožňují proniknout snadněji za bezpečnostní hranice systému.

Tab. 4.3: Přehled eliptických křivek na *Fp*

Elliptic Curve Challenges over *Fp*				
Curve	**Field size (in bits)**	**Estimated number of machine days**	**Prize (US$)**	**Status**
ECCp-79	79	146	Handbook of Applied Cryptography & Maple V software	SOLVED Dec. 1997
ECCp-89	89	4360	Handbook of Applied Cryptography & Maple V software	SOLVED Jan. 1998
ECCp-97	97	71982	$ 5,000	SOLVED March 1998
ECCp-109	109	9×10^7	$ 10,000	SOLVED Nov. 2002
ECCp-131	131	2.3×10^{10}	$ 20,000	
ECCp-163	163	2.3×10^{15}	$ 30,000	
ECCp-191	191	4.8×10^{19}	$ 40,000	
ECCp-239	239	1.4×10^{27}	$ 50,000	
ECC2p-359	359	3.7×10^{45}	$ 100,000	

4.9 Standardy

Doporučení pro výběr parametrů kryptografických systémů obsahuje celá řada standardů (SEC, FIPS 186-2, ANSI X9.62, ISO 15946, IEEE P1363), vzájemně jsou však mezi nimi odchylky. Často citovanou normou pro digitální podpis je FIPS 186-2, která zrovnoprávňuje podpis na bázi RSA, DSA i ECDSA. ECDSA, vychází z normy ANSI X9.62. Ta, stejně jako FIPS 186-2, pak těží z práce skupiny P1363 organizace IEEE, která definuje řadu asymetrických algoritmů, včetně těch na bázi eliptických křivek. ECC se zabývá i ANSI norma X9.63. Další skupinu tvoří různé normy ISO používající ECC: například ISO 14888-3 definuje digitální podpis, ISO/IEC 15946 definuje podpisy, šifrování a výměnu klíče, ISO/IEC 9798-3 autentizaci a ISO/IEC 11770-3 klíčové hospodářství. Dále jsou k dispozici různé internetové standardy IETF, využívající eliptické křivky pro internetové použití, standardy WAP fóra pro bezdrátové komunikace, zejména mobilní telefony (např. Wireless Transport Layer Security) [12].

Norma P1363

Tento materiál je základem při budování všech novějších norem pro kryptosystémy s veřejným klíčem. Byl zpracováván širokým kolektivem světových odborníků po celou řadu let a v roce 2000 byl schválen v podobě normy. Jeho součástí je rovněž široký popis aparátu eliptických kryptosystémů. V hlavním dokumentu jsou nejprve uvedeny následující základní pojmy:

- Kryptografická rodina (cryptographic family): v normě jsou prezentovány tři základní rodiny kryptografických zobrazení, která jsou založena na následujících (matematicky obtížných) problémech: diskrétní logaritmus v konečném poli (DL), diskrétní logaritmus v grupách eliptických křivek (EC) a na faktorizaci celých čísel (IF).
- Parametry systému (domain parameters): informace o matematických objektech (jako tělesa či grupy) v jejichž kontextu existují dvojice veřejný a soukromý klíč. Více dvojic klíčů může mít tytéž systémové parametry.
- Platné parametry systému (valid domain parameters): taková množina systémových parametrů, která splňuje navíc příslušné specifické matematické požadavky týkající se příslušných parametrů systému pro danou rodinu.
- Platný klíč (valid key): klíč (soukromý či veřejný), který splňuje navíc příslušné specifické matematické požadavky klíčů pro danou rodinu.
- Platná dvojice klíčů (valid key pair): dvojice veřejný a soukromý klíč, která splňuje navíc příslušné specifické matematické požadavky, kladené na dvojici klíčů pro danou rodinu.
- Ověření platnosti (validation): proces, během něhož proběhne ověření platnosti klíče, dvojice klíčů či systémových parametrů.

ANSI X9. 62 a ANSI X9.63

Tyto normy jsou určeny zejména pro finanční sféru. Vychází z P1363. Řadu otázek přitom již řeší konkrétněji (zejména např. podpisové schéma ECDSA) a s větším ohledem na praktické realizace. Součástí materiálu jsou i některé vybrané a spočtené hodnoty parametrů.

NIST

Materiál zpracovává doporučenou množinu vybraných parametrů pro eliptické křivky pro užití ve státní a veřejné správě v USA. Často citovanou normou pro digitální podpis je FIPS 186-2, která zrovnoprávňuje podpis na bázi RSA, DSA i ECDSA. Poslední zmíněný vychází z normy ANSI X9.62. Další skupinu tvoří různé normy ISO používající ECC: například ISO 14888-3 definuje digitální podpis, ISO/IEC 15946 definuje podpisy, šifrování a výměnu klíče, ISO/IEC 9798-3 autentizaci a ISO/IEC 11770-3 klíčové hospodářství.

5 KRYPTOSYSTÉMY PRO ŠIFROVÁNÍ A PODEPISOVÁNÍ

Řád křivky

Jak jsme zjistili dříve v kryptografii se používají eliptické křivky nad konečnými tělesy, která lze algebraicky klasifikovat a každé konečné těleso je pak jednoznačně určeno řádem. Řádem křivky nazýváme počet bodů na křivce. Značí se *#E*. Pro předchozí příklad je řád křivky *#E* = 28.

Schoofův algoritmus

Pro konstrukci eliptických křivek je nezbytné pro stanovení konkrétních hodnot parametrů mít k dispozici prostředek pro výpočet počtu bodů dané eliptické křivky. Obecně toto řeší tzv. Schoofův algoritmus.

Hasseho věta nám říká, že počet bodů eliptické křivky Eq je dán výrazem

$$\#E(Fq) = q + 1 - t,$$

kde $|t| \leq 2\sqrt{q}$.

Toto číslo t je však i při zadaných hodnotách základních parametrů eliptické křivky předem neznámé a je třeba výpočtem stanovit jeho hodnotu. Schoofův algoritmus to řeší tak, že počítá t mod L, kde L jsou některá prvočísla (jejich velikost je shora omezená určitou mezí L_{max}). Použitím čínské věty o zbytcích je pak jednoznačně spočtena hodnota t. Samotný algoritmus má výpočetní složitost $O(\log^8 q)$. Jeho implementace však není zdaleka triviální záležitostí. Aby byla dosažena tato (nízká) výpočetní složitost, je třeba použít složitějších technik založených na poměrně hlubokých výsledcích z teorie čísel.

Řád bodu

Řekněme, že bod P je řádu n, jestliže $nP = O$. Tato definice má velmi jednoduchou podstatu. Pokud máme bod P, vypočítáme postupně $2P$, $3P$, $4P$ atd., čímž dostáváme obecně různé body xP na křivce. Protože křivka má konečný počet bodů, po určitém počtu kroků m se musí tato posloupnost zacyklit. V bodě zacyklení mP tak platí $mP = xP$, kde xP je nějaký dřívější bod. Odtud dostáváme $mP - xP = (m - x)P = O$. Čili existuje nějaké n (kde $m - x < m$) takové, že $nP = O$.

Je tedy jasné, že v posloupnosti P, $2P$, $3P$, $4P$ atd. se vždy nakonec dostaneme do bodu O a poté cyklus začíná znovu od bodu P, neboť platí $(n + 1)P = nP + P = O + P$

= *P*. Nejmenší takové *n*, pro něž je *nP* = *O* nazýváme řád bodu *P*. Pokud použijeme bod z předchozího příkladu P = (13, 16), tak zjistíme, že řád tohoto bodu je n = 7.

Různé body na křivce E mohou mít různý řád. V kryptografické praxi vybíráme takový bod, jehož řád je roven největšímu prvočíslu v rozkladu řádu křivky nebo jeho násobku a tento bod nazýváme kofaktor *h*, jež spočítáme jako *h* = *#E* / *n*, kde *#E* je řád křivky a *n* je řád bodu [7].

U bodu řádu *n* máme tedy zaručeno, že v posloupnosti *P*, 2*P*, 3*P* atd. dojde k zacyklení až po *n*-tém kroku. Pokud je *n* velké, například řádově 2^{256} (musí platit $n > 2^{160}$ a $n > 4\sqrt{p}$), je to opravdu velmi dlouhá posloupnost. Právě při šifrování a elektronickém podepisování takovéto velké posloupnosti využíváme, a to v souvislosti s tzv. problémem diskrétního logaritmu. Zvolíme tajné číslo *k* (je to náš privátní klíč) a vypočteme bod *Q* = *kP*. Body *P* a *Q* můžeme nyní zveřejnit – budou součástí našeho veřejného klíče.

5.1 Převod dat

Pokud je k šifrování použito eliptických křivek, jedním z prvních kroků musí být převod dat, která mají být šifrována, na body eliptické křivky tak, aby příjemce byl schopen z obdržených bodů zpětně zprávu rekonstruovat. Není to proces jednoduchý. To dokládá i fakt, že zatím nebyl nalezen deterministický algoritmus, který by převod realizoval v polynomiálním čase.

Obecně platí, že je jedno jakým způsobem bude toto provedeno a jak bude "mapování" (převod znaků na body EC) navrženo. Podstatné je to, že obě strany znají stejný postup. Jedná se ve své podstatě pouze o zobrazení z prostoru abecedy zprávy do prostoru bodů křivky a opačně.

5.1.1 Možnosti převodu zprávy na E podle Koblitze

Nalezení kořenů rovnice $y^2 = f(x)$

Převod zprávy na body eliptické křivky neznamená jejich šifrování, ale jde o jednoduché zakódování dat na danou eliptickou křivku E definovanou nad konečným tělesem *Fp*. Obecná rovnice eliptické křivky má tvar viz rovnice (4.2). Po převedení pravé strany do jednodušší formy f(x) dostaneme tvar [2]:

$$y^2 = f(x).$$

Klíčovým místem převodu je právě rozhodnutí, zda výše zmíněná rovnice má pro dané x řešení (existuje y) a pokud ano, tak jej nalézt [11].

Mějme p, které je náhodné prvočíslo a f(x) je kvadratické reziduum, potom existuje y takové, že $y^2 \equiv$ f(x) mod p. V konečném tělese Fp, kde p vyhovuje vztahu $p \equiv 3 \pmod 4$, má řešení tvar:

$$y = f(x)^{\frac{(p+1)}{4}} \pmod p .$$

Algoritmus převodu

1. Mějme prvočíslo p, pro které platí $p \equiv 3 \pmod 4$ a eliptickou křivku:

$$y^2 = x^3 + \mathrm{ax} + \mathrm{b} .$$

2. Zvolme číslo K takové, že $1/2K$ vyjadřuje pravděpodobnost chyby při převodu [4]. V praxi se volí K=30 nebo K=50 .
3. Zpráva je z množiny $\{m \in Fp \mid m < (p-K)/K\}$.
4. Od $j = \{0,1,2, \ldots, K-1\}$
   - Nastavíme x_j podle vztahu: $x_j = m.K + j$
   - Spočítáme:

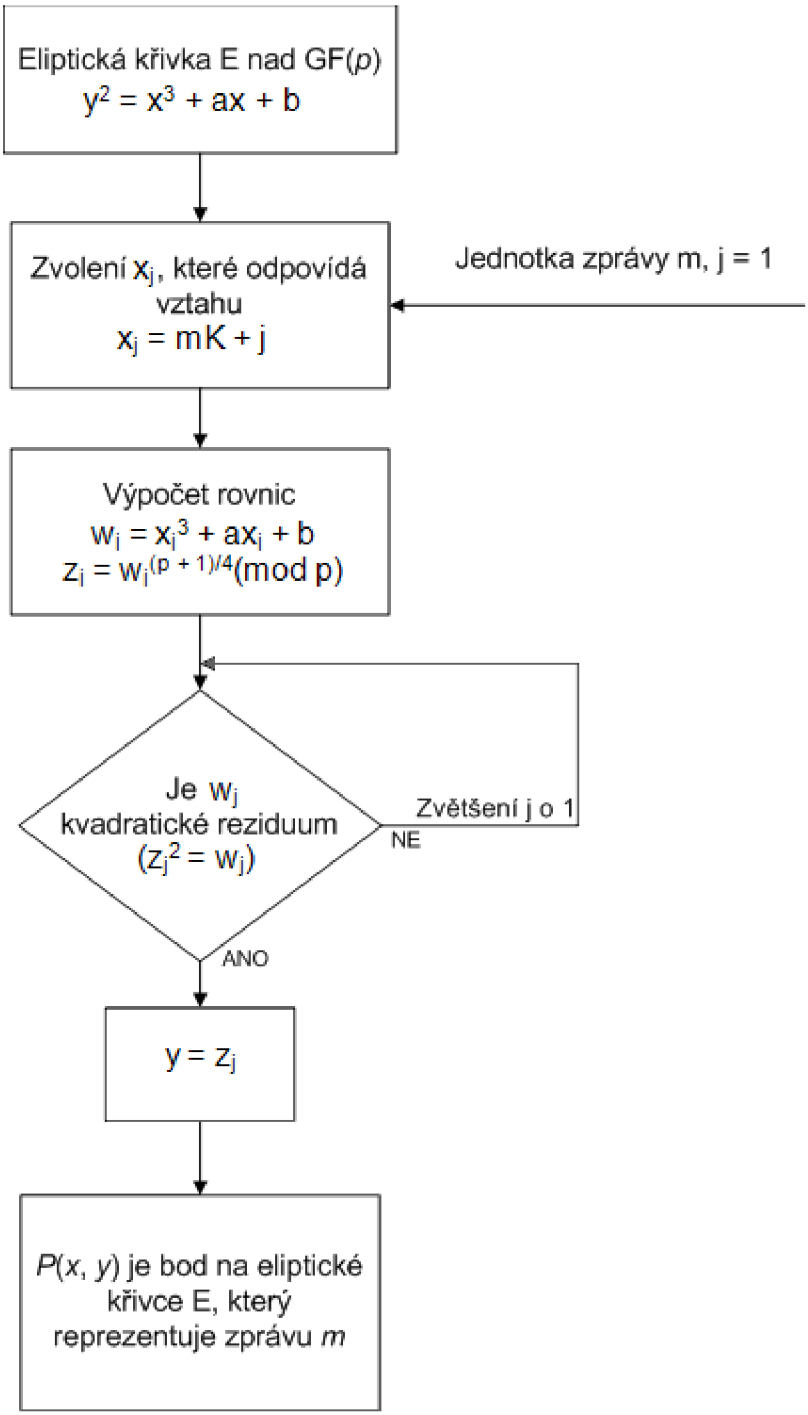
$$w_j = x_j^3 + ax_j + b ,$$

(výpočet rovnice eliptické křivky, pro jednodušší zápis $y^2 = w$) a

$$z_j = w_j^{\frac{(p+1)}{4}} \pmod p ,$$

(výpočet kořene kvadratické rovnice $y^2 =$ f(x)).

   - Pokud $z_j^2 = w_j$ (w_j je kvadratické reziduum mod p), potom (x_j, z_j) je bod na eliptické křivce E, který reprezentuje zprávu m; jinak zvyš j o jednu a výpočet opakuj, viz obr. **5.1**.
   - Pokud j (1,2,…,$K-1$) nevyhovuje, bod reprezentující zprávu m nenalezen. Pravděpodobnost, že se převod nepovede je $\leq 1/2^K$, což pro používaná K je velmi malá.
5. Z bodu (x, y) dekódujeme zprávu m vztahem $m = x/K$.



Obr. 5.1: Algoritmus pro nalezení bodu na křivce E a převedení zprávy *m*

5.2 Šifrování pomocí eliptických křivek

5.2.1 Výměna klíčů pomocí Diffie-Hellmana

Tento algoritmus řeší situaci, kdy si dvě strany, A a B chtějí vyměnit tajnou informaci přes veřejný kanál. Jak je to u všech systémů s veřejným klíčem nutné, i zde se předpokládá, že každá ze stran má k dispozici důvěryhodnou cestou získaný veřejný klíč protistrany [4].

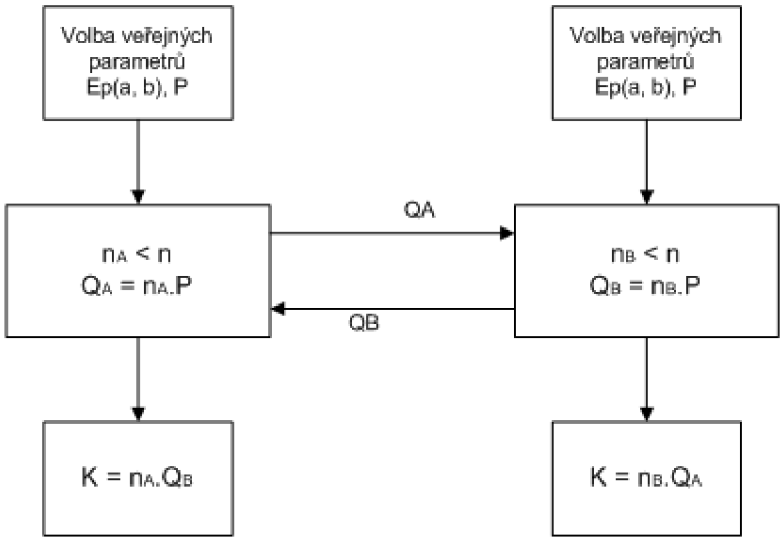
Postup ustanovení společného klíče

1. Volba veřejných parametrů: Ep(a, b) a bodu P, kde p je velké prvočíslo.
2. Generování klíčů účastníky A i B.
   a) Volba soukromého klíče: $n_A < n$, $n_B < n$ (kde n je řád bodu).
   b) Výpočet veřejného klíče: $Q_A = n_A.P$, $Q_B = n_B.P$.
3. Výměna veřejných prvků Q_A a Q_B.
4. Generování tajného klíče K:

$$K = n_A.Q_B,\ K = n_B.Q_A$$

- Pro strany A i B platí, že jsou klíče stejné. To vychází z:

$$K = n_A.Q_B = n_A.n_B.P = n_B.n_A.P = n_B.Q_A = K$$



Obr. 5.2: Algoritmus ustanovení klíče pomocí Diffie-Hellmana

Protože obě strany vycházejí ze stejného bodu P, dospějí zákonitě ke stejnému klíči K. Tento bod ovšem nezná nikdo jiný než ony, v čemž je podstata ustavení společného tajného prvku. Ani v případě, že by se na komunikačním kanálu předávaly hodnoty Q_A a Q_B, není klíč K prozrazen, protože útočník z nich není schopen určit privátní hodnoty n_A a n_B díky složitosti diskrétního logaritmu.

5.2.2 Menezes-Qu-Vanstone (MQV) protokol

MQV je ověřený protokol určený k dohodě na společném klíči založený na Diffie-Helmanově schématu. MQV je začleněn ve standardu IEEE P1363. Tento protokol poskytuje ochranu proti aktivnímu útočníkovi. Je upravený, aby pracoval nad eliptickou křivkou. Tento je známý jako ECMQV. Vlastnostmi tohoto protokolu je rychlost, odolnost a levnost hardwarové implementace [5].

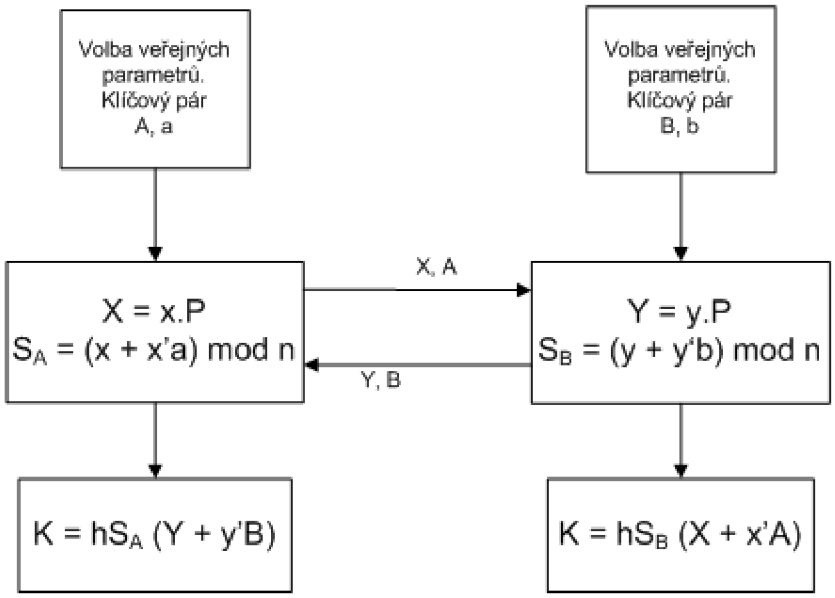
Postup ustanovení společného klíče:

1. Volba veřejných parametrů a bodu P
2. Generování klíčových párů účastníky A i B.
    - Strana A vygeneruje dvojici (A, a), kde A je veřejný klíč a a soukromý klíč.
    - Strana B vygeneruje dvojici (B, b), kde B je veřejný klíč a b soukromý klíč.
3. Obě strany spočítají klíč k sezení:
    - Strana A vypočítá dvojici (X, x) jako: $X = x.P$, kde x je celé číslo.
    - Strana B vypočítá dvojici (Y, y) jako: $Y = y.P$.
4. Obě strany si hodnoty X a Y vymění. Předpokládá se, že obě strany mají veřejné klíče protistrany (A, B).
5. Strana A spočítá S, tzv. implicitní klíč jako $S_A = (x + x'.a) \bmod n$, kde n je další generující bod P
6. Také strana B spočítá S jako $S_B = (y + y'.b) \bmod n$.
7. Obě strany spočítají tajný společný klíč K jako:

$$K = hS_A\,(Y + y'.B) = hS_B\,(X + x'.A),$$

kde h je kofaktor definovaný v P1363.

Pozn. *x'* a *y'* představují prvních *L* bitů z první složky páru *X* a *Y*, kde $L = [(\log 2n + 1) / 2]$



Obr. 5.3: Algoritmus ustanovení klíče pomocí MQV

5.2.3 ECIES (EC integrated encryption scheme)

Integrated Encryption Scheme (IES) je schéma založené na kryptografii veřejného klíče. Je to šifrovací schéma vyžívající eliptické křivky. Strana, která začíná komunikaci pomocí diffie-hellmana získá tajnou informaci, ze které klíč, který pak vygeneruje, je použit k šifrování nebo podepisování. Tento algoritmus byl představen v roce 1998 a je obsažen ve standardech ANSI X9.63, ISO/IEC 15946-3 a IEEE1363 [7].

Jsou standardizovány dvě formy: DLIES (Discrete Logarithm Integrated Encryption Scheme) a ECIES (Elliptic Curve Integrated Encryption Scheme), které je také známo jako Elliptic Curve Augmented Encryption Scheme nebo jednoduše jako Elliptic Curve Encryption Scheme (ECES).

Pokud chce strana A poslat zašifrovanou zprávu straně B pomocí ECIES potřebuje následující informace:

Jsou použity tyto kryptografické sady:

- KDF- Key derivation function (např. ANSI-X9.63-KDF s SHA-1).
- MAC (např. HMAC-SHA-1-160 s 160-bit klíčem nebo HMAC-SHA-1-80 s 160-bit klíčem)

- Symetrické šifrovací schéma E (např. 3-key TDES v CBC modu nebo XOR).
- Parametry křivky jsou: (p, a, b, G, n, h) pro křivku nad *Fp*.
- Veřejný klíč strany B: K_B, kde $K_B = k_B.G$, kde k_B je soukromý klíč, pro který platí, že k_B leží $[1, n - 1]$.
- Volitelně sdíleny informace s_1 a s_2.

Pro zašifrování zprávy *m* musí strana A udělat následující kroky:

- Vygenerovat náhodné číslo $r \in [1, n - 1]$ a vypočítat R, kde $R = rG$.
- Odvodit tajnou informaci: $S = P_x$, kde $P = (P_x, P_y) = rK_B$ a musí platit nerovnost $P \neq 0$.
- Použít KDF k odvození symetrické šifry a MAC klíčů.

$$k_E||k_M - \text{KDF}\ (S||S_1)$$

- Zašifrovat zprávu: $c = \text{E}(m, K_E)$
- Vypočítat značku ze zašifrované zprávy a S_2: $d = \text{MAC}(c||S_2;k_M)$
- Výstupem je $R||c||d$

K dešifrování šifrového textu *R*||*c*||*d* strana B musí udělat následující kroky:

- Odvodit tajnou informaci: $S = P_x$, kde $P = (P_x, P_y) = k_BR$ (je to stejné jako odvodilo A, protože platí $P = k_BR = k_BrG = rk_BG = rK_B$)
- Vypočítat klíč stejně jako strana A:

$$k_E||k_M - \text{KDF}\ (S||S_1)$$

- Použít MAC k zkontrolování značky a pokud výstup d≠ MAC($c||S_2;k_M$), nastala chyba.
- Použít symetrickou šifru k dešifrování zprávy c: $m = \text{D}(c, K_E)$

5.2.4 Šifrování jako přímý převod zprávy *m* na body EC

Právě z důvodu obtížnosti převodu zprávy na body eliptické křivky se eliptické křivky používají spíše na ustavení společného tajného klíče nebo k digitálnímu podpisu. Existuje však i další způsob, který se dá použít k šifrování. Jedná se o princip přímého převodu znaků zprávy *m* na body eliptické křivky - jako posloupnost souřadnic.

- Volí se parametry eliptické křivky $E_p(a, b)$
- Strana A si zvolí náhodné kladné k a zašifruje symbol jako dvojici bodů eliptické křivky:

$$C_m = \{ k.P, P_m + k.Q_B \}$$

- Strana B vynásobí první z dvojice bodů $(k.P)$ svým soukromým klíče n_B a tuto hodnotu odečte od druhého bodu:

$$P_m + k.Q_B - n_B(k.P) = P_m + k.(n_B.P) - n_B.(k.P) = P_m$$

Jelikož pouze A zná hodnotu k, je dešifrování i se známou hodnotou veřejného klíče B obtížné.

5.3 Digitální podpis pomocí eliptických křivek

5.3.1 ECSS (EC signature scheme)

Parametry kryptosoustavy ECSS jsou dány šesticí proměnných D = (q, a, b, G, n, h), kde $q = p$ nebo $q = 2^m$, p je prvočíslo, a i b jsou prvky určující eliptickou křivku. G = (x_P, y_P) je bod na křivce, n je řád bodu G a h je kofaktor ($h = \#E/n$). Zadané parametry jsou veřejné [13].

Vytvoření podpisu

Strana A podepisuje zprávu M pro stranu B, pomocí následujících kroků:

- Převede zprávu M do binární podoby.
- Použije hash algoritmus k spočítání hodnoty $e = H(M)$.
- Zvolí náhodné celé číslo k, kde $1 \leq k \leq n - 1$.
- Spočítá veřejný klíč $R = kP = (x_R, y_R)$
- Spořítá $r = x_R + e \bmod q$
- Použije soukromý klíč k_A k vypočítání $s = k - k_A.r \bmod n$
- Strana A pošle B zprávu M spolu s podpisem r, s.

Ověření podpisu

Strana B ověří podpis strany A (r, s) ze zprávy M pomocí následujících kroků:

- Zjistí veřejný klíč A strany A
- Spočítá bod $V = sP + rA = (x_R, y_R)$

- Spočítá hodnotu hashe $e = \mathrm{H}(M)$
- Vypočítá $r' = x_R + e \bmod q$
- Podpis je v pořádku pouze v případě, že $r = r'$.

5.3.2 Okamotovo schéma digitálního podpisu

Okamoto, Fujioka a Fujisaki roku 1992 predstavili schéma digitálního podpisu založené na eliptických křivkách nad tělesem *Fn*, kde $n = p^2q$ a kde p i q jsou prvočísla [15].

Generování klíče

- Soukromý klíč: velká prvočísla p, q ($p > q$)
- Veřejný klíč: kladné celé číslo $n = p^2q$

Vytvoření podpisu

Podpis s zprávy m je vypočítán odesílatelem:

- Výběr náhodného čísla $t \in \mathrm{F}_{pq}$
- Výpočet s tak, že:

$$w = \left\lceil \frac{h(m) - (t^k \bmod n)}{pq} \right\rceil,$$

$$u = w/(kt^{k-1}) \bmod p,$$

$$s = t + upq,$$

kde h je hashovaní funkce ($\mathrm{h}(m) \in Fn$, pro všechny kladná celá čísla m), k je cele číslé ($4 \leq k$).

Ověření podpisu

Podepsaná zpráva (s, m) je považována platnou, pokud je splněna následující podmínka:

$$h(m) \leq s^k \bmod n < h(m) + 2^{2.\lfloor n \rfloor / 3}$$

5.3.3 ECDSA (EC digital signature algorithm)

Generování klíče

- Vybereme eliptickou křivku E nad *Fp*. Počet bodů křivky #E by měl být dělitelný velkým prvočíslem *n*.
- Zvolíme bod $P \in$ E řádu *n* (poznamenejme, že ANSI X9.62 požaduje, aby řád křivky byl větší než 2^{160}) [14].
- Vybereme jedinečnou a nepredikovatelnou hodnotu privátního klíče, číslo $d \in [1, n-1]$.
- Vypočteme veřejný bod $Q = dP$.
- Veřejný klíč tvoří čtveřice (E, P, n, Q).

Vytvoření podpisu

Mějme zprávu m.

- Vybereme jedinečné a nepredikovatelné číslo $k \in [1, n-1]$.
- Vypočteme bod $kP = (x_1, y_1)$ a číslo $r = x_1 \bmod n$.
- Je-li $r = 0$, pak postup opakujeme od generování čísla *k* (to je nutné proto, aby v hodnotě *s* byl obsažen privátní klíč, viz dále).
- Vypočteme $k^{-1} \bmod n$.
- Vypočteme $s = k^{-1} \{h(m) + dr\} \bmod n$, kde *h* je hashovací funkce SHA-1.
- Je-li $s = 0$, pak opět jdeme na první bod – generování nového *k* (neexistovalo by $s^{-1} \bmod n$, viz dále proces ověření).
- Podpisem zprávy m je dvojice čísel (r, s).

Ověření podpisu

- Mějme zprávu m a její podpis (r, s).
- Důvěryhodným způsobem získáme veřejný klíč podepisujícího (E, P, n, Q).
- Ověříme, že *r* a *s* jsou z intervalu $[1, n-1]$.
- Vypočteme $w = s^{-1} \bmod n$ a $h(m)$.
- Vypočteme $u_1 = h(m)w \bmod n$ a $u_2 = rw \bmod n$.

- Vypočteme $u_1P + u_2Q = (x_0, y_0)$ a $v = x_0 \bmod n$.
- Podpis je platný právě tehdy, když $v = r$.

Uveďme si nyní, jak definuje elektronický podpis pomocí eliptických křivek standard FIPS 186-2, který zmiňuje i naše vyhláška k zákonu o elektronickém podpisu. Standard definuje více křivek, zde si vybereme tu nad tělesem *Fp* s nejmenším prvočíslem *p* (192bitovým). Toto schéma (multiplikativní grupa) se pak transformuje na eliptickou křivku (aditivní grupa) tak, že operace násobení prvků $g * g * g * g *\ldots$. (tj. gk) se převede na sčítání bodů na křivce $P + P + P + P +\ldots$. (tj. kP).

6 PODPORA VÝUKY

6.1 Metodika výuky

Při výuce klademe důraz na správné osvojení dovedností a především schopnost prakticky poznatky využít. Vyučovací jednotka (hodina) je proto konstruována tak, aby si studenti mohli nově nabyté vědomosti ihned prakticky procvičit.

6.1.1 Schéma hodiny

Motivace

Pro efektivní a optimální výuku má jednoznačný význam motivace studentů, zejména vnitřní motivace, jež je trvalejšího charakteru a je určována hodnotami a cíly, které studenti mají. Účelem motivace je připravit studenty na danou problematiku a vzbudit jejich zájem.

Prezentace nové látky

Učitel studenty seznámí se základními pojmy jako jsou (elipsa, eliptická křivka, sčítání bodů na EC, šifrování a elektronický. podpis pomocí EC), uvede je do problematiky a podá komplexní výklad. K co nejlepšímu pochopení využívá powerpointové prezentace (*v našem případě vytvořené www stránky*).

Fixace látky

Procvičení dané problematiky je nezbytné pro efektivní studium. Studenti mohou novou látku ihned prakticky vyzkoušet a projít si na svém počítači příklady vhodné k dané problematice. Tento podpůrný program slouží k lepší fixaci nové látky, neboť vycházíme z ověřeného faktu, že při praktickém procvičení si studenti zapamatují až 90% látky. V závěru cvičení mají studenti připraveny otázky nebo příklady, nad kterými se zamýšlí a tím si opět danou látku procvičují.

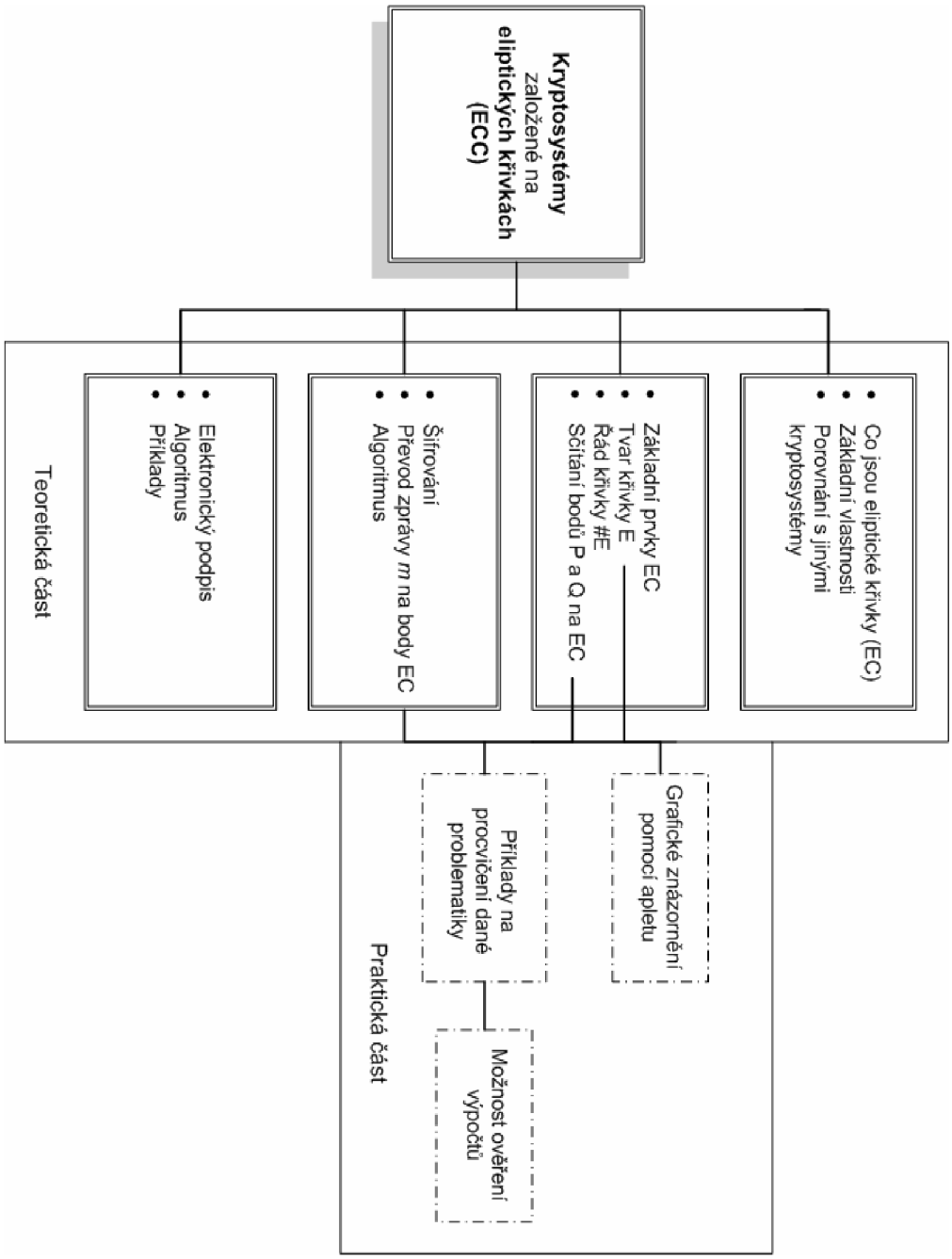
Zhodnocení hodiny

Zhodnocení hodiny má dvě fáze. Zhodnocení hodiny učitelem, kdy studentům dá najevo, jak se mu pracovalo a co ho těšilo, a co naopak ne. A zhodnocení hodiny studenty, kteří mohou takto učiteli naznačit, jakou cestou se má nadále výuka ubírat.

6.2 Obsah výuky

Popis webového rozhraní

Celá praktická část diplomové práce je zaměřena na osvojení si problematiky eliptických křivek. Dalo by se říci, že práce je rozdělena do čtyř částí. Celková struktura je zobrazena na obázku **6.1**.



Obr. 6.1: Grafické znázornění obsahu výuky

Výukový materiál je napsán pomocí xhtml a css stylů z důvodu požadavku na zobrazení na libovolném webovém prohlížeči. Webové stránky jsou rozděleny do hlavních kapitol a tyto pak následně na podkapitoly, na které jsou vytvořeny odkazy. Vzhled stránek je zobrazen na obrázku **6.2**.

Obr. 6.2: Vzhled stránek

První část je zaměřena na teoretický základ problematiky eliptických křivek. Nejprve je popsáno, co vůbec eliptické křivky jsou. Jaké mají vlastnosti, ze kterých vztahů vycházejí a za jakých podmínek můžeme eliptickou křivku sestrojit. Dále je uvedena bezpečnost, standardy a normy, ve kterých jsou eliptické křivky uvedeny a porovnání s jinými asymetrickými kryptosystémy zejména z hlediska použité délky klíčů. Pro názornější ukázku a zapamatování jsou zde uvedeny tabulky s tímto porovnáním, viz obr. **6.3** a **6.4**.

Velká část je věnována sčítání bodů na eliptické křivce, protože pro další práci s nimi je důležité. Graficky znázorněny jsou nad množinou reálných čísel, neboť pro pochopení je to snazší. V kryptografii se však používají křivky nad prvočíselným tělesem nebo nad binárním tělesem, protože se pracuje s konečným počtem bodů. Uvedeny jsou zde algoritmy určené k výměně klíčů, konkrétně ECDH – EC Diffie-Hellman a jednoduchý způsob šifrování, při kterém se znaky zprávy šifrují přímým převodem na body eliptické křivky.

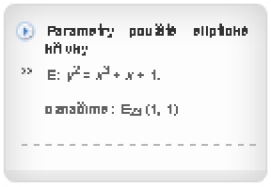
Porovnání s jinými kryptosystémy

- V tab.3 vidíme porovnání bezpečnosti symetrických systémů, u nichž se předpokládá útok hrubou silou (vyzkoušení všech možných klíčů), bezpečnost asymetrického kryptosystému RSA a bezpečnosti eliptických křivek, kde se uvažuje složitost řešení problému diskrétního logaritmu pomocí Pollardovy ρ-metody. Tuto tabulku zpracoval *NIST* jako doporučení pro federální použití v USA.
- ECC s délkou klíče 160 bitů poskytují stejnou bezpečnost jako RSA s 1024 bitovým klíčem.

Symetr. šifra délka klíče (bity)	RSA – velikost modulu (bity)	ECC nad Fp - velikost p (bity)	ECC nad F_{2^m} - číslo m (bity)	Řád generujícího bodu ECC (bity)
80	1024	192	163	160
112	2048	224	233	224
128	3072	256	283	256
192	7680	384	409	384
256	15360	521	571	512

Tab. 3 - Doporučené délky klíče podle NIST

Obr. 6.3: Ukázka části porovnání s jinými kryptosystémy



Pro názornou ukázku a vyzkoušení si závislosti tvaru E na parametrech a, b přejděte sem.

Řád bodu a křivky

- Vezměme si bod $P = (13, 16) \in E$ a vypočítejme postupně 2P, 3P, 4P, 5P, 6P atd. (viz dále), čímž dostáváme obecně různé body na křivce. Protože křivka má konečný počet bodů, po určitém počtu kroků (m) se nám musí tato posloupnost zacyklit. V bodě zacyklení (mP) tak platí $mP = xP$, kde xP je nějaký dřívější bod. Odtud ale dostáváme $mP - xP = (m-x)P = O$, čili existuje nějaké $n \rightarrow [(m - x) < m]$ takové, že $nP = O$ (tab.1), takže je jasné, že v posloupnosti P, $2P$, $3P$, $4P$ atd. se vždy nakonec dostaneme do bodu O a poté cyklus začíná znovu od bodu P, neboť $(n+1)P = nP + P = O + P = P$.

P(13,16)	R = k.P
P	R(13,16)
2P	R(5,19)
3P	R(17,20)
4P	R(17,3)
5P	R(5,4)
6P	R(13,7)

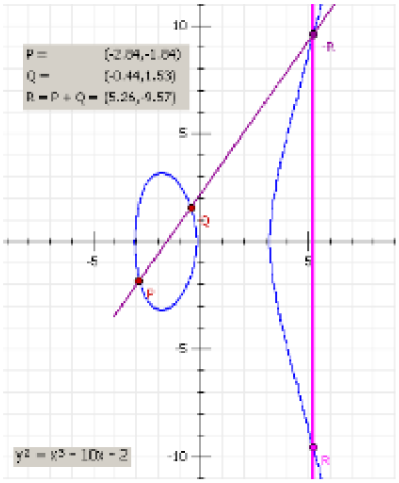
Obr. 6.4: Ukázka vlastností eliptických křivek

Nejpoužívanější digitální podpis ECDSA a jeho fáze generování parametrů, vytvoření a ověření podpisu jsou podrobně popsány. Pro názornou ukázku výpočtů se lze dostat do

sekce s vypočítanými příklady, kde je uvedeno sčítání bodů, zdvojení bodu, jak lze získat řád křivky a bodu a také příklad právě na šifrování výše uvedeným způsobem.

- Postup jak tohoto dosáhnout je následující:

  body P (x_P, y_P) a Q (x_Q, y_Q) propojíme přímkou, která protne křivku v dalším bodě, jejž označíme $-R$ a výsledkem sčítání je bod R, symetrický k $-R$ podle osy x (obr.2). Body symetrické podle osy x nazýváme opačné.



Obr. 2: Sčítání bodů P + Q = R

Obr. 6.5: Ukázka sčítání bodů na EC

Protože je splněna tato podmínka, existuje křivka E: $y^2 = x^3 + x + 1$ nad F(23). Její body jsou vypsány v tab.3 a graficky znázorněny na obr.6. Graf není spojitý, protože křivku nad Fp tvoří konečný počet bodů. Můžeme si například ověřit, že bod (13, 16) patří této křivce - platí totiž:

$$y^2 \bmod p \equiv (x^3 + ax + b) \bmod p.$$

$$16^2 \bmod 23 \equiv (13^3 + 13 + 1) \bmod 23.$$

$$256 \bmod 23 \equiv 2211 \bmod 23.$$

$$3 = 3$$

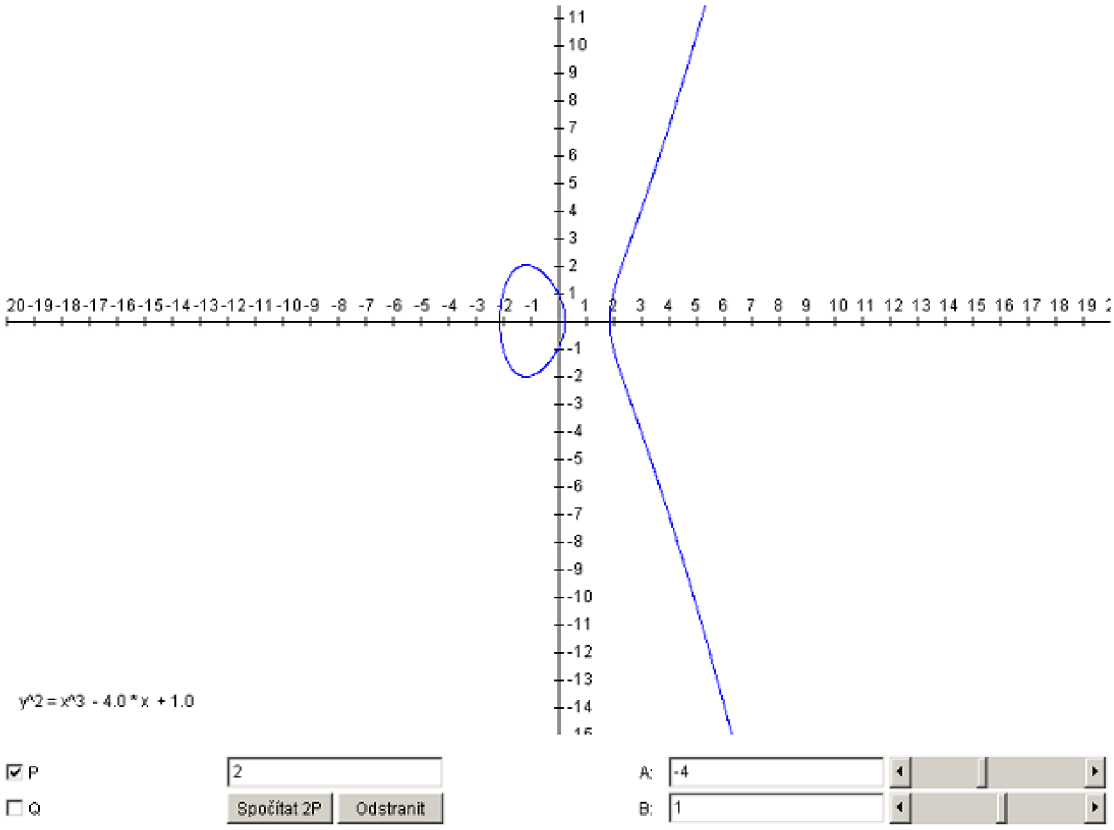
O	(4,0)	(9,7)	P+P (13,10)
(0,1)	(5,4)	(9,16)	(17,3)
(0,22)	P (5,19)	(11,3)	(17,20)
(1,7)	(6,4)	(11,20)	(18,3)
(1,16)	(6,19)	(12,4)	(18,20)
(3,10)	(7,11)	(12,19)	(19,5)
(3,13)	(7,12)	(13,7)	(19,18)

Tab. 3: Body eliptické křivky

Obr. 6.6: Ukázka řešení příkladů

Popis doplňujících java apletů

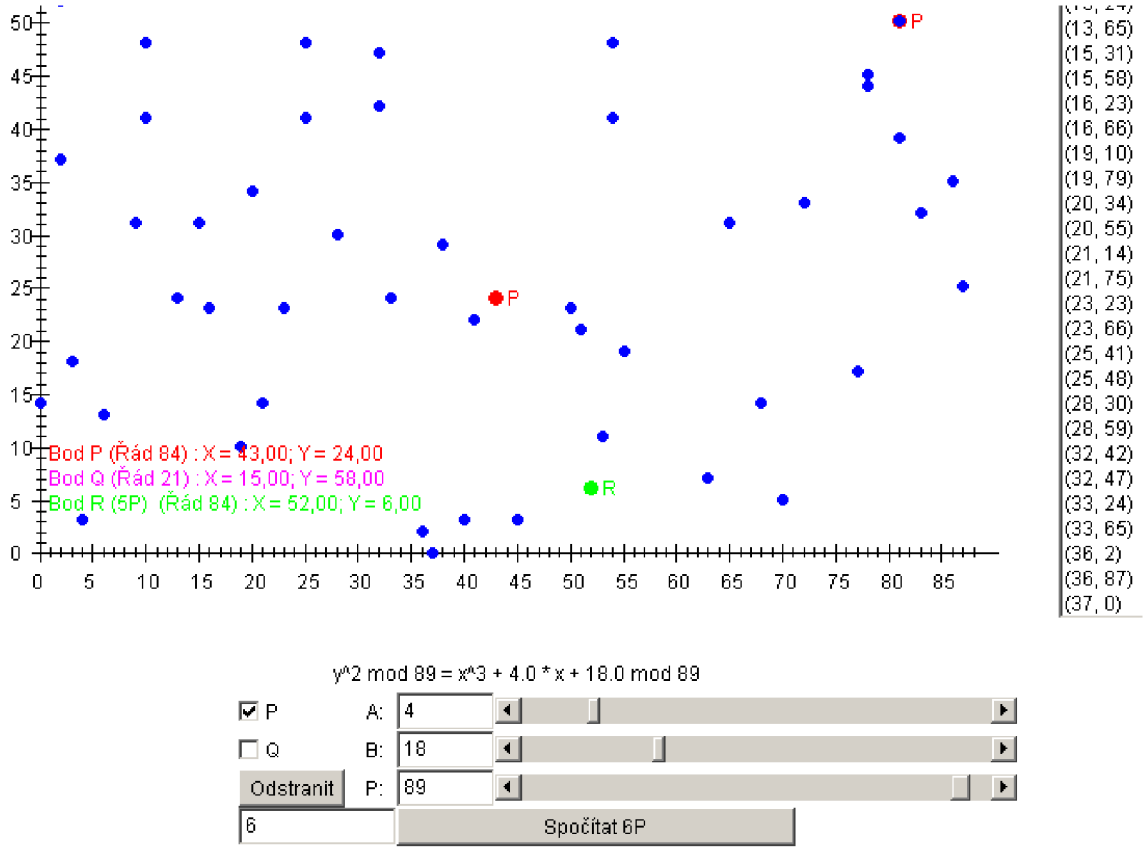
První aplet pracuje nad množinou reálných čísel. Umožňuje po zadání parametrů eliptické křivky $\{a, b\}$ vykreslit tuto křivku. Na ní je poté možno přidávat body a s těmito provádět operace součet a násobení konstantou. Je vždy vypsána aktuální rovnice a souřadnice všech bodů. Aplet je znázorněn na obrázku **6.7**.



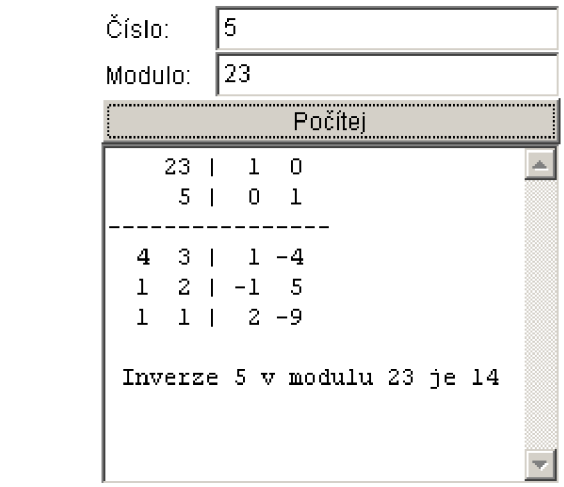
Obr. 6.7: Ukázka apletu pracujícího nad R

Druhý aplet pracuje nad prvočíselným tělesem. Nejprve je třeba zadat parametry křivky {a, b, p}. Grafem už není spojitá křivka jako v předchozím případě, ale pouze množina bodů. Proč to tak je, je vysvětleno výše. Opět je zde možnost přidání bodů, jejich součet a násobení konstantou. Aplet do tabulky vypíše i všechny body křivky. Při určení bodu na křivce, je vypočítán vždy řád tohoto bodu jak je vidět na obrázku **6.8**.

Ve třetím apletu je vytvořen jednoduchý kalkulátor, určený k výpočtu inverzního prvku, který je třeba při dílčích výpočtech součtu a násobení bodů. Aplet je zobrazen na obrázku **6.9**.

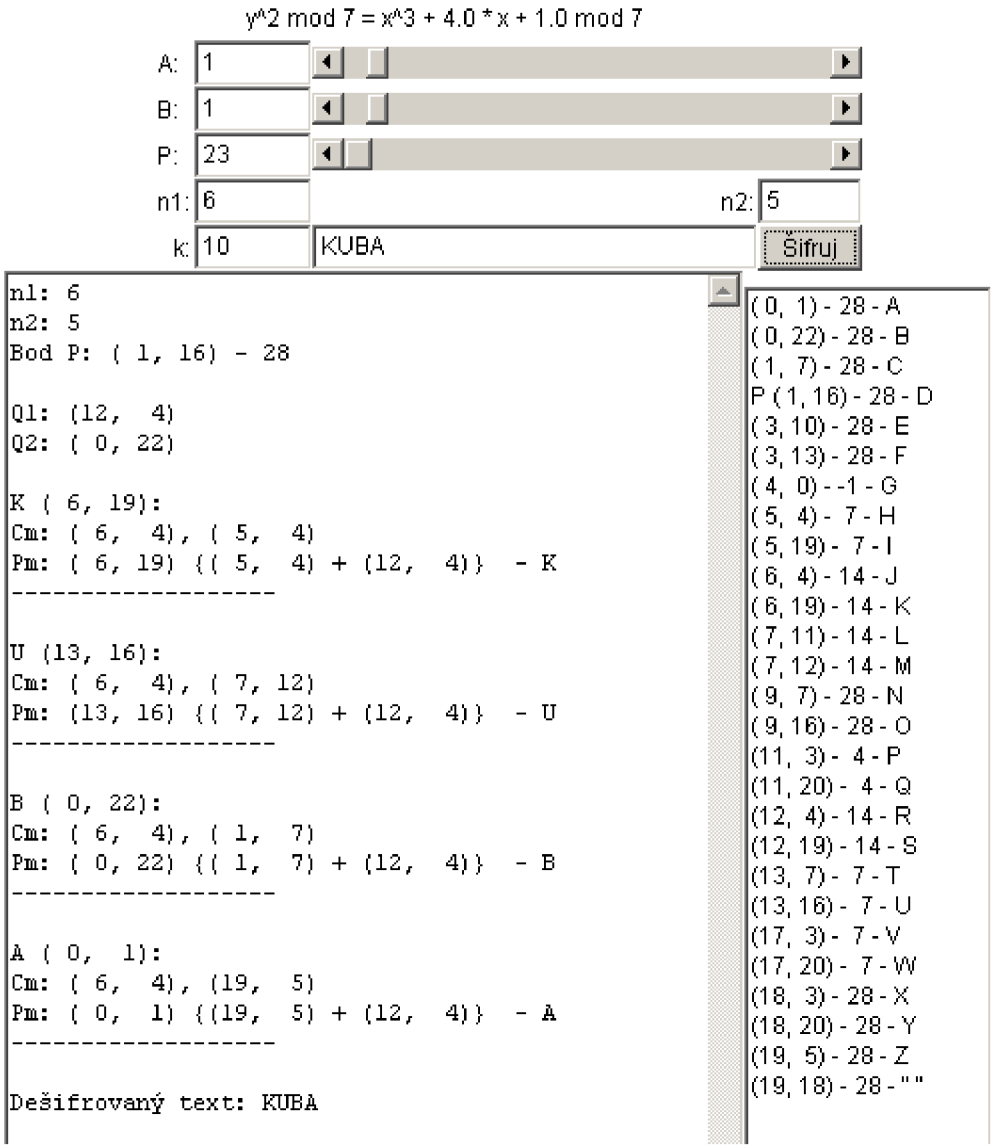


Obr. 6.8: Ukázka apletu pracujícího nad *Fp*



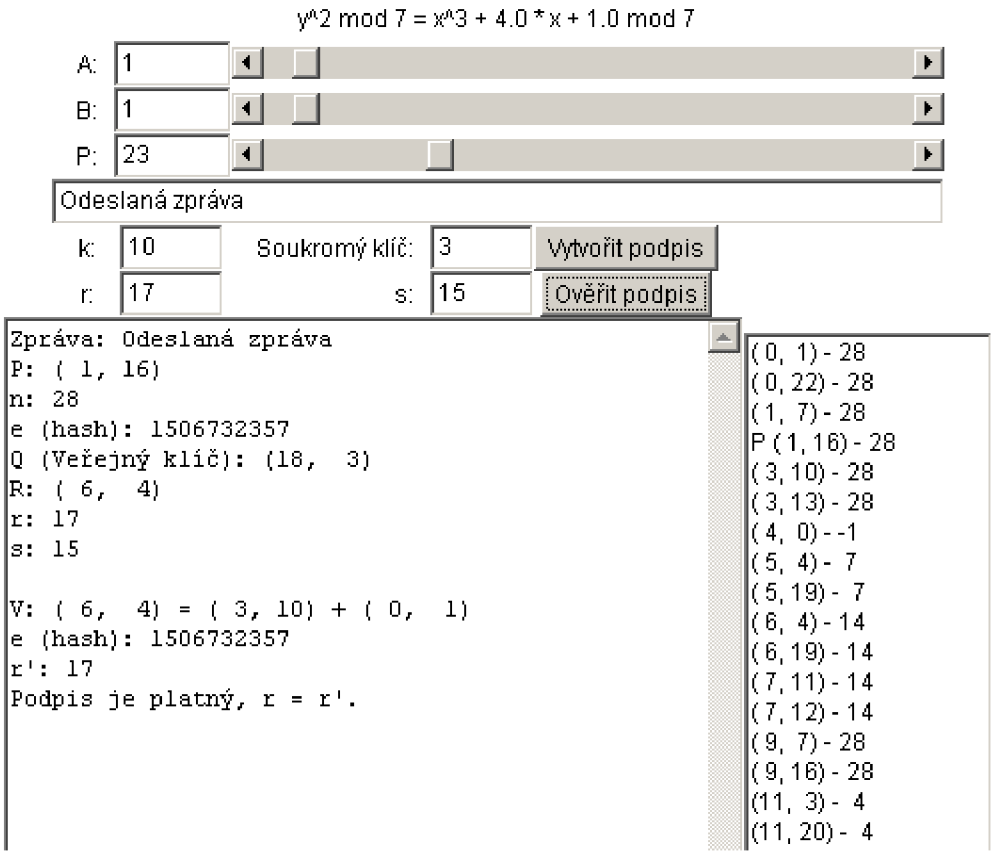
Obr. 6.9: Ukázka apletu na výpočet inverzního prvku

Předposlední aplet zobrazený na obrázku **6.10** je určený k jednoduchému způsobu šifrování, metodou přímého převodu znaků zprávy na body eliptické křivky. Pro naše potřeby je „namapování“ znaků na eliptickou křivku provedeno přímým přiřazením znaku zprávy k určitému bodu na křivce. V praxi by se volil jiný postup a algoritmus.



Obr. 6.10: Ukázka apletu k šifrování a dešifrování

Poslední vytvořený aplet slouží k digitálnímu podpisu zprávy. K tomuto je použito podepisovací schéma ECSS. Při podepsání zprávy vznikne podpis r, s. Pokud se změní jakýkoliv znak zprávy, podpis je neplatný díky hashovaní ověřovací funkci. Pokud by se změnil parametr r nebo s, je podpis znovu neplatný, protože ověřením výpočtu nebude platit rovnost r = r‘. Aplet je zobrazen na obrázku **6.11**.



Obr. 6.11: Ukázka apletu určeného k digitálnímu podepisování

7 ZÁVĚR

Kryptografie eliptických křivek je nadějný obor. Eliptické křivky se stávají součástí nejdůležitějších světových standardů. Implementaci těchto nástrojů nic nebrání, snad jen jejich nezvyklost. Také jejich bezpečnosti se věnuje značná pozornost, takže obavy tohoto druhu asi nebudou tím hlavním důvodem, proč eliptické křivky nejsou masově používány a „staré dobré" algoritmy RSA, DH a DSA ještě nevyklízejí pole. Přesto však tam, kde jsou k dispozici jen omezené hardwarové zdroje, nemají eliptické křivky konkurenci.

Cílem diplomové práce bylo se seznámit s kryptosystémy založenými na eliptických křivkách. V úvodu jsou vysvětleny pojmy nutné k orientaci v dané problematice. Konkrétně je to operace modulo, grupy, konečná tělesa a výpočet inverzního prvku. Byly popsány vlastnosti eliptických křivek, podrobně také operace s body, konkrétně sčítání dvou bodů na eliptické křivce a tzv. zdvojení bodu, neboli násobení bodu konstantou, což je základem pro šifrování nebo digitální podpis. Uvedeny jsou i normy a standardy, ve kterých jsou eliptické křivky zaneseny a uvedeny. Pro představu je uvedeno i porovnání s dalšími asymetrickými kryptosystémy zejména na základě délky použitého klíče.

Další kapitola je věnována samotnému šifrování, výměně klíčů a digitálnímu podpisu. Jako příklad dohody na společném klíči je uveden Diffie-Hellman (DH) nad eliptickou křivkou a MCQ protokol, který vychází z DH. Pro šifrování je nutné umět převést zprávu na body eliptické křivky. Možností je několik, ne vždy je však možné najít odpovídající souřadnice bodu (x, y) a poté se pohybujeme v rovině pravděpodobnosti a vybíráme souřadnici z určité množiny bodů. Uveden je algoritmus převodu podle Koblitze. Jako metoda šifrování je popsán algoritmus ECIES a metoda přímého převodu zprávy na body eliptické křivky. V podkapitole s digitálními podpisy jsou uvedeny ECSS (signature scheme) a ECDSA (digital signature algorithm), který je v současné době nejpoužívanější.

V praktické části jsem vytvořil webovou prezentaci, kde jsou shrnuty nejdůležitější základní informace z problematiky eliptických křivek. Stránky jsou napsány v xhtml a css pomocí textového editoru z důvodu požadavku na spustitelnost z libovolného webového prohlížeče. Pro snazší pochopení jsou zde vypočteny některé příklady a na těchto jsou ukázány základní pravidla a zákonitosti. Studenti, kteří s tímto materiálem přijdou do styku, mají na konci zadány příklady k samostatnému řešení. Uvedeny jsou i správné výsledky a výpočty. Ověření správnosti si však mohou studenti

provést pomocí java apletů, které jsou součástí prezentace. Konkrétně jsou to operace s křivkami nad reálnými čísly, na prvočíselným tělesem, výpočet inverzního prvku, šifrování a dešifrování zprávy a vytvoření digitálního podpisu, podle podpisového schématu ECSS.

8 SEZNAM POUŽITÉ LITERATURY A ZDROJŮ

[1] BURDA, K. Aplikovaná kryptografie – přednáška 6. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. 27 s.

[2] BURŠÍK, F. Problematika převodu zprávy na body eliptické křivky. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2007. 7 s.

[3] DEMYTKO, N. A New Elliptic Curve Based Analogue of RSA. Advances in Cryptology - EUROCRYPT '93, 1994. 40 – 49 s.

[4] DIFFIE, W., HELLMAN, M. New directions in cryptography. IEEE Transactions on Information Theory. 1976. 644 – 654 s.

[5] Elliptic curve addition. Dostupné z www: http://www.certicom.com/ecc_tutorial

[6] HANKERSON, D., VANSTONE, S., MENEZES, A. Guide to elliptic curve cryptography. Springer – Verlag, New York, 2004. 332 s. ISBN 0-387-95273-X.

[7] HUSEMOELLER, D. Elliptic curves, 2nd edition. Springer – Verlag, New York, 2004. 510 s. ISBN 0-387-95490-2

[8] KLÍMA, V., ROSA, T. Kryptologie pro praxi - použití ECC, Sdělovací technika, 11/2005, str. 14-15

[9] KLÍMA, V., ROSA, T. Kryptologie pro praxi - principy ECC, Sdělovací technika, 11/2005, str. 12-13

[10] KOBLITZ, N. A course in number theory and cryptography. Springer – Verlag, New York, 1994. 122 s. ISBN 0-387-94293-9

[11] KOBLITZ, N. Elliptic Curve Cryptosystems. Mathematics of Computation. 1987. 203 – 209 s.

[12] LASOŇ, M. Porovnání bezpečnosti kryptosystémů RSA a eliptických křivek. Březen 2005. 13 s.

[13] MENEZES, A., VASTONE, S. Elliptic Curve Cryptosystems and Their Implementation. Journal of Cryptology. 1993. 209 – 224 s.

[14] OCHODKOVÁ, E. Přínos teorie eliptických křivek k řešení moderních kryptografických systémů. Katedra informatiky, FEI, VŠB – Technická universita Ostrava. 12 s.

[15] OKAMOTO, T., FUJIOKA, A., FUJISAKI, E. An efficien digital signature scheme based on an elliptic curve over the ring *Zn*. NTT Laboratories, Japan, 1-2356, 238-03. 12 s.

[16] PINKAVA, J. Úvod do kryptografie, květen 1998. 26 s.

[17] RABAH, K. Elliptic curve cryptography over bingy finite field GF (2^m). Department of Physics, Eastern Mediterranean University, Turkey. 2006. 26 s.

[18] SZTURC, J. *Šifrátory pro mobilní telefony.* Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2007. 54 s. Vedoucí bakalářské práce doc. Ing. Václav Zeman, Ph.D.

9 SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

3DES	Tripple data encryption standard
AES	Advanced encryption standard
ANSI	American national standard institute
CBC	Cipher block chaining mode
DES	Data encryption standard
DL	Diskrétní logaritmus
DP	Digitální podpis
DSA	Digital signature algorithm
ECC	Elliptic curve cryptography
ECDH	Elliptic curve diffie-hellman
ECDSA	Elliptic curve digital signature algorithm
ECIES	Elliptic curve integrated encryption scheme
ECSS	Elliptic curve signature scheme
FIPS	Federal informatik processing standard
IDEA	International data encryption algorithm
IETF	Internet engineering task force
ISO	International organization for standardization
KDF	Key derivation function
MAC	Message authentication code
MOV	MOV attac (Menezes, Okamoto, Vastone)
MQV	Menezes-Qu-Vanstone protokol
NIST	National institute of standard and technology
RSA	Asymetrický šifrovací algoritmus (Rivest, Shamir, Adleman)
SHA	Secure hash algorithm
SK	Soukromý klíč
TK	Tajný klíč
VK	Veřejný klíč

10 SEZNAM PŘÍLOH

A: DISK CD

Příloha A: Obsah přiloženého CD

SW PODPORA VYUKY/

images

aplet.html
apletDPpopis.html
apletEDpopis.html
apletpodpis.html
apletsifrovani.html
bezpečnost.html
cojsou.html
default.css
ecc.html
ECCApplet.jar
eccpopis.html
examples.html
fp.html
FPApplet.jar
fppopis.html
index.html
inv.html
Inverze.jar
invpopis.html
obecne.html
podpis.html
Podpis.jar
porovnani.html
scitani.html
sifrovani.html
Sifrovani.jar
ukazpriklad.html
vlastnosti.html
vyresenepr.html

ZDROJOVÉ KÓDY

všechny aplety