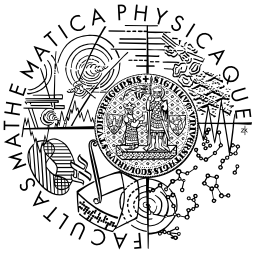

MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Jiří Březina

Rekursivní množiny a diofantické polynomy

Katedra algebry

Vedoucí bakalářské práce: doc. Mgr. Štěpán Holub, Ph.D.

Studijní program: Matematika

Studijní obor: Obecná matematika

Praha 2021

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V ..... dne ....

Podpis autora

Chtěl bych poděkovat vedoucímu práce doc. Mgr. Štěpánu Holubovi, Ph.D. za cenné rady a především ochotu při řešení problémů souvisejících s psaním této práce.

Název práce: Rekursivní množiny a diofantické polynomy

Autor: Jiří Březina

Katedra: Katedra algebry

Vedoucí bakalářské práce: doc. Mgr. Štěpán Holub, Ph.D., Katedra algebry

Abstrakt: Hlavním tématem práce jsou diofantické polynomy a diofantické množiny. V první kapitole jsou také krátce představeny rekursivní a rekursivně vyčíslitelné množiny a uvedena jejich spojitost s množinami diofantickými. Celá druhá kapitola je věnována konstruktivnímu důkazu diofantičnosti množiny prvočísel a s tím úzce spjaté nalezení polynomu, jehož obor kladných hodnot je roven množině prvočísel.

Klíčová slova: rekursivní množiny, diofantické polynomy, polynom reprezentující prvočísla.

Title: Recursive sets and diophantine equations

Author: Jiří Březina

Department: Department of Algebra

Supervisor: doc. Mgr. Štěpán Holub, Ph.D., Department of Algebra

Abstract: Main theme of this thesis are diophantine equations and diophantine sets. In the first chapter we also briefly introduce recursive sets and recursive enumerable sets and show how they are connected with diophantine sets. Whole second chapter is devoted to constructive proof that prime numbers are diophantine and also stating a polynomial which set of positives values is identical to the set of prime numbers.

Keywords: recursive sets, diophantine equations, prime representing polynomial.

Obsah

Úvod

Prvočísla i polynomy patří k základním objektům matematiky. Zajímavou otázkou je, zda dokážeme prvočísla vyjádřit pomocí nějakého polynomu. Přesněji řečeno: Lze zapsat množinu prvočísel jakožto množinu všech kladných hodnot nějakého diofantického polynomu? Kladnou odpověď na tuto otázku dal v roce 1970 ruský matematik Jurij Matijasevič, když navázal na práci Julie Robinson, Martina Davise a Hilaryho Putnama. Všichni čtyři zmínění matematici přitom nehledali primárně tento polynom, nýbrž dokazovali neexistenci řešení Hilbertova Desátého problému. Z postupu, jímž tuto neexistenci řešení prokázali, však ihned plynula existence námi hledaného polynomu.

V naší práci vyjdeme především z článku [1] a explicitně vyjádříme polynom reprezentující prvočísla. Buďto s pomocí jiných zdrojů, nebo vlastní invencí zároveň doplníme kroky, které jsou v [1] vynechány. Po cestě k nalezení polynomu ukážeme i jiné netriviální poznatky o diofantických množinách či řešení Pellovy rovnice.

V první kapitole představíme rekursivní, rekursivně vyčíslitelné a diofantické množiny. Nastíníme, jak spolu uvedené množiny souvisí a především, jak souvisí s naším problémem. Tato kapitola bude tudíž sloužit jako menší předehra ke kapitole druhé, ve které se již budeme plně věnovat hledání již zmiňovaného polynomu reprezentujícího prvočísla.

1. Rekursivní a diofantické množiny

V první kapitole představíme rekursivní, rekursivně vyčíslitelné a diofantické množiny. Uvedeme, jak spolu dané množiny souvisí a jak souvisí s polynomem reprezentujícím prvočísla. I vzhledem k tomu, že hlavní důraz práce je kladen na druhou kapitolu, budeme v této postupovat méně formálně.

Přirozenými čísly míníme nezáporná celá čísla, značíme $\mathbb{N}$. V celé práci, pokud nebude řečeno jinak, budou za všechny proměnné dosazována pouze přirozená čísla.

1.1 Rekursivní a rekursivně vyčíslitelné množiny

Rekursivní a rekursivně vyčíslitelné množiny se obvykle definují pomocí Turingova stroje. Tento stroj navrhl ve 30. letech minulého století Alan Turing jako pokus o formalizaci algoritmu. Pracovat v této kapitole s původním Turingovým strojem by pro nás bylo zbytečně obtížné, proto budeme psát kódy v jazyce Python. Oba přístupy jsou ekvivalentní v tom smyslu, že co zvládne vyřešit Turingův stroj, zvládne i Python a obráceně.

Definice 1. Množina $M \subset \mathbb{N}^k$ je *rekursivně vyčíslitelná*, pokud existuje Turingův stroj, který ukončí svůj výpočet právě tehdy, když mu je dán vstup z množiny M.

Definice 2. Množina $M \subset \mathbb{N}^k$ je *rekursivní*, pokud existuje Turingův stroj, který pro prvky M vrátí hodnotu 1 a pro ostatní prvky vrátí hodnotu 0.

Zmíníme jeden příklad množiny, která je rekursivní. Nepřekvapivě to bude právě množina prvočísel.

Příklad 1. Množina prvočísel je rekursivní.

Níže je jeden z možných programů dosvědčující rekursivitu prvočísel. Vstupem programu je přirozené číslo n. Výstupem pak jedna, nebo nula v závislosti, zda n je, či není prvočíslo.

```
def prvocislo(n):
    if n <= 1:
        return 0
    else:
        i = 2
        while i < n:
            if n%i==0:
                return 0
            i = i+1
        return 1
```

Mezi dvěma definovanými množinami existuje jednoduchý vztah. Množina všech rekursivních množin je totiž podmnožinou množiny všech rekursivně vyčíslitelných množin.

Tvrzení 1. Každá rekursivní množina je rekursivně vyčíslitelná.

Důkaz. Nechť $M \subset \mathbb{N}^k$ je rekursivní. Potom z definice rekursivní množiny existuje program rozhodující o náležení do M. Označme tento program T_1 a zkonstruujme pomocí něj program T_2 následovně. Vstupem programu T_2 je $k-$tice čísel, která je pro jednoduchost označena písmenem a.

```
def T_2(a):
    if T_1(a) == 1:
        return 1
    else:
        i=1
        while i < 2:
            print(i)
```

Program T_2 se zastaví právě tehdy, když je jeho vstupem prvek množiny M, neboli T_2 je hledaným programem z definice rekursivní vyčíslitelnosti.

□

Opačná inkluze neplatí, existují množiny, které jsou rekursivně vyčíslitelné a nejsou rekursivní. Ty jsou pro náš text však nepodstatné.

1.2 Diofantické množiny

Od teorie vyčíslitelnosti, kam patří rekursivní množiny, se přesouváme k čisté teorii čísel a takzvaným množinám diofantickým. Ačkoliv jsou tyto dva obory dost odlišné, platí, že diofantické množiny jsou právě množiny rekursivně vyčíslitelné.

V této sekci budeme místo $k-$tice přirozených čísel $(a_1, a_2, ..., a_k)$ psát stručněji $\mathbf{a}_k$, tedy $\mathbf{a}_k = (a_1, a_2, ..., a_k)$. Analogicky budeme toto značení používat pro jiné posloupnosti čísel.

Definice 3. Polynom P nazveme *diofantický*, pokud jsou všechny jeho koeficienty celá čísla.

Definice 4. Množina $D \subseteq \mathbb{N}^k$ je *diofantická*, pokud existuje diofantický polynom $P[\mathbf{x}_k, \mathbf{y}_l]$ takový, že

$$D = \{\mathbf{a}_k \in \mathbb{N}^k \mid \text{Existuje } \mathbf{b}_l \in \mathbb{N}^l \text{ takové, že } P(\mathbf{a}_k, \mathbf{b}_l) = 0.\}.$$

V případě, kdy k posloupnosti přirozených čísel $\mathbf{a}_k$ existuje posloupnost přirozených čísel $\mathbf{b}_l$ taková, že $P(\mathbf{a}_k, \mathbf{b}_l) = 0$, řekneme, že polynom $P[\mathbf{x}_k, \mathbf{y}_l]$ je řešitelný pro $\mathbf{a}_k$.

Pozor, například polynom $x+y$ není řešitelný pro $x = 1$. V definici diofantické množiny totiž bereme pouze přirozená čísla a v nich nemá rovnice $y = -1$ řešení.

Příklad 2. Množina kladných složených čísel je diofantická množina.

Číslo a je kladné složené právě tehdy, když lze zapsat jako součin dvou přirozených čísel větších než jedna. Uvažme polynom $P[x_1, y_1, y_2] = x_1 - (y_1+2)(y_2+2)$. Z definice polynomu je zjevné, že P je řešitelný v y_1, y_2 právě tehdy, když za proměnnou x_1 dosadíme složené číslo. Označíme-li množinu kladných složených čísel C, pak platí $C = \{a \in \mathbb{N} \mid$ Existují přirozená b, c taková, že $a-(b+2)(c+2) = 0.\}$.

V definici diofantických množin se nemusíme omezovat pouze na jeden polynom. Následující lemma říká, že k popisu dané množiny jich můžeme použít konečně mnoho. Samotné lemma i ideu jeho důkazu několikrát využijeme v druhé kapitole.

Lemma 2. Množina $D \subseteq \mathbb{N}^k$ je diofantická právě tehdy, když existují diofantické polynomy $P_1[\mathbf{x}_k, \mathbf{y}_l]$, $P_2[\mathbf{x}_k, \mathbf{y}_l]$, ..., $P_n[\mathbf{x}_k, \mathbf{y}_l]$ takové, že

$$D = \{\mathbf{a}_k \in \mathbb{N}^k \mid \text{Existuje } \mathbf{b}_l \in \mathbb{N}^l \text{ takové, že } P_i(\mathbf{a}_k, \mathbf{b}_l) = 0 \text{ pro každé } 1 \leq i \leq n.\}.$$

Důkaz. Implikace zleva doprava je zřejmá. K důkazu opačné implikace definujeme polynom

$$P[\mathbf{x}_k, \mathbf{y}_l] = (P_1[\mathbf{x}_k, \mathbf{y}_l])^2 + (P_2[\mathbf{x}_k, \mathbf{y}_l])^2 + ... + (P_n[\mathbf{x}_k, \mathbf{y}_l])^2.$$

Pokud platí $P_i(\mathbf{a}_k, \mathbf{b}_l) = 0$ pro všechna $1 \leq i \leq n$, poté zároveň platí také $P(\mathbf{a}_k, \mathbf{b}_l) = 0$. Naopak, pokud je $P(\mathbf{a}_k, \mathbf{b}_l) = 0$, musí být i všechny sčítance z definice P nulové, tudíž i všechny P_i musí být rovny nule po dosazení $(\mathbf{a}_k, \mathbf{b}_l)$.

Z předešlých úvah plyne, že P je řešitelný pro $\mathbf{a}_k$ právě tehdy, když jsou všechny P_i řešitelné pro $\mathbf{a}_k$. Množina D tudíž může být zadána jedním polynomem (polynomem P), tedy D je diofantická.

□

Dostáváme se k souvislosti diofantických množin s množinami rekursivně vyčíslitelnými, popřípadě rekursivními.

Tvrzení 3. Každá diofantická množina je rekursivně vyčíslitelná.

Důkaz. Nechť M je diofantická množina. Pro jednoduchost předpokládejme, že $M = \{\mathbf{a}_k \in \mathbb{N}^k \mid$ Existují přirozená b_1, b_2 taková, že $P(\mathbf{a}_k, b_1, b_2) = 0.\}$. S využitím pythonovské funkce numpy.polyval(P, $\mathbf{a}_k$, b_1, b_2), která vyhodnotí polynom P v bodě $(\mathbf{a}_k, b_1, b_2)$, sestrojíme následující program.

```
def enum(a_k):
    b1 = 0
    b2 = 0
    while numpy.polyval(P, a_k, b1, b2) != 0:
        if b1 > b2:
            b2 = b2 + 1
        else:
            b1 = b1 + 1
            b2 = 0
```

Jestliže $\mathbf{a}_k$ neleží v M, neboli neexistují b_1, b_2 taková, že $P(\mathbf{a}_k, b_1, b_2) = 0$, program nikdy nedokončí svůj výpočet. Naopak, pokud $\mathbf{a}_k$ leží v M, pak existují b_1, b_2 taková, že $P(\mathbf{a}_k, b_1, b_2) = 0$. Při těchto b_1, b_2 se program ukončí, protože while cyklus běží přes všechny dvojice přirozených čísel (b_1, b_2).

Celkově tudíž program doběhne právě tehdy, když $\mathbf{a}_k$ leží v M, z čehož vyplývá, že M je rekursivně vyčíslitelná.

Kdybychom uvažovali obecně b_1, b_2, ..., b_l namísto b_1, b_2, napsali bychom kód programu analogicky, pouze bychom použili více if-else podmínek v těle while cyklu.

□

Překvapivějším tvrzením je, že platí i opačná inkluze. Tento výsledek dokázal Jurij Matijasevič při řešení Desátého Hilbertova problému. Jelikož navázal na dlouhodobou práci Julie Robinsonové, Martina Davise a Hilaryho Putnama, říká se větě v anglické literatuře též MRDP Theorem.

Věta 4 (Matijasevičova)**.** Každá rekursivně vyčíslitelná množina je diofantická.

Důkaz věty jde naprosto mimo záběr této práce. Čtenář ho může nalézt v [2], společně s kompletním důkazem neexistence řešení Desátého Hilbertova problému. Pro nás je zajímavý především následující důsledek věty.

Tvrzení 5. Množina prvočísel je diofantická.

Důkaz. V předchozí sekci jsme dokázali, že každá rekursivní množina je rekursivně vyčíslitelná. Z Matijasevičovy věty tudíž plyne, že každá rekursivní množina je diofantická, speciálně množina prvočísel je diofantická (viz též příklad v předchozí sekci).

□

Jinými slovy, existuje diofantický polynom $M[k, y_1, ..., y_l]$, který je řešitelný právě tehdy, když je k prvočíslo. První takový polynom našel sám Jurij Matijasevič v roce 1971. V celé druhé kapitole se budeme věnovat polynomu reprezentujícímu prvočísla. Ten je s výše uvedeným polynomem M úzce svázán, ale neznačí totéž. Co pojmem polynom reprezentující prvočísla myslíme je uvedeno v následující definici.

Definice 5. Diofantický polynom $P[x_1, x_2, ..., x_k]$ nazveme *polynomem reprezentujícím prvočísla*, pokud množina jeho kladných hodnot po dosazení přirozených čísel je totožná s množinou prvočísel.

V příští sekci vyjasníme, proč intuitivně jasný pojem definujeme tak složitě. Nejprve ale uveďme souvislost diofantického polynomu M uvažovaného výše s polynomem reprezentujícím prvočísla.

Tvrzení 6. Existuje polynom reprezentující prvočísla.

Důkaz. Dle Tvrzení 5 a snadné substituce v první proměnné existuje polynom $M[k, y_1, ..., y_l]$, který je řešitelný právě tehdy, když je $k + 2$ prvočíslo. Položme

$$P[x_1, y_1, ..., y_l] = (x_1 + 2)(1 - (M[x_1, y_1, ..., y_l])^2).$$

Jelikož za proměnné dosazujeme pouze nezáporná čísla, je první faktor P vždy kladný. Pokud za x_1 dosadíme k takové, že $k+2$ je prvočíslo, budou existovat přirozená $b_1, b_2, ..., b_n$ taková, že $M(k, b_1, ..., b_n) = 0$. Z vyjádření P proto plyne, že $k + 2 = P[k, b_1, ..., b_n]$. Každé prvočíslo tudíž leží v oboru hodnot P. Pokud naopak po dosazení bude $k + 2$ složené, je polynom $M[k, y_1, ..., y_n]$ neřešitelný. Výraz $1 - (M[k, b_1, ..., b_l])^2$ je poté pro všechna přirozená $b_1, b_2, ..., b_n$ nekladný a stejně tak $P[k, b_1, ..., b_l]$. Nekladné hodnoty polynomu však neuvažujeme, což dokazuje, že nic kromě prvočísel v oboru hodnot P neleží. Dohromady dostáváme, že P je polynomem reprezentujícím prvočísla.

□

Matijasevičova věta nám tudíž zaručuje existenci polynomu reprezentujícího prvočísla. Důkaz existence je však pouze nekonstruktivní. V druhé kapitole dokážeme předchozí tvrzení konstruktivně, neboli nalezneme a explicitně uvedeme polynom P z důkazu předchozího tvrzení.

1.3 Omezující podmínky pro reprezentaci prvočísel

Vysvětlíme, proč jsme polynom reprezentující prvočísla definovali tak složitě. Konkrétně se budeme věnovat dvoum částem uvedené definice: dosazování pouze přirozených čísel a braní pouze kladných čísel z oboru hodnot.

Dosazování přirozených čísel by se díky následující slavné větě dalo obejít.

Věta 7 (Lagrangeova věta o čtyřech čtvercích)**.** Pro každé přirozené číslo m existují přirozená čísla m_1, m_2, m_3, m_4 taková, že $m = m_1^2 + m_2^2 + m_3^2 + m_4^2$.

V našem polynomu bychom mohli tudíž nahradit každou proměnnou x_i čtyřmi novými proměnnými x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}. Stačilo by pro každé i položit $x_i = x_{i_1}^2 + x_{i_2}^2 + x_{i_3}^2 + x_{i_4}^2$. Za původní proměnné by pak byla dosazována pouze přirozená čísla a díky Lagrangeově větě by obor hodnot polynomu zůstal stejný.

Polynomy reprezentující prvočísla však mají i při podmínce dosazování pouze přirozených čísel hodně proměnných (například ten, který uvedeme v druhé kapitole jich bude mít 26). Tuto podmínku proto dodáváme především kvůli přehlednosti.

Opačná situace nastává u podmínky o kladných hodnotách. Platí totiž, že žádný polynom, jehož oborem hodnot jsou právě prvočísla, neexistuje.

Tvrzení 8. Nechť $Q[x_1, x_2, ..., x_m]$ je diofantický polynom, který nabývá pouze prvočíselných hodnot. Potom je Q konstantní.

Důkaz. Tvrzení dokážeme indukcí dle m, tj. počtu proměnných.

Uvažme nejprve polynom $Q[x_1]$ v jedné proměnné. Označme $p = Q(0)$, p je z předpokladu prvočíslo. Pro každé přirozené n platí $Q(np) \equiv Q(0) \equiv 0 \mod p$, neboli p dělí $Q(np)$. Jelikož $Q(np)$ je z předpokladu také prvočíslo, musí platit $Q(np) = p$ pro všechna přirozená n. Polynom $Q[x_1] - p$ má proto nekonečně mnoho kořenů a tudíž musí být nulový. Odsud snadno dostáváme $Q[x_1] = p$, tedy Q je konstantní.

Předpokládejme nyní, že tvrzení platí pro všechny polynomy o méně než m proměnných a dokažme ho pro $Q[x_1, x_2,..., x_m]$ o m proměnných. Jelikož platí, že $Q[x_1, x_2,..., x_m] = Q[x_1, x_2,..., x_{m-1}][x_m]$, můžeme psát

$$Q[x_1, x_2,..., x_m] = \sum_{i=0}^{n} Q_i x_m^i,$$

kde n je nejvyšší mocnina x_m v polynomu Q a $Q_i \in Q[x_1, x_2,..., x_{m-1}]$. Q_n je polynom v $m-1$ proměnných. Pokud by pro všechny $(m-1)-$tice nezáporných čísel $a_1, a_2,..., a_{m-1}$ nastalo $Q_n((a_1, a_2,..., a_{m-1})) = 0$, byl by $Q_n + 2$ polynom o méně než m proměnných nabývající pouze prvočíselné hodnoty 2. Z indukčního předpokladu by poté vyplývalo, že Q_n je konstantní, tedy konstantně nulový. Musí proto existovat nezáporná $a_1, a_2,..., a_{m-1}$ taková, že $Q_n((a_1, a_2,..., a_{m-1})) \neq 0$. Polynom $R[x_m] = Q[a_1, a_2,..., a_{m-1}, x_m]$ je nenulový polynom v jedné proměnné, který stále nabývá pouze prvočíselných hodnot. Z indukčního předpokladu poté plyne, že $R[x_m]$ je konstantní, neboli $n = 0$. To znamená, že proměnná x_m se v polynomu nevyskytuje a Q je tudíž polynom o méně než m proměnných. Z indukčního předpokladu je poté Q konstantní.

□

Věta 9. Neexistuje diofantický polynom, jehož obor hodnot by byl roven množině prvočísel.

Důkaz. Snadný důsledek předchozího tvrzení.

□

2. Konstrukce polynomu

2.1 Princip konstrukce polynomu

Tato kapitola tvoří jádro práce, poněvadž na jejím konci najdeme kýžený polynom reprezentující prvočísla. Již v minulé kapitole jsme spatřili, že naprosto stěžejním při tom bude konstruktivně ukázat, že množina prvočísel je diofantická. Na rozdíl od nekonstruktivního důkazu uvedeného v minulé kapitole, nyní opravdu najdeme polynom, který je řešitelný tehdy a jen tehdy, když za jednu z jeho proměnných je dosazeno prvočíslo. Jak z tohoto polynomu odvodit polynom reprezentující prvočísla jsme již nahlédli v důkazu Tvrzení 6.

Obtížným zůstává nalezení polynomu dosvědčujícího diofantičnost množiny prvočísel. Nalézt se nám ho podaří ve Větě 30, která shrne všechny dříve dokázané poznatky dohromady a bude proto hlavní větou této práce. Budeme při tom postupovat až na malé výjimky podle článku [1]. Plán tohoto postupu nemusí být při první četbě zcela jasný. Abychom čtenáři poskytli motivaci pro některé hůře pochopitelné kroky a pomohli mu neztratit se v technických lemmatech, věnujeme následující řádky vysvětlení hlavní myšlenky našeho postupu.

První otázkou je, jak vůbec testovat pomocí polynomu, že je dané číslo prvočíslem. K tomu použijeme jeden krásný test prvočíselnosti z teorie čísel známý pod názvem Wilsonova věta. Věta tvrdí následující:

Číslo $k > 1$ je prvočíslo právě tehdy, když k dělí $(k-1)! + 1$.

Problém detekce prvočísel jsme tedy převedli na relaci obsahující dělitelnost a faktoriál. Skutečnost, že je číslo dělitelné jiným číslem, se pomocí diofantické rovnice napíše snadno. Problém nastává při určení, zda pro daná f, k platí $f = k!$. Tento problém vyřešíme ve Větě 28. Spojením Věty 28 a Wilsonovy věty pak získáme Větu 30 dokazující diofantičnost množiny prvočísel.

K důkazu výše uvedených vět bude ovšem nejdříve nutné vyslovit mnoho pomocných lemmat. Protože faktoriál roste velmi rychle, nadefinujeme si posloupnost reálných čísel, jejíž členy též rostou velmi rychle, a to exponenciálně. Tuto posloupnost budou tvořit řešení Pellovy rovnice, a tento druh rovnice nás také bude provázet velkou částí této kapitoly. V následující podkapitole připomeneme obecné znalosti o Pellově rovnici, s kterými je již mnohý čtenář obeznámen. Obecné znalosti budeme později využívat pro rovnici

$$x^2 - (a^2 - 1)y^2 = 1,$$

kde a je parametr, pro který platí $a > 1$.

Všechny poznatky o řešení této rovnice nám poté pomohou při dokazování uvedených vět.

Ještě na závěr této sekce poznamenejme, že konstrukce našeho polynomu reprezentujícího prvočísla bude mírně odlišná od té uvedené ve Tvrzení 6. Zatímco

tam jsme uvažovali obecný M řešitelný právě tehdy, když jedna z jeho proměnných nabývá prvočíselné hodnoty, náš konkrétní M bude součtem druhých mocnin. M tím pádem bude nabývat pouze nezáporných hodnot a není nutné ho mocnit na druhou, postačí položit

$$P[x_1, y_1,..., y_l] = (x_1 + 2)(1 - M[x_1, y_1,..., y_l]).$$

Zdánlivým paradoxem je, že se P, ačkoliv má nabývat pouze prvočíselných hodnot, rozkládá na dva faktory. V případě, že hodnota P je kladná, je ale druhý faktor vždy roven jedné a tento rozklad je tudíž nevlastní.

2.2 Pellova rovnice

Jak jsme zmínili v předešlé sekci, naše pouť začne u Pellovy rovnice. Na první pohled nemusí být spojitost mezi Pellovou rovnicí a naším cílem vůbec očividná. Později se však ukáže, že vlastnosti posloupnosti nezáporných řešení rovnice $x^2 - (a^2 - 1)y^2 = 1$ budou naprosto fundamentální pro důkaz diofantičnosti množiny prvočísel.

V této sekci uvedeme několik tvrzení platících pro obecnou Pellovu rovnici. Vycházíme při tom z textu Keitha Conrada, který je volně dostupný na internetu (viz [3]).

Definice 6. Diofantická rovnice tvaru $x^2 - dy^2 = 1$, kde d je kladné nečtvercové číslo, se nazývá Pellova rovice.

Pro úplnost poznamenejme, že nečtvercové číslo značí číslo, které není druhou mocninou žádného přirozeného čísla a že hledáme pouze celočíselná řešení dané rovnice.

Poznámka. Z předpisu Pellovy rovnice vidíme, že $(1, 0)$ a $(-1, 0)$ jsou vždy její řešení. Tato dvě řešení nazýváme triviální. Pokud je (a, b) řešení Pellovy rovnice, jsou řešení i $(a, -b)$, $(-a, b)$ a $(-a, -b)$. Při řešení Pellovy rovnice tedy stačí uvažovat pouze nezáporné dvojice (a, b) a zbylá řešení z těchto „nezáporných řešení" odvodit.

Na množině řešení Pellovy rovnice zavedeme grupovou strukturu. Označme M_d jakožto množinu všech řešení Pellovy rovnice $x^2 - dy^2 = 1$. Dále pro všechna $(a, b), (e, f) \in M_d$ definujme binární operaci $\cdot$ následovně

$$(a, b) \cdot (e, f) = (ae + dbf, af + be).$$

Věta 10. Nechť $x^2 - dy^2 = 1$ je Pellova rovnice. Množina řešení této rovnice spolu s binární operací $\cdot$ definovanou výše tvoří grupu.

Důkaz. Nejprve dokážeme uzavřenost množiny řešení $x^2 - dy^2 = 1$ vzhledem k výše definované operaci. Uvažme libovolná $(a, b), (e, f) \in M_d$, potřebujeme ukázat, že i $(ae + dbf, af + be) \in M_d$, neboli $(ae + dbf, af + be)$ je řešením dané rovnice. Pomocí několika algebraických úprav vypočteme, že
$(ae + dbf)^2 - d(af + be)^2 = a^2e^2 + 2aedbf + d^2b^2f^2 - da^2f^2 - 2dafbe - db^2e^2 = a^2(e^2 - df^2) - db^2(-df^2 + e^2) = (e^2 - df^2)(a^2 - db^2) = 1 \cdot 1 = 1$. V poslední

rovnosti jsme využili toho, že (a, b), (e, f) jsou řešení $x^2 - dy^2 = 1$ a dohromady tedy dokázali uzavřenost grupy na definovanou binární operaci.

Důkaz asociativity násobení je rutinní výpočet. Z definice násobení snadno nahlédneme, že (1,0) je jednotkový prvek v grupě. Zároveň díky poznámce výše tento prvek vždy v grupě leží. Podobně nahlédneme, že inverzním prvkem pro (a, b) je $(a, -b)$. Platí totiž $(a, b) \cdot (a, -b) = (a^2 - db^2, -ab + ab) = (1, 0)$. Stejně vyjde i součin $(a, -b) \cdot (a, b)$, protože definované násobení je zřejmě komutativní. Opět díky stejné poznámce platí, že inverz prvku vždy leží v M_d. Tím jsme ověřili všechny axiomy grupy, jak bylo požadováno.

□

Grupu, kterou jsme zkonstruovali označíme stejně jako její nosnou množinu, tj. M_d. Všimněme si, že definice násobení v grupě není žádná umělá konstrukce, nýbrž prosté násobení dvou prvků v $\mathbb{Z}[\sqrt{d}]$ s následným porovnáním koeficientů. V následující větě ukážeme souvislost mezi těmito dvěma strukturami. Potřebujeme k tomu definici normy v kvadratických rozšířeních a dvě snadná pozorování o této normě. Definici i následující lemma čerpáme z knihy [4]. Na rozdíl od této knihy však neuvažujeme v definici normy absolutní hodnotu. Tento rozdíl je pro naše účely zanedbatelný a navíc nám pomůže jemněji rozlišit prvky okruhu $\mathbb{Z}[\sqrt{d}]$. Námi zavedená norma pak ovšem nebude standardní normou, jak ji v matematice obvykle chápeme.

Definice 7. Nechť d je nečtvercové číslo, definujeme zobrazení $N : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}$ tak, že pro každé $a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ položíme $N(a + b\sqrt{d}) = a^2 - db^2$. Zobrazení N nazveme norma.

Lemma 11. Pro každá $\alpha, \beta \in \mathbb{Z}[\sqrt{d}]$ platí
a) $N(\alpha \cdot \beta) = N(\alpha) \cdot N(\beta)$,
b) $N(\alpha) = \pm 1$ právě tehdy, když α je invertibilní prvek v $Z[\sqrt{d}]$.

V další části pro nás bude $\mathbb{Z}^*[\sqrt{d}]$ multiplikativní grupa tvořená invertibilními prvky okruhu $\mathbb{Z}[\sqrt{d}]$. Předchozí lemma tvrdí, že $\mathbb{Z}^*[\sqrt{d}]$ obsahuje právě prvky $\mathbb{Z}[\sqrt{d}]$, jejichž norma je rovna ± 1.

Věta 12. Grupa řešení Pellovy rovnice $x^2 - dy^2 = 1$ je izomorfní podgrupě $\mathbb{Z}^*[\sqrt{d}]$ tvořené právě všemi prvky, jejichž norma je rovna jedné.

Důkaz. Definujme zobrazení $\varphi : M_d \to \mathbb{Z}^*[\sqrt{d}]$, $\varphi((a,b)) = a + b\sqrt{d}$. Obraz prvku (a, b) má normu rovnou $a^2 - db^2$. Jelikož $(a, b) \in M_d$, platí $a^2 - db^2 = 1$ a z úvah před větou proto $\varphi((a, b))$ leží v $\mathbb{Z}^*[\sqrt{d}]$. Zobrazení φ je tudíž dobře definované.

Pomocí standardního výpočtu ověříme, že φ je homomorfismus. Pro libovolná $(a, b), (e,f) \in M_d$ platí $\varphi((a, b) \cdot (e, f)) = \varphi((ae + dbf, af + be)) = ae + dbf + (af + be)\sqrt{d} = (a + b\sqrt{d})(e + f\sqrt{d}) = \varphi((a, b)) \cdot \varphi((e, f))$.

Dále výpočtem jádra φ dokážeme, že zobrazení je prosté. Buď $(a, b) \in Ker\ \varphi$, tedy $a + b\sqrt{d} = 1$. Kdyby bylo $b \neq 0$, mohli bychom rovnici ekvivalentně upravit na $\sqrt{d} = \frac{1-a}{b}$. Z tohoto vyjádření plyne, že $\sqrt{d}$ lze zapsat jako podíl dvou celých čísel. Číslo d je ovšem z předpokladu nečtvercové, tudíž je $\sqrt{d}$ iracionání, což je ve sporu s předešlými úvahami. Musí proto platit $b = 0$, což implikuje $a = 1$, neboli $(a, b) = (1, 0)$ a jádro φ je tak triviální.

Nakonec ještě dokážeme, že obrazem zobrazení je množina $X = \{\alpha \in \mathbb{Z}^*[\sqrt{d}] \mid N(\alpha) = 1\}$. Pro každé $(a, b) \in M_d$ platí $N(\varphi((a, b))) = N(a+b\sqrt{d}) = a^2-db^2 = 1$, což dokazuje, že obraz zobrazení je podmnožinou X. Pro druhou inkluzi uvažme libovolné $\alpha \in X$. Buď $\alpha = a + b\sqrt{d}$ pro a, b celočíselné. Potom $\alpha = \varphi((a, b))$ a díky požadavku na hodnotu normy α dostáváme $(a, b) \in M_d$.

Tvrzení nyní plyne z první věty o izomorfismu.

□

Víme, že izomorfní grupy jsou v podstatě stejné. Dále proto spojením $x+y\sqrt{d}$ *je řešením Pellovy rovnice* míníme, že uspořádaná dvojice (x, y) je řešením Pellovy rovnice.

Ve zbytku sekce ukážeme dva klíčové poznatky o Pellově rovnici. Prvním je, že každá Pellova rovnice má nějaké netriviální řešení. Druhým, že všechna nezáporná řešení Pellovy rovnice (tj. řešení, která mají obě složky nezáporné) se dají nagenerovat pouze pomocí jediného řešení.

První výsledek byl dokázán Lagrangem v roce 1768. Jeho důkaz zde nebudeme uvádět, je možné ho nalézt v Conradově článku [3].

Věta 13. Pro každé nečtvercové kladné d má rovnice $x^2 - dy^2 = 1$ nějaké netriviální řešení.

Druhý výsledek lze formulovat tak, že grupa řešení Pellovy rovnice je cyklická. Generátorem této grupy je minimální řešení, což je pojem, který definujeme později. V tomto duchu je i tvrzení dokázáno ve článku [3]. Nám v dalších podkapitolách bude stačit dříve uvedená konkrétnější varianta, proto se od Conradova článku mírně odkloníme. V našem důkazu avšak budeme postupovat analogicky k důkazu z tohoto článku.

Začneme s porovnáváním prvků v okruhu $\mathbb{Z}[\sqrt{d}]$. Prvky $a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ a $e + f\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ bychom buďto mohli porovnávat jako dvě reálná čísla, nebo pomocí jejich koeficientů. Následující lemma ukazuje, že pro řešení Pellovy rovnice jsou obě varianty ekvivalentní.

Lemma 14. Nechť $x^2 - dy^2 = 1$ je Pellova rovnice a (a,b), (e,f) jsou její dvě nezáporná řešení. Potom platí

a) $a < e$ právě tehdy když $b < f$,

b) $a + b\sqrt{d} < e + f\sqrt{d}$ právě tehdy, když $a < e$ a zároveň $b < f$.

Přičemž nerovností $a + b\sqrt{d} < e + f\sqrt{d}$ myslíme nerovnost reálných čísel.

Důkaz. a) Protože všechna čísla jsou nezáporná, nastává $a < e$ právě tehdy, když $a^2 < e^2$. Jelikož jsou (a, b), (e, f) řešením Pellovy rovnice, je předešlá nerovnost ekvivalentní s nerovností $1+db^2 < 1+df^2$. Tato nerovnost opět díky nezápornosti čísel nastává právě tehdy, když $b < f$.

b) Implikace zprava doleva je zřejmá. U obrácené implikace předpokládejme pro spor, že jedna z nerovností ze závěru neplatí. Z části a) poté vyplývá, že neplatí ani druhá nerovnost ze závěru, tedy nastává $a \geq e$ a zároveň $b \geq f$. Tyto dvě nerovnice dohromady dají $a + b\sqrt{d} \geq e + f\sqrt{d}$, což je ve sporu s předpokladem.

□

Před větou dokazující cykličnost Pellovy grupy si ještě uveďme jedno menší lemma, které nám pomůže v jejím důkaze. Zároveň zadefinujeme pojem *minimální*

řešení, který nám usnadní vyjadřování. Minimální řešení pro nás bude netriviální, nezáporné řešení, které má mezi všemi netriviálními, nezápornými řešeními minimální druhou složku.

Lemma 15. Pokud jsou a, b celá čísla splňující $a^2-db^2=1$ a zároveň $a+b\sqrt{d}>1$, potom jsou a i b kladná.

Důkaz. Rozkladem v reálých číslech dostáváme $(a-b\sqrt{d})(a+b\sqrt{d})=1$. Spolu s předpokladem odsud vyplývá, že $0<a-b\sqrt{d}$ a zároveň $a-b\sqrt{d}<1$. Z první nerovnosti plyne, že $a>b\sqrt{d}$. Sečtením druhé nerovnosti s nerovností z předpokladu získáme $a+b\sqrt{d}+1>a-b\sqrt{d}+1$, což implikuje, že b je kladné. Z nerovnosti $a>b\sqrt{d}$ je tudíž kladné i a.

□

Věta 16. Nechť má Pellova rovnice $x^2-dy^2=1$ nějaké netriviální řešení a buď (e,f) její minimální řešení. Potom $\{(a,b) \mid a+b\sqrt{d}=(e+f\sqrt{d})^n, n\in\mathbb{N}\}$ jsou právě všechna nezáporná řešení $x^2-dy^2=1$.

Důkaz. Všechna $(e+f\sqrt{d})^n$ jsou díky Větě 10 řešení Pellovy rovnice. Zároveň to jsou zřejmě všechno řešení nezáporná, přičemž pro $n=0$ dostáváme řešení triviální. Zbývá ukázat, že žádná jiná nezáporná řešení neexistují.

Buď $a+b\sqrt{d}$ nějaké nezáporné řešení $x^2-dy^2=1$. Minimální řešení má nutně druhou složku větší než triviální řešení, tudíž z Lemmatu 14 plyne $e+f\sqrt{d}>1$. Tím pádem je $((e+f\sqrt{d})^n)_{n=0}^{\infty}$ rostoucí posloupnost reálných čísel konvergující k nekonečnu. Existuje tudíž nezáporné n takové, že

$$(e+f\sqrt{d})^n \le a+b\sqrt{d} < (e+f\sqrt{d})^{n+1}.$$

Vydělením obou nerovností číslem $(e+f\sqrt{d})^n$ dostaneme

$$1 \le (a+b\sqrt{d})(e+f\sqrt{d})^{-n} < e+f\sqrt{d} \tag{2.1}$$

Pro přehlednost označme $o+p\sqrt{d}=(a+b\sqrt{d})(e+f\sqrt{d})^{-n}$. Z Věty 10 je $o+p\sqrt{d}$ taktéž řešením $x^2-dy^2=1$ a dosazením do (2.1) získáme

$$1 \le o+p\sqrt{d} < e+f\sqrt{d}.$$

Z druhé nerovnosti a Lemmatu 14 plyne, že $p<f$. Kdyby byla první z nerovností ostrá, neboli kdyby platilo $1<o+p\sqrt{d}$, pak by podle předchozího lemmatu byly o, p kladná čísla. Dvojice (o,p) by tudíž byla netriviálním, nezáporným řešením $x^2-dy^2=1$ s menší druhou složkou než (e,f). To by byl ovšem spor s minimalitou $e+f\sqrt{d}$.

Musí proto platit $1=o+p\sqrt{d}$ a tudíž $1=(a+b\sqrt{d})(e+f\sqrt{d})^{-n}$. Z poslední rovnosti již snadno plyne $(a+b\sqrt{d})=(e+f\sqrt{d})^n$, jak mělo být dokázáno.

□

Předchozí větu využijeme k popisu množiny nezáporných řešení Pellovy rovnice pomocí rekurentní posloupnosti. Tento pohled na množinu řešení budeme využívat téměř celou následující podkapitolu.

Tvrzení 17. Nechť má Pellova rovnice $x^2 - dy^2 = 1$ netriviální řešení a (e, f) buď její minimální řešení. Pak rekurentně zadaná posloupnost uspořádaných dvojic $(a_n, b_n)_{n=0}^{\infty}$, taková, že

$$a_{n+2} = 2ea_{n+1} - a_n, \quad a_0 = 1, \ a_1 = e$$

$$b_{n+2} = 2eb_{n+1} - b_n, \quad b_0 = 0, \ b_1 = f$$

tvoří právě všechna nezáporná řešení $x^2 - dy^2 = 1$.

Důkaz. Z předchozího věty plyne, že $\{(e + f\sqrt{d})^n\}_{n=0}^{\infty}$ jsou právě všechna nezáporná řešení $x^2 - dy^2 = 1$. Indukcí dokážeme, že pro všechna nezáporná n platí $a_n + \sqrt{d}b_n = (e + f\sqrt{d})^n$, tím bude důkaz hotov.

Pro $n = 0$, $n = 1$ rovnost platí, jelikož $a_0 + b_0\sqrt{d} = 1 = (e + f\sqrt{d})^0$ a $a_1 + b_1\sqrt{d} = (e + f\sqrt{d})^1$.

Buď $n \geq 2$, s využitím definice posloupnosti a indukčního předpokladu pro $n + 1$, n spočteme, že

$$a_{n+2} + b_{n+2}\sqrt{d} = 2ea_{n+1} - a_n + (2eb_{n+1} - b_n)\sqrt{d} = 2e(a_{n+1} + b_{n+1}\sqrt{d}) -$$

$$(a_n + b_n\sqrt{d}) = 2e(e + f\sqrt{d})^{n+1} - (e + f\sqrt{d})^n = (e + f\sqrt{d})^n(2e(e + f\sqrt{d}) - 1).$$

Nyní využijeme fakt, že $e+f\sqrt{d}$ je řešením $x^2-dy^2 = 1$, tudíž platí $e^2-df^2 = 1$. Výraz $2e(e + f\sqrt{d}) - 1$ je proto roven $2e^2 + 2ef\sqrt{d} - e^2 + df^2$, což lze upravit na tvar $(e + f\sqrt{d})^2$. Dosazením do předešlých rovností získáme

$$a_{n+2} + b_{n+2}\sqrt{d} = (e + f\sqrt{d})^n(e + f\sqrt{d})^2 = (e + f\sqrt{d})^{n+2}.$$

Tím je indukční krok dokázán.

□

Na závěr sekce ještě uveďme jedno pozorování vyplývající z důkazu předchozího tvrzení. Toto pozorování budeme často požívat ve zbytku kapitoly. Lemma čerpáme ze článku [5].

Lemma 18. Pro členy posloupnosti definované ve Tvrzení 17 a nezáporná i, j platí

a) $a_{ij} + b_{ij}\sqrt{d} = (a_i + b_i\sqrt{d})^j$,

b) $a_{i+j} = a_ia_j + b_ib_jd$, $b_{i+j} = a_jb_i + a_ib_j$.

Pokud je zároveň $i \geq j$ platí navíc také

c) $a_{i-j} = a_ia_j - b_ib_jd$, $b_{i-j} = a_jb_i - a_ib_j$.

Důkaz. Z důkazu předchozího tvrzení víme, že pro všechna nezáporná n platí $a_n + \sqrt{d}b_n = (e + f\sqrt{d})^n$. S využitím tohoto vztahu spočítáme, že

$$a_{ij} + \sqrt{d}b_{ij} = (e + \sqrt{d}f)^{ij} = ((e + \sqrt{d}f)^i)^j = (a_i + \sqrt{d}b_i)^j,$$

což dokazuje první rovnost. Podobně odvodíme, že

$$a_{i+j} + b_{i+j}\sqrt{d} = (e + \sqrt{d}f)^{i+j} = (e + \sqrt{d}f)^i(e + \sqrt{d}f)^j = (a_i + b_i\sqrt{d})(a_j + b_j\sqrt{d}).$$

Roznásobením závorek a porovnáním koeficientů dostaneme požadované vzorce.

Pro $i \geq j$ bychom obdobně odvodili, že

$$a_{i-j} + b_{i-j}\sqrt{d} = (a_i + b_i\sqrt{d})(a_j + b_j\sqrt{d})^{-1}.$$

Nyní už si jen stačí uvědomit, že (a_j, b_j) je řešením Pellovy rovnice. Platí proto $1 = a_j^2 + db_j^2 = (a_j + b_j\sqrt{d})(a_j - b_j\sqrt{d})$, neboli $(a_j + b_j\sqrt{d})^{-1} = (a_j - b_j\sqrt{d})$. Vzorce pro a_{i-j}, b_{i-j} poté opět získáme roznásobením závorek a porovnáním koeficientů. □

2.3 Diofantičnost řešení Pellovy rovnice

S nabytými znalosti o obecné Pellově rovnici z předchozí podkapitoly se nyní můžeme pustit do studia jedné speciálnější Pellovy rovnice. Tuto rovnici jsme už zmínili na začátku kapitoly, připomeňme, že se jedná o rovnici tvaru

$$x^2 - (a^2 - 1)y^2 = 1, \text{ kde } a > 1. \tag{2.2}$$

Povšimněme si, že pro každé takové a je vždy $a^2 - 1$ nečtvercové a rovnice tím pádem Pellova. Této rovnici se budeme věnovat celou sekci. Nejprve pro ni pomocí tvrzení z předchozí podkapitoly odvodíme rekurentní posloupnost jejích řešení. Poté dokážeme několik užitečných vlastností této posloupnosti. Tyto vlastnosti shrneme do Věty 23 tvrdící, že relace být *n-tou složkou posloupnosti řešení určené rovnicí* (2.2) je diofantická.

Začněme ale zkoumáním rovnice $x^2 - (a^2 - 1)y^2 = 1$. Z předpisu rovnice nahlédneme, že $(a, 1)$ je jejím řešením. Protože je druhá složka tohoto řešení rovna jedné, je toto řešení minimálním řešením rovnice.

Našli jsme minimální řešení a proto můžeme z Tvrzení 17 určit rekurentní posloupnost nezáporných řešení této rovnice. S pomocí zmíněného tvrzení vypočteme, že touto posloupností je

$$x_0(a) = 1, \quad x_1(a) = a, \quad x_{n+2}(a) = 2ax_{n+1}(a) - x_n(a),$$

$$y_0(a) = 0, \quad y_1(a) = 1, \quad y_{n+2}(a) = 2ay_{n+1}(a) - y_n(a).$$

Poznámka. Písmeno a v závorce značí závislost na a. To znamená, že například pro $a = 2$ získáme rovnici $x^2 - 3y^2 = 1$, pro jejíž posloupnost řešení platí

$$x_0(2) = 1,\ x_1(2) = 2,\ x_2(2) = 7,\ y_0(2) = 0,\ y_1(2) = 1,\ y_2(2) = 4.$$

Pokud bude a zřejmé z kontextu nebo nedůležité, budeme jeho psaní v závorce vynechávat.

Až do konce kapitoly pro nás budou $x_n(a)$, $y_n(a)$ $n-$té složky posloupností definovaných výše.

Uvedeme čtyři lemmata shrnující některé vlastnosti platící pro tato $x_n(a)$, $y_n(a)$. Důkazy lemmat většinou nejsou těžké, ale jejich spojením získáme netriviální Větu 23. Taktéž se mnohá z nich budou hodit při důkazu závěrečné věty práce a proto je lepší mít jejich znění zformulovaná. Znění i důkazy všech čtyř lemmat a následné věty jsou až na drobné odlišnosti převzaty z Davisova článku [5].

První z lemmat slouží pro odhady $x_n(a)$, $y_n(a)$. Některé z těchto odhadů jsou zřejmé a zbylé se dají snadno dokázat indukcí. Z těchto důvodu zde jejich důkaz ani neuvádíme.

Lemma 19 (Odhady pro řešení Pellovy rovnice)**.** Pro všechna $a > 1$, n, platí
a) $a^n \leq x_n(a) < x_{n+1}(a)$,
b) $x_n(a) \leq (2a)^n$,
c) $n \leq y_n < y_{n+1}$,
d) $n + y_{n-1}(a) \leq y_n(a)$,
e) $(2a-1)^n \leq y_{n+1}(a) \leq (2a)^n$.

Dále vyslovíme některé základní vlastnosti o soudělnosti $x_n(a)$, $y_n(a)$.

Lemma 20 (Soudělnost řešení Pellovy rovnice)**.** Pro všechna $a > 1$, n, $c \geq 1$ platí
a) $y_n(a) \equiv n \mod (a-1)$,
b) pokud je $a \equiv b \mod c$, potom pro všechna n platí $x_n(a) \equiv x_n(b) \mod c$, $y_n(a) \equiv y_n(b) \mod c$,
c) y_n je sudé právě tehdy, když n je sudé,
d) $\mathrm{NSD}(x_n(a), y_n(a)) = 1$.

Důkaz. Dokážeme pouze části c) a d), první dvě tvrzení se dají snadno dokázat indukcí dle n.

c) Důkaz provedeme také indukcí podle n. Z předpisu posloupnosti vidíme, že y_0 je sudé a y_1 je liché. Tvrzení tedy platí pro $n = 0$, $n = 1$.

V indukčním kroku předpokládejme, že $n > 2$ a tvrzení platí pro n, $n-1$. Z definice posloupnosti je $y_n = 2ay_{n-1} - y_{n-2}$. Podíváme-li se na paritu y_n, zjistíme, že $y_n \equiv y_{n-2} \mod 2$. Z kongruence vidíme, že y_n je sudé právě tehdy, když y_{n-2} je sudé, což z indukčního předpokladu nastane právě tehdy, když $n-2$ je sudé. Poněvadž $n-2$ a n mají zřejmě stejnou paritu, je indukční krok dokázán.

d) Jelikož $x_n(a)$, $y_n(a)$ jsou řešením (2.2), platí pro ně $x_n^2 - (a^2-1)y^2 = 1$. Uvažme libovolného dělitele d čísel $x_n(a)$ a $y_n(a)$. Potom je d nutně dělitelem i čísla $x_n^2 - (a^2-1)y_n^2$. Tento výraz je však roven jedné, jak jsme výše připomněli, tudíž $d \mid 1$, neboli $d = \pm 1$. To dokazuje, že čísla $x_n(a)$, $y_n(a)$ jsou nesoudělná. □

Zbylá dvě lemmata se opět věnují dělitelnosti a kongruentím vlastnostem posloupností. Jejich důkazy jsou však již o dost složitější.

Lemma 21 (Dělitelnost řešení Pellovy rovnice)**.** Pro všechna t, k nezáporná a n kladné platí
a) $y_n \mid y_t$ právě tehdy, když $n \mid t$,
b) $y_{nk} \equiv k y_n x_n^{k-1} \mod y_n^3$,
c) pokud $y_n^2 \mid y_t$, potom $y_n \mid t$.

Důkaz. V důkazu budeme často používat Lemma 18, v našem případě bod b) zmíněného lemmatu dává pro všechna nezáporná u, v vzorec

$$y_{u+v} = x_u y_v + x_v y_u. \tag{2.3}$$

a) Dokažme nejprve následující pozorování, že pro všechna j platí $y_n \mid y_{nj}$. Budeme postupovat indukcí podle j. Pokud je $j = 0$, je i $nj = 0$ a tedy i $y_{nj} = 0$ a tvrzení tudíž platí. Provedeme indukční krok, předpokládejme platnost tvrzení pro j a dokažme ho pro $j + 1$. Podle (2.3) platí $y_{n(j+1)} = x_{nj}y_n + x_n y_{nj}$. Z indukčního předpokladu máme, že $y_n \mid y_{nj}$ a triviálně také platí $y_n \mid y_n$. Ze vzorce pro $y_{n(j+1)}$ poté vyplývá, že $y_n \mid y_{n(j+1)}$, čímž je pozorování dokázáno.

Implikace zprava doleva je nyní díky pozorování zřejmá, stačí si uvědomit, že $t = nj$ pro nějaké nezáporné j.

Dokažme obrácenou implikaci. Předpokládáme, že $y_n \mid y_t$ a nechť $t = qn + r$, kde $0 \leq r < n$. Chceme dokázat, že $r = 0$. K tomu opět využijeme (2.3), platí totiž $y_t = y_{nq+r} = x_r y_{nq} + x_{nq} y_r$. Díky pozorování víme, že $y_n \mid y_{nq}$. V kombinaci s předpokladem a výše uvedeným vyjádřením y_t odvodíme, že $y_n \mid x_{nq} y_r$. Podle bodu d) předchozího lemmatu platí, že $\mathrm{NSD}(x_{nq}, y_{nq}) = 1$. Nyní opět využijeme vztahu $y_n \mid y_{nq}$, který implikuje, že i $\mathrm{NSD}(x_{nq}, y_n) = 1$. Z relace $y_n \mid x_{nq} y_r$ proto plyne $y_n \mid y_r$. Protože však z části c) Lemmatu 19 víme, že $y_n > y_r$, zbývá jediná možnost, a to $y_r = 0$. V tom případě je ale z definice posloupnosti i $r = 0$, přesně jak jsme chtěli ukázat.

b) Z bodu a) Lemmatu 18 plyne pro všechna n, k také vztah

$$(x_{nk} + y_{nk}\sqrt{a^2 - 1}) = (x_n + y_n\sqrt{a^2 - 1})^k.$$

Pravou stranu upravíme pomocí binomické věty a dostaneme

$$x_{nk} + y_{nk}\sqrt{a^2 - 1} = \sum_{j=0}^{k} \binom{k}{j} x_n^{k-j} y_n^j (a^2 - 1)^{\frac{j}{2}}.$$

Porovnáme koeficienty na obou stranách rovnice příslušné $\sqrt{a^2 - 1}$ a získáme

$$y_{nk} = \sum_{\substack{j=0 \\ 2j<k}} \binom{k}{2j+1} x_n^{k-(2j+1)} y_n^{2j+1} (a^2 - 1)^j$$

Nyní už si jen stačí povšimnout, že pro všechna kladná j je exponent u y_n větší než dvě. Všechny členy sumy kromě prvního jsou proto kongruentní nule modulo y_n^3 a platí tedy $y_{nk} \equiv k y_n x_n^{k-1} \mod y_n^3$.

c) Předpokládáme, že $y_n^2 \mid y_t$, což triviálně implikuje, že $y_n \mid y_t$ a proto z části a) plyne $n \mid t$, neboli $t = nk$ pro nějaké k. Z části b) dostáváme $y_t = y_{nk} \equiv k y_n x_n^{k-1} \mod y_n^3$. Tato kongruence jinými slovy tvrdí, že existuje q takové, že $k y_n x_n^{k-1} = y_t + q y_n^3$. Z předpokladu plyne, že y_n^2 dělí pravou stranu, dělí proto i levou. Čili $y_n^2 \mid k y_n x_n^{k-1}$, což implikuje $y_n \mid k x_n^{k-1}$. Protože $\mathrm{NSD}(x_n, y_n) = 1$ (viz předchozí lemma), musí nutně y_n dělit k a tudíž musí dělit i t.

□

Lemma 22 (Periodicita řešení Pellovy rovnice)**.** Nechť i, j, n jsou nezáporná čísla.

a) Pokud $j \leq 2n$, pak platí $x_{2n \pm j} \equiv -x_j \mod x_n$. Pokud $j \leq 4n$, pak platí $x_{4n \pm j} \equiv x_j \mod x_n$.

b) Nechť $x_i \equiv x_j \mod x_n$, $i \leq j \leq 2n$, $n > 0$. Potom $i = j$, až na případ, kdy $a = 2$, $n = 1$, $i = 0$ a $j = 2$.

c) Nechť $x_i \equiv x_j \mod x_n$, $0 < i \leq n$. Potom $j \equiv \pm i \mod 4n$.

Důkaz. a) K důkazu první části využijeme Lemma 18 aplikované postupně na $x_{2n\pm j}$, $x_{2n} = x_{n+n}$ a y_{2n}. S jeho pomocí spočítáme

$$x_{2n\pm j} = x_{2n}x_j \pm y_{2n}y_j = (x_n^2 + y_n^2(a^2-1))x_j \pm 2x_n y_n y_j(a^2-1).$$

Jelikož jsou x_n, y_n řešeními Pellovy rovnice, platí $x_n^2-(a^2-1)y_n^2 = 1$. Tato rovnost implikuje, že $y_n^2(a^2-1)x_j = (x_n^2-1)x_j \equiv -x_j \mod x_n$. Všechny ostatní členy výrazu jsou dělitelné x_n, z čehož vyplývá, že

$$x_{2n\pm j} \equiv -x_j \mod x_n.$$

Druhá část tvrzení se odvodí analogicky.

b) Předpokládejme pro spor, že $x_i \equiv x_j \mod x_n$ a zároveň $i < j$. Z lemmatu 19 vyplývá, že

$$0 < x_0 < x_1 < \ldots < x_{n-1} < x_n \tag{2.4}$$

Čísla $x_0, x_1, x_2, \ldots, x_{n-1}$ jsou tudíž po dvou nekongruentní modulo x_n. Snadno si tím pádem rozmyslíme, že pro platnost $x_i \equiv x_j \mod x_n$ musí být $j > n$. Dále rozlišíme, zda $i \geq n$, nebo $i < n$.

1) $i \geq n$: V části a) jsme dokázali, že $x_{2n-i} \equiv -x_i \mod x_n$ a $x_{2n-j} \equiv -x_j \mod x_n$. Ve spojení s předpokladem to dává $x_{2n-i} \equiv x_{2n-j} \mod x_n$. Protože však $2n-j < n$ a $2n-i \leq n$ a $2n-j \neq 2n-i$, nemůže kvůli (2.4) uvedená kongruence platit.

2) $i < n$: Stále platí $x_{2n-j} \equiv -x_j \mod x_n$ a proto $x_i \equiv -x_{2n-j} \mod x_n$. Předešlá kongruence tvrdí, že $x_n \mid x_i + x_{2n-j}$. Z (2.4) jsou x_i, x_{2n-j} kladná, musí proto platit $x_i + x_{2n-j} \geq x_n$. Ukážeme, že tomu tak není.

Jelikož uvažujeme $j > n$, platí $2n - j < n$. Zároveň je $i < n$, takže z (2.4) můžeme levou stranu předešlé nerovnosti odhadnout

$$x_n \leq x_i + x_{2n-j} \leq x_{n-1} + x_{n-1} = 2x_{n-1}.$$

Dále použijeme bod b) Lemmatu 18 pro x_n, platí

$$x_n = x_{(n-1)+1} = x_{n-1}x_1 + (a^2-1)y_{n-1}y_1 = x_{n-1}a + (a^2-1)y_{n-1} \geq x_{n-1}a \geq 2x_{n-1}.$$

Pohledem na obě uvedené sady nerovností, zjistíme, že všechny nerovnosti v nich musí být rovnostmi. Musí tedy platit

$$x_{n-1}a + (a^2-1)y_{n-1} = x_{n-1}a,\ x_{n-1}a = 2x_{n-1},\ x_i + x_{2n-j} = x_{n-1} + x_{n-1}.$$

Z první rovnosti plyne, že $y_{n-1} = 0$, čili $n = 1$. Z druhé plyne $a = 2$. A konečně ze třetí společně s již dokázaným $n = 1$ plyne $i = 0$ a $j = 2$. To je ovšem právě ten příklad, který jsme v tvrzení vyloučili. Pro všechny ostatní případy jsme dokázali, že tvrzení platí.

c) Předpokládáme, že $x_i \equiv x_j \mod x_n$. Nechť $j = 4nq + r$, $0 \leq r < 4n$. Platí tedy $j \equiv r \mod 4n$. Z části a) je též $x_j \equiv x_r \mod x_n$, tudíž platí $x_i \equiv x_r \mod x_n$. Dále rozlišíme dva případy hodnot, kterých může r nabývat.

1) $0 \leq r \leq 2n$: Pak z části b) kongruence $x_i \equiv x_r \mod x_n$ implikuje $r = i$, kromě oné jedné zmiňované výjimky, která však nenastane. Kdyby totiž nastala, muselo by být $n = 1$ a $r = 0$ (protože i je z předpokladu kladné). Tím pádem by se ale i muselo rovnat dvěma, což je ve sporu s tím, že i je z předpokladu menší nebo rovno než n.

2) $2n < r < 4n$. Položme $s = 4n - r$. Z části a) je $x_s \equiv x_r \mod x_n$, tudíž $x_s \equiv x_i \mod x_n$ a z části b) je $i = s$ (i, s jsou nyní obě kladné a výjimka proto nastat nemůže). V kombinaci s definicí s to dává $r \equiv -s = -i \mod 4n$

V obou případech jsme dokázali, že $r \equiv \pm i \mod 4n$, což spolu s kongruencí $j \equiv r \mod 4n$ dává požadované $j \equiv \pm i \mod 4n$.

□

Konečně jsme připraveni dokázat první větší výsledek o posloupnosti řešení rovnice (2.2). Následující věta tvrdí, že n-tý prvek dané posloupnosti je diofantická množina.

Věta 23 (Diofantičnost řešení Pellovy rovnice)**.** Pro všechna y, $a \geq 2$, $n \geq 1$ nastává $y = y_n(a)$ právě tehdy, když existují nezáporná b, c, d, e, r, s, t, u, v, x taková, že

$$\begin{aligned}
&(1)\ x^2 = (a^2-1)y^2 + 1 && (5)\ b = a + u^2(u^2 - a)\\
&(2)\ u^2 = (a^2-1)v^2 + 1 && (6)\ s = x + cu\\
&(3)\ s^2 = (b^2-1)t^2 + 1 && (7)\ t = n + 4dy\\
&(4)\ v = 4ry^2 && (8)\ y = n + e
\end{aligned}$$

Důkaz. Předpokládejme nejprve, že existují nezáporná čísla splňující rovnice (1) až (8) a dokažme, že pak $y = y_n(a)$.

Prvním pozorování je, že y, t, v jsou kladná. V tvrzení věty předpokládáme, že n je kladné. Z tohoto předpokladu a rovnic (7) (resp. (8)) plyne i kladnost t (resp. y). Pro spor předpokládejme, že $v = 0$. Potom z (2) vyplývá, že $u = 1$, tudíž z (5) $b = 1$ a proto z (3) je i $s = 1$. Z rovnice (6) vyplývá, že $x \leq s = 1$ a z (1) naopak $x \geq 1$. Dohromady tedy $x = 1$ a opět díky (1) musí být $y = 0$. To je ovšem ve sporu s našimi předchozími úvahami.

Dokázali jsme, že čísla y, t, v jsou kladná, pročež dvojice (x,y), (u,v), (s,t) jsou netriviálními řešeními Pellových rovnic ze znění věty. (Povšimněte si, že $b \geq 2$ díky kladnosti v a rovnicím (2), (5). Rovnice (3) je proto taky Pellova.) Podle Tvrzení 17 a definice posloupností $(x_m)_{m=0}^{\infty}$, $(y_m)_{m=0}^{\infty}$ existují kladná čísla i, j, k taková, že

$$x = x_i(a),\ y = y_i(a),\ u = x_k(a),\ v = y_k(a),\ s = x_j(b),\ t = y_j(b). \tag{2.5}$$

Chceme dokázat, že z osmi zadaných rovnic již plyne $i = n$ a tím pádem $y = y_i(a) = y_n(a)$. Postupně dokážeme následující tři pozorování

1) $j \equiv \pm i \mod 4y_i(a)$,

2) $j \equiv n \mod 4y_i(a)$,

3) $n + i < 4y_i(a)$.

Z prvních dvou pozorování pak vyplyne, že $4y_i(a) \mid n + i$ nebo $4y_i(a) \mid n - i$. Díky třetímu pozorování a předpokladu, že n je kladné však první vztah nemůže nastat. Tím pádem musí $4y_i(a) \mid n - i$ a opět díky třetímu pozorování si snadno rozmyslíme, že tento vztah implikuje $i = n$. Tím bude důkaz hotov, zbývá jen dokázat uvedená tři pozorování.

ad 1) Z rovnice (5) je $b \equiv a \mod u$, dosazením za u z (2.5) dostaneme kongruenci $b \equiv a \mod x_k(a)$. Lemma 20 proto implikuje $x_j(b) \equiv x_j(a) \mod x_k(a)$.

Z rovnice (6) je $s \equiv x \mod u$, neboli podle (2.5) $x_j(b) \equiv x_i(a) \mod x_k(a)$. Dohromady tudíž máme $x_i(a) \equiv x_j(a) \mod x_k(a)$. Dle rovnice (4) a kladnosti v je $y \leq v$ a proto díky (2.5) musí být $i \leq k$ (viz též Lemma 19c)). Poněvadž, jak bylo dříve řečeno, je zároveň i kladné, můžeme použít Lemma 22, z kterého plyne $j \equiv \pm i \mod 4k$. Ze (4) a (2.5) také vyplývá, že $(y_i(a))^2 \mid y_k(a)$, tudíž podle Lemmatu 21 $y_i(a) \mid k$. Ve spojení s předchozí kongruencí dostáváme požadované $j \equiv \pm i \mod 4y_i(a)$.

ad 2) V rovnici (5) nahradíme u^2 ekvivalentním výrazem z rovnice (2) a na výslednou rovnici se podíváme modulo v. Výpočtem získáme $b \equiv 1 \mod v$. Jelikož ze (4) a (2.5) $4y_i(a)$ dělí v, máme $b \equiv 1 \mod 4y_i(a)$, neboli $4y_i(a) \mid b-1$. Podle Lemmatu 20 je $y_j(b) \equiv j \mod b-1$, pročež díky předchozí kongruenci $y_j(b) \equiv j \mod 4y_i(a)$. Nakonec rovnice (7) a (2.5) tvrdí $y_j(b) \equiv n \mod 4y_i(a)$. Dohromady tudíž poslední dvě kongruence dávají kýžené $j \equiv n \mod 4y_i(a)$.

ad 3) Poslední pozorování je snadné, neboť z (8) plyne $n \leq y_i(a)$ a z Lemmatu 19 $i \leq y_i(a)$. Dohromady proto platí $n + i \leq 2y_i(a) < 4y_i(a)$.

Nyní naopak předpokládejme, že $y = y_n(a)$ a dokažme existenci nezáporných čísel vyřčených ve znění věty.

Za prvé položme $x = x_n(a)$, čímž splníme (1). Dále položme $m = 4ny_n(a)$, $u = x_m(a)$, $v = y_m(a)$, čímž je splněna (2). Povšimněme si zároveň, že v je kladné. Užitím části b) Lemmatu 21 aplikovaného na $k = y_n(a)$ získáváme

$$y_{ny_n(a)}(a) \equiv y_n(a)^2 x_n(a)^{y_n(a)-1} \mod (y_n(a))^3.$$

Pomocí této kongruence vidíme, že $y_n(a)^2 \mid y_{ny_n}(a)$, tudíž $4y_n(a)^2 \mid 4y_{ny_n}(a)$. Bod b) Lemmatu 18 implikuje

$$y_{4ny_n(a)} = y_{2ny_n(a)+2ny_n(a)} = 2x_{2ny_n(a)}y_{2ny_n(a)} = 4x_{2ny_n(a)}x_{ny_n(a)}y_{ny_n(a)}.$$

Z definice v tudíž plyne $4y_{ny_n(a)} \mid v$. Díky tranzitivitě dělitelnosti z předchozích úvah plyne, že $4y^2$ dělí v. Číslo $r = \frac{v}{4y^2}$ je proto kladné celé číslo pomocí něhož je splněna (4).

Z rovnice (2), kladnosti v a předpokladu věty je zřejmě $u^2 \geq a$, tudíž $b = a + u^2(u^2 - a)$ je nezáporné číslo splňující (5). Dále položme $s = x_n(b)$, $t = y_n(b)$, čímž je splněna (3). (Stejně jako v opačné implikaci platí, že $b \geq 2$). Z již splněných rovnic (2) a (5) zjistíme stejně jako v důkazu předchozí implikace, že $b \equiv a \mod u$. Pomocí Lemmatu 20 z tohoto pozorování plyne, že $x_n(b) \equiv x_n(a) \mod u$, neboli $s \equiv x \mod u$. Můžeme tedy nalézt c splňující rovnici (6).

Z definice t a bodu c) Lemmatu 19 je $t \geq n$ a z Lemmatu 20 je $t \equiv n \mod (b-1)$. Stejně jako v opačné implikaci jsme schopni ukázat, že $4y \mid b-1$. Dohromady tudíž předešlé úvahy implikují $4y \mid (n-t)$. Existuje proto d splňující (7).

Konečně z Lemmatu 19 plyne $y \geq n$ a proto je $y-n$ nezáporné číslo. Položíme tudíž $e = y - n$ a tím splníme i poslední zbývající rovnici (8).

□

Pomocí Lemmatu 2 a předešlé věty bychom již snadno mohli dokázat diofantičnost množiny obsahující právě $n-$tou složku posloupnosti řešení Pellovy rovnice (2.2).

Ve znění věty jsme pro jednoduchost uvedli osm rovnic. Kvůli ušetření místa a proměnných budeme dále používat její ekonomičtější zápis. Eliminací proměnných v, b, s, t ze znění věty totiž lehce získáme následující důsledek.

Důsledek. *Pro všechna y, $a \geq 2$, $n \geq 1$ nastává $y = y_n(a)$ právě tehdy, když existují nezáporná c, d, e, r, u, x taková, že*

$$x^2 = (a^2-1)y^2 + 1 \qquad u^2 = 16(a^2-1)r^2y^4 + 1$$
$$(x+cu)^2 = ((a+u^2(u^2-a^2))^2 - 1)(n+4dy)^2 + 1 \qquad y = n + e.$$

2.4 Prvočísla a faktoriál

Dostáváme se k předfinální fázi našeho snažení. V této podkapitole dokážeme dvě velká tvrzení, o kterých již byla řeč v úvodu kapitoly. Jedná se o Wilsonovu větu a Větu 28. Wilsonova věta staví do relace prvočísla s faktoriálem a druhá zmiňovaná věta pro změnu faktoriál s polynomy. Ve Větě 30 nám jen zbude tyto dvě věty šikovně spojit, přičemž toto „spojení" zajistí výsledky o Pellově rovnici z předchozích sekcí.

Připomeňme ještě jednou Wilsonovu větu, nyní i s jejím důkazem.

Věta 24 (Wilsonova)**.** Číslo $k > 1$ je prvočíslo právě tehdy, když k dělí $(k-1)!+1$.

Důkaz. Předpokládejme nejdříve, že k je složené číslo. Nechť $k = jl$ je nějaký netriviální rozklad k. Speciálně tedy platí $1 < j < k$, pročež j dělí $(k-1)!$ a tudíž j nedělí $(k-1)!+1$. Tím pádem však ani k nemůže dělit $(k-1)!+1$.

Nechť je nyní k prvočíslo. Pro dvojku tvrzení platí, dále proto budeme předpokládat, že k je liché prvočíslo. Skutečnost, že k dělí $(k-1)!+1$ je ekvivalentní tomu, že součin všech prvků grupy $\mathbb{Z}_k^*$ (tj. multiplikativní grupy tělesa $\mathbb{Z}_k$) je v této grupě roven mínus jedné. Grupa $\mathbb{Z}_k^*$ je cyklická (viz např. skripta [4], Věta 15.9), neboli existuje $a \in \mathbb{Z}_k^*$, které je generátorem této grupy. Součin všech prvků $\mathbb{Z}_k^*$ tudíž můžeme díky komutativitě násobení v $\mathbb{Z}_k^*$ vyjádřit jako

$$a^1 \cdot a^2 \cdot \ldots \cdot a^{k-1} = a^{\frac{k(k-1)}{2}} = (a^{k-1})^{\frac{k-1}{2}} \cdot a^{\frac{k-1}{2}}.$$

(V první rovnosti jsem využili vzorec pro součet prvních $k-1$ přirozených čísel.)

Prvek a má, jakožto primitivní prvek $\mathbb{Z}_k^*$, řád roven $k-1$. Proto je první činitel v součinu roven jedné v $\mathbb{Z}_k^*$ a řád druhého činitele v součinu je roven dvěma. Prvek řádu dva v $\mathbb{Z}_k^*$ je ovšem jen jeden a to -1. Součin všech prvků $\mathbb{Z}_k^*$ je proto vskutku roven mínus jedné, jak jsme chtěli dokázat. Spolu s diskuzí výše tudíž pro každé prvočíslo k platí, že k dělí $(k-1)!+1$.

□

Důkaz druhé věty bude o něco komplikovanější a budou k němu potřeba dvě následující tvrzení. První z nich je opět dost spjaté s Pellovou rovnicí. Důkazy obou tvrzení i následné věty jsou až na malé výjimky přebrány z již zmiňovaného článku [1].

Tvrzení 25. Pro každá s, $e \geq 2$ taková, že

$$e^3(e+2)(n+1)^2 + 1 = s^2, \tag{2.6}$$

platí $e^{e-2} + e - 1 \leq n$. Zároveň pro všechna kladná e, t existují nezáporná n, s taková, že platí (2.6) a zároveň $t < n$.

Důkaz. Položme $a = e + 1$. Rovnice (2.6) se pak transformuje v Pellovu rovnici $s^2 - (a^2-1)((a-1)(n+1))^2 = 1$. Povšimněme si, že $a > 2$, tudíž pro něj platí

$$1 + \frac{a-2}{(a-1)^{a-1}} < 2 < \left(2 + \frac{1}{a-1}\right)^{a-2} = \left(\frac{2a-1}{a-1}\right)^{a-2}.$$

Vynásobením obou stran nerovnice číslem $(a-1)^{a-2}$ dostaneme ekvivalentní nerovnici $(a-1)^{a-2} + (a-2)(a-1) < (2a-1)^{a-2}$. Lemma 19 část e) tvrdí, že $(2a-1)^{a-2} \leq y_{a-1}(a)$. Číslo $(a-1)(n+1)$ je kladné, tudíž z Tvrzení 17 a definice posloupnosti $(y_i)_{i=0}^{\infty}$ existuje kladné j takové, že $(a-1)(n+1) = y_j(a)$. Dále z bodu a) Lemmatu 20 plyne, že $(a-1)(n+1) \equiv j \mod a-1$, neboli $a-1 \mid j$. Ve spojení s kladností j to implikuje nerovnost $a-1 \leq j$, tudíž $y_{a-1}(a) \leq y_j(a)$ (viz též bod c) Lemmatu 19). Celkově získáváme následující sadu nerovností

$$(a-1)^{a-2} + (a-2)(a-1) < (2a-1)^{a-2} \leq y_{a-1}(a) \leq y_j(a) = (a-1)(n+1).$$

Vydělením obou stran nerovnice číslem $a-1$ a dosazením e zpět za $a-1$ dostaneme $e^{e-2} + e - 1 < n + 1$.

Zbývá ukázat existence n, s z druhé části věty. Uvažujme stále $a = e + 1$ a tedy i Pellovu rovnici $s^2 - (a^2-1)((a-1)(n+1))^2 = 1$. Podle Věty 13 existuje netriviální řešení této rovnice. Dále proto podle Věty 16 existuje nekonečná rostoucí posloupnost řešení této rovnice. Musí proto existovat řešení $(s, (a-1)(n+1))$, pro které je $n > t$.

□

Druhé tvrzení je dost technické. V jeho důkaze použijeme následující pozorování, které je snadno dokazatelné pomocí elementární matematické analýzy.

Lemma 26. Nechť $\alpha \in \mathbb{R}$, $q \in \mathbb{N}$, $q > 1$.

a) Pokud je $0 \leq \alpha < \frac{1}{q}$, potom $1 - q\alpha \leq (1-\alpha)^q$.

b) Pokud je $0 \leq \alpha \leq \frac{1}{2}$, potom $(1-\alpha)^{-1} \leq 1 + 2\alpha$.

Důkaz. a) Buď q pevné a položme $f(\alpha) = (1-\alpha)^q - (1-q\alpha)$, $\alpha \in [0, \frac{1}{q})$. K důkazu nerovnosti ze znění lemmatu stačí ukázat, že tato funkce je pro všechna α ze svého definičního oboru nezáporná.

Pro $\alpha = 0$ je $f(\alpha) = 0$ a zároveň $f(\alpha)$ je neklesající: $f'(\alpha) = q(1-(1-\alpha)^{q-1})$ a snadno si lze rozmyslet, že pro všechna α z definičního oboru funkce f je tento výraz nezáporný. Tím pádem je $f(\alpha) \geq 0$ pro všechna $\alpha \in [0, \frac{1}{q})$.

b) Zadaná nerovnice je ekvivalentní s nerovnicí $1 \leq (1+2\alpha)(1-\alpha)$. Poslední nerovnici lze ještě upravit na tvar $\alpha(1-2\alpha) \geq 0$, odkud snadno nahlédneme její platnost pro všechna $\alpha \in [0, \frac{1}{2}]$.

□

V samotném tvrzení budeme pracovat s funkcí $r(a,b)$. Tato funkce vrací zbytek a po dělení b, neboli $r(a,b) = a \mod b$.

Tvrzení 27. Pro kladná k, n, p taková, že $(2k)^k \leq n$ a zároveň $n^k < p$ platí

$$k! < \frac{(n+1)^k p^k}{r((p+1)^n, p^{k+1})} < k! + 1.$$

Důkaz. Nejprve ukážeme, že

$$r((p+1)^n, p^{k+1}) = \sum_{i=0}^{k} \binom{n}{i} p^i.$$

Z binomické věty plyne $(p+1)^n = \sum_{i=0}^{k} \binom{n}{i} p^i + \sum_{i=k+1}^{n} \binom{n}{i} p^i$. Druhá suma zřejmě nechává zbytek nula po vydělení p^{k+1}. Stačí dokázat, že první suma je menší než p^{k+1}. S využitím vzorce pro součet geometrické řady platí

$$\sum_{i=0}^{k} \binom{n}{i} p^i \leq \sum_{i=0}^{k} n^i p^i = \frac{(np)^{k+1} - 1}{np - 1}.$$

Vzorec jsme mohli využít, protože z předpokladů tvrzení je $np > 1$. Z předpokladů též odvodíme následující sadu nerovností

$$n^k p^k np - 1 \leq (p-1)p^k np - 1 = p^{k+1} np - p^{k+1} n - 1 < p^{k+1} np - p^{k+1} = (np-1)p^{k+1}.$$

Snadnou úpravou dostaneme, že $\frac{(np)^{k+1}-1}{np-1} < p^{k+1}$. To spolu s předchozími nerovnicemi dokazuje úvodní pozorování. Speciálně je $r((p+1)^n, p^{k+1})$ nenulové a tvrzení tedy dává smysl.

Můžeme se proto pustit do důkazu obou uvedených nerovnic, přičemž začneme s horním odhadem pro $k!$. Z předpokladů tvrzení je vcelku snadné nahlédnout, že pro všechna $i \in \{0, 1, \ldots, k-2\}$ platí $\binom{n}{i} p^i < \binom{n}{i+1} p^{i+1}$. Proto platí

$$\sum_{i=0}^{k} \binom{n}{i} p^i \leq k \binom{n}{k-1} p^{k-1} + \binom{n}{k} p^k \leq k \frac{n^{k-1}}{(k-1)!} p^{k-1} + \frac{n^k}{k!} p^k.$$

Po přenásobení $k!$ získáme

$$k! \sum_{i=0}^{k} \binom{n}{i} p^i \leq k^2 n^{k-1} p^{k-1} + n^k p^k.$$

Výraz na pravé straně odhadneme pomocí předpokladů tvrzení

$$k^2 n^{k-1} p^{k-1} + n^k p^k < k n^k p^{k-1} + n^k p^k < k p p^{k-1} + n^k p^k = p^k (k + n^k).$$

Jelikož je $k + n^k \leq (n+1)^k$, dostáváme spojením všech předchozích nerovnic

$$k! \sum_{i=0}^{k} \binom{n}{i} p^i < p^k (n+1)^k,$$

což je díky úvodnímu pozorování ekvivalentní s uvedeným horním odhadem pro $k!$.

K důkazu druhé nerovnosti nejprve odhadneme sumu kombinančních čísel ve jmenovateli, dostaneme

$$\frac{(n+1)^k p^k}{\sum_{i=0}^{k} \binom{n}{i} p^i} = \frac{(n+1)^k}{\sum_{i=0}^{k} \binom{n}{i} p^{i-k}} < \frac{(n+1)^k}{\binom{n}{k}}.$$

Z definice kombinačního čísla můžeme dále psát

$$\frac{(n+1)^k}{\binom{n}{k}} \leq \frac{k!}{\frac{(n+1-k)^k}{(n+1)^k}} = \frac{k!}{(1-k/(n+1))^k} < \frac{k!}{(1-k/n)^k}.$$

Nyní dvakrát použijeme předchozí lemma k odhadu jmenovatele. Pokud totiž položíme $\alpha = \frac{k}{n}$ a $q = k$, potom z části a) plyne $(1-k/n)^k \geq 1 - k^2/n$. Dále z části b) aplikované na $\alpha = \frac{k^2}{n}$ vyplývá $(1-k^2/n)^{-1} \leq 1 + 2k^2/n$. Nakonec využijeme předpoklad tvrzení a snadno ověřitelný vztah $2^{k-1}k^{k-2} \geq k!$ a získáme $2k^2/n \leq 2k^2/(2k)^k \leq 1/k!$. Dohromady všechny odvozené nerovnosti implikují $((1-k/n)^k)^{-1} \leq 1 + 1/k!$. Spolu s nerovnostmi výše dostáváme

$$\frac{(n+1)^k p^k}{\sum_{i=0}^{k} \binom{n}{i} p^i} < \frac{(n+1)^k}{\binom{n}{k}} < k!(1 + 1/(k!)) = k! + 1,$$

což je opět díky úvodnímu pozorování ekvivalentní s uvedeným dolním odhadem pro $k! + 1$.

□

Věta 28. Pro kladná f, k nastává $f = k!$ právě tehdy, když existují nezáporná čísla h, j, n, o, p, q, w, z taková, že

(1) $q = wz + h + j$,
(2) $z = f(h + j) + h$,
(3) $o^2 = (2k)^3(2k + 2)(n + 1)^2 + 1$,
(4) $p = (n + 1)^k$,
(5) $q = (p + 1)^n$,
(6) $z = p^{k+1}$.

Důkaz. Předpokládejme nejdříve existenci čísel splňujících šest zadaných rovnic. Ze třetí rovnice a Tvrzení 25 plyne, že $(2k)^{2k-2} + 2k - 1 \leq n$, pomocí čehož snadno odvodíme, že $(2k)^k \leq n$. Dále z rovnice (4) vyplývá $n^k < p$, tudíž z Tvrzení 27 plyne

$$k! < \frac{(n+1)^k p^k}{r((p+1)^n, p^{k+1})} < k! + 1.$$

Poněvadž je z rovnic (4) a (6) $z = (n+1)^k p^k$, můžeme výše uvedený vztah přepsat do tvaru

$$k! < \frac{z}{r(q, z)} < k! + 1.$$

Pokud by platilo $h + j < z$, pak bychom z rovnice (1) dostali $r(q, z) = h + j$. Předpokládejme pro spor, že rovnost $h + j < z$ neplatí, neboli $h + j \geq z$. V tom případě je ale z (2) a kladnosti f nutně $h + j = z$. Tudíž z (1) vyplývá, že $z \mid q$, neboli $p^{k+1} \mid (p + 1)^n$. V důkaze Tvrzení 27 jsme však také ukázali, že

zbytek $(p+1)^{n+1}$ po dělení číslem p^{k+1} je za daných předpokladů nenulový, čímž dostáváme spor. Platí tedy $r(q,z) = h + j$ a tudíž

$$k! < \frac{z}{h+j} < k! + 1. \tag{2.7}$$

Vydělením rovnice (2) číslem $h + j$, získáme $f = \frac{z}{h+j} - \frac{h}{h+j}$. Poněvadž (2.7) implikuje též, že $\frac{z}{h+j}$ není celé číslo, musí platit

$$f < \frac{z}{h+j} < f + 1. \tag{2.8}$$

Čísla f a $k!$ jsou celá, tudíž z (2.7) a (2.8) vyplývá $f = k!$.

Předpokládejme nyní naopak, že $f = k!$. Díky Tvrzení 25 můžeme najít n, o splňující (3), pro která zároveň platí $(2k)^k \leq n$. Položme $p = (n+1)^k$, $q = (p+1)^n$ a $z = p^{k+1}$. Tím jsou splněny rovnice (4), (5) a (6). Zároveň pro k, n, p jsou splněny předpoklady Tvrzení 27, které tudíž implikuje

$$k! < \frac{z}{r(q,z)} < k! + 1, \text{ neboli } fr(q,z) < z < (f+1)r(q,z).$$

Položme $h = z - fr(q,z)$, $j = (f+1)r(q,z) - z$. Z výše uvedených nerovností vidíme, že h, j jsou kladná čísla, pro něž dále platí $h+j = r(q,z)$ a $f(h+j)+h = z$. Tím pádem je splněna rovnice (2) a zároveň lze nalézt w splňující poslední zbývající rovnici, tj. rovnici (1).

□

2.5 Množina prvočísel je diofantická

V této podkapitole už zbývá učinit poslední krok, a to dokázat, že samotná množina prvočísel je diofantická. Většinu práce pro důkaz tohoto tvrzení jsme již učinili v minulých podkapitolách. Poslední věc, která se nám bude v důkazu věty hodit, je následující lemma. Znění věty i lemmatu jsou převzaty opět z článku [1]. Náš důkaz věty kopíruje důkaz z tohoto textu, lemma je v původním textu ponecháno bez důkazu.

Lemma 29. Pro všechna p, n, $a > 1$ platí

$$x_n(a) \equiv p^n + (a-p)y_n(a) \mod (2ap - p^2 - 1).$$

Pokud navíc $0 < p^n < a$, potom je $x_n(a) \geq p^n + (a-p)y_n(a)$.

Důkaz. Požadovanou kongruenci dokážeme indukcí podle n. Pro $n = 0$ a $n = 1$ se dané výrazy dokonce rovnají a kongruence tudíž platí. Předpokládejme platnost kongruence pro n, $n-1$ a dokažme ji pro $n+1$. Z indukčního předpokladu a vzorce pro x_{n+1} odvozeného na začátku sekce 2.3, tj. $x_{n+1}(a) = 2ax_n(a) - x_{n-1}(a)$, plyne

$$x_{n+1}(a) \equiv 2a(p^n + (a-p)y_n(a)) - (p^{n-1} + (a-p)y_{n-1}(a)) \mod (2ap - p^2 - 1).$$

Výraz na pravé straně upravíme na $p^{n-1}(2ap-1)+(a-p)(2ay_n(a)-y_{n-1}(a))$. Platí $p^2 \equiv 2ap-1 \mod (2ap-p^2-1)$, tudíž společně se vzorcem pro y_{n+1} a předešlou kongruencí dostáváme

$$x_{n+1}(a) \equiv p^{n+1}+(a-p)y_{n+1}(a) \mod (2ap-p^2-1).$$

Tím je proveden indukční krok a kongruence tudíž platí pro všechna n.

Předpokládejme nyní, že navíc $0 < p^n < a$ a dokažme uvedenou nerovnost. Pro $n = 0$ a $n = 1$ jsme již výše ukázali, že se výrazy rovnají a nerovnosti tedy platí. Předpokládejme proto dále $n \geq 2$. Připomeňme, že platí vztah $x_n^2 = (a^2-1)y_n^2+1$ a x_n, y_n jsou nezáporná. Z těchto pozorování a předpokladu lemmatu plyne

$$x_n(a) > \sqrt{a^2-1}y_n(a) \geq (a-1)y_n(a) > (a-1)y_n(a)+p^n-a.$$

Ve výše dokázaných nerovnostech lze spatřit, že $x_n(a) \geq (a-1)y_n(a)+1$, což dokazuje nerovnost pro $p = 1$. Dále proto předpokládejme $p > 1$. Bod e) Lemmatu 19 tvrdí, že $y_n(a) \geq (2a-1)^{n-1}$. Jelikož je a kladné, $p > 1$ a $n \geq 2$, platí i $(p-1)y_n(a) \geq a$, neboli $-a-y_n(a) \geq -py_n(a)$. Z tohoto pozorování tudíž plyne

$$ay_n(a)-y_n(a)+p^n-a \geq ay_n(a)-py_n(a)+p^n = p^n+(a-p)y_n(a).$$

Ve spojení s předchozími nerovnostmi dostáváme požadované tvrzení.

□

Shrnutím všech poznatků z této kapitoly jsme teď schopni konstruktivně dokázat diofantičnost množiny prvočísel a tedy i hlavní větu této práce.

Věta 30. Pro každé kladné k je $k+1$ prvočíslo právě tehdy, když existují a, b, c, d, e, f, g, h, i, j, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z taková, že

(1) $q = wz+h+j$,
(2) $z = (gk+g+k)(h+j)+h$,
(3) $f^2 = (2k)^3(2k+2)(n+1)^2+1$,
(4) $e = p+q+z+2n$,
(5) $o^2 = e^3(e+2)(a+1)^2+1$,
(6) $x^2 = (a^2-1)y^2+1$,
(7) $u^2 = 16(a^2-1)r^2y^4+1$,
(8) $l = k+i(a-1)$,
(9) $m^2 = (a^2-1)l^2+1$,
(10) $n+l+v = y$,

(11) $(x+cu)^2 = ((a+u^2(u^2-a^2))^2-1)(n+4dy)^2+1$,
(12) $m = p+l(a-n-1)+b(2a(n+1)-(n+1)^2-1)$,
(13) $x = q+y(a-p-1)+s(2a(p+1)-(p+1)^2-1)$,
(14) $pm = z+pl(a-p)+t(2ap-p^2-1)$.

Důkaz. Předpokládejme nejdříve existenci čísel splňujících rovnice (1) až (14). Z rovnice (3) a Tvrzení 25 plyne, že $2k-1+(2k)^{2k-2} \leq n$. Speciálně díky této nerovnosti a kladnosti k platí $k < n$, tedy také $1 < n$. Z rovnice (4) je proto $e \geq 2$ a tudíž opět z Tvrzení 25 a rovnice (5) plyne $e-1+e^{e-2} \leq a$. Dosazením za e z rovnice (4) získáme nerovnost

$$p+q+z+2n-1+(p+q+z+2n)^{p+q+z+2n-2} \leq a \tag{2.9}$$

Poněvadž je díky této nerovnosti zřejmě $a > n > 1$, vyplývá z důsledku za Větou 23 a rovnic (6), (7), (10), (11), že $y = y_n(a)$, $x = x_n(a)$.

Z rovnice (9) vidíme, že $m = x_{k'}(a)$ a $l = y_{k'}(a)$ pro nějaké nezáporné k'. Bod a) Lemmatu 20 implikuje $l \equiv k' \mod (a-1)$. Z (8) je také $l \equiv k \mod (a-1)$, tudíž $k \equiv k' \mod (a-1)$. Pomocí rovnice (10) nahlédneme, že $l < y$, neboť n je kladné. Tím pádem je díky Lemmatu 19 c) $k' < n$. Již dříve jsme dokázali, že $k < n$ a $n < a$, tudíž platí $k < a-1$ a $k' < a-1$. Tyto dvě nerovnice spolu s výše uvedenou kongruencí implikují $k = k'$, neboli $m = x_k(a)$, $l = y_k(a)$.

V dalších třech odstavcích ukážeme, že $p = (n+1)^k$, $q = (p+1)^n$ a $z = p^{k+1}$. V důkazu všech tří rovností budeme postupovat naprosto analogicky. Nejdříve pomocí Lemmatu 29 zjistíme, že uvažovaná dvě čísla jsou kongruentní modulo nějaké příhodné třetí číslo. Poté díky několika nerovnostem odhadneme, že toto třetí číslo je větší než obě uvažovaná čísla. Dohromady tato dvě pozorování dají kýženou rovnost uvažovaných čísel.

Díky Lemmatu 29 platí $m \equiv (n+1)^k + (a-n-1)l \mod (2a(n+1) - (n+1)^2 - 1)$. Ve spojení s (12) je proto $p \equiv (n+1)^k \mod (2a(n+1) - (n+1)^2 - 1)$. Víme, že $n > k > 0$, z čehož díky (2.9) snadno odvodíme, že $p < a$ a $(n+1)^k < a$. Zároveň z (2.9) vyplývá $(n+1)^2 < a$, pročež $a \leq 2a - (n+1)^2 - 1 < 2a(n+1) - (n+1)^2 - 1$. Všechny nerovnosti společně s kongruencí dávají $p = (n+1)^k$.

Díky Lemmatu 29 platí $x \equiv (p+1)^n + (a-p-1)y \mod (2a(p+1) - (p+1)^2 - 1)$. Ve spojení s (13) je proto $q \equiv (p+1)^n \mod (2a(p+1) - (p+1)^2 - 1)$. Z (2.9) je opět vidět, že $q < a$ a $(p+1)^n < a$. Poslední nerovnice také implikuje $(p+1)^2 < a$, odkud analogicky jako v předchozím odstavci odvodíme $a < 2a(p+1) - (p+1)^2 - 1$. Všechny nerovnosti pak společně s kongruencí dávají $q = (p+1)^n$.

Ještě jednou díky Lemmatu 29 platí $m \equiv p^k + (a-p)l \mod (2ap - p^2 - 1)$. Tím pádem platí $pm \equiv p^{k+1} + p(a-p)l \mod (2ap - p^2 - 1)$. Ve spojení se (14) je proto $z \equiv p^{k+1} \mod (2ap - p^2 - 1)$. Dále z (2.9) plyne $z < a$ a $p^{k+1} < a$. Tyto nerovnosti implikují $a \leq 2a - p^2 - 1 \leq 2ap - p^2 - 1$. (Poslední nerovnost platí, protože p je kladné, neboť $p = (n+1)^k$.) Stejnou úvahou jako v předchozích dvou odstavcích dostáváme $z = p^{k+1}$.

Výše dokázané tři rovnosti spolu s rovnicemi (1), (2), (3) a Větou 28 implikují $k! = gk + g + k$. Rovnost upravíme do tvaru $k! + 1 = (g+1)(k+1)$, z kterého vyplývá, že $k+1 \mid k!+1$. Z Wilsonovy věty pak plyne, že $k+1$ musí být prvočíslo.

Předpokládejme nyní, že $k+1$ je prvočíslo. Z Wilsonovy věty plyne, že $(k+1)$ dělí $k! + 1$. Jinými slovy existuje nezáporné g takové, že $(g+1)(k+1) = k! + 1$, neboli $gk + g + k = k!$. Poté dle Věty 28 existují nezáporná f, h, j, n, p, q, w, z splňující (1), (2), (3) a zároveň rovnosti $p = (n+1)^k$, $q = (p+1)^n$, $z = p^{k+1}$.

Položme $e = p + q + z + 2n$, čímž bude splněna (4). Poněvadž $e \geq 2$, plyne z Tvrzení 25 existence čísel $a > 1$, o splňujících (5). Díky tomu, že $a > 1$ můžeme položit $y = y_n(a)$. Podle důsledku za Větou 23 potom existují c, d, l', r, u, x splňující (6), (7), (11), pro něž zároveň platí $n + l' = y$.

Dále položme $m = x_k(a)$, $l = y_k(a)$, tím je splněna (9). Z bodu a) Lemmatu 20 plyne $y_k(a) \equiv k \mod a-1$ a z bodu c) Lemmatu 19 plyne $y_k(a) \geq k$. Díky těmto pozorováním je možno nalézt i splňující (8). Lemma 19 d) tvrdí, že $n + y_{n-1}(a) \leq y_n(a)$. V minulé implikaci jsme ukázali, že z (3) plyne $k < n$. To spolu s předchozí nerovností značí, že $n + y_k(a) \leq n + y_{n-1}(a) \leq y_n(a)$. (V první nerovnosti jsme využili část c) Lemmatu 19). Nerovnost tedy tvrdí, že $n + l \leq y = n + l'$, tudíž je možné nalézt v splňující $l + v = l'$ a splnit tak (10).

Důkaz existence zbývajících čísel b, s, t se opírá o stejné úvahy, jaké jsme vyu-

živali v předešlé implikaci při důkazech $p = (n+1)^k$, $q = (p+1)^n$ a $z = p^{k+1}$. Pro názornost předvedeme důkaz existence b splňující (12), zbylé dva případy se řeší analogicky. Na začátku důkazu této implikace jsme ukázali, že $p = (n+1)^k$. Dle Lemmatu 29 je $m \equiv (n+1)^k + (a-n-1)l \mod (2a(n+1) - (n+1)^2 - 1)$. Z rovnic (4) a (5) plyne $(n+1)^k < a$ (viz důkaz předchozí implikace), tudíž Lemma 29 též tvrdí, že $m \geq (n+1)^k + (a-n-1)l$. To spolu se zmíněnou kongruencí zaručuje existenci b splňujícího rovnici (12). Podobně se dokáže existence s splňující (13) a t splňující (14). Tím je dokázána požadovaná existence čísel splňujících daných čtrnáct rovnic.

□

Pomocí Lemmatu 2 a substituce $k+1$ místo k získáme následující důsledek.

Důsledek. *Množina prvočísel je diofantická.*

2.6 Polynom reprezentující prvočísla

Na závěr shrneme veškeré poznatky a explicitně vyjádříme polynom reprezentující prvočísla. Na konci minulé sekce jsme uvedli čtrnáct rovnic o 26 proměnných, které jsou řešitelné v nezáporných číslech právě tehdy když $k+1$ je prvočíslem. Přesuneme-li všechny členy rovnic vždy na jednu stranu dané rovnice, získáme na těchto „nenulových stranách" čtrnáct polynomů v 26 proměnných. Klasickým trikem pomocí součtu čtverců těchto čtrnácti polynomů (formálněji uvedeným v Lemmatu 2) získáme jediný polynom řešitelný právě pro prvočíselné hodnoty $k+1$. Zároveň zřejmě platí, že tento polynom nabývá pouze nezáporných hodnot. Jak z takového polynomu zkonstruovat polynom reprezentující prvočísla jsme již diskutovali na začátku kapitoly. Posledním menším zádrhelem je, že jsme v předchozí větě předpokládali k kladné a přitom za proměnné dosazujeme obecně nezáporné hodnoty. Tento zádrhel však snadno vyřešíme nahrazením k proměnnou $k+1$. Tím získáme polynom řešitelný tehdy a jen tehdy, když $k+2$ je prvočíslo, přesně jako jsme tvrdili na začátku kapitoly. Naším polynomem reprezentující prvočísla tudíž je

$$(k+2)\{1 - [wz + h + j - q]^2 - [(gk + 2g + k + 1)(h + j) + h - z]^2 - [2n + p + q + z - e]^2 - [16(k+1)^3(k+2)(n+1)^2 + 1 - f^2]^2 - [e^3(e+2)(a+1)^2 + 1 - o^2]^2 - [(a^2-1)y^2 + 1 - x^2]^2 - [16r^2y^4(a^2-1) + 1 - u^2]^2 - [((a + u^2(u^2 - a))^2 - 1)(n + 4dy)^2 + 1 - (x + cu)^2]^2 - [n + l + v - y]^2 - [(a^2-1)l^2 + 1 - m^2]^2 - [ai + k + 1 - l - i]^2 - [p + l(a-n-1) + b(2an + 2a - n^2 - 2n - 2) - m]^2 - [q + y(a-p-1) + s(2ap + 2a - p^2 - 2p - 2) - x]^2 - [z + pl(a-p) + t(2ap - p^2 - 1) - pm]^2\}.$$

Tím je náš úkol hotov. Problém polynomů reprezetnujících prvočísla tím ale nekončí. Otevřenými problémy zůstavají například minimální počet proměnných potřebných ke konstrukci polynomu nebo minimální stupeň polynomu. V oblasti proměnných bylo dokázáno, že existuje polynom reprezentující prvočísla o 12 proměnných. Důkaz tohoto faktu je postaven na metodách Juriho Matijaseviče a Julie Robinsonové a je ho možno nalézt v [1]. Co se týče stupně, je znám polynom reprezentující prvočísla stupně 5. Tento polynom není těžké sestrojit, stačí využít takzvanou Skolemovu substituční metodu rozebíranou například v [5]. O této

metodě už zde hovořit nebudeme, jen uveďme, že pro náš polynom by její použití způsobilo vzrůst proměnných z 26 na 42.

Závěr

V první kapitole jsme stručne nastínili, co to jsou rekursivní, rekursivně vyčíslitelné a diofantické množiny. Uvedli jsme několik snadných pozorování, která spíše sloužila jako menší vsuvka před kapitolou druhou.

V té už jsme se věnovali důkazu diofantičnosti mnohých množin. Vrcholem práce byl konstruktivní důkaz diofantičnosti množiny prvočísel a nalezení polynomu reprezentujícího prvočísla. Vycházeli jsme přitom především z článku [1]. Snažili jsme se ale postupovat více podrobněji a pro případného čtenáře přívětivěji. Doplnili jsme obecný úvod o Pellově rovnici, některé snazší kroky, které jsou ve článku [1] vynechány a podrobněji jsme vysvětlili a motivovali kroky obtížnější.

Na závěr jsme pak ještě uvedli několik otevřených otázek ohledně polynomů reprezentujících prvočísla. Jednalo se o problémy ohledně maximálního stupně polynomu a maximálního počtu jeho proměnných. Tyto problémy by mohly být zajímavým rozšířením práce.

Literatura

[1] J. P. Jones, D. Sato, H. Wada, and D. Wiens, "Diophantine representation of the set of prime numbers," *The American Mathematical Monthly, 83, str. 449-464.*, 1976.

[2] Y. V. Matiyasevich, *Hilbert's Tenth Problem.* The MIT Press, 1993.

[3] K. Conrad, "Pell's equation I,II," *https://kconrad.math.uconn.edu/blurbs/.*

[4] D. Stanovský, *Základy algebry. První vydání.* Matfyzpress, Praha, 2010.

[5] M. Davis, "Hilbert's Tenth Problem is Unsolvable," *The American Mathematical Monthly, 80(3), str. 233-269*, 1973.