Универзитет у Београду
Математички факултет

Димитрије Д. Чвокић

Гоморијеве одсијецајуће равни — развој и примјене

Мастер рад

Београд, септембар 2012.

Универзитет у Београду
Математички факултет

Димитрије Д. Чвокић

Гоморијеве одсијецајуће равни — развој и примјене

Мастер рад

Ментор
проф. др Ђорђе Дугошија

Београд, септембар 2012.

Садржај

Списак слика

Ознаке

$\mathbb{R}$	поље реалних бројева
$\mathbb{Z}$	прстен цијелих бројева
$\mathbb{Q}$	прстен рационалних бројева
$\times$	Декартов производ
S^n	$S \times S \times \ldots \times S$
S^n_+	$\{x \in S^n \ : \ x_i \geqslant 0, \ i = \overline{1,n}\}$
$(x_1, x_2, \ldots, x_n)$	уређена n-торка реалних бројева
$\emptyset$	празан скуп
$\mathbf{0}$	нула вектор
$\mathbf{1}$	јединични вектор
$\mathrm{convh}(S)$	конвексни омотач скупа S
$\mathrm{cone}(S)$	конусни (ненегативни) омотач скупа S
$\mathrm{lin}(P)$	простор линеарности полиедра P
$\mathrm{recc}(P)$	рецесивни (карактеристични) конус полиедра P
$\mathrm{corner}(\boldsymbol{\beta})$	корнерски полиедар за базу $\boldsymbol{\beta}$
$\dim(S)$	димензија скупа S
$\mathrm{vol}(P)$	запремина скупа P
$\min, \max$	минимум, максимум
$=$	једнако
$\neq$	различито
$<, \leqslant$	релације поретка мање и мање-или-једнако на скупу $\mathbb{R}$
$>, \geqslant$	релације поретка веће и веће-или-једнако на скупу $\mathbb{R}$
$\lvert S \rvert$	моћ (број елемената) скупа S
$\cap$	операција скуповног пресјека
$\in$	скуповна релације припадања
$\notin$	скуповна релација неприпадања
$\subseteq, \subset$	релација подскупа и правог подскупа
S°	унутрашњост скупа S
$\emptyset$	празан скуп
$\wedge, \vee, \veebar, \neg$	логичка конјукција, дисјункција, ексклузивна дисјункција и негација
$\top, \bot$	логичке вриједности тачно и нетачно
$\sum$	сума бројева
$\prod$	производ бројева
$A_{m \times n}, \ [a_{ij}], \ [B]$	матрица
a_{*j}	j-та колона матрице
a_{i*}	i-та врста матрице
$\det(B)$	детерминанта матрице B
c^T	транспоновани вектор c

$\lfloor a \rfloor$	(доњи) цио дио од a
$\lceil a \rceil$	горњи цио дио (строп) од a
$r(a)$	функција разломљеног дијела реалног броја a, тј. $a - \lfloor a \rfloor$
$\prec$	релација лексикографског поретка
n!	факторијел од n
$\|t\|$	норма вектора t
*	у теорији формалних језика означава Клинијево затворење
$\lvert x \rvert$	у теорији формалних језика означава дужину ријечи x
$\trianglelefteq, \triangleleft$	релације сводљивости и дјелотворне сводљивости
∞	бесконачно

Сажетак

Многи привредни проблеми могу се моделовати као цјелобројни (линеарни) програми. Теорија у позадини линеарног програмирања није довољна за њихово рјешавање, штавише није довољна ни за опис допустиве области цјелобројног програма, а доказано је и да су цјелобројни програми NP-комплетни. Развијено је више различитих метода за њихово рјешавање, међу којима је први Гоморијев метод, представљен давне 1958. године. Гоморијев метод као главно оруђе користи такозване Гоморијеве одсијецајуће равни, које до дана данашњег не престају да заокупљају пажњу математичара (научника). У овом раду представљене су изворне Гоморијеве одсијецајуће равни, али и различите верзије које су током времена настајале — потпуно цјелобројне Гоморијеве одсијецајуће равни, агрегиране Гоморијеве одсијецајуће равни, Гоморијеве мјешовито-цјелобројне одсијецајуће равни, одсијецајуће равни добијене различитим техникама појачавања и на крају одсијецајуће равни добијене из Гоморијевог корнерског полиедра.

Кључни појмови: цјелобројни програм, Гоморијеве одсијецајуће равни, корнерски полиедар

Keywords: integer program, Gomory's cutting plane, corner polyhedra

Глава 1

Цјелобројно програмирање

> *„Има области математике у којима се дошло до тако значајних резултата да се не смије чекати хиљаду година на развој интересовања за њих. Једна од тих области је и цјелобројна оптимизација која је данас главно математичко оруђе операционих истраживања.”*
>
> Томас Л. Сати

Од самог свог настанка човјек је природно тежио доношењу најбољих (тзв. оптималних) одлука, односно да за најмања улагања и најмањи утрошен рад оствари највећи могућ добитак. Самим тим оптимизација као научна (математичка) дисциплина има кључну улогу у рјешавању стварних животних проблема. Имајући то у виду, природно је да су се још у древна времена, када се математика рађала као наука, јавили проблеми оптимизације, али и да су се наставили јављати на читавом потезу историје као „вјечито” савремена проблематика.

У скоријој историји, као један од споредних производа II свјетског рата, увидјело се да је могуће математички моделовати разне људске дјелатности као што су транспорт, производња и пројектовање путне (или неке друге) мреже. Уочило се да се међусобно испреплетени услови, које дати модели укључују, могу „заробити” у систем једначина и неједначина, а потом и да се може пронаћи најбоље рјешење (оптимизовати) према неком задатом критеријуму (циљу). Најпростији такав модел, сачињен само од линеарних функција, постао је познат као *линеаран програм* (ЛП). Наравно, може се одмах увидјети да постоји више различитих облика линеарног програма с обзиром на то какве су дате линеарне функције, односно да ли су у програму присутне једнакости или неједнакости, да ли се тражи минимум или пак максимум, и томе слично. Сви ти облици се увијек могу превести једни у друге, па их због тога све уопштено зовемо линеарним програмима. *Стандардним обликом* зваћемо запис сљедећег изгледа:

$$\max \sum_{j=1}^{n} c_j x_j$$

уз ограничења

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i, \text{ гдје је } i = 1, 2, \dots, m;$$

$$x_j \geqslant 0, \text{ гдје је } j = 1, 2, \dots, n.$$

Дати програм можемо свакако краће представити, користећи се матричним записом, на сљедећи начин [4]:

$$\max\ c^T x$$

$$Ax \leqslant b,\ x \geqslant 0$$

гдје су нам $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, а $A \in \mathbb{R}^{m \times n}$ је позната нам матрица ранга m. Функција $x \mapsto c^T x$ је *функција циља*, вриједности вектора $c^T x$ су *вриједности циља*, док нам је x вектор промјенљивих које требамо израчунати.

Читајући различиту литературу увиђа се, интересантно, да сам „стандардни облик” и није нешто стандардизован. Често се среће, поготово у старијој литератури, да су главна (функционална) ограничења стандарног облика уствари једнакости, а не неједнакости, што се, опет, у литератури на нашем језику назива *канонским обликом* [4, 2].

По открићу линеарног програмирања, теорија се проширила на системе са нелинеарним конвексним функцијама. Значај конвексности се огледа у чињеници да она осигурава постојање само једног јединог екстрема (оптимума).

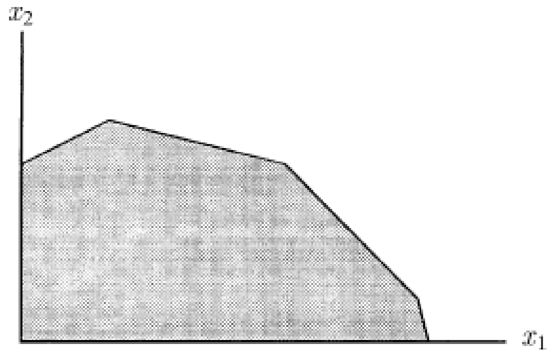
У међувремену, увидјело се да је запањујуће велик број ситуација и дјелатности које се на одговарајући начин могу моделовати као линеарни програми или конвексни нелинеарни програми. Но, како свијет у ком живимо није (искључиво) линеаран, нити конвексан, главни недостатак ових модела је немогућност да се њима представе неконвексности, или пак прекиди различитих врста. У великом броју таквих ситуаца „на сцену ступа” цјелобројно програмирање — једно од главних оруђа за моделовање и обраду разних неконвексности и прекида у оптимизацији.

1.1 О цјелобројним линеарним програмима

> *„Модел — основна форма која спаја математику и спољни свијет.”*
> Јуриј А. Митропољски

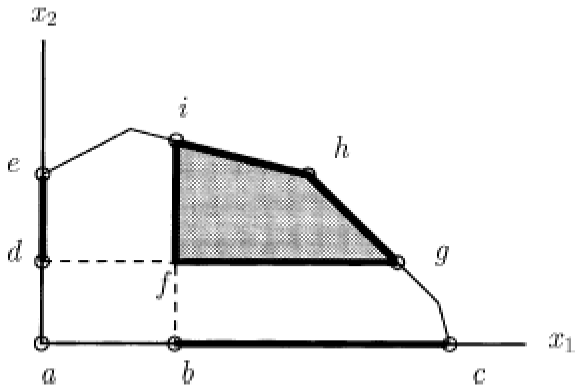
Ради сликовитијег приказа, замислимо да нека творница производи двије врсте робе, чије капацитете одређују четири линеарна ограничења представљена на Слици 1.1. Ове неједнакости, заједно са неједнакостима $x_1 \geqslant 0$, $x_2 \geqslant 0$ (количина робе која се произведе не може бити негативан број), одређују такозвани допустиви домен (скуп), на слици приказан као осијенчена област. Уколико је добитак творнице линеарна функција количине која се произведе, онда ће се оптималан (тј. добитно-максимизујући) план производње, због тога што је дата област конвексан полиедар, подударати са једним од тјемена датог полигона.

Замислимо да је додатно наметнут сљедећи услов (ограничење) — за сваку робу (ставку) постоји праг (норма) испод ког је не вриједи производити. Нека је праг за



Слика 1.1 Допустив скуп без (позитивне) доње границе за количину робе

робу 1 једнак b јединица, а за робу 2 нека је једнак d јединица. Поред тога, бар једна од двије робе се мора произвести. Допустив скуп, као што је то приказано на Слици 1.2, се у овом случају састоји од три раздвојена дијела — двије линије $[b, c]$ и $[d, e]$ и осијенчене области (полигона) чији су врхови f, g, h, i. Оптимум може сад да буде, у зависности од тога каква је функција циља, у тачкама b, c, d, e, g, h, i (тјемена конвексног омотача).



Слика 1.2 Допустив скуп са (позитивном) доњом границом за количину робе

Такође, примјетно је да је проблем постао и квалитативно другачији. Услови „прага” који су облика „$x_1 = 0 \veebar x_1 \geqslant b$” и „$x_2 = 0 \veebar x_2 \geqslant d$”, као и „$x_1 = 0 \veebar x_2 = 0$” се не могу моделовати техникама линеарног нити конвексног програмирања. Истовјетна ситуација је и када се задавају разни други „логички” услови:

- дисјункције („ово или противно”, „мора важити бар једно од ових неколико ограничења”, „највише једна од неколико промјенљивих смије бити позитивна”),
- импликације („ако је предузета дата радња, онда се мора иврšити ова радња”),
- услови поретка („дати догађај мора претходити овом догађају”, „ова радња не може почети све док се неке друге не заврше”), и многи други.

Свакако да услови овог типа нису ни на који начин изузеци, штавише, они просто прожимају читаво мноштво стварних животних ситуација и стварних дешавања.

Проблем линеарног (нелинеарног) програмирања чије су промјенљиве додатно ограничене на цјелобројне вриједности зовемо проблем цјелобројног линеарног (нелинеарног) програмирања, или просто *цјелобројним линеарним (нелинеарним) програмом.* У литератури се пак чешће користи краћи назив — цјелобројни програм (ЦП), пошто се подразумјева да су функције линеарне (нелинеарност се обично посебно нагласи). Цјелобројни програми се могу записати на сљедећи начин:

$$\max \sum_{j=1}^{n} c_j x_j$$

$$\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, \text{ за } i = 1, 2, \ldots, m;$$

$$x_j \geqslant 0, \text{ за } j = 1, 2, \ldots, n;$$

$$x_j \in \mathbb{Z}, \text{ за } j = 1, 2, \ldots, n.$$

Конкретно, у овом запису се ради о *потпуно* цјелобројном програму (такође се краће означава са ЦП).

Уколико неке промјењљиве (али не све) не би биле ограничене на цјелобројне вриједности, онда би се радило о *мјешовито-цјелобројном програму* (МЦП). Такав проблем се може записати на сљедећи начин:

$$\max \sum_{j=1}^{n} c_j x_j$$

$$\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, \text{ за } i = 1, 2, \ldots, m;$$

$$x_j \geqslant 0, \text{ за } j = 1, 2, \ldots, n;$$

$$x_j \in \mathbb{Z} \text{ за } j \in I \subseteq \{1, \ldots, n\}.$$

Потпуно цјелобројан програм је, у ствари, посебан случај мјешовито-цјелобројног програма када је $I = \{1, \ldots, n\}$. У индустрији се најчешће јављају управо проблеми мјешовито-цјелобројног програмирања. Ипак, у овом раду ће нагласак бити стављен на потпуно цјелобројне програме, пошто се већи дио теорије за рјешавање линеарних и цјелобројних програма може на природан начин проширити до теорије за рјешавање мјешовито-цјелобројних програма.

Поред претходних разматрања, битно је издвојити и такозване *0-1 (бинарне) програме* (потпуне и мјешовите). У тим програмима, цјелобројне промјењљиве су додатно ограничене да од вриједности могу узимати или 0 или 1, то јест, то су такозване изборне промјењљиве — рецимо, нешто је или урађено или није урађено. Према мишљењу многих аутора, од којих су неки Џорџ Нојмхаузер, Елис Џонсон и Егон Бала [42], 0-1 програми чине најзначајнију класу цјелобројних програма. Један он разлога за то је што се разне неконвексности, као на примјер оне из претходног проблема, или пак оне које су додатно наведене, могу представити помоћу 0-1 промјењљивих. На примјер, у предоченом проблему планирања производње, ако је скуп ограничења, како смо га приказали на Слици 1.1,

$$\begin{aligned} &a_{i1}x_1 + a_{i2}x_2 \leqslant b_i, \text{ за } i = 1, \ldots, 4, \\ &x_1 \geqslant 0, \; x_2 \geqslant 0, \end{aligned}$$

онда се услови

$$\begin{aligned} x_1 \leqslant 0 \ &\vee \ x_1 \geqslant b, \\ x_2 \leqslant 0 \ &\vee \ x_2 \geqslant d, \\ x_1 > 0 \ &\vee \ x_2 > 0, \end{aligned}$$

наметнути у варијанти предоченој на Слици 1.2, могу формулисати увођењем двије 0-1 промјенљиве — δ_1 и δ_2, на сљедећи начин

$$\begin{aligned} & b\delta_1 \leqslant x_1 \leqslant c\delta_1, \\ & d\delta_2 \leqslant x_2 \leqslant e\delta_2, \\ & \delta_1 + \delta_2 \geqslant 1, \quad \delta_1, \delta_2 \in \{0, 1\}. \end{aligned}$$

Опет, ограничења типа $0 \leqslant x_i \leqslant b_i$ се могу замијенити ограничењима $0 \leqslant x_{i1} + x_{i2} + \ldots + x_{i\lfloor b_i \rfloor} \leqslant b_i$, гдје су x_{ij} $(j = 1, \ldots, b_i)$ 0-1 промјенљиве.

С друге стране, међу математичким моделима у цјелобројном програмирању, посебно мјесто заузима такозвани *проблем ранца.*

Проблем 1.1 (Проблем ранца)**.** Претпоставимо да је дат неки ранац запремине $b \geqslant 0$ и неки скуп предмета. За сваки предмет j се знају његова запремина $a_j \geqslant 0$ и његова вриједност $c_j \geqslant 0$. Напунити ранац тако да је укупна вриједност ствари у њему највећа могућа.

Сам појам „ранац" треба уопштено схватити, то јест под њим се може подразумијевати, на примјер, складиште теретног брода или пак простор космичке летјелице, такође ни вриједност не мора имати уобичајено (буквално) значење. Слиједи формулација проблема у облику цјелобројног програма:

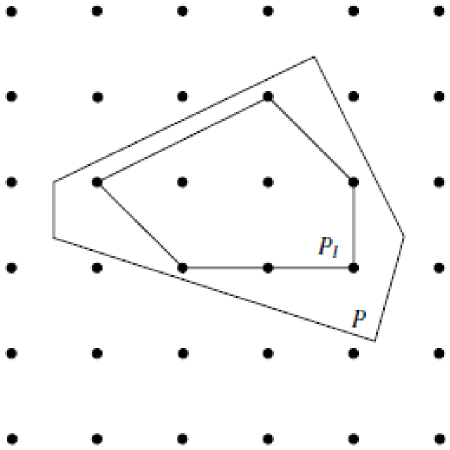
$$\begin{aligned} \max \sum_{j=1}^{n} c_j x_j & \\ \sum_{j=1}^{n} a_j x_j \leqslant b, & \quad \text{(само једно ограничење)} \\ x_j \geqslant 0, \ x_j \in \mathbb{Z}. & \end{aligned}$$

Овако формулисан проблем се назива *неомеђен проблем ранца*, због тога што не постоји горња међа на број умножака предмета x_j, сваке врсте j, којим се треба напунити ранац. У *омеђеном проблему ранца* број умножака предмета x_j је омеђен одозго неким цијелим бројем w_j, то јест $x_j \in \{0, ..., w_j\}$. Наравно, постоји и 0-1 формулација, која је иначе једна од најчешће кориштених. У *0-1 проблему ранца* се сваки предмет x_j може највише једном искористити за попуњавање (не посматрају се предмети по врстама).

Значај овог проблема у теорији цјелобројног програмирања се огледа у чињеници да се сваки рационалан цјелобројни програм може *агрегацијом* свести на одговарајући проблем ранца [3]. Под агрегацијом система линеарних једначина са цјелобројним коефицијентима подразумијевамо поступак добијања једне једначине са цијелим ненегативним коефицијентима која је истовјетна почетном систему у смислу да су им једнаки скупови ненегативних цјелобројних рјешења. Главни недостатак агрегације је експоненцијална зависност броја промјењљивих у добијеној једначини од броја једначина у систему (ограничења програма) [3].

Природно, сваком цјелобројном програму је придружен одговарајући линеарни програм, који се добије уклањањем услова цјелобројности. Дати линеарни програм се назива *релаксација цјелобројног програма у линеарни*, или краће — *ЛП-релаксација.*

У линеарним програмима, скуп допустивих рјешења $P = \{x \in \mathbb{R}^n \ : \ Ax \leqslant b, \ x \geqslant 0\}$ је политоп, па је природно, због захтјеване цјелобројности рјешења, посматрати конвексан омотач цјелобројних вектора унутар P, односно $P_I = \mathrm{convh}(P \cap \mathbb{Z}^n)$. Очито је да је $P_I \subseteq P$. P_I називамо *цјелобројним омотачем* унутар P. Јасно је да је за омеђен P, цјелобројни омотач P_I политоп (Сл. 1.3), као конвексан омотач коначног скупа тачака (вектора).



Слика 1.3 Цјелобројни омотач P_I унутар политопа P

Иначе, показано је да важи и јаче тврђење, то јест сљедећа теорема.

Теорема 1.1 (Шрајвер [56]). *За сваки рационалан полиедар P, цјелобројан омотач P_I је такође рационалан полиедар.*

Ово тврђење, не важи у општијем случају, то јест када се узму у обзир ирационални полиедри.

Примјер 1.1. За полиедар $P = \{(x, y) \in \mathbb{R}^2 \ : \ y \leqslant \sqrt{2}x, \ x \geqslant 0\}$ важи да је $P_I = \sup\{P_k \ : \ k \in \mathbb{N}_0\}$, гдје је

$$P_k = \{(x, y) \in \mathbb{R}^2 \ : \ y \leqslant (\prod_{i=0}^{k}(1 + \frac{1}{4i+1})(1 - \frac{1}{4i+3}))x, \ x \geqslant 0\},$$

односно важи $P_I = P$. ■

Помало изненађујуће, из претходне теореме непосредно слиједи да се сваки цјелобројни програм може записати као $\max\{x \ : \ x \in P_I\}$, што није ништа друго до линеарни програм. Другим ријечима, да би се могле примјенити технике линеарног програмирања довољно је окарактерисати, бар код цјелобројног рјешења, P_I помоћу линеарних неједнакости. Испоставља се, међутим, да то није нимало лак задатак.

1.2 Уобичајене погрешне претпоставке о цјелобројним програмима

> *„Свака добра теорема мора имати добар контрапримјер.”*
> Франческо Севери

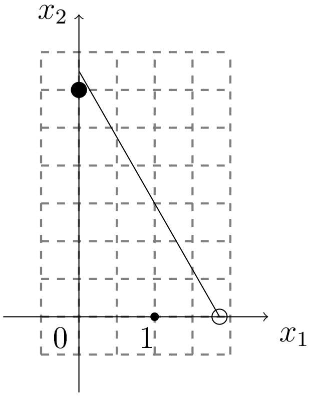
У наставку су дати контрапримјери за неке уобичајене погрешне претпоставке приликом интуитивног сагледавања цјелобројног програмирања, а од којих су многе у непосредној вези са линеарним програмирањем. Наравно, самим тим се још више расвјетљује разлика и бива јаснији однос између линеарног и цјелобројног програмирања.

Претпоставка 1. *Рјешење цјелобројног програма се може добити заокруживањем рјешења његове ЛП-релаксације.*

Контрапримјер. Напротив, могуће је да рјешење буде доста далеко од рјешења линеарног програма, штавише могуће је да заокружено рјешење уопште не буде допустиво. Погледајмо сљедећи програм (Сл. 1.4)

$$\max\{21x_1 + 11x_2 \ :\ 7x_1 + 4x_2 \leqslant 13,\ x_1 \geqslant 0,\ x_2 \geqslant 0\}.$$

Оптимално рјешење ЛП-релаксације се достиже у тачки $(\frac{13}{7}, 0)$, док се с друге стране оптимално цјелобројно рјешење достиже у тачки $(0, 3)$. ■



Слика 1.4

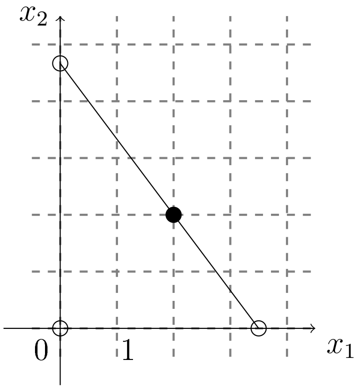
Гледајући претходни контрапримјер, могло би се помислити да се цјелобројно рјешење може добити заокруживањем неког од врхова политопа ЛП-релаксације, као на примјер $(0, \frac{13}{4})$ на $(0, 3)$ у нашем случају.

Претпоставка 2. *Рјешење цјелобројног програма се може добити заокруживањем неког од допустивих базних рјешења (врхова политопа) одговарајуће ЛП-релаксације.*

Контрапримјер. Ипак, ни ова претпоставка није тачна, што се врло лако увиђа на сљедећем програму (Сл. 1.5)

$$\max\{13x_1 + 19x_2 \ :\ 4x_1 + 3x_2 \leqslant 14,\ x_1 \geqslant 0,\ x_2 \geqslant 0\}.$$

Врхови политопа ЛП-релаксације су у $(\frac{14}{4}, 0)$, $(0, \frac{14}{3})$ и $(0, 0)$, а максимум за цјелобројне тачке се достиже у $(2, 2)$. ■



Слика 1.5

Наравно, поред теоријских контрапримјера, постаља се питање, колико често се у пракси срећу такви случајеви у којима заокруживање није ни од какве помоћи? Да ли пракса, да тако кажемо, дозвољава заокруживање? Гловер и Сомер су у [30] представили (контра)примјере из праксе, штавише представили су класу проблема који се јављају у пракси и у којима заокруживање даје јако лоше резултате.

Претпоставка 3. *Рјешење ЛП-релаксације је рјешење одговарајућег цјелобројног програма ако и само ако је оптимално базно рјешење ЛП-релаксације цјелобројно.*

Контрапримјер. Довољан услов је увијек тачан (испуњен), међутим потребан услов важи (колико је до сада познато [39]) само за бинарне програме, код којих су ограничења ЛП-релаксације једнакости, (у такозваном *канонском облику*). Генерално тврдња није тачна. Допустива област програма

$$\max\{0x \ : \ 0 \leqslant x \leqslant \begin{bmatrix} b \\ \vdots \\ b \end{bmatrix}, \ x \in \mathbb{Z}, \ b \notin \mathbb{Z}\},$$

има нецјелобројне екстремне тачке, а цјелобројну оптималну тачку у унутрашњости. Рјешење ЛП-релаксације за овај програм је било која допустива тачка. За $b > 1$, тачка (јединични вектор) $\mathbf{1}$ је допустива, па је стога и оптимална тачка датог ЛП-а. Самим тим, $\mathbf{1}$ је оптимална тачка цјелобројног програма, али то ипак није оно што тражимо — базна (врх политопа) оптимална тачка ЛП-релаксације. ■

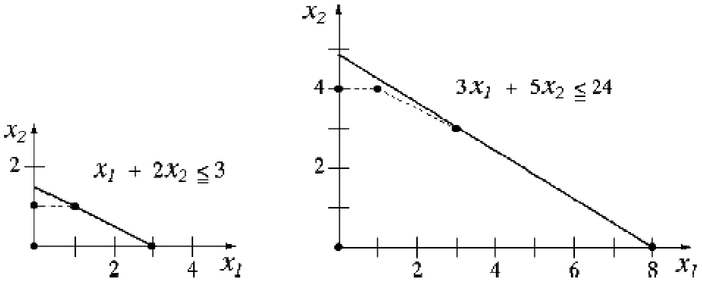
Као што се може примјетити, постоје озбиљне тешкоће у проналажењу, па и карактерисању рјешења цјелобројног програма помоћу рјешења одговарајуће ЛП-релаксације. Природно, поставља се питање да ли је онда могуће некако окарактерисати бар цјелобројни омотач на основу допустиве области ЛП-релаксације.

Претпоставка 4. *Број врхова цјелобројног омотача је омеђен неким умношком броја врхова ЛП-релаксације, гдје умножак зависи од броја промјенљивих, као и броја ограничења.*

Контрапримјер. Можда неочекивано на први поглед, али цјелобројни омотач може имати било који број екстремних тачака и то уз само једно ограничење. Да би то показали, узмимо у разматрање сљедећи полигон који има само двије промјенљиве и једно ограничење (уз услов ненегативности):

$$P = \{x \in \mathbb{R}^2_+ \ : \ a_1x_1 + a_2x_2 \leqslant b\}$$

гдје су нам $a, b > 0$. Овај полигон има три екстремне тачке. Можемо да изаберемо a и b такве да њихов цјелобројни омотач, $\mathrm{convh}(P \cap \mathbb{Z}^2)$, има n екстремних тачака за било које $n \geqslant 3$. Слика 1.6 приказује два таква полигона.



(a) $N = 4$ (b) $N = 5$

Слика 1.6

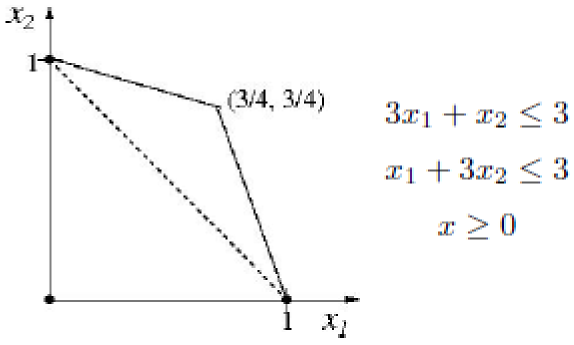
У сљедећој табели је приказано још неколико избора a и b за које бројеви врхова одговарајућег цјелобројног омотача редом расту, а и лако се уочава одговарајући образац који је Рубин представио у [53]:

n	a_1	a_2	b
4	1	2	3
5	3	5	24
6	8	13	168
7	21	34	1155
8	55	89	7920
9	144	233	54288
10	377	610	372099

Нека је F_k k-ти Фибоначијев број и $P = \{x \in \mathbb{R}^2_+ : \ F_{2k-1}x_1 + F_{2k}x_2 \leqslant F^2_{2k} - 1\}$. Тада convh$(P \cap \mathbb{Z}^2)$ има $k+3$ екстремне тачке. Рубин у свом раду даје још неколико начина за производење полигона помоћу једног ограничења у $\mathbb{R}^2_+$, који при том имају произвољан број цјелобројних врхова. Иначе, исти ови контрапримјери су у прихватљиви да смо којим случајем, у претпоставци, врхове у теореми замијенили страницама. ■

Претпоставка 5. *Број врхова цјелобројног омотача је већи или пак једнак броју врхова полиедра ЛП-релаксације.*

Контрапримјер. Полиедар ЛП-релаксације може да има области без цјелобројних тачака, као што је приказано на Слици 1.7. Цјелобројни омотач има три екстремне тачке, док полиедар ЛП-релаксације има четири (и може их имати произвољно много). Као и у претходном контрапримјеру, слично се може показати и за странице одговарајућих политопа. ■



Слика 1.7

Из теорије линераног програмирања познато је да се оптимална рјешења траже међу врховима политопа, па је због тога број могућих рјешења $\binom{n}{m}$ (ако је матрица ЛП-а димензија $m \times n$). Другим ријечима, свако оптимално рјешење можемо свести на допустиво базно оптимално рјешење. Међутим, тако нешто нам није од помоћи ако нас интересују цјелобројна рјешења.

Претпоставка 6. *Ако постоји оптимално рјешење цјелобројног програма записаног у канонском облику, онда постоји оптимално рјешење са највише $\phi(m)$ позитивних вриједности, гдје је ϕ функција која зависи само од m, то јест не зависи од n.*

Контрапримјер. Зна се да за ЛП важи $\phi(m) = m$. Нека су $p_1, \dots, p_n$ неки међусобно различити прости бројеви. Узевши да је $a_j = \prod_{i \neq j} p_i$, размотримо цјелобројни програм

$$\max\{cx \ : \ \sum_{j=1}^{n} a_j x_j = \sum_{j=1}^{n} a_j, \ x \in \mathbb{Z}^n_+\}.$$

Једино допустиво рјешење је $x = \mathbf{1}$, па је стога број позитивних промјенљивих једнак n.

Према [39], контрапримјер припада Полу Вилијамсу и Леслију Тротеру. ■

Претпоставка 7. *Уколико цјелобројни програм има неомеђену ЛП-релаксацију, онда је и он сам, такође, неомеђен.*

Контрапримјер. Размотримо програм

$$\max\{x_1 \ : \ x \in \mathbb{Z}^4_+, \ x_3 - \sqrt{2}(x_1 - x_2) = 0, \ x_2 + x_4 = 1\}.$$

Скуп одређен ограничењима ЛП-релаксације садржи зрак $\{(t, 0, t\sqrt{2}, 1) \ : \ t \geqslant 0\}$, па је ЛП-релаксација неомеђена. Међутим, за цјелобројно рјешење мора да важи да је $x_1 = x_2$, $x_3 = 0$ и $x_1, x_2, x_4 \in \{0, 1\}$. Слиједи да су једина два допустива рјешења цјелобројног програма $(0, 0, 0, 1)$ и $(1, 1, 0, 0)$. ■

Примједба 1.2. Мејер је у [51] показао да претпоставка ипак важи ако су, додатно, испуњени сљедећи услови (подразумјева се да су програми допустиви и да су функције циља ограничене):

1. довољан услов за потпуно цјелобројан програм у канонском облику је да су сви коефицијенти *рационални*
2. за мјештовито-цјелобројан програм довољан услов је да су сви коефицијенти *рационални.*

Примједба 1.3. Са становишта рачунарске обраде цјелобројни програм је омеђен и то само са рационалним подацима, ако се при том дозволе SOS (енгл. Special Ordered Sets) одломци, као у CPLEX-у. Наредни примјер који то појашњава је дао Ед Клоц [39].

$$\max\{x_1 \ : \ x \in \mathbb{Z}^3_+, \ x_2 \leqslant 1, \ x_3 - 1.41421x_1 + 1.41421x_2 = 0\}$$

```
SOS:
S2:: x_1: 1 x_2: 2 x_3: 3
End
```

У ЛП-релаксацији, зрак $(t, 0, 1.41421t)$ је допустив. Ипак, SOS захтјев дозвољава да само двије узастопне промјенљиве у SOS скупу узимају ненула вриједности, па је софтвер „приморан на сјечу” овог зрака. Као резултат тога, цјелобројни програм има ограничено оптимално рјешење, $x = (1, 1, 0)$.

Претпоставка 8. *Ако цјелобројни програм има неомеђену ЛП-релаксацију, онда је он допустив.*

Контрапримјер. Полиедар је неомеђен ако садржи допустив зрак (полуправу) — другим рјечима, $\{x^0 + th \ : \ t \geqslant 0\} \subset P$, гдје $x^0 \in P$ и $h \neq \mathbf{0}$. Посматрајмо полиедар

$P = \{(x, y) \ : \ x \geqslant \frac{3}{4}, \ y \geqslant \frac{1}{2}, \ x - y = \frac{1}{4}\} = \{(\frac{3}{4}, \frac{1}{2}) + t(1, 1) \ : \ t \geqslant 0\}$. Увиђа се да он не садржи ни једну једину цјелобројну тачку, пошто

$$(x, y) = (\frac{3}{4} + t, \frac{1}{2} + t) \ \Rightarrow \ x - y = \frac{1}{4}. \qquad \blacksquare$$

Претпоставка 9. *Цјелобројни програм се може ријешити у коначно много корака додавањем одсијецајућих хиперравни облика*

$$\sum_{j \in N^k} x_j \geqslant 1$$

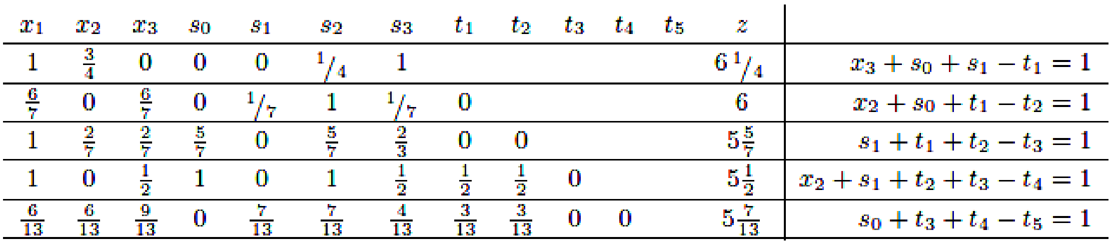
гдје је N^k скуп небазних промјенљивих у k-тој ЛП-релаксацији за коју постоји нецјелобројно базно рјешење.

Контрапримјер. Ове хиперравни су познате под називом *Данцигове одсијецајуће равни.* Ралф Гомори и Ален Хофман су сљедећим контрапримјером доказали да њихово додавање не мора да конвергира у коначном броју корака ка оптималном рјешењу.

У цјелобројном програму

$$\max\{4x_1 + 3x_2 + 3x_3 \ : \ x \in \{0, 1\}^3, \ 3x_1 + 4x_2 + 4x_3 \leqslant 6\},$$

максимум циљне функције је једнак 4 и достиже се у цјелобројној тачки $x = (1, 0, 0)$. Нека је s_j изједначавајућа промјенљива за горње међе вриједности промјенљивих, $x_j + s_j = 1$ и $j \in \{1, 2, 3\}$, а нека је s_0 изједначавајућа промјенљива за главно ограничење програма, то јест $3x_1 + 4x_2 + 4x_3 + s_0 = 6$. Рјешење ЛП-релаксацију је једнако $6\frac{1}{4}$ и достиже се у $x = (1, \frac{3}{4}, 0)$, док је $s = (0, 0, \frac{1}{4}, 1)$.



x_1	x_2	x_3	s_0	s_1	s_2	s_3	t_1	t_2	t_3	t_4	t_5	z	
1	$\frac{3}{4}$	0	0	0	$^1/_4$	1						$6\,^1/_4$	$x_3 + s_0 + s_1 - t_1 = 1$
$\frac{6}{7}$	0	$\frac{6}{7}$	0	$^1/_7$	1	$^1/_7$	0					6	$x_2 + s_0 + t_1 - t_2 = 1$
1	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{5}{7}$	0	$\frac{5}{7}$	$\frac{2}{3}$	0	0				$5\frac{5}{7}$	$s_1 + t_1 + t_2 - t_3 = 1$
1	0	$\frac{1}{2}$	1	0	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0			$5\frac{1}{2}$	$x_2 + s_1 + t_2 + t_3 - t_4 = 1$
$\frac{6}{13}$	$\frac{6}{13}$	$\frac{9}{13}$	0	$\frac{7}{13}$	$\frac{7}{13}$	$\frac{4}{13}$	$\frac{3}{13}$	$\frac{3}{13}$	0	0		$5\frac{7}{13}$	$s_0 + t_3 + t_4 - t_5 = 1$

Претходна табела приказује првих пет корака, уводећи допунске промјенљиве t^k, када је конструисан k-ти рез.

Одсијецајуће равни се додају једна за другом, али се програм, ипак, не рјешава у коначном броју корака [31].

Џозеф Боуман и Џорџ Нојмахузер су доказали конвергенцију поступка са измијењеним Данциговим одсијецајућим равнима, а двије године касније њихов резултат су додатно побољшали Рубин и Грејвс. ■

Као што се може примјетити, теорија у позадини линеарног програмирања није довољна за рјешавање и нити за потпуну карактеризацију цјелобројних програма, па се природно поставља питање да ли су цјелобројни програми, рачунски гледано, тежи од линеарних (без обзира на технике за рјешавање)?

1.3 Рачунска сложеност цјелобројног програма

„P у односу на NP — поклон математици од информатике.”
Стивен Смејл
"Тврд је орах, воћка чудновата, не сломи га, ал' зубе поломи."
Петар II Петровић Његош, *Горски вијенац*

Што се тиче рачунске тежине цјелобројног програма, чињеница је да нису познати такозвани дјелотворни алгоритми за његово рјешавање. У суштини, то и није тако зачуђујуће пошто је могуће формулистати велики број (рачунски) јако тешких проблема (нпр. многе познате проблеме комбинаторне оптимизације) у облику цјелобројних програма.

Тјурингова машина (недетерминистичка, детерминистичка, са улазом и излазом), као основна апстрактна машина чији се ресурси процјењују током рада програма, је најчешћи формални модел израчунавања, који се користи у анализи сложености. Њене предности су релативна једноставност и могућност дјелотворног подражавања других апстрактних машина, тако да се основне класе сложености као што су P, NP и $PSPACE$ не мијењају промјеном саме машине. Ресурси израчунавања, чија искориштеност осликава сложеност алгоритма, су *вријеме* (број корака у извршавању алгоритма) и *простор* (меморијски простор који алгоритам користи).

Предмет проучавања теорије сложености израчунавања чине такозвани одлучиви проблеми, као појединачне формуле. Неформално, то су они проблеми који су алгоритамски рјешиви, то јест постоје програми за њихово рјешавање на апстрактним машинама. Обично су окарактерисани питањима на која се даје одговор *да*, односно *не*. На примјер, провјера да је број прост је један проблем. Ако претходно питање поставимо за неки конкретан број, онда је то *примјерак* проблема.

Само представљање проблема се врши на неком формалном језику (тј. кодира се) који је изграђен над азбуком неке Тјурингове машине.

Проблем I за који се испитује сложеност је подскуп скупа свих ријечи неке азбуке, односно сам проблем је у ствари формалан језик. *Комплемент проблема* I, у ознаци $\overline{I}$ над неком азбуком је скуп свих ријечи над датом азбуком које нису у I.

За дату азбуку $\mathcal{A}$ и проблем I, Тјурингова машина M прихвата улазни податак (ријеч) x, ако постоји израчунавање (обрада податка) у ком се полазећи од ријечи x уписане на улазној траци у почетном стању q_0, долази до завршног стања $q_{\text{да}}$, а одбацује x, ако увијек долази до завршног стања $q_{\text{не}}$. Уколико M прихвата све ријечи x које припадају проблему I, а одбацује сваку ријеч која није у проблему I, каже се да M *одлучује* проблем I.

Ако Тјурингова машина M за улазни податак x у израчунавању долази до стања q_z, онда је садржај посљедње траке резултат рада машине и означава се са $M(x)$. Са $I(x)$ означаваћемо примјерак проблема I за улазни податак x, односно питање да ли је $x \in I$.

За дати одлучив проблем, горња граница потребног времена и простора одређује се разматрањем конкретног алгоритма који га рјешава. Додатно, да би се доказала оптималност алгоритма, потребно је имати и одговарајуће доње границе сложености проблема (овдје то неће бити разматрано).

У анализи сложености израчунавања, обично се разматра асимптотско понашање, за шта се користи такозвана O-нотација.

Ако постоје позитивна константа c и природан број n_0 такви да за функције f и g над природним бројевима важи

$$f(n) \leqslant c \cdot g(n) \text{ за све вриједности } n \text{ веће од } n_0$$

Онда пишемо

$$f = O(g)$$

и читамо „f је у великом О од g".

Само O није функција, већ означава класу функција. Сабирање или множење са константом не утичу на класу којој функција припада.

Ако за $T(n)$ као вријеме извршавања алгоритма A (чија величина улаза је окарактерисана природним бројем n) важи да је $T = O(g)$, онда кажемо да је алгоритам A *временске сложености (или реда)* g и да припада класи $O(g)$.

Аналогно се дефинише *просторна сложеност* алгоритма.

Гледано у односу на Тјурингову машину, ако је приликом самог израчунавања промијењено t конфигурација машине, онда је t *вријеме* тог израчунавања. Под *простором* израчунавања подразумјевамо број поља машине којима се приступа током израчунавања. Тјурингова машина *прихвата проблем у времену (простору)* $F(n)$, ако за сваку улазну вриједност коју прихвата (тј. за коју је одговор потврдан) и чија величина може бити описана природним бројем n, постоји израчунавање које је прихвата у времену (простору) које не прелази $F(n)$.

Са *ДВРИЈЕМЕ*$(F(n))$ означавамо скуп проблема за које постоје детерминистичке Тјурингове машине које их одлучују, а за које је временска граница сложености $F(n)$. *НВРИЈЕМЕ*$(F(n))$ се аналогно дефинише у односу на недетерминистичке Тјурингове машине.

ДПРОСТОР$(F(n))$ је скуп проблема за које постоје детерминистичке Тјурингове машине које их одлучују, а за које је просторна граница сложености $F(n)$. Скуп проблема *НПРОСТОР*$(F(n))$ се дефинише аналогно, у односу на недетерминистичке Тјурингове машине.

У наставку излагања, подразумјева се предочена сличност при дефинисању просторне сложености у односу на временску сложеност.

За дати проблем одлучивања са улазном вриједношћу n, гдје је n величина репрезентације улазног податка x, кажемо да је *полиномске временске сложености* ако је његово вријеме извршавања на детерминистичкој Тјуринговој машини $O(P(n))$ гдје је $P(n)$ полином по n. Класу полиномских алгоритама означавамо са P, то јест

$$P = \bigcup_{i \in \mathbb{N}_0} \text{\textit{ДВРИЈЕМЕ}}(n^i).$$

Генерално, под појмом класа (сложености израчунавања) се подразумјева скуп алгоритама са заједничком временском границом.

Проблеми за чије рјешавање постоје алгоритми из класе P сматрамо *дјелотворно рјешивим.*

У анализи сложености проблема линеарног и цјелобројног програмирања наредне двије теореме имају темељну улогу (за доказе погледати [28, 26]), но прије њих дефинисаћемо величину бинарно кодираног записа рационалног броја, рационалног вектора и рационалне матрице.

Ако је $r = p/q$ рационални број у канонској представи гдје је $p \in \mathbb{Z}$ и $q \in \mathbb{N}$, *величина бинарно кодираног записа* броја r је

$$\text{величина} = 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(q + 1) \rceil.$$

Величина бинарно кодираног записа рационалног вектора $v \in \mathbb{Q}^n$ и рационалне матрице $A \in \mathbb{Q}^{m \times n}$ је дефинисана помоћу

$$\text{величина}(v) = n + \sum_{j=1}^{n} \text{величина}(v_i),$$

$$\text{величина}(A) = mn + \sum_{i=1}^{m} \sum_{j=1}^{n} \text{величина}(a_{ij})$$

Теорема 1.4. *За рационалну квадратну матрицу A, величина њене детерминанте је полиномски ограничена, конкретније,* $\text{величина}(\det(A)) \leqslant 2 \cdot \text{величина}(A)$.

Слиједи да ако постоји рјешење система рационалних линеарних једначина $Ax = b$, онда је оно полиномски ограничено величинама матрице A и вектора b.

Теорема 1.5 (Едмондс [26]). *За систем рационалних линеарних једначина $Ax = b$ постоји алгоритам полиномског времена (заснован на Гаусовој или Гаус-Жордановој елиминацији) који или проналази рјешење система или пак даје увјерење да рјешење система не постоји.*

Класа свих проблема одлучивања за чије рјешавање постоје алгоритми за недетерминистичке Тјурингове машине, који имају полиномско вријеме извршавања (по величини улаза), назива се *класа NP*, односно

$$NP = \bigcup_{i \in \mathbb{N}_0} \textit{НВРИЈЕМЕ}(n^i).$$

Лако се увиђа да важи $P \subseteq NP$, али се још увијек не зна да ли важи $P = NP$, што је и један од миленијумских проблема математичког института Клеј. Напоменимо да ако неки проблем из класе NP има полиномско рјешење, онда то још увијек не значи да је $P = NP$.

Поред претходне дефиниције, користи се и сљедећа (еквивалентна) дефиниција класе NP.

Језик $J \subseteq \{0,1\}^*$ припада *класи NP* ако постоји полином $p : \mathbb{N} \mapsto \mathbb{N}$ и детерминистичка Тјурингова машина полиномског времена M тако да за свако $x \in \{0,1\}^*$, важи да је

$$x \in J \iff \exists u \in \{0,1\}^{p(|x|)} \text{ тако да је } M(x, u) = 1.$$

Ако $x \in J$ и $u \in \{0,1\}^{p(|x|)}$ задовољавају да је $M(x, u) = 1$ онда u називамо *увјерењем* за x (у односу на језик J и машину M). Другим ријечима класа NP садржи проблеме чије се рјешење може провјерити детерминистички у полиномском времену. У литератури се може пронаћи и назив *потврда* или пак *свједок* умјесто увјерења.

Проблем I_1 је *сводљив* на проблем I_2, у ознаци $I_1 \trianglelefteq I_2$ ако постоји израчунљива функција f, таква да је одговор на $I_1(x)$ потврдан ако и само ако је потврдан одговор на $I_2(f(x))$. Функција f се тада назива *функција сводљивости (редукције).*

Функција $f : \{0,1\}^* \to \{0,1\}^*$ је израчунљива у полиномском времену ако постоји полином p и детерминистичка Тјурингова машина M која кад почне са радом у ω, завршава са $f(\omega)$ (на траци), при чему је вријеме рада машине мање од $p(|\omega|)$.

Функција сводљивости f проблема I_1 на проблем I_2 је *дјелотворна*, а проблем I_1 је *дјелотворно сводљив* на I_2, у ознаци $I_1 \triangleleft I_2$, ако је сложеност функције f у класи P. Иначе, пошто су полиноми затворени под композицијом, релација дјелотворне сводљивости је транзитивна.

За проблем I кажемо да је NP*-тежак* ако је сваки проблем из класе NP дјелотворно сводљив на I, при чему он сам не мора да припада класи NP.

За проблем I кажемо да је NP*-комплетан* ако је NP-тежак и ако припада класи NP.

Општије, ако проблем I припада некој класи $\mathcal{K}$ и ако је сваки проблем из класе $\mathcal{K}$ сводљив у полиномском времену на I, онда за проблем I кажемо да је $\mathcal{K}$*-комплетан.*

Имајући у виду претходно речено, ако неки NP-тежак проблем припада класи P, тада важи да је $P = NP$. Стога, да би се доказало да је $P = NP$, довољно је за један NP-комплетан проблем доказати да припада класи P. Значи, „само” је потребно пронаћи и размотрити неки погодан NP-комплетан проблем. Ипак, доказивање да је неки проблем NP-комплетан, непосредно сљедећи дефиницију, је поприлично тежак задатак. Наредна теорема омогућује да то доказивање буде једноставније (доказ тврђења је тривијалан).

Теорема 1.6. *Проблем I је NP-комплетан ако:*

- *I припада класи NP;*
- *J је дјелотворно сводљив на I, гдје је J неки NP-комплетан проблем (тј. показујемо да је I NP-тежак).*

Наведена теорема омогућује утврђивање да је неки проблем NP-комплетан, ако је за неки други већ познато да је NP-комплетан. Међутим, поставља се питање да ли уопште постоји иједан NP-комплетан проблем. Стивен Кук је први, 1971. године, доказао (непосредно, користећи формализам Тјурингове машине) да постоје NP-комплетни проблеми, тачније, доказао је да је проблем SAT (енгл. Boolean SATisfability problem — Проблем задовољивости логичких исказа) NP-комплетан. Дато тврђење се сматра једним од најзначајнијих резултата теоријског рачунарства (за доказ видјети [19]).

Иначе, Стивен Кук је први доказао да постоје и P-комплетни проблеми. Један од њих је CVP (енгл. Circuit Value Problem) — Проблем процјене излаза Буловог кола.

Проблем 1.2 (Проблем SAT)**.** Посматрамо логичке изразе који су записани само помоћу оператора $\wedge, \vee, \neg$, логичких промјенљивих (као вриједности примају или $\top$ или $\bot$) и заграда. Поставља се питање: „Ако је дат израз, да ли постоји нека додјела вриједности $\top$ и $\bot$ промјенљивим, тако да цио израз има вриједност $\top$?”. Ако је одговор на питање потврдан, каже се да је израз *задовољив.*

Теорема 1.7 (Кукова теорема)**.** *SAT је NP-комплетан.*

Непосредна посљедица Кукове теореме је да је $P = NP$ ако и само ако $SAT \in P$.

Да би доказали да је цјелобројан програм NP-комплетан треба доказати да он припада класи NP и да је NP-тежак (нпр. дјелотворно сводљив на SAT).

Прије него што кренемо са доказом да цјелобројан програм припада класи NP, навешћемо неколико корисних појмова и тврђења из теорије полиедара. Докази теорема се могу пронаћи у [28, 1].

Скупу S додјељујемо скупове $S_{[k]}$ конвексних комбинација (за све природне k):

$$S_{[k]} = \left\{ \sum_{i=1}^{k} \lambda_i x^i \ : \ x^i \in S, \ \sum_{i=1}^{k} \lambda_i = 1, \lambda_i \geqslant 0 \right\}.$$

Теорема 1.8 (Каратеодоријева теорема). *Ако је $S \subseteq \mathbb{R}^n$ непразан скуп, вриједи*

$$\mathrm{convh}(S) = S_{[n+1]}.$$

За скуп $\{v_1, \ldots, v_k\}$ вектора у $\mathbb{R}^n$, дефинишемо њихов *конусни (или ненегативни) омотач* тако да је

$$\mathrm{cone}(\{v_1, \ldots, v_k\}) = \{x \ : \ x = \sum_i \lambda_i v_i, \text{ гдје је } \lambda_i \geqslant 0 \text{ за } \forall i = \overline{1,k}\}.$$

За $P, Q \subseteq \mathbb{R}^n$, њихова *сума Минковског* је дефинисана тако да је

$$P + Q = \{p + q \ : \ p \in P \text{ и } q \in Q\}.$$

Теорема 1.9. *За $P \subseteq \mathbb{R}^n$, сљедећа тврђења су истовјетна:*

(а) P је полиедар, то јест постоји $A \in \mathbb{R}^{m \times n}$ и $b \in \mathbb{R}^m$ за неко $m \in \mathbb{N}$ тако да је $P = \{x \ : \ Ax \leqslant b\}$;

(б) P је коначно генерисан, то јест постоји коначно много вектора v_i и r_j из $\mathbb{R}^n$ таквих да је $P = \mathrm{convh}(\{v_1, \ldots, v_k\}) + \mathrm{cone}(\{r_1, \ldots, r_t\})$.

Простор линеарности непразног полиедра $P \subseteq \mathbb{R}^n$, у ознаци $\mathrm{lin}(P)$, дефинишемо као

$$\mathrm{lin}(P) = \{y \in \mathbb{R}^n \ : \ x + \lambda y \in P, \ \forall x \in P \text{ и } \forall \lambda \in \mathbb{R}\}.$$

Рецесивни конус непразног полиедра $P \subseteq \mathbb{R}^n$, у ознаци $\mathrm{recc}(P)$, дефинишемо као

$$\mathrm{recc}(P) = \{y \in \mathbb{R}^n \ : \ x + \lambda y \in P, \ \forall x \in P \text{ и } \forall \lambda \geqslant 0\}.$$

У литератури је познат и под називом *карактеристични конус*. Вектор y називамо *смијер рецесије* полиедра P.

Оба скупа садрже исходиште и генерално важи да је $\mathrm{lin}(P) \subseteq \mathrm{recc}(P)$. Полиедар P се назива *заоштреним* ако има екстремну тачку (врх).

Теорема 1.10 (Нека структурна својства полиедра). *Нека је P непразан полиедар у $\mathbb{R}^n$, онда важе сљедећи искази:*

(а) Ако важи да је $P = Q + C$ за неки политоп Q и неки полиедарски конус C, онда је $C = \mathrm{recc}(P)$;

(б) Ако важи да је $P = \{x : Ax \leqslant b\}$, онда је $\mathrm{recc}(P) = \{x : Ax \leqslant \mathbf{0}\}$ *и* $\mathrm{lin}(P) = \{x : Ax = \mathbf{0}\}$*;*

(в) P је заоштрен ако и само ако је $\text{lin}(P) = \{\mathbf{0}\}$*;*

(г) P је политоп (омеђен полиедар) ако и само ако је $\text{recc}(P) = \{\mathbf{0}\}$.

За $a \in \mathbb{R}^n$ и $b \in \mathbb{R}$, неједнакост $a^T x \leqslant b$ се назива *ваљаном* за полиедар P ако је задовољена (важи) за све $x \in P$. Подскуп F полиедра P се назива *лице* полиедра P ако се може представити као $F = P \cap \{x \ : \ a^T x = b\}$, гдје су a и b из неке ваљане неједнакости $a^T x \leqslant b$.

Примијетимо да су $\emptyset$ и P такође лица, и то такозвана *проста* лица. Врхови (ектремне тачке) су лица димензије 0. Лица димензије $\dim(P) - 1$ се називају *стране*, а у литератури су познате и под називом *фацети* (од (старо)француског *facette*, што је деминутив од *face* — лице, појава, изглед).

Теорема 1.11. *Ако је $P = \{x \in \mathbb{R}^n \ : \ Ax \leqslant b\}$, онда је F као непразан подскуп од P лице полиедра P ако и само ако важи*

$$F = \{x \ : \ A^1 x = b^1 \ \text{и} \ A^2 x \leqslant b^2\},$$

гдје је $A = \begin{bmatrix} A^1 \\ A^2 \end{bmatrix}$ *и* $b = \begin{bmatrix} b^1 \\ b^2 \end{bmatrix}$.

Посљедица 1.12. *Свако минимално непразно лице полиедра P је афини потпростор облика $\{x \ : \ A^1 x = b^1\}$ гдје је $A^1 x = b^1$ подсистем система $Ax = b$.*

Посљедица 1.13. *Свако непразно лице полиедра P садржи тачку која је полиномски ограничена величинама матрице A и вектора b.*

Ова посљедица указује на то да је величина сваког врха полиедра полиномски ограничена.

Теорема 1.14. *Сваки рационални полиедар $P = \{x : A \leqslant b\}$ има генераторско представљање* $P_I = \text{convh}(\{v_1, \ldots, v_s\}) + \text{cone}(\{r_1, \ldots, r_t\})$ *тако да је величина сваког генератора — v_i или r_j, полиномски ограничена величином матрице $[A \ b]$.*

Примједба 1.15. Доказано је да се минимално представљање полиедра P састоји од сљедећих скупова:

- скупа тачака од којих је свака из минималног непразног лица полиедра P
- скупа вектора који генеришу простор линеарности.
- скупа вектора који генеришу конус са исходиштем који настаје као пресјек полиедра P и линеарног простора ортогоналног на простор линеарности полиедра P

Теорема 1.16. *За непразан цјелобројни омотач P_I у рационалном полиедру P, важи да је $P_I = Q + C$, гдје је Q неки цјелобројан политоп, а* $C = \text{recc}(P)$.

Слиједи основно тврђење на основу ког ће се доказати да се цјелобројни програм налази у класи NP.

Теорема 1.17. *Цјелобројни омотач P_I у рационалном полиедру $P = \{x \ : \ A \leqslant b\}$ има генераторску представу* $P_I = \text{convh}(\{z_1, \ldots, z_k\}) + \text{cone}(\{r_1, \ldots, r_h\})$ *тако да је величина сваког генератора — z_i или r_j, полиномски ограничена величином матрице $[A \ b]$.*

Доказ. Означимо са δ величину матрице $[A \ b]$. Према Теореми 1.14, полиедар P се може представити збиром $Q + C$ гдје је $Q = \operatorname{convh}(\{v_1, \ldots, v_s\})$ и $C = \operatorname{cone}(\{r_1, \ldots, r_h\})$, при чему је величина сваког v_i и сваког r_j полиномски ограничена по δ. Без губљења општости, може се претпоставити да су r_j цјелобројни вектори. Према претходној теореми $\operatorname{recc}(P_I) = \{r_1, \ldots, r_h\}$. Показаћемо да је

$$P_I = \operatorname{convh}(\{z_1, \ldots, z_k\}) + C, \tag{1.1}$$

гдје су $z_1, \ldots, z_k$ цјелобројне тачке у скупу $Q + Y$ и

$$Y = \{y \ : \ y = \sum_{j=1}^{h} \lambda_j r_j, \ 0 \leqslant \lambda_j \leqslant 1, \ j = \overline{1, h}, \text{ и највише } n \text{ коеф. } \lambda_j \text{ је позитивно}\}.$$

Величина сваког вектора z_i је полиномски ограничена по δ. Пошто су сви r_j цјелобројни, највише n их се користи за представљање одговарајућег вектора z_i, па је величина сваке цјелобројне тачке у Q је полиномски ограничена по δ.

Са овим поставкама, остала је још (1.1) једнакост да се докаже. Довољно је доказати да свако минимално непразно лице F полиедра P_I садржи бар једну тачку из $\{z_1, \ldots, z_k\}$ (погледати примједбу на Теорему 1.14). Нека је z цјелобројна тачка у F. Пошто је $z \in P$, важи да је

$$z = q + \sum_{j=1}^{h} \mu_j r_j,$$

за неко $\mu_j \geqslant 0$. Према Каратеодоријевој теореми, можемо претпоставити да је међу коефицијентима μ_j највише n позитивно (па је из скупа $Q + Y$). Нека је z' вектор

$$z' = q + \sum_{j=1}^{h} (\mu_j - \lfloor \mu_j \rfloor) r_j.$$

Слиједи да је z' цјелобројан вектор и стога један од вектора из $\{z_1, \ldots, z_k\}$. Лако се увиђа да $z' \in F$. Овим је доказ завршен. □

Теорема 1.18. *Цјелобројни програм припада класи NP.*

Доказ. За доказ тврдње је довољно разматрати сложеност полиномски истовјетног проблема из претходне теореме, односно језика:

$$J_{\text{ЦП}} = \{(A, b) \ : \ \exists x \in \mathbb{Z}^n \text{ тако да је } Ax \leqslant b\},$$

при чему претпостављамо да је $A \in \mathbb{Q}^{m \times n}$ и $b \in \mathbb{Q}^n$ (практично гледано, интересују нас само рационални програми). Другим ријечима, треба одлучити да ли је скуп

$$\mathbb{Z}^n \cap \{x \in \mathbb{R}^n \ : \ Ax \leqslant b\} \tag{1.2}$$

непразан. Случај када је

$$P = \{x \in \mathbb{Z}^n \ : \ Ax \leqslant b, \ x \geqslant 0\}$$

је полиномски сводљив на (1.2). Запис програма без захтјева за одређивањем максимума, односно уз измјену функције циља у једнакост

$$\begin{aligned} \sum_{j=1}^{n} c_j x_j &= v \\ \sum_{j=1}^{n} a_{ij} x_j &\leqslant b_i, \text{ за } i = 1, 2, \ldots, m; \\ x_j &\geqslant 0, \text{ за } j = 1, 2, \ldots, n; \\ x_j &\in \mathbb{Z}, \text{ за } j = 1, 2, \ldots, n, \end{aligned}$$

гдје је v нека вриједност, се може преобликовати у (1.2). Означимо са ω оптимално рјешење допустивог и ограниченог ЦП-а, а M вриједност одговарајућег полинома у δ (величином матрице $[A\ b]$). Важи да је $\omega \leqslant M \cdot \sum_{j=1}^{n} |c_j|$. За доњу границу, лако увиђамо да је $-|\omega_{\text{ЛП}}| \leqslant \omega$, гдје је $\omega_{\text{ЛП}}$ оптимално рјешење ЛП-релаксације цјелобројног програма. Ипак, пошто је сâмо $\omega_{\text{ЛП}}$ добијено за неко допустиво базно рјешење, важи $|\omega_{\text{ЛП}}| \leqslant M \cdot \sum_{j=1}^{n} |c_j|$. Самим тим је $|\omega| \leqslant M \cdot \sum_{j=1}^{n} |c_j|$. Сада се може рјешавати одговарајући допустив ЦП за сваку цјелобројну вриједност ω у размаку

$$-M \cdot \sum_{j=1}^{n} |c_j| - 1 \leqslant \omega \leqslant M \cdot \sum_{j=1}^{n} |c_j|.$$

Додатно се може искористити алгоритам бинарног претраживања ради добијања бољег времена у анализи.

С друге стране, јасно је да се рјешење цјелобројног програма може узети за увјерење и да се оно при том може провјерити у полиномском времену (мањем од или једнаком M).

Уколико је проблем неомеђен, на основу претходне теореме, постоје допустива тачка z и зрак r, обоје полиномски ограничени величином матрице $[A\ b]$, такви да је $c^T r > 0$. Самим тим они сачињавају одговарајуће увјерење. □

Теорема 1.19. *Цјелобројни програм је NP-тежак.*

Доказ. Довољно је доказати да је $SAT \lhd$ ЦП. За сваку логичку промјенљиву p_i у примјерку проблема SAT, уводимо цјелобројну промјенљиву x_i тако да је:

- $x_i = 1$ ако је $p_i = \top$,
- $x_i = 0$ ако је $p_i = \bot$.

Тако да одмах имамо неједнакости $0 \leqslant x_i \leqslant 1$ за свако p_i. У наставку посматрамо проблем SAT записан у конјунктивној форми, пошто се сваки логички израз може превести у одговарајућу конјуктивну форму (наравно, формално је доказано да је и SAT у конјуктивној форми NP-комплетан). Било која елементарна дисјункција се претвара у одговарајућу неједнакост на сљедећи начин:

- p_i претварамо у x_i
- $\neg p_i$ претварамо у $(1 - x_i)$
- $\bigvee q_i$ претварамо у $\sum y_i \geqslant 1$, гдје је $y_i = x_i$ ако је $q_i = p_i$, а $y_i = (1 - x_i)$ ако је $q_i = \neg p_i$.

Јасно је да се претварање извршава у полиномском времену, пошто се за дату величину улаза — дужину логичког израза, претварање врши у времену једнаком највише четири дужине улаза.

Даље, остало је да се докаже да је логички израз задовољив ако је одговарајући ЦП допустив. Ако постоји задовољиво придруживање вриједности $\top$ и $\bot$ логичким промјенљивим примјерку проблема SAT у конјуктивној форми, онда важи да је задовољена и свака елементарна дисјункција (због конјуктивних веза), а то значи да x задовољава сваку неједнакост која одговара елементарној дисјункцији. Стога, ЦП је допустив, штавише вриједи и обрнута тврдња.

Конкретно, доказано је да је 0-1 програм NP-тежак, но самим тим то важи и за општији случај. ☐

Посљедица 1.20. *Цјелобројни програм је NP-комплетан.*

1.4 О начинима рјешавања цјелобројних програма

> *„Ако ћемо — лако ћемо! Ако нећемо — како ћемо?”*
> Народна пословица

Проучавање цјелобројних програма, због карактеристичне тежине и сложености, одвија се у два смијера. У првом, проучавају се посебни модели (цјелобројни програми) и за неке од њих су до данас развијени дјелотворни алгоритми, док се у другом ради о општим алгоритмима који не зависе од посебне структуре цјелобројног програма.

1.4.1 Дјелотворно рјешиви програми

Интуитивно, намеће се питање: „Под којим околностима ЛП-релаксација има цјелобројно рјешење?”. Природно, имаће цјелобројно рјешење ако је полиедар — скуп допустивих рјешења програма цјелобројан, то јест ако има само цјелобројне врхове.

Матрица са елементима $-1, 0, 1$ је *потпуно унимодуларна* ако детерминанта сваке квадратне подматрице може да има само вриједности из скупа {-1, 0, 1}.

Имајући у виду потпуну унимодуларност Хофман и Краскал су у [40] доказали сљедеће тврђење.

Теорема 1.21 (Хофман-Краскал)**.** *За матрицу $A \in \mathbb{Z}^{m\times n}$, сљедеће тврдње су истовјетне:*

- *A је потпуно унимодуларна.*
- *За сваки вектор $b \in \mathbb{Z}^m$, полиедар $P = \{x \in \mathbb{R}^n \ : \ Ax \leqslant b, \ x \geqslant \mathbf{0}\}$ је цјелобројан.*

Из чињенице да је матрица A потпуно унимодуларна ако и само ако је A^T потпуно унимодуларна и претходне теореме, слиједи да цјелобројан програм са цјелобројним подацима и потпуно унимодуларном матрицом својих ограничења има цјелобројна примална и дуална оптимална рјешења.

Посљедица 1.22. *За потпуно унимодуларну матрицу A величине $m \times n$, вектор $b \in \mathbb{Z}^m$ и $c \in \mathbb{Z}^n$ важи да оба оптимума у једначини ЛП-дуалности*

$$\max\{c^T x \ : \ Ax \leqslant b\} = \min\{y^T b \ : \ y^T A = c^T, \ y \geqslant \mathbf{0}\},$$

имају цјелобројна оптимална рјешења (ако су коначна).

Посљедица 1.23. *Ако претпоставимо да за матрицу A величине $m \times n$, вектор $b \in \mathbb{Z}^m$ и вектор $c \in \mathbb{R}^n$ важи да програм*

$$\max\{c^T x \ : \ Ax \leqslant b, x \geqslant \mathbf{0}\}$$

има оптимално рјешење x^* и то такво да колоне матрице A које одговарају позитивним компонентама вектора x^* формирају потпуно унимодуларну матрицу, онда дати програм има цјелобројно рјешење.

Карактеристичан проблем за који је матрица ограничења проблема потпуно унимодуларна је Проблем најјефтинијег мрежног протока.

Проблем 1.3 (Проблем најјефтинијег мрежног протока)**.** Потребно је пронаћи најјефтинији проток робе кроз мрежу. Мрежа (усмјерен граф) се састоји од скупа чворова (спојева) S и скупа (продукто)вода V који повезују чворове. Вод v_{ij} из скупа V се односи на уређен пар (i, j) (проток иде од i до j) гдје су $i, j \in S$. Не морају нужно сви чворови бити међусобно повезани. Један пар чворова може бити повезан са више водова, другим ријечима могу постојати v_{ij}^1 и v_{ij}^2 као два различита вода из скупа V.

Сваком воду се придружују четири вриједности:

- количина протока p_v, при чему су дозвољене негативне вриједности (указују да проток иде од j ка i)
- доња граница d_v, која одређује најмањи дозвољен проток кроз вод v (0 је подразумијевана вриједност).
- горња граница g_v, која одређује највећи дозвољен проток кроз v (∞ је подразумијевана вриједност).
- цијена (трошак, циљна вриједност) протока кроз вод v, у ознаци c_v.

Сваком чвору s је придружено o_s, то јест количина опскрбљивања тог чвора. Ако је $o_s > 0$, чвор се назива отпремником (извором, пумпом), ако је $o_s < 0$ чвор се назива пријемником (одливним чвором), а ако је $o_s = 0$, онда се назива проточним чвором. Збир свих отпремања мора бити једнак збиру свих пријема, иначе проблем није допустив.

I_s је скуп свих водова у којима се проток креће од чвора s (излазни водови из чвора s), а U_n је скуп свих водова у којима се проток креће ка чвору s (улазни водови за чвор s).

Проблем се може записати као линеарни програм

$$\min \sum_{v \in V} c_v p_v$$

$$\sum_{v \in U_s} p_v - \sum_{v \in I_s} p_v = o_s, \forall s \in S$$

$$d_v \leqslant p_v \leqslant g_v, \forall v \in V$$

Уколико програм ојачамо условом цјелобројности, претходна тврђења говоре да ће допустива базна рјешења бити цјелобројна, па ће се самим тим и даље моћи користити методи линеарног програмирања за његово рјешавање.

У пракси (индустрији) се јавља неколико различитих и битних варијанти овог проблема:

- Проблем најкраћег пута — означени су полазни чвор a и одредишни чвор b; уведена су ограничења $p_v \in \{0,1\}$, $o_a = 1$, $o_b = -1$, $o_s = 0$ за $s \notin \{a,b\}$ и $g_v = 1$ за $v \in V$.
- Проблем транспорта — по један чвор за сваког снадбјевача и по један чвор за сваког наручиоца робе; водови (ивице графа) одговарају путевиме доставе од снадбјевача до наручиоца; $g_v = \infty$; $o_i \in \{S_1, \ldots, S_m\}$ за снадбјеваче и $o_j \in \{N_1, \ldots, N_n\}$ за наручиоце.
- Проблем максималног протока — означени су полазни чвор a и одредишни чвор b; уведен је јединствен помоћни вод v_{ba}, за којег важи да је $c_{v_{ba}} = -1$, а за све остале је $c_v = 0$; $o_s = 0$; $d_s = 0$.

Полиедар одређен системом линеарних ограничења је *потпуно дуално цјелобројан* (ПДЦ) ако за сваку функцију циља која има цјелобројне коефицијенте, дуални линеарни програм има оптимално рјешење (кад год оно постоји).

Теорема 1.24 (Едмондс-Џајлс [25]). *Ако је $P = \{x : Ax \leqslant b\}$ ПДЦ и b цјелобројан вектор, онда је P цјелобројан.*

Поред потпуно унимодуларних матрица, важну улогу у теорији полиедара имају балансиране матрице Клода Бержа. 0-1 матрица је *балансирана* ако не садржи квадратну подматрицу непарног реда која има по двије јединице у врстама и колонама. Показало се да имају битну улогу у проблемима паковања и покривања.

Теорема 1.25 ([16, 29]). *Ако је A балансирана 0-1 матрица, онда су сљедећи полиедри цјелобројни:*

- $P = \{x \;:\; Ax \leqslant \mathbf{1},\ x \geqslant \mathbf{0}\}$ *(проблем паковања)*
- $Q = \{x \;:\; Ax \geqslant \mathbf{1},\ \mathbf{0} \leqslant x \leqslant \mathbf{1}\}$ *(проблем покривања)*
- $R = \{x \;:\; Ax = \mathbf{1},\ x \geqslant \mathbf{0}\}$ *(проблем партиционисања)*

Трумпер је проширио дефиницију балансираних матрица на $\{-1,0,1\}$-матрице [58], а Конфорти и Корнуижол су у [17] доказали одговарајуће проширење претходне теореме.

Лако се уочава да је $\{-1,0,1\}$-матрица A балансирана ако и само ако је A^T уравнотежена. Штавише, A је балансирана (потпуно унимодуларна) ако и само ако је свака њена подматрица балансирана (потпуно унимодуларна), па се претходна тврђења могу примијенити на све њене подматрице. Класа потпуно унимодуларних матрица чини поткласу балансираних матрица, то јест потпуно унимодуларне матрице су увијек балансиране [14]. Иначе, постоји алгоритам полиномског времена за препознавање балансираних матрица [60].

Занимљиво је да је још током XIX вијека њемачки физичар Густав Кирхоф, познат по Кирхофовим законима, описао класу потпуно унимодуларних матрица, коју су Хофман, Краскал, Харолд Кун и Алберт Такер поново „открили” 50-тих година XX вијека [40].

С друге стране, не разматрајући програме са посебном структуром, Хендрик В. Ленстра је у [47] доказао да је цјелобројни програм са фиксираним бројем промјенљивих рјешив у полиномском времену, док је Христос Х. Пападимитриоу у [52] доказао да постоји алгоритам псеудополиномског времена за рјешавање цјелобројног програма са фиксираним бројем ограничења.

Алгоритам се извршава у *псеудополиномском времену* ако је његово вријеме извршавања полиномско по вриједностима улазних података и величини проблема, али не по дужини записа улаза. Такви алгоритми су технички експоненцијалне функције свога улаза и стога се не сматра да су полиномског времена.

1.4.2 О општим методима

Први егзактан алгоритам за рјешавање општег цјелобројног програма, заснован на такозваном методу одсијецајућих равни, пронашао је 1958. године Ралф Е. Гомори. Идеју одсијецајућих равни су први предложили Данциг, Реј Фулкерсон и Селмер Џонсон за чувени проблем трговачког путника још 1954. год. [57], но према контрапримјеру за Претпоставку 9, дати метод је имао ману — нађени су примјери за које не конвергира.

До данас је развијено много различитих егзактних метода за рјешавање општег цјелобројног програма и сви они имају, међусобно их сагледавајући, одређене предности и мане. Савремени алгоритми цјелобројног програмирања су засновани на:

- методима одсијецајућих равни
- методима гранања и омеђивања
- методима гранања и одсијецања
- методима гранања и оцјењивања

У овом раду биће представљене и разматране Гоморијеве одсијецајуће равни, као оруђе које се користи у више различитих метода за рјешавање цјелобројних програма и које је од свог увођења, давне 1958. године, још увијек предмет пажње и проучавања математичара.

Глава 2

Методи одсијецајућих равни

> *А кад зачу Краљевићу Марко,*
> *Трже сабљу, одсијече му главу,*
> *Па је баци у воду Марицу*
>
> Народни пјесник

Методи одсијецајућих равни служе за рјешавање општих проблема конвексне оптимизације (не само цјелобројних програма). Засновани су на употреби хиперравни помоћу којих се дати проблем апроксимира низом боље одређених проблема. Главно својство тих хиперравни је да се посматрана тачка налази са једне, а све тачке које задовољавају почетне услове, са друге стране равни. Обично су мање дјелотворни од других метода, али имају бројне предности због којих постају одличан избор у одређеним ситуацијама:

- не захтјевају диференцијабилност
- некад је могуће искористити карактеристичну структуру проблема
- не захтјевају израчунавање вриједности функције циља и функција ограничења у сваком кораку (итерацији). Због тога су корисни за проблеме са великим бројем ограничења.
- помоћу њих се проблем може подијелити на мање проблеме који се затим могу рјешавати појединачно

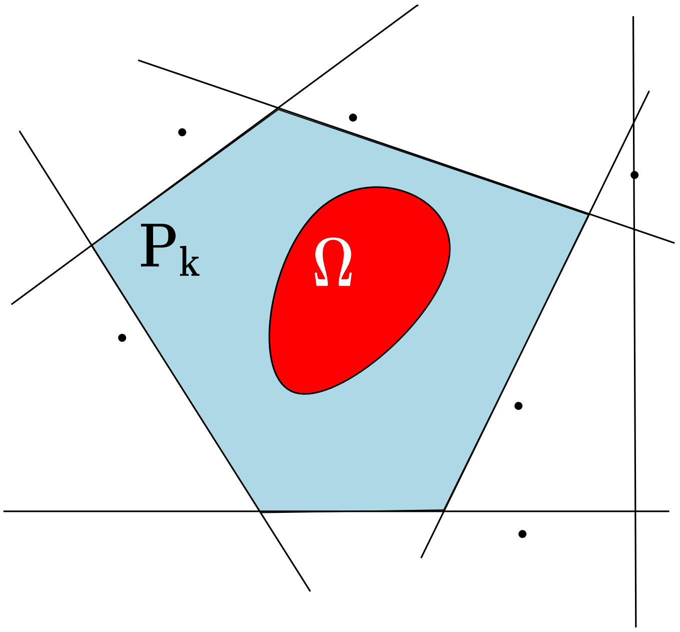
2.1 Основни алгоритам

Треба пронаћи тачку из скупа $\Omega \subseteq \mathbb{R}^n$, који се назива *циљни скуп*, или треба показати да је Ω празан. На примјер, ако се тражи тачка максимума функције, скуп Ω може бити скуп тачака максимума.

У првом кораку се одређује тачка $x \in \mathbb{R}^n$ и провјерава се да ли се она налази у скупу Ω. У случају да $x \in \Omega$ наш задатак је ријешен, док се у противном одређује хиперраван са својством да се тачка x налази с једне стране, а све тачке скупа Ω с друге стране равни. На примјер, нађемо $a \neq 0$ и b такве да

$$a^T y \leqslant b \text{ за } y \in \Omega, \;\; a^T x > b.$$

Ова хиперраван се зове *одсијецајућа раван* пошто одсијеца (избацује) скуп $\{y \,:\, a^T x > b\}$ из даље претраге (тачке одсјеченог скупа не могу припадати скупу Ω).



Слика 2.1

Алгоритам 2.1 Основни алгоритам метода одсијецајућих равни у псевдокоду

```
нађи полиедар $P_0$ који садржи $\Omega$
$k \leftarrow 0$
пронађеноРјешење $\leftarrow$ нетачно
празан $\leftarrow$ нетачно
све док је (пронађеноРјешење $=$ нетачно) или (празан $=$ нетачно)
    изабери тачку $x_k \in P_k$
    ако је $x_k \in \Omega$
        пронађеноРјешење $\leftarrow$ тачно
    иначе
        пронађи $a_k$ и $b_k$ т.д. је $a_k^T y \leqslant b_k$ (за $y \in \Omega$) и $a_k^T x_k > b_k$
        $P_{k+1} \leftarrow P_k \cap \{y : a_k^T y \leqslant b_k\}$
        ако је $P_{k+1} = \emptyset$ онда
            празан $\leftarrow$ тачно
    $k \leftarrow k + 1$
ако је (пронађеноРјешење $=$ тачно)
    резултат је $x_k$
иначе
    резултат је $\emptyset$
```

Претходни алгоритам производи низ угнијежђених полиедара $\mathcal{P}$ такав да за све $P_i \in \mathcal{P}$ и $i \in \mathbb{N}_0$ важи да

$$P_0 \supset P_1 \supset \ldots \supset P_k \supset \ldots \supset \Omega,$$

то јест који све боље и боље, споља, карактеришу циљни скуп Ω. Апроксимација се побољшава око тренутно изабране тачке која највише обећава у датој етапи.

Уопштено гледано, ова процедура може бити бесконачна (нпр. x постоји и $x \notin \Omega$ за свако $P \in \mathcal{P}$). За бесконачан алгоритам одсијецајућих равни кажемо да *конвергира* ако производи низ изабраних тачака x_k у ком свака тачка нагомилавања $\overline{x}$ припада Ω. У пракси, алгоритам се зауставља када се x_k довољно добро приближи скупу Ω. Као што се може примјетити, метод одсијецајућих равни и не мора да конвергира, но ако конвергира његов ред конвергенције је 1. Узевши у обзир да не захтјева диференцијабилност, сасвим је прихватљив.

Полиедар P_k представља наше знање о могућем положају циљне (оптималне) тачке након k-тог корака метода одсијецајућих равни. Мјера P_k представља оцјену наше несигурности у погледу положаја дате тачке, па ако је P_k велик, још увијек смо поприлично несигурни гдје се налази наша циљна тачка.

Једна очита мјера омеђеног скупа (полиедра) P_k је пречник d најмање кугле $\{\hat{x}+t : ||t|| \leqslant \frac{d}{2}\}$ која садржи P_k. У том случају знамо да нам удаљеност рјешења проблема од тачке $\hat{x}$ може бити највише $\frac{d}{2}$ од $\hat{x}$, па коришћењем ове оцјене можемо пратити напредовање поступка посматрајући одговарајуће пречнике кугла са центром у $\hat{x}$ које садрже полиедре P_{k+i} (пошто је $P_{k+1} \subset P_k$, тј. пречник се не повећава).

Разматрањем односа запремина омеђених полиедара P_k

$$\frac{\text{vol}(P_{k+1})}{\text{vol}(P_k)}$$

у k-том кораку поступка је још један од начина који може да послужи за праћење напретка.

С друге стране, није тешко примјетити да су нека питања остала без одговора, а међу њима и два веома важна:

1. Како изабрати тачку $x_k \in P_k$?
2. Како одредити раван која раздваја тачку x_k од скупа Ω?

Сваки нови начин избора тачке нам даје нови метод одсијецајућих равни, међу којма сваки има своје добре и лоше стране у зависности од тога колико је тешко изабрати сљедећу тачку, каква је теоријска сложеност метода и као и његова практична примјена.

2.2 Генерички метод за цјелобројне програме

У генеричком методу начин рјешавања цјелобројних програма се ослања на рјешавање низа линеарних програма (избор тачке), који један за другим пружају све боље и боље апроксимације рјешења цјелобројног програма.

Генерички метод одсијецајућих равни за цјелобројне програме

1. Наћи x^* — оптимално рјешење ЛП-релаксације.
2. Ако је x^* цјелобројно, прекинути поступак пошто је x^* тражено рјешење
3. Ако x^* није у потпуности цјелобројно, наћи неједнакост која:
    - задовољава сва допустива рјешења ЦП-а;
    - нарушена је у x^*.
4. Додати дату неједнакост ЛП-релаксацији и врати се на корак 1.

Додавањем неједнакости (тј. одсијецајућих равни) низу ЛП-релаксација, низ оптималних вриједности узастопних линеарних програма је нерастући низ горњих ограничења оптималне вриједности цјелобројног програма. Главни проблем у примјењивању генеричког метода одсјецајућих равни лежи у проналажењу неједнакости из трећег корака, чији се смисао огледа у покушају карактеризације цјелобројног омотача програма, бар у близини цјелобројног оптималног рјешења.

Глава 3

Гоморијеве одсијецајуће равни

> *„У Морнарици сам и даље био савјетник, па сам наставио да радим на њиховим проблемима током својих мјесечних путовања у Вашингтон. На једном од ових путовања, група је представила линеаран програм који је био модел Морнаричке Оперативне Групе (енгл. Navy Task Force). Један од излагача је прокоментарисао како би било лијепо да се имају цјелобројни резултати, пошто 1,3 носача авиона, као резултат, не значи ништа (конкретно)."*
>
> Ралф Е. Гомори

3.1 Хватал-Гоморијеве одсијецајуће равни

Имајући у виду цјелобројни програм

$$\max \sum_{j=1}^{k} c_j x_j$$

$$\sum_{j=1}^{k} a_{ij} x_j \leqslant b_i, \text{ за } i = 1, 2, \ldots, m; \tag{i}$$

$$x_j \geqslant 0, \text{ за } j = 1, 2, \ldots, k; \tag{ii}$$

$$x_j \in \mathbb{Z}, \text{ за } j = 1, 2, \ldots, k, \tag{iii}$$

за $\tau \in \mathbb{R}_+^m$ важи да за свако x које задовољава неједнакост (i), вриједи и неједнакост

$$\sum_{i=1}^{m} \tau_i \sum_{j=1}^{n} a_{ij} x_j \leqslant \sum_{i=1}^{m} \tau_i b_i, \tag{i'}$$

или, што је истовјетно,

$$\sum_{j=1}^{n} \left\lfloor \sum_{i=1}^{m} \tau_i a_{ij} \right\rfloor x_j + \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \tau_i a_{ij} - \left\lfloor \sum_{i=1}^{m} \tau_i a_{ij} \right\rfloor \right) x_j \leqslant \sum_{i=1}^{m} \tau_i b_i.$$

Слиједи да сва рјешења која задовољавају (*i*) и (*ii*) задовољавају и

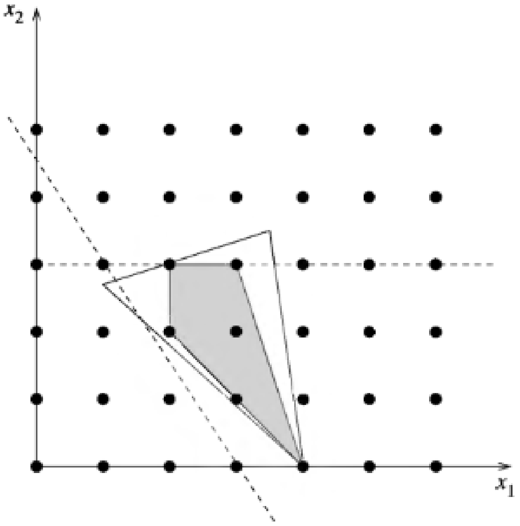
$$\sum_{j=1}^{n}\left\lfloor\sum_{i=1}^{m}\tau_i a_{ij}\right\rfloor x_j \leqslant \sum_{i=1}^{m}\tau_i b_i. \qquad (ii')$$

Важно је примјетити да још увијек није кориштен услов (*iii*). Коначно, из (*iii*) слиједи да сва рјешења по (*ii*′) и (*iii*) задовољавају *Хватал-Гоморијеву одсијецајућу раван*

$$\sum_{j=1}^{n}\left\lfloor\sum_{i=1}^{m}\tau_i a_{ij}\right\rfloor x_j \leqslant \left\lfloor\sum_{i=1}^{m}\tau_i b_i\right\rfloor. \qquad (iii')$$

Формално, једначина описује полупростор, али је суштина у хиперравни којом је тај полупростор на неки начин одређен, односно суштина је у одређивању њених коефицијената.

Да би била јаснија улога Хватал-Гоморијевих равни и да би се увидјела област коју оне покривају (гдје се могу налазити за конкретан програм), битно је запазити да је (*iii*′) задовољено за сва рјешења која задовољавају услове (*ii*′) и (*iii*), али није, пак, за сва она рјешења која задовољавају (*i*) и (*ii*). Другим ријечима, Хватал-Гоморијеве одсијецајуће равни су ваљане неједнакости за цјелобројни омотач у полиедру одговарајуће ЛП-релаксације цјелобројног програма.



Слика 3.1 Хватал-Гоморијеве одсијецајуће равни

Примјер 3.1 (Хватал-Гоморијеве одсијецајуће равни). Размотримо програм

$$\max 2x_1 + x_2$$

уз услове:

$$\begin{aligned} 7x_1 + x_2 &\leqslant 28; \\ -x_1 + 3x_2 &\leqslant 7; \\ -8x_1 - 9x_2 &\leqslant -32; \\ x_1, x_2 &\geqslant 0; \\ x_1, x_2 &\in \mathbb{Z}. \end{aligned}$$

Избор $\tau_1 = 0$, $\tau_2 = 1/3$, $\tau_3 = 1/3$ даје раван одсијецања $-3x_1 - 2x_2 \leqslant -9$, док избор $\tau_1 = 1/21$, $\tau_2 = 7/22$, $\tau_3 = 0$ даје раван одсијецања $x_2 \leqslant 3$ (Сл. 3.1). ■

Иако примјер предочава дјеловање одсијецајућих равни, још увијек није дат систем (метод) за њихово генерисање, који би довео, примјеном генеричког метода одсијецајућих равни, до цјелобројног рјешења. Такав метод биће представљен у наставку, када ће се и ријешити проблем из примјера.

3.2 Гоморијеве одсијецајуће равни

Гоморијеве одсијецајуће равни су равни опште сврхе, то јест не искориштавају неку посебну структуру, нити су пак везане за неки тип цјелобројног програма, него настају од допустивих базних рјешења одговарајуће ЛП-релаксације. За почетак, претпоставиће се да су a_{ij}, b_i, и c_j, природно, сви цјелобројни. Нека је x_0 дефинисано на сљедећи начин

$$x_0 - \sum_{j=1}^{n} c_j x_j = 0.$$

За $i = 1, 2, \ldots, m$, дефинишу се ненегативне *изједначавајуће* промјенљиве x_{n+i} на сљедећи начин

$$x_{n+i} + \sum_{j=1}^{n} a_{ij} x_j = b_i.$$

Примјетно је да су у оваквој поставци изједначавајуће промјенљиве и x_0 цјелобројни, управо због претпоставке да су a_{ij}, b_i, и c_j сви цјелобројни. Важна посљедица овога је да се у увећаном облику програма и даље разматра цјелобројни програм.

За све $1 \leqslant i, j \leqslant m$, дефинише се

$$a_{i,(n+j)} = \begin{cases} 1 & \text{ако је } i = j \\ 0 & \text{ако је } i \neq j \end{cases}$$

и нека је $a_{i,0} = 0$ за све $1 \leqslant i \leqslant m$. Узевши да је $k = n + m$, за $0 \leqslant j \leqslant k$ дефинише се још

$$a_{0j} = \begin{cases} 1 & \text{ако је } j = 0 \\ -c_j & \text{ако је } 1 \leqslant j \leqslant n \\ 0 & \text{ако је } n + 1 \leqslant j \leqslant n + m, \end{cases}$$

и на крају $b_0 = 0$. Значи, ако је A матрица главних ограничења програма, онда је

$$T = \begin{bmatrix} 1 & -c^T & \mathbf{0} \\ \mathbf{0} & A & I \end{bmatrix}$$

матрица програма у *увећаном облику*. Базно рјешење ЛП-релаксације се може добити одговарајућом обрадом система једначина

$$\sum_{j=0}^{k} a_{ij}x_j = b_i, \text{ за } \ i = 0, 1, 2, \ldots, m.$$

Скуп индекса промјенљивих $\{0, 1, 2, \ldots, k\}$ дијелимо на скуп базних индекса $\boldsymbol{\beta} = \{\beta_0 = 0, \beta_1, \beta_2, \ldots, \beta_m\}$ и скуп небазних индекса $\boldsymbol{\eta} = \{\eta_1, \eta_2, \ldots, \eta_{k-m}\}$ тако да систем једначина

$$\sum_{j=0}^{m} a_{i\beta_j}x_{\beta_j} + \sum_{j=1}^{k-m} a_{i\eta_j}x_{\eta_j} = b_i, \ i = 0, 1, 2, \ldots, m \tag{3.1}$$

има јединствено ненегативно рјешење x^* за $x^*_{\eta_1} = x^*_{\eta_2} = \ldots = x^*_{\eta_{k-m}} = 0$. Даље, једначине са базним промјенљивим се могу ријешити по небазним промјенљивим

$$x_{\beta_i} + \sum_{j=1}^{k-m} \overline{a}_{\beta_i\eta_j}x_{\eta_j} = x^*_{\beta_i}, \ i = 0, 1, 2, \ldots, m \tag{3.2}$$

гдје су $\overline{a}_{\beta_i\eta_j}$ елементи матрице $\overline{A} = T_{\boldsymbol{\beta}}^{-1}T_{\boldsymbol{\eta}}$. Ова једначина у суштини представља i-ти ред симплекс-табеле. Коначно, имајући у виду ненегативност и цјелобројност x_j-а, добијају се тражене *Гоморијеве одсијецајуће равни*

$$x_{\beta_i} + \sum_{j=1}^{k-m} \left\lfloor \overline{a}_{\beta_i\eta_j} \right\rfloor x_{\eta_j} \leqslant \left\lfloor x^*_{\beta_i} \right\rfloor, \ i = 0, 1, 2, \ldots, m. \tag{G_{β_i}}$$

Може се примијетити да је неједнакост одсијецајуће равни (G_{β_i}) нарушена у базном рјешењу сваки пут када $x^*_{\beta_i} \notin \mathbb{Z}$. Штавише, Гоморијеве равни имају цјелобројне коефицијенте, а та чињеница је изузетно важна, јер је управо због тога могуће поновити претходни поступак. На крају, одузимањем (3.2) од (G_{β_i}) добијају се у суштини исте (истовјетне) неједнакости, такозване *разломљене Гоморијеве одсијецајуће равни*,

$$\sum_{j=1}^{k-m} (\lfloor \overline{a}_{\beta_i\eta_j} \rfloor - \overline{a}_{\beta_i\eta_j})x_{\eta_j} \leqslant \lfloor x^*_{\beta_i} \rfloor - x^*_{\beta_i}. \tag{$\overline{G}_{\beta_i}$}$$

гдје $i = 0, 1, 2, \ldots, m$.

Примједба 3.1 (Гоморијеве равни су Хватал-Гоморијеве равни)*.* Ако су $\lambda_{t1}, \lambda_{t2}, \ldots, \lambda_{tm}$ реални бројеви такви да је према (3.1) и (3.2)

$$\sum_{i=1}^{m} \lambda_{ti} \cdot b_i = x^*_t,$$

дата агрегација се може записати, не обазирући се конкретно на распоред базних и небазних промјенљивих, као

$$\sum_{i=1}^{m} \lambda_{ti} \sum_{j=1}^{k} a_{ij}x_j = x^*_t.$$

Разломљена Гоморијева одсијецајућа раван, која се изводи из реда t симплексне табеле, у овом случају је облика

$$(G_t) - \sum_{j=1}^{k}\sum_{i=1}^{m}\lfloor\lambda_{ti}\rfloor a_{ij}x_j \leqslant \left\lfloor \sum_{i=1}^{m}\lambda_{ti}b_i \right\rfloor - \sum_{i=1}^{m}\lfloor\lambda_{ti}\rfloor b_i.$$

Уз чињеницу да је за $\lambda \in \mathbb{R},\ a \in \mathbb{Z}$

$$\lfloor(\lambda - \lfloor\lambda\rfloor)a\rfloor = \lfloor\lambda a\rfloor - \lfloor\lambda\rfloor a,$$

из претходног се добија да је

$$\sum_{j=1}^{k}\left\lfloor \sum_{i=1}^{m}(\lambda_{ti} - \lfloor\lambda_{ti}\rfloor)a_{i}j \right\rfloor x_j \leqslant \left\lfloor \sum_{i=1}^{m}(\lambda_{ti} - \lfloor\lambda_{ti}\rfloor)b_i \right\rfloor,$$

односно ако се узме да је $u_{ti} = \lambda_{ti} - \lfloor\lambda_{ti}\rfloor$, Гоморијеве одсијецајуће равни се могу записати у облику

$$\sum_{j=1}^{n}\left\lfloor \sum_{i=1}^{m}u_{ti}a_{ij} \right\rfloor x_j \leqslant \lfloor u_{ti}b_i\rfloor,$$

а то је ништа друго до облик за Хватал-Гоморијеве одсијецајуће равни.

3.3 Гоморијев метод

Настављајући са рјешавањем проблема из Примјера 3.1 (Хватал-Гоморијеве одсијецајуће равни), уводе се изједначавајуће промјенљиве $x_3, x_4,$ и x_5. Оптимална база ЛП-релаксације се састоји од базних индекса $\{0, 1, 2, 5\}$ као што се може видјети из одговарајуће симплекс-табеле

$$\begin{array}{cccccccc}
x_0 & & & +\frac{7}{22}x_3 & +\frac{5}{22}x_4 & & = & \frac{21}{2} \\
 & x_1 & & +\frac{3}{22}x_3 & -\frac{1}{22}x_4 & & = & \frac{7}{2} \\
 & & x_2 & +\frac{1}{22}x_3 & +\frac{7}{22}x_4 & & = & \frac{7}{2} \\
 & & & \frac{3}{2}x_3 & +\frac{5}{2}x_4 & +x_5 & = & \frac{55}{2}.
\end{array}$$

Из претходних једначина добијају се сљедеће Гоморијеве одсијецајуће равни

$$x_0 \leqslant 10; \tag{0}$$

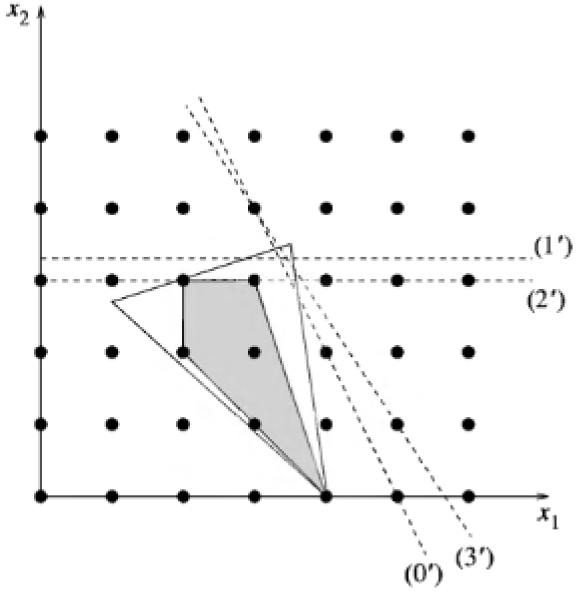
$$x_1 - x_4 \leqslant 3; \tag{1}$$

$$x_2 \leqslant 3; \tag{2}$$

$$x_3 + 2x_4 + x_5 \leqslant 27, \tag{3}$$

чији је облик у простору почетних промјенљивих

$$2x_1 + x_2 \leqslant 10; \tag{0'}$$
$$3x_2 \leqslant 10; \tag{1'}$$
$$x_2 \leqslant 3; \tag{2'}$$
$$3x_1 + 2x_2 \leqslant 17. \tag{3'}$$



Слика 3.2 Гоморијеве одсијецајуће равни

У простору небазних промјенљивих, разломљене Гоморијеве одсијецајуће равни су облика

$$-\frac{7}{22}x_3 - \frac{5}{22}x_4 \leqslant -\frac{1}{2}; \tag{$\bar{0}$}$$
$$-\frac{3}{22}x_3 - \frac{21}{22}x_4 \leqslant -\frac{1}{2}; \tag{$\bar{1}$}$$
$$-\frac{1}{22}x_3 - \frac{7}{22}x_4 \leqslant -\frac{1}{2}; \tag{$\bar{2}$}$$
$$-\frac{1}{2}x_3 - \frac{1}{2}x_4 \leqslant -\frac{1}{2}. \tag{$\bar{3}$}$$

Гоморијев метод се заснива на идеји да се у генеричком методу, приликом произвођења одсијецајуће равни, одабере одговарајући *исходни ред* симплексне табеле са нецјелобројним базним промјенљивим, а онда да се из датог реда производе одговарајућа Гоморијева одсијецајућа раван.

У наставку завршавамо примјер Гоморијевим методом. Прикључујемо Гоморијеву одсијецајућу раван нашој симплексној табели послије њеног записивања помоћу небазних промјенљивих и увођења (базне) изједначавајуће промјенљиве. Након тога израчунавамо нови оптимум за нову почетну симплексну табелу дуалним симплексним методом (пивоти су уоквирени правоугаоним заградама).

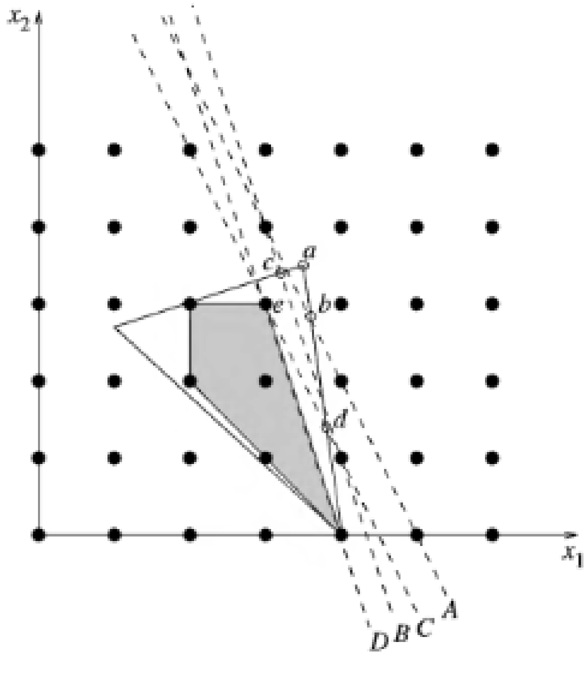
$10\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}$	0	0	$27\frac{1}{2}$		
x_0	x_1	x_2	x_3	x_4	x_5	ДС	
1	0	0	$\frac{7}{22}$	$\frac{5}{22}$	0	$\frac{21}{2}$	←
0	1	0	$\frac{3}{22}$	$-\frac{1}{22}$	0	$\frac{7}{2}$	
0	0	1	$\frac{1}{22}$	$\frac{7}{22}$	0	$\frac{7}{2}$	
0	0	0	$\frac{3}{2}$	$\frac{5}{2}$	1	$\frac{55}{2}$	

$10\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}$	0	0	$27\frac{1}{2}$	$-\frac{1}{2}$	
x_0	x_1	x_2	x_3	x_4	x_5	x_6	ДС
1	0	0	$\frac{7}{22}$	$\frac{5}{22}$	0	0	$\frac{21}{2}$
0	1	0	$\frac{3}{22}$	$-\frac{1}{22}$	0	0	$\frac{7}{2}$
0	0	1	$\frac{1}{22}$	$\frac{7}{22}$	0	0	$\frac{7}{2}$
0	0	0	$\frac{3}{2}$	$\frac{5}{2}$	1	0	$\frac{55}{2}$
0	0	0	$-\frac{7}{22}$	$[-\frac{5}{22}]$	0	1	$-\frac{1}{2}$

10	$3\frac{3}{5}$	$2\frac{4}{5}$	0	$2\frac{1}{5}$	22	0		
x_0	x_1	x_2	x_3	x_4	x_5	x_6	ДС	
1	0	0	0	0	0	1	10	
0	1	0	$\frac{1}{5}$	0	0	$-\frac{1}{5}$	$\frac{18}{5}$	←
0	0	1	$-\frac{2}{5}$	0	0	$\frac{7}{5}$	$\frac{14}{5}$	
0	0	0	-2	0	1	11	22	
0	0	0	$-\frac{7}{5}$	1	0	$-\frac{22}{5}$	$\frac{11}{5}$	

Понављајући претходни поступак долазимо до посљедње табеле.

9	3	3	4	1	19	1	1	0	0	
x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	ДС
1	0	0	0	0	0	0	0	1	0	9
0	1	0	0	0	0	0	0	-1	1	3
0	0	1	0	0	0	0	0	3	-2	3
0	0	0	0	0	1	0	0	19	-10	19
0	0	0	0	0	0	1	0	-1	0	1
0	0	0	0	0	0	0	1	0	-1	1
0	0	0	0	1	0	0	0	-10	7	1
0	0	0	1	0	0	0	0	4	-5	4



Слика 3.3 Рјешавање проблема Гоморијевим методом

У простору почетних промјенљивих, низ Гоморијевих равни одсијецања је

$$2x_1 + x_2 \leqslant 10; \tag{A}$$
$$3x_1 + x_2 \leqslant 13; \tag{B}$$
$$2x_1 + x_2 \leqslant 9; \tag{C}$$
$$3x_1 + x_2 \leqslant 12, \tag{D}$$

а низ оптималних рјешења линеарних програма (у почетним промјенљивим) је

$$x_1 = 3\frac{1}{2}, \qquad x_2 = 3\frac{1}{2}; \tag{a}$$
$$x_1 = 3\frac{3}{5}, \qquad x_2 = 2\frac{4}{5}; \tag{b}$$
$$x_1 = 3\frac{1}{5}, \qquad x_2 = 3\frac{2}{5}; \tag{c}$$
$$x_1 = 3\frac{4}{5}, \qquad x_2 = 1\frac{2}{5}; \tag{d}$$
$$x_1 = 3, \qquad x_2 = 3. \tag{e}$$

У наставку описивања конвергенције Гоморијевог метода у коначном броју корака, треба разумјети нешто о неомеђеним цјелобројним програмима. Пошто се разматрају само рационални програми, имајући у виду Мејеров резултат наведен у примједби за Претпоставку 7, контрапозицијом, добијамо да ако цјелобројни програм има оптимално рјешење, онда ЛП-релаксација не може бити неомеђено. Дакле, увиђамо да има смисла додати још једно ограничење формулацији (проблема) и произвести дуалну допустиву ЛП-базу.

Управо под овим претпоставкама, Гомори је показао да његов метод у сваком кораку све боље и боље апроксимира оптимално рјешење и да га проналази у коначном броју корака.

Теорема 3.2 (Конвергенција Гоморијевог метода [46]). *Под претпоставком да цјелобројни програм има оптимално рјешење, Гоморијев метод одсијецајућих равни конвергира уз услове:*

(а) *исходни ред одговара нецјелобројној базној промјенљивој најмањег индекса;*

(б) *изједначавајућој промјенљивој, за сваку прикључену једнакост, придружен је наредни доступан индекс;*

(в) *ϵ-пертурбовани дуални симплекс метод се користи за реоптимизацију послије сваке прикључене једнакости.*

Доказ. По прикључењу Гоморијеве одсијецајуће равни, у облику $(\overline{G}_{\beta_i})$, оптималној симплексној табели, вриједност нове базне промјенљиве за дати ред је негативна. Коефицијент циља који придружујемо изједначавајућој промјенљивој $x_{\beta_{k+1}}$ је $\epsilon^{\beta_{k+1}}$, према ϵ-пертурбованом дуалном симплекс-методу [4]. Тиме се вриједност циља за дати пертурбован проблем, прије нашег пивотирања, мијења за $(\lfloor x^*_{\beta_i} \rfloor - x^*_{\beta_i})\epsilon^{k+1}$, односно опада. Потом се врши реоптимизација дуалним симплекс методом, а вриједност циља за пертурбовани проблем наставља да опада у сваком новом кораку дуалног симплексног метода.

Размотримо први корак дуалног симплекс метода, управо послије прикључења $(\overline{G}_{\beta_i})$ симплексној табели. Нека је небазна промјенљива, на примјер x_{η_j}, замијењена са промјенљивом $x_{\beta_{k+1}}$. Вриједност x_{β_i} се мијења од $x^*_{\beta_i}$ до

$$\tilde{x}^*_{\beta_i} = x^*_{\beta_i} - \overline{a}_{\beta_i\eta_j} \frac{x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor}{\overline{a}_{\beta_i\eta_j} - \lfloor \overline{a}_{\beta_i\eta_j} \rfloor}.$$

Пошто је

$$\frac{x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor}{\overline{a}_{\beta_i \eta_j} - \lfloor \overline{a}_{\beta_i \eta_j} \rfloor}$$

позитивно, смањење у x_{β_i} значи да је $\overline{a}_{\beta_i \eta_j} > 0$.[1] Стога, уз

$$\frac{\overline{a}_{\beta_i \eta_j}}{\overline{a}_{\beta_i \eta_j} - \lfloor \overline{a}_{\beta_i \eta_j} \rfloor} \geqslant 1,$$

закључујемо да је

$$\tilde{x}^*_{\beta_i} \leqslant \lfloor x^*_{\beta_i} \rfloor \tag{3.3}$$

Претпоставимо да је z^* оптимална вриједност циља ЛП-релаксације за наш цјелобројан програм. Нека нам је $x^{\text{ЛП}}$ оптимално рјешење ЛП-релаксације засновано на ϵ-пертурбованом дуалном симплекс методу.

Размотримо сад квадар цјелобројних тачака

$$K = \{x \in \mathbb{Z}^{n+1} \ : \ z^* \leqslant x_0 \leqslant x_0^{\text{ЛП}} \text{ и } 0 \leqslant x_j \leqslant x_j^{\text{ЛП}}, \text{ за } j = 1, 2, \ldots, n\},$$

у ком, ради досљедности, одбројавање промјенљивих почиње од нуле. Коначан скуп K се може уредити *лексикографски* — што се рачунски може представити као $x^1 \prec x^2$ ако $\sum_{j=0}^{n} x^1_j \epsilon^j < \sum_{j=0}^{n} x^2_j \epsilon^j$ за неко мало $\epsilon > 0$. Послије сваке (узастопне) реоптимизације (ϵ-пертурбованим дуалним симплексним методом) у којој је одабран исходни ред са почетним промјенљивим (нпр. $x_0,\ x_1, \ldots, x_n$) према услову (а), добија се рјешење које је лексикографски мање од рјешења претходне оптимизације. Због (3.3) дато рјешење реоптимизације је мање и од првог елемента скупа K који је лексикографски мањи од претходног рјешења. Зато на крају, послије коначног броја корака, све почетне промјенљиве попримају цјелобројне вриједности. ☐

Алгоритми засновани искључиво на методима одсијецања нису нашли неку примјену у пракси због разних нумеричких проблема који су се јављали. Генерално, постоји тенденција да се генерисане одсијецајуће равни паралелизују са функцијом циља (феномен испраћен дуалном дегнерацијом), да се повећава детерминанта базе, а и да се грешке заокруживања толико нагомилају, да утичну на тачност добијеног резултата. Због ових разлога, методи одсијецајућих равни се обично користе у споју са методима гранања и омеђивања.

С друге стране, тежећи да се продуби разумијевање лоших својстава метода одсијецајућих равни, у недавној студији [61], Занете *et al.* је показао да је лексикографски симплексни метод вјероватно најбољи међу симплексним методима, као платформа за реализацију Гоморијевог метода и производњу самих Гоморијевих одсијецајућих равни.

У раду који се надовезао на дато истраживање [10], разматрана је и енумеративна (бројачка) природа Гоморијевог метода и представљен је практичан значај енумеративне интерпретације — Гоморијев метод је практично калуп у свом енумеративном радном оквиру и природан начин за његово побољшање је одбацивање ригидности енумеративне љуске. Један од начина је третирање циљне функције имплицитно, при чему се лексикографско уређење редефинише у ходу како би се опонашала класична стратегија гранања (основ енумеративних метода).

---

[1] У исходном реду $\overline{a}_{\beta_i \eta_j} > 0$ с обзиром да то у ствари слиједи из правила избора пивота. У противном, да није могућа таква ситуација, програм не би био допустив.

3.4 Агрегирање Гоморијевих одсијецајућих равни

Имајући у виду да је оптимално рјешење програма, такође рјешење Диофантових једначина — једначи̂на симплексне табеле у било којој етапи метода, за свако $i \in \{0, 1, \ldots, m\}$ вриједност x_{β_i} из (3.2) може се сматрати цјелобројном.

Самим тим, свака цјелобројна линеарна комбинација $\sum_{i=1}^{m} \tau_i x_{\beta_i}$, за $\tau \in \mathbb{Z}^m$, је једнака неком цјелом броју, а вриједност израза

$$\sum_{i=0}^{m} \tau_i x^*_{\beta_i} - \sum_{i=0}^{m} \tau_i \sum_{j=1}^{k-m} \overline{a}_{\beta_i \eta_j} x_{\eta_j}$$

је цјелобројна. Лако се уочава сличност са произвођењем Хватал-Гоморијевих одсијецајућих равни.

Издвајајући разломљен дио, важи да је вриједност сљедећег израза

$$\left(\sum_{i=0}^{m} \tau_i x^*_{\beta_i} - \left\lfloor \sum_{i=0}^{m} \tau_i x^*_{\beta_i} \right\rfloor \right) - \left(\sum_{j=1}^{k-m} \sum_{i=0}^{m} \tau_i \overline{a}_{\beta_i \eta_j} x_{\eta_j} - \sum_{j=1}^{k-m} \left\lfloor \sum_{i=0}^{m} \tau_i \overline{a}_{\beta_i \eta_j} \right\rfloor x_{\eta_j} \right) \tag{3.4}$$

такође цио број. Имајући у виду природу разломљеног дијела броја, то јест да важи

$$0 \leqslant \sum_{i=0}^{m} \tau_i x^*_{\beta_i} - \left\lfloor \sum_{i=0}^{m} \tau_i x^*_{\beta_i} \right\rfloor < 1,$$

односно

$$0 \leqslant \sum_{j=1}^{k-m} \sum_{i=0}^{m} \tau_i \overline{a}_{\beta_i \eta_j} x_{\eta_j} - \sum_{j=1}^{k-m} \left\lfloor \sum_{i=0}^{m} \tau_i \overline{a}_{\beta_i \eta_j} \right\rfloor x_{\eta_j}$$

при чему је $x_{\eta_j} \geqslant 0$, израз (3.4) не може бити позитиван. На основу овога добијају се такозване *агрегиране Гоморијеве одсијецајуће равни*

$$\sum_{j=1}^{k-m} \left\lfloor \sum_{i=1}^{m} \tau_i \overline{a}_{\beta_i \eta_j} \right\rfloor x_{\eta_j} - \sum_{j=1}^{k-m} \sum_{i=1}^{m} \tau_i \overline{a}_{\beta_i \eta_j} x_{\eta_j} \leqslant \left\lfloor \sum_{i=1}^{m} \tau_i x^*_{\beta_i} \right\rfloor - \sum_{i=1}^{m} \tau_i x^*_{\beta_i}. \tag{3.5}$$

Генерално, различите Гоморијеве одсијецајуће равни одговарају различитим векторима $\tau \in \mathbb{Z}^m$. Ако се дефинише операција $*$ над скупом Гоморијевих равни, тако да је за

$$g_{\tau^1} \; : \; \sum_{j=1}^{k-m} \left\lfloor \sum_{i=1}^{m} \tau_i^1 \overline{a}_{\beta_i \eta_j} \right\rfloor x_{\eta_j} - \sum_{j=1}^{k-m} \sum_{i=1}^{m} \tau_i^1 \overline{a}_{\beta_i \eta_j} x_{\eta_j} \leqslant \left\lfloor \sum_{i=1}^{m} \tau_i^1 x^*_{\beta_i} \right\rfloor - \sum_{i=1}^{m} \tau_i^1 x^*_{\beta_i} \tag{3.6}$$

$$g_{\tau^2} \; : \; \sum_{j=1}^{k-m} \left\lfloor \sum_{i=1}^{m} \tau_i^2 \overline{a}_{\beta_i \eta_j} \right\rfloor x_{\eta_j} - \sum_{j=1}^{k-m} \sum_{i=1}^{m} \tau_i^2 \overline{a}_{\beta_i \eta_j} x_{\eta_j} \leqslant \left\lfloor \sum_{i=1}^{m} \tau_i^2 x^*_{\beta_i} \right\rfloor - \sum_{i=1}^{m} \tau_i^2 x^*_{\beta_i} \tag{3.7}$$

израз $g_{\tau^1} * g_{\tau^2}$ једнак

$$\begin{aligned} g_{\tau^1 + \tau^2} \; : \; \sum_{j=1}^{k-m} \left\lfloor \sum_{i=1}^{m} (\tau_i^1 + \tau_i^2) \overline{a}_{\beta_i \eta_j} \right\rfloor x_{\eta_j} &- \sum_{j=1}^{k-m} \sum_{i=1}^{m} (\tau_i^1 + \tau_i^2) \overline{a}_{\beta_i \eta_j} x_{\eta_j} \\ &\leqslant \left\lfloor \sum_{i=1}^{m} (\tau_i^1 + \tau_i^2) x^*_{\beta_i} \right\rfloor - \sum_{i=1}^{m} (\tau_i^1 + \tau_i^2) x^*_{\beta_i}. \end{aligned} \tag{3.8}$$

онда скуп свих Гоморијевих одсијецајућих равни чини Абелову групу $(\mathcal{G}, *)$, пошто је, просто, $(\mathbb{Z}^m, +)$ Абелова група, а $(\mathcal{G}, *)$ је хомоморфна са $(\mathbb{Z}^m, +)$. Група Гоморијевих одсијецајућих равни је хомоморфна групи $(\mathbb{Z}^m, +)$ пошто вектору $\tau \in \mathbb{Z}^m$ одговара само једна одсијецајућа раван, но с друге стране, једној одсијецајућој равни одговара бесконачно много вектора из $\mathbb{Z}^m$.

Примједба 3.3. Нулти ред симплексне табеле се може изоставити, пошто је хиперраван, која му одговара, сувишна с обзиром на допустиву област одређену осталим хиперравнима симплексне табеле.

Примједба 3.4. Скуп Хватал-Гоморијевих равни за $\tau \in \mathbb{Z}_+^m$ чини комутативан моноид под операцијом $*$.

Примједба 3.5. Хватал-Гоморијеве равни чине Абелову групу, за $\tau \in \mathbb{R}_+^m$, под операцијом $\circ$ дефинисаном тако да је $h_{\tau^1} \circ h_{\tau^2} = h_{\tau^1 \tau^2}$, за неке Хватал-Гоморијеве одсијецајуће равни h_{τ^1}, h_{τ^2}.

3.4.1 Ред групе агрегираних равни

Треба одредити укупан број различитих одсијецајућих равни које се могу добити помоћу вектора $\tau \in \mathbb{Z}^m$, или бар неку горњу границу тог броја.

Теорема 3.6 ([43]). *За сваку матрицу $M_{p\times q}$, постоји придружена дијагонална матрица $\Delta_{p\times q}$ у Смитовој нормалној форми, добијена трансформацијом*

$$U_{p\times p} A_{p\times q} V_{q\times q} = \Delta_{p\times q},$$

гдје су U и V регуларне унимодуларне матрице.

За регуларну матрицу T_β, формата $m \times m$, постоје двије регуларне унимодуларне матрице $U_{m\times m}$ и $V_{m\times m}$, као и матрица $\Delta_{m\times m}$, такве да је

$$UT_\beta V = \Delta$$

Сматрајући T_β регуларном, Δ је такође регуларна (U и V су регуларне, производ три регуларне матрице је регуларна матрица). Стога, редукцијом на Смитову нормалну форму добијамо матрицу Δ такву да је

$$\Delta = \begin{bmatrix} \delta_1 & 0 & \dots & 0 \\ 0 & \delta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \delta_m \end{bmatrix},$$

при чему δ_i дијели δ_{i+1} за све $i = 1, 2, \dots, m-1$.

Множећи с лијева са U^{-1} и с десна са V^{-1}, добија се да је $T_\beta = U^{-1}\Delta V^{-1}$. Инвертовањем имамо да је $T_\beta^{-1} = V\Delta^{-1}U$, при чему је

$$\Delta^{-1} = \begin{bmatrix} \frac{1}{\delta_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\delta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\delta_m} \end{bmatrix}.$$

Агрегиране Гоморијеве одсијецајуће равни могу се према (3.5) матрично записати у облику

$$(\lfloor \tau^T T_\beta^{-1} T_\eta \rfloor - \tau^T T_\beta^{-1} T_\eta)^T x_\eta \leqslant \lfloor \tau^T T_\beta^{-1} x_\beta^* \rfloor - \tau^T T_\beta^{-1} x_\beta^*,$$

при чему је

$$\lfloor M \rfloor = \begin{bmatrix} \lfloor a_{11} \rfloor & \lfloor a_{12} \rfloor & \dots & \lfloor a_{1q} \rfloor \\ \lfloor a_{21} \rfloor & \lfloor a_{22} \rfloor & \dots & \lfloor a_{2q} \rfloor \\ \vdots & \vdots & \ddots & \vdots \\ \lfloor a_{p1} \rfloor & \lfloor a_{p2} \rfloor & \dots & \lfloor a_{pq} \rfloor \end{bmatrix}$$

за неку матрицу M, формата $p \times q$.

Даље, у датој неједнакости, замјеном инвертоване матрице T_β^{-1}, добијамо неједнакост

$$(\lfloor \tau^T V \Delta^{-1} U T_\eta \rfloor - \tau^T V \Delta^{-1} U T_\eta)^T x_\eta \leqslant \lfloor \tau^T V \Delta^{-1} U x_\beta^* \rfloor - \tau^T V \Delta^{-1} U x_\beta^*$$

Пошто је V унимодуларна и формирана од цијелих бројева, V^{-1} је такође формирана од цијелих бројева, па узевши да је

$$v = \tau^T V$$

придружена вектор-врста за вектор τ, поново замјењујући, добија се неједнакост

$$\begin{aligned}(\lfloor v\Delta^{-1} U T_\eta - v\Delta^{-1} U T_\eta \rfloor) x_\eta \\ \leqslant \lfloor v\Delta^{-1} U x_\beta^* \rfloor - v\Delta^{-1} U x_\beta^*.\end{aligned}$$

Настављајући на исти начин, то јест узимајући да је

$$u = U x_\beta^*$$

$$N = U T_\eta$$

претходна неједнакост постаје

$$\begin{aligned}(\lfloor v\Delta^{-1} N \rfloor - v\Delta^{-1} N) x_\eta \\ \leqslant \lfloor v\Delta^{-1} u \rfloor - v\Delta^{-1} u.\end{aligned}$$

Опет уводећи смјену (нотацијску)

$$\mu = v\Delta^{-1} = [\frac{v_1}{\delta_1}\ \frac{v_2}{\delta_2}\ \dots\ \frac{v_m}{\delta_m}].$$

добијамо сљедећи облик неједнакости

$$(\lfloor \mu N \rfloor - \mu N) x_\eta \leqslant \lfloor \mu u \rfloor - \mu u.$$

Не заборављајући да су вектор u и матрица N сачињени од цијелих бројева, коришћењем својства функције цио дио $ab - \lfloor ab \rfloor = (a - \lfloor a \rfloor)b - \lfloor (a - \lfloor a \rfloor)b \rfloor$ (за $a, b \in \mathbb{R}$), неједнакост се може записати као

$$(\lfloor (\mu - \lfloor \mu \rfloor) N \rfloor - (\mu - \lfloor \mu \rfloor) N) x_\eta \leqslant \lfloor (\mu - \lfloor \mu \rfloor) u \rfloor - (\mu - \lfloor \mu \rfloor) u.$$

Максимални број одсијецајућих равни које се могу добити одговара броју различитих одсијецајућих равни које можемо добити за различите векторе

$$\mu - \lfloor \mu \rfloor = v\Delta^{-1} - \lfloor v\Delta^{-1} \rfloor.$$

Пошто је у питању група остатака по модулу, постоји:

- $|\delta_1|$ вриједности за $\mu_1 - \lfloor \mu_1 \rfloor = v_1/\delta_1 - \lfloor v_1/\delta_1 \rfloor$,
- $|\delta_2|$ вриједности за $\mu_2 - \lfloor \mu_2 \rfloor = v_2/\delta_2 - \lfloor v_2/\delta_2 \rfloor$,
- $\ldots$
- $|\delta_m|$ вриједности за $\mu_m - \lfloor \mu_m \rfloor = v_m/\delta_m - \lfloor v_m/\delta_m \rfloor$.

Стога, постоји

$$|\delta_1 \delta_2 \cdot \ldots \cdot \delta_m|$$

могућих вектора $v\Delta^{-1} - \lfloor v\Delta^{-1} \rfloor$, а изузимајући међу њима вектор $\mathbf{0}$, добијамо да их различитих, практично, највише може бити

$$|\delta_1 \delta_2 \ldots \delta_m| - 1.$$

Пошто је матрица Δ регуларна и дијагонална, њена детерминанта је $\det \Delta = \delta_1 \delta_2 \cdot \ldots \cdot \delta_m$, то јест једнака је детерминанти матрице T_β (пошто су U и V регуларне и унимодуларне) по апсолутној вриједности, јер вриједност $\det U$, односно $\det V$, може бити једнака -1. Одатле је број Гоморијевих одсијецајућих равни које се могу добити за дату симплексну таблицу

$$|\det T_\beta| - 1.$$

Примјер 3.2. Разматрајући у Примјеру 3.1 симплексну таблицу добијену у првој етапи Гоморијевог метода, имамо да је

$$T_\beta^{-1} = \begin{bmatrix} \frac{3}{22} & \frac{-1}{22} & 0 \\ \frac{1}{22} & \frac{7}{22} & 0 \\ \frac{3}{2} & \frac{5}{2} & 1 \end{bmatrix}.$$

Наравно,

$$T_\beta = \begin{bmatrix} 7 & 1 & 0 \\ -1 & 3 & 0 \\ -8 & -9 & 1 \end{bmatrix},$$

при чему је онда $\det T_\beta = 22$, то јест, могуће је произвести 22 различите Гоморијеве одсијецајуће равни у првој етапи метода.

Налазећи Смитову нормалну форму матрице T_β, добијамо

$$U = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 6 & 0 \\ 1 & 7 & 0 \end{bmatrix} \quad V = \begin{bmatrix} 0 & 1 & -19 \\ 0 & 0 & 1 \\ 1 & 8 & -143 \end{bmatrix} \quad \Delta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 22 \end{bmatrix}. \qquad \blacksquare$$

Иначе, матрице U и V нису јединствене, док је Δ, због цјелобројности матрице T_β, у редукованом облику.

3.4.2 Цикличка група агрегираних равни

Ако је матрица Δ таква да је $\delta_1 = \delta_2 = \ldots = \delta_{m-1} = 1$, онда се $|\det \Delta| - 1$ вектора $\mu - \lfloor \mu \rfloor$ могу записати као

$$\mu - \lfloor \mu \rfloor = [0\ 0\ \ldots\ 0\ (\frac{v_m}{\delta_m} - \lfloor \frac{v_m}{\delta_m} \rfloor)].$$

У том случају, изглед Гоморијевих одсијецајућих равни је облика

$$(\lfloor \alpha_m N_{m1} \rfloor - \alpha_m N_{m1}) x_{\eta_1} + \ldots + (\lfloor \alpha_m N_{mn} \rfloor - \alpha_m N_{mn}) x_{\eta_n} \\ \leqslant \lfloor \alpha_m u_m \rfloor - \alpha_m u_m,$$

при чему смо са α_m означили $\frac{v_m}{\delta_m} - \lfloor \frac{v_m}{\delta_m} \rfloor$.

Као што се може примијетити, све одсијецајуће равни су добијене примјеном операције $*$ (дефинисаној у (3.5)-(3.7)), почевши са

$$(\lfloor \frac{1}{\delta_m} N_{m1} \rfloor - \frac{1}{\delta_m} N_{m1}) x_{\eta_1} + \ldots + (\lfloor \frac{1}{\delta_m} N_{mn} \rfloor - \frac{1}{\delta_m} N_{mn}) x_{\eta_n} \\ \leqslant \lfloor \frac{1}{\delta_m} u_m \rfloor - \frac{1}{\delta_m} u_m,$$

Стога, група Гоморијевих одсијецајућих равни формира цикличку групу у којој је претходна неједнакост генератор.

Примјер 3.3. Јасно је да је група Гоморијевих одсијецајућих равни, разматрана у претходном примјеру, цикличка, односно да су ненула вектори $\mu - \lfloor \mu \rfloor$ облика $[0\ 0\ \frac{t}{22}]$, за $t = 1, 2, \ldots, 22$.

У наставку, по израчунавању

$$N = \begin{bmatrix} 0 & 0 \\ 1 & 6 \\ 1 & 7 \end{bmatrix} \qquad u = \begin{bmatrix} -32 \\ 70 \\ 77 \end{bmatrix},$$

добијамо генеричку одсијецајућу раван

$$-\frac{1}{22} x_3 - \frac{7}{22} x_4 \leqslant -\frac{1}{2},$$

а то је раван $(\overline{2})$ из Примјера 3.1. ■

3.5 Потпуно цјелобројан Гоморијев метод

> *„Онда сам се окренуо нечему чему сам одувијек тежио, потпуно цјелобројном програмском алгоритму — алгоритму који би подражавао симплексни метод, али у ком би коефицијенти увијек остајали цјелобројни.”*
>
> Ралф Е. Гомори

У рачунарској реализацији Гоморијевог метода јављају се нумерички проблеми при дијељењу (цијелих) бројева. На примјер, дијелећи 1 са 3, добија се количник $0,3333\ldots3$, гдје број тројки зависи од величине регистра за похрањивање података. Помножи ли се добијени количник са 3 (дјелилац од малоприје), неће се више добити првобитни дијељеник 1, него нешто налик на $0,999\ldots9$. Овакве грешке заокруживања се нагомилавају у сваком кораку метода, па цјелобројност резултата и међурезултата може постати тешко достижна. Дате нумеричке потешкоће су подстакле конструкцију такозваног потпуно цјелобројног метода за рјешавање цјелобројних програма.

Нека је симплексна табела T, која одговара некој бази цјелобројног програма, дуално допустива и цјелобројна, то јест (користећи ознаке из Поглавља 3.2)

$$\begin{aligned}&\overline{a}_{ij} \in \mathbb{Z} \text{ за } i = 0, 1, \ldots, m+n;\ j = 0, 1, \ldots, m;\\&\overline{a}_{0j} \geqslant 0 \text{ за } j = 1, 2, \ldots, n.\end{aligned}$$

Уколико је $x^*_{\beta_i} \geqslant 0$ $(i = 1, 2, \ldots, m+n)$, онда је T оптимална симплексна табела за цјелобројни програм.

У противном, нека је $\overline{a}_{rs}$ пивот, изабран према правилима дуалног симплексног метода. Ако је $\overline{a}_{rs} = -1$, онда се послије једног корака дуалног симплекс метода поново добија дуално допустива цјелобројна таблица.

Идеја потпуно цјелобројног метода се састоји у формирању одсијецајућих равни које омогућују избор пивота једнаког -1.

Лема 3.7. *Ако је*

$$x = a_0 + \sum_{j=1}^{n} a_j(-x_j) \tag{3.9}$$

при чему $x \in \mathbb{Z},\ x \geqslant 0,\ x_j \in \mathbb{Z},\ x_j \geqslant 0,\ a_0 \in \mathbb{Z},\ a_j \in \mathbb{Z}\ (j = 1, \ldots, n)$, *онда за било које* $\lambda \in \mathbb{Z}$ *такво да је* $\lambda \geqslant 1$, *важи да је сљедећа неједнакост ваљана за цјелобројни омотач у полиедру ЛП-релаксације*

$$\lfloor \frac{a_0}{\lambda} \rfloor + \sum_{j=1}^{n} \lfloor \frac{a_j}{\lambda} \rfloor (-x_j) \geqslant 0. \tag{3.10}$$

Доказ. Из (3.9) слиједи да је

$$\sum_{j=1}^{n} a_j x_j \leqslant a_0.$$

Одатле је

$$\sum_{j=1}^{n} \frac{a_j}{\lambda} x_j \leqslant \frac{a_0}{\lambda}.$$

Из посљедње неједнакости добијамо

$$\sum_{j=1}^{n} \lfloor \frac{a_j}{\lambda} \rfloor x_j \leqslant \frac{a_0}{\lambda}.$$

Лијева страна посљедње неједнакости је цјелобројна, па самим тим није већа ни од $\lfloor \frac{a_0}{\lambda} \rfloor$, што је и требало доказати. □

Нека је r исходни ред дуално допустиве табеле T, којој одговарају индекси небазних промјенљивих $T_\eta = \{\eta_1, \eta_2, \ldots, \eta_n\}$. Тада је

$$x_{\beta_r} = x^*_{\beta_r} + \sum_{j=1}^{n} \overline{a}_{\beta_r \eta_j}(-x_{\eta_j}).$$

Узевши да је
$$\lambda = \max\{|\overline{a}_{\beta_r\eta_j}| \ : \ \overline{a}_{\beta_r\eta_j} < 0 \text{ и } j = 1, \ldots, n\}, \tag{3.11}$$
нека је
$$x_{m+n+1} = \lfloor \frac{x^*_{\beta_r}}{\lambda} \rfloor + \sum_{j=1}^{n} \lfloor \frac{\overline{a}_{\beta_r\eta_j}}{\lambda} \rfloor (-x_{\eta_j}). \tag{3.12}$$
Према Леми 3.7 неједнакост $x_{m+n+1} \geqslant 0$ је ваљана одсијецајућа раван, па има смисла прикључити је табели.

Иначе, иако је раван добијена из исходног реда r, имајући у виду речено у Поглављу 3.4, произвољна линеарна комбинација може послужити за произвођење нове одсијецајуће равни.

Пошто је $x^*_{\beta_r} < 0$, онда је и $\lfloor x^*_{\beta_r} \rfloor < 0$, па сљедствено томе, у новом реду се може бирати пивот, рецимо у колони s, према правилу дуалног симплекс-метода. Очито је да је $\overline{a}_{(m+n+1)s} = -1$. Послије једне етапе симплексног метода с претходно добијеним пивотом, нови ред се може избацити, а поступак поновити, без обзира на то што промјенљива која одговара том реду није базна.

Овако реконструисан Гоморијев метод назива се *потпуно цјелобројним.*

Иначе, избор вриједности λ може бити побољшан (што не утиче на израз за одсијецајућу раван (3.12). Да би први ред симплексне табеле био ненегативан (за сљедбену итерацију), мора, при избору исходног реда r, као пивот-реда, и колоне s, као пивот-колоне, важити да је
$$\min_j \left(\frac{\overline{a}_{0j}}{-\lfloor \frac{\overline{a}_{rj}}{\lambda} \rfloor} \right) = \frac{\overline{a}_{0s}}{-\lfloor \frac{\overline{a}_{rs}}{\lambda} \rfloor}, \tag{3.13}$$
гдје је j такво да је $\lfloor \overline{a}_{rj}/\lambda \rfloor < 0$.

С друге стране, са циљем да симплексна табела остане и даље цјелобројна, λ мора бити изабрано тако да је
$$\left\lfloor \frac{\overline{a}_{rs}}{\lambda} \right\rfloor = -1 \tag{3.14}$$
одакле се, замјењујући (3.14) у (3.13), добија
$$\overline{a}_{0s} \leqslant \frac{\overline{a}_{0j}}{-\lfloor \frac{\overline{a}_{rj}}{\lambda} \rfloor} \leqslant \overline{a}_{0j}, \tag{3.15}$$
гдје је j такво да је $\lfloor \overline{a}_{rj}/\lambda \rfloor < 0$.

Промјена вриједности циљне функције у сљедбеној итерацији, без обзира на који начин одредили λ, једнака је
$$\overline{a}_{0s} \left\lfloor \frac{x^*_{\beta_r}}{\lambda} \right\rfloor. \tag{3.16}$$

Пошто у проблему максимизације, рјешавајући га дуалним симплекс-методом, вриједност циљне функције опада у свакој итерацији, онда у свакој итерацији треба максимизовати апсолутну вриједност тог опадања израженог у (3.16). Значи, треба узети најмању могућу вриједност λ сагласну са (3.14) и (3.15), то јест
$$\lambda = \max \left(-\overline{a}_{rs}, \left\lfloor \frac{-\overline{a}_{rj}}{\left\lfloor \frac{\overline{a}_{0j}}{\overline{a}_{0s}} \right\rfloor} \right\rfloor \right), \tag{3.17}$$

при чему су s и j такви да је $\overline{a}_{rs} < 0$ и $\overline{a}_{rj} < 0$, за $j = 1, 2, \ldots, n$.

На основу досад реченог, слиједи да се прво бира пивот-колона s, тако да важи

$$\overline{a}_{0s} = \min_j \overline{a}_{0j} \quad \text{и} \quad \overline{a}_{rs} < 0,\ \overline{a}_{rj} < 0,\ j = 1, 2, \ldots, n, \tag{3.18}$$

па се потом, по избору вриједности λ, одређује одсијецајућа раван којој су сви коефицијенти цјелобројни.

Потпуно цјелобројан Гоморијев метод

0. Почети са дуално допустивом цјелобројном симплексном табелом T.
1. Ако је $x^*_{\beta_i} \geqslant 0$ за све $i = 0, 1, \ldots, m + n$, онда је оптимални вектор пронађен и прекида се спровођење метода. У противном, бира се најмањи $r \in \{1, 2, \ldots, m\}$ за који је $x^*_{\beta_r} < 0$.
2. Бира се пивот-колона s према (3.18).
3. Одређује се λ према (3.17).
4. Табели T се прикључује ред са коефицијентима одсјечења (3.12). Уколико је $\lambda = 1$, нови ред се подудара са исходним.
5. Гаусовим поступком се врши пивотирање, то јест промјенљива са индексом s се производи базном. Потом се враћамо на корак 1.

Наравно, може се и користити усправна (ступчана) форма дуалног симплексног метода, штавише, уколико би се λ одређивало према (3.11), редослијед, а и сами кораци, би се мало измијенили. Прво би се одређивала вриједност λ, па би се, по прикључењу новог реда симплексној таблици, одредио пивот према правилу лексикографског дуалног симплексног метода

$$\frac{\overline{a}_{0s}}{|\overline{a}_{(m+n+1)s}|} = \min \left\{ \frac{\overline{a}_{0j}}{|\overline{a}_{(m+n+1)j}|} \ :\ \overline{a}_{(m+n+1)j} < 0 \quad (j = 1, 2, \ldots, n) \right\}.$$

Теорема 3.8. *Ако цјелобројни програм има оптимално рјешење, тада претходно описани метод потпуно цјелобројног Гоморијевог метода конвергира у коначном броју корака.*

Доказ. Претпоставимо супротно, то јест да је рачунски поступак бесконачан. Ако се користи лексикографски дуални симплекс-метод, то јест коначна модификација симплексног метода, онда поступак може бити бесконачан само ако се генерише бесконачно много одсијецајућих равни.

Даље, вектор десне колоне симплексне табеле, у ознаци x^*_i ($i \in \{0, 1, \ldots, m\}$), се монотоно лексикографски смањује. Штавише, ако вриједност циљне функције монотоно опада, због природе измјена, она опада за цјелобројне вриједности. Због постојања оптималног рјешења, слиједи да ће се вриједност циља у једном тренутку престати мијењати.

Размотримо сад промјенљиву x^*_1. Ако се њена вриједност смањује, она се смањује за неки цио број. Када вриједност постане негативна, први ред постаје исходни.

Према опису метода, за неко s важи да је $\overline{a}_{1s} < 0$ које одређује пивот-колону. Правило за одређивање вриједности λ нам гарантује да је

$$\left\lfloor \frac{\overline{a}_{1s}}{\lambda} \right\rfloor = -1 \quad \text{и} \quad \left\lfloor \frac{x_1^*}{\lambda} \right\rfloor < 0.$$

Због тога се вриједност x_1^* строго увећава, па слиједи да се циљна вриједност мора промијенити (имајући у виду лексикографско уређење), што противрјечи тврдњи о њеној неизмјењивости.

Одатле, вриједност x_1^* је постојана у свим наредним итерацијама метода, а слично расуђивање можемо провести и за другу, трећу, односно све остале промјенљиве и њима одговарајуће редове.

Значи, послије коначног броја корака све вриједности x_i^* $(i = 1, 2, \ldots, m)$ ће постати ненегативне и цјелобројне. □

Примјер 3.4. Ријешити програм

$$\begin{aligned} \max \ & -3x_1 - 8x_2 \\ & -4x_1 - 5x_2 \leqslant -2 \\ & -3x_1 - 7x_2 \leqslant -2 \\ & x_1, x_2 \in \mathbb{N}. \end{aligned}$$

Конструише се почетна симплексна табела

0	0	0	-2	-2		
x_0	x_1	x_2	x_3	x_4	дс	
1	3	8	0	0	0	
0	-4	-5	1	0	-2	←
0	-3	-7	0	1	-2	

имајући у виду изједначавајуће промјенљиве x_3 и x_4. Пошто посљедња колона није ненегативна, табела не предочава рјешење програма. Према алгоритму, бира се исходни ред (означен са ←), ради произвођења потпуно цјелобројне одсијецајуће равни. Колоне које одговарају промјенљивим x_1 и x_2 су кандидати за тражење пивота, другим ријечима, x_1 или x_2 могу постати базне промјенљиве у сљедећој етапи, пошто су у исходном реду -4 и -5 једине негативне вриједности. У реду који одговара промјенљивој x_0, кореспондирајуће вриједности за x_1 и x_2 су 3 и 8, а пошто је 3 много мање од 8, пивот се бира из колона промјељиве x_1. Потом се бира λ,

$$\begin{aligned} \lambda &= \max\left(-(-4), \left\lfloor \frac{-(-5)}{\lfloor \frac{8}{3} \rfloor} \right\rfloor\right) \\ &= 4. \end{aligned}$$

Сада је могуће произвести одсцијецајућу раван, то јест

$$\left\lfloor \frac{-2}{\lambda} \right\rfloor - \left\lfloor \frac{-4}{\lambda} \right\rfloor x_1 - \left\lfloor \frac{-5}{\lambda} \right\rfloor \leqslant 0,$$

односно за $\lambda = 4$,

$$-1 + x_1 + 2x_2 \leqslant 0.$$

Добија се табела

0	0	0	-2	-2	-1	
x_0	x_1	x_2	x_3	x_4	x_5	дс
1	3	8	0	0	0	0
0	-4	-5	1	0	0	-2
0	-3	-7	0	1	0	-2
0	[-1]	-2	0	0	1	-1

гдје је пивот означен помоћу [...], и очекивано једнак је -1. У наставку, по пивотирању добија се оптимална табела

-3	1	0	2	1	0	
x_0	x_1	x_2	x_3	x_4	x_5	дс
1	0	2	0	0	8	-3
0	0	3	1	0	-4	2
0	0	-1	0	1	-3	1
0	1	2	0	0	-1	1

из које се очитава цјелобројно рјешење

$$x_1 = 1, \qquad x_2 = 0. \qquad \blacksquare$$

Примједба 3.9. Може се примјетити да не требају обавезно сви коефицијенти програма бити цјелобројни. Довољно је да су десне стране (главних) ограничења програма цјелобројне уколико се користи дуални симплекс-метод у положеној форми, односно да су коефицијенти функције циља цијели, ако се користи дуални симплекс-метод у усправној форми. Без обзира да ли су остали коефицијенти цјелобројни или не, коефицијенти одсијецајуће равни (3.12) су увијек цјелобројни, а пивот је једнак -1. Доказ конвергенције метода се суштински не разликује од претходно датог у Теореми 3.4.

3.6 Појачавање Гоморијевих одсијецајућих равни

На почетку биће дефинисано надјачавање међу неједнакостима. Ако су $a_\sigma^T x \leqslant b_\sigma$ и $a_\pi^T x \leqslant b_\pi$ двије ваљане неједнакости, онда сматрамо другу неједнакост *јачом* од прве, ако постоји скалар $v > 0$, такав да је $va_\sigma < a_\pi$ и $vb_\sigma \geqslant b_\pi$, при чему је

$(va_\sigma, vb_\sigma) \neq (a_\pi, b_\pi)$. Наравно, у датом случају, прву неједнакост сматрамо *слабијом (надјачаном)* у односу на другу.

У суштини, јача неједнакост посматрана као одсијецајућа раван, одсијеца више допустиве области него слабија.

Примједба 3.10. Наравно, јача одсијецајућа раван (према претходној дефиницији) не мора бити и бољa у датом тренутку за дати цјелобројни програм.

Изворне Гоморијеве одсијецајуће равни могу се појачати на неколико различитих начина. Прва техника је описана у сљедећој теореми.

Теорема 3.11 (Прва техника). *За сваки цио број t, одсијецајућа раван*

$$\sum_{j=1}^{n} (\lfloor t\overline{a}_{\beta_i,\eta_j} \rfloor - t\overline{a}_{\beta_i,\eta_j}) x_{\eta_j} \leqslant \lfloor t x^*_{\beta_i} \rfloor - t x^*_{\beta_i} \tag{3.19}$$

је ваљана за све цјелобројне векторе за које је ваљана и раван $(\overline{G}_{\beta_i})$. Штавише, ако је t позитивно, $x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor < 1/2$ и $1/2 \leqslant t(x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor) < 1$, онда је одсијецајућа раван (3.19) јача или пак, у најгорем случају, једнака изворној разломљеној Гоморијевој одсијецајућој равни $(\overline{G}_{\beta_i})$.

Доказ. Ова техника појачавања одсијецајућих равни је у суштини посебан случај агрегираних Гоморијевих одсијецајућих равни (3.5) за вектор $\tau = [0 \ \ldots \ t \ \ldots \ 0]^T$, гдје је t на мјесту β_i, одакле слиједи ваљаност.

Што се тиче надјачавања, (3.19) је јача или, у најгорем случају, једнака $(\overline{G}_{\beta_i})$ ако и само ако вриједи, с обзиром да су у питању мање-или-једнако неједнакости,

$$\frac{\lfloor \overline{a}_{\beta_i\eta_j} \rfloor - \overline{a}_{\beta_i\eta_j}}{\lfloor x^*_{\beta_i} \rfloor - x^*_{\beta_i}} \geqslant \frac{\lfloor t\overline{a}_{\beta_i\eta_j} \rfloor - t\overline{a}_{\beta_i\eta_j}}{\lfloor t x^*_{\beta_i} \rfloor - t x^*_{\beta_i}} \tag{3.20}$$

за свако $j = 1, \ldots, n$. Значи, довољно је показати да (3.20) вриједи под условима теореме.

Пошто је $t(x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor) < 1$, вриједи да је

$$0 = \lfloor t(x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor) \rfloor = \lfloor t x^*_{\beta_i} \rfloor - t\lfloor x^*_{\beta_i} \rfloor,$$

одакле слиједи да је $\lfloor t x^*_{\beta_i} \rfloor = t\lfloor x^*_{\beta_i} \rfloor$. Уз посљедњу једнакост, (3.20) је истовјетно неједнакости

$$\frac{\lfloor \overline{a}_{\beta_i\eta_j} \rfloor - \overline{a}_{\beta_i\eta_j}}{\lfloor x^*_{\beta_i} \rfloor - x^*_{\beta_i}} \geqslant \frac{\lfloor t\overline{a}_{\beta_i\eta_j} \rfloor - t\overline{a}_{\beta_i\eta_j}}{t(\lfloor x^*_{\beta_i} \rfloor - x^*_{\beta_i})}, \tag{3.21}$$

што је еквивалентно неједнакости $t\lfloor \overline{a}_{\beta_i\eta_j} \rfloor \leqslant \lfloor t\overline{a}_{\beta_i\eta_j} \rfloor$ за позитивно t, а то и стоји у услову теореме.

Да би се избјегао случај $t = 1$, због услова $t(x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor) < 1$, $x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor$ мора бити строго мање од $1/2$, а онда и $1/2 \leqslant t(x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor)$. □

Примједба 3.12. Иако је условима теореме искључен случај када је $t = 1$, јасно је да не мора значити да је тиме онемогућен случај да се претходном техником добије истовјетна одсијецајућа раван. На примјер, за исходни ред симплексне табеле

$$x_2 + \frac{4}{15}x_4 + \frac{4}{15}x_5 = \frac{16}{3},$$

добија се Гоморијева одсијецајућа раван

$$-\frac{4}{15}x_4 - \frac{4}{15} \leqslant -\frac{1}{3}.$$

Пошто је $x^*_{\beta_i} = \frac{1}{3}$, а то је мање од $\frac{1}{2}$, првом техником за $t = 2$ се добија одсијецајућа раван

$$-\frac{8}{15}x_4 - \frac{8}{15}x_5 \leqslant -\frac{2}{3},$$

која је у ствари истовјетна претходној, само је скалирана са два. Разлог томе је што за одговарајуће a, може да постоји цјелобројно $t \geqslant 2$, такво да $t\lfloor a \rfloor = \lfloor ta \rfloor$

Што се тиче друге технике, за њен доказ биће потребна сљедећа теорема.

Теорема 3.13 (Волси [59]). *Ако је* $Q = \{(x, y) \in \mathbb{R}_+ \times \mathbb{Z} \ : \ y \leqslant b + x\}$, *и* $b - \lfloor b \rfloor > 0$, *онда је неједнакост*

$$y \leqslant \lfloor b \rfloor + \frac{x}{1 - (b - \lfloor b \rfloor)}$$

ваљана за Q.

Слиједи резултат који се односи на другу технику.

Теорема 3.14 (Друга техника [32]). *Означимо са* $r_{\beta_i\eta_j} = \overline{a}_{\beta_i\eta_j} - \lfloor \overline{a}_{\beta_i\eta_j} \rfloor$ *и* $r^*_{\beta_i} = x^*_{\beta_i} - \lfloor x^*_{\beta_i} \rfloor$. *Неједнакост*

$$r^*_{\beta_i} \leqslant \sum_{j=1}^{n} \min\left\{ r_{\beta_i\eta_j},\ r^*_{\beta_i}\frac{1 - r_{\beta_i\eta_j}}{1 - r^*_{\beta_i}} \right\} x_{\eta_j} \tag{3.22}$$

је ваљана, штавише она је јача или пак, у најгорем случају, једнака изворној Гоморијевој одсијецајућој равни ($\overline{G}_{\beta_i}$).

Доказ. Разбијајући небазне промјенљиве на два дисјунктна скупа тако да је

$$x_{\beta_i} + \sum_{r_{\beta_i\eta_j} \leqslant r^*_{\beta_i}} \overline{a}_{\beta_i\eta_j} x_{\eta_j} + \sum_{r_{\beta_i\eta_j} > r^*_{\beta_i}} \overline{a}_{\beta_i\eta_j} x_{\eta_j} = x^*_{\beta_i}, \tag{3.23}$$

добијена једнакост се може записати као

$$x_{\beta_i} + \sum_{r_{\beta_i\eta_j} \leqslant r^*_{\beta_i}} \lfloor \overline{a}_{\beta_i\eta_j} \rfloor x_{\eta_j} + \sum_{r_{\beta_i\eta_j} \leqslant r^*_{\beta_i}} r_{\beta_i\eta_j} x_{\eta_j} + \sum_{r_{\beta_i\eta_j} > r^*_{\beta_i}} \lceil \overline{a}_{\beta_i\eta_j} \rceil x_{\eta_j} - \sum_{r_{\beta_i\eta_j} > r^*_{\beta_i}} (1 - r_{\beta_i\eta_j}) x_{\eta_j} = x^*_{\beta_i},$$

а преуређењем операнада, пребацујући цјелобројне дијелове на лијеву страну, а остале на десну, једнакост се трансформише у

$$x_{\beta_i} + \sum_{r_{\beta_i\eta_j} \leqslant r^*_{\beta_i}} \lfloor \overline{a}_{\beta_i\eta_j} \rfloor x_{\eta_j} + \sum_{r_{\beta_i\eta_j} > r^*_{\beta_i}} \lceil \overline{a}_{\beta_i\eta_j} \rceil x_{\eta_j} = x^*_{\beta_i} - \sum_{r_{\beta_i\eta_j} \leqslant r^*_{\beta_i}} r_{\beta_i\eta_j} x_{\eta_j} + \sum_{r_{\beta_i\eta_j} > r^*_{\beta_i}} (1 - r_{\beta_i\eta_j}) x_{\eta_j}.$$

Из посљедње једнакости се добија сљедећа

$$x_{\beta_i} + \sum_{r_{\beta_i\eta_j} \leqslant r^*_{\beta_i}} \lfloor \overline{a}_{\beta_i\eta_j} \rfloor x_{\eta_j} + \sum_{r_{\beta_i\eta_j} > r_{\beta_i}} \lceil \overline{a}_{\beta_i\eta_j} \rceil x_{\eta_j} \leqslant x^*_{\beta_i} + \sum_{r_{\beta_i\eta_j} > r^*_{\beta_i}} (1 - r_{\beta_i\eta_j}) x_{\eta_j}.$$

Примјењујући на њу Теорему 3.13, добија се ваљана неједнакост

$$x_{\beta_i} + \sum_{r_{\beta_i\eta_j}\leqslant r^*_{\beta_i}} \lfloor \bar{a}_{\beta_i\eta_j}\rfloor x_{\eta_j} + \sum_{r_{\beta_i\eta_j}>r^*_{\beta_i}} \lceil \bar{a}_{\beta_i\eta_j}\rceil x_{\eta_j} \leqslant \lfloor x^*_{\beta_i}\rfloor + \sum_{r_{\beta_i\eta_j}>r^*_{\beta_i}} \frac{1-r_{\beta_i\eta_j}}{1-r^*_{\beta_i}} x_{\eta_j}.$$

Базно x_{β_i} из (3.23) се може уврстити у посљедњу неједнакост, послије чега треба настали израз упростити да се добије

$$r^*_{\beta_i} \leqslant \sum_{r_{\beta_i\eta_j}\leqslant r^*_{\beta_i}} r_{\beta_i\eta_j} x_{\eta_j} + \sum_{r_{\beta_i\eta_j}>r^*_{\beta_i}} \frac{r^*_{\beta_i}(1-r_{\beta_i\eta_j})}{1-r^*_{\beta_i}} x_{\eta_j}. \tag{3.24}$$

Пошто је за $r_{\beta_i\eta_j} \leqslant r^*_{\beta_i}$

$$r_{\beta_i\eta_j} \leqslant r^*_{\beta_i} = r^*_{\beta_i}\frac{1-r^*_{\beta_i}}{1-r^*_{\beta_i}} \leqslant r^*_{\beta_i}\frac{1-r_{\beta_i\eta_j}}{1-r^*_{\beta_i}}, \tag{3.25}$$

а за $r_{\beta_i,\eta_j} > r^*_{\beta_i}$

$$r^*_{\beta_i}\frac{1-r_{\beta_i\eta_j}}{1-r_{\beta_i}} \leqslant r^*_{\beta_i}\frac{1-r^*_{\beta_i}}{1-r^*_{\beta_i}} = r^*_{\beta_i} < r_{\beta_i\eta_j}, \tag{3.26}$$

неједнакост (3.24) се може изнова написати као тражена (3.22), која је самим тим ваљана и јача од изворне Гоморијеве $(\overline{G}_{\beta_i})$. □

Примједба 3.15. Да би се одсијецајућа раван (3.22) коректно прикључила симплексној таблици, потребно ју је помножити са -1.

Примједба 3.16. Иначе, неједнакост (3.22) је посебан случај Гоморијевих мјешовито-цјелобројних одсијецајућих равни, о којима ће бити ријечи у наредном поглављу.

Пошто нису сви коефицијенти једнаки $r_{\beta_i\eta_j}$, прва техника се не може надовезати на другу, но с друге стране, друга техника се може надовезати на прву, односно

$$\sum_{j=1}^{n} \min\left\{ r(t\bar{a}_{\beta_i,\eta_j}),\ r(t\bar{x}^*_{\beta_i})\frac{1-r(t\bar{a}_{\beta_i,\eta_j})}{1-r(t\bar{x}_{\beta_i^*})} \right\} x_{\eta_j} \geqslant r(t\bar{x}^*_{\beta_i}). \tag{3.27}$$

Ово ће представљати *трећу технику* појачавања Гоморијевих одсијецајућих равни.

Коришћењем истог доказа да друга техника даје јаче одсијецајуће равни од изворних Гоморијевих, може се показати и да су равни добијене трећом техником јаче од равни добијених првом техником. С друге стране, друга и трећа техника нису међусобно упоредиве. Испоставља се да је друга техника доста добра, а и са рачунског гледишта корисна, када је $r^*_{\beta_i}$ јако мало, пошто је већа вјероватноћа да ће коефицијенти у (3.22) бити већи од $r^*_{\beta_i}$. Код треће технике пак, прво се увећава $r^*_{\beta_i}$ што је могуће више, без прекорачења прага цјелобројности, а потом примјењује друга техника која у том случају не мора бити нешто дјелотворна (мања је вјероватноћа). Када је $r^*_{\beta_i}$ велико (веће од $\frac{1}{2}$) онда се не користи прва техника, па је трећа техника једнако добра као и друга.

Пред крај, биће представљена и *четврта техника* појачавања Гоморијевих одсијецајућих равни. Ради краћег записа уводи се функција разломљеног дијела $r(x) = x - \lfloor x\rfloor$.

Теорема 3.17 (Четврта техника [48]). *Претпоставимо да је* $\lfloor x^*_{\beta_i} \rfloor > 0$ *и нека је* $k \geqslant 1$ *јединствени цио број, такав да је* $\frac{1}{k+1} \leqslant r(x^*_{\beta_i}) < \frac{1}{k}$. *Партиционисавши* η *на класе* $\eta^0, \eta^1, \dots, \eta^k$ *тако да је*

$$\eta^0 = \{j \in \eta \ : \ r(\overline{a}_{\beta_i,j}) \leqslant r(x^*_{\beta_i})\},$$

и за $p = 1, 2, \dots, k$ *је*

$$\eta^p = \{j \in \eta \ : \ r(x^*_{\beta_i}) + (p-1)\frac{1 - r(x^*_{\beta_i})}{k} < r(\overline{a}_{\beta_i,j}) \leqslant r(x^*_{\beta_i}) + p\frac{1 - r(x^*_{\beta_i})}{k}\},$$

сљедећа неједнакост,

$$\sum_{j\in\eta^0} r(\overline{a}_{\beta_i,j})x_j + \sum_{p=1}^{k}\sum_{j\in\eta^p}(r(\overline{a}_{\beta_i,j}) - \frac{p}{k+1})x_j \geqslant r(x^*_{\beta_i}), \tag{3.28}$$

је ваљана за цјелобројни омотач у ЛП-релаксацији цјелобројног програма. Штавише, дата одсијецајућа раван је јача или пак, у најгорем случају, једнака одговарајућој изворној Гоморијевој одсијецајућој равни $(\overline{G}_{\beta_i})$*.*

Доказ. Из Гоморијеве одсијецајуће равни G_{β_i}, имајући у виду да је функција $\lfloor\cdot\rfloor$ суперадитивна, увиђа се да је $\lfloor kx^*_{\beta_i}\rfloor = k\lfloor x^*_{\beta_i}\rfloor$ и да је $\lfloor k\overline{a}_{\beta_i\eta_j}\rfloor \geqslant k\lfloor\overline{a}_{\beta_i\eta_j}\rfloor + p$ за партиционисање $q^p = \{r(\overline{a}_{\beta_i\eta_j}) \ : \ p/k \leqslant r(\overline{a}_{\beta_i\eta_j}) < (p+1)/k\}$ $(i = 0, 1, \dots, k-1)$. На основу тога добијамо сљедећу неједнакост

$$\sum_{p=0}^{k-1}\sum_{i\in q^p}(k\lfloor\overline{a}_{\beta_i\eta_j}\rfloor + p)x_{\eta_j} \leqslant k\lfloor x^*_{\beta_i}\rfloor. \tag{3.29}$$

Ако се узме мало $e > 0$, такво да је $1 - e \geqslant \frac{r(x^*_{\beta_i})}{\overline{a}_{\beta_i\eta_j}}$ за све $i \in \eta\setminus\eta_0$, онда важи (лако се показује) да су $\frac{1-e}{r(x^*_{\beta_i})}$ и $1 + (1 - \frac{1-e}{r(x^*_{\beta_i})})/k$ ненегативни. Стога, множећи $(\overline{G}_{\beta_i})$ са $\frac{1-e}{r(x^*_{\beta_i})}$, (3.29) са $1 + (1 - \frac{1-e}{r(x^*_{\beta_i})})/k$, па их потом сабирајући, добијамо ваљану неједнакост

$$\sum_{p=0}^{k-1}\sum_{i\in q^p}\left((k+1)\lfloor\overline{a}_{\beta_i\eta_j}\rfloor + p + \frac{(1-e)(r(\overline{a}_{\beta_i\eta_j} - p/k))}{r(x^*_{\beta_i})} + p/k\right)x_{\eta_j}.$$

Лако се провјерава да је терм

$$\frac{(1-e)(r(\overline{a}_{\beta_i\eta_j} - p/k))}{r(x^*_{\beta_i})} + p/k$$

(са лијеве стране) ненегативан и мањи од два, а премашује јединицу ако и само ако је $r(\overline{a}_{\beta_i\eta_j}) > r(x^*_{\beta_i}) + p(1 - r(x^*_{\beta_i}))/k$. Примијењујући одсијецање разломљеног дијела и препартиционисавши на скупове из Теореме, добијамо неједнакост

$$(k+1)x_{\beta_i} + \sum_{p=1}^{k}\sum_{i\in\eta^p}((k+1)\lfloor\overline{a}_{\beta_i\eta_j}\rfloor + p)x_{\eta_j} \leqslant (k+1)\lfloor x^*_{\beta_i}\rfloor.$$

Подијеливши посљедњу добијену неједнакост са $k+1$, одузевши добијено од $(\overline{G}_{\beta_i})$, а потом преуредивши према партиционисању из услова теореме, добијамо тражену неједнакост (3.28)

Пошто су, очито, коефицијенти неједнакости (3.28) мало већи од коефицијената неједнакости $(\overline{G}_{\beta_i})$, онда је (3.28) јача неједнакост. □

Обично се ове одсијецајуће равни називају *јаким Гоморијевим одсијецајућим равнима.*

Примјер 3.5. За исходни ред симплексне таблице оличен у једнакости

$$x_2 - \frac{1}{15}x_4 + \frac{4}{15}x_5 = \frac{16}{3}.$$

Имамо сљедеће одсијецајуће равни:

- изворна Гоморијева одсијецајућа раван
$$\frac{14}{15}x_4 + \frac{4}{15}x_5 \geqslant \frac{1}{3},$$
- одсијецајућа раван добијена првом техником (јер је $x^*_{\beta_i} = \frac{1}{3}$ што је мање од $\frac{1}{2}$) за, на примјер, $t = 2$
$$\frac{13}{15}x_4 + \frac{8}{15}x_5 \geqslant \frac{2}{3}$$
- одсијецајућа раван добијена другом техником
$$\begin{aligned} &\min\{\frac{14}{15}, \frac{1}{3}\frac{\frac{1}{15}}{\frac{2}{3}}\}x_4 + \min\{\frac{4}{15}, \frac{1}{3}\frac{\frac{11}{15}}{\frac{2}{3}}\}x_5 \geqslant \frac{1}{3} \\ \Leftrightarrow &\frac{1}{30}x_4 + \frac{4}{15}x_5 \geqslant \frac{1}{3} \end{aligned}$$
- одсијецајућа раван добијена трећом техником
$$\begin{aligned} &\min\{\frac{13}{15}, \frac{2}{3} \cdot \frac{\frac{2}{15}}{\frac{1}{3}}\}x_4 + \min\{\frac{8}{15}, \frac{2}{3} \cdot \frac{\frac{7}{15}}{\frac{1}{3}}\}x_5 \geqslant \frac{2}{3} \\ \Leftrightarrow &\frac{2}{15}x_4 + \frac{4}{15}x_5 \geqslant \frac{1}{3} \end{aligned}$$
- одсијецајућа раван добијена четвртом техником — $\eta^0 = \{5\}$, $\eta^1 = \emptyset$, $\eta^2 = \{4\}$.
$$\begin{aligned} &\frac{4}{15}x_5 + (\frac{14}{15} - \frac{2}{3})x_4 \geqslant \frac{1}{3} \\ \Leftrightarrow &\frac{4}{15}x_4 + \frac{4}{15}x_5 \geqslant \frac{1}{3} \end{aligned}$$
■

Као што можемо примијетити, четврта техника није упоредива у општем случају са другом и трећом техником.

Као и за трећу технику, могуће је надовезати четврту технику на прву, што би била *пета техника.* За пету технику важе слична разматрања у односу на остале технике, као и за трећу у односу на прве двије.

Јаче одсијецајуће равни могу се, свакако, добити искориштавањем структуре конкретног цјелобројног програма.

3.6.1 Појачавање Хватал-Гоморијевих одсијецајућих равни

Што се тиче Хватал-Гоморијевих одсијецајућих равни, природно, важе слични резултати (биће наведена само тврђења, доказе погледати у [48]). Слиједи аналогон прве технике (Теореме 3.5) за Гоморијеве одсијецајуће равни.

Теорема 3.18. *Ако је* $r(\tau^T b) < 1/2$, t *било који позитиван цио број такав да је* $1/2 \leqslant t \cdot r(\tau^T b) < 1$, *онда, замјењујући* τ *са* $r(t\tau)$ *добијамо јачу или пак, у најгорем случају, једнако јаку Хватал-Гоморијеву одсијецајућу раван.*

За представљање аналогона другој техници, згодно је увести додатне ознаке. Нека је $N = \{1, \dots, n\}$ и $a_0 = \tau^T b$, а за $i \in N$ нека је a_i једнако i-тој компоненти вектор-врсте матрице $\tau^T A$. Очито је да се ваљана неједнакост $\sum_{i \in N} (\tau^T A) x_i \leqslant \tau^T b$ може записати као

$$\sum_{i \in N} a_i x_i \leqslant a_0,$$

а Хватал-Гоморијева одсијецајућа раван као

$$\sum_{i \in N} \lfloor a_i \rfloor x_i \leqslant \lfloor a_0 \rfloor.$$

Слиједи аналогон друге технике (Теореме 3.7).

Теорема 3.19. *Неједнакост*

$$\sum_{i=1}^{n} \left(\lfloor a_i \rfloor + \max \left\{ 0, \frac{r(a_i) - r(a_0)}{1 - r(a_0)} \right\} \right) \leqslant \lfloor a_0 \rfloor. \tag{3.30}$$

је ваљана и јача је или пак, у најгорем случају, једнака изворној Хватал-Гоморијевој одсијецајућој равни.

Поново, технике се могу надовезати (друга на прву), при чему се τ замјењује са $r(t\tau)$ да би се добио аналогон треће технике.

На крају, биће представљен аналогон четвртој техници (Теорема 3.8) појачавања Гоморијевих одсијецајућих равни.

Теорема 3.20. *Претпоставимо да је* $r(x^*_{\beta_i}) > 0$ *и нека је* $k \geqslant 1$ *јединствени цио број такав да је*

$$\frac{1}{k+1} \leqslant r(x^*_{\beta_i}) < \frac{1}{k}.$$

Партиционишимо η *на класе* $\eta_\mathbf{0}, \dots, \eta_\mathbf{k}$ *тако да је* $\eta_\mathbf{0} = \{j \in \eta \ : \ r(\alpha_i) \leqslant r(x^*_{\beta_i})\}$ *и за* $p = 1, \dots, k$ *нека је* $\eta_p = \{j \in \eta \ : \ r(x^*_{\beta_i}) + (p-1)(1 - x^*_{\beta_i})/k < r(\alpha_j) \leqslant x^*_{\beta_i} + p(1 - r(x^*_{\beta_i}))/k\}$. *Сљедећа појачана разломљена одсијецајућа раван је ваљана за цјелобројни омотач у ЛП-релаксацији цјелобројног програма и јача је или пак, у најгорем случају, једнака изворној Хватал-Гоморијевој одсијецајућој равни:*

$$\sum_{j \in \eta_0} \lfloor a_i \rfloor x_i + \sum_{p=1}^{k} \sum_{j \in \eta_p} (\lfloor a_i \rfloor + \frac{p}{k+1}) x_i \leqslant \lfloor a_0 \rfloor.$$

3.7 Гоморијеве мјешовито-цјелобројне одсијецајуће равни

Аргументација за извођење Гоморијевих или Хватал-Гоморијевих одсијецајућих равни се заснива на чињеници да све промјенљиве морају бити цјелобројне, стога се дата аргументација не може непосредно примјенити када треба ријешити мјешовито-цјелобројан програм. Другим ријечима, није примјенљива на мјешовито-цјелобројан полиедар програма

$$P_I = \{(x, z) \in \mathbb{Z}^n_+ \times \mathbb{R}^p_+ \ : \ Ax + Dy \leqslant b\}.$$

Додавајући изједначавајуће промјенљиве, добијамо сљедећи скуп

$$P'_I = \{(x, y) \in \mathbb{Z}^n_+ \times \mathbb{R}^n_+ \ : \ Ax + Gy = b\},$$

гдје је $G = [D\ I]$ и $y = [z\ s]^T$. Једнакости се, поново, множе са вектором $u \in \mathbb{R}^m$ да би се добио скуп

$$P'_I(u) = \{(x, y) \in \mathbb{Z}^n_+ \times \mathbb{R}^{p+m}_+ \ : \ \overline{a}x + \overline{g}y = \overline{b}\},$$

у ком је $\overline{a} = uA$, $\overline{g} = uG$ и $\overline{b} = ub$, при чему се још претпоставља да је разломљени дио $r(b) > 0$. Узевши да је $r_i = r(\overline{a})$ и $r_0 = r(\overline{b})$, $P'_I(u)$ се може записати као

$$\sum_{i:r_i \leqslant r_0} r_i x_i + \sum_{i:r_i > r_0} (r_i - 1)x_i + \overline{g}y = \left(\lfloor \overline{b} \rfloor - \sum_{i=1}^{n} \lfloor \overline{a} \rfloor x_i - \sum_{i:r_i > r_0} x_i\right) + r_0.$$

Пошто је $\lfloor \overline{b} \rfloor - \sum_{i=1}^{n} \lfloor \overline{a}_i \rfloor x_i - \sum_{i:r_i > r_0} x_i = k \in \mathbb{Z}$, онда је или $k \leqslant -1$, па је

$$-\sum_{i:r_i \leqslant r_0} \frac{r_i}{1 - r_0} x_i + \sum_{i:r_i > r_0} \frac{1 - r_i}{1 - r_0} x_i - \frac{\overline{g}}{1 - r_0} y \geqslant 1,$$

или је пак $k \geqslant 0$, па је у том случају

$$\sum_{i:r_i \leqslant r_0} \frac{r_i}{r_0} x_i + \sum_{i:r_i > r_0} \frac{r_i - 1}{r_0} x_i + \frac{\overline{g}}{r_0} y \geqslant 1.$$

Како су обе одсијецајуће равни облика $a^1 \omega \geqslant 1$ и $a^2 \omega \geqslant 1$, при жему важи да је $\omega > 0$, онда одсијецајућа раван

$$\sum_{j} \max\{a^1_j, a^2_j\} \omega \geqslant 1$$

је ваљана за $P'_I(u)$, а тиме и за P'_I. Дакле, неједнакост

$$\sum_{i:r_i \leqslant r_0} \frac{r_i}{r_0} x_i + \sum_{r_i > r_0} \frac{1 - r_i}{1 - r_0} x_i + \sum_{j:\overline{g}_j \geqslant 0} \frac{\overline{g}_j}{r_0} y_j - \sum_{j:\overline{g}_j < 0} \frac{\overline{g}_j}{1 - r_0} y_j \geqslant 1.$$

звана *Гоморијева мјештовито-цјелобројна одсијецајућа раван*, је ваљана за P'_I.

Примједба 3.21. За потпуно цјелобројни случај, Гоморијева мјешовито-цјелобројна одсијецајућа раван је и даље ваљана и облика је

$$\sum_{i:r_i \leqslant r_0} \frac{r_i}{r_0} x_i + \sum_{i:r_i > r_0} \frac{1 - r_i}{1 - r_0} x_i \geqslant 1,$$

односно, одговара одсијецајућој равни добијеној другом техником појачавања.

3.8 Нумеричко-аритметички аспект

Већ је споменуто да је линеарна алгебра, потребна за конструкцију реда симплексне табеле из ког се изводи Гоморијева одсијецајућа раван, подложна грешкама заокруживања када се равни рачунају у коначној аритметици, односно у аритметици покретне запете (стандардној рачунској платформи). Ове нумеричке потешкоће могу у пракси довести до генерисања одсијецајућих равни које нису ваљане за одговарајући примјерак цјелобројног програма. На примјер, одсијецајућа раван за неједнакост $x_1 + x_2 \leqslant 2.005$ (подразумијевајући да $x_1, x_2 \in \mathbb{Z}$) је $x_1 + x_2 \leqslant 2$. Ипак, ако је присутна грешка величине 0.001, у прорачунима можемо добити сљедећу неједнакост $x_1 + x_2 \leqslant 1.995$, па стога и одсијецајућу раван $x_1 + x_2 \leqslant 1$, која уопште не мора бити ваљана. У веома пажљивој студији [50] је документовано не тако ријетко јављање одсијецајућих равни које нису ваљане у оквиру одговарајућих метода за рјешавање (мјештовито-)цјелобројних програма.

3.8.1 Гоморијеве одсијецајуће равни у егзактној аритметици

Разматрање Гоморијевог метода са становишта аритметике у којој се прорачун врши може се, природно, подијелити на два дијела — одређивање тачног рјешења ЛП-релаксације и произвођење Гоморијевих одсијецајућих равни.

Комерцијални софтвери, као на примјер CPLEX, се служе неком толеранцијом у својим прорачунима. Ово доводи до тога да се бројеви блиски нули (испод толеранције) сматрају једнаким нули. Поставља се питање да ли пивотирати над колонама код којих је на примјер негативни редуковани трошак близак нули, или просто претпоставити да је редуковани трошак једнак нули? У једном случају има се посла са дуалном дегенерисаном базом, а у другом случају се не проналази оптимално рјешење.

Примјер 3.6. Размотримо програм

$$\begin{aligned} \min\ & -1.0001x_1 - 2x_2 \\ & x_1 + 2x_2 + x_3 = 10000 \\ & x_1, x_2, x_3 \in \mathbb{N}_0. \end{aligned}$$

Уводећи x_2 у базу добија се сљедећа табела

-0.0001	0	1	10000
$\frac{1}{2}$	1	$\frac{1}{2}$	5000

Ако је толеранција 10^{-4}, онда би претпоставили да је -0.0001 у ствари нула и стога не би пивотирали по тој колони. Ово би дало сљедећи резултат

$$x_1 = 0 \qquad x_2 = 5000 \qquad x_3 = 0,$$

при чему је вриједност циља -10000. Наравно, резултат није оптималан. Оптималан резултат је

$$x_1 = 10000 \qquad x_2 = 0 \qquad x_3 = 0$$

са циљном вриједношћу -10001. Уколико би x_1 представљало мјерице за кукуруз, а x_2 килограме меда које би пољопривредник требао дао прода, онда би ова разлика и те како била значајна. ■

Штавише, ситуација може бити и гора, односно CPLEX може дати погрешан резултат (тачку која није допустива) због толеранције, као у сљедећем примјеру.

Примјер 3.7. Размотримо програм

$$\begin{aligned} \min\ -x_1 - x_2 \\ 0.499999x_1 + 0.5x_2 = 1 \\ x_1 + 2x_2 = 3 \\ x_1, x_2 \geqslant 0. \end{aligned}$$

CPLEX даје као рјешење тачку $x_1 = x_2 = 1$ и вриједност циља -2, међутим, ова тачка није допустива. ■

У егзактној аритметици, сви подаци су представљени у облику разломака, при чему се засебно води рачуна о бројиоцу и имениоцу (тј. њиховој похрани). Све аритметичке операције се изводе у оквиру симболичког рачуна (операције са разломцима), то јест не разматрају се апроксимације резултата аритметичких операција, као што је то случај у запису са запетом (било фиксном, било покретном).

Ипак, егзактна аритметика, иако би рјешила споменуте нумеричке проблеме, има једну јако лошу страну. Главни рачунски проблем егзактне аритметике је невјероватно велик број основних аритметичких операција. За сабирање (одузимање) два броја (улазна податка у нашем смислу), практично су потребна три множења и једно сабирање, док је за множење (дијељење) два броја практично потребно извршити двије операције множења. За поређење два броја потребно је извршити два множење и једну стандарну операцију поређења. Даље, још једна ствар које је проблематична јесте одржавање бројиоца и имениоца разломка узајамно простим, управо да би разломци заузимали што мање меморије. Просто, без скраћивања разломака, бројеви би могли да расту експоненцијално, пошто свака операција у егзактној аритметици садржи множење. Сам поступак скраћивања разломака може бити, временски гледано, најзахтјевнија операција међу свим до сад споменутим. Један од бржих начина (али ипак недовољно брз) за скраћивање разломака је коришћење Еуклидовог алгоритма за налажење највећег заједничког дјелиоца (НЗД). Дати алгоритам се заснива на наредној теореми.

Теорема 3.22. *НЗД* (a, b) = *НЗД* $(\min(a, b),\ |a - b|)$.

Према предоченом у [6], наивно рјешавање линеарног програма у егазктној аритметици, практично гледано, нема смисла, па је потребан другачији приступ.

Дифлауи *et al.* је у [24] примјенио нешто другачију стратегију (провјери-па-поправи), допуштајући рјешавање линеарног програма у аритметици покретне запете, које би потом служило као почетна тачка за добијање тачног рјешења уколико оно није добијено. Умјесто разматрања прималног и дуалног рјешења добијеног рачуном у покретној запети, Дифлауи *et al.* узима опис предочене (оптималне) базе за коју очекује да је оптимална или бар блиска оптималној. Дату база се провјерава рачуном у егзактној аритметици (одређује да ли је примално или дуално допустива), послије чега прелази на израчунавање оптималног рјешења, такође у егзактној аритметици, почевши, свакако, од те саме базе (уколико рјешење заиста није оптимално). Међутим, Дифлауи није успио ријешити све проблеме из NETLIB библиотеке, јер се у његовом приступу и даље јавља исти рачунски проблем превеликог броја операција, односно великог времена извршавања алгоритма.

С друге стране, Кох је у [45], исте године, извјестио о рјешавању свих проблема NETLIB библиотеке, модификујући приступ Дифлаоуија, просто прерачунавајући одговарајуће захтјевније проблеме користећи двоструко већу прецизност записа са покретном запетом и тиме заобилазећи симболички рачун. У [6] Кохова стратегија је унапријеђена динамичким повећавањем прецизност рачунања у покретној запети све док се не добије база која даје рационална примална и дуална рјешења, а при чему су испуњени услови оптималности, неомеђености, или пак недопустивости.

Метод тачног рјешавања линарног програма према [6]

```
1. $p \leftarrow$ најбоља уграђена прецизност записа са покретном запетом
2. $\mathcal{B} \leftarrow \emptyset$
3. Израчунати апроксимације $\overline{c}, \overline{b}, \overline{A}$ почетног улаза у текућој прецизности
4. Ријеши $\max\{\overline{c}x \ : \ \overline{A}x \leqslant \overline{b}, \ x \geqslant 0\}$ користећи $\mathcal{B}$ (ако није празан скуп)
5. $\mathcal{B} \leftarrow$ крајња база симплексног алгоритма.
6. Тестирај резултат у егзактној аритметици.
7. Ако је тест негативан
      - повећај прецизност $p$
      - иди на корак 3
8. Испиши резултат (стоп)
```

Имајући у виду претходне начине за добијање тачног рјешења ЛП-релаксације, из одговарајуће базе се може произвести Гоморијева одсијецајућа раван, штавише она се у случају потребе може произвести са цјелобројним коефицијентима. Неједнакост само треба помножити са највећим имениоцем међу коефицијентима, па потом подијелити са највећим заједничким дјелиоцем новодобијених коефицијената. На примјер, за

$$-\frac{4}{15}x_3 - \frac{2}{15}x_4 \leqslant -\frac{2}{3}$$

добијамо сљедећи облик одсијецајуће равни

$$-2x_3 - x_4 \leqslant -5.$$

3.8.2 Поуздане Гоморијеве одсијецајуће равни у покретној запети

Користећи рутине програмског језика C могуће је произвести ваљане и нумерички поуздане Гоморијеве одсијецајуће равни у аритметици покретне запете, а које се због тога могу инкорпорирати у одговарајуће софтвере. Другим ријечима, могуће је доћи до тачног оптималног рјешења, уз доста мање коришћења егзактне аритметике, или пак повећавања прецизности записа.

Запис броја у покретној запети према IEEE 754 стандарду, састоји се од знака, значајног дијела (мантисе), степена и основе бројчаног система (увијек мора бити парна). На примјер, за запис броја у двострукој прецизности, 11 бита се одваја за степен, а 52 бита за значајни дио. Узећемо да је $\mathbb{M} \subseteq \mathbb{Q}$ скуп рационалних бројева репрезентабилних у запису са покретном запетом. $\mathbb{M}$ је коначан, и ако је $q \in \mathbb{M}$, онда важи да су и $-q$, $r(q) \in \mathbb{M}$. Такође, ако су a, $b \in \mathbb{M}$, онда су и $\min\{a, b\}$, $\max\{a, b\} \in \mathbb{M}$. Скуп $\mathbb{M}$ није нешто посебно структуиран, пошто операција сабирања није асоцијативна над њим.

За дату функцију $f : \mathbb{R}^n \to \mathbb{R}$, каже се да функција $f^{\text{одозго}} : \mathbb{M}^n \to \mathbb{M}$ *одозго апроксимира функцију f на скупу* $\mathbb{M}$, ако

$$f^{\text{одозго}}(x) \geqslant f(x), \quad x \in \mathbb{M}^n,$$

а за функцију $f^{\text{одоздо}} : \mathbb{M}^n \to \mathbb{M}$ каже се да *одоздо апроксимира функцију f на* $\mathbb{M}$, ако важи

$$f^{\text{одоздо}}(x) \leqslant f(x), \quad x \in \mathbb{M}^n.$$

Можемо примијетити да су основне аритметичке операције $+$ и $*$ функције са $\mathbb{R}^2 \to \mathbb{R}$. Означавамо респективно са $\overline{a+b}$ и $\underline{a+b}$ вриједности апроксимација одозго/одоздо збира $a+b$ за било које $a, b, a+b \in \mathbb{M}$ и користимо слично означавање за одузимање и множење.

У програмском језику C (ISO/IEC ’99), конвенција о заокруживању у покретној запети се, према стандарду IEEE, може подесити да даје споменуте апроксимације одозго и одоздо (умјесто подразумијеваног заокруживања на најближу вриједност) помоћу fesetround функције. Апроксимације одозго се добијају позивом функције са аргументом FE_UPWARD, послије чега се врше уобичајене аритметичке операције. Слично, апроксимације одоздо се добијају коришћењем FE_DOWNWARD аргумента.

Нека је $k \geqslant 3$ цио број и нека су a_i, $\pi_i \in \mathbb{M}$ за $i = 1, \ldots, k$. За $\sum_{i=1}^{k} \pi_i a_i$ рекурзивно дефинишемо апроксимацију одозго, односно одоздо,

$$\overline{\sum_{i=1}^{k} \pi_i a_i} = \overline{\left(\overline{\sum_{i=1}^{k-1} \pi_i a_i}\right) + \overline{\pi_k a_k}}$$

$$\underline{\sum_{i=1}^{k} \pi_i a_i} = \underline{\left(\underline{\sum_{i=1}^{k-1} \pi_i a_i}\right) + \underline{\pi_k a_k}}.$$

Битно је примијетити да ће било који поредак елемената $a_1, \ldots, a_k$ дати апроксимацију одозго/одоздо суме $\sum_{i=1}^{k} \pi_i a_i$, али различити поретци могу дати различите резултате. Стога, операција није комутативна.

Такође, битно је примијетити да апроксимације одозго/одоздо неће бити исправне ако се јави прекорачење/поткорачење, па када се год то деси, може се узети да је резултат недефинисан.

Разматрајући Хватал-Гоморијеве одсијецајуће равни, може се одмах примијетити да је за скуп множиоца $\tau \in \mathbb{M}^m$, већих од нуле, неједнакост $\tau^T Ax \leqslant \tau^T b$ задовољена у свим тачкама цјелобројног омотача P_I у полиедру ЛП-релаксације P. Ипак, дата неједнакост не мора бити $\mathbb{M}$-репрезентабилна, а израчунавање $\tau^T A$ и $\tau^T b$ у аритметици покретне запете, не обезбјеђује ваљаност резултујуће неједнакости.

Подразумијева се да је $A \in \mathbb{M}^{m \times n}$ и $b \in \mathbb{M}^{n}$. Без обзира на то, ако је $x \geqslant 0$, важи да је

$$\sum_{j \in \eta} \underline{\left(\sum_{i=1}^{m} \tau_i a_{ij} \right)} x_j \leqslant \overline{\sum_{i=i}^{m} \tau_i b_i}$$

$\mathbb{M}$-репрезентабилна ваљана неједнакост, од које се може добити ваљана одсијецајућа раван.

У случају једнакости, као када се производи агрегирана Гоморијева одсијецајућа раван, једнакости се морају опустити до неједнакости, па се онда неједнакости агрегирају коришћењем множиоца τ да би се добила ваљана $\mathbb{M}$-репрезентабилна неједнакост. Иначе, тада τ не мора бити ограничено на ненегативне вриједности. Очито, ако желимо да добијемо мање-или-једнако $\mathbb{M}$-репрезентабилну неједнакост, а множилац неког конкретног ограничења је негативан, релаксација мора бити супротна, то јест веће-или-једнако неједнакост.

Остале Гоморијеве одсијецајуће равни се на сличан начин трансформишу у нумерички поуздане одсијецајуће равни. Јасно је да је тако добијена одсијецајућа раван слабија у односу на изворну.

На крају, као што је већ напоменутно на неки начин, произвођење нумерички поузданих одсијецајућих равни није јемац да софтвер неће одсијећи оптимално рјешење, јер је потребно да се остатак рачунског софтвера трансформише да ради у нумерички поузданом режиму (нпр. поуздане дуалне границе, претпроцесирање, провјера о неомеђености, итд.).

Глава 4

Корнерски полиедар и одсијецајуће равни

Дат је сљедећи полиедар

$$\begin{aligned} &Ax = b \\ &x_j \geqslant 0,\ j = 1, \dots, n \\ &x_j \in \mathbb{Z},\ j = 1, \dots, p, \end{aligned} \tag{4.1}$$

гдје је $p \leqslant n$, матрица $A \in \mathbb{Q}^{m \times n}$ (ранга m) и вектор-колона $b \in \mathbb{Q}^m$. Скуп базних индекса (тј. база) биће, као и досад, означаван са $\boldsymbol{\beta}$, а скуп индекса небазних промјенљивих са $\boldsymbol{\eta} = \{1, \dots, n\} \backslash \boldsymbol{\beta}$. Претходни систем једначина и неједначина може се записати у облику

$$x_i = x_i^* - \sum_{j \in \boldsymbol{\eta}} \overline{a}_{ij} x_j, \quad i \in \boldsymbol{\beta} \tag{4.2}$$

гдје је $x_i^* \geqslant 0$ за $i \in \boldsymbol{\beta}$. Одговарајуће базно рјешење је $\overline{x}_i = x_i^*$ ($i \in \boldsymbol{\beta}$), $\overline{x}_j = 0$ ($j \in \boldsymbol{\eta}$). Ако је $x_i^* \in \mathbb{Z}$ за све $i \in \boldsymbol{\beta} \cap \{1, \dots, p\}$, онда $\overline{x}$ припада скупу (4.1), у противном треба наћи ваљане одсијецајуће равни за скуп (4.1) које су нарушене у тачки $\overline{x}$.

Једна, мало другачија идеја за проналажење одсијецајућих равни је да се умјесто полиедра (4.1) користи Гоморијев корнерски полиедар [33, 34], који се добија из (4.1) одбацивањем ограничења ненегативности по свим базним промјенљивим x_i ($i \in \boldsymbol{\beta}$). У таквој релаксацији могу се одбацити и ограничења $x_i = x_i^* - \sum_{j \in \boldsymbol{\eta}} \overline{a}_{ij} x_j$ за све $i \in \boldsymbol{\beta} \cap \{p+1, \dots, n\}$ зато што промјенљиве x_i:

- не морају бити цјелобројне
- јављају се тачно једном у систему, а ни једна друга базна промјенљива не зависи од њих.

Због тога ће се подразумијевати да су све базне промјенљиве у (4.2) цјелобројне, то јест $\boldsymbol{\beta} \subseteq \{1, \dots, p\}$.

Опуштење за (4.1) које је увео Гомори је

$$\begin{aligned} &x_i = x_i^* - \sum_{j \in \boldsymbol{\eta}} \overline{a}_{ij} x_j && i \in \boldsymbol{\beta} \\ &x_j \geqslant 0 && j \in \boldsymbol{\eta} \\ &x_i \in \mathbb{Z} && i = 1, \dots, p. \end{aligned} \tag{4.3}$$

Конвексни омотач допустивих рјешења (4.3) се назива *корнерски полиедар за базу* $\boldsymbol{\beta}$ и означава се са $\mathrm{corner}(\boldsymbol{\beta})$. Свака ваљана одсијецајућа раван за корнерски полиедар је ваљана и за скуп (4.1).

Ради краћег записа, са $P(\boldsymbol{\beta})$ означаваће се ЛП-релаксација корнерског полиедра (4.3). Корнерски полиедар је боље структуиран у односу на одговарајући полиедар мјешовито-цјелобројног програма, па се врхови и екстремални зраци релаксације $P(\boldsymbol{\beta})$ могу простије описати, што се, наравно, показало корисним у произвођењу ваљаних одсијецајућих равни.

Полиедар $P(\boldsymbol{\beta})$ има само један врх и то у тачки са координатама $x_i = x_i^*$ за $i \in \boldsymbol{\beta}$ и $x_j = 0$ за $j \in \boldsymbol{\eta}$.

Рецесивни конус полиедра $P(\boldsymbol{\beta})$ је одређен сљедећим системом

$$\begin{aligned} x_i &= -\sum_{j\in\boldsymbol{\eta}} \overline{a}_{ij} x_j \qquad && \text{за } i \in \boldsymbol{\beta} \\ x_j &\geqslant 0 && \text{за } j \in \boldsymbol{\eta}. \end{aligned}$$

Пошто је пројекција овог конуса на $\mathbb{R}^{\boldsymbol{\eta}}$ одређена неједнакостима $x_j \geqslant 0$, $(j \in \boldsymbol{\eta})$, а промјенљиве x_i, $i \in \boldsymbol{\beta}$ су одређене горњим једначинама, екстремални зраци конуса су вектори који, осим једног, задовољавају сва остала ограничења ненегативности као једнакости. Стога, постоји $|\boldsymbol{\eta}|$ екстремалних зрака $\overline{z}^j$ за $j \in \boldsymbol{\eta}$, дефинисаних као

$$\overline{z}_h^j = \begin{cases} -\overline{a}_{hj} & \text{за } h \in \boldsymbol{\beta}, \\ 1 & \text{за } j = h, \\ 0 & \text{за } h \in \boldsymbol{\eta}\backslash\{j\}. \end{cases} \tag{4.4}$$

Примједба 4.1. Вектори $\overline{z}^j$, при чему $j \in \boldsymbol{\eta}$, су линеарно независни. Отуда је $P(\boldsymbol{\beta})$ $|\boldsymbol{\eta}|$-димензионални полиедар чији је афини омотач одређен једначинама $x_i = x_i^* - \sum_{j\in\boldsymbol{\eta}} \overline{a}_{ij} x_j$ за $i \in \boldsymbol{\beta}$.

Лема 4.2. *Ако афини омотач ЛП-релаксације* $P(\boldsymbol{\beta})$ *садржи тачку у* $\mathbb{Z}^p \times \mathbb{R}^{n-p}$, *онда је* $\mathrm{corner}(\boldsymbol{\beta})$ $|\boldsymbol{\eta}|$*-димензионални полиедар. У противном,* $\mathrm{corner}(\boldsymbol{\beta})$ *је празан.*

Доказ. Пошто је $\mathrm{corner}(\boldsymbol{\beta})$ садржан у афином омотачу полиедра $P(\boldsymbol{\beta})$, $\mathrm{corner}(\boldsymbol{\beta})$ је празан када афини омотач полиедра $P(\beta)$ не садржи нити једну тачку из $\mathbb{Z}^p \times \mathbb{R}^{n-p}$.

Остаје да се докаже да је $\mathrm{corner}(\boldsymbol{\beta})$ $|\boldsymbol{\eta}|$-димензионални полиедар, ако афини омотач полиедра $P(\boldsymbol{\beta})$ садржи тачку из $\mathbb{Z}^p \times \mathbb{R}^{n-p}$.

За почетак, треба доказати да је $\mathrm{corner}(\boldsymbol{\beta})$ непразан. Ако $x' \in \mathbb{Z}^p \times \mathbb{R}^{n-p}$ припада афином омотачу полиедра $P(\boldsymbol{\beta})$, онда је $x'_i = x_i^* - \sum_{j\in\boldsymbol{\eta}} \overline{a}_{ij} x'_j$ за $i \in \boldsymbol{\beta}$.

Узевши да је $\boldsymbol{\eta}^-$ подскуп индекса из $\boldsymbol{\eta}$ таквих да је $x'_j < 0$, ако је $\boldsymbol{\eta}^-$ празно, онда x'_j припада $\mathrm{corner}(\boldsymbol{\beta})$. Даље, постоји $D \in \mathbb{Z}_+$ такво да је $D\overline{a}_{ij} \in \mathbb{Z}$ за све $i \in \boldsymbol{\beta}$ и $j \in \boldsymbol{\eta}^-$. На основу претходног, може се увести тачка x'', дефинисана на сљедећи начин

$$x''_j = \begin{cases} x'_j & j \in \boldsymbol{\eta}\backslash\boldsymbol{\eta}^- \\ x'_j - D\lfloor \frac{x'_j}{D} \rfloor & j \in \boldsymbol{\eta}^- \\ x_i^* - \sum_{j\in\boldsymbol{\eta}} \overline{a}_{ij} x''_j & i \in \boldsymbol{\beta} \end{cases}.$$

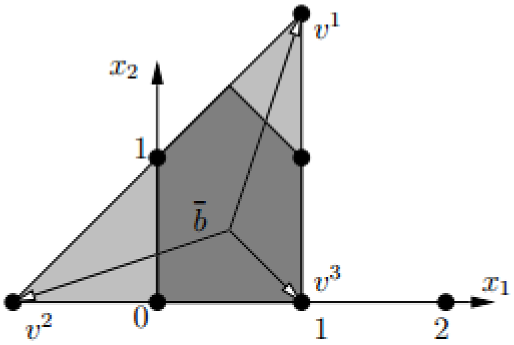
Према конструкцији, $x''_j \geqslant 0$ за све $j \in \boldsymbol{\eta}$ и x''_i је цио број за сваки $i \in \boldsymbol{\beta}$. Пошто x'' задовољава $x''_i = x_i^* - \sum_{j\in\boldsymbol{\eta}} \overline{a}_{ij} x''_j$, слиједи да x'' припада $\mathrm{corner}(\boldsymbol{\beta})$, односно да је $\mathrm{corner}(\boldsymbol{\beta})$ непразан.

Имајући у виду да је $P(\boldsymbol{\beta})$ рационални полиедар, на основу Теореме 1.16, рецесивни конуси полиедара $P(\boldsymbol{\beta})$ и corner$(\boldsymbol{\beta})$ се поклапају. Како је димензија $P(\boldsymbol{\beta})$ и његовог рецесивног конуса $|\boldsymbol{\eta}|$, а corner$(\boldsymbol{\beta}) \subseteq P(\boldsymbol{\beta})$, димензија corner$(\boldsymbol{\beta})$ је $|\boldsymbol{\eta}|$. □

Примјер 4.1. Дат је потпуно цјелобројан програм

$$\begin{aligned}
&\min \frac{1}{2}x_2 + x_3 \\
&x_1 + x_2 + x_3 &&\leqslant 2 \\
&x_1 - \frac{1}{2}x_3 &&\geqslant 0 \\
&x_2 - \frac{1}{2}x_3 &&\geqslant 0 \\
&x_1 + \frac{1}{2}x_3 &&\geqslant 0 \\
&-x_1 + x_2 + x_3 &&\leqslant 1 \\
&x_1, x_2, x_3 \in \mathbb{Z} \\
&x_1, x_2, x_3 \geqslant 0 &&.
\end{aligned} \tag{4.5}$$

Овај проблем има четири допустива рјешења $(0,0,0)$, $(1,0,0)$, $(0,1,0)$ и $(1,1,0)$, при чему сва задовољавају $x_3 = 0$. Дате четири тачке су показане у (x_1, x_2)-простору на Слици 4.1.



Слика 4.1 Пресјек корнерског полиедра са равни $x_3 = 0$

На самом почетку, програм се преводи у облик (4.1) уводећи одговарајуће нецјелобројне изједначавајуће промјенљиве (позитивне или негативне) $x_4, \ldots, x_8$. По рјешавању ЛП-релаксације, добија се

$$\begin{aligned}
x_1 &= \tfrac{1}{2} &+\tfrac{1}{4}x_6 &&-\tfrac{3}{4}x_7 &&+\tfrac{1}{4}x_8 \\
x_2 &= \tfrac{1}{2} &+\tfrac{3}{4}x_6 &&-\tfrac{1}{4}x_7 &&-\tfrac{1}{4}x_8 \\
x_3 &= 1 &-\tfrac{1}{2}x_6 &&-\tfrac{1}{2}x_7 &&-\tfrac{1}{2}x_8 \\
x_4 &= 0 &-\tfrac{1}{2}x_6 &&+\tfrac{3}{2}x_7 &&+\tfrac{1}{2}x_8 \\
x_5 &= 0 &-\tfrac{1}{2}x_6 &&-\tfrac{1}{2}x_7 &&+\tfrac{1}{2}x_8.
\end{aligned}$$

Оптимално базно рјешење је $x_1 = x_2 = \frac{1}{2},\ x_3 = 1,\ x_4 = \ldots = x_8 = 0$.

Одбацујући ненегативност базних промјенљивих и два ограничења која се односе на непрекидне базне промјенљиве x_4 и x_5, добија се (4.3)-формулација за овај примјер

$$\begin{aligned} x_1 &= \tfrac{1}{2} &+\tfrac{1}{4}x_6 &\ -\tfrac{3}{4}x_7 &\ +\tfrac{1}{4}x_8 \\ x_2 &= \tfrac{1}{2} &+\tfrac{3}{4}x_6 &\ -\tfrac{1}{4}x_7 &\ -\tfrac{1}{4}x_8 \\ x_3 &= 1 &-\tfrac{1}{2}x_6 &\ -\tfrac{1}{2}x_7 &\ -\tfrac{1}{2}x_8 \\ x_1, x_2, x_3 &\in \mathbb{Z} \\ x_6, x_7, x_8 &\geqslant 0. \end{aligned} \tag{4.6}$$

На Слици 4.1 унија осијенчених области је пресјек $P(\boldsymbol{\beta})$ са равни $x_3 = 0$. Нека је P полиедар дефинисан неједнакостима (4.5) које су према једначинама (4.6) задовољене у тачки $x^* = (\frac{1}{2}, \frac{1}{2}, 1)$. Пресјек полиедра P са равни $x_3 = 0$ је тамно осијенчена област. Увиђа се да је P строго садржан у $P(\boldsymbol{\beta})$, а то се обично јавља када је база дегенерисана, што овдје и јесте случај и што се, иначе, често јавља у рјешавању цјелобројних програма. ■

За полиедар corner($\boldsymbol{\beta}$), неједнакост $\sum_{j\in\boldsymbol{\eta}} \gamma_j x_j \geqslant \delta$ se назива *тривијалном* ако слиједи непосредно из ограничења ненегативности $x_j \geqslant 0$ $(j \in \boldsymbol{\eta})$, то јест, ако важи да је $\gamma_j \geqslant 0$ за све $j \in \boldsymbol{\eta}$, а $\delta \leqslant 0$. У противном, неједнакост се назива *нетривијалном.*

Лема 4.3. *Ако је* corner($\boldsymbol{\beta}$) *непразан, свака нетривијална ваљана неједнакост за* corner($\boldsymbol{\beta}$) *се може записати у облику* $\sum_{j\in\boldsymbol{\eta}} \gamma_j x_j \geqslant 1$, *гдје је* $\gamma_j \geqslant 0$ *за* $j \in \boldsymbol{\eta}$.

Доказ. Пошто је свака базна промјенљива линерана комбинација небазних, свака ваљана неједнакост за *corner*($\boldsymbol{\beta}$) се може записати као $\sum_{j\in\boldsymbol{\eta}} \gamma_j x_j \geqslant \delta$, то јест само помоћу небазних промјенљивих x_j за $j \in \boldsymbol{\eta}$.

Даље, ако је $\gamma_l < 0$ за неко $l \in \boldsymbol{\eta}$, разматрајући $\overline{z}^l$ дефинисано у (4.4) слиједи да је $\sum_{j\in\boldsymbol{\eta}} \gamma_j \overline{z}^l_j = \gamma_l < 0$, па је скуп $\{\sum_{j\in\boldsymbol{\eta}} \gamma_j x_j \ :\ x \in \text{corner}((\boldsymbol{\beta}))\}$ неомеђен одоздо, јер је $\overline{z}^l$ у рецесивном конусу полиедра corner($\boldsymbol{\beta}$). Слиједи да је $\gamma_j \geqslant 0$ за све $j \in \boldsymbol{\eta}$.

Ако је $\delta \leqslant 0$, неједнакост $\sum_{j\in\boldsymbol{\eta}} \gamma_j x_j \geqslant \delta$ је тривијална, стога је $\delta > 0$, а због скалирања можемо претпоставити без губитка општости да је $\delta = 1$. □

4.1 Бала́ове одсијецајуће равни

У овом поглављу биће представљена Бала́ова парадигма [9] за конструкцију неједнакости које су ваљане за корнерски полиедар и које одсијецају базно рјешење $\overline{x}$.

За почетак, сматра се да је дат затворен конвексан скуп $C \subset \mathbb{R}^n$ такав да унутрашњост скупа C садржи тачку $\overline{x}$ (значи кодимензија скупа C је 0), а не садржи тачку из скупа $\mathbb{Z}^p \times \mathbb{R}^{n-p}$, односно да C не садржи допустиву тачку релаксације (4.3) у својој унутрашњости. За сваки од $|\boldsymbol{\eta}|$ екстремалних зрака полиедра corner($\boldsymbol{\beta}$), дефинише се

$$\alpha_j = \max\{\alpha \geqslant 0 \ :\ \overline{x} + \alpha \overline{z}^j \in C\}. \tag{4.7}$$

Пошто је $\overline{x}$ у унутрашњости скупа C, важи да је $\alpha_j > 0$. У случају да полуправа $\{\overline{x} + \alpha \overline{z}^j \ : \ \alpha \geqslant 0\}$ пресјеца обод (границу) скупа C, онда је α_j коначно, тачка $\overline{x} + \alpha_j \overline{z}^j$ припада ободу скупа C, а полуотворени интервал $\{\overline{x} + \alpha \overline{z}^j \ : \ 0 \leqslant \alpha < \alpha_j\}$ се налази у унутрашњости скупа C. Уколико пак $\overline{z}^j$ припада рецесивном конусу скупа C, онда је $\alpha_j = +\infty$ и узима се да је $\frac{1}{+\infty} = 0$. Неједнакост

$$\sum_{j \in \boldsymbol{\eta}} \frac{x_j}{\alpha_j} \geqslant 1 \tag{4.8}$$

се зове *Бала́ова одсијецајућа раван* полиедра corner($\boldsymbol{\beta}$) одређена скупом C. У литератури на енглеском језику дате равни се називају *intersection cutting planes.*

Теорема 4.4 (Бала́ [8]). *Ако је $C \subset \mathbb{R}^n$ затворен конвексан скуп у чијој се унутрашњости налази тачка $\overline{x}$ и ниједна од тачака из скупа $\mathbb{Z}^p \times \mathbb{R}^{n-p}$, онда је Балаова одсијецајућа раван (4.8) одређена скупом C ваљана неједнакост за* corner($\boldsymbol{\beta}$).

Доказ. Скуп свих тачака ЛП-релаксације $P(\boldsymbol{\beta})$ полиедра $\boldsymbol{\beta}$ које су одсјечене Балаовом одсијецајућом равни је

$$S = \{x \in \mathbb{R}^n \ : \ x_i = x_i^* - \sum_{j \in \boldsymbol{\eta}} \overline{a}_{ij} x_j \ (\text{за } i \in \boldsymbol{\beta});\ x_j \geqslant 0 \ (\text{за } j \in \boldsymbol{\eta});\ \sum_{j \in \boldsymbol{\eta}} \frac{x_j}{\alpha_j} < 1\}$$

Довољно је показати да се S налази у унутрашњости скупа C, пошто унутрашњост скупа C не садржи тачку из $\mathbb{Z}^p \times \mathbb{R}^{n-p}$.

Размотримо полиедар

$$\overline{S} = \{x \in \mathbb{R}^n \ : \ \ x_i = x_i^* - \sum_{j \in \boldsymbol{\eta}} \overline{a}_{ij} x_j \ (\text{за } i \in \boldsymbol{\beta});\ x_j \geqslant 0 \ (\text{за } j \in \boldsymbol{\eta});\ \sum_{j \in \boldsymbol{\eta}} \frac{x_j}{\alpha_j} \leqslant 1\}.$$

Према Примједби 4.1 $\overline{S}$ је $|\boldsymbol{\eta}|$-димензионални полиедар са врховима $\overline{x}$ и $\overline{x} + \alpha_j \overline{z}^j$ за α_j коначно и екстремалним зрацима $\overline{z}^j$ за $\alpha_j = +\infty$. Пошто су врхови скупа $\overline{S}$ који леже на хиперравни $\{x \in \mathbb{R}^n \ : \ \sum_{j \in \boldsymbol{\eta}} \frac{x_j}{\alpha_j} = 1\}$ тачке $\overline{x} + \alpha_j \overline{z}^j$ за α_j коначно, свака тачка из S се може изразити као конвексна комбинација тачака из сегмената $\{\overline{x} + \alpha \overline{z}^j,\ 0 \leqslant \alpha < \alpha_j\}$ (за α_j коначно) сабрана са конусном комбинацијом екстремалних зрака $\overline{z}^j$ (за $\alpha_j = +\infty$). Пошто, према дефиницији α_j, унутрашњост скупа C садржи сегменте $\{\overline{x} + \alpha \overline{z}^j,\ 0 \leqslant \alpha < \alpha_j\}$ за α_j коначно, а зраци $\overline{z}^j$ за $\alpha_j = +\infty$ припадају рецесивном конусу скупа C, па слиједи да се скуп S налази у унутрашњости скупа C. □

Примједба 4.5. Нека су C_1 и C_2 два затворена конвексна скупа чије унутрашњости садрже $\overline{x}$, али не садрже тачку из скупа $\mathbb{Z}^p \times \mathbb{R}^{n-p}$. Ако је C_1 садржан у C_2, онда је Балаова одсијецајућа раван одређена скупом C_2 јача од Балаове одсијецајуће равни одређене скупом C_1.

Затворен конвексан скуп C чија унутрашњост садржи $\overline{x}$, али не садржи тачку $\mathbb{Z}^p \times \mathbb{R}^{n-p}$ је *максималан* ако не постоји његов надскуп са истим, претходно наведеним својствима. Наравно, било који затворен конвексан скуп чија унутрашњост садржи $\overline{x}$, а не садржи тачку $\mathbb{Z}^p \times \mathbb{R}^{n-p}$ је подскуп неког максималног скупа [11].

Из својства максималности и Примједбе 4.5 слиједи да је довољно разматрати Балаове одсијецајуће равни које су одређене максималним затвореним конвексним скуповима чије унутрашњости садрже $\overline{x}$ и не садрже нити једну тачку из скупа $\mathbb{Z}^p \times \mathbb{R}^{n-p}$.

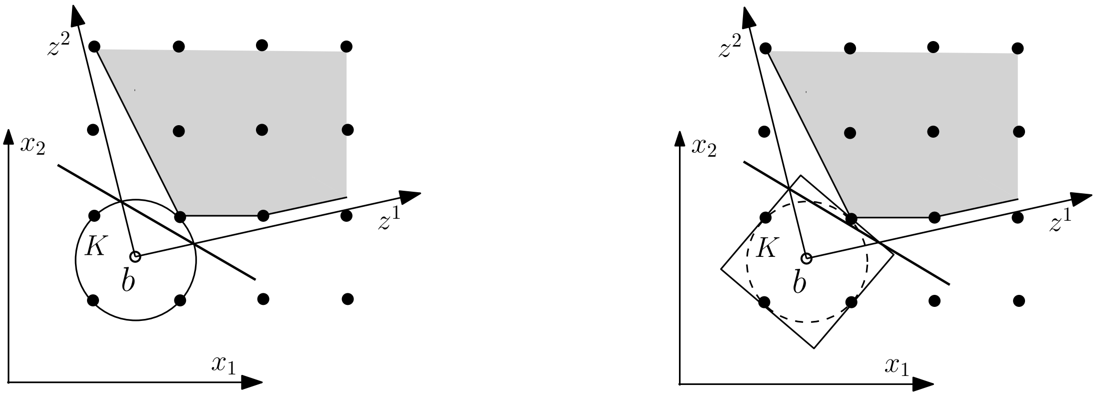
Скуп $K \subset \mathbb{R}^p$ који не садржи нити једну тачку из скупа $\mathbb{Z}^p$ у својој унутрашњости назива се *$\mathbb{Z}^p$-слободним.*

Примједба 4.6. Један начин да се конструише затворен конвексан скуп C у чијој се унутрашњости налази $\overline{x}$, а који при том не садржи нити једну тачку скупа $\mathbb{Z}^p \times \mathbb{R}^{n-p}$, је да се у простору $\mathbb{R}^p$ прво конструише $\mathbb{Z}^p$-слободан затворен конвексан скуп K чија унутрашњост садржи ортогоналну пројекцију $\overline{x}$ на $\mathbb{R}^p$ и тада је тражени скуп цилиндар $C = K \times \mathbb{R}^{n-p}$.

Примјер 4.2. Дат је мјештовито-цјелобројан скуп

$$\begin{aligned} x_1 &= b_1 + a_{11}y_1 + a_{12}y_2 \\ x_2 &= b_2 + a_{21}y_1 + a_{22}y_2 \\ x &\in \mathbb{Z}^2 \\ y &\geqslant 0 \end{aligned} \tag{4.9}$$

код кога зраци $z^1 = \binom{a_{11}}{a_{21}}$, $z^2 = \binom{a_{12}}{a_{22}} \in \mathbb{R}^2$ нису колинеарни и $b = \binom{b_1}{b_2} \notin \mathbb{Z}^2$.



Слика 4.2 Балаове одсијецајуће равни

На Слици 4.2 је проблем (4.9) у представљен у простору промјенљивих x_1 и x_2. Скуп допустивих тачака $x \in \mathbb{R}^2$ за ЛП-релаксацију проблема је конус са тјеменом b и ектремалним зрацима z^1 и z^2. Црне тачке унутар конуса представљају допустиве тачке $x \in \mathbb{Z}^2$. Осијенчена област представља корнерски полиедар у (x_1, x_2)-простору. Слика приказује два примјера $\mathbb{Z}^2$-слободна конвексна скупа $K \subset \mathbb{R}^2$ која садрже b у својој унутрашњости.

Пошто у овом примјеру имају двије небазне промјенљиве, Балаова одсијецајућа раван се може представити правом у простору базних промјељивих, наиме, правом која пролази кроз тачке p^1 и p^2 као пресјеке обода скупа K и полуправих $\{b + \alpha z^1 : \alpha \geqslant 0\}$, $\{b + \alpha z^2 : \alpha \geqslant 0\}$.

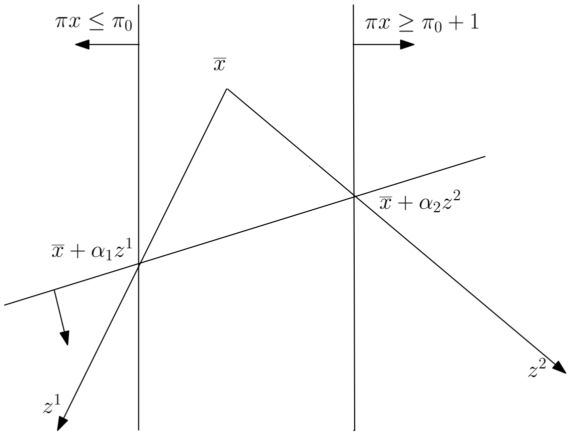
Коефицијенти α_1 и α_2 који одређују одсијецајућу раван $\frac{y_1}{\alpha_1} + \frac{y_2}{\alpha_2} \geqslant 1$ су добијени као $\alpha_j = \frac{\|p^j - b\|}{\|z^j\|}$, $j = 1, 2$. Примјетимо да је одсијецајућа раван на десној страни јача од одсијецајуће равни на лијевој страни, пошто $\mathbb{Z}^2$-слободан скуп с десна садржи $\mathbb{Z}^2$-слободан скуп с лијева (коефицијенти $\frac{1}{\alpha_j}$ су мањи). ■

Примјер 4.3 (Балаове одсијецајуће равни одређене расцјепом)**.** За дате $\pi \in \mathbb{Z}^p$ и $\pi_0 \in \mathbb{Z}$, нека је $K = \{x \in \mathbb{R}^p \ : \ \pi_0 \leqslant \pi x \leqslant \pi_0 + 1\}$. K је $\mathbb{Z}^p$-слободан конвексан скуп, пошто за свако $\overline{x} \in \mathbb{Z}^p$ важи да је или $\pi\overline{x} \leqslant \pi_0$ или $\pi\overline{x} \geqslant \pi_0 + 1$. Уколико су елементи вектора π узајамно прости, обе хиперравни $\{x \in \mathbb{R}^p \ : \ \pi x = \pi_0\}$ и $\{x \in \mathbb{R}^p \ : \ \pi x = \pi_0 + 1\}$ садрже тачке из $\mathbb{Z}^p$. Слиједи да је K максималан $\mathbb{Z}^p$-слободан затворен конвексан скуп. Према Примједби 4.6 цилиндар $C = K \times \mathbb{R}^{n-p} = \{x \in \mathbb{R}^n \ : \ \pi_0 \leqslant \sum_{j=1}^{p} \pi_j x_j \leqslant \pi_0 + 1\}$ је максималан $\mathbb{Z}^p \times \mathbb{R}^{n-p}$-слободан затворен конвексан скуп, који се назива *расцјеп* (енгл. *split*).

За дати корнерски полиедар corner($\boldsymbol{\beta}$), нека је $\overline{x}$ јединствен врх своје ЛП-релаксације $P(\boldsymbol{\beta})$. Ако $\overline{x}_j \notin \mathbb{Z}$ за неко $j \in \{1, \ldots, p\}$, при чему постоје π, π_0 такви да је $\pi_0 < \sum_{j=1}^{p} \pi_j \overline{x}_j < \pi_0 + 1$, онда $\overline{x}$ припада унутрашњости расцјепа C. Уводећи $\epsilon = \pi\overline{x} - \pi_0$, важи да је $0 < \epsilon < 1$, пошто је $\pi_0 < \pi\overline{x} < \pi_0 + 1$. Примјењујући формулу (4.7) на C, за $j \in \boldsymbol{\eta}$ добија се формула за скаларе

$$\alpha_j = \begin{cases} -\frac{\epsilon}{\pi\overline{z}^j} & \pi\overline{z}^j < 0 \\ \frac{1-\epsilon}{\pi\overline{z}^j} & \pi\overline{z}^j > 0 \\ +\infty & \text{иначе,} \end{cases} \tag{4.10}$$

гдје је $\overline{z}^j$ дефинисано у (4.4).



Слика 4.3 Балаова одсијецајућа раван одређена расцјепом

Скалари α_j имају исти смисао и улогу као и код (4.7) и (4.8), то јест одсијецајућа раван

$$\sum_{j \in \boldsymbol{\eta}} \frac{x_j}{\alpha_j} \geqslant 1. \tag{4.11}$$

је Балаова одсијецајућа раван одређена расцјепом C (види Сл. 4.3). ■

Примјер 4.4 (Гоморијеве мјешовито-цјелобројне одсијецајуће равни)**.** Треба да се покаже да су Гоморијеве мјешовито-цјелобројне одсијецајуће равни произведене из редова симплексне табеле (4.2) у ствари Балаове одсијецајуће равни одређене расцјепима.

Нека је дат корнерски полиедар corner($\boldsymbol{\beta}$) описан системом (4.3), при чему x_i^* није цио број. Дефинишимо $r_0 = x_i^* - \lfloor x_i^* \rfloor$, $r_j = \overline{a}_{ij} - \lfloor \overline{a}_{ij} \rfloor$, $\pi_0 = \lfloor x_i^* \rfloor$, а за $j = 1, \ldots, p$,

$$\pi_j = \begin{cases} \lfloor \overline{a}_{ij} \rfloor & j \in \boldsymbol{\eta} \text{ и } r_j \leqslant r_0, \\ \lceil \overline{a}_{ij} \rceil & j \in \boldsymbol{\eta} \text{ и } r_j > r_0, \\ 1 & j = i, \\ 0 & \text{иначе.} \end{cases} \tag{4.12}$$

Уколико је $j = p = 1, \ldots, n$ узимамо да је $\pi_j = 0$. Запазимо да је $\pi_0 < \pi\overline{x} < \pi_0 + 1$.

Сада треба, као у претходном примјеру, произвести Балаову одсијецајућу раван одређену расцјепом $C = \{x \in \mathbb{R}^n \ : \ \pi_0 \leqslant \pi x \leqslant \pi_0 + 1\}$. Израчунавајући α_j помоћу формуле (4.10) за $j \in \boldsymbol{\eta}$, добија се да је

$$\epsilon = \pi\overline{x} - \pi_0 = \sum_{i \in \boldsymbol{\beta}} \pi_i \overline{x}_i - \pi_0 = \overline{x}_i - \lfloor \overline{x}_i \rfloor = r_0.$$

Из (4.4) и (4.12) слиједи да је $\pi\overline{r}^j = \pi_j \overline{z}_j^j + \pi_i \overline{z}_i^j$, пошто је $\overline{z}_h^j = 0$ за све $h \in \boldsymbol{\eta} \backslash \{j\}$ и $\pi_h = 0$ за све $h \in \boldsymbol{\beta} \backslash \{i\}$. Вриједи

$$\pi\overline{r}^j = \begin{cases} \lfloor \overline{a}_{ij} \rfloor - \overline{a}_{ij} = -r_j & 1 \leqslant j \leqslant p \text{ и } r_j \leqslant r_0 \\ \lceil \overline{a}_{ij} \rceil - \overline{a}_{ij} = 1 - r_j & 1 \leqslant j \leqslant p \text{ и } r_j > r_0, \\ -\overline{a}_{ij} & j \geqslant p + 1. \end{cases}$$

Скалар α_j се добија из формуле (4.10), па је одсијецајућа раван (4.11) одређена расцјепом C

$$\begin{aligned} &\sum_{\substack{j \in \boldsymbol{\eta}, \\ j \leqslant p, \\ r_j \leqslant r_0}} \frac{r_j}{r_0} x_j + \sum_{\substack{j \in \boldsymbol{\eta}, \\ j \leqslant p, \\ r_j > r_0}} \frac{1 - r_j}{1 - r_0} x_j + \\ &\qquad \sum_{\substack{p+1 \leqslant j \leqslant n, \\ \overline{a}_{ij} > 0}} \frac{\overline{a}_{ij}}{r_0} x_j - \sum_{\substack{p+1 \leqslant j \leqslant n, \\ \overline{a}_{ij} < 0}} \frac{\overline{a}_{ij}}{1 - r_0} x_j \geqslant 1, \end{aligned} \tag{4.13}$$

а то је управо Гоморијева мјешовито-цјелобројна одсијецајућа раван. ■

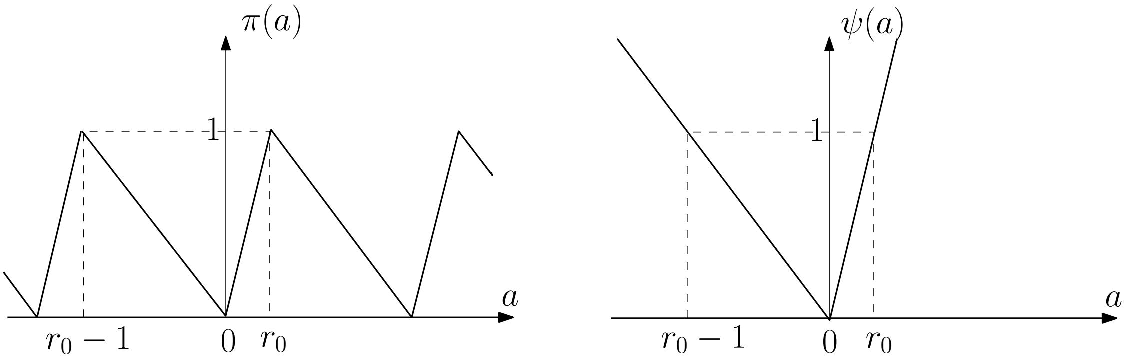
Пошто је запис Гоморијеве формуле доста компликован, згодније је о њој размишљати као о неједнакости облика

$$\sum_{j=1}^{p} \pi(\overline{a}_{ij}) x_j + \sum_{j=p+1}^{n} \psi(\overline{a}_{ij}) x_j \geqslant 1,$$

гдје су π и ψ функције придружене цјелобројним и нецјелобројним промјенљивим респективно, а дефинисане као (погледати Сл. 4.1)

$$\begin{aligned} \pi(a) &= \min\{\frac{r}{r_0}, \frac{1 - r}{1 - r_0}\} \quad \text{при чему је } r = a - \lfloor a \rfloor \\ \psi(a) &= \max\{\frac{a}{r_0}, \frac{-a}{1 - r_0}\}. \end{aligned} \tag{4.14}$$

Теорема 4.4 казује да су Балаове одсијецајуће равни ваљане за corner($\boldsymbol{\beta}$). Сљедећа теорема даје пак обратну тврдњу (претпоставља се да је corner($\boldsymbol{\beta}$) непразан).



Слика 4.4 Гоморијеве функције π и ψ

Теорема 4.7. *Свака нетривијална ваљана неједнакост за полиедар* $\mathrm{corner}(\boldsymbol{\beta})$ *је Балаова одсијецајућа раван.*

Доказ. Овдје ће теорема бити доказана само за потпуно цјелобројни случај (видјети [15] за општији случај). Према Леми 4.3, нетривијална ваљана неједнакост за $\mathrm{corner}(\boldsymbol{\beta})$ је облика $\sum_{j\in\boldsymbol{\eta}}\gamma_j x_j \geqslant 1$, при чему је $\gamma_j \geqslant 0,\ j \in \boldsymbol{\eta}$. Треба показати да је она Балаова одсијецајућа раван.

Размотримо полиедар $S = P(\boldsymbol{\beta}) \cap \{x \in \mathbb{R}^n \ :\ \sum_{j\in\boldsymbol{\eta}}\gamma_j x_j \leqslant 1\}$. Пошто је $\sum_{j\in\boldsymbol{\eta}}\gamma_j x_j \geqslant 1$ ваљана неједнакост за $corner(\boldsymbol{\beta})$, све тачке у $\mathbb{Z}^p \cap S$ задовољавају једнакост $\sum_{j\in\boldsymbol{\eta}}\gamma_j x_j = 1$.

$P(\boldsymbol{\beta})$ је рационалан полиедар, па постоје цјелобројна матрица A' и цјелобројни вектор b' тако да је $P(\boldsymbol{\beta}) = \{x \in \mathbb{R}^n \ :\ A'x \leqslant b'\}$. Нека је

$$T = \{x \in \mathbb{R}^n \ :\ A'x \leqslant b'+1,\ \sum_{j\in\boldsymbol{\eta}}\gamma_j x_j \leqslant 1\}.$$

Сад треба за T доказати да је $\mathbb{Z}^p$-слободан конвексан скуп. Претпоставимо да унутрашњост T садржи цјелобројну тачку $\tilde{x}$, то јест $\tilde{x}$ задовољава све неједнакости које дефинишу T строго. Како је $A'x \leqslant b'+1$ цјелобројни систем, слиједи да је $A'\tilde{x} \leqslant b'$ и $\sum_{j\in\boldsymbol{\eta}}\gamma_j\tilde{x}_j < 1$. Ово противрјечи чињеници да све тачке $\mathbb{Z}^p \cap S$ задовољавају $\sum_{j\in\boldsymbol{\eta}}\gamma_j x_j = 1$.

Пошто $\overline{x}$ припада S и $\sum_{j\in\boldsymbol{\eta}}\gamma_j\overline{x}_j = 0$, T је $\mathbb{Z}^p$-слободан затворен конвексан скуп који садржи $\overline{x}$ у својој унутрашњости. Балаова одсијецаућа раван одређена полиедром T је $\sum_{j\in\boldsymbol{\eta}}\gamma_j x_j \geqslant 1$. □

4.2 Максимални $\mathbb{Z}^p$-слободни конвексни скупови

Према Примједби 4.5, набоље одсијецајуће равни су одређене *максималним* $\mathbb{Z}^p \times \mathbb{R}^{n-p}$*-слободним затвореним конвексним скуповима* кодимензије 0.

Лема 4.8. *Ако је* C *максималан* $\mathbb{Z}^p \times \mathbb{R}^{n-p}$*-слободан затворен конвексан скуп и ако је* K *његова пројекција на* $\mathbb{R}^p$*, онда је* K *максималан* $\mathbb{Z}^p$*-слободан затворен конвексан скуп и* $C = K \times \mathbb{R}^{n-p}$.

Доказ. Нека је K' пројекција скупа C на $\mathbb{R}^p$. Онда је C садржан у скупу $K' + (\{0\}^p \times \mathbb{R}^{n-p})$. Пошто је C $\mathbb{Z}^p \times \mathbb{R}^{n-p}$-слободан затворен конвексан скуп, K' је $\mathbb{Z}^p$-слободан затворен конвексан скуп. Нека је K максималан $\mathbb{Z}^p$-слободан затворен конвексан скуп који садржи K'. $K \times \mathbb{R}^{n-p}$ је тада $\mathbb{Z}^p \times \mathbb{R}^{n-p}$-слободан затворен конвексан скуп и $C \subseteq K' \times \mathbb{R}^{n-p} \subseteq K \times \mathbb{R}^{n-p}$. Пошто је C максималан, ова три скупа се поклапају одакле слиједи тврдња. $\square$

Претходна теорема казује да је довољно разматрати максималне $\mathbb{Z}^p$-слободне затворене конвексне скупове. Ласло Ловас је додатно окарактерисао дате скупове у сљедећој теореми (наводимо је без доказа).

Теорема 4.9 (Ловаш [49]). *Скуп* $K \subset \mathbb{R}^p$ *је* $\mathbb{Z}^p$*-слободан затворен конвексан скуп кодимензије* 0 *ако и само ако је* K *полиедар облика* $K = P + L$ *гдје је* P *неки политоп,* L *рационалан линеаран простор,* $\dim(K) = \dim(P) + \dim(L) = p$, K *не садржи нити једну тачку скупа* $\mathbb{Z}^p$ *у својој унутрашњости и постоји тачка скупа* $\mathbb{Z}^p$ *у релативној унутрашњости сваке од страна полиедра* K.

С друге стране, Скарф је дате скупове на другачији начин окарактерисао, односно доказао је да важи наредна тврдња.

Теорема 4.10 (Скарф [54]). *Сваки* $K \subset \mathbb{R}^p$ *максималан* $\mathbb{Z}^p$*-слободан затворен конвексан скуп чија је кодимензија* 0 *има највише* 2^p *страна.*

Доказ. Према претходној теореми, свака страна S садржи цјелобројну тачку x^S у својој релативној унутрашњости. Ако има више од 2^p страна, онда постоје двије различите стране S и S', такве да су x^S и $x^{S'}$ конгруентне по модулу 2. Њихова аритметичка средина $\frac{1}{2}(x^S + x^{S'})$ је цјелобројна и налази се у унутрашњости скупа K, што противрјечи чињеници да је K $\mathbb{Z}^p$-слободан. $\square$

У $\mathbb{R}^2$, ова теорема имплицира да максимални $\mathbb{Z}^2$-слободни затворени конвексни скупови чија је кодимензија 0 могу имати највише 4 стране. Кориштењем Теореме 4.9, може се показати да су то сљедећи скупови (Сл. 4.5):

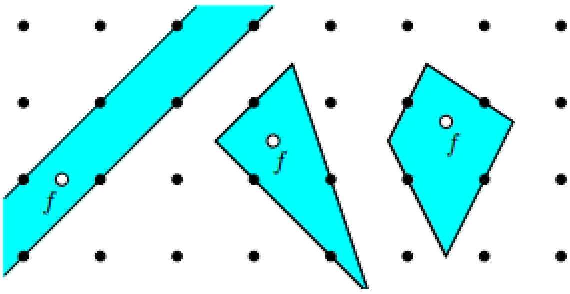
- расцјепи — скупови облика $\{x \in \mathbb{R}^2 \ : \ \pi_0 \leqslant \pi_1 x_1 + \pi_2 x_2 \leqslant \pi_0 + 1\}$, гдје су $\pi_0, \pi_1, \pi_2 \in \mathbb{Z}$ и π_1, π_2 узајамно прости;
- троуглови са цјелобројном тачком у релативној унутрашњости сваке од страна и нецјелобројним тачкама у унутрашњости троугла;
- четвороуглови са цјелобројним тачкама у релативној унутрашњости сваке од страна и без цјелобројне тачке у унутрашњости четвороугла.

Андерсен *et al.* [5] је доказао да за случај када је $|\boldsymbol{\beta}| = p = 2$, Балаове одсијецајуће равни одређене расцјепима, троугловима и четвороугловима у потпуности описују $\operatorname{corner}(\boldsymbol{\beta})$.

Корнуижол и Маргот [21] су окарактерисали тачно какви расцјепи, троуглови и четвороуглови производе Балаове одсијецајуће равни које су стране $\operatorname{corner}(\boldsymbol{\beta})$, опет за случај $|\boldsymbol{\beta}| = p = 2$.

Аверков *et al.* је у [7] доказао да за фиксирану димензију, у оквиру унимодуларних трансформација, постоји коначан број максималних полиедара (гледано у односу на инклузију) међу цјелобројним $\mathbb{Z}^p$-слободним полиедрима.

Скарф [55] је доказао да у $\mathbb{R}^3$, максимални $\mathbb{Z}^3$-слободни затворени конвексни скупови са тачно једном цјелобројном тачком у релативној унутрашњости сваке од



Слика 4.5 Максимални $\mathbb{Z}^2$-слободни затворени конвексни скупови са непразном унутрашњошћу у $\mathbb{R}^2$

страна, имају својство да све дате цјелобројне тачке леже на двије сусједне паралелне хиперравни, које су саме одређене цјелобројним тачкама (у суштини, ради се о другачијој формулацији необјављене теореме Хоуа).

Према Примједби 4.5 најјаче Балаове одсијецајуће равни су одређене максималним $\mathbb{Z}^p \times \mathbb{R}^{n-p}$-слободним затвореним конвексним скуповима који садрже тачку $\overline{x}$ у својој унутрашњости. На основу Леме 4.8 и Теореме 4.9, ови скупови су полиедри типа $K \times \mathbb{R}^{n-p}$, гдје је K $\mathbb{Z}^p$-слободан полиедар у $\mathbb{R}^p$. Сљедећа теорема показује како се рачунају коефицијенти Балаових одсијецајућих равни из описа страна полиедра K.

Теорема 4.11. *Нека је* $K = \{x \in \mathbb{R}^p \ : \ \sum_{h=1}^{p} d_h^i(x_h - \overline{x}_h) \leqslant 1, i = 1, \ldots, t\}$ $\mathbb{Z}^p$*-слободан полиедар. Коефицијенти* α_j*,* $(j \in \boldsymbol{\eta})$ *Балаове одсијецајућу равни (4.8) одређене полиедром* $K \times \mathbb{R}^{n-p}$*, дати су са*

$$\frac{1}{\alpha_j} = \max_{i=1,\ldots,t} \sum_{h=1}^{p} d_h^i \overline{z}_h^j.$$

Доказ. Пошто је $\alpha_j = \max\{\alpha \geqslant 0 \ : \ \overline{x} + \alpha\overline{z}^j \in K \times \mathbb{R}^{n-p}\}$ и $K \times \mathbb{R}^{n-p} = \{x \in \mathbb{R}^n \ : \ \sum_{h=1}^{p} d_h^i(x_h - \overline{x}_h) \leqslant 1, \ i = 1, \ldots, t\}$, онда је $\frac{1}{\alpha_j} = \max\{0, \sum_{h=1}^{p} d_h^i \overline{z}_h^j \ (i = 1, \ldots, t)\}$.

Пошто је $K \times \mathbb{R}^{n-p}$ садржано у максималном $\mathbb{Z}^p \times \mathbb{R}^{n-p}$-слободном затвореном конвексном скупу, према Теореми 4.9, рецесивни конус политопа K има димензију мању од p, па самим тим има празну унутрашњост. Стога, систем строгих неједначина $\sum_{h=1}^{p} d_h^i z_h < 0, \ i = 1, \ldots, t$ нема рјешења, одакле слиједи

$$\max_{i=1,\ldots,t} \sum_{h=1}^{p} d_h^i \overline{z}_h^j \geqslant 0.$$

□

Нека $z^j \in \mathbb{R}^p$ означава рестрикцију $\overline{z}^j \in \mathbb{R}^n$ на првих p компонената. Претходна теорема каже да су Балаове одсијецајуће равни облика $\sum_{j \in \boldsymbol{\eta}} \psi(z^j) x_j \geqslant 1$, гдје је $\psi : \mathbb{R}^p \to \mathbb{R}_+$ дефинисана као

$$\psi(z) = \max_{i=1,\ldots,t} \sum_{h=1}^{p} d_h^i z_h. \tag{4.15}$$

За фиксни позитивни цијели број p и x_i^* ($i \in \boldsymbol{\beta}$) из (4.3), дефинише се *ваљана функција* као она функција $\psi : \mathbb{R}^p \to \mathbb{R}_+$ за коју важи да је неједнакост $\sum_{j\in\boldsymbol{\eta}} \psi(z^j)x_j \geqslant 1$ је ваљана за corner($\boldsymbol{\beta}$) за било који број непрекидних промјенљивих и било који избор $\overline{a}_{ij}$ ($i \in \boldsymbol{\beta}, j \in \boldsymbol{\eta}$), гдје је z^j поменута рестрикција вектора $\overline{z}^j$ дефинисаног као у (4.4) за $i = 1, \ldots, p$. Ваљана функција ψ је *минимална* ако не постоји ψ' различита од ψ таква да је $\psi' \leqslant \psi$.

Пошто су најјаче Балаове одсијецајуће равни одређене максималним $\mathbb{Z}^p \times \mathbb{R}^{n-p}$-слободним затвореним конвексним скуповима који садрже $\overline{x}$ у својој унутрашњости, минималне ваљане функције су облика (4.15).

Функција $\psi : \mathbb{R}^p \to \mathbb{R}$ је *дио-по-дио линеарна* ако се $\mathbb{R}^p$ може партиицонисати на коначан број полиедарских области тако да је рестрикција ψ на унутрашњост сваке од тих области афина функција.

Посљедица 4.12. *Свака минимална ваљана функција је позитивно хомогена, субадитивна и дио-по-дио линеарна.*

4.3 Бесконачна релаксација

Теорема 4.11 даје формулу за израчунавање коефицијената Балаове одсијецајуће равни, то јест $\frac{1}{\alpha} = \psi(z^j)$, гдје је ψ функција дефинисана у (4.15). Можемо примијетити да функција ψ не зависи од броја небазних промјенљивих, нити од вектора z^j. Свака функција са тим својствима се може искористити као „црна кутија" за произвођење одсијецајућих равни из табеле билоког цјелобројног програма. Прича у наредних пар поглавља се односи на боље разумијевање ових функција. Гомори и Џонсон [35] су увели одговарајући концепт за проучавање ових функција, који ћемо представити у наставку.

Проблем (4.3) када су све промјенљиве x_j цјелобројне за $j \in \boldsymbol{\eta}$ може се представити као

$$\begin{aligned} f_i + \sum_{j\in\boldsymbol{\eta}} r_i^j x_j \in \mathbb{Z} \qquad & i = 1, \ldots, q \\ x_j \in \mathbb{Z}_+ \qquad & j \in \eta. \end{aligned} \tag{4.16}$$

Гомори и Џонсон су предложили релаксацију простора промјенљивих x_j, ($j \in \boldsymbol{\eta}$) до бесконачно-димензионалног простора, гдје су промјенљиве x_r дефинисане за било које $r \in \mathbb{R}^q$ — такозване *бесконачне релаксације*

$$\begin{aligned} f + \sum_{r\in\mathbb{R}^q} r x_r \in \mathbb{Z}^q & \\ x_r \in \mathbb{Z}_+ \quad & \text{за све } r \in \mathbb{R}^q. \end{aligned} \tag{4.17}$$

при чему се узима у обзир да је $x_r > 0$ за коначан број $r \in \mathbb{R}^q$. Природно, (4.16) се може добити из (4.17) тако што ће се све промјенљиве, осим њих коначно много, изједначити са 0. Мотивација за рад са G_f је што постоји само један параметар, то јест f, па су стога резултати чистији. Још једна специфишност G_f је да није затворен. На примјер, за низ x^k ($k = 1, 2, \ldots$), дефинисан тако да је

$$x_r^k = \begin{cases} \frac{1}{k} & \text{ако } r = -kf \\ 0 & \text{иначе} \end{cases},$$

припада G_f и конвергира ка 0, али $0 \notin G_f$, јер $f \in \mathbb{Q}^q \backslash \mathbb{Z}^q$.

Скуп свих допустивих тачака за (4.17) биће означаван са G_f. Лако се увиђа да је $G_f \neq \infty$, пошто је, на примјер, $x_r = 1$ за $r = -f$ и $x_r = 0$ у осталим случајевима, допустиво рјешење за (4.17). Функција $\pi : \mathbb{R}^q \to \mathbb{R}$ је *ваљана* ако је $\pi \geqslant 0$ и ако линеарна неједнакост

$$\sum_{r \in \mathbb{R}^q} \pi(r) x_r \geqslant 1 \tag{4.18}$$

вриједи за све тачке из G_f.

Релевантност претходне дефиниције се огледа у чињеници да било која ваљана функција π даје, рестрикцијом неједнакости (4.18) на простор r^j ($j \in \boldsymbol{\eta}$), ваљану неједнакост за почетни (цјелобројни) полиедар (4.16).

Претпоставка о ненегативности у дефиницији ваљане функције није вјештачка, као што то може да изгледа. Уколико би се уклонила та претпоставка ваљана функција би могла узимати негативне вриједности, но свака ваљана функција требала би, ипак, да буде ненегативна над рационалним векторима. Ради појашњења, претпоставимо да важи супротно, то јест да је π таква функција да важи (4.18) за свако x из G_f и $\pi(\tilde{r}) < 0$ за неко $\tilde{r} \in \mathbb{Q}^q$. Постоји $D \in \mathbb{Z}_+$ такво да је $D\tilde{r}$ цјелобројан вектор и нека је $\overline{x}$ из G_f (на примјер $\overline{x}_r = 1$ за $r = -f$ и $\overline{x}_r = 0$ у осталим случајевима). Нека је $\tilde{x}$ дефинисано као $\tilde{x} = \overline{x}_{\tilde{r}} + MD$ гдје је M позитиван цио број и $\tilde{x} = \overline{x}_r$ за $r \neq \tilde{r}$. Слиједи да је $\tilde{x}$ припада G_f, при чему је

$$\sum_{r \in \mathbb{R}^q} \pi(r) \tilde{x}_r = \sum_{r \in \mathbb{R}^q} \pi(r) \overline{x}_r + \pi(\tilde{r}) MD.$$

Број M се може изабрати довољно великим, односно

$$M > \frac{\sum_{r \in \mathbb{R}^q} \pi(r) \overline{x}_r - 1}{D|\pi(\tilde{r})|},$$

па да важи $\sum \pi(r) \tilde{x}_r < 1$, што противријечи тврдњи да је $\tilde{x}$ допустиво.

Пошто су подаци у мјешовито-цјелобројним програмима рационални, а ваљане функције морају бити ненегативне над рационалним векторима, природно је да се пажња усредсреди на ненегативне ваљане функције.

Ваљана функција $\pi : \mathbb{R}^q \to \mathbb{R}_+$ је *минимална* ако не постоји ваљана функција $\pi' \neq \pi$ таква да је $\pi'(r) \leqslant \pi(r)$ за све $r \in \mathbb{R}^q$.

Функција $\pi : \mathbb{R}^q \to \mathbb{R}_+$ је *периодична* ако је $\pi(r) = \pi(r + w)$ за свако $w \in \mathbb{Z}^q$. Стога, дате периодичне функције су у потпуности дефинисане на основу вриједности из интервала $[0, 1)^q$.

Лема 4.13. *Ако је π минимална ваљана функција, онда је π периодична и $\pi(0) = 0$.*

Доказ. Претпоставимо да π није периодична. Тада је $\pi(\tilde{r}) > \pi(\tilde{r} + w)$ за неко $\tilde{r} \in \mathbb{R}^q$ и $w \in \mathbb{Z}^q$. Дефинишимо функцију π' на сљедећи начин:

$$\pi'(r) = \begin{cases} \pi(\tilde{r} + w) & \text{ако } r = \tilde{r} \\ \pi(r) & \text{ако } r \neq \tilde{r} \end{cases}.$$

Показаћемо да је π' ваљана. Узмимо неко $\overline{x} \in G_f$. Нека је $\tilde{x}$ дефинисано на сљедећи начин:

$$\tilde{x}_r = \begin{cases} \overline{x}_r & \text{ако } r \neq \tilde{r}, \tilde{r}+w \\ 0 & \text{ако } r = \tilde{r} \\ \overline{x}_{\tilde{r}} + \overline{x}_{\tilde{r}+w} & \text{ако } r = \tilde{r}+w \end{cases}.$$

Пшто је $\overline{x} \in G_f$ и $w\overline{x}_{\tilde{r}+w} \in \mathbb{Z}^q$, важи да је $\tilde{x} \in G_f$. Штавише $\sum_{r\in\mathbb{R}^q} \pi'(r)\overline{x}_r = \sum_{r\in\mathbb{R}^q} \pi(r)\tilde{x} \geqslant 1$. Овим смо доказали да је π периодична.

Ако је $\overline{x} \in G_f$, онда и $\tilde{x}$ припада G_f, узимајући да је $\tilde{x}_r = \overline{x}_r$ за $r \neq 0$ и $\tilde{x}_0 = 0$. Стога, ако је π ваљана, онда је π' дефинисана као $\pi'(r) = \pi(r)$ за $r \neq 0$ и $\pi'(0) = 0$ такође ваљана. Пошто је π минимална и ненегативна, слиједи да је $\pi(0) = 0$. □

Лема 4.14. *Ако је π минимална ваљана функција, онда је π субадитивна.*

Доказ. Нека су $r^1, r^1 \in \mathbb{R}^q$. Треба показати да је $\pi(r^1) + \pi(r^2) \geqslant \pi(r^1 + r^2)$. Дефинишимо функцију π' на сљедећи начин.

$$\pi'(r) = \begin{cases} \pi(r^1) + \pi(r^2) & \text{ако } r = r^1 + r^2 \\ \pi(r) & \text{иначе} \end{cases}.$$

Показаћемо да је π' ваљана. Узмимо неко $\overline{x} \in G_f$. Дефинишимо $\tilde{x}$

$$\tilde{x}_r = \begin{cases} \overline{x}_{r^1} + \overline{x}_{r^1+r^2} & \text{ако } r = r^1 \\ \overline{x}_{r^2} + \overline{x}_{r^1+r^2} & \text{ако } r = r^2 \\ 0 & \text{ако } r = r^1 + r^2 \\ \overline{x}_r & \text{иначе} \end{cases}.$$

Користећи дефиниције π' и $\tilde{x}$, лако је провјерити да вриједи

$$\sum_{r\in\mathbb{R}^q} \pi'(r)\overline{x}_r = \sum_{r\in\mathbb{R}^q} \pi(r)\tilde{x}_r. \tag{4.19}$$

Штавише, вриједи да $f + \sum_{r\in\mathbb{R}^q} r\overline{x}_r = f + \sum_{r\in\mathbb{R}^q} r\tilde{x}_r \in \mathbb{Z}^q$. Пошто је $\tilde{x} \geqslant 0$, слиједи да је $\tilde{x} \in G_f$.

Пошто је π ваљана, слиједи да је $\sum_{r\in\mathbb{R}^q} \pi(r)\tilde{x}_r \geqslant 1$, па је стога, према (4.19), $\sum_{r\in\mathbb{R}^q} \pi'(r)\overline{x}_r \geqslant 1$. Слиједи да је π' ваљана, а пошто је π минимална, добијамо да је $\pi(r^1 + r^2) \leqslant \pi'(r^1 + r^2) = \pi(r^1) + \pi(r^2)$. □

Иначе, за сваку минималну ваљану функцију π важи да је $\pi(r) \leqslant 1$ за све $r \in \mathbb{R}^q$. Узимајући у разматрање допустиву тачку $x_{-f} = 1$ и $x_r = 0$ $(r \in \mathbb{R}^q \backslash \{-f\})$, добија се да је $\pi(-f) \geqslant 1$, па на крају слиједи да је $\pi(-f) = 1$ за сваку минималну ваљану функцију π.

Функција $\pi : \mathbb{R}^q \to \mathbb{R}_+$ задовољава *услов симетричности* ако $\pi(r) + \pi(-f-r) = 1$ за све $r \in \mathbb{R}^q$.

Лема 4.15. *Ако је π минимална ваљана функција, онда π задовољава услов симетричности.*

Доказ. Претпоставимо да постоји $\tilde{r} \in \mathbb{R}^q$ такво да $\pi(\tilde{r}) + \pi(-f - \tilde{r}) \neq 1$. Пошто је π ваљана, $\pi(\tilde{r}) + \pi(-f - \tilde{r}) = 1 + \delta$, при чему је $\delta > 0$. Пошто је $\pi(r) \leqslant 1$ за све $r \in \mathbb{R}^q$, слиједи да је $\pi(\tilde{r}) > 0$.

Дефинишимо функцију π' на сљедећи начин:

$$\pi'(r) = \begin{cases} \frac{1}{1+\delta}\pi(\tilde{r}) & \text{ако } r = \tilde{r} \\ \pi(r) & \text{иначе} \end{cases},$$

гдје $r \in \mathbb{R}^q$. Показаћемо да је π' ваљана. Размотримо неко $\overline{x} \in G_f$ и запазимо да вриједи

$$\sum_{r \in \mathbb{R}^q} \pi'(r)\overline{x}_r = \sum_{\substack{r \in \mathbb{R}^q,\\ r \neq \tilde{r}}} \pi(r)\overline{x}_r + \frac{1}{1+\delta}\pi(\tilde{r})\overline{x}_{\tilde{r}}.$$

Ако је $\overline{x}_{\tilde{r}} = 0$ онда вриједи $\sum_{r \in \mathbb{R}^q} \pi'(r)\overline{x}_r = \sum_{r \in \mathbb{R}^q} \pi(r)\overline{x}_r \geqslant 1$, пошто је π ваљана. Даље, ако је $\overline{x}_{\tilde{r}} \geqslant (1+\delta)/\pi(\tilde{r})$, онда вриједи $\sum_{r \in \mathbb{R}^q} \pi'(r)\overline{x}_r \geqslant 1$. Стога, можемо претпоставити да је $1 \leqslant \overline{x}_{\tilde{r}} < (1+\delta)/\pi(\tilde{r})$. Слиједи да је

$$\begin{aligned} \sum_{r \in \mathbb{R}^q} \pi'(r)\overline{x}_r &= \sum_{\substack{r \in \mathbb{R}^q,\\ r \neq \tilde{r}}} \pi(r)\overline{x}_r + \pi(\tilde{r})(\overline{x}_{\tilde{r}} - 1) + \pi(\tilde{r}) - \frac{\delta}{1+\delta}\pi(\tilde{r})\overline{x}_{\tilde{r}} \\ &\geqslant \pi(-f - \tilde{r}) + \pi(\tilde{r}) - \delta \\ &\geqslant 1 + \delta - \delta = 1, \end{aligned}$$

гдје неједнакост $\sum_{r \in \mathbb{R}^q, r \neq \tilde{r}} \pi(r)\overline{x}_r + \pi(\tilde{r})(\overline{x}_{\tilde{r}} - 1) \geqslant \pi(-f - \tilde{r})$ слиједи из субадитивности функције π (претходна лема). Стога, π' је ваљана, што противријечи минималности функције π. □

Теорема 4.16 (Гомори и Џонсон [35]). *Функција $\pi : \mathbb{R}^q \to \mathbb{R}_+$ је минимална ваљана функција ако и само ако је периодична, субадитивна, ако важи да је $\pi(0) = 0$ и ако задовољава услов симетричности.*

Доказ. Претходне три леме дају потребан услов. Претпоставимо сад да $\pi(0) = 0$, π је периодична, субадитивна и задовољава услов симетричности.

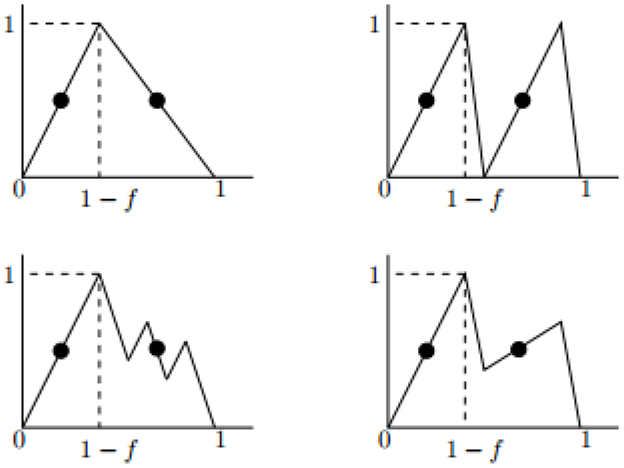
Прво доказујемо да је π ваљана. Услов симетричности имплицира да је $\pi(0) + \pi(-f) = 1$. Пошто је $\pi(0) = 0$, вриједи $\pi(-f) = 1$. Свако $\overline{x} \in G_f$ задовољава једнакост $\sum_{r \in \mathbb{R}^q} r\overline{x}_r = -f + w$ за неко $w \in \mathbb{Z}^q$. Имамо да је

$$\sum_{r \in \mathbb{R}^q} \pi(r)\overline{x}_r \geqslant \pi(\sum_{r \in \mathbb{R}^q} r\overline{x}_r) = \pi(-f + w) = \pi(-f) = 1,$$

гдје неједнакост слиједи из субадитивности, а претпосљедња једнакост слиједи из периодичности. Стога, π је ваљана.

Ако π није минимална, онда постоји ваљана функција $\pi' \leqslant \pi$ таква да је $\pi'(\tilde{r}) < \pi(\tilde{r})$ за неко $\tilde{r} \in \mathbb{R}^q$. У том случају, $\pi(\tilde{r}) + \pi(-f - \tilde{r}) = 1$ имплицира да је $\pi'(\tilde{r}) + \pi'(-f - \tilde{r}) < 1$, што противријечи ваљаности функције π'. □

Примјер 4.5. Слика 4.6 приказује пар примјера минималних ваљаних функција за (4.17), при чему је $q = 1$. Функције су представљене на сегменту $[0, 1]$ (на остатку домена су дефинисане периодом). Уочљива је симетрија у односу на тачке $(\frac{1-f}{2}, \frac{1}{2})$ и $(1 - \frac{f}{2}, \frac{1}{2})$. ■



Слика 4.6 Примјери минималних ваљаних функција ($q = 1$)

Гомори *et al.* је у [38] доказао да је за непрекидне ненегативне дио-по-дио линеарне функције довољно провјерити да је $\pi(a) + \pi(b) \geqslant \pi(a+b)$ и $\pi(a) + \pi(b-a) \geqslant \pi(b)$ у свим тачкама прелома функције.

Нека $E(\pi)$ представља скуп свих могућих неједнакости $\pi(r^1) + \pi(r^2) \geqslant \pi(r^1 + r^2)$ које су задовољене као једнакости, под условом да је π минимална ваљана функција. Важи сљедеће тврђење.

Лема 4.17. *Ако је π минимална ваљана функција, таква да је $\pi = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$, гдје је π_1 и π_2 ваљане функције, онда су π_1 и π_2 минималне ваљане функције и $E(\pi) \subseteq E(\pi_1) \cap E(\pi_2)$.*

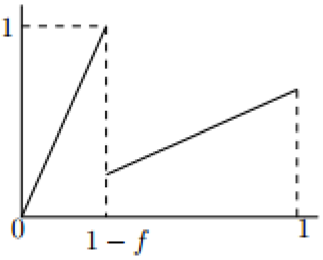
Доказ. Претпоставимо да π_1 није минимално. Нека је $\pi_1' \neq \pi$ ваљана функција, таква да је $\pi_1' \leqslant \pi_1$. Онда $\pi' = \frac{1}{2}\pi_1' + \frac{1}{2}\pi_2$ је ваљана функција, различита од π и $\pi' \leqslant \pi$. Ово противрјечи минималности функције π.

Претпоставимо да $E(\pi) \nsubseteq E(\pi_1) \cap E(\pi_2)$. Можемо претпоставити да је $E(\pi) \nsubseteq E(\pi_1)$, односно да постоје r_1, r_2 такви да је $\pi(r_1)+\pi(r_2) = \pi(r_1+r_2)$ и $\pi_1(r_1)+\pi_1(r_2) > \pi(r_1 + r_2)$. Пошто је π_2 минимална, она је субадитивна и стога је $\pi_2(r_1) + \pi_2(r_2) \geqslant \pi_2(r_1 + r_2)$. Ово противрјечи претпоставци да је $\pi = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$. □

Ваљана функција π је *екстремна* ако се не може изразити као конвексна комбинација двије различите ваљане функције, односно из $\pi = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$ слиједи да је $\pi = \pi_1 = \pi_2$. Екстремној функцији π одговара *екстремна* ваљана неједнакост $\sum_{r\in\mathbb{R}^q} \pi(r)x_r \geqslant 1$ за G_f. Смисао екстремних функција је у томе да оне одређују неједнакости које дефинишу стране бесконачног полиедра. Овај концепт се адекватно уводи и код осталих релаксација, што ће се моћи видјети у наредним поглављима.

Примјер 4.6. Четири функције приказане на Слици 4.6 су екстремне. Што се тиче прве три, то се лако увиђа из теореме о два нагиба (у наставку), док је доказ ектремности посљедње функције нешто сложенији (погледати [37]). Иначе, екстремне функције нису увијек непрекидне. Диј *et al.* [23] је доказао да је за $0 < 1 - f < 0.5$ сљедећа функција (Сл. 4.7) екстремна.

$$\pi(r) = \begin{cases} \frac{r}{1-f} & 0 \leqslant r \leqslant 1-f \\ \frac{r}{2-f} & 1-f < r < 1 \end{cases}$$ ■



Слика 4.7 Екстремна ваљана функција са прекидом

Гомори и Џонсон су поставили хипотезу да су екстремне ваљане функције увијек дио-по-дио линеарне. Басу *et al.* у [13] је нашао контрапримјер, но патолошки случајеви настају само када је број цјелобројних промјенљивих бесконачан.

Корисна оруђа у доказивању да је нека функција екстремна су наредне двије тврдње које наводимо без доказа (погледати [18]).

Лема 4.18 (Интервална лема). *Нека је функција $f : \mathbb{R} \to \mathbb{R}$ омеђена на сваком омеђеном интервалу и нека су $a_1 < a_2$ и $b_1 < b_2$ четири рационална броја који одређују скупове $A = [a_1, a_2]$, $B = [b_1, b_2]$ и $A+B = [a_1+b_1, a_2+b_2]$. Ако је $f(a)+f(b) = f(a+b)$ за свако $a \in A$ и $b \in B$, онда је f афина функција у сваком од скупова A, B и $A+B$ и има исти нагиб у сваком од њих.*

Из дефиниције слиједи да је функција $\pi : [0,1] \to \mathbb{R}_+$ дио-по-дио линеарна ако постоји коначно много вриједности $0 = r_0 < r_1 < \ldots < r_k = 1$ таквих да је функција облика $\pi(r) = a_j r + b_j$ на интервалу (r_{j-1}, r_j) за $j = 1, \ldots, k$. *Нагиби* дио-по-дио линеарне функције су различите вриједности a_j за $j = 1, \ldots, k$. Дио-по-дио линеарна функција $\pi : [0,1] \to \mathbb{R}_+$ је непрекидна ако и само ако за $j = 1, \ldots, k-1$ важи $a_j r_{j+1} + b_j = a_{j+1} r_{j+1} + b_{j+1}$, односно π је непрекидна ако и само ако је $\pi(r) = a_j r_j + b_j$ у сегменту $[r_{j-1}, r_j]$ за $j = 1, \ldots, k$.

Теорема 4.19 (Гомори-Џонсонова теорема о два нагиба). *Ако је рестрикција минималне ваљане функције $\pi : \mathbb{R} \to \mathbb{R}_+$ на сегмент $[0,1]$ непрекидна дио-по-дио линеарна функција са само два нагиба, онда је π екстремна.*

На основу претходне теореме слиједи да су Гоморијеве мјешовито-цјелобројне одсијецајуће равни екстремне.

Иначе, треба имати у виду да је теорема примјењљива само на једноредни проблем ($q = 1$ у (4.17)). Корнуижол и Молинаро су у [20] доказали теорему о три нагиба за дворедне проблеме ($q = 2$ у (4.17)).

4.4 Непрекидна бесконачна релаксација

Разматра се модел у ком небазне промјенљиве y_j ($j \in \boldsymbol{\eta}$) не морају бити цјелобројне.

$$\begin{aligned} f_i + \sum_{j \in \boldsymbol{\eta}} r_i^j y_j \in \mathbb{Z} \qquad & \text{за } i = 1, \ldots, q \\ y_j \geqslant 0 \qquad & \text{за } j \in \boldsymbol{\eta}. \end{aligned} \tag{4.20}$$

Од датог модела се прави *непрекидна бесконачна релаксација* увећавањем простора промјењивих y_j $(j \in \boldsymbol{\eta})$ до бесконачно-димензионалног простора $\{y_r \ : \ r \in \mathbb{R}^q\}$, односно

$$\begin{aligned} f + \sum_{r \in \mathbb{R}^q} r y_r \in \mathbb{Z}^q \\ y_r \geqslant 0 \quad \text{за све } r \in \mathbb{R}^q, \end{aligned} \tag{4.21}$$

при чему је $y_r > 0$ само за коначан број $r \in \mathbb{R}^q$.

У наставку, скуп свих допустивих тачака за (4.21) биће означаван са R_f. Наравно, функција $\psi : \mathbb{R}^q \to \mathbb{R}$ је *ваљана* за релаксацију (4.21) ако линеарна неједнакост

$$\sum_{r \in \mathbb{R}^q} \psi(r) y_r \geqslant 1 \tag{4.22}$$

вриједи у свим тачкама скупа R_f. Свака овако дефинисана ваљана функција ψ, рестрикцијом неједнакости (4.22) на промјењиве y_{r^j} $(j \in \boldsymbol{\eta})$, даје ваљану неједнакост за скуп (4.20).

Ваљана функција $\psi : \mathbb{R}^q \to \mathbb{R}$ за (4.21) је *минимална* ако не постоји ваљана функција $\psi' \neq \psi$ таква да је $\psi'(r) \leqslant \psi(r)$ за све $r \in \mathbb{R}^q$.

Лема 4.20. *Ако је $\psi : \mathbb{R}^q \to \mathbb{R}$ минимална ваљана функција за R_f, онда важи да је ψ позитивно хомогена, субадитивна и ненегативна.*

Доказ. Доказ да је ψ субадитивна је у суштини идентичан доказу Леме 4.14.

Сад, требамо доказати да је ψ позитивно хомогена. Претпоставимо да постоји $\tilde{r} \in \mathbb{R}^q$ и $\lambda > 0$ такви да је $\psi(\lambda \tilde{r}) \neq \lambda \psi(\tilde{r})$. Без губљења општости можемо претпоставити да је $\psi(\lambda \tilde{r}) < \lambda \psi(\tilde{r})$, иначе може размотрити $\lambda \tilde{r}$ умјесто $\tilde{r}$ и λ^{-1} умјесто λ. Дефинишемо функцију ψ' на сљедећи начин:

$$\psi'(r) = \begin{cases} \lambda^{-1} \psi(\lambda \tilde{r}) & \text{ако } r = \tilde{r} \\ \psi(r) & \text{иначе} \end{cases}.$$

Показаћемо да је ψ' ваљана. Размотримо неко $\overline{y} \in R_f$. Дефинишимо $\tilde{y}$ на сљедећи начин:

$$\tilde{y}_r = \begin{cases} 0 & \text{ако } r = \tilde{r} \\ \overline{y}_{\lambda \tilde{r}} + \lambda^{-1} \overline{y}_{\tilde{r}} & \text{ако } r = \lambda \tilde{r} \\ \overline{y}_r & \text{иначе} \end{cases}.$$

Користећи дефиницију функција ψ' и $\tilde{y}$, лако је провјерити да вриједи

$$\sum_{r \in \mathbb{R}^q} \psi'(r) \overline{y}_r = \sum_{r \in \mathbb{R}^q} \psi(r) \overline{y}_r.$$

Штавише, имамо да је $f + \sum_{r \in \mathbb{R}^q} r \overline{y}_r = f + \sum_{r \in \mathbb{R}^q} r \tilde{y}_r \in \mathbb{Z}^q$. Пошто је $\tilde{y} \geqslant 0$, слиједи да је $\tilde{y}$ допустиво за (4.21). Из ваљаности ψ слиједи да је $\sum_{r \in \mathbb{R}^q} \psi(r) \tilde{y}_r \geqslant 1$, па је тиме $\sum_{r \in \mathbb{R}^q} \psi'(r) \overline{y}_r \geqslant 1$. Значи ψ' је ваљана, што противрјечи минималности функције ψ, одакле слиједи позитивна хомогеност.

Остала је још ненегативност да се докаже. Прво доказујемо да је $\psi(r) \geqslant 0$ за свако $r \in \mathbb{Q}^q$. Претпоставимо да је $\psi(\tilde{r}) < 0$ за неко $\tilde{r} \in \mathbb{Q}^q$. Нека је $D \in \mathbb{Z}_+$ такво да је $D\tilde{r}$ цјелобројан вектор и нека је $\overline{y}$ допустиво рјешење за (4.17) (нпр. $\overline{y}_r = 1$ за

$r = -f$, а у осталим случајевима је $\overline{y}_r = 0$). Нека је $\tilde{y}$ дефинисано са $\tilde{y}_{\tilde{r}} = \overline{y}_{\tilde{r}} + MD$, при чему је M природан број и $\tilde{y}_r = \overline{y}_r$ за $r \neq \tilde{r}$. Слиједи да је $\tilde{y}$ допустиво за (4.17). Имамо да је

$$\sum_{r \in \mathbb{R}^q} \psi(r)\tilde{y}_r = \sum_{r \in \mathbb{R}^q} \psi(r)\overline{y}_r + \psi(\tilde{r})MD.$$

M се може изабрати довољно великим, тачније таквим да је

$$M > \frac{\sum_{r \in \mathbb{R}^q} \psi(r)\overline{y}_r - 1}{D|\psi(\tilde{r})|},$$

па да вриједи $\sum_{r \in \mathbb{R}^q} \psi(r)\tilde{y}_r < 1$ што противрјечи тврдњи да је $\tilde{y}$ допустиво.

Из претходног, пошто је ψ непрекидна функција и ненегативна на $\mathbb{Q}^q$, а $\mathbb{Q}^q$ је густ у $\mathbb{R}^q$, слиједи да је ψ ненегативна на читавом $\mathbb{R}^q$. □

Стандардан концепт у конвексној анализи је функција Минковског и испоставља се да се он може искористити у карактеризацији ваљаних функција, другим ријечима, Балаове одсијецајуће равни имају лијеп опис на језику конвексне анализе. За скуп $K \subseteq \mathbb{R}^q$ који је затворен, конвексан и садржи исходиште у својој унутрашњости, *функција Минковског* μ_K се дефинише као

$$\mu_K(r) = \inf\{t > 0 \ : \ \frac{r}{t} \in K\}, \text{ за све } r \in \mathbb{R}^q.$$

У сљедећој леми су дата нека својства функције Минковског.

Лема 4.21. *За дати затворен конвексан скуп K са исходиштем у унутрашњости, функција Минковског μ за K је ненегативна, позитивно хомогена и субадитивна функција.*

Доказ. Из дефиниције функције Минковског слиједи да је она позитивно хомогена и ненегативна.

Треба доказати да је субадитивна. Пошто је K затворен конвексан скуп, μ је конвексна функција. Вриједи

$$\mu(r^1) + \mu(r^2) = 2(\mu(\frac{r^1}{2}) + \mu(\frac{r^2}{2})) \geqslant 2\mu(\frac{r^1 + r^2}{2}) = \mu(r^1 + r^2),$$

гдје једнакости слиједе из позитивне хомогености, а неједнакост из конвексности. □

Сада је могуће доказати сљедећу лему.

Лема 4.22. *Ако је B $\mathbb{Z}^q$-слободан затворен конвексан скуп са тачком f у својој унутрашњости и ако је ψ функција Минковског за $B - f$, онда је ψ ваљана функција.*

Доказ. Према претходној леми, ψ је субадитивна и позитивно хомогена. Нека је $y \in R_f$. Тада је $\sum_{r \in \mathbb{R}^q} ry_r = \overline{x} - f$ за неко $\overline{x} \in \mathbb{Z}^q$. Важи

$$\sum \psi(r)y_r = \sum \psi(ry_r) \geqslant \psi(\sum ry_r) = \psi(\overline{x} - f) \geqslant 1,$$

гдје прва једнакост слиједи из позитивне хомогености, прва неједнакост лиједи из субадитивности, а посљедња неједнакост слиједи из чињенице да је B $\mathbb{Z}^q$-слободан затворен конвексан скуп. □

Примједба 4.23. Позитивно хомогена и субадитивна функција $g \ : \ \mathbb{R}^q \to \mathbb{R}$ је конвексна (па тиме и непрекидна).

Лема 4.24. *Ако је $g \ : \ \mathbb{R}^q \to \mathbb{R}$ ненегативна, позитивно хомогена и субадитивна функција и ако је $K = \{x \in \mathbb{R}^n \ : \ g(x) \leqslant 1\}$, онда је K затворен конвексан скуп са исходиштем у својој унутрашњости, а g је функција Минковског за K.*

Доказ. Према претходној примједби, g је конвексна, стога K је затворен конвексан скуп. Пошто је унутрашњост скупа K $\{x \in \mathbb{R}^n \ : g(x) < 1\}$ и $g(0) = 0$, исходиште је у унутрашњости скупа K.

Нека је $x \in \mathbb{R}^q$. Ако зрак $\{tx \ : \ t \geqslant 0\}$ пресијеца обод скупа K, нека је онда $t^* > 0$ такво да је $g(t^*x) = 1$. Пошто је g позитивно хомогена, $g(x) = \frac{1}{t^*} = \inf\{t > 0 \ : \ \frac{x}{t} \in K\}$. Ако зрак $\{tx \ : \ t \geqslant 0\}$ не пресијеца обод скупа K, пошто је g ненегативна и позитивно хомогена, онда је $g(tx) = 0$ за све $t > 0$. Због тога је $g(x) = 0 = \inf\{t > 0 \ : \ \frac{x}{t} \in K\}$. ☐

За ненегативну, позитивно хомогену и субадитивну функцију ψ дефинише се

$$B_\psi = \{x \in \mathbb{R}^q \ : \ \psi(x - f) \leqslant 1\}.$$

Теорема 4.25. *Ако је ψ минимална ваљана функција, онда је B_ψ максимални $\mathbb{Z}^q$-слободан затворен конвексан скуп који садржи f у својој унутрашњости и ψ је функција Минковског за $B_\psi - f$.*

Доказ. Пошто је ψ минимална ваљана функција, према Леми 4.20 ψ је ненегативна, позитивно хомогена и субадитивна функција. Према Леми 4.24, B_ψ је затворен конвексан скуп са f у својој унутрашњости и ψ је функција Минковског за $B_\psi - f$.

Пошто је ψ ваљана, $\psi(\overline{x} - f) \geqslant 1$ за свако $\overline{x} \in \mathbb{Z}^q$ и стога је B_ψ $\mathbb{Z}^q$-слободан затворен конвексан скуп.

Још само треба доказати да је B_ψ максималан $\mathbb{Z}^q$-слободан затворен конвексан скуп. Претпоставимо да није и нека је B' $\mathbb{Z}^q$-слободан затворен конвексан скуп који садржи као прави подскуп B_ψ. Нека је ψ' функција Минковског $B' - f$. Онда је према дефиницији функције $\psi' \leqslant \psi$ и $\psi' \neq \psi$, јер је $B' \neq B_\psi$. Према претходној леми ψ' је ваљана функција, што противријечи минималности функције ψ. ☐

Примједба 4.26. Пошто је B_ψ скуп нецјелобројних тачака око f које су одсјечене одговарајућим неједнакостима, слиједи да се запремина скупа B_ψ може искористити као квалитативна мјера за јачину произведене одсијецајуће равни.

Теорема 4.27. *Ако је B максималан $\mathbb{Z}^q$-слободан затворен конвексан скуп у чијој унутрашњости се налази f и ако је $\psi : \mathbb{R}^q \to \mathbb{R}$ функција Минковског за $B - f$, онда је ψ минимална ваљана функција за (4.21).*

Доказ. Према Леми 4.22 ψ је ваљана. Ако се претпостави да постоји минимална ваљана функција ψ' таква да је $\psi' \leqslant \psi$ и $\psi' \neq \psi$, онда је B'_ψ $\mathbb{Z}^q$-слободан затворен конвексан скуп и $B_{\psi'} \supset B_\psi$. Пошто је $B_{\psi'} = B_\psi$, претходна тврдња противријечи тврдњи о максималности скупа B. ☐

Као и у претходном поглављу, ваљана функција ψ је *екстремна* ако се не може изразити као конвексна комбинација двије различите ваљане функције. Одговарајућа ваљана неједнакост се такође назива *екстремном.*

Сљедећа теорема говори о подударности између екстремних неједнакости за бесконачни модел (4.21) и екстремних неједнакости за коначан проблем (4.20) [12].

Теорема 4.28. *Нека је B максималан $\mathbb{Z}^q$-слободан конвексан скуп у $\mathbb{R}^q$ са f у својој унутрашњости. Нека је $L = \operatorname{lin}(B)$ и нека је $P = B \cap (f + L^{\perp})$. Онда је $B = P + L$, L је рационални простор, а P је политоп. Нека су $v^1, \ldots, v^k$ врхови политопа P и $r^{k+1}, \ldots, r^{k+h}$ рационална база простора L. Дефинишимо $r^j = v^j - f$ за $j = 1, \ldots, k$. Нека $R_f(r^1, \ldots, r^{k+h})$ означава скуп рјешења за (4.20), при чему је $\boldsymbol{\eta} = \{1, \ldots, k+h\}$. Онда неједнакост $\sum_{r\in\mathbb{R}^q} \psi_B(r)s_r \geqslant 1$ екстремна за R_f ако и само ако је неједнакост $\sum_{j=1}^{k} s_j \geqslant 1$ екстремна за $\operatorname{conv}(R_f(r^1, \ldots, r^{k+h}))$.*

4.5 Мјешовито-цјелобројна бесконачна релаксација

Сада се разматра сљедећи мјешовито-цјелобројан скуп

$$\begin{aligned} f + \sum_{r\in\mathbb{R}^q} r x_r + \sum_{r\in\mathbb{R}^q} r y_r &\in \mathbb{Z}^q \\ x_r \in \mathbb{Z}_+ \qquad & \text{за све } r \in \mathbb{R}^q \\ y_r \geqslant 0 \qquad & \text{за све } r \in \mathbb{R}^q \end{aligned} \tag{4.23}$$

при чему је само коначан број промјенљивих x_r и y_r већи од нуле.

Са M_f означаваћемо скуп допустивих тачака за (4.23). Може се примијетити да је бесконачна релаксација G_f у ствари скуп $\{x \ : (x, 0) \in M_f\}$, а непрекидна бесконачна релаксација R_f скуп $\{y \ : \ (0, y) \in M_f\}$.

Функција (π, ψ), таква да је $\pi \ : \ \mathbb{R}^q \to \mathbb{R}$ и $\psi \ : \ \mathbb{R}^q \to \mathbb{R}$ је *ваљана* за M_f ако $\pi \geqslant 0$ и линеарна неједнакост

$$\sum_{r\in\mathbb{R}^q} \pi(r)x_r + \sum_{r\in\mathbb{R}^q} \psi(r)y_r \geqslant 1 \tag{4.24}$$

вриједи за све векторе из M_f. Ако је (π, ψ) ваљана функција, онда је π ваљана функција за G_f, а ψ је ваљана за R_f.

Ваљана функција (π, ψ) за M_f је *минимална* ако не постоји ваљана функција (π', ψ') различита од (π, ψ), таква да је $\pi' \leqslant \pi$ и $\psi' \leqslant \psi$.

Лема 4.29. *Ако је (π, ψ) минимална ваљана функција за M_f, онда:*

- *$\pi(r) \leqslant \psi(r)$ за сваки $r \in \mathbb{R}^q$*
- *ψ је ненегативна, позитивно хомогена, субадитивна функција.*

Доказ. Нека је (π, ψ) минимално ваљана функција за M_f. Претпоставимо да је $\pi(r^*) > \psi(r^*)$ за неко $r^* \in \mathbb{R}^q$. Нека је (π', ψ) функција дефинисана тако да је

$$\pi'(r) = \begin{cases} \psi(r) & \text{ако је } r = r^* \\ \pi(r) & \text{ако је } r \in \mathbb{R}^q \backslash \{r^*\} \end{cases}$$

За дато (x, y) из скупа допустивих тачака дефинишемо

$$x'_r = \begin{cases} 0 & \text{за } r = r^* \\ x_r & \text{за } r \in \mathbb{R}^q \backslash \{r^*\} \end{cases} \qquad\qquad y'_r = \begin{cases} x_r + y_r & \text{за } r = r^* \\ y_r & \text{за } r \in \mathbb{R}^q \backslash \{r^*\} \end{cases}.$$

Непосредно се провјерава да је тачка $(x', y') \in M_f$ и да вриједи

$$\sum_{r\in\mathbb{R}^q} \pi'(r)x_r + \sum_{r\in\mathbb{R}^q} \psi(r)y_r = \sum_{r\in\mathbb{R}^q} \pi(r)x'_r + \sum_{r\in\mathbb{R}^q} \psi(r)y'_r \geqslant 1.$$

Из овога слиједи да је (π', ψ) ваљана функција, а то противrијечи минималности функције (π, ψ). Овим је доказана прва тачка теореме.

Што се тиче друге тачке теореме, у суштини, истом аргументацијом као у доказу Леме 4.20 показује се да је $\psi : \mathbb{R}^q \to \mathbb{R}$ ненегативна, позитивно хомогена и субадитивна функција и да за такву функцију важи $\psi(0) = 0$. $\square$

Сљедећу теорему је доказао Џонсон у [41] и она говори о томе да је у минималној ваљаној функцији (π, ψ) за M_f, функција ψ јединствено одређена функцијом π.

Теорема 4.30. *Ваљана функција (π, ψ) је минимална за M_f ако и само ако је π минимална ваљана функција за G_f и ψ је дефинисано на сљедећи начин*

$$\psi(r) = \limsup_{\epsilon \to 0+} \frac{\pi(\epsilon r)}{\epsilon} \quad \text{за све } r \in \mathbb{R}^q. \tag{4.25}$$

Доказ. Користећи исту аргументацију као у доказу Лема 4.13, 4.14, 4.15 може се показати да ако је (π, ψ) минимална, онда функција $\pi : \mathbb{R}^q \to \mathbb{R}$ периодична, субадитивна, задовољава услов симетричности и $\pi(0) = 0$. Према Теореми 4.16 π је минимална ваљана функција за G_f.

Остало је само да се покаже да за дату ваљану функцију (π, ψ) за M_f такву да је π минимална ваљана за G_f, важи да је (π, ψ) минимална ваљна за M_f ако и само ако је ψ дефинисана према (4.25).

Дефинишимо функцију ψ' тако да је

$$\psi'(r) = \limsup_{\epsilon \to 0+} \frac{\pi(\epsilon r)}{\epsilon} \quad \text{за свако } r \in \mathbb{R}^q.$$

Показаћемо да је ψ' добро дефинисана, (π, ψ') је ваљана за M_f и $\psi' \leqslant \psi$. Из овога слиједи да је (π, ψ) минимална ако и само ако је $\psi = \psi'$, што у суштини и треба доказати.

Да би показали да је ψ' добро дефинисана треба показати да је $\limsup$ у (4.25) увијек коначан. Подсјетимо се да је

$$\limsup_{\epsilon \to 0+} \frac{\pi(\epsilon r)}{\epsilon} = \lim_{\alpha \to 0+} \sup\{\frac{\pi(\epsilon r)}{\epsilon} \ : \ 0 < \epsilon \leqslant \alpha\} = \inf_{\alpha > 0} \sup\{\frac{\pi(\epsilon r)}{\epsilon} \ : \ 0 < \epsilon \leqslant \alpha\}.$$

Нека је ψ'' функција таква да је $\psi'' \leqslant \psi$ и (π, ψ'') је минимална ваљана функција за M_f (као што је прије поменуто, таква функција постоји).

Према Леми 4.29, $\pi \leqslant \psi''$ и ψ'' је субадитивна, позитивно хомогена и $\psi''(0) = 0$. Стога, за свако $\epsilon > 0$ и свако $r \in \mathbb{R}^q$, вриједи

$$\frac{\pi(\epsilon r)}{\epsilon} \leqslant \frac{\psi''(\epsilon r)}{\epsilon} = \psi''(r),$$

односно

$$\limsup_{\epsilon \to 0+} \frac{\pi(\epsilon r)}{\epsilon} \leqslant \psi''(r).$$

Ово показује да је ψ' добро дефинисана и да је $\psi' \leqslant \psi'' \leqslant \psi$.

Лако се увиђа, из дефиниције функције ψ' и дефиниције $\limsup$, да је ψ' субадитивна, позитивно хомогена и $\psi'(0) = 0$.

Доказ завршавамо показујући да је (π, ψ') ваљана за M_f. Нека је $(\overline{x}, \overline{y}) \in M_f$. Даље, претпоставимо да је

$$\sum_{r \in \mathbb{R}^q} \pi(r)\overline{x}_r + \sum_{r \in \mathbb{R}^q} \psi'(r)\overline{y}_r = 1 - \delta,$$

гдје је $\delta > 0$.

Нека је $\overline{r} = \sum_{r \in \mathbb{R}^q} r\overline{y}_r$. Из дефиницији функције ψ' слиједи да је за неко $\overline{\alpha} > 0$ довољно мало,

$$\frac{\pi(\epsilon\overline{r})}{\epsilon} < \psi'(\overline{r}) + \delta \text{ за све } 0 < \epsilon \leqslant \overline{\alpha}. \tag{4.26}$$

Можемо изабрати $D \in \mathbb{Z}$ такво да је $D \geqslant \overline{\alpha}$ и дефинисати за све $r \in \mathbb{R}^q$

$$\tilde{x}_r = \begin{cases} \overline{x}_r & r \neq \frac{\overline{r}}{D} \\ \overline{x}_r + D & r = \frac{\overline{r}}{D} \end{cases}.$$

Запазимо да су сви елементи $\tilde{x}$ ненегативни цијели бројеви и да $\sum_{r \in \mathbb{R}^q} r\tilde{x}_r = \sum_{r \in \mathbb{R}^q} r\overline{x}_r + \sum_{r \in \mathbb{R}^q} r\overline{y}_r$, па је стога $\tilde{x} \in G_f$. Сада имамо да је

$$\begin{aligned} \sum_{r \in \mathbb{R}^q} \pi(r)\tilde{x}_r &= \sum_{r \in \mathbb{R}^q} \pi(r)\overline{x}_r + \frac{\pi(D^{-1}\overline{r})}{D^{-1}} \\ &< \sum_{r \in \mathbb{R}^q} \pi(r)\overline{x}_r + \psi'(\overline{r}) + \delta \\ &\leqslant \sum_{r \in \mathbb{R}^q} \pi(r)\overline{x}_r + \sum_{r \in \mathbb{R}^q} \psi'(r)\overline{y}_r + \delta = 1, \end{aligned}$$

гдје прва (строга) неједнакост слиједи из (4.26) јер је $D^{-1} \leqslant \overline{\alpha}$, а друга неједнакост слиједи из субадитивности и позитивне хомогености функције ψ. Слиједи да π није ваљана за G_f, што противрјечи претпоставци теореме. ☐

Лема 4.31. *Ако је (π, ψ) минимална ваљана функција за (4.23) и $B_\psi = \{x \in \mathbb{R}^q : \psi(x - f) \leqslant 1\}$, онда је ψ функција Минковског за $\mathbb{Z}^q$-слободан затворен конвексан скуп B_ψ.*

Доказ. Према Леми 4.29 ψ је ненегативна позитивно хомогена и субадитивна функција и $\psi(0) = 0$, а према Леми 4.24 ψ је функција Минковског за B_ψ. Пошто је ψ ваљана функција за R_f, слиједи да је $\psi(\overline{x} - f) \geqslant 1$, за свако $\overline{x} \in \mathbb{Z}^q$. Стога, B_ψ је $\mathbb{Z}^q$-слободан затворен конвексан скуп. ☐

Ако је (π, ψ) минимална ваљана функција, према Теореми 4.30, π је минимална ваљана функција за G_f, но ψ није, у општем случају, минимална ваљана функција за R_f.

Размотримо четири функције $\pi_1, \ldots, \pi_r$ са Слике 4.6, које су екстремне за G_f и нека s_i^+ означава позитиван нагиб функције π_i у 0, а s_i^- негативан нагиб у 1 (или пак у 0, пошто је функција периодична). Према Теореми 4.30, за сваку од тих функција, функција ψ_i за коју је (π_i, ψ_i) минимална ваљна за M_f је дефинисана као

$$\psi_i(r) = \begin{cases} s_i^+ & \text{за } r \geqslant 0 \\ s_i^- & \text{за } r \leqslant 0 \end{cases}.$$

Позитивни нагиби су идентични, док је највећи негативни нагиб s_1^-, па је ψ_1, тачка-по-тачка гледано, мања него друге функције.

За више информација о својствима ваљаних функција за различите бесконачне релаксације погледати [17, 35, 36, 38, 41].

4.6 О неким потокласама ваљаних неједнакости

Проналажење минималне ваљане неједнакости за M_f своди се на проналажење одговарајућих максималних $\mathbb{Z}^q$-слободних полиедра $B_\psi \subset \mathbb{R}^q$. Иако је позната карактеризација свих максималних $\mathbb{Z}^q$-слободних затворених конвексних скупова у равни, није позната таква карактеризација за више-димензионалне просторе. Данијел Еспиноза је у [27] издвојио три простије фамилије са којима је он, за сада једини са становишта научног израчунавања, постигао одређен успјех. Иначе, у научној литератури, као на примјер у [27], Балаове одсијецајуће равни за бесконачне релаксације се називају *вишередним Гоморијевим одсијецајућим равнима.*

Сљедећи нотацију из [27] прва класа скупова носи ознаку $T1_n$ и дефинише се тако да је

$$T1_q = \{x \in \mathbb{R}^q \ : \ x \geqslant 0, \ \sum_{i=1}^{q} x_i \leqslant q\}.$$

$T1_q$ је одређен са $q+1$ неједнакошћу, а свака од њих садржи једну цјелобројну тачку у својој релативној унутрашњости. Такође је $T1_q^\circ \cap \mathbb{Z}^q = \emptyset$ и његова запремина је $q^q/q!$, то јест он је омеђен, конвексан и затворен максималан $\mathbb{Z}^q$-слободан скуп у $\mathbb{R}^q$.

Друга класа скупова, у ознаци G_q, дефинише се тако да је

$$G_q = [\frac{1}{2} \dots \frac{1}{2}]^T + \{x \ : \ \delta^T x \leqslant q/2, \ \forall \delta \in \{-1, 1\}^q\}.$$

Политоп G_q је пресјек укрштених расцјепа у $\mathbb{R}^q$ који садржи 0-1 хиперкоцку. Одређен је са 2^q неједнакости, при чему свака од њих садржи тачно једну цјелобројну тачку у својој релативној унутрашњости, а запремина политопа G_q је једнака $q^q/q!$. G_q је такође омеђен, конвексан и затворен максималан $\mathbb{Z}^q$-слободан скуп у $\mathbb{R}^q$.

Трећа класа скупова се дефинише на сљедећи начин,

$$T2'_q = \left\{ x : \begin{array}{lll} (R_j) & \displaystyle\sum_{i=1}^{j-1} x_i \leqslant x_j, & j = 1, \dots, q \\ (R_{q+1}) & \displaystyle\sum_{i=1}^{q} x_i \leqslant 2^q - 1 & \end{array} \right\},$$

при чему се дефинишу и вектори $\{v_{k,q}\}_{k=1}^{q+1}$, гдје је $(v_{k,q})_i = 0$ ако је $i < k, (v_{k,q})_k = 2^k(1 - 2^{-q})$ и $(v_{k,q})_i = 2^{i-1}(1 - 2^{-q})$ за $i > k$.

Важи да је $v_{k,q}$ врх политопа $T2'_q$ одређен пресјеком свих ограничења осим (R_k). Штавише, запремина политопа је $(2^{\frac{q+1}{2}} - 2^{\frac{1-q}{2}})^q/q!$, што је знатно већа запремина од запремине политопа $T1_q$ када q расте.

Теорема 4.32. *Унутрашњост скупа $T2'_q$ не садржи нити једну цјелобројну тачку, док, с друге стране, свака страна политопа $T2'_q$ садржи цјелобројну тачку у својој релативној унутрашњости.*

Доказ. Уведимо нову фамилију скупова $T2_q = T2'_q + v_o$, гдје је $(v_o)_i = 1-2^{i-1}$. Пошто је v_o цјелобројан вектор, довољно је доказати да је $T2_q$ $\mathbb{Z}^q$-слободан и да свака његова страна има цјелобројну тачку у својој релативној унутрашњости.

Запазимо да је $T2_q$ скуп тачака $x \in \mathbb{R}^q$ које задовољавају $\sum_{i=1}^{j-1} x_i - x_j \leqslant j-1$, за $j = 1, \ldots, q$ и $\sum_{i=1}^{q} x_i \leqslant q$.

Настављамо доказ индукцијом по q. Према дефиницији $T2_1 = \text{convh}\{0, 1\}$, стога тврдња вриједи у случају да је $q = 1$.

Претпоставимо да тврдња вриједи за $q-1$. Пошто је $(v_o + v_{k,q})_1 \in \{0, 2(1-2^{-q})\}$, онда је $\text{proj}_{x_1}(T2_q) = [0, 2-2^{1-q}]$. Дефинишимо $S_1 = \{x \ : \ x_1 = 0\} \cap T2_n \cap \mathbb{Z}^q$ и $S_2 = \{x \ : \ x_1 = 1\} \cap T2_n \cap \mathbb{Z}^q$. Пресјек $T2_q \cap \mathbb{Z}^q$ се може записати као $S_1 \cup S_2$. Све тачке у S_1 задовољавају (R_1) као једнакост, штавише, све тачке у $\{(0, y) \ : \ y \in T2_{q-1} \cap \mathbb{Z}^{q-1}\}$ задовољавају (R_1) као једнакост. За тачке у S_2 важи да је $\{x \ : \ x_1 = 1\} \cap T2_q = \{x \in \mathbb{R}^q \ : \ x = (1, y), \ y \in T2_{q-1}\}$, па према хипотези, све цјелобројне тачке у S_2 задовољавају неко ограничење скупа $T2_q$. □

Као што је већ поменуто, за случај када је $q = 2$, Корнуижол и Маргот су у [21] окарактерисали тачно какви максимални $\mathbb{Z}^q$-слободни полиедри дају екстремне неједнакости. Испоставља се да међу дате полиедре спадају $T1_2$, $T2_2$, а што није случај са полиедром G_2. Ипак, мала пертурбација било које од неједнакости која дефинише G_2 омогућује да резултујуће неједнакости буду екстремне. Ово запажање, као и ограничена нумеричка прецизност записа бројева у покретној запети, оправдавају са практичног становишта превиђање чињенице да G_2 не даје екстремне неједнакости.

За $q > 2$, могу се уопштити резултати из [21] и доказати да скупови $T1_q$ и $T2_q$ дају екстремне неједнакости. Ипак, да ли је слично уопштење могуће за G_q још увијек није познато.

Можемо такође примијетити да скупови $T1_q$, $T2_q$, G_q садрже 0-1 хиперкоцку у $\mathbb{R}^q$, па ако ротирамо било коју осу за 180 степени око тачке $(\frac{1}{2}, \ldots, \frac{1}{2})$, или ако пак пермутујемо промјенљиве, такође добијамо максималне $\mathbb{Z}^q$-слободне скупове. На G_q не утиче нити једна од ових трансформација, $T1_q$ генерише 2^n различитих скупова, а $T2_n$ генерише $2^n(n-1)!$ различитих скупова. Сваки од датих скупова појединачно гледано називамо *конфигурација*, а фамилије скупова сад можемо означити, за фиксирано q, са $\overline{G}_q$, $\overline{T1}_q$, $\overline{T2}_q$ респективно.

За добијену одсијецајућу раван, која је облика $ax \geqslant 1$, можемо узети за мјеру напретка вриједност $\frac{1}{\|a\|}$. Стога, за дато f и фамилију скупова $\mathcal{S}$, проблем налажења најбоље неједнакости се своди на израчунавање вектора a (тј. вриједности минималне ваљане функције) за све конфигурације, уз задржавање најбоље одсијецајуће равни. За фамилију $\overline{G}_q$ рачунски посао није проблематичан, док је за друге двије фамилије експоненцијалан по q. Уколико вриједност броја q није велика, то не представља проблем у практичном израчунавању. За веће вриједности, итерацијски процес по $\{-1,1\}^q$ се може остварити коришћењем неке шеме за пребројевање засноване на Грејовом коду [44], а за пермутације се може искористити Штајнхаус-Џонсон-Тротеров алгоритам [44]. Главна предност ових алгоритама је да су разлике међу сусједним конфигурацијама мале (не више од два елемента), што омогућава да се сачува дио претходног прорачуна.

Што се тиче рачунских експеримената Еспинозе, у већини случајева резултати добијени помоћу фамилија $\overline{G}_2$, $\overline{T1}_2$, $\overline{T2}_2$ су били бољи у односу на резултате добијене помоћу фамилија $\overline{G}_q$, $\overline{T1}_q$, $\overline{T2}_q$. Ипак, пошто то није увијек био случај, остаје отворено питање у којим случајевима се добијају бољи резултати за веће q.

Друга интересантна ствар је што се у општем случају фамилија $\overline{G}_q$ показала бољом у односу на фамилију $\overline{T2}_q$, која се опет показала бољом у односу на $\overline{T1}_q$.

Имајући претходно у виду, остало је доста могућности да се испита, као што је на примјер додавање неколико одсијецајућих равни у свакој итерацији, али из различитих релаксација, онда избор редова табеле, или пак тражење неких других фамилија скупова B_ψ. Штавише, још увијек није дат одговор на питање колико је добра сама бесконачна релаксација.

Литература

[1] Милан Јовановић. Математичко програмирање. Наставни материјали, стр. 2-4.

[2] В. Вујчић and М. Ашић и Н. Миличић. *Математичко програмирање.* Математички институт, 11000 Београд, Кнез Михаилова 35, 1980.

[3] С. И. Веселов и В. Н. Шевченко. Об экспоненциальном росте коэффициентов агрегирующего уравнения. *Кибернетика*, (4):78–79, 1978.

[4] Ђорђе Дугошија. *Линеарно програмирање.* Завод за уџбенике, Београд, 1 edition, 2011.

[5] Kent Andersen, Quentin Louveaux, Robert Weismantel, and Laurence A. Wolsey. Inequalities from two rows of a simplex tableau. In *IPCO*, pages 1–15, 2007.

[6] David Applegate, William J. Cook, Sanjeeb Dash, and Daniel G. Espinoza. Exact solutions to linear programming problems. *Oper. Res. Lett.*, 35(6):693–699, 2007.

[7] Gennadiy Averkov, Christian Wagner, and Robert Weismantel. Maximal lattice-free polyhedra: Finiteness and an explicit description in dimension three. *Math. Oper. Res.*, 36(4):721–742, November 2011.

[8] Egon Balas. Integer programming and convex analysis: Intersection cuts from outer polars. *Mathematical Programming 2*, pages 330–382, 1972.

[9] Egon Balas. Intersection cuts — a new type of cutting planes for integer programming. *Operations Research*, 19:19–39, 1972.

[10] Egon Balas, Matteo Fischetti, and Arrigo Zanette. On the enumerative nature of gomory's dual cutting plane method. *Math. Program.*, 125(2):325–351, October 2010.

[11] Amitabh Basu, Manoel B. Campêlo, Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. On lifting integer variables in minimal inequalities. In *IPCO*, pages 85–95, 2010.

[12] Amitabh Basu, Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. Maximal lattice-free convex sets in linear subspaces. *Math. Oper. Res.*, 35(3):704–720, 2010.

[13] Amitabh Basu, Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. A counterexample to a conjecture of gomory and johnson. *Math. Program.*, 133(1-2):25–38, 2012.

[14] Paul Camion. Characterization of Totally Unimodular Matrices. In *Proceedings of the American Mathematical Society*, volume 16, pages 1068–1073, 1965.

[15] M. Campelo, G. Cornuejols, and G. Zambelli. Stable sets, corner polyhedra and the chavatal closure. *Operations Research Letters*, 38:375–378, 2009.

[16] Berge Claude. Balanced matrices. *Mathematical Programming*, 2:19–31, 1972.

[17] Gerard Conforti, Michele; Cornuejols. Balanced 0,-1,+1-matrices, bicoloring and total dual integrality. *Mathematical Programming*, 71:249–258, 1995.

[18] Michele Conforti, Gerard Cornuejols, and Giacomo Zambelli. Corner polyhedron and intersection cuts. *Surveys in operations research and management science*, 16(2):105–120, 2011.

[19] Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, STOC '71, pages 151–158, New York, NY, USA, 1971. ACM.

[20] G. Cornuejols and M. Molinaro. A 3-slope theorem for the infinite relaxation in the plane. Technical report, Yorktown Heights, New York, 2011.

[21] Gérard Cornuéjols and François Margot. On the facets of mixed integer programs with two integer variables and two constraints. *Math. Program.*, 120(2):429–456, 2009.

[22] D. Cvetković, M. Čangalović, Đ. Dugošija, V. Kovačević-Vujčić, S. Simić, and J. Vuleta. *Kombinatorna optimizacija: matematička teorija i algoritmi.* Biblioteka Operaciona istraživanja i informacioni sistemi. Društvo operacionih istraživača Jugoslavije, 1996.

[23] Santanu S. Dey, Jean-Philippe P. Richard, Yanjun Li, and Lisa A. Miller. On the extreme inequalities of infinite group problems. *Math. Program.*, 121(1):145–170, 2010.

[24] Marcel Dhiflaoui, Stefan Funke, Carsten Kwappik, Kurt Mehlhorn, Michael Seel, Elmar Schömer, Ralph Schulte, and Dennis Weber. Certifying and repairing solutions to large lps how good are lp-solvers? In *SODA*, pages 255–256, 2003.

[25] J. Edmonds and R. Giles. A min-max relation for submodular functions on graphs. *Annals of Discrete mathematics*, 1:185–204, 1977.

[26] Jack Edmonds. Systems of distinct representatives and linear algebra. *Journal of Research of the National Bureau of Standards*, 718(4):241–245, 1967.

[27] Daniel G. Espinoza. Computing with multi-row gomory cuts. *Oper. Res. Lett.*, 38(2):115–120, 2010.

[28] Jan Foniok and Fukuda Komei. Integer programming. Lecture notes, May 2010.

[29] D. R. Fulkerson, A. Hoffman, and R. Oppenheim. On balanced matrices. *Mathematical Programming Study*, 1:120–132, 1974.

[30] Fred Glover and David C. Sommer. Pitfalls of rounding in discrete management decision problems. *Decision Sciences*, 6(2):211–220, 1975.

[31] R. E. Gomory, A. J. Hoffman, and Thomas J. Watson. On the convergence of an integer-programming process. *Naval Research Logistics Quarterly*, 10(1):121–123, 1963.

[32] Ralph E. Gomory. An algorithm for the mixed-integer problem. Report RM-2597.

[33] Ralph E. Gomory. Relation between integer and non-integer solutions to linear programs. In *Proc. Nat. Acad. Sci.*, volume 53, pages 260–265, 1965.

[34] Ralph E. Gomory. Some polyhedra related to combinatorial problems. *Linear Algebra and Applications*, 2:451–558, 1969.

[35] Ralph E. Gomory and Ellis L. Johnson. Some continuous functions related to corner polyhedra i. *Mathematical Programming*, 3:23–85, 1972. 10.1007/BF01584976.

[36] Ralph E. Gomory and Ellis L. Johnson. Some continuous functions related to corner polyhedra ii. *Mathematical Programming*, 3:359–389, 1972.

[37] Ralph E. Gomory and Ellis L. Johnson. T-space and cutting planes. *Math. Program.*, 96(2):341–375, 2003.

[38] Ralph E. Gomory, Ellis L. Johnson, and Lisa Evans. Corner polyhedra and their connection with cutting planes. *Math. Program.*, 96(2):321–339, 2003.

[39] Harvey J. Greenberg. *Myths and Counterexamples in Myths and Counterexamples in.* INFORMS Computing Society, `http://glossary.computing.society.informs.org/myths/CurrentVersion/mythsIP.pdf`.

[40] A. Hoffman and J. Kruskal. Integral boundary points of convex polyhedra, in Linear Inequalities and Related Systems (H. Kuhn and A. Tucker, Eds.). *Annals of Maths. Study*, 38:223–246, 1956.

[41] Ellis L. Johnson. On the group problem for mixed integer programming. In M. L. Balinski and et al., editors, *Approaches to Integer Programming*, volume 2 of *Mathematical Programming Studies*, pages 137–179. Springer Berlin Heidelberg, 1974. 10.1007/BFb0120692.

[42] Ellis L. Johnson and George L. Nemhauser. Recent developments and future directions in mathematical programming. *IBM Systems Journal*, 31(1):79–93, 1992.

[43] Arnold Kaufmann and Arnaud Henry-Labordere. *Methodes et modeles de la recherche operationnelle*, volume 137 of *Mathematics in science and engineering.* New York: Academic Press, 1977. (стр. 353.).

[44] D. E. Knuth. *The Art of Computer Programming*, volume 4. Addison-Wesley, 1st. edition, February 2005.

[45] Thorsten Koch. The final netlib-lp results. *Op. Res. Letters*, 32:2003.

[46] Jon Lee. *A First Course in Combinatorial Optimization.* Cambridge University Press, 2004.

[47] H. W. jr. Lenstra. Integer programming with a fixed number of variables. *Math. Oper. Res.*, 8:538–548, 1983.

[48] Adam N. Letchford and Andrea Lodi. Strengthening chvatal-gomory cuts and gomory fractional cuts. *Operations Research Letters*, 30:2002, 2001.

[49] Laslo Lovasz. Geometry of numbers and integer programming. In M. Iri and K. Tanabe, editors, *Mathematical Programming: Recent Developments and Applications*, pages 177–201. Kluwer, 1989.

[50] François Margot. Testing cut generators for mixed-integer linear programming. *Mathematical Programming Computation*, 1(1):69–95, July 2009.

[51] R. R. Meyer. On the existence of optimal solutions to integer and mixed-integer programming problems. *Mathematical Programming*, 7:223–235, 1974.

[52] Christos H. Papadimitriou. On the complexity of integer programming. *J. ACM*, 28(4):765–768, October 1981.

[53] David S. Rubin. On the unlimited number of faces in integer hulls of linear programs with a single constraint. *Operations Research*, 18(5):940–946, 1970.

[54] H. E. Scarf. An observation on the structure of production sets with indivisibilities. In *Proceedings of the National Academy of Sciences of the United States of America*, volume 74, pages 3637–3641, 1977.

[55] H. E. Scarf. Integral polyhedra in three space. *Mathematics of Operations Research*, 10:403–438, 1985.

[56] Alexander Schrijver. On cutting planes. *Combinatorics 79, Part II (M. Deza and I. G. Rosenberg, eds.), Annals of Discrete Mathematics*, 9:291–296, 1980.

[57] Alexander Schrijver. *On the History of Combinatorial Optimization (till 1960)*, pages 1–68. 2005.

[58] Klaus Truemper. Alpha-balanced graphs and matrices and *gf*(3)-representability of matroids. *J. Comb. Theory, Ser. B*, 32(2):112–139, 1982.

[59] Laurence A. Wolsey. *Integer programming.* Wiley-Interscience, 1 edition, September 1998.

[60] Giacomo Zambelli. A polynomial recognition algorithm for balanced matrices. *Journal of Combinatorial Theory, Series B*, 95:2005, 2003.

[61] Arrigo Zanette, Matteo Fischetti, and Egon Balas. Can pure cutting plane algorithms work? In *Proceedings of the 13th international conference on Integer programming and combinatorial optimization*, IPCO'08, pages 416–434, Berlin, Heidelberg, 2008. Springer-Verlag.