Selected Topics in Modern Mathematics Edition 2014

Monograph Edited by:
Grzegorz Gancarzewicz
Marcin Skrzyński

Publishing House AKAPIT

Selected Topics in Modern Mathematics Edition 2014

Monograph Edited by:
Grzegorz Gancarzewicz
Marcin Skrzyński

Publishing House AKAPIT
Kraków 2014

Reviewers:
Prof. dr hab. Anatolij Prykarpatski (AGH University of Science and Technology)
Prof. dr hab. Kamil Rusek (Pedagogical University of Cracow)
Prof. dr hab. Tadeusz Stanisz (Cracow University of Economics)

ISBN 978-83-63663-54-4

Typeset using LaTeX.

Wydawnictwo Naukowe "Akapit", Kraków
Phone +48 608 024 572; www.akapit.krakow.pl
e-mail: wn@akapit.krakow.pl

Contents

Chapter 1

Introduction

The monograph **"Selected Topics in Modern Mathematics — Edition 2014"** has been prepared to summarize results of the current research conducted at the Cracow University of Technology's Institute of Mathematics. Occasion to this summary are reorganization of the institute consisting on creating research teams that work on various branches of mathematics and its applications as well as the jubilee of 70th anniversary of the Cracow University of Technology.

70 lat
Politechniki
Krakowskiej

This jubilee is celebrated during the whole academic year 2014-2015. More information about the anniversary can be found at the following site.

`www.jubileusz.pk.edu.pl`

There are presently five research teams at the Cracow University of Technology's Institute of Mathematics:

- Abstract algebra and geometry,

- Differential equations,
- Differential geometry,
- Functional analysis,
- Statistics and stochastic processes.

Several researchers also work in the field of discrete mathematics and graph theory, foundations of mathematics, logic, measure theory and topology.

This is not only occasion to recall and summarize the results achieved, but also to look ahead, take on new tasks for further development. The editors hope that the monograph will stimulate exchange of ideas both among authors of individual chapters and between the authors and readers. The exchange of ideas is the best way to stimulate scientific research. We hope that the monograph will gather scientists involved in various branches of mathematics to exchange their experience and accelerate their research.

The volume consists of thirteen chapters that can be grouped into seven parts:

- Algebra (Chapter 9, *Derivations on rings*, and Chapter 12, *Rank functions in algebra, matrix theory, and geometry*),
- Approximation theory (Chapter 6, *Approximation of continuous and unbounded functions*),
- Differential equations (Chapter 1, *On some applications of the Banach contraction theorem*, and Chapter 2, *Existence of a nontrivial solution for hemivariational inequality involving $p(x)$-Laplacian*),
- Foundations of Mathematics (Chapter 4, *An outline of the theory of* $\mathrm{Aut}(\mathcal{P}(\omega)/\mathrm{FIN})$, and Chapter 11, *Finite axiomatization of logical matrices*),
- Geometry (Chapter 8, *Lipschitz cell decomposition with a parameter in o–minimal structures*, and Chapter 13, *Conformal Killing forms and conformal Killing tensors in Riemannian geometry*),
- Graph theory (Chapter 5, *Cycles through matchings in bipartite graphs*, and Chapter 7, *A survey on known values and bounds on the Shannon capacity*),
- Statistics and mathematical finance (Chapter 3, *Estimation of stability index*, and Chapter 10, *Simple chooser options with some risk reducing derivatives*).

The editors would like to express their thanks to the scientific reviewers of the monograph, Prof. dr hab. Anatolij Prykarpatski (AGH University of Science and Technology), Prof. dr hab. Kamil Rusek (Pedagogical University of Cracow), and Prof. dr hab. Tadeusz Stanisz (Cracow University of Economics) for their great work and valuable comments.

We would like also to express our thanks to Professors Dariusz Cichoń (Jagiellonian University), Barbara Fryc (University of Information Technology and Management in Rzeszów), Włodzimierz Jelonek (Cracow University of Technology), Marta Kosek (Jagiellonian University), Marcin Kulczycki (Jagiellonian University), Piotr Niemiec (Jagiellonian University) and Eugeniusz Wachnicki (Pedagogical University of Cracow) for reviewing selected chapters of the monograph. Their work and comments helped to improve this book.

Grzegorz Gancarzewicz
Marcin Skrzyński

December, 2014

Chapter 2

On Some Applications of the Banach Contraction Theorem

Ludwik Byszewski

Ludwik Byszewski
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: lbyszews@pk.edu.pl

Abstract

The aim of this paper is to prove theorems on the existence and uniqueness of solutions for some nonlocal differential problems, by applying mainly the Banach contraction theorem.

Contents

Part I: Existence and Uniqueness of Solutions of Evolution Nonlocal Cauchy Problem

I.1. Introduction to Part I

In this part of the paper we prove two theorems on the existence and uniqueness of mild and classical solutions of a semilinear integrodifferential evolution nonlocal Cauchy problem for a first order equation. To do it the method of semigroups, the Banach contraction theorem and the Winiarska theorem will be applied. The results of this part of the paper are based also on those from [2, 4, 3, 6, 8, 9].

In this part of the paper we shall assume that E is a Banach space with norm $\|\cdot\|$, $-A$ is the infinitesimal generator of a C_0 semigroup $\{T(t)\}_{t\geq 0}$ on E and $D(A)$ is the domain of A.

Throughout this part of the paper we shall use the notations:

$$J = [t_0, t_0 + a], \text{ where } t_0 \geq 0 \text{ and } a > 0,$$

$$\Delta = \{(t, s) : t_0 \leq s \leq t \leq t_0 + a\},$$
$$M = \sup\{\|T(t)\| : t \in [0, a]\}$$

and

$$X = C(J, E).$$

The nonlocal Cauchy problem considered here is of the form

$$u' + Au(t) = f(t, u(t), u(b_1(t)), \ldots, u(b_m(t)))+$$
$$+\int_{t_0}^{t} f_1(t, s, u(s))ds + \int_{t_0}^{t_0+a} f_2(t, s, u(s))ds, \quad t \in J \setminus \{t_0\}, \tag{1}$$
$$u(t_0) + g(u) = u_0, \tag{2}$$

where $f : J \times E^{m+1} \longrightarrow E$, $f_i : \Delta \times E \longrightarrow E$ $(i = 1, 2)$, $g : X \longrightarrow E$, $b_i : J \longrightarrow J$ $(i = 1, \ldots, m)$ are given functions satisfying some assumptions, and $u_0 \in E$.

I.2 The Winiarska Theorem

The results of this section were obtained by Professor Teresa Winiarska (see [9]).

Let us consider the Cauchy problem

$$u'(t) + Au(t) = k(t), \quad t \in J \setminus \{t_0\}, \tag{3}$$

$$u(t_0) = x. \tag{4}$$

A function $u : J \longrightarrow E$ is said to be a classical solution of problem (3)-(4), if

(i) u is continuous on J and continuously differentiable on $J \setminus \{t_0\}$,

(ii) $u'(t) + Au(t) = k(t)$ for $t \in J \setminus \{t_0\}$,

(iii) $u(t_0) = x$.

Theorem 1. *Assume that E is a reflexive Banach space, $k : J \longrightarrow E$ is Lipschitz continuous on J and $x \in D(A)$. Then the Cauchy problem (3)-(4) has the only one classical solution u given by the formula*

$$u(t) = T(t-t_0)x + \int_{t_0}^{t} T(t-s)k(s)ds, \quad t \in J.$$

Proof. See [9, Section 4.2].

I.3 Theorem about a Mild Solution

A function $u \in C(J, E)$ satisfying the integral equation

$$u(t) = T(t-t_0)u_0 - T(t-t_0)g(u) + \int_{t_0}^{t} T(t-s)\Big(f(s, u(s), u(b_1(s)), \ldots, u(b_m(s))) +$$

$$+ \int_{t_0}^{s} f_1(s,\tau,u(\tau))d\tau + \int_{t_0}^{t_0+a} f_2(s,\tau,u(\tau))d\tau\Big)ds, \quad t \in J$$

is said to be a mild solution of the integrodifferential evolution nonlocal Cauchy problem (1)-(2).

Theorem 2. *Assume that:*

(i) $f : J \times E^{m+1} \longrightarrow E$ is continuous with respect to the first variable in J, $f_i : \Delta \times E \longrightarrow E$ $(i = 1, 2)$ are continuous with respect to the first and second variables on Δ, $g : X \longrightarrow E$, $b_i : J \longrightarrow J$ $(i = 1, \ldots, m)$ are continuous on J, and there exist positive constants L, L_i $(i = 1, 2)$ and K such that

$$\|f(s, z_0, z_1, \ldots, z_m) - f(s, \tilde{z}_0, \tilde{z}_1, \ldots, \tilde{z}_m)\| \le L \sum_{i=0}^{m} \|z_i - \tilde{z}_i\| \tag{5}$$

for $s \in J$, $z_i, \tilde{z}_i \in E$ $(i = 0, 1, \ldots, m)$,

$$\|f_i(s,\tau,z) - f_i(s,\tau,\tilde{z})\| \le L_i\|z - \tilde{z}\| \quad (i = 1, 2) \tag{6}$$

for $(s,\tau) \in \Delta$, $z, \tilde{z} \in E$

and

$$\|g(w) - g(\tilde{w})\| \le K\|w - \tilde{w}\|_X \text{ for } w, \tilde{w} \in X, \tag{7}$$

(ii) $M[a((m+1)L + aL_1 + aL_2) + K] < 1,$

(iii) $u_0 \in E.$

Then the integrodifferential evolution nonlocal Cauchy problem (1)-(2) has a unique mild solution.

Proof. Introduce the operator $F : X \longrightarrow X$ given by the formula

$$(Fw)(t) := T(t-s)u_0 - T(t-s)g(w) +$$
$$+ \int_{t_0}^{t} T(t-s)\Big(f(s, w(s), w(b_1(s)), \ldots, w(b_m(s))) +$$
$$+ \int_{t_0}^{s} f_1(s,\tau,w(\tau))d\tau + \int_{t_0}^{t_0+a} f_2(s,\tau,w(\tau))d\tau \Big) ds, \quad t \in J. \tag{8}$$

From (8) and (5)-(7), we have

$$\|(Fw)(t) - (F\tilde{w})(t)\| \leq \tag{9}$$
$$\leq \|T(t-t_0)\| \|g(w) - g(\tilde{w})\| +$$
$$+ \int_{t_0}^{t} \|T(t-s)\| \|f(s, w(s), w(b_1(s)), \ldots, w(b_m(s))) -$$
$$- f(s, \tilde{w}(s), \tilde{w}(b_1(s)), \ldots, \tilde{w}(b_m(s)))\| ds +$$
$$+ \int_{t_0}^{t} \Big(\|T(t-s)\| \int_{t_0}^{s} \|f_1(s,\tau,w(\tau)) - f_1(s,\tau,\tilde{w}(\tau))\| d\tau \Big) ds +$$
$$+ \int_{t_0}^{t} \Big(\|T(t-s)\| \int_{t_0}^{t_0+a} \|f_2(s,\tau,w(\tau)) - f_2(s,\tau,\tilde{w}(\tau))\| d\tau \Big) ds \leq$$
$$\leq MK\|w - \tilde{w}\|_X + ML \int_{t_0}^{t} \Big(\|w(s) - \tilde{w}(s)\| + \sum_{i=1}^{m} \|w(b_i(s)) - \tilde{w}(b_i(s))\| \Big) ds +$$
$$+ ML_1 \int_{t_0}^{t} \Big(\int_{t_0}^{s} \|w(\tau) - \tilde{w}(\tau)\| d\tau \Big) ds + ML_2 \int_{t_0}^{t} \Big(\int_{t_0}^{t_0+a} \|w(\tau) - \tilde{w}(\tau)\| d\tau \Big) ds \leq$$
$$\leq M[a((m+1)L + aL_1 + aL_2) + K] \|w - \tilde{w}\|_X$$

for $w, \tilde{w} \in X$ and $t \in J$. If we define

$$q = M[a((m+1)L + aL_1 + aL_2) + K],$$

then by (9) and assumption (ii),

$$\|Fw - F\tilde{w}\|_X \leq q\|w - \tilde{w}\|_X \text{ for } w, \tilde{w} \in X \tag{10}$$

with $0 < q < 1$. This shows that F is a contraction on X. Consequently, by (10) the operator F satisfies all the assumptions of the Banach contraction theorem. Therefore, in the space X there is exactly one fixed point of F and this point is a mild solution of problem (1)-(2). So, the proof of the theorem is complete.

I.4 Theorem about a Classical Solution

A function $u : J \longrightarrow E$ is said to be a classical solution of the integrodifferential evolution nonlocal Cauchy problem (1)-(2), if

(i) u is continuous on J and continuously differentiable on $J \setminus \{t_0\}$,

(ii) $u'(t) + Au(t) = f(t, u(t), u(b_1(t)), \ldots, u(b_m(t)))+$
$$+\int_{t_0}^{t} f_1(t, s, u(s))ds + \int_{t_0}^{t_0+a} f_2(t, s, u(s))ds \text{ for } t \in J \setminus \{t_0\},$$

(iii) $u(t_0) + g(u) = u_0$.

Theorem 3. *Assume that:*

(i) E is a reflexive Banach space and $u_0 \in E$,

(ii) $f : J \times E^{m+1} \longrightarrow E$, $f_i : \Delta \times E \longrightarrow E$ $(i = 1, 2)$ are continuous with respect to the first and second variables on Δ, $g : X \longrightarrow E$, $b_i : J \longrightarrow J$ $(i = 1, \ldots, m)$ are continuous on J, and there exist positive constants C, C_i $(i = 1, 2)$ and K such that

$$\|f(s, z_0, z_1, \ldots, z_m) - f(\tilde{s}, \tilde{z}_0, \tilde{z}_1, \ldots, \tilde{z}_m)\| \leq \tag{11}$$
$$\leq C\Big(|s - \tilde{s}| + \sum_{i=0}^{m} \|z_i - \tilde{z}_i\|\Big) \text{ for } s, \tilde{s} \in J,\ z_i, \tilde{z}_i \in E\ (i = 0, 1, \ldots, m),$$

$$\|f_i(s, \tau, z) - f_i(\tilde{s}, \tau, \tilde{z})\| \leq C_i(|s - \tilde{s}| + \|z - \tilde{z}\|) \quad (i = 1, 2) \tag{12}$$
$$\text{for } (s, \tau), (\tilde{s}, \tau) \in \Delta,\ z, \tilde{z} \in E$$

and

$$\|g(w) - g(\tilde{w})\| \leq K\|w - \tilde{w}\|_X \text{ for } w, \tilde{w} \in X, \tag{13}$$

(iii) $M[a((m + 1)C + aC_1 + aC_2) + K] < 1$.

Then the integrodifferential evolution nonlocal Cauchy problem (1)-(2) has a unique mild solution (which will be denoted by) u. Moreover, if $u_0 \in D(A)$, $g(u) \in D(A)$ and there is a positive constant κ such that

$$\|u(b_i(s)) - u(b_i(\tilde{s}))\| \leq \kappa\|u(s) - u(\tilde{s})\| \text{ for } s, \tilde{s} \in J\ (i = 1, \ldots, m), \tag{14}$$

then u is the unique classical solution of problem (1)-(2).

Proof. Since all the assumptions of Theorem 2 are satisfied, it is easy to see that problem (1)-(2) possesses a unique mild solution, denoted by u.

Now, we shall show that u is a classical solution of problem (1)-(2). To this end introduce

$$N := \max_{s\in J} \|f(s, u(s), u(b_1(s)), \ldots, u(b_m(s)))\|, \tag{15}$$

$$N_i := \max_{(s,\tau)\in\Delta} \|f_i(s, \tau, u(\tau))\| \quad (i = 1, 2) \tag{16}$$

and observe that

$$u(t + h) - u(t) = \tag{17}$$

$$= (T(t + h - t_0)u_0 - T(t - t_0)u_0) - (T(t + h - t_0)g(u) - T(t - t_0)g(u))+$$

$$+ \int_{t_0}^{t+h} T(t + h - s)\Big(f(s, u(s), u(b_1(s)), \ldots, u(b_m(s)))+$$

$$+ \int_{t_0}^{s} f_1(s, \tau, u(\tau))d\tau + \int_{t_0}^{t_0+a} f_2(s, \tau, u(\tau))d\tau\Big)ds-$$

$$- \int_{t_0}^{t} T(t - s)\Big(f(s, u(s), u(b_1(s)), \ldots, u(b_m(s)))+$$

$$+ \int_{t_0}^{s} f_1(s, \tau, u(\tau))d\tau + \int_{t_0}^{t_0+a} f_2(s, \tau, u(\tau))d\tau\Big)ds =$$

$$= T(t - t_0)(T(h) - I)u_0 - T(t - t_0)(T(h) - I)g(u)+$$

$$+ \int_{t_0}^{t_0+h} T(t + h - s)\Big(f(s, u(s), u(b_1(s)), \ldots, u(b_m(s)))+$$

$$+ \int_{t_0}^{s} f_1(s, \tau, u(\tau))d\tau + \int_{t_0}^{t_0+a} f_2(s, \tau, u(\tau))d\tau\Big)ds+$$

$$+ \int_{t_0}^{t} T(t - s)\Big(f(s + h, u(s + h), u(b_1(s + h)), \ldots, u(b_m(s + h)))-$$

$$-f(s, u(s), u(b_1(s)), \ldots, u(b_m(s)))\Big)ds+$$

$$+ \int_{t_0}^{t} T(t - s)\Big(\int_{t_0}^{s} (f_1(s + h, \tau, u(\tau)) - f_1(s, \tau, u(\tau)))d\tau\Big)ds+$$

$$+ \int_{t_0}^{t} T(t - s)\Big(\int_{s}^{s+h} f_1(s + h, \tau, u(\tau))d\tau\Big)ds+$$

$$+\int_{t_0}^{t} T(t-s)\Big(\int_{t_0}^{t_0+a} (f_2(s+h,\tau,u(\tau)) - f_2(s,\tau,u(\tau)))d\tau\Big)ds$$

for $t \in [t_0, t_0+a)$, $h > 0$ and $t+h \in (t_0, t_0+a]$. Consequently, by (17), (15), (16), (11), (12) and (14),

$$\|u(t+h)-u(t)\| \leq \tag{18}$$

$$\leq hM\|Au_0\| + hM\|Ag(u)\| + hM(N + aN_1 + aN_2) + MCah +$$

$$+MC\int_{t_0}^{t}\Big(\|u(s+h)-u(s)\| + \sum_{i=1}^{m}\|u(b_i(s+h)) - u(b_i(s))\|\Big)ds +$$

$$+Ma^2C_1h + MaN_1h + Ma^2C_2h =$$

$$= C_*h + MC\int_{t_0}^{t}\Big(\|u(s+h)-u(s)\| + \sum_{i=1}^{m}\|u(b_i(s+h)) - u(b_i(s))\|\Big)ds \leq$$

$$\leq C_*h + MC(1+m\kappa)\int_{t_0}^{t}\|u(s+h)-u(s)\|ds$$

for $t \in [t_0, t_0+a)$, $h > 0$ and $t+h \in (t_0, t_0+a]$. It follows from (18) and Gronwall's inequality that

$$\|u(t+h)-u(t)\| \leq C_* e^{aMC(1+m\kappa)}h$$

for $t \in [t_0, t_0+a)$, $h > 0$ and $t+h \in (t_0, t_0+a]$. Hence u is Lipschitz continuous on J. The Lipschitz continuity of u on J combined with the Lipschitz continuity of f on $J \times E^{m+1}$ and f_i $(i = 1, 2)$ with respect to the first variables in J implies that the function

$$J \ni t \longrightarrow f(t, u(t), u(b_1(t)), \dots, u(b_m(t))) + \int_{t_0}^{t} f_1(t,s,u(s))ds +$$

$$+\int_{t_0}^{t_0+a} f_2(t,s,u(s))ds \in E$$

is Lipschitz continuous on J. This property and the assumptions of Theorem 3 yield, by the Winiarska theorem and Theorem 2, that the linear Cauchy problem

$$v'(t) + Av(t) = f(t, u(t), u(b_1(t)), \dots, u(b_m(t))) +$$

$$+\int_{t_0}^{t} f_1(t,s,u(s))ds + \int_{t_0}^{t_0+a} f_2(t,s,u(s))ds, \quad t \in J \setminus \{t_0\},$$

$$v(t_0) = u_0 - g(u)$$

has a unique classical solution v given by

$$v(t) = T(t-t_0)u_0 - T(t-t_0)g(u)+$$

$$+\int_{t_0}^{t} T(t-s)\Big(f(s,u(s),u(b_1(s)),\ldots,u(b_m(s))) + \int_{t_0}^{s} f_1(s,\tau,u(\tau))d\tau+$$

$$+\int_{t_0}^{t_0+a} f_2(s,\tau,u(\tau))d\tau\Big)ds = u(t) \quad (t \in J).$$

Consequently, u is the unique classical solution of problem (1)-(2) and, therefore, the proof of Theorem 3 is complete.

Part II: Existence and Uniqueness of Solutions of the Neumann Nonlocal Problem together with Nonlocal Initial Condition

II.1 Introduction to Part II

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Moreover, let T be a fixed positive number and $k \in \mathbb{R} \setminus \{0\}$. We will need the following assumption.

Assumption (H) (see [1]). $J \in C(\mathbb{R}^n, \mathbb{R})$ is a nonnegative radial function with $J(0) > 0$ and

$$\int\limits_{\mathbb{R}^n} J(x)dx = 1.$$

In [1], the existence and uniqueness of a solution of the following nonlocal Neumann boundary value problem is studied:

$$u_t(x,t) = \int\limits_{\Omega} J(x-y)(u(y,t) - u(x,t))dy+$$

$$+\int\limits_{\mathbb{R}^n\setminus\Omega} G(x,x-y)g(y,t)dy, \quad x \in \Omega,\ t > 0,$$

$$u(x,0) = u_0(x), \quad x \in \Omega.$$

The Banach contraction theorem is applied to the study. Kamont [10], Muszyński and Myszkis [5], and Pelczar and Szarski [7] also studied the existence and uniqueness of solutions of differential problems using the Banach contraction theorem.

The aim of this part of the paper is to give a theorem on the existence and uniqueness of a solution of the following nonlocal Neumann boundary value problem together with a nonlocal initial condition:

$$u_t(x,t) = \int\limits_{\Omega} J(x-y)(u(y,t)-u(x,t))dy+ \tag{19}$$

$$+\int\limits_{\mathbb{R}^n\setminus\Omega} G(x,x-y)g(y,t)dy, \quad x\in\Omega,\, t\in(0,T),$$

$$u(x,0)+kTu(x,T)=u_0(x), \quad x\in\Omega.$$

For this purpose we will also apply the Banach contraction theorem.

II.2 Existence and Uniqueness of Solutions

Let assumption (H) be satisfied throughout this section. Moreover, let $G \in L^\infty(\Omega\times\mathbb{R}^n)$, $g\in L^\infty([0,T);L^1(\mathbb{R}^n\setminus\Omega))$ and $u_0\in L^1(\Omega)$.

A function $u\in C([0,T];L^1(\Omega))$ is said to be a solution of nonlocal problem (19), if

$$u(x,t) = u_0(x) - kTu(x,T) + \int_0^t\int_\Omega J(x-y)(u(y,s)-u(x,s))dyds+$$

$$+\int_0^t\int_{\mathbb{R}^n\setminus\Omega} G(x,x-y)g(y,s)dyds, \;\; x\in\Omega,\, t\in[0,T].$$

Consider the Banach space

$$X_T = C([0,T];L^1(\Omega))$$

with the norm

$$|||w||| = \max_{0\le t\le T}\|w(\cdot,t)\|_{L^1(\Omega)}.$$

The solution of problem (19) will be obtained as a fixed point of the operator

$$\mathcal{T}_{u_0,g} : X_T \longrightarrow X_T$$

defined by the formula

$$\mathcal{T}_{u_0,g}(w)(x,t) = u_0(x) - kTw(x,T)+ \tag{20}$$

$$+\int_0^t\int_\Omega J(x-y)(w(y,s)-w(x,s))dyds+$$

$$+\int_0^t \int_{\mathbb{R}^n\setminus\Omega} G(x,x-y)g(y,s)dyds \quad (x\in\Omega,\, t\in[0,T]).$$

To prove the existence and uniqueness of the solution of problem (19) we will need the following lemma.

Lemma 4. *Let $G \in L^\infty(\Omega\times\mathbb{R}^n)$, $g,h \in L^\infty((0,T);L^1(\mathbb{R}^n\setminus\Omega))$ and $u_0,v_0 \in L^1(\Omega)$. Then there is a constant*

$$C = \max\{|k|, k_1, k_2\}, \tag{21}$$

where $k_i > 0$ ($i=1,2$), depending only on Ω, J and G such that

$$|||\mathcal{T}_{u_0,g}(w) - \mathcal{T}_{v_0,h}(z)||| \leq \tag{22}$$

$$\leq \|u_0 - v_0\|_{L^1(\Omega)} + 2CT\big(|||w-z||| + \|g-h\|_{L^\infty((0,T);L^1(\mathbb{R}^n\setminus\Omega))}\big)$$

for all $w,z\in X_T$.

Proof. Observe that

$$\int_\Omega |\mathcal{T}_{u_0,g}(w)(x,t) - \mathcal{T}_{v_0,h}(z)(x,t)|dx \leq$$

$$\leq \int_\Omega |u_0(x)-v_0(x)|dx + |k|T\int_\Omega |w(x,T)-z(x,T)|dx +$$

$$+\int_\Omega \Big| \int_0^t \int_\Omega J(x-y)[(w(y,s)-z(y,s)) - (w(x,s)-z(x,s))]dyds \Big| dx +$$

$$+\int_\Omega \int_0^t \int_{\mathbb{R}^n\setminus\Omega} |G(x,x-y)||g(y,s)-h(y,s)|dydsdx \leq$$

$$\leq \|u_0-v_0\|_{L^1(\Omega)} + |k|T|||w-z||| + k_1T|||w-z||| + k_2T\|g-h\|_{L^\infty((0,T);L^1(\mathbb{R}^n\setminus\Omega))} \leq$$

$$\leq \|u_0 - v_0\|_{L^1(\Omega)} + 2CT\big(|||w-z||| + \|g-h\|_{L^\infty((0,T);L^1(\mathbb{R}^n\setminus\Omega))}\big),$$

where C is defined by (21). This yields inequality (22).

Applying Lemma 4 we will prove the existence and uniqueness of the solution of problem (19).

Theorem 5. *Let $G \in L^\infty(\Omega\times\mathbb{R}^n)$. Then, for every $u_0\in L^1(\Omega)$ and $g \in L^\infty((0,T);L^1(\mathbb{R}^n\setminus\Omega))$, there is a unique solution of problem (19) provided that $2CT < 1$, where C is given by (21).*

Proof. Let $\mathcal{T} = \mathcal{T}_{u_0,g}$. We first check that $\mathcal{T}$ maps X_T into X_T. By (20), we have

$$\|\mathcal{T}(w)(t_2) - \mathcal{T}(w)(t_1)\|_{L^1(\Omega)} \leq A \int_{t_1}^{t_2} \int_{\Omega} |w(y,s)| dyds +$$

$$+ B \int_{t_1}^{t_2} \int_{\mathbb{R}^n \setminus \Omega} |g(y,s)| dyds$$

for $0 \leq t_1 \leq t_2 \leq T$. The above estimate gives that $\mathcal{T}(w) \in C([0,T]; L^1(\Omega))$. Hence $\mathcal{T}$ maps X_T into X_T.

Now, choose T such that $2CT < 1$. Lemma 4 implies that $\mathcal{T}$ is a strict contraction in X_T. Therefore, the existence and uniqueness of the solution of problem (19) follow from the Banach contraction theorem on the interval $[0,T]$. The proof is complete.

References

[1] F. Andreu-Vailo et al. *Nonlocal Diffusion Problems.* American Mathematical Society, Providence, R. I., 2010.

[2] K. Balachandran and S. Ilamaran. "Existence and uniqueness of mild and strong solutions of a semilinear evolution equation with nonlocal conditions". In: *Indian J. Pure Appl. Math.* 25, No. 4 (1994), pp. 411–418.

[3] L. Byszewski. "Existence and uniqueness of mild and classical solutions of semilinear functional-differential evolution nonlocal Cauchy problem". In: *Selected Problems of Mathematics, Cracow University of Technology, Anniversary Issue 6.* 1995, pp. 25–33.

[4] L. Byszewski and T. Winiarska. "Integrodifferential evolution nonlocal problem for the first order equation". In: *Technical Transactions* 1-NP (2011), pp. 15–21.

[5] J. Muszyński and A. D. Myszkis. *Ordinary Differential Equations.* Państwowe Wydawnictwo Naukowe, Warszawa [in Polish], 1984.

[6] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer-Verlag, New York etc., 1983.

[7] A. Pelczar and J. Szarski. *Introduction to the Theory of Differential Equations.* Państwowe Wydawnictwo Naukowe, Warszawa [in Polish], 1987.

[8] H. L. Tidke. "Existence of solutions of nonlinear mixed integrodifferential equations of Sobolev type". In: *Nonlinear Funct. Anal. Appl.* 14, No. 4 (2009), pp. 605–618.

[9] T. Winiarska. *Differential Equations with Parameter, Monograph 68.* Technical University of Cracow, 1988.

[10] Kamont. Z. *Ordinary Differential Equations*. Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk [in Polish], 1999.

Chapter 3

Existence of a nontrival solution for hemivariational inequality involving $p(x)$-Laplacian

Sylwia Barnaś

Sylwia Barnaś
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: sbarnas@pk.edu.pl

Abstract

We study the nonlinear elliptic problem with $p(x)$-Laplacian (hemivariational inequality) where the weighted function λ may change sign. By using nonsmooth critical point theory for locally Lipschitz functional and the properties of the generalized Sobolev spaces, we obtain conditions which ensure the existence of a solution for our problem.

Contents

3.1 Introduction

Let $\Omega \subseteq \mathbb{R}^N$ be a bounded domain with a smooth boundary $\partial\Omega$. We consider a differential inclusion in Ω involving a $p(x)$-Laplacian of the type

$$\begin{cases} -\Delta_{p(x)}u - \lambda|u(x)|^{p(x)-2}u(x) \in \partial j(x, u(x)) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{P}$$

where $p : \overline{\Omega} \to \mathbb{R}$ is a continuous function satisfying

$$1 < p^- := \inf_{x\in\overline{\Omega}} p(x) \leqslant p(x) \leqslant p^+ := \sup_{x\in\overline{\Omega}} p(x) < \widehat{p}^*,$$

where

$$\widehat{p}^* := \begin{cases} \frac{Np^-}{N-p^-} & p(x) < N, \\ \infty & p(x) \geqslant N. \end{cases} \tag{3.1}$$

A functional $j(x, t)$ is a measurable in the first variable and locally Lipschitz in the second variable. By $\partial j(x, t)$ we denote the subdifferential of $j(x, \cdot)$ in the sense of Clarke [6]. The operator

$$\Delta_{p(x)}u := \operatorname{div}\big(|\nabla u(x)|^{p(x)-2}\nabla u(x)\big)$$

is the so-called $p(x)$-Laplacian, which becomes p-Laplacian when $p(x) \equiv p$.

In problem (P) appears λ, for which we will assume that

$$\lambda < \frac{p^-}{p^+}\lambda_* \quad \text{and} \quad \lambda < \frac{(p^- - 1)p^+}{(p^+ - 1)p^-}\lambda_*, \tag{3.2}$$

where λ_* is defined by

$$cit\lambda_* = \inf_{u\in W_0^{1,p(x)}(\Omega)\setminus\{0\}} \frac{\int_\Omega |\nabla u(x)|^{p(x)}dx}{\int_\Omega |u(x)|^{p(x)}dx}. \tag{3.3}$$

When we consider

$$\widetilde{p} := \min\Big\{\frac{(p^- - 1)p^+}{(p^+ - 1)p^-}, \frac{p^-}{p^+}\Big\},$$

then

$$\lambda < \widetilde{p}\lambda_*. \tag{3.4}$$

It may happen that $\lambda_* = 0$ (see Fan-Zhang [7]).

3.2 Mathematical preliminaries

Let X be a Banach space and X^* its topological dual. By $\|\cdot\|$ we will denote the norm in X and by $\langle\cdot,\cdot\rangle$ the duality brackets for the pair (X, X^*).

3.2.1 Clarke subdifferential and Cerami condition

In analogy with the directional derivative of a convex function, we define the generalized directional derivative of a locally Lipschitz function f at $x \in X$ in the direction $h \in X$ by

$$f^0(x,h) = \limsup_{x' \to 0, \lambda \searrow 0} \frac{f(x+x'+\lambda h) - f(x+x')}{\lambda}.$$

The function $h \longmapsto f^0(x,h) \in \mathbb{R}$ is sublinear, continuous so it is the support function of a nonempty, convex and w^*-compact set

$$\partial f(x) = \{x^* \in X^* : \langle x^*, h\rangle \leqslant f^0(x,h) \text{ for all } h \in X\}.$$

The set $\partial f(x)$ is known as generalized or Clarke subdifferential of f at x. If f is strictly differentiable at x (in particular if f is continuously Gâteaux differentiable at x), then $\partial f(x) = \{f'(x)\}$.

A point $x \in X$ is said to be a critical point of the locally Lipschitz function $f : X \to \mathbb{R}$, if $0 \in \partial f(x)$. If $x \in X$ is local minimum or local maksimum of f, then x is critical point, and moreover this time the value $c = f(x)$ is called a critical value of f. From more details on the generalized subdifferential we refer to Clarke [6] and Gasiński-Papageorgiou [10] .

The critical point theory for smooth functions uses a compactness type condition known as the Cerami condition. In our nonsmooth setting, the condition takes the following form.

Definition 3.2.1. *We say that f satisfies the "nonsmooth Cerami condition" (nonsmooth C-condition for short), if any sequence $\{x_n\}_{n\geqslant 1} \subseteq X$ such that $\{f(x_n)\}_{n\geqslant 1}$ is bounded and $(1+\|x_n\|)m(x_n) \to 0$ as $n \to \infty$, where $m(x_n) = \min\{\|x^*\|_* : x^* \in \partial f(x_n)\}$, has a strongly convergent subsequence.*

3.2.2 Mountain pass theorem

The first theorem is due to Chang [5] and extends to a nonsmooth setting the well known "mountain pass theorem" due to Ambrosetti -Rabinowitz [1] .

Theorem 3.2.1. *If X is a reflexive Banach space, $R : X \to \mathbb{R}$ is a locally Lipschitz functional satisfying C-condition and for some $\rho > 0$ and $y \in X$ such that $\|y\| > \rho$, we have*

$$\max\{R(0), R(y)\} < \inf_{\|x\|=\rho} \{R(x)\} =: \eta,$$

then R has a nontrivial critical point $x \in X$ such that the critical value $c = R(x) \geqslant \eta$ is characterized by the following minimax principle

$$c = \inf_{\gamma\in\Gamma} \max_{0\leqslant\tau\leqslant1} \{R(\gamma(\tau))\},$$

where $\Gamma = \{\gamma \in \mathcal{C}([0, 1], X) : \gamma(0) = 0, \gamma(1) = y\}$.

3.2.3 Generalized Lebesgue-Sobolev spaces

By $S(\Omega)$ we denote the set of all measurable real-valued function defined on $\mathbb{R}^N$. We define

$$L^{p(x)}(\Omega) = \{u \in S(\Omega) : \int_\Omega |u(x)|^{p(x)} dx < \infty\}.$$

We furnish $L^{p(x)}(\Omega)$ with the following norm (known as the Luxemburg norm)

$$\|u\|_{p(x)} = \|u\|_{L^{p(x)}(\Omega)} = \inf \Big\{\lambda > 0 : \int_\Omega \Big|\frac{u(x)}{\lambda}\Big|^{p(x)} dx \leqslant 1\Big\}.$$

Also we introduce the variable exponent Sobolev space

$$W^{1,p(x)}(\Omega) = \{u \in L^{p(x)}(\Omega) : |\nabla u| \in L^{p(x)}(\Omega)\},$$

and we equip it with the norm

$$\|u\| = \|u\|_{W^{1,p(x)}(\Omega)} = \|u\|_{p(x)} + \|\nabla u\|_{p(x)}.$$

By $W_0^{1,p(x)}(\Omega)$ we denote the closure of $C_0^\infty(\Omega)$ in $W^{1,p(x)}(\Omega)$ (see f.e. Fan-Zhao [8, 9]).

Lemma 3.2.1 (Fan-Zhao [8])**.** *If $\Omega \subset \mathbb{R}^N$ is an open domain, then*
(a) the spaces $L^{p(x)}(\Omega)$, $W^{1,p(x)}(\Omega)$ and $W_0^{1,p(x)}(\Omega)$ are separable and reflexive Banach spaces;
(b) the space $L^{p(x)}(\Omega)$ is uniformly convex;
(c) if $1 \leqslant q(x) \in \mathcal{C}(\overline{\Omega})$ and $q(x) \leqslant p^*(x)$ (respectively $q(x) < p^*(x)$) for any $x \in \overline{\Omega}$, where

$$p^*(x) = \begin{cases} \frac{Np(x)}{N-p(x)} & p(x) < N \\ \infty & p(x) \geqslant N, \end{cases}$$

then $W^{1,p(x)}(\Omega)$ *is embedded continuously (respectively compactly) in* $L^{q(x)}(\Omega)$*;*
(d) Poincaré inequality in $W_0^{1,p(x)}(\Omega)$ *holds i.e., there exists a positive constant* c *such that*

$$\|u\|_{p(x)} \leqslant c\|\nabla u\|_{p(x)} \qquad \text{for all } u \in W_0^{1,p(x)}(\Omega);$$

(e) $(L^{p(x)}(\Omega))^* = L^{p'(x)}(\Omega)$*, where* $\frac{1}{p(x)} + \frac{1}{p'(x)} = 1$ *and for all* $u \in L^{p(x)}(\Omega)$ *and* $v \in L^{p'(x)}(\Omega)$*, we have*

$$\int_\Omega |uv|dx \leqslant \Big(\frac{1}{p^-} + \frac{1}{p'^-}\Big)\|u\|_{p(x)}\|v\|_{p'(x)}.$$

3.2.4 Energy functional

Consider the following function

$$J(u) = \int_\Omega \frac{1}{p(x)}|\nabla u|^{p(x)}dx, \qquad \text{for all } u \in W_0^{1,p(x)}(\Omega).$$

We know that $J \in \mathcal{C}^1(W_0^{1,p(x)}(\Omega))$ and operator $-\text{div}(|\nabla u|^{p(x)-2}\nabla u)$, is the derivative operator of J in the weak sense. We denote

$$A = J' : W_0^{1,p(x)}(\Omega) \to (W_0^{1,p(x)}(\Omega))^*,$$

then

$$\langle Au, v\rangle = \int_\Omega |\nabla u(x)|^{p(x)-2}(\nabla u(x), \nabla v(x))dx, \tag{3.5}$$

for all $u, v \in W_0^{1,p(x)}(\Omega)$.

Lemma 3.2.2. *If* A *is the operator defined above, then* A *is a continuous, bounded, strictly monotone and maximal monotone operator of type* (S_+) *i.e., if* $u_n \to u$ *weak in* $W_0^{1,p(x)}(\Omega)$ *and* $\limsup_{n\to\infty}\langle Au_n, u_n - u\rangle \leqslant 0$*, implies that* $u_n \to u$ *in* $W_0^{1,p(x)}(\Omega)$*.*

We introduce locally Lipschitz energy functional $R : W_0^{1,p(x)}(\Omega) \to \mathbb{R}$ defined by

$$R(u) = \int_\Omega \frac{1}{p(x)}|\nabla u(x)|^{p(x)}dx - \int_\Omega \frac{\lambda}{p(x)}|u(x)|^{p(x)}dx - \int_\Omega j(x, u(x))dx,$$

for all $u \in W_0^{1,p(x)}(\Omega)$.

3.3 Existence of solution

3.3.1 Assumptions

We start by introducing our assumptions for the nonsmooth potential $j(x,t)$.

Let us consider the hypothesis
$H(j)$ $j : \Omega \times \mathbb{R} \to \mathbb{R}$ be a function such that $j(x,t)$ satisfies $j(x,0) = 0$ almost everywhere on Ω and
(i) for all $t \in \mathbb{R}$, the function $\Omega \ni x \to j(x,t) \in \mathbb{R}$ is measurable;
(ii) for almost all $x \in \Omega$, the function $\mathbb{R} \ni t \to j(x,t) \in \mathbb{R}$ is locally Lipschitz;
(iii) for almost all $x \in \Omega$ and all $v \in \partial j(x,t)$, we have

$$|v| \leq a(x) + c_1 |t|^{r(x)-1},$$

with $a \in L^\infty_+(\Omega), c_1 > 0$ and $r \in \mathcal{C}(\overline{\Omega})$ such that $p^+ < r^- := \min_{x\in\overline{\Omega}} r(x) \leqslant r(x) \leqslant r^+ := \max_{x\in\overline{\Omega}} r(x) < \widehat{p}^*$;
(iv) there exists $\mu > 0$ such that

$$\limsup_{t\to 0} \frac{j(x,t)}{|t|^{p(x)}} \leqslant -\mu,$$

uniformly for almost all $x \in \Omega$;
(v) we have

$$\limsup_{|t|\to\infty} \frac{v^* t - j(x,t)}{|t|^{p(x)}} \leqslant 0,$$

uniformly for almost all $x \in \Omega$ and all $v^* \in \partial j(x,t)$;
(vi) there exists $\overline{u} \in W_0^{1,p(x)}(\Omega) \setminus \{0\}$, such that

$$\frac{1}{p^-}\int_\Omega |\nabla \overline{u}(x)|^{p(x)} dx + \frac{\lambda_-}{p^-}\int_\Omega |\overline{u}(x)|^{p(x)} dx \leqslant \int_\Omega j(x,\overline{u}(x))dx,$$

where $\lambda_- := \max\{0, -\lambda\}$.

3.3.2 Cerami condition

Lemma 3.3.1. *If hypothesis $H(j)$ hold and $\lambda \in (-\infty, \frac{(p^- -1)p^+}{(p^+ -1)p^-}\lambda_*)$, then R satisfies the nonsmooth C-condition.*

Proof. Let $\{u_n\}_{n\geqslant 1} \subseteq W_0^{1,p(x)}(\Omega)$ be a sequence such that $\{R(u_n)\}_{n\geqslant 1}$ is bounded and $(1 + \|u_n\|)m(u_n) \to 0$ as $n \to \infty$. We will show that $\{u_n\}_{n\geqslant 1} \subseteq W_0^{1,p(x)}(\Omega)$ is bounded.

Because $|R(u_n)| \leqslant M$ for all $n \geqslant 1$, we have

$$-M \leqslant \int_\Omega \frac{1}{p(x)}|\nabla u_n(x)|^{p(x)}dx - \int_\Omega \frac{\lambda}{p(x)}|u_n(x)|^{p(x)}dx - \int_\Omega j(x,u_n(x))dx. \tag{3.6}$$

Since $\partial R(u_n) \subseteq (W_0^{1,p(x)}(\Omega))^*$ is weakly compact, nonempty and the norm functional is weakly lower semicontinuous in a Banach space, then we can find $u_n^* \in \partial R(u_n)$ such that $\|u_n^*\|_* = m(u_n)$, for $n \geqslant 1$.

Consider the operator $A : W_0^{1,p(x)}(\Omega) \to (W_0^{1,p(x)}(\Omega))^*$ defined by (3.5). Then, for every $n \geqslant 1$, we have

$$u_n^* = Au_n - \lambda|u_n|^{p(x)-2}u_n - v_n^*, \tag{3.7}$$

where $v_n^* \in \partial\psi(u_n) \subseteq L^{p'(x)}(\Omega)$, for $n \geqslant 1$, with $\frac{1}{p(x)} + \frac{1}{p'(x)} = 1$ and $\psi : W_0^{1,p(x)}(\Omega) \to \mathbb{R}$ is defined by $\psi(u_n) = \int\limits_\Omega j(x,u_n(x))dx$.

From the choice of the sequence $\{u_n\}_{n\geqslant 1} \subseteq W_0^{1,p(x)}(\Omega)$, at least for a subsequence, we have

$$|\langle u_n^*, w\rangle| \leqslant \frac{\varepsilon_n\|w\|}{1+\|u_n\|} \quad \text{for all } w \in W_0^{1,p(x)}(\Omega), \tag{3.8}$$

with $\varepsilon_n \searrow 0$. Putting $w = u_n$ in (3.8) and using (3.7), we obtain

$$-\varepsilon_n \leqslant -\int_\Omega |\nabla u_n(x)|^{p(x)}dx + \lambda\int_\Omega |u_n(x)|^{p(x)}dx + \int_\Omega v_n^*(x)u_n(x)dx. \tag{3.9}$$

Now, let us consider two cases.

Case 1.

Let $\lambda < 0$. We define $\lambda_- := \max\{0, -\lambda\}$.
From (3.6) and (3.9), we have

$$\begin{aligned} -M - \varepsilon_n \quad &\leqslant \quad \Big(\frac{1}{p^-} - 1\Big)\int_\Omega |\nabla u_n(x)|^{p(x)}dx + \lambda_-\Big(\frac{1}{p^-} - 1\Big)\int_\Omega |u_n(x)|^{p(x)}dx \\ &\quad + \int_\Omega v_n^*(x)u_n(x)dx - \int_\Omega j(x,u_n(x))dx. \end{aligned} \tag{3.10}$$

By virtue of hypotheses $H(j)(v)$, we know that for some sufficiently small constant $c > 0$ there exist constant $L_1 > 0$, such that

$$v^*t - j(x,t) \leqslant c|t|^{p(x)}, \tag{3.11}$$

uniformly for almost all $x \in \Omega$ and all t such that $t \geqslant L_1$.

On the other hand, from the Lebourg mean value theorem and hypothesis $H(j)(iii)$, for almost all $x \in \Omega$ and all $t \in \mathbb{R}$ such that $|t| < L_1$, we have

$$|j(x,t)| \leqslant c_3, \tag{3.12}$$

for some $c_3 > 0$. Therefore, from (3.11) and (3.12) it follows that for almost all $x \in \Omega$ and all $t \in \mathbb{R}$, we have

$$v^* t - j(x,t) \leqslant c|t|^{p(x)} + c_4, \tag{3.13}$$

for some $c_4 > 0$. From (3.10), we obtain

$$\lambda_-\Big(1 - \frac{1}{p^-}\Big)\int_\Omega |u_n(x)|^{p(x)}dx \leqslant M + \varepsilon_n + \int_\Omega v_n^*(x)u_n(x)dx - \int_\Omega j(x, u_n(x))dx.$$

If we use (3.13), we get

$$\lambda_-\Big(1 - \frac{1}{p^-}\Big)\int_\Omega |u_n(x)|^{p(x)}dx \leqslant M + \varepsilon_n + c\int_\Omega |u_n(x)|^{p(x)}dx + \int_\Omega c_4\ dx,$$

for all $n \geqslant 1$, which leads to

$$\Big[\lambda_-\Big(1 - \frac{1}{p^-}\Big) - c\Big]\int_\Omega |u_n(x)|^{p(x)}dx \leqslant c_5,$$

for some $c_5 := M + \varepsilon_1 + c_4|\Omega| > 0$. Because c is some sufficiently small positive constant, so we know that $\lambda_-\Big(1 - \frac{1}{p^-}\Big) - c > 0$ and

$$\text{the sequence } \{u_n\}_{n\geqslant 1} \subseteq L^{p(x)}(\Omega) \text{ is bounded.} \tag{3.14}$$

Now, consider again (3.10), we obtain

$$\Big(1 - \frac{1}{p^-}\Big)\int_\Omega |\nabla u_n(x)|^{p(x)}dx \leqslant M + \varepsilon_n + \int_\Omega v_n^*(x)u_n(x)dx - \int_\Omega j(x, u_n(x))dx.$$

In a similar way, by using (3.13) we have

$$\Big(1 - \frac{1}{p^-}\Big)\int_\Omega |\nabla u_n(x)|^{p(x)}dx \leqslant M + \varepsilon_n + c\int_\Omega |u_n(x)|^{p(x)}dx + \int_\Omega c_4\ dx,$$

for all $n \geqslant 1$. From (3.14) we know that $\{u_n\}_{n\geqslant 1} \subseteq L^{p(x)}(\Omega)$ is bounded, so we have that

$$\text{the sequence } \{\nabla u_n\}_{n\geqslant 1} \subseteq L^{p(x)}(\Omega; \mathbb{R}^N) \text{ is bounded.} \tag{3.15}$$

From (3.14) and (3.15), we have that

$$\text{the sequence } \{u_n\}_{n\geqslant 1} \subseteq W_0^{1,p(x)}(\Omega) \text{ is bounded.}$$

Case 2.

Now, let $\lambda \geqslant 0$.
Again from (3.6) and (3.9), we have

$$-M - \varepsilon_n \leqslant \Big(\frac{1}{p^-} - 1\Big) \int_\Omega |\nabla u_n(x)|^{p(x)} dx + \lambda\Big(1 - \frac{1}{p^+}\Big) \int_\Omega |u_n(x)|^{p(x)} dx + \int_\Omega v_n^*(x) u_n(x) dx - \int_\Omega j(x, u_n(x)) dx. \quad (3.16)$$

From the definition of λ_*, we get

$$\lambda_* \int_\Omega |u_n(x)|^{p(x)} dx \leqslant \int_\Omega |\nabla u_n(x)|^{p(x)} dx, \quad (3.17)$$

for all $n \geqslant 1$. Using this fact in (3.16), we have

$$\Big[\lambda_*\Big(1 - \frac{1}{p^-}\Big) + \lambda\Big(\frac{1}{p^+} - 1\Big)\Big] \int_\Omega |u_n(x)|^{p(x)} dx \leqslant M + \varepsilon_n + \int_\Omega v_n^*(x) u_n(x) dx - \int_\Omega j(x, u_n(x)) dx. \quad (3.18)$$

In a similar way like in Case 1, by using (3.13) in (3.18), we obtain

$$\Big[\lambda_*\Big(1 - \frac{1}{p^-}\Big) + \lambda\Big(\frac{1}{p^+} - 1\Big) - c\Big] \int_\Omega |u_n(x)|^{p(x)} dx \leqslant c_6,$$

for some $c_7 := M + \varepsilon_1 + c_4|\Omega| > 0$. Fro sufficiently small $c > 0$, we know that $\lambda_*\Big(1 - \frac{1}{p^-}\Big) + \lambda\Big(\frac{1}{p^+} - 1\Big) - c > 0$, hence

$$\text{the sequence } \{u_n\}_{n\geqslant 1} \subseteq L^{p(x)}(\Omega) \text{ is bounded.} \quad (3.19)$$

In analogous way, from (3.13), (3.16) and (3.19), we have that

$$\text{the sequence } \{\nabla u_n\}_{n\geqslant 1} \subseteq L^{p(x)}(\Omega; \mathbb{R}^N) \text{ is bounded.} \quad (3.20)$$

From (3.19) and (3.20), we have that

$$\text{the sequence } \{u_n\}_{n\geqslant 1} \subseteq W_0^{1,p(x)}(\Omega) \text{ is bounded.}$$

From Cases 1 and 2, we have that

$$\text{the sequence } \{u_n\}_{n\geq 1} \subseteq W_0^{1,p(x)}(\Omega) \text{ is bounded.}$$

Hence, by passing to a subsequence if necessary, we may assume that

$$\begin{array}{ll} u_n \to u & \text{weakly in } W_0^{1,p(x)}(\Omega), \\ u_n \to u & \text{in } L^{p(x)}(\Omega), \end{array} \quad (3.21)$$

for some $u \in W^{1,p(x)}(\Omega)$. Putting $w = u_n - u$ in (3.8) and using (3.7), we obtain

$$\Big| \langle Au_n, u_n - u\rangle - \lambda \int_\Omega |u_n(x)|^{p(x)-2} u_n(x)(u_n - u)(x)dx -$$

$$\int_\Omega v_n^*(x)(u_n - u)(x)dx \Big| \leqslant \varepsilon_n,$$

with $\varepsilon_n \searrow 0$. If we pass to the limit as $n \to \infty$, we get

$$\limsup_{n\to\infty} \langle Au_n, u_n - u\rangle \leqslant 0$$

(for details see Barnaś [4, 3]). So from Lemma 3.2.2, we have that $u_n \to u$ in $W_0^{1,p(x)}(\Omega)$ as $n \to \infty$. Thus R satisfies the C-condition. □

We have the following Lemma

Lemma 3.3.2. *If hypothesis $H(j)$ hold, $\lambda < \frac{p^-}{p^+}\lambda_*$ and $\theta \in (r^+, \widehat{p}^*)$, then there exist $\beta_1, \beta_2 > 0$ such that for all $u \in W_0^{1,p(x)}(\Omega)$ with $\|u\| < 1$, we have*

$$R(u) \geqslant \beta_1 \|u\|^{p^+} - \beta_2 \|u\|^\theta.$$

The proof you can find in Barnaś [2]).

3.3.3 Main theorem

Using Lemmata 3.3.1 and 3.3.2, we can prove the following existence theorem for problem (P).

Theorem 3.3.1. *If hypothesis $H(j)$ hold and $\lambda < \widetilde{p}\lambda_*$, then problem (P) has a nontrival solution.*

Proof. From Lemma 3.3.2 we know that there exist $\beta_1, \beta_2 > 0$, such that for all $u \in W_0^{1,p(x)}(\Omega)$ with $\|u\| < 1$, we have

$$R(u) \geqslant \beta_1 \|u\|^{p^+} - \beta_2 \|u\|^\theta = \beta_1 \|u\|^{p^+} \Big(1 - \frac{\beta_2}{\beta_1} \|u\|^{\theta - p^+}\Big).$$

Since $p^+ < \theta$, if we choose $\rho \in (0,1)$ small enough, we will have that $R(u) \geqslant L$, for all $u \in W_0^{1,p(x)}(\Omega)$, with $\|u\| = \rho$ and some $L > 0$.

Now, let $\overline{u} \in W_0^{1,p(x)}(\Omega)$ be as in hypothesis $H(j)(vi)$. We have

$$\begin{aligned} R(\overline{u}) &= \int_\Omega \frac{1}{p(x)}|\nabla\overline{u}(x)|^{p(x)}dx - \int_\Omega \frac{\lambda}{p(x)}|\overline{u}(x)|^{p(x)}dx - \int_\Omega j(x,\overline{u}(x))dx \\ &\leqslant \frac{1}{p^-}\int_\Omega |\nabla\overline{u}(x)|^{p(x)}dx + \frac{\lambda_-}{p^-}\int_\Omega |\overline{u}(x)|^{p(x)}dx - \int_\Omega j(x,\overline{u}(x))dx. \end{aligned}$$

From hyphothesis $H(j)(vi)$, we get $R(\overline{u}) \leqslant 0$. This permits the use of Theorem 3.2.1 which gives us $u \in W_0^{1,p(x)}(\Omega)$ such that $R(u) > 0 = R(0)$ and $0 \in \partial R(u)$. From the last inclusion we obtain

$$0 = Au - \lambda|u|^{p(x)-2}u - v^*,$$

where $v^* \in \partial\psi(u)$. Hence

$$Au = \lambda|u|^{p(x)-2}u + v^*,$$

so for all $v \in \mathcal{C}_0^\infty(\Omega)$, we have $\langle Au, v\rangle = \lambda\langle|u|^{p(x)-2}u, v\rangle + \langle v^*, v\rangle$. So we have

$$\begin{aligned} &\int_\Omega |\nabla u(x)|^{p(x)-2}(\nabla u(x), \nabla v(x))_{\mathbb{R}^N}\,dx \\ = &\int_\Omega \lambda|u(x)|^{p(x)-2}u(x)v(x)dx + \int_\Omega v^*(x)v(x)dx, \end{aligned}$$

for all $v \in \mathcal{C}_0^\infty(\Omega)$. From the definition of the distributional derivative we have

$$\begin{cases} -\Delta_{p(x)}u(x) - \lambda|u(x)|^{p(x)-2}u(x) \in \partial j(x,u(x)) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Therefore $u \in W_0^{1,p(x)}(\Omega)$ is a nontrivial solution of (P). □

Remark 3.3.1. *A nonsmooth potential satisfying hypothesis $H(j)$ is for example the one given by the following function:*

$$j_1(x,t) = \begin{cases} -\mu|t|^{p(x)} & \text{if} \quad |t| \leqslant 1, \\ (\mu + \sigma - |2|^{p(x)})|t| - 2\mu - \sigma + |2|^{p(x)} & \text{if} \quad 1 < |t| \leqslant 2, \\ \ln|t| + \sigma - |2|^{p(x)} - \ln 2 & \text{if} \quad |t| > 2, \end{cases}$$

with $\mu, \sigma > 0$ and continuous function $p : \overline{\Omega} \to \mathbb{R}$ which satisfies $1 < p^- \leqslant p(x) \leqslant p^+ < \widehat{p}^*$.

References

[1] A. Ambrosetti and P. H. Rabinowitz. “Dual variational methods in critical point theory and applications”. In: *J. Funct. Anal.* 14 (1973), pp. 349–381.

[2] S. Barnaś. “Existence of a nontrival solution for Dirichlet problem involving $p(x)$–Laplacian”. In: *Discuss. Math., Differ. Incl. Control Optim.* 34, No. 1 (2014), pp. 15–39.

[3] S. Barnaś. “Existence result for differential inclusion with $p(x)$-Laplacian”. In: *Schedae Informaticae* 21 (2012), pp. 41–55.

[4] S. Barnaś. “Existence result for hemivariational inequality involving $p(x)$-Laplacian”. In: *Opusc. Math.* 32, No. 3 (2012), pp. 439–454.

[5] K. C. Chang. “Variational methods for non-differentiable functionals and their applications to partial differential equations”. In: *J. Math. Anal. Appl.* 80 (1981), pp. 102–129.

[6] F. H Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1993.

[7] X. L. Fan and Q. H. Zhang. “Existence of solutions for $p(x)$-Laplacian Dirichlet problem”. In: *Nonlinear Anal., Theory Methods Appl., Ser. A, Theory Methods* 52, No. 8 (2003), pp. 1843–1852.

[8] X. L. Fan and D. Zhao. “On the generalized Orlicz-Sobolev space $W^{k,p(x)}(\Omega)$”. In: *J. Gansu Educ. College (Natural Science)* 12 (1998), pp. 1–6.

[9] X. L. Fan and D. Zhao. “On the spaces $L^{p(x)}(\Omega)$ and $W^{m,p(x)}(\Omega)$”. In: *J. Math. Anal. Appl.* 263, No. 2 (2001), pp. 424–446.

[10] L. Gasiński and N. S. Papageorgiou. *Nonlinear Analysis*. Chapman & Hall/CRC, Boca Raton, FL., 2006.

Chapter 4

Estimation of stability index

ANTONI LEON DAWIDOWICZ AND BARBARA WODECKA

Antoni Leon Dawidowicz
Wydział Matematyki i Informatyki
Uniwersytet Jagielloński
e-mail: antoni.dawidowicz@im.uj.edu.pl

Barbara Wodecka
Instytut Matematyki
Uniwersytet Jana Kochanowskiego
e-mail: barbara.wodecka@ujk.edu.pl

Abstract

Stable distributions are an extension of normal distribution but, on the contrary, are able to describe heavy tails. A lack of closed form formulas for probability density can make a problem of estimating distribution parameters. In this paper we introduce an estimator of stability index α based on k-records values. First, we want to remind some properties of stable distribution, as well defined k-record values and their distributions. Secondly, we study the asymptotic behavior of presented estimator and we give some numerical results.

Contents

4.1 Stable distribution

Stable distribution was characterized by Paul Lévy in his study of sums of independent identically distributed terms in the 1920's. This family are a rich class of probability distributions that have many intriguing mathematical properties. Additionally, they generalize the normal distribution. Therefore they have been proposed as model for many types of physical and economic systems. Unfortunately, there is a major drawback to use them by practitioners, that is the lack of closes formulas for densities for almost all stable distributions (the exceptions are: normal, Cauchy and Levy distributions). Nevertheless, nowadays there are reliable computer programs to compute stable densities, distribution functions and quantiles. With these programs, it is possible to use α-stable models in a variety of practical problems.

Due to a great interest of α-stable distribution, there are several equivalent definitions of this distribution. The main definition includes an important property of normal random variable:

Property 4.1. *If X is normal, then for X_1 and X_2 independent copies of X and any positive numbers A and B,*

$$AX_1 + BX_2 \stackrel{\mathcal{D}}{=} CX + D, \tag{4.1}$$

for some positive number C and some $D \in \mathbb{R}$, where $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution.

This property allows us to define stable distribution in following way.

Definition 4.2. *A random variable X is said to have a stable distribution if for any positive numbers A and B, there is a positive number C and a real number D* (4.1) *holds, where X_1 and X_2 independent copies of X. The random variable is strictly stable if* (4.1) *holds with $D = 0$ for all choices of A and B.*

The equivalent definition of stable distribution is as follow.

Definition 4.3. *A random variable X is said to have a stable distribution if for any $n \geq 2$, there is a positive number C_n and a real number D_n such that*

$$X_1 + \ldots + X_n \stackrel{\mathcal{D}}{=} C_n X + D_n, \tag{4.2}$$

where $X_1, X_2, \ldots, X_n$ are independent copies of X.

The only possible choice for the scaling constants is $C_n = n^{1/\alpha}$ for some $\alpha \in (0, 2\rangle$.

The third definition states that stable distribution are the only distribution that can be obtain as limits or normalized sums of independent identically distributed (i.i.d.) random variables.

Definition 4.4. *A random variable X is said to have a stable distribution if it has a domain of attraction, i.e., if there is a sequence of i.i.d. random variables $Y_1, Y_2, \ldots$ and sequence of positive numbers $\{d_n\}$ and real numbers $\{a_n\}$, such that*

$$\frac{Y_1 + Y_2 + \ldots + Y_n}{d_n} + a_n \xrightarrow{\mathcal{D}} X, \tag{4.3}$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution.

The next definition describes all possible α-stable distributions in the most precise way.

Definition 4.5. *A random variable X is said to have a stable distribution if there are parameters $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\sigma > 0$, and $\mu \in \mathbb{R}$ such that its characteristic function has the following form:*

$$\mathbb{E}\exp(itX) = \begin{cases} \exp\left(-|t|^{\alpha}\left(1 - i\beta\,\mathrm{sgn}(t)\tan\left(\frac{\pi\alpha}{2}\right)\right)\right), & \alpha \neq 1 \\ \exp\left(-|t|\left(1 + i\beta\frac{2}{\pi}\,\mathrm{sgn}(t)\ln|t|\right)\right), & \alpha = 1 \end{cases}. \tag{4.4}$$

Definition 4.5 shows that a general stable distribution requires four parameters to describe: stability index (tail index, tail exponent or characteristic exponent) $\alpha \in (0, 2\rangle$, a skewness parameters $\beta \in \langle -1, 1\rangle$, a scale parameter $\sigma > 0$ and a location parameter $\mu \in \mathbb{R}$. Therefore, we will denote stable distributions by $S(\alpha, \beta, \mu, \sigma)$ and write

$$X \sim S(\alpha, \beta, \mu, \sigma)$$

to indicate that X has the stable distribution $S(\alpha, \beta, \mu, \sigma)$. The index of stability α is treated as a constant that characterized distribution. Accordingly such distributions are called α-stable distributions.

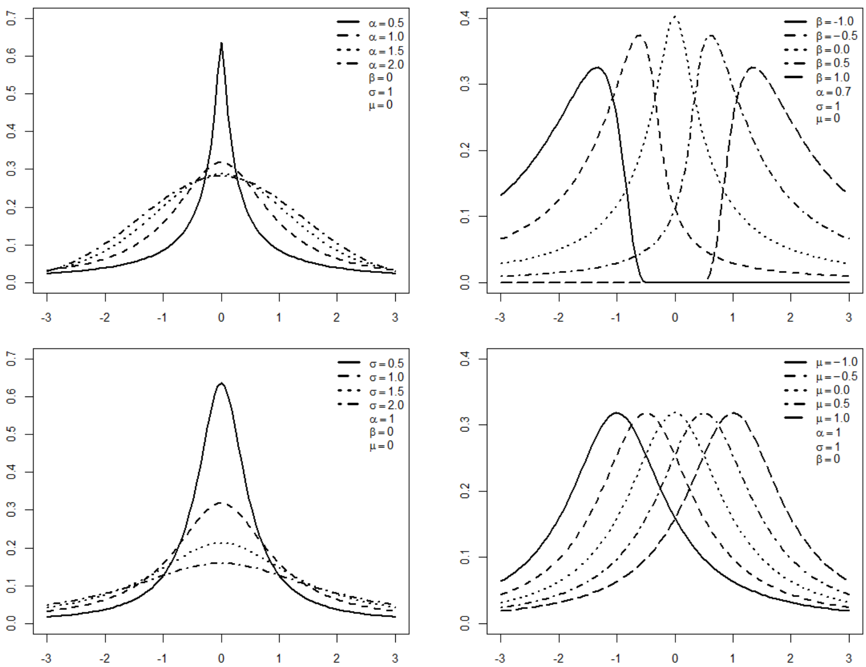


Figure 4.1: Influence of parameters $(\alpha, \beta, \mu, \sigma)$ on the probability density of α-stable distributions.

The probability density of α-stable random variables exist and are continuous but, with a few exceptions, the are not known in closed form. The exceptions are:

1. Normal distribution $\mathcal{N}(\mu, 2\sigma^2) = S(2, \beta, \mu, \sigma)$, whose density is:

$$f(x) = \frac{1}{2\sigma\sqrt{\pi}} \exp\left\{-\frac{(x-\mu)^2}{4\sigma^2}\right\}.$$

   Note that although the value β in not specified, one typically associates the normal distribution with the choice $\beta = 0$.

2. Cauchy distribution $\mathcal{C}(\mu, \sigma) = S(1, 0, \mu, \sigma)$, whose density is:

$$f(x) = \frac{1}{\pi} \cdot \frac{\sigma}{\sigma^2 + (x-\mu)^2}.$$

3. Lévy distribution $\mathcal{L}(\mu, \sigma) = S(\frac{1}{2}, 1, \mu, \sigma)$, whose density is:

$$f(x) = \sqrt{\frac{\sigma}{2\pi}} \cdot \frac{1}{(x-\mu)^{3/2}} \cdot \exp\left\{-\frac{\sigma}{2(x-\mu)}\right\} \cdot \mathbb{1}_{(\mu,\infty)}(x).$$

Stable densities are supported on either the whole real line or a half line. The latter situations can only occur when $\alpha < 1$ and $|\beta| = 1$.

Lemma 4.6. *The support of a stable distribution is*

$$\text{supp} f(\cdot|\alpha, \beta, \mu, \sigma) = \begin{cases} \langle \mu, \ +\infty) & \alpha < 1 \textit{ and } \beta = 1 \\ (-\infty, \ \mu\rangle & \alpha < 1 \textit{ and } \beta = -1 \\ (-\infty, \ +\infty) & \textit{otherwise} \end{cases}. \tag{4.5}$$

When $\alpha = 2$, in other words for normal distribution, asymptotic tail properties are well known. But what if $\alpha < 2$?

For $\alpha < 2$ stable distributions have one tail (when $\alpha < 1$ and $|\beta| = 1$) or both tails (otherwise), that are asymptotically equivalent to a Pareto law. The following property describes the tail property behave of stable distributions.

Property 4.7. *Let $X \sim S(\alpha, \beta, \mu, \sigma)$ with $0 < \alpha < 2$. Then*

$$\begin{aligned} \lim_{x\to\infty} x^\alpha \mathbb{P}(X > x) &= C_\alpha \frac{1+\beta}{2} \sigma^\alpha, \\ \lim_{x\to\infty} x^\alpha \mathbb{P}(X < -x) &= C_\alpha \frac{1-\beta}{2} \sigma^\alpha, \end{aligned} \tag{4.6}$$

where

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x \, dx\right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)}, & \textit{if} \quad \alpha \neq 1 \\ \frac{2}{\pi}, & \textit{if} \quad \alpha = 1 \end{cases}. \tag{4.7}$$

Tail behavior (4.6) is a widely used property of α-stable distributions. Since $\mathbb{E}|X|^r = \int_0^\infty \mathbb{P}(|X|^r > x)dx$, we obtain:

Property 4.8. *Let $X \sim S(\alpha, \beta, \mu, \sigma)$ with $0 < \alpha < 2$. Then*

$$\begin{aligned} \mathbb{E}|X|^p &< \infty \quad \textit{for any} \quad 0 < p < \alpha, \\ \mathbb{E}|X|^p &= \infty \quad \textit{for any} \quad p \geq \alpha. \end{aligned}$$

The fact that α-stable distributions with $0 < \alpha < 2$ do not have finite second moments or variance causes some to immediately dismiss stable distributions as being irrelevant to any practical problem. When $\alpha \leq 1$ the mean of X is undefined, because $\mathbb{E}|X| = \infty$. On the other hand, when $1 < \alpha \leq 2$, $\mathbb{E}|X| < \infty$ and the mean of X is given below.

Property 4.9. *When $1 < \alpha \leq 2$, the shift parameter μ equals the mean.*

4.2 Records

Record values and associated statistics arise naturally in many real life applications involving data related to economics, sports, weather, meteorology, hydrology, largest insurance claims, and others wherein only record values may be recorded. This was the main cause of creating the record theory. The first person who introduced the study of records values and documented many of basic properties of records was K. N. Chandler in 1952. His achievement attracted the attention of many researchers and inspired many new publications. The theory of records cannot be separate from the theory of order statistics because records are especially closely related to extremal order statistics. Now we introduce the formal definitions of record values.

Let $X_1, X_2, \dots$ be an infinite sequence of i.i.d. random variables with cumulative distribution function F.

Definition 4.10. *An observation X_j is called an upper record value (or simply record) if its value exceeds that of all previous observations. The record time sequence $\{T_n, n \geq 0\}$ is defined in the following manner:*

$$T_0 = 1 \quad \textit{with probability} \;\; 1$$

and, for $n \geq 1$

$$T_n = \min\{j : X_j > X_{T_{n-1}}\}. \tag{4.8}$$

The record value sequence $\{R_n\}$ is defined by:

$$R_n = X_{T_n}, \quad n = 0,\ 1,\ 2,\ \dots\ . \tag{4.9}$$

An analogous definition deals with lower record value.

Now we assume that F is absolutely continuous with corresponding probability function f.

Theorem 4.11. *The distribution function of nth record value has the form*

$$F_{R_n}(r) = \mathbb{P}(R_n \leq r) = 1 - (1 - F(r)) \sum_{j=0}^{n} \frac{1}{j!} \left(-\ln(1 - F(r))\right)^n. \tag{4.10}$$

We may differentiate the above expression to obtain the probability function for R_n.

Corollary 4.12. *The probability density of the record value is of the form*

$$f_{R_n}(x) = \frac{1}{n!} \left(-\ln(1 - F(x))\right)^n f(x). \tag{4.11}$$

There are several situations where the second and third largest values are of special interest, e.g. insurance claims some non-life insurance. Thus, the k-th record values have become an extension of the records theory. In 1976 Dziubdziela and Kopociński defined the notion of the k-th record values as follows.

Let $X_1, X_2, \ldots$ be an infinite sequence of i.i.d. random variables with cumulative distribution function F, continuous with density f. Denote by $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ the order statistics in the considered sequence.

Definition 4.13. *Let k be a positive integer. The sequences of the k-th times are defined by*

$$T_{0(k)} = k \quad \textit{with probability} \;\; 1$$

and, for $n \geq 1$

$$T_{n(k)} = \min\{j : \; j > T_{(n-1)(k)}, \; X_j > X_{T_{(n-1)(k)}-k+1:T_{(n-1)(k)}}\}, \tag{4.12}$$

and then the sequence of k-th records values $\{R_{n(k)}, n \geq 0\}$ by

$$R_{n(k)} = X_{T_{n(k)}-k+1:T_{n(k)}} \quad n = 0,\ 1,\ 2,\ \ldots\ . \tag{4.13}$$

In other words, by eliminating repetitions in the non-decreasing sequence of k-th order statistics a strictly increasing subsequence is obtained and it is called a sequence of k-th record values

$$X_{1:k} < X_{T_{2(k)}-k+1:T_{2(k)}} < X_{T_{3(k)}-k+1:T_{3(k)}} < \ldots\ . \tag{4.14}$$

Nevzorov (1986) extended the above definition to such k-th record values that $\{k = k_n, n \geq 1\}$ is a sequence of integers.

Let us consider the sequence of random variables $\{R_{n(k)}, n \geq 0\}$.

Theorem 4.14. *The probability density of the k-th record value is given by*

$$f_{R_n(k)}(r) = \frac{k}{n!}\left(-\ln(1-F(r))^k\right)^n (1-F(r))^{k-1} f(r). \tag{4.15}$$

Corollary 4.15. *The cumulative distribution function of k-th record value has the form*

$$F_{R_n(k)}(r) = \mathbb{P}(R_n(k) \leq r) = \int_0^{-k\ln(1-F(r))} \frac{w^n}{n!} e^{-w} dw. \tag{4.16}$$

4.3 Defining the estimator

The estimation of stability index α is difficult because of the lack of closed formulas for densities functions. Many traditional methods (e.g. method of moments) cannot be used due to the non-existence of moments. However, there are numerical methods of estimation of α useful in practice. One of them does not assume a parametric form for the entire distribution, but focuses only on the tail behavior. In this paper we want to present the estimator which is based on the relation $\alpha = \gamma^{-1}$, where γ is the extreme value index (EVI)[1], but only for $\alpha < 2$ since tails of α-stable distributions are asymptotically equivalent to tails of Pareto distribution (see Property 4.7).

Let $X_1, X_2, \ldots$ be an infinite sequence of i.i.d. random variables with cumulative distribution function F, continuous with density f. Denote by: $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ the order statistics associated to the sample $X_1, X_2, \ldots, X_n$, and $R_{1(k)}, R_{2(k)}, \ldots, R_{n(k)}$ a sequence of k-th record values in the sequence $X_1, X_2, \ldots$. Assume that F belongs to the domain of attraction of the extreme value distributions

$$G_\gamma(x) = \exp\left(-(1+\gamma x)^{-1/\gamma}\right), \quad 1+\gamma x > 0, \tag{4.17}$$

with γ real and where for $\gamma = 0$ the right-hand side is interpreted as $\exp(-e^{-x})$. The parameter γ is called the extreme value index.

We define the estimator of stability index α as follows.

Definition 4.16. *Let $\{k_n, n \geq 1\}$ be a sequence of positive integers such that $1 \leq k_n < n$. The estimator of the tail index α based on the k-th record values has the form*

$$\hat{\alpha}_n = (\hat{\gamma}_n)^{-1} = \left(\ln \frac{R_{n(k)} - R_{(n-k)(k)}}{R_{(n-k)(k)} - R_{(n-2k)(k)}}\right)^{-1}, \tag{4.18}$$

where $k = k_n$.

Before proofing the asymptotic behavior of $\hat{\alpha}_n$ let us define necessary terms.

Denote by $h(x) = -\ln(1-F(x))$, $x \in \mathbb{R}$, the hazard function associated to F and by $H(t) = h^{\leftarrow}(t) = inf\{s : h(s) \geq t,\ t \in \mathbb{R}_+\}$, its generalized inverse function.

We shall also need the notation of Π-variation.

Definition 4.17. *A measurable function $g : \mathbb{R}_+ \to \mathbb{R}$ is Π-varying ($g \in \Pi(a)$) if there exists a function $a : \mathbb{R}_+ \to \mathbb{R}$ such that for $x > 0$,*

$$\lim_{t\to\infty} \frac{g(tx) - g(t)}{a(t)} = \ln x. \tag{4.19}$$

[1]Definition of the extreme value index can be found in [5]

Throughout this section we use the following assumptions on the sequence $\{k_n, n \geq 1\}$:

$$k_n \to \infty, \qquad k_n/n \to 0, \qquad \text{as } n \to \infty \tag{4.20}$$
$$k_n/\ln n \to \infty, \qquad k_n/n \to 0, \qquad \text{as } n \to \infty \tag{4.21}$$

In order to simpify the formulas, we write k instead of k_n whenever (4.20) or (4.21) are assumed. Our main results are the following.

Theorem 4.18. *Under the assumption* (4.17) *and* (4.20), $\hat{\alpha}_n \xrightarrow[n\to\infty]{\mathbb{P}} \alpha$, *where* $\xrightarrow{\mathbb{P}}$ *denotes convergence in probability.*

Theorem 4.19. *Under the assumption* (4.17) *and* (4.21), $\hat{\alpha}_n \xrightarrow[n\to\infty]{a.s.} \alpha$, *where* $\xrightarrow{a.s.}$ *denotes convergence almost surely.*

Theorem 4.20. *Assume that the function H has a positive derivative H'. Then under the assumption $\pm t^{-\gamma}H'(\ln t) \in \Pi(a)$ for some positive function a*

$$k^{1/2}(\hat{\alpha}_n - \alpha) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{(e^{2/\alpha}+1)\alpha^2}{(e^{1/\alpha}-1)^2}\right) \tag{4.22}$$

as $n \to \infty$, for any sequence $k = k_n \to \infty$ satisfying $k_n/n \cdot R^{\leftarrow}(n) \to 0$ $(n \to \infty)$, where $R(n) = t\,(e^{-t\gamma}H'(t)/a(t))^2$.

Now we shall use proofs consider in article [1] to proof the above theorems.

Proof of Theorem 4.18. From Theorem 2.1. in [1] we have $\hat{\gamma}_n \xrightarrow[n\to\infty]{\mathbb{P}} \gamma$.
In that case, for any $\varepsilon > 0$, we get $\mathbb{P}(|\hat{\gamma}_n - \gamma| > \varepsilon) \xrightarrow[n\to\infty]{} 0$. Thus, we need to check if $\mathbb{P}(|\hat{\alpha}_n - \alpha| > \varepsilon) \xrightarrow[n\to\infty]{} 0$ to proof the thesis.
Let $\varepsilon > 0$.

$$\mathbb{P}(|\hat{\alpha}_n - \alpha| > \varepsilon) = \mathbb{P}(|\frac{1}{\hat{\gamma}_n} - \frac{1}{\gamma}| > \varepsilon) = 1 - \mathbb{P}(-\varepsilon \leq \frac{1}{\hat{\gamma}_n} - \frac{1}{\gamma} \leq \varepsilon) =$$
$$1 - \mathbb{P}(\frac{1}{\gamma} - \varepsilon \leq \frac{1}{\hat{\gamma}_n} \leq \frac{1}{\gamma} + \varepsilon) = 1 - \mathbb{P}(\frac{\gamma}{1+\varepsilon\gamma} \leq \hat{\gamma}_n \leq \frac{\gamma}{1-\varepsilon\gamma}) =$$
$$1 - \mathbb{P}(\frac{-\varepsilon\gamma}{1+\varepsilon\gamma} \leq \hat{\gamma}_n - \gamma \leq \frac{\varepsilon\gamma}{1-\varepsilon\gamma}) =$$
$$\mathbb{P}(|\hat{\gamma}_n - \gamma| > \frac{\varepsilon\gamma}{1+\varepsilon\gamma}) \xrightarrow[n\to\infty]{} 0$$

□

Proof of Theorem 4.19. From Theorem 2.2. in [1] we have $\hat{\gamma}_n \xrightarrow[n\to\infty]{a.s.} \gamma$.
In that case, we get

$$
\begin{aligned}
1 =&\mathbb{P}(\{\omega \in \Omega : \lim_{n\to\infty} \hat{\gamma}_n(\omega) = \gamma\}) = \mathbb{P}(\{\omega \in \Omega : \lim_{n\to\infty} \frac{1}{\hat{\gamma}_n(\omega)} = \frac{1}{\gamma}\}) = \\
&\mathbb{P}(\{\omega \in \Omega : \lim_{n\to\infty} \hat{\alpha}_n(\omega) = \alpha\})
\end{aligned}
$$

□

Proof of Theorem 4.20. From Theorem 2.3. in [1] we have

$$k^{1/2}(\hat{\gamma}_n - \gamma) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \frac{(e^{2\gamma}+1)\gamma^2}{(e^{\gamma}-1)^2}\right).$$

We write

$$k^{1/2}(\hat{\alpha}_n - \alpha) = k^{1/2}(\frac{1}{\hat{\gamma}_n} - \frac{1}{\gamma}) = k^{1/2} \cdot \frac{\gamma - \hat{\gamma}_n}{\gamma\hat{\gamma}_n} = -\frac{1}{\gamma\hat{\gamma}_n} \cdot k^{1/2}(\hat{\gamma}_n - \gamma).$$

Using Slucky's theorem we obtain $1/(\gamma\hat{\gamma}_n) \to 1/\gamma^2$ as $n \to \infty$.
This leads us to

$$k^{1/2}(\hat{\alpha}_n - \alpha) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \frac{(e^{2/\alpha}+1)\alpha^2}{(e^{1/\alpha}-1)^2}\right),$$

which completes the proof. □

We shall notice that the expected number of k-records values and its variance are related to the sequence $k = k_n$. This relation in given by following lemma.

Lemma 4.21. *Denote by $N^{(k_n)}(n)$ the number of k-record values in the sequence $X_1, X_2, \ldots, X_n$. Then*

$$
\begin{aligned}
\mathbb{E}\left(N^{(k_n)}(n)\right) &= k_n \sum_{i=k_n}^{n} \frac{1}{i}, \\
Var\left(N^{(k_n)}(n)\right) &= k_n \sum_{i=k_n}^{n} \frac{1}{i} - k_n^2 \sum_{i=k_n}^{n} \frac{1}{i^2},
\end{aligned}
$$

for $n \geq 1$.

4.4 Simulation research

All pictures and simulations in this paper was done by means of R software ([6],[2],[7]).

To illustrate the finite-sample behavior of the proposed estimator we give some simulation result for the standard symmetric α-stable distribution, i.e. $\beta = 0$, $\mu = 0$, $\sigma = 1$, with $\alpha \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9\}$.
We generate pseudorandom independent and identically distributed samples of size $n = 5000$ and replicate them $J = 1000$ times independently for each value of α. The series of stability index estimators are calculated for all k in rang from 1 to 500 with respect to each simulated series. For each independent estimates $\hat{\alpha}_1, \ldots, \hat{\alpha}_J$ obtained from the estimator of α, quantile lines in orders 0.1, 0.3, 0.5, 0.7, 0.9 are determined.

Figures 4.2, 4.3, and 4.4 present simulation research results in graphical form for stable distribution with several selected values of α.

Table 4.1 presents arithmetic means and medians for each independent estimates $\hat{\alpha}_1, \ldots, \hat{\alpha}_J$ obtained from the estimator of α.

4.5 Conclusion

The carried simulation study leads to the following conclusions.

For $\alpha < 1$ values of each estimators are underestimated, but with the increase of the estimated parameter estimators values start to be significantly overestimated. And the closer to $\alpha = 2$ we are the larger values we get. This fact should not be surprising. In this paper we propose an estimator based on the relation $\alpha = \gamma^{-1}$. We have to keep in mind that for normal distribution the extreme value index is equal to 0. Hence, stable distribution with α close to 2 are close to normal distribution and the extreme value index tends to 0. Therefore α tends to infinity. However, this fact should not worry us too much. Most practitioners want a tool for heavy-tailed distributions that is for $\gamma > 0$. In other words, the estimator proposed in this paper can be an useful tool for detection tail heaviness of distribution.

Problematic becomes the optimal choice of k. We cannot choose to large value of k (the larger value of k the greater values and instability of the estimator) or to small (the smaller value of k the greater instability of the estimator). Averaging α's estimators seems to be a good idea, but there is a problem too. Namely, what rang of k we should choose. The selection of the optimal k has been subject of studies, but it has not yet been completely solved; see for instance [3] or [4].

	$k = 40$		$k = 80$		$k = 120$	
α	mean	median	mean	median	mean	median
0.1	0.0987	0.1004	0.0963	0.0972	0.0952	0.0957
0.2	0.1948	0.1983	0.1946	0.1961	0.1909	0.1923
0.3	0.2927	0.2999	0.2914	0.2947	0.2855	0.2881
0.4	0.3899	0.4018	0.3869	0.3924	0.3842	0.3872
0.5	0.4844	0.5020	0.4915	0.4982	0.4873	0.4908
0.6	0.5856	0.6159	0.5882	0.6014	0.5839	0.5897
0.7	0.6883	0.7222	0.6815	0.7039	0.6869	0.6975
0.8	0.7866	0.8392	0.7913	0.8133	0.8026	0.8115
0.9	0.8857	0.9594	0.8956	0.9233	0.9125	0.9286
1.0	0.9658	1.0656	0.9974	1.0367	1.0257	1.0530
1.1	1.0649	1.1999	1.1098	1.1572	1.1662	1.2019
1.2	1.1753	1.3051	1.2367	1.3005	1.3094	1.3559
1.3	1.2725	1.5387	1.3641	1.3486	1.4910	1.5765
1.4	1.3669	2.0498	1.5089	1.6552	1.7495	1.9017
1.5	1.5186	1.9137	1.7897	2.0272	2.2649	2.5652
1.6	1.5676	1.9619	2.2186	10.0336	3.4623	5.7992
1.7	1.8618	2.4652	3.2917	5.4600	4.9301	5.5004
1.8	2.4332	0.5085	4.5366	14.7054	-2.8281	1.1272
1.9	3.0725	1.0131	-4.2608	-2.9002	-5.2600	-2.5021

Table 4.1: Simulation results of the estimation of the stability index α for some fixed value of k.

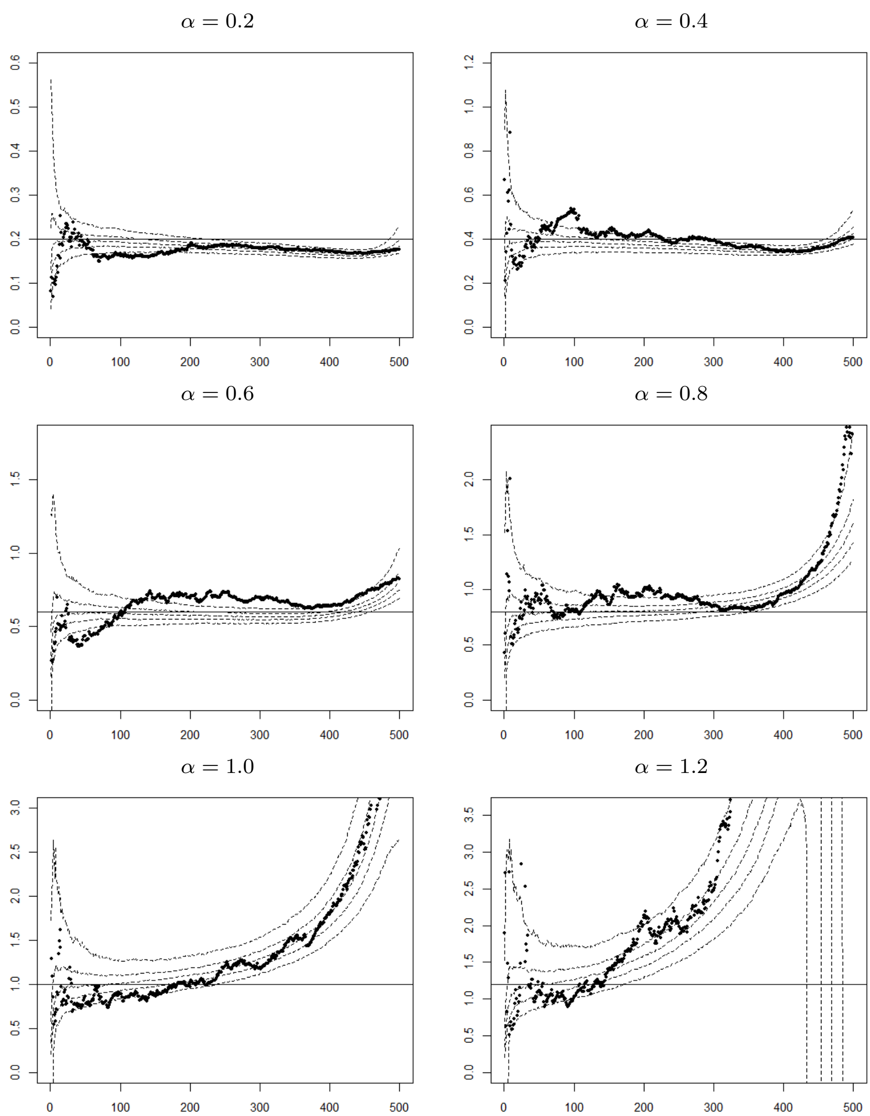


Figure 4.2: Panels present quantile lines (dashed lines) and one randomly selected series for the exemplary values of $\hat{\alpha}_n$ (black points) for all k going on horizontal axes. The horizontal lines represents the true value of α.

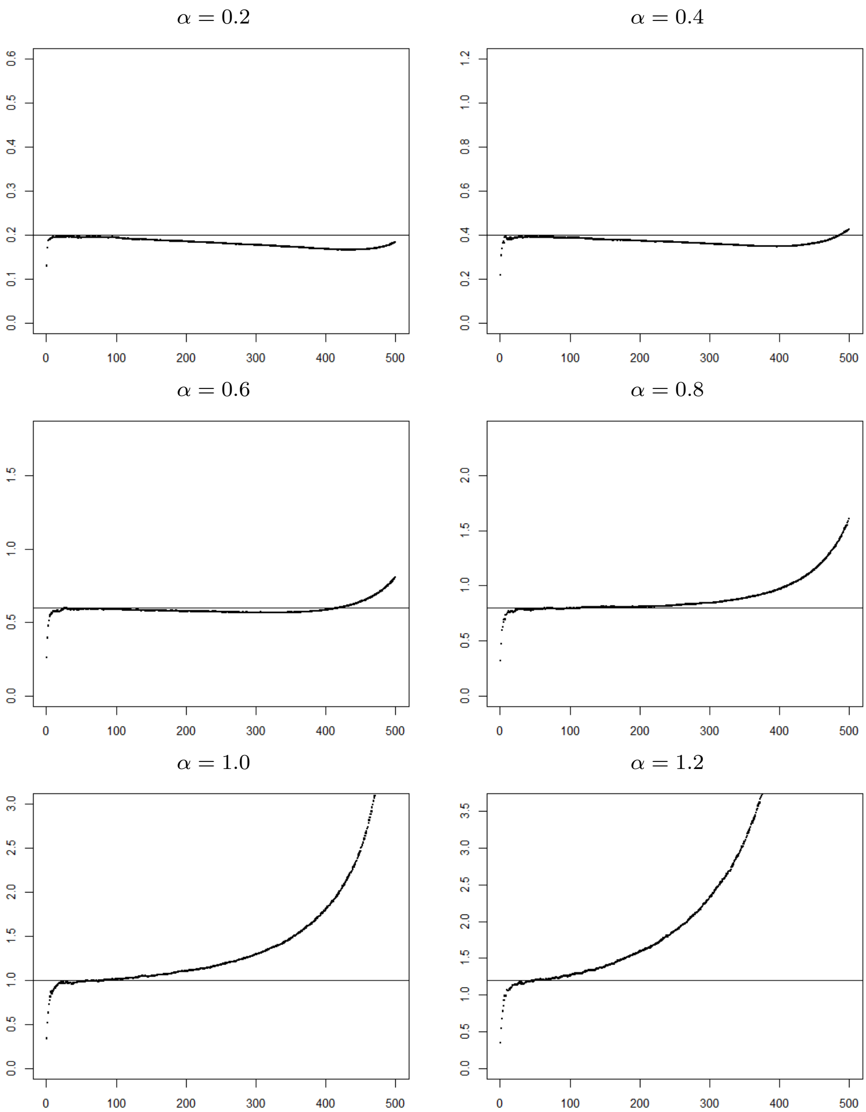


Figure 4.3: Plots of the median of $\hat{\alpha}_n$ (black points) for all k going on horizontal axes. The horizontal lines represents the true value of α.

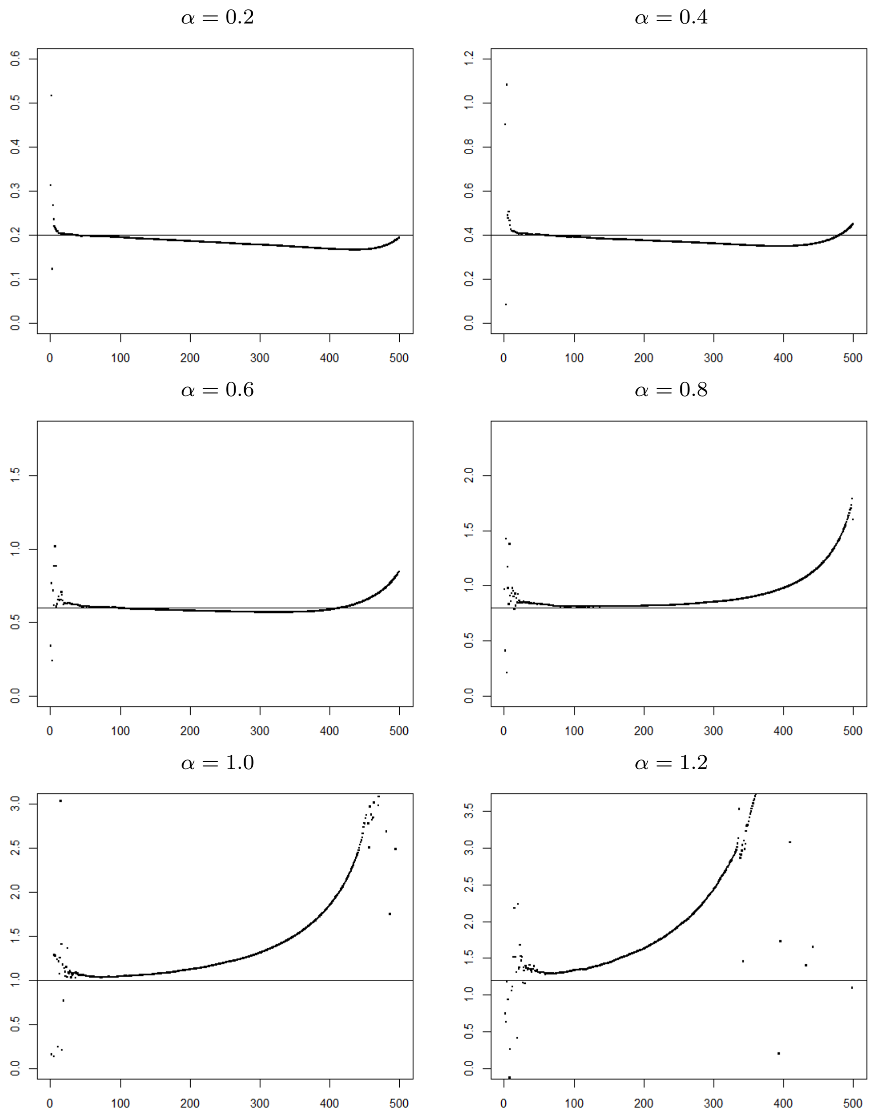


Figure 4.4: Plots of the mean value of $\hat{\alpha}_n$ (black points) for all k going on horizontal axes. The horizontal lines represents the true value of α.

References

[1] M. Berred. "K-record values and the extreme-value index". In: *J. Stat. Plann. Inference* 45, No. 1-2 (1995), pp. 49–63.

[2] M. Chrapek. *Records: Record Values and Record Times*. R package version 1.0. 2012. URL: http://CRAN.R-project.org/package=Records.

[3] M. Chrapek, M. Stachura, and B. Wodecka. "Estymacja indeksu ekstremalnego w oparciu o k-te wartości rekordowe – sugestia poprawy jakości estymacji". In: *Metody Matematyczne, Ekonometryczne i Komputerowe w Finansach i Ubezpieczeniach 2010*. Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, 2012, pp. 9–28.

[4] W. Dziubdziela, M. Stachura, and B. Wodecka. "Extreme value index of left and right tails for financial time series". In: *Acta Universitatis Lodziensis, Folia Oeconomica* 255 (2011), pp. 227–238.

[5] L. de Haan and A. Ferreira. *Extreme Value Theory. An Introduction*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006.

[6] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2012. URL: http://www.R-project.org/.

[7] D. Wuertz, M. Maechler, and Rmetrics core team members. *stabledist: Stable Distribution Functions*. R package version 0.6-6. 2013. URL: http://CRAN.R-project.org/package=stabledist.

Chapter 5

Cycles through matchings in bipartite graphs

Grzegorz Gancarzewicz

Grzegorz Gancarzewicz
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: ggancarzewicz@gmail.com

Abstract

We give several results on hamiltonian cycles through matchings in bipartite graphs. We are focused on conditions with degree sum of two independent vertices.

Contents

5.1 Introduction

We consider only finite simple graphs i.e. graphs without loops and multiple edges. By V or V(G) we denote the vertex set of the graph G and respectively by E or E(G) the edge set of G. By d(x) we denote *the degree of a vertex x in the graph G* and by d(x, y) or $\mathrm{d}_G(x, y)$ *the distance between x and y in G.*

In a graph G a path that contains every vertex of G is called *a hamiltonian path of G* and similarly a cycle that contains every vertex of G is called *a hamiltonian cycle of G.* A graph is *hamiltonian* iff it contains a hamiltonian cycle.

Such paths and cycles are named in memory of Hamilton[1], who in 1856 described a mathematical game on the dodecahedron in which one person sticks five pins in any five consecutive vertices and the other person have to complete the so formed path to a spanning cycle.

Next, the mathematicians were investigating a similar problem in arbitrary graphs and they called it *a hamiltonian problem.* In general the hamiltonian problem is a problem concerning with finding sufficient conditions under which the graph has a hamiltonian path or cycle.

The hamiltonian problem is one of the most known problems in graph theory.

Research works are extended to find cycles of given length (hamiltonian or not), containing given edges or vertices.

We are particularly interested in finding cycles (not necessarily hamiltonian) containing a given family of independent edges (*a matching*) for the special case of bipartite graphs.

5.2 Preliminary results

Let G be a graph, V(G) denotes the vertex set of G and E(G) denotes the edge set of G. By $\mathrm{d}_G(x)$ or d(x) we denote *the degree of the vertex x in the graph G* and δ or $\delta(G)$ denotes *the minimal degree of the graph G.*

We are particularly interested in degree conditions under which a graph G is hamiltonian. One of the first conditions of this type was obtained by G. A. Dirac [7] in 1952. He obtained the following result:

Theorem 5.2.1 ([7]). *Let G be a graph on n vertices. If*

$$\delta(G) \geq \frac{n}{2}, \tag{5.1}$$

[1] The description of the game is taken from [4].

then G is hamiltonian.

The condition (5.1) is called a *Dirac condition* for minimal degree.

Let us introduce the σ_k.

Definition 5.2.1. *Let G be a graph and* $k \geq 0$.
$$\sigma_k = \min\{\sum_{i=1}^{k} \mathrm{d}(x_i) : \{x_1, \ldots, x_k\} \subset \mathrm{V}(G) \text{ and independent } \}$$

Eight years after Dirac, O. Ore [10] proved the following:

Theorem 5.2.2 ([10]). *If G is a graph on* $n \geq 3$ *vertices satisfying*

$$\sigma_2 \geq n, \tag{5.2}$$

then G is hamiltonian.

The condition (5.2) is called *Ore condition.* Later many *Ore type theorems* and *Dirac type theorems* have been established.

Ore condition for a graph G :

$$\sigma_2 \geq l$$

can be also written as:

$$\text{If } x, y \in \mathrm{V}(G),\ xy \notin \mathrm{E}(G), \quad \text{then:} \quad \mathrm{d}(x) + \mathrm{d}(y) \geq l.$$

Note that Ore condition is weaker than Dirac condition i.e.

Remark 5.2.1. *Let G be a graph and* $k \geq 0$. *If G satisfies Dirac condition*

$$\delta(G) \geq \frac{k}{2},$$

then G also satisfies Ore condition:

$$\sigma_2 \geq k.$$

Hence any Ore type theorem has as corollary a Dirac type theorem.

Let G_1 and G_2 be two graphs. The graph $G_1 * G_2$ is a graph obtained from G_1 and G_2 by adding all edges between $\mathrm{V}(G_1)$ and $\mathrm{V}(G_2)$, i.e. $\mathrm{V}(G_1 * G_2) = \mathrm{V}(G_1) \cup \mathrm{V}(G_2)$ and $\mathrm{E}(G_1 * G_2) = \mathrm{E}(G_1) \cup \mathrm{E}(G_2) \cup \{uv : u \in \mathrm{V}(G_1),\ v \in \mathrm{V}(G_2)\}$.
The graph $G_1 * G_2$ is called *the join of graphs G_1 and G_2*.

In [9] Ore gave two examples of nonhamiltonian graphs of maximal size (i.e. with as many edges as possible):

$$G = K_1 * (K_1 \cup K_{n-2}), \quad n \geq 3$$

and

$$G = \overline{K}_3 * K_2, \quad n = 5.$$

Next J.A. Bondy [3] and V. Chvátal [5] proved that this is a complete list of nonhamiltonian graphs of maximal size i.e.

Theorem 5.2.3 ([3, 5]). *If G is a nonhamiltonian graph on n ($n \geq 3$) vertices of maximal size then:*

$$G = K_1 * (K_1 \cup K_{n-2}), \quad n \geq 3$$

or

$$G = \overline{K}_3 * K_2, \quad n = 5.$$

Note that for any $n \geq 3$ the graph G from Theorem 5.2.3 satisfies $\sigma_2 = n - 1$, hence the bound for σ_2 in Theorem 5.2.2 is best possible.

There is also a similar condition, proved by V. Chvátal in [6], under which the graph G has a hamiltonian path.

Theorem 5.2.4 ([6]). *If G is a graph on $n \geq 3$ vertices satisfying*

$$\sigma_2 \geq n - 1, \tag{5.3}$$

then G has a hamiltonian path.

A graph $G = (\mathrm{V}, \mathrm{E})$ is *a bipartite graph* if its vertex set can be divided into two disjoint sets B and W (B and W are independent sets) such that every edge is joining a vertex from B to one in W. Vertex set B and W are called *the partite sets.* In case of bipartite graph we will write $G = (B, W; \mathrm{E})$.

Similar results where proved for the special case of bipartite graphs and the bound for degree sum in the case of bipartite graphs is about half of its value for graphs.

Let $G = (B, W; \mathrm{E})$ be a bipartite graph and let $S \subset \mathrm{V}(G)$, then $S_B = S \cap B$ and $S_W = S \cap W$. We shall say that S is *balanced* iff $|S_B| = |S_W|$. A bipartite graph $G = (B, W; \mathrm{E})$ is *balanced* iff $\mathrm{V}(G)$ is balanced i.e. $|B| = |W|$.

An analogue of Theorem 5.2.2 for bipartite graphs was proved by J. Moon and M. Moser [8] in 1963.

Theorem 5.2.5 ([8])**.** *Let $G = (B, W; \mathrm{E})$ be a balanced bipartite graph of order $2n$. If for all nonadjacent vertices $x \in B$ and $y \in W$ we have:*

$$\mathrm{d}(x) + \mathrm{d}(y) \geq n + 1\,, \tag{5.4}$$

then G is hamiltonian.

The condition (5.4) is a generalization of Ore condition for bipartite graphs.

Note that the bound for $\mathrm{d}(x) + \mathrm{d}(y)$ is best possible, since the graph $G = K_{\frac{n}{2},\frac{n}{2}} + K_{\frac{n}{2},\frac{n}{2}}$ is a balanced nonhamiltonian bipartite graph such that for all nonadjacent vertices x and y from different partite sets we have $\mathrm{d}(x) + \mathrm{d}(y) \geq n$.

5.3 Cycles through matchings in simple graphs

Let G be a graph and let $k \geq 1$. We shall call a set of k independent edges a *k-matching* or a *matching.* A matching M of a graph G is a *perfect matching of G* iff every vertex from $\mathrm{V}(G)$ is incident to M.

R. Häggkvist in [11] proved the first Ore type theorem about hamiltonian cycles through perfect matchings:

Theorem 5.3.1 ([11])**.** *Let G be graph of even order n satisfying*

$$\sigma_2 \geq n + 1.$$

Then every perfect matching lies in a hamiltonian cycle.

He conjectured that if $\sigma_2 \geq n+1$, then any matching lies in a cycle. Häggkvist's conjecture was proved by K.A. Berman in [2].

Theorem 5.3.2 ([2]). *Let G be graph of order n satisfying*

$$\sigma_2 \geq n+1.$$

Then every matching lies in a cycle.

From Theorem 5.3.1 or Theorem 5.3.2, if a graph G satisfies $\sigma_2 \geq n+1$ then any perfect matching or a matching which can be extended to a perfect matching is contained in a hamiltonian cycle. However the condition $\sigma_2 \geq n+1$ does not imply that every matching is extendable to a perfect matching.

Let n and s be to integers such that $1 \leq s < n$ and $(n-s)$ is even. Consider a family of graphs:

$$G_{n,s} = \overline{K}_s * K_{n-s} \tag{5.5}$$

with a $\frac{n-s}{2}$-matching M contained in K_{n-s}. (M is a perfect matching in the graph K_{n-s}.) The graph G satisfies $\sigma_2 = 2(n-s)$ and M is a maximal matching which cannot be extended to a perfect matching of G. Hence for $\frac{n+2}{3} \leq s \leq \frac{n-1}{2}$ we have a family of graphs satisfying $\sigma_2 \geq n+1$ with a non-extendable matching and the cycle containing M from Theorem 5.3.2 is not certainly hamiltonian.

On matchings and the condition σ_2 we can cite two results of M. Las Vergnas [12].

Theorem 5.3.3 ([12]). *Let G be a graph on $n \geq 2$ vertices and let k be an integer satisfying $1 \leq k \leq \frac{n}{2}$. If G satisfies*

$$\sigma_2 \geq 2k-1\,,$$

then G contains a k-matching.

Theorem 5.3.4 ([12]). *Let G be a graph on $n \geq 2$ vertices and let k and l be two integers satisfying $0 \leq l < k \leq \frac{n}{2}$. If G satisfies*

$$\sigma_2 \geq 2(l+k)-1\,,$$

then every l-matching of G lies in a k-matching of G (or in other words every l-matching can be extended to a k-matching).

From Theorem 5.3.3 we know that if G satisfies $\sigma_2 \geq n+1$, then there is a perfect matching in G, but using Theorem 5.3.4 we can extend to a perfect matching only a 1-matching.

About graphs with every k-matching in a hamiltonian cycle or path Las Vergnas obtained the following two results:

Theorem 5.3.5 ([12]). *Let G be a graph on $n \geq 3$ vertices and let k be an integer such that $0 \leq k \leq \frac{n}{2}$. If G satisfies*

$$\sigma_2 \geq n + k - 1\,,$$

then every k-matching of G lies in a hamiltonian path.

Theorem 5.3.6 ([12]). *Let G be a graph on $n \geq 3$ vertices and let k be an integer such that $0 \leq k \leq \frac{n}{2}$. If G satisfies*

$$\sigma_2 \geq n + k\,,$$

then every k-matching of G lies in a hamiltonian cycle.

In 1983 Wojda [13] proved the following generalization of a result of Häggkvist [11]:

Theorem 5.3.7 ([13]). *Let G be a graph on $n \geq 3$ vertices, satisfying*

$$\sigma_2 \geq \frac{4n-4}{3}\,.$$

Then every matching of G lies in a hamiltonian cycle or $G \in \mathcal{G}_n$, where $\mathcal{G}_n$ is defined above.

Note that if a graph G from $\mathcal{G}_n$ satisfies $\sigma_2 \geq \frac{4n-4}{3}$, then in fact for this graph we have $\sigma_2 = \frac{4n-4}{3}$, and so we have a the following corollary:

Corollary 5.3.1. *Let G be a graph on $n \geq 3$ vertices, satisfying*

$$\sigma_2 > \frac{4n-4}{3}\,.$$

Then every matching of G lies in a hamiltonian cycle.

Note that if M is a perfect matching or a matching which can be extended to a perfect matching and the graph G satisfies $\sigma_2 \geq n+1$, then from Theorem 5.3.2 M lies in hamiltonian cycle. So if G satisfies $\sigma_2 \geq \frac{4n-4}{3}$ then from Theorem 5.3.3 there is a perfect matching in G and from Theorem 5.3.4 for $l \leq \frac{n+5}{6}$, every l-matching can be extended to a perfect matching but for big l it is not sure.

If $n > 7$ and $\frac{n+2}{3} < s < n$, then graph $G_{n,s}$ defined by (5.5) satisfies $\sigma_2 > \frac{4n-4}{3}$ and has a matching of size at least $\frac{n-1}{3}$ which cannot be extended to a perfect matching. Thus Theorems 5.3.2 and 5.3.7 are independent.

Obviously for $k \geq \frac{n-2}{3}$ Theorem 5.3.7 is better than Theorem 5.3.6.

5.4 Results for bipartite graphs

We consider only simple bipartite graphs i.e. bipartite graphs without loops and multiple edges. In 1972, M. Las Vergnas [12] obtained several results on matchings in bipartite graphs.

Theorem 5.4.1 ([12]). *Let* $G = (B, W; E)$ *be a bipartite graph and* $0 \leq k \leq \max\{|B|, |W|\}$. *If for any* $x \in B$, $y \in W$, $xy \notin E$ *we have* $\mathrm{d}(x) + \mathrm{d}(y) \geq k$, *then there is a* k*-matching in* G.

Theorem 5.4.2 ([12]). *Let* $G = (B, W; E)$ *be a bipartite graph and* $0 \leq l \leq k \leq \max\{|B|, |W|\}$. *If for any* $x \in B$, $y \in W$, $xy \notin E$ *we have* $\mathrm{d}(x) + \mathrm{d}(y) \geq l + k$, *then every* l*-matching is contained in a* k*-matching in* G.

For the existence of a perfect matching, he gave the sufficient condition:

Theorem 5.4.3 ([12]). *Let* $G = (B, W; E)$ *be a balanced bipartite graph of order* $2n$ *and let* $q \geq 2$. *If for any* $x \in B$, $y \in W$, $xy \notin E$ *we have* $\mathrm{d}(x) + \mathrm{d}(y) \geq n + q$, *then every matching of cardinality* q *is contained in a perfect matching.*

and an analogue of Theorem 5.3.2 for perfect matchings in bipartite graphs:

Theorem 5.4.4 ([12]). *Let* $G = (B, W; E)$ *be a balanced bipartite graph of order* $2n$. *If for any* $x \in B$, $y \in W$, $xy \notin E$ *we have* $\mathrm{d}(x) + \mathrm{d}(y) \geq n + 2$, *then every perfect matching in* G *is contained in a hamiltonian cycle.*

Using these two results he obtained the following Corollary:

Corollary 5.4.1 ([12]). *Let* $G = (B, W; E)$ *be a balanced bipartite graph of order* $2n$ *and let* $q \geq 2$. *If for any* $x \in B$, $y \in W$, $xy \notin E$ *we have* $\mathrm{d}(x) + \mathrm{d}(y) \geq n + q$, *then every matching of cardinality* q *is contained in a hamiltonian cycle.*

Theorem 5.4.2 is an analogue of Theorem 5.3.4 and 5.4.1 is an analogue of Theorem 5.3.7.

D. Amar, E. Flandrin, G. Gancarzewicz and A.P. Wojda gave sufficient conditions in a balanced bipartite graph for a matching to be contained in a hamiltonian cycle or a cycle not necessarily hamiltonian [1]. Moreover, for the hamiltonian case the condition is almost best possible.

Results are proved in Sections 5.5.2 and 5.5.3.

Theorem 5.4.5 ([1]). *Let* $G = (B, W; \mathrm{E})$ *be a balanced bipartite graph of order* $2n$.

1. *If for any $x \in B$, $y \in W$, $xy \notin E$ we have*

$$\mathrm{d}(x) + \mathrm{d}(y) > \frac{4n}{3}\,,$$

then every matching M in G is contained in a hamiltonian cycle.

2. *If $n > 4$ and for any $x \in B$, $y \in W$, $xy \notin E$ we have*

$$\mathrm{d}(x) + \mathrm{d}(y) \geq \frac{5n}{4}\,,$$

then every matching M in G is contained in a cycle of G.

The first part of the theorem is almost best possible in the sense that if one decreases the sum of degrees of more than $\frac{2}{3}$ then the theorem is no more true. By $\overline{K}_{l,l}$ we denote the balanced bipartite graph of order $2l$ with empty edge set.

Let $\overline{K}_{p+1,p+1} = (B_{p+1}, W_{p+1}, \mathrm{E}_{p+1})$ and $K_{2p+1,2p+1} = (B_{2p+1}, W_{2p+1}, \mathrm{E}_{2p+1})$.

Consider the following bipartite graph $G = (B, W; \mathrm{E})$ with $B = B_{p+1} \cup B_{2p+1}$, $W = W_{p+1} \cup W_{2p+1}$ and $\mathrm{E} = \mathrm{E}_{2p+1} \cup \{uv : u \in B_{p+1}, v \in W_{2p+1}\} \cup \{uv : u \in W_{p+1}, v \in B_{2p+1}\}$.

Note that G is a balanced bipartite graph of order $2n = 2(3p+2)$. Let M be a perfect matching of $K_{2p+1,2p+1}$. It is evident that there is no hamiltonian cycle containing M and that the minimum sum of degrees of two nonadjacent vertices is $\frac{4n-2}{3}$.

Let now G' be the graph obtained from G by replacing $\overline{K}_{p+1,p+1}$ by $\overline{K}_{p,p}$. Then G' is a balanced bipartite graph which satisfies the hypothesis of the part (1) of Theorem 5.4.5 and by consequence there is a hamiltonian cycle which contains M. Notice however that M is not contained in any perfect matching of G', and the degree constraint in part (2) of Theorem 5.4.5 is clearly not sufficient to imply that any matching can be extended into a perfect matching.

Thus Theorems 5.4.5 and 5.4.4 are independent.

5.4.1 Conjecture

During works on the proof of Theorem 5.4.5, D. Amar posed the following conjecture:

Conjecture:
Let $G = (B, W; E)$ be a balanced bipartite graph of order $2n$. If for any $x \in B$, $y \in W$, $xy \notin E$ we have

$$\mathrm{d}(x) + \mathrm{d}(y) \geq n + 2\,,$$

then every matching M in G is contained in a cycle of G.

It is not difficult to show the following:

Remark: If $|M| = n-1$ and for any $x \in B$, $y \in W$, $xy \notin E$ $\mathrm{d}(x)+\mathrm{d}(y) \geq n+2$, then M is contained in a hamiltonian cycle.

Suppose that G is not a complete graph (if G is complete then **remark** is true.) Let $M \cup (pq)$, with $p \in B$, $q \in W$, $pq \notin E$ be a perfect matching containing M. From Theorem 5.4.4 it is contained in a hamiltonian cycle C. Let $D: \quad qu_1u_2...u_{2l}p$ be a hamiltonian path in G obtained from C by deleting the edge pq. The edges $u_1u_2, \dots, u_{2i+1}u_{2i+2}, \dots, u_{2l-1}u_{2l}$ are edges of the matching M. Since $\mathrm{d}(p)+\mathrm{d}(q) \geq n+2$ then there exists a k, such that $qu_{k+1} \in E$ and $pu_k \in E$. Note that $p \in B$, $q \in W$, $u_k \in W$ and then k is even. The edge u_ku_{k+1} is not in M. The cycle $C: \quad qu_1, ...u_kpu_lu_{l-1}...u_{k+1}q$ is a hamiltonian cycle of G which contains M.

5.5 Proof

We present the proof of Theorem 5.4.5 [1], as in this proof we use a concept of Θ-graph, which is very useful in proofs of results concerning cycles through specified edges of a graph.

5.5.1 Basic definitions and notation

We will give now some definitions and notation, not introduced in the previous section, that we will use in the proof.

Let G be a graph and H a subgraph of G.

Definition 5.5.1. *We use* $\mathrm{N}_G(H)$ *to denote the set of all vertices of the graph* G *which are adjacent to a vertex of the subgraph* H, *i.e.*

$$\mathrm{N}_G(H) = \{u \in \mathrm{V}(G) \,:\, \exists v \in \mathrm{V}(H) \text{ such that } uv \in \mathrm{E}(G)\}.$$

Consider an arbitrary vertex $x \in \mathrm{V}(G)$. We use $\mathrm{N}(x)$ to denote the set of all neighbors of the vertex x in G, i.e. $\mathrm{N}(x) = \{u \in \mathrm{V}(G) :\ xu \in \mathrm{E}(G)\}$.

If H is a subgraph of G or simply $H \subset \mathrm{V}(G)$, then $\mathrm{N}_H(x)$ denotes the set of all neighbors of the vertex x in H, i.e. $\mathrm{N}_H(x) = \{u \in \mathrm{V}(H) :\ xu \in \mathrm{E}(G)\}$ or $\mathrm{N}_H(x) = \{u \in H :\ xu \in \mathrm{E}(G)\}$.

Similarly the symbol $\mathrm{d}_H(x)$ denotes the number of neighbors of x in H i.e. $\mathrm{d}_H(x) = |\mathrm{N}_H(x)|$. Note that if H is a subgraph, then $\mathrm{d}_H(x)$ is *the degree of the vertex x in the subgraph H*.

Let D and F be two subgraphs of G or $D, F \subset \mathrm{V}(G)$. Then $e(D,F)$ denotes the number of edges of G with one vertex in D and the other in F. In other words, it is the number of edges between D and F. Let A be a subgraph of G and v a vertex of G and also $e(D,F) = \sum_{v \in D} \mathrm{d}_F(v) = \sum_{v \in F} \mathrm{d}_D(v)$.

In the proofs we will only use cycles and paths with a given orientation. For a cycle $C : c_1 \dots c_k$ or a path $P : p_1 \dots p_l$ we will use implicit orientation. Let C be a cycle or a path with a given orientation and let $x \in \mathrm{V}(C)$, then x^- is *the predecessor of* x and x^+ is its *successor* according to the orientation of C.

Definition 5.5.2. *Let C be a cycle or a path with a given orientation and let $x \in \mathrm{V}(G)$. Then $\mathrm{N}_C(x)^+ = \{u \in \mathrm{V}(C) : \ u^- \in \mathrm{N}_C(x)\}$.*

Let P be a path $p_1 \dots p_k$ and u, $v \in \mathrm{V}(G)$ such that up_1, $vp_k \in \mathrm{E}(G)$. Then uPv is the path $up_1 \dots p_k v$ and vPu is the path $vp_k \dots p_1 u$.

Definition 5.5.3. *Let $C : c_1 \dots c_l$ be a cycle in G with a given orientation. For any pair of vertices c_i, $c_j \in \mathrm{V}(C)$ with $i < j$ we can define four intervals:*

- $]c_i, c_j[$ *is the path* $c_{i+1} \dots c_{j-1}$.
- $[c_i, c_j[$ *is the path* $c_i \dots c_{j-1}$.
- $]c_i, c_j]$ *is the path* $c_{i+1} \dots c_j$.
- $[c_i, c_j]$ *is the path* $c_i \dots c_j$.

Observe that these four intervals are subgraphs of the cycle C.

A path P is *an even path* if $|V(P)|$ is even and is *an odd path* if $|V(P)|$ is odd.

Let G_1 and G_2 be subgraphs of G.

We say that G_1 and G_2 are *disjoint* if they have no vertex in common.

The union $G_1 \cup G_2$ of G_1 and G_2 is the subgraph with vertex set $\mathrm{V}(G_1 \cup G_2) = \mathrm{V}(G_1) \cup \mathrm{V}(G_2)$ and edge set $\mathrm{E}(G_1 \cup G_2) = \mathrm{E}(G_1) \cup \mathrm{E}(G_2)$.
If G_1 and G_2 are disjoint we sometimes denote their union by $G_1 + G_2$.

The intersection $G_1 \cap G_2$ of G_1 and G_2 is defined similarly, but in this case we assume that $\mathrm{V}(G_1) \cap \mathrm{V}(G_2) \neq \emptyset$.

Let $G = (B, W; E)$ be a balanced bipartite graph and M a matching in G.

A subgraph H of G is said to be a Θ-*graph compatible with* M if H is a union of two cycles C_1 and C_2 satisfying the conditions:

1. The intersection of C_1 and C_2 is a path R of length at least one.
2. Every edge of M is an edge of H.
3. Every edge of M incident with a vertex of R lies in R.
4. $|V(R)|$ is even and the end vertices, say x and y, of R are in different partite sets.

We denote $P: \quad x\, C_1 \setminus C_2\, y$, $Q: \quad x\, C_2 \setminus C_1\, y$ and $H = (P, Q, R)$.

The notion of the Θ-graph is based on the paper of K.A. Berman [2]. On Figure (5.1) there is an example of a Θ-graph.

A subgraph H of G is said to be *a strict* Θ-*graph compatible with* M if H is a Θ-graph (P, Q, R) such that if we label the vertices of the paths

$$\begin{aligned} P &: \quad xp_1 \;...\; p_\alpha y \\ Q &: \quad xq_1 \;...\; q_\beta y \\ R &: \quad xr_1 \;...\; r_\gamma y \end{aligned}$$

then $q_1 \in V(H) \setminus V(M)$, $p_\alpha \in V(H) \setminus V(M)$, $xr_1 \in M$, and $r_\gamma y \in M$.

On Figure (5.2) there is an example of a strict Θ-graph.

If on a path $\pi: \quad x_1x_2...x_k$ of $G = (B, W; E)$ is given an orientation from x_1 to x_k, π is said to be a BB-path if $x_1 \in B$, $x_k \in B$, a WW-path if $x_1 \in W$, $x_k \in W$, a BW-path if $x_1 \in B$, $x_k \in W$ and a WB-path if $x_1 \in W$, $x_k \in B$.

5.5.2 Proof of the part (1) of Theorem 5.4.5

Let $G = (B, W; E)$ be a bipartite graph satisfying the conditions of part (1) of Theorem 5.4.5 and let us suppose that there is a matching M in G such that there is no hamiltonian cycle through M. Without loss of generality we may suppose that:

(i) M is maximal, i.e. M is the only matching which contains M.

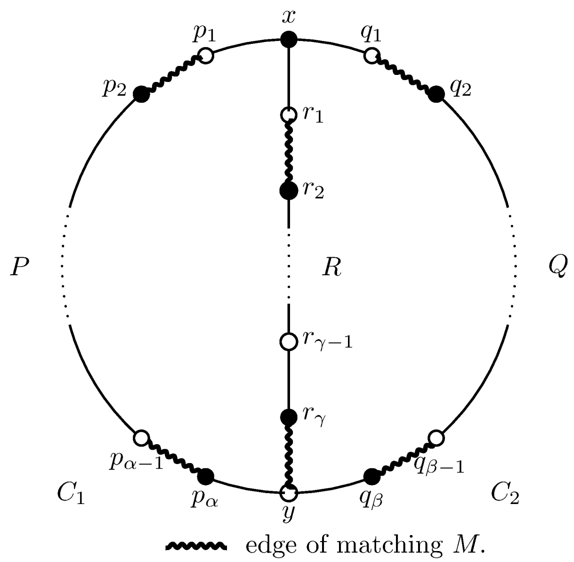


Figure 5.1: A Θ-graph compatible with M and containing all the vertices of the graph G.

(ii) G is maximal without a hamiltonian cycle through M (any addition of an edge uv, $u \in B$, $v \in W$, $uv \notin E$ creates a hamiltonian cycle containing M)

So we have a hamiltonian path P_H : $up_1 ... p_{2n-2}v$ containing M. Since $uv \notin E$, we have $\mathrm{d}(u) + \mathrm{d}(v) > \frac{4n}{3}$ and this implies that we have at least two vertices p_i, p_{i+1} satisfying up_{i+1}, $vp_i \in E$. Then the hamiltonian cycle:

$$C' : \quad up_{i+1}p_{i+2}...vp_ip_{i-1}...u$$

contains all edges of the path P_H except p_ip_{i+1}. Since there is no hamiltonian cycle containing M in G we have $p_ip_{i+1} \in M$. Now take the cycles: C_1 : $up_{i+1}p_{i-1}...u$ and C_2 : $vp_ip_{i+1}p_{i+2}...v$. The subgraph $H = C_1 \cup C_2$ is a Θ-graph compatible with M and containing all the vertices of the graph G. We can see an example of such Θ-graph which in not a strict Θ-graph in the Figure (5.1).

Following the notations from Section 5.5.1, label the vertices of the paths P, Q and R as follows:

$$\begin{aligned} P &: \; xp_1 \dots p_\alpha y \\ Q &: \; xq_1 \dots q_\beta y \\ R &: \; xr_1 \dots r_\gamma y \end{aligned}$$

and denote P_i, $i = 1, ..., n_P$, Q_j, $j = 1, ..., n_Q$ and R_k, $k = 1, ..., n_R$ the paths obtained respectively from P, Q, R by removal of the edges of M. Without loss of generality we may assume $x \in B$, $y \in W$.

We may assume that $H = (P, Q, R)$ is a $\Theta-$graph compatible with M such that $|V(R)|$ is maximum.

Remark Since M is maximal, for any i, j and k we have $2 \leq |V(P_i)| \leq 3$, $2 \leq |V(Q_j)| \leq 3$, and $1 \leq |V(R_k)| \leq 3$.

From the assumption that every edge of M incident with a vertex of R lies in R, if one of the edges $p_\alpha q_1$, $p_1 q_\beta$ exists then there is a hamiltonian cycle in G containing every edge of M, so we may assume $p_\alpha q_1 \notin E$ and $p_1 q_\beta \notin E$ and then we have:

$$\mathrm{d}(p_1) + \mathrm{d}(q_1) + \mathrm{d}(p_\alpha) + \mathrm{d}(q_\beta) > \frac{8n}{3} \tag{5.6}$$

Neighbors of $p_1, p_\alpha, q_1, q_\beta$ on Q and P.

Claim 5.5.1. *If $p_1 q_l \in E$ and $l > 1$ (p_1 and q_1 are in the same partite set), then $q_l q_{l+1} \in M$. Moreover for $i = 2, ...n_Q$, $e(p_1, Q_i) \leq 1$ and if $e(p_1, Q_i) = 1$, then $e(q_\beta, Q_i) = 0$.*

Proof of Claim 5.5.1:

In fact if $p_1 q_l \in E$, then $H' = (P', Q', R')$ with $P' : \; q_l p_1 p_2 ... p_\alpha y$, $Q' : q_l q_{l+1} ... q_\beta y$ and $R' : \; q_l q_{l-1} ... q_1 x r_1 r_2 ... r_\gamma y$ is a $\Theta-$graph compatible with M with $|V(R')| > |V(R)|$ unless $q_l q_{l+1} \in M$.

So let us suppose that $p_1 q_l \in E$ and $q_l q_{l+1} \in M$, with $q_l \in Q_{i_0}$. Then $q_l \in B$ for $p_1 \in W$. The vertex q_{l-1} is the only vertex of $V(Q_{i_0})$ in W.

If $q_\beta q_{l-1} \in E$ then the cycle:

$$C' : \quad q_{l-1} q_{l-2} ... q_1 x r_1 ... r_\gamma y p_\alpha p_{\alpha-1} ... p_1 q_l q_{l+1} ... q_\beta q_{l-1} \tag{5.7}$$

is a hamiltonian cycle of G containing M and Claim 5.5.1 is proved.

□

Claim 5.5.2. *$1 \le e(p_1, Q_1) \le 2$ and if $e(p_1, Q_1) = 1$ then $e(q_\beta, Q_1) \le 1$. If $e(p_1, Q_1) = 2$ then $e(q_\beta, Q_1) = 0$.*

Proof of Claim 5.5.2:

Since $x \in N(p_1) \cap Q_1$ we have $e(p_1, Q_1) \ge 1$. Note that $|V(Q_1)| = 2$ or $|V(Q_1)| = 3$. When $|V(Q_1)| = 2$ then q_β may be adjacent to q_2 and $e(q_\beta, Q_1) \le 1$. If $|V(Q_1)| = 3$ and $p_1 q_2 \in E$ then $e(q_\beta, Q_1) = 0$, because otherwise the cycle C' given by (5.7) for $l = 2$ is a hamiltonian cycle of G containing M and Claim 5.5.2 is proved.

□

Claim 5.5.3. *1. If Q_{i_0} is a $BB-$path and Q_{j_0} is a $WW-$path, $2 \le i_0,\ j_0 \le n_Q$, then*

$$e(\{p_1, q_\beta\}, Q_{i_0} \cup Q_{j_0}) \le 3 = \frac{|V(Q_{i_0})| + |V(Q_{j_0})|}{2}. \tag{5.8}$$

2. If Q_k, $2 \le k \le n_Q$ is a BW-path or a WB-path then

$$e(\{p_1, q_\beta\}, Q_k) \le 1 = \frac{|Q_k|}{2}. \tag{5.9}$$

3. In any case

$$e(\{p_1, q_\beta\}, Q_1) \le 2. \tag{5.10}$$

Proof of Claim 5.5.3

For any i, since the matching M is maximal we have $|Q_i| = 3$, if and only if Q_i is a $BB-$path or a $WW-$path and $|Q_i| = 2$, if and only if Q_i is a $BW-$path or a $WB-$path. Consider a BB-path Q_{i_0} and a $WW-$path Q_{j_0}, ($2 \le i_0,\ j_0 \le n_Q$). From Claim 5.5.1 for $2 \le i_0,\ j_0 \le n_Q$, we have $e(\{p_1, q_\beta\}, Q_{i_0}) \le 1$ and $e(\{p_1, q_\beta\}, Q_{j_0}) \le 2$. These proves the inequality (5.8). If $|Q_i| = 2$ from Claim 5.5.1 we have (5.9). The inequality (5.10) is an immediate consequence of Claim 5.5.2.

□

Let us denote $\nu_3(Q)$ the number of odd paths Q_i and $\nu_2(Q)$ the number of even paths Q_k, $1 \le i, k \le n_Q$.

As $|V(Q)|$ is even, the number of $BB-$paths is equal to the number of WW-paths and so $\nu_3(Q)$ is even i.e. $\nu_3(Q) = 2\mu$. Clearly $|V(Q)| = \beta + 2 = 3\,\nu_3(Q) + 2\,\nu_2(Q) = 6\mu + 2\,\nu_2(Q)$.

Now we shall estimate $e(\{p_1, q_\beta\}, Q)$. From Claims 5.5.1 — 5.5.3 we have:

$$
\begin{aligned}
e(\{p_1, q_\beta\}, Q) &= \sum_{|V(Q_i)|=3} e(\{p_1, q_\beta\}, Q_i) + \sum_{|V(Q_k)|=2} e(\{p_1, q_\beta\}, Q_k) \\
&\leq 3\,\mu + \nu_2(Q) + 1 = \frac{\beta}{2} + 2. \qquad (5.11)
\end{aligned}
$$

Similarly we obtain the following three inequalities:

$$
\begin{aligned}
e(\{q_1, p_\alpha\}, Q) &\leq \frac{\beta}{2} + 2, \qquad (5.12)\\
e(\{p_1, q_\beta\}, P) &\leq \frac{\alpha}{2} + 2, \qquad (5.13)\\
e(\{q_1, p_\alpha\}, P) &\leq \frac{\alpha}{2} + 2. \qquad (5.14)
\end{aligned}
$$

Neighbors of $p_1, p_\alpha, q_1, q_\beta$ on R

Note that, for any $k = 1, ..., n_R$, we have $1 \leq |V(R_k)| \leq 3$. If $xr_1 \in M$ then $R_1 = \{x\}$ and $|V(R_1)| = 1$. If $r_\gamma y \in M$ then $R_\gamma = \{x\}$ and $|V(R_\gamma)| = 1$. For $k = 2, ..., n_R - 1$ we have $2 \leq |V(R_k)| \leq 3$.

It is easy to check that if $|V(R_i)| = 2$ then $e(\{p_1, p_\alpha\}, R_i) \leq 1$ and if $|V(R_j)| = 3$ then $e(\{p_1, p_\alpha\}, R_j) \leq 2$.

If $|V(R_j)| = 1$ then $e(\{p_1, p_\alpha\}, R_j) = 1$.

Denote by $\nu_3(R)$ the number of paths R_i with three vertices, by $\nu_2(R)$ the number of paths R_i with two vertices and by $\nu_1(R)$ the number of paths R_k with one vertex.

Note that $\nu_1(R) + \nu_3(R)$ is even and $\gamma + 2 = 3\,\nu_3(R) + 2\,\nu_2(R) + \nu_1(R)$

We have:

$$
\begin{aligned}
e(\{p_1, p_\alpha\}, R) &= \sum_{|V(R_j)|=3} e(\{p_1, p_\alpha\}, R_j) + \sum_{|V(R_i)|=2} e(\{p_1, p_\alpha\}, R_i) \\
&+ \sum_{|V(R_k)|=1} e(\{p_1, p_\alpha\}, R_k) \\
&\leq 2\,\nu_3(R) + \nu_2(R) + \nu_1(R) \\
&= \frac{2\gamma + 4 + \nu_1(R) - \nu_2(R)}{3} \\
&\leq \frac{2\gamma + 6}{3}\,. \qquad (5.15)
\end{aligned}
$$

Similarly we have:

$$e(\{q_1, q_\beta\}, R) \le \frac{2\gamma + 6}{3}\,. \tag{5.16}$$

Now we shall estimate the sum $\mathrm{d}(p_1) + \mathrm{d}(p_\alpha) + \mathrm{d}(q_1) + \mathrm{d}(q_\beta)$.
From (5.11) — (5.16) we have:

$$\mathrm{d}(p_1) + \mathrm{d}(p_\alpha) + \mathrm{d}(q_1) + \mathrm{d}(q_\beta) =$$

$$e(\{p_1, q_\beta\}, Q) + e(\{q_1, p_\alpha\}, Q) + e(\{q_1, p_\alpha\}, P) + e(\{p_1, q_\beta\}, P)$$

$$+e(\{p_1, p_\alpha\}, R) + e(\{q_1, q_\beta\}, R) - 2\, e(\{p_1, q_1, p_\alpha, q_\beta\}, \{x, y\})$$

$$\le \alpha + \beta + 8 + \frac{4\gamma + 12}{3} - 8 = \frac{3\alpha + 3\beta + 4\gamma}{3} + 4.$$

As $\alpha \ge 2$ and $\beta \ge 2$, we obtain the following inequality:

$$\mathrm{d}(p_1) + \mathrm{d}(p_\alpha) + \mathrm{d}(q_1) + \mathrm{d}(q_\beta) \le \frac{4(\alpha + \beta + \gamma) + 8}{3} = \frac{8n}{3},$$

which contradicts (5.6) and the proof is finished.

□

5.5.3 Proof of the part (2) of Theorem 5.4.5

Let $G = (B, W; E)$ be a balanced bipartite graph with $|B| = |W| = n$, $n > 4$ satisfying the conditions of Theorem 5.4.5.

For $n \ge 8$ we have $\frac{5n}{4} \ge n + 2$ and so from assumptions of Theorem 5.4.5 we have:

$$\mathrm{d}(x) + \mathrm{d}(y) \ge \frac{5n}{4} \ge n + 2, \tag{5.17}$$

for any $x \in B$, $y \in W$, $xy \notin E$.

Note that for $n = 5$, 6 and 7 $\frac{5n}{4}$ is not an integer so in fact $\mathrm{d}(x)+\mathrm{d}(y) \ge \left\lceil \frac{5n}{4} \right\rceil$.
It is easy to verify that for $n = 5$, 6 and 7 we have $\left\lceil \frac{5n}{4} \right\rceil = n + 2$.

From the above and (5.17) we have:

$$\mathrm{d}(x) + \mathrm{d}(y) \ge n + 2, \tag{5.18}$$

for any $x \in B$, $y \in W$, $xy \notin E$.

Let M be a matching in G. We may assume that M is a maximal matching. If M is a perfect matching, then Theorem 5.4.4 implies that M is contained in a hamiltonian cycle. We can assume that M is not a perfect matching and we consider a maximal counterexample, i.e. a balanced bipartite graph G and a maximal matching M such that:

1. There is no cycle in G containing M.
2. For every pair of vertices (p, q), $p \in B$, $q \in W$, $pq \notin E$, $p, q \notin V(M)$, then M is contained in a cycle in $G \cup (pq)$.

Note that since M is not a perfect matching then we have at least two vertices p, q such that $p, q \notin V(M)$.

Thus there is a path:

$$D: \quad qu_1u_2...u_lp \tag{5.19}$$

containing M and oriented from q to p.

Since $qp \notin E$ then from (5.18) there exists an i such that $1 \leq i \leq l-1$, $qu_{i+1} \in E$ and $pu_i \in E$.

The cycle:

$$C': \quad qu_{i+1}u_{i+2}...u_lpu_iu_{i-1}...u_1q$$

cannot contain the matching M, so $u_iu_{i+1} \in M$.

Consider paths:

$$\begin{aligned} P &: u_ipu_l \ ... \ u_{i+2}u_{i+1} \\ Q &: u_iu_{i-1} \ ... \ u_1qu_{i+1} \\ R &: u_iu_{i+1} \end{aligned}$$

and note that $H = (P, Q, R)$ is a strict $\Theta-$graph compatible with the matching M. (For an example cf. Figure (5.2).)

Let u_s, $u_r \in V(D)$, $s < r$ be such that $pu_s \in E$, $qu_r \in E$, $u_su_{s+1} \in M$, $u_{r-1}u_r \in M$ (note that $s = i$, $r = i+1$ satisfy these conditions) and $r - s$ is maximal.

The graph $H = (P, Q, R)$:

$$\begin{aligned} P &: u_spu_lu_{l-1} \ ... \ u_{r+1}u_r \\ Q &: u_su_{s-1} \ ... \ u_1qu_r \\ R &: u_s \ ... \ u_r \end{aligned}$$

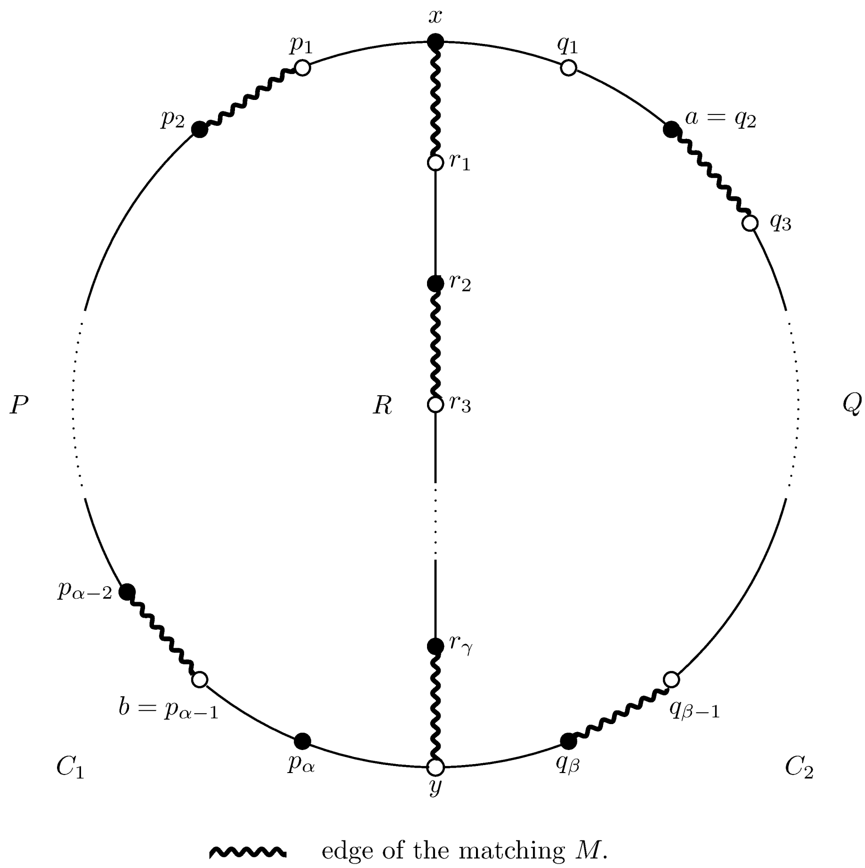


Figure 5.2: A strict Θ-graph compatible with M.

is a strict $\Theta-$graph compatible with the matching M such that $|V(R)|$ is maximum.

Since there is no cycle containing M we have $E(P) \cap M \neq \emptyset$, $E(Q) \cap M \neq \emptyset$ and since H is a strict Θ-graph $|V(P)|$, $|V(Q)| \geq 6$.

We label the vertices of H as follows:

$$\begin{aligned} P &: \ xp_1 \ ... \ p_\alpha y \\ Q &: \ xq_1 \ ... \ q_\beta y \\ R &: \ xr_1 \ ... \ r_\gamma y \end{aligned}$$

We assume that $x \in B$, $y \in W$, $q = q_1 \in W$, $a = q_2 \in B$, $p = p_\alpha \in B$ and $b = p_{\alpha-1} \in W$.

Let G_M be the subgraph of G induced by $V(G) \setminus V(M)$ and let Z be the subgraph of G induced by $V(G) \setminus V(D)$. Subgraphs G_M and Z are independent i.e. $e(V(G_M), Z) = 0$.

Since $V(G) = V(P - \{y\}) \cup V(Q - \{x\}) \cup V(R - \{x, y\}) \cup V(Z)$ and the sets $V(P - \{y\})$, $V(Q - \{x\})$, $V(R - \{x, y\})$ and $V(Z)$ are vertex-disjoint for every vertex $v \in V(G)$, we have:

$$\mathrm{d}(v) = \mathrm{d}_{P-\{y\}}(v) + \mathrm{d}_{Q-\{x\}}(v) + \mathrm{d}_{R-\{x,y\}}(v) + \mathrm{d}_Z(v). \tag{5.20}$$

Let $|M| = m$, $|V(M)| = 2m$, $|V(D \backslash M)| = 2\delta$ and $|V(Z)| = 2t$, then $n = m + \delta + t$.

Remark: As $p_\alpha \notin V(M)$, $q_1 \notin V(M)$, $|V(P)|$ and $|V(Q)|$ are even, then $\delta \geq 2$. (There are at least two vertices of $V(G) \backslash V(M)$ on P and on Q.)

Denote P_i $i = 1, ..., n_P$, Q_j, $j = 1, ..., n_Q$ and R_k, $k = 1, ..., n_R$ the paths obtained respectively from P, Q and $R \setminus \{x, y\}$ by removal of the edges of M.

Take an $i \in \{1, ..., n_P\}$. Note that since M is maximal then if P_i is an odd path then $|V(P_i)| = 3$ and if P_i is an even path then $|V(P_i)| = 2$. Moreover if $|V(P_i)| = 3$ then P_i is a BB-path or a WW-path. If $|V(P_i)| = 2$ then P_i is a BW-path or WB-path. As $|V(P)|$ is even, the number of $BB-$paths is equal to the number of WW-paths. Let $\nu_3(P)$ be the number of odd paths P_i, $\nu_2^{BW}(P)$ the number of BW-paths P_i, $\nu_2^{WB}(P)$ the number of WB-paths P_i and $\nu_2(P) = \nu_2^{BW}(P) + \nu_2^{WB}(P)$ the number of even paths P_i.

The paths Q_i $i = 1, ..., n_Q$ and R_i, $i = 1, ..., n_R$ have the same properties as the paths P_i and in the same way as above, we define ν_2^{BW}, ν_2^{BW}, $\nu_2 = \nu_2^{BW} + \nu_2^{BW}$ and ν_3 for paths Q and R (in which the number of $BB-$paths is also equal to the number of WW-paths).

From the maximality of G and M the graph induced by $V(D) \setminus V(M)$ is independent. Thus since $bp_\alpha y$ is a WW-path we have:

$$n_P = \nu_3(P) + \nu_2(P) = |M \cap E(P)| + 1. \tag{5.21}$$

Similarly since xq_1a is a BB-path we have:

$$n_Q = \nu_3(Q) + \nu_2(Q) = |M \cap E(Q)| + 1. \tag{5.22}$$

Note that on the path $R \setminus \{x, y\}$ we have:

$$n_R = \nu_3(R) + \nu_2(R) = |M \cap E(R)| - 1. \tag{5.23}$$

From (5.21) — (5.23) we have:

$$\sum_{i=2}^{3} (\nu_i(P) + \nu_i(Q) + \nu_i(R)) = m + 1. \tag{5.24}$$

In every path odd P_i, there is one vertex of $V(D) \setminus V(M)$ and since $|V(R)|$ is even we have:

$$\nu_3(P) = |V(P \setminus M)|. \tag{5.25}$$

Similarly we have:

$$\nu_3(Q) = |V(Q \setminus M)|. \tag{5.26}$$

$$\nu_3(R) = |V(R \setminus M)|. \tag{5.27}$$

From (5.25) — (5.27) we have:

$$\nu_3(P) + \nu_3(Q) + \nu_3(R) = 2\delta. \tag{5.28}$$

Lower bound of the sums of degrees

If one of the edges ab, $p_\alpha q_1$, $p_1 q_\beta$ exists, we have a cycle in G containing every edge of M. For example if $p_1 q_\beta \in E$ then the cycle:

$$C: \quad p_1 q_\beta q_{\beta-1} ... q_1 x r_1 ... r_\gamma y p_\alpha .... p_1$$

is containing M.

We may assume $ab \notin E$, $p_\alpha q_1 \notin E$, $p_1 q_\beta \notin E$ and then:

$$\mathrm{d}(a) + \mathrm{d}(b) \geq \frac{5n}{4}, \tag{5.29}$$

$$\mathrm{d}(q_\beta) + \mathrm{d}(p_1) \geq \frac{5n}{4}, \tag{5.30}$$

$$\mathrm{d}(p_\alpha) + \mathrm{d}(q_1) \geq \frac{5n}{4}. \tag{5.31}$$

Upper bound of the sum of degrees

Neighbors of $a,\ b,\ p_\alpha,\ q_1,\ q_\beta,\ p_1$ on $R \setminus \{x, y\}$

1. Consider a $WB-$path $R_i :\ vu$ on R, $u \in B$, $v \in W$, $v = u^-$, $uv \notin M$. Since there is no cycle containing every edge of M, the following inequalities are satisfied: $e(\{p_\alpha, p_1\}, R_i) \leq 1$, $e(\{q_1, q_\beta\}, R_i) \leq 1$ and $e(\{a, b\}, R_i) \leq 1$.

   Suppose that $e(\{a, b\}, R_i) = 2$, then av, $bu \in E$ and the following cycle C :

   $$C : \quad avv^- ... r_1 x p_1 ... p_{\alpha-2} b u u^+ ... r_\gamma y q_\beta ... a$$

   contains M, a contradiction.

   Now suppose that $e(\{p_1, p_\alpha\}, R_i) = 2$. In this case $p_1 u$, $p_\alpha v \in E(G)$ and the following cycle C :

   $$C : \quad p_\alpha v v^- ... r_1 x q_1 ... q_\beta y r_\gamma ... u p_1 p_2 ... p_\alpha$$

   contains M, a contradiction.

   The case $e(\{q_1, q_\beta\}, R_i) = 2$ is the same as $e(\{p_1, p_\alpha\}, R_i) = 2$ and so we have:

   $$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, R_i) \leq 3. \tag{5.32}$$

2. Consider a $BW-$path $R_i :\ uv$ on R, $u \in B$, $v \in W$, $v = u^+$, $uv \notin M$. The following inequalities holds: $e(\{p_1, p_\alpha\}, R_i) \leq 1$, $e(\{q_1, q_\beta\}, R_i) \leq 1$, $e(\{a, b\}, R_i) \leq 2$.

   Since a, $u \in B$ and b, $v \in W$ it is clear that $e(\{a, b\}, R_i) \leq 2$.

   Suppose that $e(\{p_1, p_\alpha\}, R_i) = 2.$, then $p_1 u$, $v p_\alpha \in E$ and the following cycle C :

   $$C : \quad p_\alpha v v^+ ... r_\gamma y q_\beta ... q_1 x r_1 .... u p_1 ... p_\alpha$$

   contains M, a contradiction.

   The case $e(\{q_1, q_\beta\}, R_i) = 2$ is the same as $e(\{p_1, p_\alpha\}, R_i) = 2$. Thus:

   $$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, R_i) \leq 4. \tag{5.33}$$

3. Consider a $WW-$path $R_i: \ v_1uv_2$, $u \in B$, v_1, $v_2 \in W$, $u \in V(D \setminus M)$, $u = v_1^+ = v_2^-$. As $q_1 \notin V(M)$, $u \notin V(M)$ and M is maximal, we have $q_1u \notin E$. Since there is no cycle containing M, the following inequalities holds: $e(\{p_1, p_\alpha\}, R_i) \leq 2$, $e(\{a, q_\beta\}, R_i) \leq 2$.

   We will start to compute $e(\{p_1, p_\alpha\}, R_i)$.

   If $p_1u \notin E$ then $e(\{p_1, p_\alpha\}, R_i) \leq 2$.

   Suppose now that $p_1u \in E$ and $e(p_\alpha, R_i) \neq 0$. $e(p_\alpha, R_i) \neq 0$ implies that $p_\alpha v_1 \in E$ or $p_\alpha v_2 \in E$.

   If $p_\alpha v_1 \in E$, then the following cycle C :

$$C: \quad p_\alpha v_1 v_1^- ... r_1 x q_1 .... q_\beta y r_\gamma .... u p_1 ... p_\alpha$$

   contains M, a contradiction.

   If $p_\alpha v_2 \in E$, then the following cycle C :

$$C: \quad p_\alpha v_2 v_2^+ ... r_\gamma y q_\beta ... q_1 x r_1 .... u p_1 ... p_\alpha$$

   contains M, a contradiction. So if $p_1u \in E$ we have $e(\{p_1, p_\alpha\}, R_i) = 1$.

   Thus in any case we have $e(\{p_1, p_\alpha\}, R_i) \leq 2$.

   Now, we shall compute $e(\{a, q_\beta\}, R_i)$. Note that a and q_β cannot be adjacent to two different vertices on R_i. Since a, u, $q_\beta \in B$ and v_1, $v_2 \in W$, we shall consider the existence of four edges: av_1, $q_\beta v_1$, av_2 and $q_\beta v_2$.

   Suppose that av_1, $q_\beta v_2 \in E$, then the following cycle C :

$$C: \quad a v_1 v_1^- ... r_1 x p_1 .... p_\alpha y r_\gamma .... v_2 q_\beta ... a$$

   contains M, a contradiction.

   If av_2, $q_\beta v_1 \in E$, then the following cycle C :

$$C: \quad a v_2 v_2^+ ... r_\gamma y p_\alpha .... p_1 x r_1 ... v_1 a$$

   contains M, a contradiction.

   So we have $e(\{a, q_\beta\}, R_i) \leq 2$ and since it may happen that $bu \in E$, we have:

$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, R_i) \leq 5. \tag{5.34}$$

4. Consider a $BB-$path R_i : $u_1 v u_2$, u_1, $u_2 \in B$, $v \in W$, $v \in V(D \setminus M)$, $v = u_1^+ = u_2^-$. Since $p_\alpha \notin V(M)$, $v \notin V(M)$ and M is maximal, we have $p_\alpha v \notin E$. Using the same arguments as in the case 3, since there is no cycle containing M, the following inequalities hold: $e(\{q_1, q_\beta\}, R_i) \leq 2$, $e(\{b, p_1\}, R_i) \leq 2$ and since it may happen that $av \in E$, we have:

$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, R_i) \leq 5. \tag{5.35}$$

By summing over all the paths R_i from (5.32) — (5.35) we have:

$$e(\{a, b, p_\alpha, q_1, q_\beta, p_1\}, R - \{x, y\}) \leq 3\nu_2(R) + \nu_2^{BW}(R) + 5\nu_3(R). \tag{5.36}$$

Neighbors of a, b, p_α, q_1, q_β, p_1 on $Q \setminus \{x\}$.

1. Consider the vertices $\{q_1, a\}$. Since there is no cycle containing M we have $e(\{p_1, q_\beta\}, \{q_1, a\}) \leq 1$, $aq_1 \in E$, $p_\alpha q_1$, $ab \notin E$ and thus:

$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, \{q_1, a\}) \leq 3. \tag{5.37}$$

2. Consider a BW-path $Q_i : uv$, $u \in B$, $v \in W$, $v = u^+$, $uv \notin M$. Since there is no cycle containing M we have $e(\{p_1, p_\alpha\}, Q_i) \leq 1$ and $e(\{a, b\}, Q_i) \leq 1$.

   Suppose that $e(\{p_1, p_\alpha\}, Q_i) = 2$, then $p_1 u$, $p_\alpha v \in E$ and the following cycle C :

$$C: \quad p_1 u u^- \ldots q_1 x r_1 \ldots r_\gamma y q_\beta \ldots v p_\alpha p_{\alpha-1} \ldots p_1$$

   contains M, a contradiction.

   If $e(\{a, b\}, Q_i) = 2$, then bu, $av \in E$ and the following cycle C :

$$C: \quad b u u^- \ldots a v v^+ \ldots q_\beta y r_\gamma \ldots r_1 x p_1 \ldots b$$

   contains M, a contradiction.

   Note that $e(\{q_1, q_\beta\}, Q_i) \leq 2$ and thus:

$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, Q_i) \leq 4. \tag{5.38}$$

3. Consider a WB-path $Q_i : vu$, $u \in B$, $v \in W$, $u = v^+$, $vu \notin M$. Since there is no cycle containing M, using similar arguments as in the case 2, for the vertices a, b, we have $e(\{q_1, p_\alpha\}, Q_i) \leq 1$, $e(\{p_1, q_\beta\}, Q_i) \leq 1$ and thus:
$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, Q_i) \leq 4. \tag{5.39}$$

4. Consider a WW-path $Q_i : v_1uv_2$, $v_2 \neq y$, $u \in B$, v_1, $v_2 \in W$, $u = v_1^+ = v_2^-$.

   As $q_1 \notin V(M)$, $u \notin V(M)$ and M is maximal, we have $q_1u \notin E$ and from this: $e(\{q_1, q_\beta\}, Q_i) \leq 2$.

   Since $v_2 \neq y$ and as R is maximal $p_\alpha v_2 \notin E$. Suppose that $p_\alpha v_2 \in E$ then the graph $H' = (P', Q, R')$ with:
$$\begin{aligned} P' &: xp_1 \ldots p_\alpha v_2 \\ Q' &: xq_1 \ldots v_2 \\ R' &: xr_1 \ldots r_\gamma y q_\beta \ldots . v_2 \end{aligned}$$
   is a strict Θ-graph compatible with M with $|V(R')| > |V(R)|$.

   Since there is no cycle containing M, using similar arguments as in the case 2, we have $e(\{p_1, p_\alpha\}, \{v_1, u\}) \leq 1$, $e(\{a, b\}, \{u, v_2\}) \leq 1$. Hence $e(\{p_1, p_\alpha\}, Q_i) \leq 1$ and since it is possible that $av_1 \in E$ we have $e(\{a, b\}, Q_i) \leq 2$.

   From these inequalities we have:
$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, Q_i) \leq 5. \tag{5.40}$$

5. In the case 4 we have assumed that $v_2 \neq y$. If $v_2 = y$, then $i = N_Q$ and the path Q_{N_Q} is a WW-path $Q_{N_Q} : q_{\beta-1}q_\beta y$. In fact, it is the same case as the case 4, but since $p_\alpha y \in E$, we have:
$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, Q_i) \leq 6. \tag{5.41}$$

6. Consider a BB-path $Q_i : u_1vu_2$, u_1, $u_2 \in B$, $v \in W$, $v = u_1^+ = u_2^-$. Note that since p_α, $v \notin V(M)$ and since M is maximal we have $p_\alpha v \notin E$.

   Since there is no cycle containing M we have $e(\{a, b\}, \{u_1, v\}) \leq 1$, $e(\{p_1, q_\beta\}, \{v, u_2\}) \leq 1$.

   Suppose that $e(\{a, b\}, \{u_1, v\}) = 2$, then av, $bu_1 \in E$ and the following cycle C :

$$C: \quad bu_1u_1^-....avv^+....q_\beta yr_\gamma...r_1xp_1...b$$

contains M, a contradiction.

Suppose that $e(\{p_1, q_\beta\}, \{v, u_2\}) = 2$, then p_1u_2, $q_\beta v \in E$ and the following cycle C :

$$C: \quad p_1u_2u_2^+....q_\beta vu_1....q_1xr_1...r_\gamma yp_\alpha...p_1$$

contains M, a contradiction.

Note that $p_1u_1 \notin E$, because if $p_1u_1 \in E$, then the graph $H' = (P', Q, R')$ with:

$$\begin{aligned} P' &: \; u_1p_1 \;...\; p_\alpha y \\ Q' &: \; u_1vu_2 \;...\; q_\beta y \\ R' &: \; u_1u_1^- \;...\; q_1xr_1 \;...\; r_\gamma y \end{aligned}$$

is a strict Θ-graph compatible with M with $|V(R')| > |V(R)|$, a contradiction.

From the above we have: $e(\{p_1, q_\beta\}, Q_i) \leq 1$, $e(\{a, b\}, Q_i) \leq 2$ and since $e(\{q_1\}, Q_i) \leq 2$ we have

$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, Q_i) \leq 5. \tag{5.42}$$

By summing over all the paths Q_i from (5.37) — (5.42) we have:

$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, Q \setminus \{x\}) \leq 4\nu_2(Q) + 5\nu_3(Q) - 1. \tag{5.43}$$

Neighbors of $a, b, p_\alpha, q_1, q_\beta, p_1$ on $P \setminus \{y\}$

Using the similar arguments as in Section 5.5.3 we have:

$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, P \setminus \{y\}) \leq 4\nu_2(P) + 5\nu_3(P) - 1. \tag{5.44}$$

Neighbors of $a, b, p_\alpha, q_1, q_\beta, p_1$ in Z

Since G_M and Z are independent we have:

$$\mathrm{d}_Z(p_\alpha) = \mathrm{d}_Z(q_1) = 0$$

and thus:

$$e(\{a, b, p_1, p_\alpha, q_1, q_\beta\}, Z) \leq 4t. \tag{5.45}$$

Neighbors of p_α and q_1 on $R \cup Q \cup P$

Using similar methods as those in Sections 5.5.3 — 5.5.3 we get the following inequalities:

$$e(\{p_\alpha, q_1\}, R \setminus \{x, y\}) \le \nu_2^{BW}(R) + 2\nu_2^{WB}(R) + 2\nu_3(R). \tag{5.46}$$

$$e(\{p_\alpha, q_1\}, Q \setminus \{x\}) \le \nu_2(Q) + 2(\nu_3(Q) - 1) + 1 = \nu_2(Q) + 2\nu_3(Q) - 1. \tag{5.47}$$

$$e(\{p_\alpha, q_1\}, P \setminus \{y\}) \le \nu_2(P) + 2(\nu_3(P) - 1) + 1 = \nu_2(P) + 2\nu_3(P) - 1. \tag{5.48}$$

Now we shall estimate the sum of degrees. From (5.46) — (5.48) we have

$$\mathrm{d}(p_\alpha) + \mathrm{d}(q_1) \le \nu_2^{BW}(R) + m + 2\delta - 1 = \nu_2^{WB}(R) + n - t + \delta - 1. \tag{5.49}$$

Conclusion

From (5.36), (5.43), (5.44) and (5.45) we have:

$$\begin{aligned} \mathrm{d}(a) + \mathrm{d}(b) + \mathrm{d}(p_1) + \mathrm{d}(p_\alpha) + \mathrm{d}(q_1) + \mathrm{d}(q_\beta) \le \\ 4\Big(\sum_{i=2}^{3} (\nu_i(P) + \nu_i(Q) + \nu_i(R))\Big) \\ + \nu_3(P) + \nu_3(Q) + \nu_3(R) - 2 + 4t - \nu_2^{WB}(R). \end{aligned} \tag{5.50}$$

From (5.24), (5.28) and (5.50) we deduce:

$$\begin{aligned} \mathrm{d}(a) + \mathrm{d}(b) + \mathrm{d}(p_1) + \mathrm{d}(p_\alpha) + \mathrm{d}(q_1) + \mathrm{d}(q_\beta) \le \\ -2 + 4(m+1) + 2\delta + 4t - \nu_2^{WB}(R). \end{aligned} \tag{5.51}$$

Since $n = m + \delta + t$, from (5.51) we have:

$$\begin{aligned} \mathrm{d}(a) + \mathrm{d}(b) + \mathrm{d}(p_1) + \mathrm{d}(p_\alpha) + \mathrm{d}(q_1) + \mathrm{d}(q_\beta) \le \\ 4n + 2 - \nu_2^{WB}(R) - 2\delta. \end{aligned} \tag{5.52}$$

From (5.49) and (5.52) we can deduce that:

$$\mathrm{d}(a) + \mathrm{d}(b) + \mathrm{d}(q_\beta) + \mathrm{d}(p_1) + 2\mathrm{d}(p_\alpha) + 2\mathrm{d}(q_1) \leq 5n - t - \delta + 1. \tag{5.53}$$

Note that $\delta \geq 2$ and from (5.53) we have:

$$\mathrm{d}(a) + \mathrm{d}(b) + \mathrm{d}(q_\beta) + \mathrm{d}(p_1) + 2\mathrm{d}(p_\alpha) + 2\mathrm{d}(q_1) \leq 5n - 1. \tag{5.54}$$

Now we shall give the lower bound of the sum of degrees. From (5.29) — (5.31) we have

$$4\frac{5n}{4} \leq \mathrm{d}(q_\beta) + \mathrm{d}(p_1) + \mathrm{d}(a) + \mathrm{d}(b) + 2\mathrm{d}(p_\alpha) + 2\mathrm{d}(q_1) \tag{5.55}$$

Assuming that there does not exist a cycle in G which contains every edge of the matching M, we have obtained (5.54) and (5.55). Hence,

$$5n \leq 5n - 1,$$

a contradiction. The part (2) of Theorem 5.4.5 is proved. □

References

[1] D. Amar et al. "Bipartite graphs with every matching in a cycle". In: *Discrete Math.* 307, No. 11-12 (2007), pp. 1525–1537.

[2] K. A. Berman. "Proof of a conjecture of Häggkvist on cycles and independent edges". In: *Discrete Math.* 46 (1983), pp. 9–13.

[3] J. A. Bondy. "Variations on the Hamiltonian theme". In: *Can. Math. Bull.* 15 (1972), pp. 57–62.

[4] J. A. Bondy and U. S. R. Murty. *Graph theory with applications.* American Elsevier Publishing Co., New York, 1976.

[5] V. Chvátal. "New directions in Hamiltonian graph theory". In: *New directions in the theory of graphs, F. Harary (ed.)* Academic Press, New-York-London, 1972, pp. 65–95.

[6] V. Chvátal. "On Hamilton's ideals". In: *J. Comb. Theory* 12 (1972), pp. 163–168.

[7] G. A. Dirac. "Some theorems on abstract graphs". In: *Proc. Lond. Math. Soc., III. Ser.* 2 (1952), pp. 69–81.

[8] J. Moon and L. Moser. "On Hamiltonian bipartite graphs". In: *Isr. J. Math.* 1 (1963), pp. 163–165.

[9] O. Ore. "Arc coverings of graphs". In: *Ann. Mat. Pura Appl., IV. Ser.* 55 (1961), pp. 315–321.

[10] O. Ore. "Note on Hamiltonian circuits". In: *Amer. Math. Mon.* 67 (1960), p. 55.

[11] R. Haggkvist. *On F-Hamiltonian graphs*. 1979.

[12] M. Las Vergnas. "Problèmes de couplages et problèmes hamiltoniens en théorie des graphes". PhD thesis. Université Paris XI, 1972.

[13] A. P. Wojda. "Hamiltonian cycles through matchings". In: *Demonstr. Math.* 21, No. 2 (1988), pp. 547–553.

Chapter 6

An outline of the theory of $\mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$

MAGDALENA GRZECH

Magdalena Grzech
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: magdag@pk.edu.pl

Abstract

A survey of properties of $\mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ in a set-theoretical context is presented.

Contents

6.1 Introduction

The aim of this paper is to give a survey of results concerning the group of automorphisms of the Boolean algebra $\mathcal{P}(\omega)/\mathsf{FIN}$. The algebra plays an important role in foundation of mathematics. Moreover, the results about $\mathcal{P}(\omega)/\mathsf{FIN}$ have a wide implications in various fields of mathematics. Many mathematical problems can be reduced to questions on properties of $\mathcal{P}(\omega)/\mathsf{FIN}$. The algebra and its Stone space $\omega^\star$ are well studied both from the topological and set-theoretical point of view. This paper has been written from the perspective of a set-theorist. We do not aim to be complete. Several interesting results will not be mentioned.

We begin with reminding some basic facts and definitions. By ω we denote the set of all natural numbers. The symbol FIN stands for the ideal of all finite subsets of ω. The ideal determines the following equivalence relation:

$$\text{for } A,\ B \subseteq \omega,\ A =_\star B \quad \text{if and only if} \quad A \div B \in \mathsf{FIN},$$

$\mathcal{P}(\omega)/\mathsf{FIN}$ is its factor algebra. An order in $\mathcal{P}(\omega)/\mathsf{FIN}$ is defined in the following way:

$$A \subseteq_\star B \text{ if } A \setminus B \in \mathsf{FIN}.$$

The space $\beta[\omega]$ consists of all ultrafilters on ω and is equipped with a topology generated by basic open sets of the form

$$\mathsf{V}_A = \{\mathsf{p} \in \beta[\omega] :\ A \notin \mathsf{p}\}, \qquad \text{for } A \subseteq \omega.$$

This space is the Čech-Stone compactification of the countable discrete space ω. The space ω can be embedded into $\beta[\omega]$ as follows: we identify each natural number n with a principal ultrafilter $\mathsf{p}_n = \{A \subseteq \omega : n \in A\}$. A symbol $\omega^\star$ denotes the remainder (or growth) space of ω i.e. the space $\beta[\omega] \setminus \omega$ which consists of all free ultrafilters on ω. For $A \subseteq \omega \subseteq \beta[\omega]$, the closure of A (in $\beta[\omega]$) is equal to $\overline{A} = \{\mathsf{p} \in \beta[\omega] :\ A \in \mathsf{p}\}$. Note, that

$$\overline{A} = \beta[\omega] \setminus \mathsf{V}_A = \mathsf{V}_{\omega\setminus A} = \beta[\omega] \setminus \overline{\omega \setminus A},$$

thus $\overline{A}$ is clopen set in $\beta[\omega]$. Let $A^\star = \overline{A} \setminus A$. Thus $A^\star$ is a clopen set in $\omega^\star$. Thus for infinite $A,\ B \subseteq \omega$, $A =_\star B$ if and only if the basic sets V_A,V_B, are equal and we have

$$A^\star = B^\star \text{ if and only if } A =_\star B.$$

It follows that the Boolean algebra of clopen subsets of $\omega^\star$ is isomorphic to the Boolean algebra $\mathcal{P}(\omega)/\mathsf{FIN}$. To simplify notation we shall identify the equivalence class $[A]_\star$ (with respect to the relation $=_\star$) of A with its growth $A^\star$. ($\omega^\star$ and $\mathcal{P}(\omega)/\mathsf{FIN}$ will be used almost interchangeably. As well, subsets of ω will routinely be confused with clopen sets of $\omega^\star$.)

It is well known that the $\omega^\star$ is compact zero-dimentional space of weight $\mathfrak{c}$ (the continuum, $\mathfrak{c} = 2^{\omega_0}$). Moreover it satisfies the following conditions:

1.T The space has no isolated points.

2.T Nonempty G_δ sets have nonempty interiors.

3.T $\omega^\star$ is an F-space (Recall that a topological space X is an *F-space* provided that X is Tychonoff space and each bounded real-valued continuous function on the complement of a functionally open subset of X can be extended continuously to the whole of X.)

A compact zero-dimensional space is called a *Parovičenko space* if it satisfies conditions 1.T, 2.T and 3.T.

Every separable compact space as well as every compact space of weight at most ω_1 is a continuous image of $\omega^\star$. Under CH (the Continuum Hypothesis) every Parovičenko space is homeomorphic to $\omega^\star$ ([5])

The above properties can be translated in algebraical language. Let λ, κ be ordinals. *A gap of the type (λ, κ)* in a Boolean algebra $(\mathbb{A}, +, \cdot, 0, 1)$ is a pair $(\{a_\gamma : \gamma < \lambda\}\ ,\ \{b_\beta : \beta < \kappa\})$ of subsets of $\mathbb{A}$ such that $a_\gamma \cdot b_\beta = 0$. If for every $\gamma_1 < \gamma_2 < \lambda$, $\beta_1 < \beta_2 < \kappa$, $a_{\gamma_1} \cdot a_{\gamma_2} = a_{\gamma_1}$ and $b_{\beta_1} \cdot b_{\beta_2} = b_{\beta_1}$, the gap is said to be *incresingly ordered.* An element $c \in \mathbb{A}$ *fills (separates)* the gap if $a_\gamma \cdot c = a_\gamma$ and $b_\beta \cdot c = 0$ for every $\gamma < \lambda$, $\beta < \kappa$. If there is no such an element, the gap is called *non-separable.* A (strictly) decreasing sequence $(a_\beta : \beta < \gamma)$ of elements of the $\mathbb{A}$ of length γ is called *γ-limit* if there is no non-zero element $a \in \mathbb{A}$ such that for every $\beta < \gamma$, $a \cdot a_\beta = a$. A non-zero element a of $\mathbb{A}$ is *an atom* if 0 is a unique element smaller than a. An algebra $\mathbb{A}$ is *atomless* provided it has no atoms at all.

The algebra $\mathcal{P}(\omega)/\mathsf{FIN}$ has cardinality $\mathfrak{c}$. Moreover

1.A $\mathcal{P}(\omega)/\mathsf{FIN}$ is atomless.

2.A $\mathcal{P}(\omega)/\mathsf{FIN}$ has no ω-limits (i.e. γ-limits with $\mathrm{card}(\gamma) = \omega$).

3.A In the algebra $\mathcal{P}(\omega)/\mathsf{FIN}$ every countable gap (i.e. $\mathrm{card}(\lambda) = \mathrm{card}(\kappa) = \omega$) is filled.

Let f be a autohomeomorphism of of $\omega^\star$. Then $\mathsf{h}_f : \mathcal{P}(\omega)/\mathsf{FIN} \ni [A]_\star \to f[A^\star] = [B]_\star \in \mathcal{P}(\omega)/\mathsf{FIN}$ is an automorphism of $\mathcal{P}(\omega)/\mathsf{FIN}$, where

$$[A]_\star \approx A^\star \to f\left[\{\mathsf{p} \in \omega^\star : A \in \mathsf{p}\}\right] \text{ is clopen in } \omega^\star,$$

thus $f\left[\{\mathsf{p} \in \omega^\star : A \in \mathsf{p}\}\right] = [B]_\star$, for some B. On the other hand each automorphism h of $\mathcal{P}(\omega)/\mathsf{FIN}$ determines an autohomeomorphism of $\omega^\star$. Indeed, we set

$$f_{\mathsf{h}} : \omega^\star \ni \mathsf{p} \to f_{\mathsf{h}}(\mathsf{p}),$$

where $f_{\mathsf{h}}(\mathsf{p})$ is the unique element of the set

$$\bigcap_{A^\star \in \mathsf{B}_\mathsf{p}} \mathsf{h}\left[[A]_\star\right], \text{ for } \mathsf{B}_\mathsf{p} = \{A^\star : \mathsf{p} \in A^\star\}.$$

Thus the group of automorphisms of $\mathcal{P}(\omega)/\mathsf{FIN}$ is isomorphic to the group of autohomemorphisms of $\omega^\star$.

Assume that g is *almost permutation of* ω (i.e. an injective function from ω to ω whose domain and range are both cofinite). Then g induces an autohomeomorphism $g^\star$ of $\omega^\star$ and automorphism $\mathsf{h}_{g^\star}$ of $\mathcal{P}(\omega)/\mathsf{FIN}$:

$$g^\star : \omega^\star \ni \mathsf{p} \to \{g[A] : A \in \mathsf{p}\} \in \omega^\star, \quad \mathsf{h}_{g^\star} : \mathcal{P}(\omega)/\mathsf{FIN} \ni A^\star \to g^\star[A]^\star \in \mathcal{P}(\omega)/\mathsf{FIN}.$$

It is easy to see that for almost permutations f and g, $f^\star \neq g^\star$ if and only if $\{n \in \omega : f(n) \neq g(n)\} \in \mathsf{FIN}$. Since there are continuum almost permutations of ω, it follows

$$\mathrm{card}\left(\mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})\right) \geq \mathfrak{c}.$$

An automorphism induced by almost permutation is called *trivial*. An automorphism h is *somewhere trivial* if there is an infinite set $A \subseteq \omega$ such that the restriction of h to $\mathcal{P}(A)/\mathsf{FIN}$ is trivial. An automorphism is *nowhere trivial* if it is not somewhere trivial.

We finish this section with the Sikorski's theorem (on extending homomorphism) which will be frequently used.

Theorem 1 (Sikorski [11], [10]). *Let $\mathbb{A}$, $\mathbb{B}$ be Boolean algebras, $\mathbb{A}_0$ a subalgebra of $\mathbb{A}$ and $a_0 \in \mathbb{A} \setminus \mathbb{A}_0$. Assume that $\mathsf{h} : \mathbb{A}_0 \to \mathbb{B}$ is a homomorphism. If there exists an element $b \in \mathbb{B}$ which fills a gap*

$$\mathcal{L} = \left(\{\mathsf{h}(x) : x \leq a_0\}, \{\mathsf{h}(x) : x \cdot a_0 = 0\}\right),$$

then h can be extended to a homomorphism $\overline{\mathsf{h}} : \mathbb{A}_1 \to \mathbb{B}$ (where $\mathbb{A}_1$ is a subalgebra generated by $\mathbb{A}_0 \cup \{a_0\}$) with $\bar{f}(a_0) = b$.
Morover if h is monomorphism then $\overline{\mathsf{h}}$ is monomorphism if and only if the following condition holds:

$$\text{for all } x, y \in \mathbb{A}_0 \ [\ (\ x \leq a_0 \iff \mathsf{h}(x) \leq b\ ) \text{ and } (\ y \geq a_0 \iff \mathsf{h}(y) \geq b\ )].$$

6.2 Automorphisms of $\mathcal{P}(\omega)/\mathsf{FIN}$ under CH.

We shall show that, under CH, $\mathrm{card}\left(\mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})\right) = 2^{\mathfrak{c}}$. We shall point out a Parovičenko space homeomorphic to $\omega^\star$ such that cardinality of group of its automorphism can be easily estimate. The following lemma is the starting point of the proof ([2],[3]):

Lemma 1 (L. Gillman, M. Henriksen). *If X is non-compact but locally compact and $\sigma-$compact space, then the growth $X^\star = \beta[X] \setminus X$ is a compact F-space.*

We shall use the following corollary of the lemma: if Y is a zero-dimensional, compact space, then $(\omega \times Y)^\star = \beta[\omega \times Y] \setminus (\omega \times Y)$ is a Parovičenko space. Take $Y = 2^\mathfrak{c}$, where $2^\mathfrak{c}$ is the Cantor cube of weight $\mathfrak{c}$. Then $(\omega \times 2^\mathfrak{c})^\star$ is a Parovičenko space of weight $\mathfrak{c}$ thus, under CH, is homeomorphic to $\omega^\star$. Since the cube $2^\mathfrak{c}$ is a topological group of cardinality $2^\mathfrak{c}$, it follows that $\omega^\star$ has at least $2^\mathfrak{c}$ autohomeomorpisms. Since $\omega^\star$ has weight $\mathfrak{c}$, it cannot have more than $2^\mathfrak{c}$ autohomeomorphisms. Thus card $\big(\mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})\big) = 2^\mathfrak{c}$. $\square$

A question arises if for every two ultrafilters p, $\mathsf{q} \in \omega^\star$ there exist an automorphism $\mathsf{h} \in \mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ such that $f_\mathsf{h}(\mathsf{p}) = \mathsf{q}$, where f_h is the autohomeomorphism induced by h. In order to answer this question we use a notion of P-point. An ultrafilter $\mathsf{p} \in \omega^\star$ is *a P-point* if for every countable subfamily $\{V_n : n \in \omega\} \subseteq \mathsf{p}$ there exists $V \in \mathsf{p}$ such that $V \subseteq_\star V_n$, for every $n \in \omega$. Equivalentely, an element $\mathsf{p} \in \omega^\star$ is called *a P-point* if for every countable family $\{V_n^\star : n \in \omega\}$ of neighbourhoods of p, p belongs to the interior of the intersection $\bigcap_{n\in\omega} V_n^\star$. It is obvious that a homeomorphic image of P-point have to be P-point. Unlike the ultrafilters which are not P-poins and which exist in every model of ZFC, P-points do not exist in some models ([7]). Under CH the set of all P-points is dense in $\omega^\star$. Thus there are points p, $\mathsf{q} \in \omega^\star$ (p is a P-point, q is not a P-point) such that for each $\mathsf{h} \in \mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$, $f_\mathsf{h}(\mathsf{p}) \neq \mathsf{q}$. On the other hand the following holds:

Theorem 2 ([6]). *Suppose that* p, q *are P-points. Then there exists an automorphism* h *of* $\mathcal{P}(\omega)/\mathsf{FIN}$ *such that an autohomeomorphism* f_h *induced by* h *maps* p *to* q.

Sketch of the proof. We construct a required automorphism by induction using Sikorski's theorem. Take two sets of generators of $\mathcal{P}(\omega)/\mathsf{FIN}$: $\mathsf{p} \subseteq \{C_\alpha : \alpha < \omega_1\}$, $\mathsf{q} \subseteq \{D_\alpha : \alpha < \omega_1\}$. Define by induction two increasing sequences $(\mathbb{A}_\beta)_{\beta<\omega_1}$, $(\mathbb{B}_\beta)_{\beta<\omega_1}$ of subalgebras of $\mathcal{P}(\omega)/\mathsf{FIN}$ and an incresing sequence of isomorphisms between them such that:

1. $\{C_\alpha : \alpha < \beta\} \subseteq \bigcup_{\alpha\leq\beta} \mathbb{A}_\alpha, \quad \{D_\alpha : \alpha < \beta\} \subseteq \bigcup_{\alpha\leq\beta} \mathbb{B}_\alpha$,
2. $\mathsf{h}_\alpha : \mathbb{A}_\alpha \to \mathbb{B}_\alpha$ is an isomorphism for $\alpha \leq \beta$, $\mathsf{h}_\gamma \subseteq \mathsf{h}_\alpha$, for $\gamma \leq \alpha \leq \beta$,
3. $\mathbb{A}_\beta = \bigcup_{\alpha<\beta} \mathbb{A}_\alpha$, $\mathbb{B}_\beta = \bigcup_{\alpha<\beta} \mathbb{B}_\alpha$, $\mathsf{h}_\beta = \bigcup_{\alpha<\beta} \mathsf{h}_\alpha$, if β is limit,
4. $\bigcup_{\alpha<\beta} \mathbb{A}_\alpha$ is at most countable for $\alpha < \beta$,
5. if $V \in \mathsf{p} \cap \mathbb{A}_\beta$, $W \in \mathsf{q} \cap \mathbb{B}_\beta$ then $\mathsf{h}_\beta(V) \in \mathsf{q}$, $\mathsf{h}^{-1}(W) \in \mathsf{p}$.

At step $\beta + 1$ consider generators C_β and D_β. If $C_\beta \in \mathbb{A}_\beta$, $D_\beta \in \mathbb{B}_\beta$ then $A_{\beta+1} := \mathbb{A}_\beta$, $B_{\beta+1} := \mathbb{B}_\beta$, $\mathsf{h}_{\beta+1} := \mathsf{h}_\beta$. Assume that $C_\beta \notin \mathbb{A}_\beta$. Then construct (using the standard method described in e.g. [10]) a countable gap (L, U) in $\mathcal{P}(\omega)/\mathsf{FIN}$ such that

1. $\{C \in \mathbb{A}_\beta : C \subseteq_\star C_\beta\} \subseteq \mathsf{L}$, $\{D \in \mathbb{A}_\beta : D \cap C_\beta =_\star\} \subseteq \mathsf{U}$,
2. $\{\mathsf{h}_\beta(C) : C \in \mathsf{L} \cap \mathsf{p}\} = \{\mathsf{h}_\beta(C) : C \in \mathsf{L}\} \cap \mathsf{q}$,
   $\{\mathsf{h}_\beta(C) : C \in \mathsf{U} \cap \mathsf{p}\} = \{\mathsf{h}_\beta(C) : C \in \mathsf{U}\} \cap \mathsf{q}$,
3. each $D \notin B_\beta$, which fills the gap $(\{\mathsf{h}_\beta(C) : C \in \mathsf{L}\}, \{\mathsf{h}_\beta(C) : C \in \mathsf{U}\})$ satisfies the following condition: for all $C \in \mathbb{A}_\beta$
   $(\ C \subseteq_\star C_\beta \iff \mathsf{h}_\beta(C) \subseteq_\star D\ )$ and $(\ C_\beta \subseteq C \iff D \subseteq_\star \mathsf{h}_\beta(C);\ )$.

Moreover, if $C_\beta \in \mathsf{p}$ the condition 2., and the assumption that q is a P-point implies that there is an element $D \in \mathsf{q} \setminus \mathbb{B}_\beta$, which fills the gap $(\{\mathsf{h}_\beta(C) : C \in \mathsf{L}\}$, $\{\mathsf{h}_\beta(C) : C \in \mathsf{U}\})$. Let $\overline{\mathbb{A}_\beta}$ $(\overline{\mathbb{B}_\beta})$ denote the subalgebra generated by $\mathbb{A}_\beta$ nad C_β ($\mathbb{B}_\beta$ nad D). Extend h_β to the isomorphism $\overline{\mathsf{h}}_\beta$ from $\overline{\mathbb{A}_\beta}$ onto $\overline{\mathbb{B}_\beta}$ with $\overline{\mathsf{h}}_\beta(C_\beta) = D$. If $D_\beta \in \mathbb{B}_\beta$ put $\mathbb{A}_{\beta+1} = \overline{\mathbb{A}_\beta}$, $\mathbb{B}_{\beta+1} = \overline{\mathbb{B}_\beta}$ and $\mathsf{h}_{\beta+1} = \overline{\mathsf{h}}_\beta$. Otherwise, repeat the construction for the generator D_β, $\mathbb{B}_\beta$ and h_β^{-1}.

After ω_1 steps we obtain a required isomorphism. □

Under CH, the structure of $\mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ is rather complicated. Indeed, if $\mathbb{A}$ is an arbitrary Boolean algebra of cardinality at most $\mathfrak{c}$, then there exists a subgroup of $\mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ which is isomorphic to the group of automorphisms of $\mathbb{A}$ ([4]).

6.3 Shelah's model

As one expect the case of negation of the CH is much more difficult. Investigations of the case we begin with the following result due to S. Shelah ([7]):

Theorem 3. *There exists a model in which $\mathfrak{c} = \omega_2$ and each automorphism of $\mathcal{P}(\omega)/\mathsf{FIN}$ is trivial.*

We shortly sketch out construction of the model. As a groud model take a constructible universe $\mathbb{L}$ ($\mathbb{V} = \mathbb{L}$). Thus GCH is satisfied and moreover the principle $\diamondsuit_{\omega_2}$ holds in the following form:

$\diamondsuit_{\omega_2}$. *There is a sequence $(\mathsf{D}_\alpha : \mathrm{cf}(\alpha) = \omega_1,\ \alpha < \omega_2)$, $\mathsf{D}_\alpha \subseteq \alpha$ such that for every $A \subseteq \omega_2$ the set $\{\alpha : \mathrm{cf}(\alpha) = \omega_1,\ A \cap \alpha = \mathsf{D}_\alpha\}$ is stationary in ω_2.*

For an automorphism $\mathsf{h} \in \mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ fix a function $\underset{*}{\mathsf{h}} : P(\omega) \to P(\omega)$ such that the following diagram commutes:

$$\begin{array}{ccc} P(\omega) & \xrightarrow{\underset{*}{\mathsf{h}}} & P(\omega) \\ \downarrow i & & \downarrow i \\ \mathcal{P}(\omega)/\mathsf{FIN} & \xrightarrow{\mathsf{h}} & \mathcal{P}(\omega)/\mathsf{FIN} \end{array} \qquad i : P(\omega) \ni A \to A^\star \in \mathcal{P}(\omega)/\mathsf{FIN}$$

The ground model will be extended by using forcing. We shall define, by induction, a finite support iteration $(\mathbb{P}_\alpha : \alpha < \omega_2)$ such that

1. $\mathrm{card}(\mathbb{P}_\alpha) < \omega_2$,
2. $\mathbb{P}_\alpha$ satisfies the countable chain condition,
3. in the resulting model all automorphisms are trivial.

The principle $\diamondsuit_{\omega_2}$ implies the following.

Claim 1. *Let $\mathbb{V} = \mathbb{L}$ and let $(\mathbb{P}_\alpha : \alpha < \omega_2)$ be a finite support iteration of forcings such that $\mathrm{card}(\mathbb{P}_\alpha) < \omega_2$ and $\mathbb{P}_\alpha$ satisfies the countable chain condition. Then there exists a sequence $(\mathsf{h}_\alpha : \alpha < \omega_2)$ with $\mathsf{h}_\alpha \in \mathbb{V}^{\mathbb{P}_\alpha}$ such that for arbitrary $\mathsf{h} \subseteq [P(\omega) \times P(\omega)]^{\mathbb{V}^{\mathbb{P}}}$, $\mathsf{h} \in \mathbb{V}^{\mathbb{P}}$ the set $\left\{\alpha < \omega_2 : \mathrm{cf}(\alpha) = \omega_1,\ \mathsf{h}|_{[P(\omega)]^{\mathbb{V}^{\mathbb{P}_\alpha}}} = \mathsf{h}_\alpha\right\}$ is stationary in ω_2, where $\mathbb{P} = \bigcup_{\alpha<\omega_2} \mathbb{P}_\alpha$.*

We intend to define $(\mathbb{P}_\alpha : \alpha < \omega_2)$ as in the above claim, so we have the set h_α predicting any automorphism $\mathsf{h} \in \mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ in $\mathbb{V}^{\mathbb{P}_{\omega_2}}$. Put

$$\mathbb{P}_0 = \{0\}\,,\ \mathbb{P}_\lambda = \bigcup_{\alpha<\lambda} \mathbb{P}_\alpha,\ \textit{for a limit ordinal } \lambda < \omega_2.$$

In order to construct $\mathbb{P}_{\alpha+1}$ we consider the set h_α. If $\mathrm{cf}(\alpha) \neq \omega_1$ or $\mathsf{h}_\alpha \in \mathbb{V}^{\mathbb{P}_\alpha}$ is not a non-trivial automorphism then $\widehat{\mathbb{Q}}_\alpha$ is a $\mathbb{P}_\alpha$-name for Cohen's forcing and $\mathbb{P}_{\alpha+1} = \mathbb{P}_\alpha\star = \mathbb{P}_\alpha \star \widehat{\mathbb{Q}}_\alpha$, otherwise we apply the following lemma:

Lemma 2. *Suppose that $\mathbb{V}$ is a model in which CH holds. Assume that h is non-trivial automorphism of $\mathcal{P}(\omega)/\mathsf{FIN}$. Then there exists a forcing notion $\mathbb{Q}$ of cardinality at most ω_1 which adds a subset $X \subset \omega$ such that if $\mathbb{C}$ is a Cohen's forcing then for every $Y \subset \omega$ added by $\mathbb{Q} \star \mathbb{C}$ there are (in the ground model) $A,\ B \subset \omega$ such that $\left(X \cap A =_\star B \text{ and } Y \cap \underset{*}{\mathsf{h}}(A) \neq_\star \underset{*}{\mathsf{h}}(B)\right)$ or $\left(X \cap A \neq_\star B \text{ and } Y \cap \underset{*}{\mathsf{h}}(A) = \underset{*}{\mathsf{h}}(B)\right)$*

Let $\widehat{\mathbb{Q}}_\alpha$ be the $\mathbb{P}_\alpha$-name for the forcing $\mathbb{Q}$, and $\mathbb{P}_{\alpha+1} = \mathbb{P}_\alpha \star \widehat{\mathbb{Q}}_\alpha$. Note that according to the lemma h_α cannot be extended onto $[P(\omega)]^{\mathbb{V}^{\mathbb{P}_{\alpha+1}}}$. Indeed, a new real X_α added by the forcing $\mathbb{Q}_\alpha$ witnesses that no $Y \in [P(\omega)]^{\mathbb{V}^{\mathbb{P}_{\alpha+1}}}$ is an appropriate candidate for the image under h_α of X_α. However, care must be taken to prevent adding such a candidate at some further step of inductive construction. To do this we use ω_1-oracle. $\bar{\mathrm{M}} = \langle M_\delta : \delta < \omega_1,\ \delta \text{ is limit ordinal}\rangle$ is called *an ω_1-oracle* if

- M_δ is a countable model of ZFC (or sufficiently large part of ZFC)

- $\delta \subseteq M_\delta$,
- $M_\delta \models' \delta$ *is countable'*,
- $(\forall\, A \subseteq \omega_1) \quad \{\delta : A \cap \delta \in M_\delta\}$ is stationary.

Denote by $\mathcal{D}_M$ the filter generated by sets $\{\delta < \omega_1 : A \cap \delta \in M_\delta\}$, $A \subseteq \omega_1$.

Let $\mathbb{P}$ be a forcing. We say that $\mathbb{P}$ *satisfies the* $\bar{\mathrm{M}}$*-c.c.* if one of the following conditions holds:

- (1) $\mathbb{P}$ is countable
- (2) $\mathrm{card}(\mathbb{P}) = \aleph_1$ and for some (every) injection $h : \mathbb{P} \to \omega_1$ the set of all $\delta < \omega_1$ which satisfies the following
  $(\forall\, A \subseteq \omega_1)$ *if* $A \in M_\delta,\ A \subseteq \delta,\ h^{-1}[A]$ *is predense in* $h^{-1}[\delta]$
  then $h^{-1}[A]$ *is predense in* $\mathbb{P}$
  belongs to the $\mathcal{D}_M$
- $\mathrm{card}(\mathbb{P}) > \aleph_1$ and for all $\mathbb{P}_0 \subseteq \mathbb{P}$, with $\mathrm{card}(\mathbb{P}_0) \leq \aleph_1$, there exists $\mathbb{P}_1 \subseteq \mathbb{P}$ such that $\mathbb{P}_0 \subseteq \mathbb{P}_1$ and $\mathbb{P}_1$ satisfies the condition (2)

At each stage α, we define (for $\mathbb{P}_\alpha$) an ω_1-oracles $\bar{\mathrm{M}}_\alpha \in \mathbb{V}^{\mathbb{P}_\alpha}$ such that

- If $\alpha \leq \beta < \omega_2$ and $\widehat{\mathbb{Q}}$ is $\mathbb{P}_\beta$-name for forcing such that $\mathrm{card}(\widehat{\mathbb{Q}}) < \omega_2$ and $\widehat{\mathbb{Q}}$ satisfies $\bar{\mathrm{M}}_\alpha$-c.c. then $\widehat{\mathbb{Q}}$ does not add any candidate for the image under h_α of X_α.
- If $\alpha \leq \beta < \omega_2$ then $\widehat{\mathbb{Q}}_\alpha$ satisfies $\bar{\mathrm{M}}_\alpha$-c.c.

Let $\mathbb{P} = \bigcup_{\alpha<\omega_2} \mathbb{P}_\alpha$. $\mathbb{P}$ satisfies the countable chain condition and has cardinality ω_2. Thus there are at most $\omega_2^{\omega_0} = \omega_2$ $\mathbb{P}$-names for reals. On the other hand we add ω_2 new reals thus $\mathbb{V}^{\mathbb{P}} \Vdash 2^{\omega_0} = \omega_2$. Assume that $\mathsf{h} \in \mathbb{V}^{\mathbb{P}} \cap \mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$. We show that h is trivial. Assume to the contrary that h is non-trivial. The set

$$\left\{\alpha < \omega_2 : \mathrm{cf}(\alpha) \neq \omega_1,\ or\ \mathsf{h}|_{[P(\omega)]^{\mathbb{V}^{\mathbb{P}_\alpha}}}\ is\ a\ non-trivial\ automorphism\ in\ \mathbb{V}^{\mathbb{P}_\alpha}\right\}$$

is closed and unbounded. Thus there exists $\alpha < \omega_2$ such that $\mathsf{h}|_{[P(\omega)]^{\mathbb{V}^{\mathbb{P}_\alpha}}} = \mathsf{h}_\alpha$. So at stage $\alpha+1$ we added X_α and an ω_1-oracle such that the image of X_α under h_α cannot exist in generic extension of $\mathbb{V}^{\mathbb{P}_\beta}$, $\beta > \alpha$ with forcing satisfying $\bar{\mathrm{M}}_\alpha$-c.c. $\mathsf{h} \in \mathbb{V}^{\mathbb{P}}$ is such an extension, a contradiction. □

6.4 Automorphisms under PFA

The forcing used by Shelah in his construction is proper. Shelah and Steprans in [9] showed that

Theorem 4. *PFA (the Proper Forcing Axiom) implies that all automorphisms of* $\mathcal{P}(\omega)/\mathsf{FIN}$ *are trivial.*

Recall that forcing $(\mathbb{P}, \prec)$ is *proper* if for every uncountable cardinal λ every stationary subset of $[\lambda]^\omega = \{C \subseteq \lambda : \mathrm{card}(C) = \omega\}$ remains stationary in the generic extension.

The Proper Forcing Axiom. [***PFA***] *If* $(\mathbb{P}, \prec)$ *is a proper forcing and if* $\mathcal{D}$ *is a collection of* ω_1 *dense subsets of* $\mathbb{P}$ *then there exists a* $\mathcal{D}$*-generic filter on* $\mathbb{P}$.

Proof of the theorem consists of three steps:
1. Show that under PFA all automorphisms are somewhere trivial.
2. Show that for a given automorphism $\mathsf{h} \in \mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ the set

$$T(\mathsf{h}) = \{A \subseteq \omega : \mathsf{h} \text{ is trivial on } A\}$$

is a P-ideal (i.e. $\{\omega \setminus A : A \in T(\mathsf{h})\}$ is a P-filter).
3. Show that PFA implies that every somewhere trivial automorphism is actually trivial.

The crucial part of the proof is the first step. Let $\mathbb{V}$ be a model in which PFA holds true and $2^{\omega_0} = \omega_2$. Assume that $\mathsf{h} \in \mathbb{V}$ is nowhere trivial automorphism which induces $\underset{*}{\mathsf{h}} : P(\omega) \to P(\omega)$ (like in the Shelah's construction). By $\mathbb{R}$ denote a contably closed forcing which forces $\diamondsuit$. Note that since $\mathbb{R}$ does not add any new reals the h remains nowhere trivial after forcing with $\mathbb{R}$.

Lemma 3. *Let* h *be nowhere trivial automorphism of* $\mathcal{P}(\omega)/\mathsf{FIN}$. *In the extension by the forcing* $\mathbb{R}$ *there is a family* $\{(A_\alpha, B_\alpha) : \alpha < \omega_1\}$*, a countable chain condition forcing* $\mathbb{P}$ *and a* $\mathbb{P}$*-name* X *such that:*

1. $\{A_\alpha : \alpha < \omega_1\}$ *is almost disjoint family (i.e.* $A_\alpha \cap A_\beta \in \mathsf{FIN}$ *if* $\alpha \neq \beta$*),*

2. $\mathbb{P} \Vdash' X \cap A_\alpha = B'_\alpha$,

3. $\mathbb{P} \Vdash' (\forall Y \subseteq \omega)\, (\exists \beta \in \omega_1)\, (\forall \gamma > \beta) \Big(Y \cap \underset{*}{\mathsf{h}}(A_\gamma) \neq_\star \underset{*}{\mathsf{h}}(B_\gamma)\Big)'$.

The family $\{(A_\alpha, B_\alpha) : \alpha < \omega_1\}$ is constructed by induction (in the construction assumption that the h is nowhere trivial is used) and is applied to define the forcing $\mathbb{P}$. $\mathbb{P}$ consists of the functions $f \in \{0,1\}^\omega$ such that:

(a) $\mathrm{dom}(f) =_\star \bigcup_{\alpha \in \mathsf{s}} A_\alpha$, where s is a finite subset of ω_1,

(b) if $\alpha \in \mathsf{s}$ then $f^{-1}(1) \cap A_\alpha =_\star B_\alpha$.

$\mathbb{P}$ is ordered by inverse inclusion. To see how the forcing can be applied suppose that $\mathbb{G}$ is $\mathbb{P}$-generic and let $X = \bigcup_{f \in \mathbb{G}} f^{-1}(1)$. Then X satisfies conditions 2 and 3. According to these conditions, X determines in $\mathcal{P}(\omega)/\mathsf{FIN}$ an (ω_1, ω_1) gap, namely $(\{A_\alpha \setminus B_\alpha : \alpha < \omega_1\}, \{B_\alpha : \alpha < \omega_1\})$. The element X separates the gap. To ensure that there is no element Y which fills the image of the gap we use the following lemma:

Lemma 4. *Let $\mathbb{R}$, $\mathbb{P}$ be as above. Then $\mathbb{R} \star \mathbb{P}$ forces that there is a proper forcing $\mathbb{Q}$ and a name Γ such that*

$\mathbb{Q} \Vdash'\, \Gamma$ *is an uncountable subset of* ω_1 *and if* $\alpha,\ \beta \in \Gamma,\ \alpha \neq \beta$ *then either* $\left(\underset{*}{\mathsf{h}}(A_\alpha) \setminus \underset{*}{\mathsf{h}}(B_\alpha)\right) \cap \underset{*}{\mathsf{h}}(B_\beta) \neq \emptyset$ *or* $\left(\underset{*}{\mathsf{h}}(A_\beta) \setminus \underset{*}{\mathsf{h}}(B_\beta)\right) \cap \underset{*}{\mathsf{h}}(B_\alpha) \neq \emptyset$.'

To check that $\left(\{\underset{*}{\mathsf{h}}(A_\alpha) \setminus \underset{*}{\mathsf{h}}(B_\alpha) : \alpha < \omega_1\}, \{\underset{*}{\mathsf{h}}(B_\alpha : \alpha < \omega_1)\}\right)$ is unfilled assume the opposite and derive a contradiction. Suppose that Y fills the gap ($\underset{*}{\mathsf{h}}(B_\alpha) \subseteq_\star Y$ and $\underset{*}{\mathsf{h}}(A_\alpha \setminus B_\alpha) \cap Y =_\star \emptyset$, for $\alpha < \omega_1$). By Δ lemma there are an uncountable $\Gamma_1 \subseteq \Gamma$, and finite subsets s, t of ω such that

$$\text{for every } \alpha \in \Gamma_1 \quad Y \cap (\underset{*}{\mathsf{h}}(A_\alpha) \setminus \underset{*}{\mathsf{h}}(B_\alpha)) = \mathsf{s} \quad \text{and} \quad \underset{*}{\mathsf{h}}(B_\alpha) \setminus Y = \mathsf{t}.$$

Fix $\alpha,\ \beta \in \Gamma_1,\ \alpha \neq \beta$. Since $(\underset{*}{\mathsf{h}}(A_\alpha) \setminus \underset{*}{\mathsf{h}}(B_\alpha)) \cap \underset{*}{\mathsf{h}}(B_\alpha) = \emptyset$ it follows that $\underset{*}{\mathsf{h}}(B_\alpha) \subseteq (Y \setminus \mathsf{s}) \cup \mathsf{t}$. Thus

$$(\underset{*}{\mathsf{h}}(A_\alpha) \setminus \underset{*}{\mathsf{h}}(B_\alpha)) \cap \underset{*}{\mathsf{h}}(B_\beta) \subseteq (\underset{*}{\mathsf{h}}(A_\alpha) \setminus \underset{*}{\mathsf{h}}(B_\alpha)) \cap ((Y \setminus \mathsf{s}) \cup \mathsf{t}) = \emptyset,$$

a contradiction.

The forcing $\mathbb{R} \star \mathbb{P} \star \mathbb{Q}$ is proper and for each $\mathsf{h} \in \mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ adds $X \subseteq \omega$ and $\Gamma_1 \subseteq \omega_1$ which witness that h have to be somewhere trivial. Slightly modifying the lemma 3 by adding new assumption:

(4) $\{A_\alpha : \alpha < \omega_1\} \subseteq T(\mathsf{h})$

and using the fact that $T(\mathsf{h})$ is a P-ideal we conclude that h is trivial. □

6.5 Automorphisms under OCA+MA(ω_1)

In this section we present a diffrent approach to $\mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ by studying the properties of the group using Ramsey forcing axioms which assert the existence of large homogeneous subsets in certain kinds of partitions. We shall outline the following result due to B. Velickovic ([13]):

Theorem 5. *Under OCA and MA(ω_1) every automorphism of $\mathcal{P}(\omega)/\mathsf{FIN}$ is trivial.*

The following version of the OCA will be used: Let X be a set of reals and $[\mathsf{X}]^2 = \{\{x, y\} : x, y \in \mathsf{X},\ x \neq y\}$. We say that $\mathsf{K} \subseteq [\mathsf{X}]^2$ is *open* if the set $\{(x, y) : \{x, y\} \in \mathsf{K}\}$ is open in the space $\mathsf{X} \times \mathsf{X}$ ($\subseteq \mathbb{R} \times \mathbb{R}$ with the product topology).

The Open Coloring Axiom (OCA). *Assume that X is a set of reals. For any partition $[\mathsf{X}] = \mathsf{K}_0 \times \mathsf{K}_1$ with K_0 open, either there is an uncountable $\mathsf{Y} \subseteq \mathsf{X}$ such that $[\mathsf{Y}]^2 \subseteq \mathsf{K}_0$ or there exist sets H_n, $n \in \omega$, such that $\mathsf{X} = \bigcup_{n\in\omega} \mathsf{H}_n$ and $[\mathsf{H}_n]^2 \subseteq \mathsf{K}_1$ for all n.*

A set Y such that $[\mathsf{Y}]^2 \subseteq \mathsf{K}_i$ is called K_i-homogeneous. Recall also the Martin's Axiom:

The Martin's Axiom (MA(ω_1)). *If $(\mathbb{P}, \prec)$ is a partially ordered set that satisfies the countable chain condition and $\mathcal{D}$ is a collection of at most ω_1 dense subsets of $\mathbb{P}$, then there exists a $\mathcal{D}-$generic filter on $\mathbb{P}$.*

We begin with a remark which holds in ZFC. Note that each subset A of ω can be identified wich its characteristic function χ_A. So $\mathcal{P}(\omega)$ can be identified with the Cantor cube $\{0, 1\}^\omega$ i.e. the set of all zero-one sequences(with the product topology). $\{0, 1\}^\omega$ is a complete separable metric space with no isolated points.

Remark 1. *([13]) Let $\mathsf{h} \in \mathcal{Aut}(\mathcal{P}(\omega)/\mathsf{FIN})$. Suppose that there are Borel functions $\mathsf{H}_n : \{0, 1\}^\omega \to \{0, 1\}^\omega$ such that for every $A \subseteq \omega$ there exists $n \in \omega$ such that $\underset{*}{\mathsf{h}}(A) =_\star \mathsf{H}_n (\chi_A)^{-1} (1)$. Then h is trivial.*

The proof of the theorem 1 consists of a sequence of lemmas. Fix an $\mathsf{h} \in \mathcal{Aut}(\mathcal{P}(\omega)/\mathsf{FIN})$ and $\underset{*}{\mathsf{h}} : P(\omega) \to P(\omega)$ (induced by h).

Lemma 5. *(OCA+MA(ω_1)) If h is non-trivial then there exists an uncountable almost disjoint family $\mathcal{A} = \{\mathsf{A}_\alpha : \alpha\Gamma\}$ such that h is non-trivial on every $\mathsf{A}_\alpha \in \mathcal{A}$.*

In order to show that each automorphism is trivial we construct a subfamily of $\mathcal{B}$ of $\mathcal{A}$ such that h is trivial on all elements of $\mathcal{B}$. It will be done in two steps: firstly, we decomposed elements of the family $\mathcal{A}$ to obtain two families which are *neat*, secondly, we show that h is trivial on all but countable many elements of a neat family.

We call a family $\mathcal{C}$ of almost disjoint subsets of ω *neat* if there is an injective map $\varphi : \omega \to \{0, 1\}^{<\omega}$ such that for each $\mathsf{C} \in \mathcal{C}$ and $k,\ l \in \mathsf{c}$, $\varphi(k) \subseteq \varphi(l)$ or $\varphi(l) \subseteq \varphi(k)$ (where $\{0, 1\}^{<\omega}$ denotes the set of all finite zero-one sequences).

Lemma 6. *(MA(ω_1)) Let $\mathcal{A} = \{\mathsf{A}_\alpha : \alpha \in \Gamma\}$ be an uncountable almost disjoint family of infinite subsets of ω. Then there is an uncountable $\Gamma_1 \subseteq \Gamma$ and for every $\alpha \in \Gamma_1$ a decomposition $\mathsf{A}_\alpha = \mathsf{A}^0_\alpha \cup \mathsf{A}^1_\alpha$ such that the family $\mathcal{C}^i \left\{\mathsf{A}^i_\alpha : \alpha \in \Gamma_1\right\}$ is neat for $i \in \{0, 1\}$.*

The following lemma finishes the proof:

Lemma 7. *(OCA) Let $\mathcal{C}$ be a neat family. Then h is trivial on all but countably many elements of $\mathcal{C}$.*

We sketch out the proof of the lemma (and show how OCA is applied). Fix a function $\varphi : \omega \to \{0, 1\}^\omega$ witnessing that $\mathcal{C}$ is neat. A topological space X (occuring in OCA) is defined as follows:

$$\mathsf{X} = \{(A, B) : \ (\exists C \in \mathcal{C}) \ B \subseteq A \subseteq C\}$$

and is eqquiped with the separable metric topology τ obtained by identifying (A, B) with $(A, B, \underset{*}{\mathsf{h}}(A), \underset{*}{\mathsf{h}}(B))$. Define the partition $[\mathsf{X}]^2 = \mathsf{K}_0 \cup \mathsf{K}_1$ by the condition:

$$\Big((A, B), (\tilde{A}, \tilde{B})\Big) \in \mathsf{K}_0 \text{ iff } \begin{cases} (a)\ \{\varphi(n) : n \in A\} \neq \{\varphi(n) : n \in \tilde{A}\}, \\ (b)\ A \cap \tilde{B} = \tilde{A} \cap B, \\ (c)\ \underset{*}{\mathsf{h}}(A) \cap \underset{*}{\mathsf{h}}(\tilde{B}) \neq \underset{*}{\mathsf{h}}(\tilde{A}) \cap \underset{*}{\mathsf{h}}(B). \end{cases}$$

K_0 is open in the topology of the $[\mathsf{X}]^2$. Firstly it is showed that there is no uncountable K_0-homogeneous subset of X (We skip the part of the proof). Thus OCA implies that X is countable union of pairwise disjoint the K_1-homogeneous sets X_n. For each an X_n fix a dense countable subset $\mathsf{D}_n \subseteq \mathsf{X}_n$. Define a family $\mathcal{B} = \{C \in \mathcal{C} : (\exists n \in \omega)(\exists (A, B) \in \mathsf{D}_n) \ B \subseteq A \subseteq C\}$. $\mathcal{B}$ is at most countable. We show that h is trivial on every $C \in \mathcal{C} \setminus \mathcal{B}$. So fix such a C. Using the assertion that the X_n is K_1-homogeneous we prove that C can be decomposed into two disjoint sets C_0, C_1 in such a way that for every $i \in \{0, 1\}$, $n \in \omega$, and $(A, B) \in \mathsf{X}_n$, if $A \subseteq C_i$ then for every $k \in \omega$ there exists $(\tilde{A}, \tilde{B}) \in \mathsf{D}_n$ such that:

(a) $A \cap \tilde{B} = \tilde{A} \cap B$,

(b) $A \cap \{0, 1, \ldots, m\} = \tilde{A} \cap \{0, 1, \ldots, m\}$ and $B \cap \{0, 1, \ldots, m\} = \tilde{B} \cap \{0, 1, \ldots, m\}$,

(c) $\underset{*}{\mathsf{h}}(A) \cap \{0, 1, \ldots, m\} = \underset{*}{\mathsf{h}}(\tilde{A}) \cap \{0, 1, \ldots, m\}$ and $\underset{*}{\mathsf{h}}(B) \cap \{0, 1, \ldots, m\} = \underset{*}{\mathsf{h}}(\tilde{B}) \cap \{0, 1, \ldots, m\}$.

Now define the function $\mathsf{H}_n : P(C_0) \to P(\omega)$, for $n \in \omega$, by putting:

$$\mathsf{H}_n(B) = \bigcup \left\{ \underset{*}{\mathsf{h}}(C_0) \cap \underset{*}{\mathsf{h}}(\tilde{B}) : (\tilde{A}, \tilde{B}) \in \mathsf{D}_n \text{ and } \tilde{A} \cap B = C_0 \cap \tilde{B} \right\}.$$

The H_n is a Borel function and conditions (a)-(c) imply that if $(C_0, B) \in \mathsf{X}_n$ then $\mathsf{H}_n(B) =_\star \underset{*}{\mathsf{h}}(B)$. Thus by the Remark 1 H_n is trivial on C_0. Using similar argument we show that H_n is trivial on C_1.

To sum up: we have shown that h is trivial on every element of uncoutable subfamily of arbitrary almost disjoint family of infinite subsets of ω, a contradiction. Thus h is trivial.

6.6 Non-trivial automorphisms

It has been shown that under CH there exist non-trivial automorphisms of $\mathcal{P}(\omega)/\mathsf{FIN}$ while several set-theoretical axioms imply that all automorphisms are trivial. One can ask if existence of non-trivial automorphism of $\mathcal{P}(\omega)/\mathsf{FIN}$ is equivalent to CH. Actually, it is not the case. One of the first results towards constructing non-trivial automorphism without CH was due to Frolik ([1]). He proved that

Lemma 8. *The set of fixed points of any injective function from an extremaly disconnected space to itself is clopen.*

We shall present construction of a non-trivial automorphism due to S. Shelah and J. Steprans. They apply the Frolik's result and the following lemma ([8]):

Lemma 9. *Suppose that $\mathcal{I}$ is an ideal on ω generating by an increasing (with respect to $\subseteq_\star$) sequence $(A_\alpha : \alpha < \kappa)$. Suppose further that f_α is almost permutation on A_α $(\alpha < \kappa)$ and $f_\alpha \subseteq_\star f_\beta$, if $\alpha < \beta$. Then the functions $\{f_\alpha : \alpha < \kappa\}$ induce an isomorphism h on subalgebra of $\mathcal{P}(\omega)/\mathsf{FIN}$ generated by $\mathcal{I}$ defined by*

$$\mathsf{h}(B) = \begin{cases} (f_\alpha[B])^\star, & \textit{if there is } \alpha \textit{ such that } B \subseteq_\star A_\alpha \\ (\omega \setminus f_\alpha[\omega \setminus B])^\star, & \textit{if there is } \alpha \textit{ such that } \omega \setminus B \subseteq_\star A_\alpha. \end{cases}$$

Assume that p is a P-point in $\omega^\star$. Take $(V_\alpha : \alpha < \omega_1)$ a neighbourhood base of clopen sets for p. Construct increasing sequence of permutations f_α on $A_\alpha = \omega \backslash V_\alpha$ such that homeomorphism of $A^\star_{\alpha+1} \setminus A^\star_\alpha$ is not the identity. By the lemma there is an automorphism $\mathsf{h} \in \mathcal{A}ut(\mathcal{P}(\omega)/\mathsf{FIN})$ and an autohomeomorphism f_h of $\omega^\star$ induced by the f_α's. This automorphim is non-trivial. To show this note that p is isolated fixed point of f_h. If f_h was trivial, then either it or its inverse would extend to an injective continuous function $\bar{f}_\mathsf{h}$ from $\beta[\omega]$ to $\beta[\omega \backslash \mathsf{s}]$, for $\mathsf{s} \in \mathsf{FIN}$. But p would remain an isolated fixed point of $\bar{f}_\mathsf{h}$. Since $\beta[\omega]$ and $\beta[\omega \backslash \mathsf{s}]$ are extremally disdonected the set $\{\mathsf{p}\}$ would be clopen, a contradiction. It follows that in any model containing P-point of character ω_1 there is a non-trivial automorphism. Note that there are such models with $\mathfrak{c} > \omega_1$. (A similar proof due to Baumgartner appears in [12]).

In [8] S. Shelah and J. Steprans have presented two another proofs of existence of non-trivial automorphisms in models in which the CH does not hold.

Theorem 6. *Assume that* $\mathbb{V} \Vdash$ CH *and enlarge* $\mathbb{V}$ *by adding* ω_2 *Cohen's reals. Then in the resulting model there is a non-trivial automotphism of* $\mathcal{P}(\omega)/\mathsf{FIN}$.

Note that the Shelah's model (described in the section 3) is constructed in such a way that each non-trivial automorphism is eliminated by adding new real to which it is impossible to extend the automorphism. One can ask whether it is possible to eliminate a non-trivial automorphism by adding a new generic almost permutation which turns the non-trivial automorphism to a trivial one. Of course, in some cases it can be possible but there are automorphisms which are in a sense absolutely non-trivial ([8]):

Theorem 7. *Assume that in the ground model* $\mathbb{V}$ *there is* $\mathcal{F} \subseteq \mathcal{P}(\omega)$ *a P-filter of character* ω_1. *Let* $\mathcal{I}$ *be its dual ideal. Then there is in* $\mathbb{V}$ *an automorphism of the Boolean algebra* $\mathbb{B} \subseteq \mathcal{P}(\omega)/\mathsf{FIN}$ *generated by* $\mathcal{I}$ *which is not induced by any function from* ω *to* ω. *Moreover, in any extension* $\mathbb{V}_1$ *of* $\mathbb{V}$, *which preserving* ω_1, *there is no almost permutation which turns the* h *to a trivial automorphism.*

Existence of non-trivial automorphism (in some models in which CH fails) also follows from the following theorem ([4]).

Theorem 8. *It is consistent with* ZFC+MA*(σ-linked) that the cardinality of the continuum is arbitrarily large and for aech Boolean algebra* $\mathbb{A}$ *of cardinality* $\leq \mathfrak{c}$, *there is an embedding* $\mathsf{i} : \mathbb{A} \to \mathcal{P}(\omega)/\mathsf{FIN}$ *such that each automorphism of* $\mathsf{i}[\mathbb{A}]$ *can be extended to an automorphism of* $\mathcal{P}(\omega)/\mathsf{FIN}$.

We conclude the this paper with the following result due to the author (in preparation).

Theorem 9. *Suppose that in a ground model* $\mathbb{V}$, $\kappa_1,\ \kappa_2,\ \kappa_3$ *are regular cardinals such that*

1. $\omega_2 < \kappa_1 \leq \kappa_2 \leq \kappa_3$,
2. $(\forall \lambda < \kappa_1) \quad 2^\lambda = \lambda^+$ *and if* $\kappa_1 = \lambda^+$ *then* $\mathrm{cf}(\lambda) > \omega$.
3. $2^{\kappa_1} = \kappa_3$
4. $\diamondsuit_{\kappa_1}$ *holds in the following form:*
   Let $H(\kappa_1)$ *denote the family of all sets of hereditary power* $< \kappa_1$ *and let* $(H_\alpha : \alpha < \kappa_1)$ *be a continuously increasing sequence of sets of cardinality* $< \kappa_1$ *such that*
   $$H(\kappa_1) = \bigcup_{\alpha<\kappa_1} H_\alpha.$$

Then there is a sequence $(T_\alpha : \alpha < \kappa_1, \operatorname{cf}(\alpha) = \omega_1)$ *such that for every set* X *of hereditary power* $< \kappa_1$*, the set*

$$\{\alpha < \kappa_1 : \operatorname{cf}(\alpha) = \omega_1,\, X \cap H_\alpha = T_\alpha\}$$

is stationary in κ_1*.*

5. $\mathbb{B}$ *is a Boolean algebra of cardinality* κ_1 *and* $\operatorname{card}(\mathcal{A}ut(\mathbb{B})) = \kappa_2$

Then there exists a generic extension $\mathbb{V}[\mathbb{G}]$*, which satisfies the following*

1. κ_1*,* κ_2*,* κ_3 *are regular cardinals (in* $\mathbb{V}[\mathbb{G}]$ *),* $2^{\omega_0} = \kappa_1$ *and* $2^{\kappa_1} = \kappa_3$*,*

2. *there exists a subalgebra* $\mathsf{P}_\mathbb{B}$ *of* $\mathcal{P}(\omega)/\mathsf{FIN}$ *which is isomorphic to* $\mathbb{B}^\mathbb{V}$ *and every automorphism of* $\mathcal{P}(\omega)/\mathsf{FIN}$ *can be (unquely) extended to an automorphism of* $\mathcal{P}(\omega)/\mathsf{FIN}$,

3. *each automorphism of* $\mathcal{P}(\omega)/\mathsf{FIN}$ *is a composition of trivial automorphisms and elements of* $\mathsf{P}_\mathbb{B}$.

We shortly outline how non-trivial automorphism are eliminated in the during construction of $\mathbb{V}[\mathbb{G}]$. Take a ground model $\mathbb{V}$ which satisfies the assumptions. Define a finite support iteration $(\mathbb{P}_\alpha : \alpha < \kappa_1)$. We use two (standard) notions of forcing: $\mathbb{Q}(\mathcal{L})$ which adds a real which separates an unfilled gap $\mathcal{L}$ and $\mathbb{E}(\mathcal{L})$ which guarantees that no forcing satifying the countable chain condition adds a generic real which separates $\mathcal{L}$ of the type $(\omega_1, \omega_1.)$ (If $\mathcal{L}$ is unfilled gap of the type (ω_1, ω_1) then $\mathbb{E}(\mathcal{L})$ adds un uncountable antichain to the $\mathbb{Q}(\mathcal{L})$.)

In the ground model $\mathbb{V}$ we choose the algebra $\mathbb{B}$. We enumerate the set of its generators, $\{b_\gamma : \gamma < \kappa_1\}$. We shall construct an embedding of $\mathbb{B}$ into $\mathcal{P}(\omega)/\mathsf{FIN}$ in such a way that images of generators b_γ will be generic sets added (by $\mathbb{Q}(\mathcal{L})$) to $\mathcal{P}(\omega)/\mathsf{FIN}$. Simultaneously automorphisms of the image of $\mathbb{B}$ will be extended. The sequence in which these automorphisms will appear in our construction is specified by $\diamondsuit_{\kappa_1}$. To extend the automorphism we will also use the forcing $\mathbb{Q}(\mathcal{L})$ and Sikorski's theorem.

Assume, that for some $\alpha < \kappa_1$

$$\mathbb{V}^{\mathbb{P}_\alpha} \Vdash' T_\alpha \textit{ is a non}-\textit{trivial automorphism} \text{ and } \mathsf{T}_\alpha \notin \mathcal{A}ut(\mathsf{P}_\mathbb{B}).'$$

We start to build an increasingly ordered gap

$$\mathcal{L} = (\{X_\beta : \beta < \omega_1\}, \{X_\beta : \beta < \omega_1\})$$

such that each trivial automorphism from $\mathbb{V}^{\mathbb{P}_\alpha}$ differs from any (non-trivial) extension of T_α on all (but at most countable many) elements of this gap. Moreover the further construction is done in such a way that if that after ω_1 steps at stage β some extension $\bar{\mathsf{T}}$ of T_α appears i.e.

$$\mathbb{P}_\beta \Vdash' \textit{There exists } \bar{\mathsf{T}},\ \mathsf{T}_\alpha \subseteq \bar{\mathsf{T}} \textit{such that } X_\beta,\ Y_\beta \in \mathrm{dom}(\bar{\mathsf{T}})'.$$

then the gap will satisfy the following conditions.

1. For each almost permutation f there exist $\beta < \omega_1$ and $A_f \subseteq X_\beta$ such that $\mathsf{T}(A_f^\star) \subseteq [f(Y_\beta)]^\star$,
2. $\mathsf{T}(\mathcal{L})$ does not "depend on" $\mathcal{L}$ which means that
   $\mathbb{Q}(\mathcal{L}) \nVdash$ "*$T(\mathcal{L})$ is separated*"
   and $\mathbb{E}(\mathcal{L}) \nVdash$ "*$\mathbb{Q}(T(\mathcal{L}))$ has an uncountable antichain*".

Then there are three posibilities:

1. There exists $\tilde{\mathcal{L}}$ such that $\mathbb{Q}(\tilde{\mathcal{L}}) \Vdash$ "*$T(\mathcal{L})$ is separated*" and we are going to force with $\mathbb{Q}(\tilde{\mathcal{L}})$, then $\mathbb{P}_{\beta+1} = \mathbb{P}_\beta \star \mathbb{E}(\mathcal{L})$;
2. There exists $\tilde{\mathcal{L}}$ such that $\mathbb{E}(\tilde{\mathcal{L}}) \Vdash$ "*$T(\mathcal{L})$ has an uncountable antichain*" then $\mathbb{P}_{\beta+1} = \mathbb{P}_\beta \star \mathbb{Q}(\mathcal{L})$;
3. Otherwise $\mathbb{P}_{\beta+1} = \mathbb{P}_\beta \star \mathbb{Q}(\mathcal{L})$.

This finishes the proof. □

References

[1] Z. Frolik. "Fixed points of maps of extremally disconnected spaces and complete Boolean algebras". In: *Bull. Acad. Pol. Sci. Sér. Sci. Math. Astron. Phys.* 16 (1968), pp. 269–275.

[2] L. Gillman and M. Henriksen. "Concerning rings of continuous functions". In: *Trans. Am. Math. Soc.* 77 (1954), pp. 340–362.

[3] L. Gillman and M. Henriksen. "Rings of continuous functions in which every finitely generated ideal is principal". In: *Trans. Am. Math. Soc.* 82 (1956), pp. 366–391.

[4] M. Grzech. "On automorphisms of Boolean algebras embedded in $\mathrm{P}(\omega)/\mathrm{fin}$". In: *Fundam. Math.* 150 (1996), pp. 127–147.

[5] I. I. Parovičenko. "A universal bicompact of weight ℵ". In: *Soviet Math. Dokl.* 4 (1963), pp. 592–595.

[6] W. Rudin. "Homogeneity problems in the theory of Čech compactifications". In: *Duke Math. J.* 23 (1956), pp. 409–419.

[7] S. Shelah. *Proper Forcing.* Springer-Verlag, Berlin etc., 1982.

[8] S. Shelah and J. Steprans. “Non-trivial homeomorphisms of $\beta\mathbb{N}\setminus\mathbb{N}$ without the continuum hypothesis”. In: *Fundam. Math.* 132, No. 2 (1989), pp. 135–141.

[9] S. Shelah and J. Steprans. “PFA implies all automorphisms are trivial”. In: *Proc. Am. Math. Soc.* 104, No. 4 (1988), pp. 1220–1225.

[10] R. Sikorski. *Boolean Algebras.* Springer, 1960.

[11] R. Sikorski. “On an analogy between measures and homomorphisms”. In: *Ann. Soc. Polon. Math.* 23 (1950), pp. 1–20.

[12] B. Velickovic. “Definable automorphisms of $\mathrm{P}(\omega)/\mathrm{fin}$”. In: *Proc. Am. Math. Soc.* 96 (1986), pp. 130–135.

[13] B. Velickovic. “OCA and automorphisms of $\mathrm{P}(\omega)/\mathrm{fin}$”. In: *Topology Appl.* 49, No. 1 (1993), pp. 1–13.

Chapter 7

Approximation of continuous and unbounded functions

MONIKA HERZOG

Monika Herzog
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: mherzog@pk.edu.pl

Abstract

In the chapter we study approximation properties of modified Szász-Mirakyan operators for real-valued functions, continuous and unbounded on $[0, \infty)$. The function to be approximated must satisfy some growth condition, so we consider weighted spaces. We present some direct theorems on approximation order of functions by these operators. We also consider a truncated version of Szász-Mirakyan operators which is more useful from the computational point of view.

Contents

7.1 Introduction

The approximation theory is a branch of mathematical analysis which is focused on approximation of complex functions by simpler and more easily calculated ones. It is obvious that the basis of the theory is the Weierstrass approximation theorem (1885), which states that every continuous function defined on a closed interval $[a, b]$ can be uniformly approximated by a sequence of polynomials. In 1912 S. N. Bernstein [4] gave a simple and effective method for constructing the sequence of polynomials that converges uniformly to the function, now called the sequence of Bernstein operators. These operators belong to the class of linear, positive operators. We present below some primary examples of such operators.

Let us denote by $C(I)$ the set of all real-valued functions continuous on I, where I is an interval in $\mathbb{R}$.

1. The Bernstein operators $B_n : C([0,1]) \to C([0,1])$ are defined in [4] by

$$B_n(f;x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right).$$

2. Let $0 \le \alpha \le \beta$ be real numbers.
The Bernstein-Stancu operators $P_n^{(\alpha,\beta)} : C([0,1]) \to C([0,1])$ are defined in [30] by

$$P_n^{(\alpha,\beta)}(f;x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k+\alpha}{n+\beta}\right).$$

3. The Landau operators $L_n : C([0,1]) \to C([0,1])$ are defined in [21] by

$$L_n(f;x) = \frac{\int_0^1 (1-(u-x)^2)^n f(u)du}{2\int_0^1 (1-u^2)^n du}.$$

4. Let $\alpha > 0$, D_4 be the set of all real-valued functions f, continuous on $[0,\infty)$ and such that $f(x) = O(x^\alpha)$ as $x \to \infty$.
The Szász-Mirakjan operators $S_n : D_4 \to C([0,\infty))$ are defined in [24, 31] by

$$S_n(f;x) = e^{-nx} \sum_{k=0}^{\infty} \frac{(nx)^k}{k!} f\left(\frac{k}{n}\right).$$

5. Let $\alpha > 0$, D_5 be the set of all real-valued functions f, continuous on $[0,\infty)$ and such that $f(x) = O(e^{\alpha x})$ as $x \to \infty$.
The Baskakov operators $V_n : D_5 \to C([0,\infty))$ are defined in [3] by

$$V_n(f;x) = \sum_{k=0}^{\infty} \binom{n+k-1}{k} \frac{x^k}{(1+x)^{n+k}} f\left(\frac{k}{n}\right).$$

6. Let $\alpha > 0$, D_6 be the set of all real-valued functions f, continuous on $[0,1)$ and such that $f(x) = O((1-x)^{-\alpha})$ as $x \to 1^-$.
   The Meyer-König and Zeller operators $M_n : D_6 \to C([0,1))$ are defined in [23] by

$$M_n(f;x) = \sum_{k=0}^{\infty} \binom{n+k}{k} x^k (1-x)^{n+1} f\left(\frac{k}{n+k+1}\right).$$

7. Let D_7 be the set of all real-valued functions, bounded and uniformly continuous on $[0,\infty)$.
   The Bleimann-Butzer-Hahn operators $H_n : D_7 \to C([0,\infty))$ are defined in [5] by

$$H_n(f;x) = \frac{1}{(1+x)^n} \sum_{k=0}^{n} \binom{n}{k} x^k f\left(\frac{k}{n-k+1}\right).$$

8. Let D_8 be the set of all real-valued functions, bounded and continuous on $\mathbb{R}$.
   The Gauss-Weierstrass operators $W_n : D_8 \to C(\mathbb{R})$ are defined in [35] by

$$W_n(f;x) = \sqrt{\frac{n}{2\pi}} \int_{\mathbb{R}} e^{-\frac{n}{2}(u-x)^2} f(u)du$$

9. Let D_9 be the set of all real-valued functions, bounded and continuous on $(0,\infty)$.
   The Post-Widder operators $P_n : D_9 \to C((0,\infty))$ are defined in [27] by

$$P_n(f;x) = \frac{1}{(n-1)!} \left(\frac{x}{n}\right)^n \int_0^{\infty} e^{-\frac{n}{x}u} u^{n-1} f(u)du.$$

10. Let D_{10} be the set of all real-valued functions, bounded and continuous on $(0,\infty)$.
    The gamma operators $G_n : D_{10} \to C((0,\infty))$ are defined in [22] by

$$G_n(f;x) = \frac{x^{n+1}}{n!} \int_0^{\infty} e^{-ux} u^n f\left(\frac{n}{u}\right)du.$$

11. Let $\alpha > 0$. $D_{11} = \{(x,t) : x > 0, 0 < t < \frac{1}{4\alpha}\}$, f is measurable function in $[0,\infty)$ and such that $f(x) = O(e^{\alpha x})$ as $x \to \infty$.
    The Wachnicki operators T_n are defined in [32] by

$$T_n(f;x,t) = f(0)e^{-\frac{x}{4t}} + \int_0^{\infty} f(s)H(s,x,t)ds,$$

where $H(s,x,t) = \frac{1}{4t}\exp(-\frac{s+x}{4t})\sqrt{\frac{x}{s}}I_1\left(\frac{\sqrt{sx}}{2t}\right)$, I_1 is the modified Bessel function of the first kind.

In the 20th century there were a lot of mathematicians who studied approximation properties of these operators, direct approximation theorems giving the degree of approximation by these operators and inverse theorems for them.
In 1950s T. Popoviciu [26], H. Bohman [6] and P. P. Korovkin [18] independently proposed sufficient conditions for the uniform convergence of the sequence of linear, positive operators.

Theorem 7.1.1. *If (A_n) is a sequence of positive, linear operators on $C([a,b])$ such that*

$$A_n(e_k) \to e_k \quad \textit{uniformly on } [a,b] \quad \textit{as} \quad n \to \infty,$$

where $e_k(x) = x^k$, $k = 0,1,2$, then for every $f \in C([a,b])$

$$A_n(f) \to f \quad \textit{uniformly on } [a,b] \quad \textit{as} \quad n \to \infty.$$

If the domain of the continuous function is unbounded but f has the finite limit at infinity then the convergence remains valid. In this case the test functions e_k are replaced, for instance, by $q_k(x) = e^{-kx}$, $k = 0,1,2$.
In 1974 A. D. Gadziev [10, 11] introduced the weighted spaces $C_p(I)$ to extend the Bohman-Korovkin theorem for continuous and unbounded functions defined on $[0,\infty)$.
The operators and many generalizations of them still are under consideration, [1, 2] [7, 8, 9] [15, 13, 14, 12, 16, 17], [20, 19, 25, 29, 28, 34, 33, 32].
In the end of the 20th century E. Wachnicki proposed a new modification of the well-known Szász-Mirakjan operators - A_n^ν defined below.

Let us define $\mathbb{R}_0 = [0,\infty)$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. In papers [15, 13] we consider operators of Szász-Mirakjan type defined as follows

$$A_n^\nu(f;x) = \begin{cases} \sum_{k=0}^{\infty} p_{n,k}^\nu(x) f(\frac{2k}{n}), & x > 0; \\ f(0), & x = 0 \end{cases}$$

and

$$p_{n,k}^\nu(x) = \frac{1}{I_\nu(nx)} \frac{(nx)^{2k+\nu}}{2^{2k+\nu}k!\Gamma(k+\nu+1)},$$

where Γ is the gamma function and I_ν stands for the modified Bessel function, i.e.,

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{z^{2k+\nu}}{2^{2k+\nu}k!\Gamma(k+\nu+1)}.$$

We study the operators in exponential weight spaces

$$E_q = \{f \in C(\mathbb{R}_0) : v_p\, f \text{ is uniformly continuous and bounded on } \mathbb{R}_0\},$$

where v_q is the exponential weight function defined as follows

$$v_q(x) = e^{-qx}, \quad q \in \mathbb{R}_0$$

for $x \in \mathbb{R}_0$. The spaces E_q are normed spaces with the norm

$$\|f\|_q = \sup_{x \in \mathbb{R}_0} v_q(x)|f(x)|.$$

The main aim of this survey paper is to present a brief overview of useful methods and tools for proving approximating theorems for these operators.

7.2 Auxiliary results

In papers [15, 13] we presented among other thing some preliminary properties of the operators A_n^ν. The purpose of this paragraph is to remind the reader of some basic results which we apply to prove main theorems.

Lemma 7.2.1. *For all $n \in \mathbb{N}$, $\nu \in \mathbb{R}_0$ and $x \in \mathbb{R}_0$*

$$A_n^\nu(e_0; x) = 1, \quad A_n^\nu(e_1; x) = x\frac{I_{\nu+1}(nx)}{I_\nu(nx)},$$

$$A_n^\nu(e_2; x) = x^2\frac{I_{\nu+2}(nx)}{I_\nu(nx)} + \frac{2x}{n}\frac{I_{\nu+1}(nx)}{I_\nu(nx)},$$

where $e_r(t) = t^r$, $r = 0, 1, 2$.

Lemma 7.2.2. *For all $n \in \mathbb{N}$, $\nu \in \mathbb{R}_0$ and $x \in \mathbb{R}_0$*

$$A_n^\nu(\phi_{x,1}; x) = x\left(\frac{I_{\nu+1}(nx)}{I_\nu(nx)} - 1\right),$$

$$A_n^\nu(\phi_{x,2}; x) = x^2\left(\frac{I_{\nu+2}(nx)}{I_\nu(nx)} + 2\frac{I_{\nu+1}(nx)}{I_\nu(nx)} + 1\right) + \frac{2x}{n}\frac{I_{\nu+1}(nx)}{I_\nu(nx)},$$

where $\phi_{x,r}(t) = (t - x)^r$, $r = 1, 2$.

Lemma 7.2.3. *For all $\nu \in \mathbb{R}_0$ there exists a positive constant $M(\nu)$ such that for all $n \in \mathbb{N}$ and $x \in \mathbb{R}_0$ we have*

$$\left|\frac{I_{\nu+1}(nx)}{I_\nu(nx)}\right| \le M(\nu), \quad nx\left|\frac{I_{\nu+1}(nx)}{I_\nu(nx)} - 1\right| \le M(\nu).$$

Lemma 7.2.4. *For all $\nu \in \mathbb{R}_0$ there exists a positive constant $M(\nu)$ such that for $x \in \mathbb{R}_0$ and $n \in \mathbb{N}$ we have*

$$|A_n^\nu(\phi_{x,1};x)| \le \frac{M(\nu)}{n}, \quad |A_n^\nu(\phi_{x,2};x)| \le M(\nu)\frac{x+1}{n}.$$

Moreover, from paper [13] we have the following lemmas.

Lemma 7.2.5. *For all $n \in \mathbb{N}$, $\nu, q \in \mathbb{R}_0$ and $x \in \mathbb{R}_0$ we have*

$$A_n^\nu\left(\frac{1}{v_q};x\right) = \frac{I_\nu(nxe^{\frac{q}{n}})}{I_\nu(nx)}e^{-\nu\frac{q}{n}}, \quad A_n^\nu\left(\frac{e_1}{v_q};x\right) = x\frac{I_{\nu+1}(nxe^{\frac{q}{n}})}{I_\nu(nx)}e^{(1-\nu)\frac{q}{n}},$$

$$A_n^\nu\left(\frac{e_2}{v_q};x\right) = x^2\frac{I_{\nu+2}(nxe^{\frac{q}{n}})}{I_\nu(nx)}e^{(2-\nu)\frac{q}{n}} + \frac{2x}{n}\frac{I_{\nu+1}(nxe^{\frac{q}{n}})}{I_\nu(nx)}e^{(1-\nu)\frac{q}{n}}.$$

Lemma 7.2.6. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(\nu, q)$ such that for all $n \in \mathbb{N}$ and $z \in \mathbb{R}_0$ we have*

$$\frac{I_\nu(ze^{\frac{q}{n}})}{I_\nu(z)} \le M(\nu,q)\exp\left(z(e^{\frac{q}{n}}-1)-\frac{q}{2n}\right).$$

Lemma 7.2.7. *For all $\nu \in \mathbb{R}_0$, $s > q > 0$ there exists a positive constant $M(\nu, q, s)$ such that for $n > n_0$ and $f \in E_q$ we have*

$$v_s(x)\left|A_n^\nu\left(\frac{\phi_{x,2}}{v_q};x\right)\right| \le M(\nu,q,s)\frac{x(x+1)}{n}.$$

In the same paper we prove

Lemma 7.2.8. *For all $\nu \in \mathbb{R}_0$ and $s > q > 0$ there exists a positive constant $M(\nu, q)$ such that for $n > n_0$ we have*

$$\left\|A_n^\nu\left(\frac{1}{v_q};\cdot\right)\right\|_s \le M(\nu,q),$$

where $n_0 = q(\ln\frac{s}{q})^{-1}$.

Theorem 7.2.1. *For all $\nu \in \mathbb{R}_0$, $s > q > 0$ there exists a positive constant $M(\nu, q)$ such that for $n > n_0$ and $f \in E_q$ we have*

$$\|A_n^\nu(f;\cdot)\|_s \le M(\nu,q)\|f\|_q.$$

This means that the operators A_n^ν are linear, positive, bounded and transform the space E_q into E_s for some $s > q$.
Notice that the operators A_n^ν do not transform the space E_q into itself. This follows from the following remark.

Remark 7.2.1. *For all $\nu \in \mathbb{R}_0$, $q > 0$ and $n \in \mathbb{N}$ we have*

$$\sup_{x \in \mathbb{R}_0} v_q(x) A_n^\nu \left(\frac{1}{v_q}; x \right) = \infty.$$

A certain modification of A_n^ν ([14]) yields a linear, positive and bounded operator $\overline{A}_n^\nu$ mapping the space E_q into itself, namely

$$\overline{A}_n^\nu(f;x) = \begin{cases} \sum_{k=0}^\infty p_{n,k}^\nu(x) f(\frac{2k}{n+q}), & x > 0; \\ f(0), & x = 0. \end{cases}$$

All the basic properties of the modification are still valid because of the following lemmas.

Lemma 7.2.9. *For all $n \in \mathbb{N}$, $\nu, q \in \mathbb{R}_0$ and $x \in \mathbb{R}_0$*

$$\overline{A}_n^\nu(e_0;x) = 1, \quad \overline{A}_n^\nu(e_1;x) = \frac{nx}{n+q} \frac{I_{\nu+1}(nx)}{I_\nu(nx)},$$

$$\overline{A}_n^\nu(e_2;x) = \left(\frac{nx}{n+q} \right)^2 \frac{I_{\nu+2}(nx)}{I_\nu(nx)} + \frac{2nx}{(n+q)^2} \frac{I_{\nu+1}(nx)}{I_\nu(nx)},$$

where $e_r(t) = t^r$, $r = 0, 1, 2$.

Lemma 7.2.10. *For all $n \in \mathbb{N}$, $\nu, q \in \mathbb{R}_0$ and $x \in \mathbb{R}_0$*

$$\overline{A}_n^\nu(\phi_{x,1};x) = x \left(\frac{n}{n+q} \frac{I_{\nu+1}(nx)}{I_\nu(nx)} - 1 \right),$$

$$\overline{A}_n^\nu(\phi_{x,2};x) = x^2 \left(\left(\frac{n}{n+q} \right)^2 \frac{I_{\nu+2}(nx)}{I_\nu(nx)} - \frac{2n}{n+q} \frac{I_{\nu+1}(nx)}{I_\nu(nx)} + 1 \right)$$

$$+ \frac{2nx}{(n+q)^2} \frac{I_{\nu+1}(nx)}{I_\nu(nx)},$$

where $\phi_{x,r}(t) = (t-x)^r$, $r = 1, 2$.

Lemma 7.2.11. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(q,\nu)$ such that for all $x \in \mathbb{R}_0$ and $n \in \mathbb{N}$ we have*

$$|\overline{A}_n^\nu(\phi_{x,1};x)| \le M(q,\nu) \frac{x+1}{n}, \quad |\overline{A}_n^\nu(\phi_{x,2};x)| \le M(q,\nu) \frac{x(x+1)}{n}.$$

Lemma 7.2.12. *For all $n \in \mathbb{N}$, $\nu, q \in \mathbb{R}_0$ and $x \in \mathbb{R}_0$ we have*

$$\overline{A}_n^{\nu}\left(\frac{1}{v_q}; x\right) = \frac{I_\nu(nxe^{\frac{q}{n+q}})}{I_\nu(nx)} e^{-\nu\frac{q}{n+q}}, \ \overline{A}_n^{\nu}\left(\frac{e_1}{v_q}; x\right) = \frac{nx}{n+q}\frac{I_{\nu+1}(nxe^{\frac{q}{n+q}})}{I_\nu(nx)} e^{\frac{(1-\nu)q}{n+q}},$$

$$\overline{A}_n^{\nu}\left(\frac{e_2}{v_q}; x\right) = \left(\frac{nx}{n+q}\right)^2 \frac{I_{\nu+2}(nxe^{\frac{q}{n+q}})}{I_\nu(nx)} e^{\frac{(2-\nu)q}{n+q}} + \frac{2nx}{(n+q)^2}\frac{I_{\nu+1}(nxe^{\frac{q}{n+q}})}{I_\nu(nx)} e^{\frac{(1-\nu)q}{n+q}}.$$

Lemma 7.2.13. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(\nu, q)$ such that for all $n \in \mathbb{N}$ and $z \in \mathbb{R}_0$ we have*

$$\frac{I_\nu(ze^{\frac{q}{n+q}})}{I_\nu(z)} \leq M(\nu, q)\exp\left(z(e^{\frac{q}{n+q}} - 1) - \frac{q}{2(n+q)}\right).$$

Lemma 7.2.14. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(q, \nu)$ such that for all $x \in \mathbb{R}_0$ and $n \in \mathbb{N}$ we have*

$$v_q(x)\left|\overline{A}_n^{\nu}\left(\frac{\phi_{x,2}}{v_q}; x\right)\right| \leq M(q, \nu)\frac{x(x+1)}{n}.$$

Lemma 7.2.15. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(q, \nu)$ such that for all $x \in \mathbb{R}_0$ and $n \in \mathbb{N}$ we have*

$$\left\|\overline{A}_n^{\nu}\left(\frac{1}{v_q}; \cdot\right)\right\|_q \leq M(q, \nu).$$

These lemmas allow one to write

Theorem 7.2.2. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(q, \nu)$ such that for all $x \in \mathbb{R}_0$, $n \in \mathbb{N}$ and $f \in E_q$ we have*

$$\left\|\overline{A}_n^{\nu}(f; \cdot)\right\|_q \leq M(q, \nu)\|f\|_q.$$

Hence the operators transfom the space E_q into itself.

7.3 Approximation theorems

In the paragraph we present approximation theorems, the inequalities which estimate the weighted error of approximation for function belonging to the spaces

$$E_q^{(k)} = \{f \in E_q : f^{(i)} \in E_q, i = 1, ..., k\},$$

where $f^{(i)}$ we denote the i-th derivative of f. We apply the weighted modulus of continuity of the first and the second order defined as follows,

$$\omega_q(f, E_q; t) = \sup\{\|\Delta_h f\|_q : h \in [0, t]\},$$

and

$$\omega_q^2(f, E_q; t) = \sup\{\|\Delta_h^2 f\|_q : h \in [0, t]\}, \tag{1}$$

respectively, where

$$\Delta_h f(x) = f(x+h) - f(x), \quad \Delta_h^2 f(x) = \Delta_h(\Delta_h f(x)) = f(x+2h) - 2f(x+h) + f(x)$$

for $x, h \in \mathbb{R}_0$.

We consider the case of the exponential weight spaces E_q ([16]) but we can assume that f is from polynomial weight spaces

$$C_p = \{f \in C(\mathbb{R}_0) : w_p f \text{ is uniformly continuous and bounded on } \mathbb{R}_0\},$$

where w_p is the polynomial weight function defined as follows

$$w_p(x) = \begin{cases} 1, & p = 0; \\ \frac{1}{1+x^p}, & p \in \mathbb{N}. \end{cases}$$

In this case we have analogous results.
As we mentioned in the end of Paragraph 1.1 we demonstrate typical proofs of several approximating theorems.

Theorem 7.3.1. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(\nu, q)$ such that for all $x \in \mathbb{R}_0$, $n \in \mathbb{N}$ and $g \in E_q^{(1)}$ we have*

$$v_q(x)|\overline{A}_n^\nu(g; x) - g(x)| \le M(q, \nu)\|g'\|_q \sqrt{\frac{x(x+1)}{n}}.$$

Theorem 7.3.2. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(\nu, q)$ such that for all $x \in \mathbb{R}_0$, $n \in \mathbb{N}$ and $f \in E_q$ we have*

$$v_q(x)|\overline{A}_n^\nu(f; x) - f(x)| \le M(q, \nu)\omega\left(f, E_q; \sqrt{\frac{x(x+1)}{n}}\right).$$

Proof. Let f_h be the Steklov mean of $f \in E_q$, i.e.,

$$f_h(x) = \frac{1}{h}\int_0^h f(x+t)dt$$

for $x \in \mathbb{R}_0$, $h > 0$. We have

$$f_h(x) - f(x) = \frac{1}{h}\int_0^h (f(x+t) - f(x))dt, \quad f'_h(x) = \frac{1}{h}(f(x+h) - f(x)).$$

It is easy to notice that if $f \in E_q$ then $f_h \in E_q^{(1)}$ for every fixed $h > 0$. Moreover, for $h > 0$

$$\|f_h - f\|_q \le \sup_{x\in\mathbb{R}_0}\left(\frac{1}{h}\int_0^h v_q(x)|f(x+t) - f(x)|dt\right) \le \omega(f, E_q; h)$$

and

$$\|f'_h\|_q \le \frac{1}{h}\omega(f, E_q; h)$$

hold. From the above inequalities, linearity of $\overline{A}_n^\nu$, Theorem 1.2.2 and Theorem 1.3.1 we get

$$\begin{aligned} & v_q(x)|\overline{A}_n^\nu(f;x) - f(x)| \\ \le\ & v_q(x)(|\overline{A}_n^\nu(f - f_h;x)| + |\overline{A}_n^\nu(f_h;x) - f_h(x)| + |f_h(x) - f(x)|) \\ \le\ & M(q,\nu)\left(\|f - f_h\|_q + \|f'_h\|_q\sqrt{\frac{x(x+1)}{n}}\right) \\ \le\ & M(q,\nu)\omega(f, E_q; h)\left(1 + \frac{1}{h}\sqrt{\frac{x(x+1)}{n}}\right) \end{aligned}$$

for $x \in \mathbb{R}_0$, $n \in \mathbb{N}$ and $h > 0$. Setting $h = \sqrt{\frac{x(x+1)}{n}}$ in the last expression we obtain the desired estimation.

The above theorem implies

Corollary 7.3.1. *If $\nu, q \in \mathbb{R}_0$ and $f \in E_q$, then for all $x \in \mathbb{R}_0$*

$$\lim_{n\to\infty}\{\overline{A}_n^\nu(f;x) - f(x)\} = 0.$$

Moreover, the above convergence is uniform on every interval $[x_1, x_2]$ with $0 \le x_1 < x_2$.

It is interesting to see a probabilistic approach to the proof of Corollary 1.3.1 which we present in paper [12]. The idea comes from [4].

Proof of Corollary 1.3.1. Let $f \in E_q$ with some $q \in \mathbb{R}_0$. Pick $x \in \mathbb{R}_0$ and $\varepsilon > 0$. There exists a number δ such that

$$|f(t) - f(x)| < \frac{\varepsilon}{2} \tag{2}$$

for $|t-x|<\delta$, $t\in\mathbb{R}_0$. By linearity of $\overline{A}_n^\nu$ and Lemma 1.2.9 we get

$$|\overline{A}_n^\nu(f;x)-f(x)|\le\overline{A}_n^\nu(|f-f(x)|;x)$$

$$=\sum_{|\frac{2k}{n+q}-x|<\delta}p_{n,k}^\nu(x)\left|f\left(\frac{2k}{n+q}\right)-f(x)\right|+\sum_{|\frac{2k}{n+q}-x|\ge\delta}p_{n,k}^\nu(x)\left|f\left(\frac{2k}{n+q}\right)-f(x)\right|$$

$$=I_1+I_2.$$

Hence by (2) we obtain $I_1<\frac{\varepsilon}{2}$. Furthermore, we have

$$I_2\le\frac{\|f\|_q}{\delta}\overline{A}_n^\nu\left(\frac{|\phi_{x,1}|}{v_q};x\right)+\frac{\|f\|_q}{\delta v_q(x)}\overline{A}_n^\nu(|\phi_{x,1}|;x).$$

Applying Hölder's inequality, Lemma 1.2.11, Lemmas 1.2.14, 1.2.15 and Theorem 1.2.2 we can write

$$I_2\le\frac{\|f\|_q}{\delta}\left[\overline{A}_n^\nu\left(\frac{\phi_{x,2}}{v_q};x\right)\right]^{\frac{1}{2}}\left[\overline{A}_n^\nu\left(\frac{1}{v_q};x\right)\right]^{\frac{1}{2}}+\frac{\|f\|_q}{\delta v_q(x)}[\overline{A}_n^\nu(\phi_{x,2};x)]^{\frac{1}{2}}$$

$$\le\frac{\|f\|_q}{\delta}\left[M(q,\nu)\frac{x(x+1)}{nv_q(x)}\right]^{\frac{1}{2}}\left[M(q,\nu)\frac{1}{v_q(x)}\right]^{\frac{1}{2}}+\frac{\|f\|_q}{\delta v_q(x)}\left[M(q,\nu)\frac{x(x+1)}{n}\right]^{\frac{1}{2}}<\frac{\varepsilon}{2}$$

The above estimations imply the convergence in Corollary 1.3.1.

Analogously as in paper [20] we define operators H_n^ν to estimate the error of approximation by the second modulus of continuity.

$$H_n^\nu(f;x)=\overline{A}_n^\nu(f;x)-f(\overline{A}_n^\nu(e_1;x))+f(x) \tag{3}$$

for $\nu,q\in\mathbb{R}_0$, $f\in E_q$ and $x\in\mathbb{R}_0$. Considering the basic properties of $\overline{A}_n^\nu$ we can write

$$H_n^\nu(f;x)=\overline{A}_n^\nu(f;x)-f\left(\frac{nx}{n+q}\frac{I_{\nu+1}(nx)}{I_\nu(nx)}\right)+f(x).$$

Observe that the operators are linear and possess the following properties

$$H_n^\nu(e_0;x)=1,\quad H_n^\nu(\phi_{x,1};x)=0.$$

In paper [16] we prove

Lemma 7.3.1. *For all $\nu,q\in\mathbb{R}_0$ there exists a positive constant $M(\nu,q)$ such that for all $x\in\mathbb{R}_0$, $n\in\mathbb{N}$ and $g\in E_q^{(2)}$ we have*

$$v_q(x)|H_n^\nu(g;x)-g(x)|\le M(\nu,q)\|g''\|_q\frac{x(x+1)}{n}.$$

Theorem 7.3.3. *For all $\nu, q \in \mathbb{R}_0$ there exists a positive constant $M(\nu, q)$ such that for all $x \in \mathbb{R}_0$, $n \in \mathbb{N}$ and $f \in E_q$ we have*

$$v_q(x)|\overline{A}_n^\nu(f;x) - f(x)|$$

$$\leq M(q,\nu)\omega_q^2\left(f, E_q; \sqrt{\frac{x(x+1)}{n}}\right) + \omega_q(f, E_q; |\overline{A}_n^\nu(\phi_{x,1}, x)|).$$

Proof. Let $x \in \mathbb{R}_0$ and f_h be the second order Steklov mean of $f \in E_q$, i.e.,

$$f_h(x) = \frac{4}{h^2}\int_0^{\frac{h}{2}}\int_0^{\frac{h}{2}}(2f(x+s+t) - f(x+2(s+t)))dsdt$$

for $x \in \mathbb{R}_0$ and $h > 0$. Notice that

$$f(x) - f_h(x) = \frac{4}{h^2}\int_0^{\frac{h}{2}}\int_0^{\frac{h}{2}}\Delta_{s+t}^2 f(x)dsdt$$

for $x \in \mathbb{R}_0$ and $h > 0$. By definition (1) we get the following estimations

$$\|f - f_h\|_q \leq \omega_q^2(f, E_q; h)$$

and since

$$f_h''(x) = \frac{1}{h^2}\left(8\Delta_{h/2}^2 f(x) - \Delta_h^2 f(x)\right)$$

we can write

$$\|f_h''\|_q \leq \frac{9}{h^2}\omega_q^2(f, E_q; h).$$

The above inequalities imply that the Steklov mean f_h and f_h'' belong to E_q. Moreover, by linearity of $\overline{A}_n^\nu$ and connection (3) we have

$$|\overline{A}_n^\nu(f;x) - f(x)|$$

$$\leq H_n^\nu(|f - f_h|;x) + |f(x) - f_h(x)| + |H_n^\nu(f_h;x) - f_h(x)|$$
$$+|f(\overline{A}_n^\nu(e_1;x)) - f(x)|.$$

Applying above estimations, Theorem 1.2.2 and Lemma 1.3.1 we obtain

$$v_q(x)|\overline{A}_n^\nu(f;x) - f(x)|$$

$$\leq v_q(x)H_n^\nu(|f - f_h|;x) + v_q(x)|f(x) - f_h(x)|$$

$$+v_q(x)|H_n^\nu(f_h;x) - f_h(x)| + v_q(x)|f(\overline{A}_n^\nu(e_1;x)) - f(x)|$$

$$\leq M(q,\nu)\|f-f_h\|_q + M(q,\nu)\|f_h''\|_q \frac{x(x+1)}{n} + v_q(x)|f(\overline{A}_n^\nu(e_1;x)) - f(x)|$$
$$\leq M(q,\nu)\omega_q^2(f,E_q;h)\left(1+\frac{1}{h^2}\frac{x(x+1)}{n}\right) + \omega_q(f,E_q;|\overline{A}_n^\nu(e_1;x)|),$$

where $\overline{A}_n^\nu(e_1;x) = x\left(\frac{n}{n+q}\frac{I_{\nu+1}(nx)}{I_\nu(nx)} - 1\right)$. Substituting $h = \sqrt{\frac{x(x+1)}{n}}$ we get the assertion of our theorem.

In paper [25] a certain modification of the operators A_n^ν is introduced

$$B_n^\nu(f,a_n;x) = \begin{cases} \sum_{k=0}^{[n(x+a_n)]} p_{n,k}^\nu(x) f(\frac{2k}{n}), & x>0; \\ f(0), & x=0, \end{cases}$$

for $f \in C_p$, where (a_n) is a sequence of positive numbers such that

$$\lim_{n\to\infty} a_n\sqrt{n} = \infty \tag{4}$$

and $[n(x+a_n)]$ we denote the integral part of $n(x+a_n)$.

In the paper there is deduced among other things that

$$\lim_{n\to\infty}\{B_n^\nu(f,a_n;x) - f(x)\} = 0$$

for every $f \in C_p$, uniformly on every interval $[x_1,x_2]$, $x_2 > x_1 \geq 0$.
In that case the crucial assumption is that the sequence of positive numbers (a_n) satisfy condition (4).

Similar problems for Baskakov type operators are discussed in paper [34] and some generalization of truncated operators can be found in [33].

We propose ([12]) the following class of partial operators for functions $f \in E_q$

$$\overline{B}_n^\nu(f,a_n;x) = \begin{cases} \sum_{k=0}^{[(n+q)(x+a_n)]} p_{n,k}^\nu(x) f(\frac{2k}{n+q}), & x>0; \\ f(0), & x=0, \end{cases}$$

where we replace the infinite summing in $\overline{A}_n^\nu$ by the finite one and we still have assumption (4).

Theorem 7.3.4. *If $\nu, q \in \mathbb{R}_0$ and $f \in E_q$ then for all $x \in \mathbb{R}_0$*

$$\lim_{n\to\infty}\{\overline{B}_n^\nu(f,a_n;x) - f(x)\} = 0.$$

Moreover, the above convergence is uniform on every set $[x_1,x_2]$, $x_2 > x_1 \geq 0$.

As we can see both operators B_n^ν and $\overline{B}_n^\nu$ have the same degree of approximation, namely $O(n^{-1/2})$.

References

[1] A. Aral. “A generalization of Szász-Mirakyan operators based on q-integers”. In: *Math. Comput. Modelling* 47, No. 9-10 (2008), pp. 1052–1062.

[2] A. Aral and V. Gupta. “On the Durrmeyer type modification of the q-Baskakov type operators”. In: *Nonlinear Anal., Theory Methods Appl., Ser. A, Theory Methods* 72, No. 3-4 (2010), pp. 1171–1180.

[3] V. A. Baskakov. “An instance of a sequence of linear positive operators in the space of continuous functions”. In: *Dokl. Akad. Nauk SSSR* 113 (1957), pp. 249–251.

[4] S. N. Bernstein. “Démonstration du théorème de Weierstrass fondée sur le calcul de probabilités”. In: *Charkow Ges. (2)* 13 (1912), pp. 1–2.

[5] G. Bleimann, P. L. Butzer, and L. Hahn. “A Bernstein-type operator approximating continuous functions on the semi-axis”. In: *Indag. Math.* 42 (1980), pp. 255–262.

[6] H. Bohman. “On approximation of continuous and of analytic functions”. In: *Ark. Mat.* 2 (1952), pp. 43–56.

[7] A. Ciupa. “Positive linear operators for the approximation in polynomial weighted spaces”. In: *Bull. Allahabad Math. Soc.* 24, No. 2 (2009), pp. 233–240.

[8] B. Della Vecchia, G. Mastroianni, and J. Szabados. “Weighted approximation of functions by Szász-Mirakyan-type operators”. In: *Acta Math. Hung.* 111, No. 4 (2006), pp. 325–345.

[9] S. Ersan and O. Doğru. “Statistical approximation properties of q-Bleimann, Butzer and Hahn operators”. In: *Math. Comput. Modelling* 49, No. 7-8 (2009), pp. 1595–1606.

[10] A. D. Gadziev. “The convergence problem for a sequence of positive linear operators on unbouded sets, and theorems analogous to that of P. P. Korovkin”. In: *Sov. Math., Dokl.* 15 (1974), pp. 1433–1436.

[11] A. D. Gadziev. “Theorems of Korovkin type”. In: *Math. Notes* 20 (1977), pp. 995–998.

[12] M. Herzog. “A note on the convergence of partial Szász-Mirakyan type operators”. In: *Ann. Univ. Paedagog. Crac. Stud. Math.* 13 (2014), pp. 45–50.

[13] M. Herzog. “Approximation of functions from exponential weight spaces by operators of Szász-Mirakyan type”. In: *Ann. Soc. Math. Pol., Ser. I, Commentat. Math.* 43, No. 1 (2003), pp. 77–94.

[14] M. Herzog. "Approximation of functions of two variables from exponential weight spaces". In: *Technical Transactions. Fundamental Sciences* 1-NP (2012), pp. 3–10.

[15] M. Herzog. "Approximation theorems for modified Szász-Mirakjan operators in polynomial weight spaces". In: *Matematiche* 54, No. 1 (1999), pp. 77–90.

[16] M. Herzog. "The degree of approximation of functions from exponential weight spaces". In: *Technical Transactions. Fundamental Sciences (to appear)* ().

[17] A. Holhoş. "Uniform Approximation by Positive Linear Operators on Noncompact Intervals". In: *Automat. Comput. Appl. Math.* 18, No. 1 (2009), pp. 121–132.

[18] P. P. Korovkin. "On the convergence of linear positive operators in the space of continuous functions". In: *Dokl. Akad. Nauk SSSR, n. Ser.* 90 (1953), pp. 961–964.

[19] G. Krech. "On the rate of convergence for modified Gamma operators". In: *Rev. Unión Mat. Argent.* (). (to appear).

[20] G. Krech. "Some approximation results for operators of Szász-Mirakyan-Durrmeyer type". In: *Math. Slovaca (to appear)* ().

[21] E. Landau. "Über die Approximation einer stetiegen Funktion durch eine ganze rationale Funktion". In: *Palermo Rend.* 25 (1908), pp. 337–346.

[22] A. Lupas and M. Müller. "Approximationseigenschaften der Gammaoperatoren". In: *Math. Z.* 98 (1967), pp. 208–226.

[23] W. Meyer-König and K. L. Zeller. "Bernsteinsche Potenzreihen". In: *Stud. Math.* 19 (1960), pp. 89–94.

[24] G. M. Mirakjan. "Approximation des fonctions continues au moyen de polynomes de la forme $e^{-nx}\sum_{k=0}^{m_n} C_{k,n}x^k$". In: *Comptes rendus de l'Académie des sciences de l'URSS* 31 (1941), pp. 201–205.

[25] R. N. Mohapatra and Z. Walczak. "Remarks on a class of Szász-Mirakyan type operators". In: *East J. Approx.* 15 (2) (2009), pp. 197–206.

[26] T. Popoviciu. "Asupra demonstra'tiei teoremei lui Weierstrass cu ajutorul polinoamelor de interpolare". In: *Lucrarile Sesiunii Generale Stiintifice, Editura Academiei Republicii Populare Romane, Bucuresti* 2-12 (1951), pp. 1664–1667.

[27] E. L. Post. "Generalized differentiation". In: *Transactions A. M. S.* 32 (1930), pp. 723–781.

[28] L. Rempulska and M. Skorupka. "Approximation by generalized MKZ-operators in polynomial weighted spaces". In: *Anal. Theory Appl.* 23, No. 1 (2007), pp. 64–75.

[29] L. Rempulska and Z. Walczak. "On some properties of Szász-Mirakyan operators in Hölder spaces". In: *Turk. J. Math.* 27, No. 3 (2003), pp. 435–446.

[30] D. D. Stancu. "On a generalization of Bernstein polynomials (Romanian)". In: *Stud. Univ. Babeş-Bolyai, Ser. Math.-Phys.* 14, No. 2 (1969), pp. 31–44.

[31] O. Szász. "Generalization of S. Bernstein's Polynomials to the Infinite Interval". In: *Journal of Research of the National Bureau of Standards* 45 (1950), pp. 239–245.

[32] E. Wachnicki. "A generalized Mazhar-Totik operators". In: *Demonstratio Math.* 38 (1) (2005), pp. 143–151.

[33] Z. Walczak. "Approximation theorems for a general class of truncated operators". In: *Appl. Math. Comput.* 217, No. 5 (2010), pp. 2142–2148.

[34] Z. Walczak and V. Gupta. "A note on the convergence of Baskakov type operators". In: *Appl. Math. Comput.* 202, No. 1 (2008), pp. 370–375.

[35] K. Weierstrass. "Über die analytische Darstellbarkeit sogenannter willkürlicher Functionen einer reellen Veränderlichen". In: *Berl. Ber.* (1885), pp. 633–639, 789–806.

Chapter 8

A survey on known values and bounds on the Shannon capacity

Marcin Jurkiewicz

Marcin Jurkiewicz[1]
Gdansk University of Technology
Department of Algorithms and System Modelling
Gabriela Narutowicza 11/12, 80-233 Gdansk, Poland
e-mail: marcin.jurkiewicz@eti.pg.gda.pl

Abstract

In this survey we present exact values and bounds on the Shannon capacity for different classes of graphs, for example for regular graphs and Kneser graphs. Additionally, we show a relation between Ramsey numbers and Shannon capacity.

Contents

[1]Supported in part by the Polish National Science Centre grant 2011/02/A/ST6/00201.

8.1 Introduction

By a *simple graph* G we mean an ordered pair of sets V and E such that the set E consists of two-element subsets of V, that is $E \subset [V]^2$. The elements of the set V we call *vertices* of a simple graph G, and the elements of the set E are called its *edges*. A simple graph we will shortly call a *graph*. If, at the same time, we want to indicate that the given set is the set of vertices of the graph G, then we write $V(G)$. Analogically we write for the set of edges.

Let us consider a graph G. If $v, w \in V(G)$ and $\{v, w\} \in E(G)$, then we say that v *is adjacent to* w. The *degree* of a vertex $v \in V(G)$ is the number of all vertices adjacent to v.

We can perform many operations on graphs. By analogy to the algebra of sets, in graph theory we introduce the relation of containing. We have $H \subset G$, if $V(H) \subset V(G)$ and $E(H) \subset E(G)$. Then we say that H is a *subgraph* of G. In the case when $V(H) \subset V(G)$ and $E(H) = E(G) \cap [V(H)]^2$, the graph H is an *induced subgraph* of G, and then we note this property by $H \sqsubset G$. A next operation that we will use in this survey is the operation of complement. A *complement* of a graph $G = (V, E)$ is a graph $\overline{G}$ on a set of vertices V, the set of edges of which is $[V]^2 \setminus E$. A graph G is called a self-complementary graph if $\overline{G} = G$. A next operation that we will describe is the operation of exponentiation of a graph. The *k-th power* of a graph G is a graph $G^k = (V, E_k)$ such that for each pair of vertices v and w we have $\{v, w\} \in E_k$ if and only if $\mathrm{dist}_G(v, w) \le k$, where dist_G is the distance between vertices of a graph. Another operation creating a graph from already existing graphs is the so called strong product of two graphs. Given two graphs $G = (V(G), E(G))$ and $H = (V(H), E(H))$, the strong product $G \boxtimes H$ is defined as follows. The vertices of $G \boxtimes H$ are all pairs of the Cartesian product $V(G) \times V(H)$. There is an edge between (v, v') and (w, w') if and only if one of the following holds, $\{v, w\} \in E(G)$ and $\{v', w'\} \in E(H)$, or $v = w$ and $\{v', w'\} \in E(H)$, or $v' = w'$ and $\{v, w\} \in E(G)$. Shortly, we write $G^{\boxtimes p}$, what means $G \boxtimes G \boxtimes \cdots \boxtimes G$, where G appears p times.

Among graphs, we can distinguish graphs of special structure, that is, from all graphs we can extract some classes of graphs. One of such classes is the class of *complete graphs*. A graph $G = (V, E)$ is a complete graph, if its set of vertices is $[V]^2$. Complete graphs are denoted by K_n, where n is the number of vertices of such graph. Complete graphs and complements of complete graphs are closely related to the concept of a clique and the concept of an independent set, respectively. A *clique* in a graph $G = (V, E)$ is a subset $V^{'} \subset V$ such that for every two vertices $u, v \in V^{'}$ we have $\{u, v\} \in E$. If for every two vertices $u, v \in V$ we have $\{u, v\} \notin E$, then $V^{'}$ is called an *independent set*. Other classes of graphs the definitions of which will now be given are paths and cycles. A *path* is a non-empty graph $P = (V, E)$, in which $V = \{v_1, \ldots, v_n\}$,

$E = \{\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{n-1}, v_n\}\}$, where v_i for $i = 1, \ldots, n$ are pairwise distinct. A path on n vertices we denote by P_n. Let us consider a path P_n, where $n \geq 3$. A graph obtained from adding a single edge $\{v_n, v_1\}$ is called a *cycle* on n vertices. Let us observe that cycles are so called *regular graphs*, that is, graphs in which all vertices have the same degree. Among regular graphs we distinguish vertex-transitive graphs. A graph is *vertex-transitive*, if for each pair of vertices (i, j) there exists an automorphism π such that $\pi(i) = j$. Analogically, we define *edge-transitive graphs*. Interestingly, the aforementioned cycles are graphs, which are vertex-transitive and edge-transitive. Another subclass belonging to those two classes are Kneser graphs. Let q and r be integers such that $q \geq 2r$. A *Kneser graph* $K(q, r)$ is a graph such that its set of vertices is a set of r-element subsets of the set $[q]$. The set of edges of this graph we define in such a way that between two vertices of this graph there is an edge if and only if the corresponding sets are disjoint. A next important class of graphs that will be discussed in this survey are the so called circulant graphs. Let $D = \{d_1, \ldots, d_l\}$ be a subset of the set $[n]$, where $n \geq 3$. A *circulant graph* $C_n(D)$ is a graph on n vertices such that $\{i, j\} \in E$ if and only if there exists $k \in [l]$ such that $j \equiv i \pm d_k \pmod{n}$. If $G = C_n(\{d_1, \ldots, d_l\})$, then we will shortly write $G = C_n(d_1, \ldots, d_l)$. If $G = C_n(\{1, \ldots, p\})$, then we will shortly write $G = C_{n,p}$.

The size of a largest independent set in a graph G is called the *independence number* of G, and we denote it by $\alpha(G)$. Another function defined on graphs is the *fractional independence number* [17], which is given by the below formula

$$\alpha^*(G) = \max_f \sum_{v \in V(G)} f(v), \tag{8.1}$$

where f is a non-negative function such that for each clique C of a graph G we have

$$\sum_{v \in C} f(v) \leq 1. \tag{8.2}$$

For graphs we also define the so called Lovász number [15]. Consider a graph G of order n, and let $\mathcal{A}_G$ be the set of real symmetric matrices A indexed by $V(G)$ that satisfy $A_{vw} = 1$, if $v = w$ or if v and w are non-adjacent. The *Lovász number* $\vartheta(G)$ is defined by

$$\vartheta(G) = \inf_{A \in \mathcal{A}_G} \lambda_1(M), \tag{8.3}$$

where $\lambda_1(M)$ denotes the largest eigenvalue of M. It is noteworthy that $\alpha(G) \leq \vartheta(G) \leq \alpha^*(G)$.

To the above measures is related the following measure, the so called *Shannon capacity* [18], defined as follows.

$$\Theta(G) = \lim_{i\to\infty} \sqrt[i]{\alpha(G^{\boxtimes i})}. \tag{8.4}$$

More information on this measure can be found in [14, 11].

We next give fundamental theorems concerning the strong product. Three first theorems that will be given, in the literature are treated as fundamental properties of graph products.

Theorem 8.1.1 (**Folklore**). *Let G' and G be graphs such that $G' \subset G$ and $p \geq 1$. Then*

$$G'^{\boxtimes p} \subset G^{\boxtimes p}. \tag{8.5}$$

Theorem 8.1.2 (**Folklore**). *Let G be a graph and let $n \geq 1$. Then*

$$\alpha(K_n \boxtimes G) = \alpha(G) \tag{8.6}$$

Theorem 8.1.3 (**Folklore**). *Let G and H be graphs. Then*

$$\alpha(G \boxtimes H) \geq \alpha(G)\alpha(H). \tag{8.7}$$

The following two theorems are the commonly known upper bounds on the independence number of a strong product of graphs.

Theorem 8.1.4 (**Rosenfeld [17]**). *For every graphs G and H we have*

$$\alpha(G \boxtimes H) \leq \alpha(G)\alpha^*(H). \tag{8.8}$$

Theorem 8.1.5 (**Lovász [15]**). *For every graph G we have*

$$\alpha(G) \leq \Theta(G) \leq \vartheta(G). \tag{8.9}$$

Let us observe that the Lovász number is multiplicative, what is given in the following theorem.

Theorem 8.1.6 (**Lovász [15]**). *For every graphs G and H we have*

$$\vartheta(G \boxtimes H) = \vartheta(G)\vartheta(H). \tag{8.10}$$

The next two theorems can be obtained directly from the definition of the Shannon capacity.

Theorem 8.1.7 (**Shannon [18]**). *For every graphs G and H we have*

$$\Theta(G \boxtimes H) \geq \Theta(G)\Theta(H). \tag{8.11}$$

Theorem 8.1.8 (**Shannon [18]**). *Let G be a graph and let i be a positive integer. Then*

$$\Theta(G) \geq \sqrt[i]{\alpha(G^{\boxtimes i})}. \tag{8.12}$$

Interestingly, there exist graphs for which $\Theta(G) > \alpha(G)$. From the information theory point of view, the graphs achieving the above inequality describe channels creating an advantage in the zero-error communication [14].

8.2 Regular graphs

The Shannon capacity is upper bounded by the Lovász number, that is, for every graph G we have $\Theta(G) \leq \vartheta(G)$ (Theorem 8.1.5). We now give an upper bound on the Lovász number for regular graphs.

Theorem 8.2.1 (**Hoffman [15]**). *Let G be a regular graph on n vertices, and let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of its adjacency matrix A. Then*

$$\vartheta(G) \leq \frac{-n\lambda_n}{\lambda_1 - \lambda_n}. \tag{8.13}$$

The equality is satisfied if the graph G is edge-transitive.

Lovász, in the acknowledgments of his work [15], states that Hoffman is the author of the above theorem. It is worth to note that edge-transitive graphs do not have to be regular, for example a path on three vertices. However, vertex-transitive graphs are always, by definition, regular. For the latter, Lovász gave the following theorem.

Theorem 8.2.2 (**Lovász [15]**). *Let G be a vertex-transitive graph on n vertices. Then*

$$\vartheta(G)\vartheta(\overline{G}) = n. \tag{8.14}$$

From Theorems 8.1.5 and 8.2.2 we easily get the following inequality,

$$\Theta(G)\Theta(\overline{G}) \leq n, \tag{8.15}$$

where G is a vertex-transitive graph. The equality is satisfied if $G = \overline{G}$, that is, for self-complementary vertex-transitive graphs. For these graphs the value of the Shannon capacity is known.

Theorem 8.2.3 (**Lovász [15]**). *Let G be a self-complementary vertex-transitive graph on n vertices. Then*

$$\Theta(G) = \vartheta(G) = \sqrt{n}. \tag{8.16}$$

Additionally, Sonnemann, for vertex-transitive graphs, proved a theorem that helps to determine the value of the lower bound on the Shannon capacity.

Theorem 8.2.4 (Sonnemann [19]). *Let G be a vertex-transitive graph. Then for any graph H,*

$$\alpha(G \boxtimes H) = \min_{G' \sqsubset G} \frac{|V(G)|}{|V(G')|}\alpha(G' \boxtimes H) \tag{8.17}$$

If in Theorem 8.2.4 the subgraph G' will be a maximal clique of the graph G, then using Theorem 8.1.2 we will get the following useful corollary.

Corollary 8.2.1. *Let G be a vertex-transitive graph. Then for any graph H,*

$$\alpha(G \boxtimes H) \leq \left\lfloor \frac{|V(G)|}{\omega(G)}\alpha(H) \right\rfloor. \tag{8.18}$$

We will now present more detailed results concerning subclasses of regular graphs.

8.3 Circulant graphs

Circulant graphs are a large subclass of the class of vertex-transitive graphs. To that class belong, among others, cycles, for which Lovász [15] gave the following result.

Theorem 8.3.1 (Lovász [15]). *For odd n,*

$$\vartheta(C_n) = \frac{n\cos(\pi/n)}{1 + cos(\pi/n)}. \tag{8.19}$$

Next, Brimkov et al. [5] determined the Lovász number for circulant graphs of the form $C_n(1, d)$, where $d = 2, 3$.

Theorem 8.3.2 (Brimkov et al. [5]). *For odd n,*

$$\vartheta(C_n(1,2)) = n\left(1 - \frac{0.5 - \cos(ax) - \cos(x(a+1))}{(\cos(ax) - 1)(\cos(ax+1) - 1)}\right) \tag{8.20}$$

and

$$\vartheta(C_n(1,3)) = n\left(1 - \frac{\cos^2 y - \cos y \cos(bx) + \cos^2(bx) - 1)}{(\cos y + 1)(\cos(bx) - 1)(1 - \cos y + \cos(bx))}\right) \tag{8.21}$$

for odd n, where $x = \frac{2\pi}{n}$, $y = \frac{\pi}{n}$, $a = \lfloor \frac{n}{3} \rfloor$ and $b = \lceil \frac{n-3}{6} \rceil$.

Directly from Theorem 8.2.2 we get the value of the Lovász number for complements of graphs appearing in the last two theorems.

There are many results concerning the independence number of strong products of graphs of the form $C_n(1, \dots, k)$. One of the fundamental results is the following theorem generalizing a result independently obtained by Hales [10] and Baumert et al. [2] for cycles.

Theorem 8.3.3 (**Badalyan et al. [1]**). *Let $m, n \geq 3$ and $k, l \geq 1$. Then*

$$\alpha(C_m^l \boxtimes C_n^k) = \min\{\lfloor \alpha^*(C_m^l)\alpha(C_n^k)\rfloor, \lfloor \alpha(C_m^l)\alpha^*(C_n^k)\rfloor\}, \tag{8.22}$$

particularly

$$\alpha(C_n^k \boxtimes C_n^k) = \left\lfloor \frac{n}{k+1} \left\lfloor \frac{n}{k+1} \right\rfloor \right\rfloor. \tag{8.23}$$

Let us observe that it is easy to obtain an upper bound on the maximal independent set following from formula (8.23), using Corollary 8.2.1 and formulas $\alpha(C_n^k) = \lfloor n/(k+1) \rfloor$. Jurkiewicz et al. [12] obtained partial results for the third power of circulant graphs $C_n(1, \dots, k)$, which are given in Table 8.1.

In the case when $k = 1$, that is, for cycles, several other exact formulas are known. The first one is a consequence of fundamental properties, which were given earlier and of the fact that $\alpha(C_5^2) = 5$.

Theorem 8.3.4 (**Lovász [15]**). *Let $i \geq 1$. Then*

$$\alpha(C_5^{\boxtimes 2i}) = \vartheta^{2i}(C_5) = 5^i. \tag{8.24}$$

Despite the Shannon capacity was determined for C_5 and it is $\Theta(C_5) = \sqrt{5}$ (from Theorem 8.2.3), the number $\alpha(C_5^{\boxtimes d})$ is not known for odd integers $d \geq 5$. Sonnemann et al. [19] showed that $\alpha(C_5^{\boxtimes 3}) = 10$. It is worth to note that this result was inspired by Hales [10]. Independently, the latter and Baumert et al. [2] gave the following important result.

Theorem 8.3.5 (**Baumert et al. [2], Hales [10]**). *Let $d \geq 1$ and $k \geq 1$. Then*

$$\alpha(C_{k2^d+1}^{\boxtimes d}) = k(k2^d + 1)^{d-1}. \tag{8.25}$$

Another result for the independence number of strong powers of cycles on $k2^d + r$ vertices is given in the following theorem.

Theorem 8.3.6 (**Baumert et al. [2]**). *Let $d \geq 1$ and $k \geq 1$. Then*

$$\alpha(C_{k2^d+3}^{\boxtimes d}) = \frac{k(k2^d + 3)^d + 1}{k2^d + 1}. \tag{8.26}$$

$n \setminus k$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
4	8	1	1	1	1	1	1	1	1	1
5	10	1	1	1	1	1	1	1	1	1
6	27	8	1	1	1	1	1	1	1	1
7	33	8	1	1	1	1	1	1	1	1
8	64	12	8	1	1	1	1	1	1	1
9	81	27	8	1	1	1	1	1	1	1
10	125	30	10	8	1	1	1	1	1	1
11	148	40 36	13	8	1	1	1	1	1	1
12	216	64	27	8	8	1	1	1	1	1
13	247	73 69	29 27	10	8	1	1	1	1	1
14	343	84 79	33	14	8	8	1	1	1	1
15	390 382	125	41 36	27	10	8	1	1	1	1
16	512	138 133	64	28 27	12	8	8	1	1	1
17	578	158 149	70 64	33 30	14	9	8	1	1	1
18	729	216	81	36	27	10	8	8	1	1
19	807	240 224	90 82	41 36	28 27	10	9	8	1	1
20	1000	266 247	125	64	30	14	10	8	8	1

Table 8.1: Exact values and bounds on $\alpha((C_n^k)^{\boxtimes 3})$. The values known earlier are given in the first column [2, 7]. These values are partially present in Table 8.2. The remaining non-trivial cases were recently determined in a work of [12]

Sonnemann proved the above theorem performing a different reasoning [19]. Recently, Bohman et al. [4] gave a precise result for cycles on $k2^d + r$ vertices.

Theorem 8.3.7 (Bohman et al. [4]). *Let $k \geq 1$. If $2k+1$ is prime, then*

$$\alpha(C_{8k+5}^{\boxtimes 3}) = (8k+5)\frac{(2k+1)(8k+5)-1}{2}. \tag{8.27}$$

Generally, the following inequality is satisfied.

$$\alpha(C_{8k+5}^{\boxtimes 3}) \leq t_k + 2, \tag{8.28}$$

where t_k is the right hand side of the equation (8.27).

It is worth to mention one of the most wanted values of the combinatorial information theory, that is, concerning the problem of Shannon capacity for channels described by the cycle C_7. Despite it has been over 50 years from the time when Shannon defined his measure, the value of the Shannon capacity for a small graph C_7 and larger odd cycles has not been determined yet. Although, some steps have been made. Baumert et al. [2] numerically determined that $\alpha(C_7^{\boxtimes 3}) = 33$. Later, Vesel and Žerovnik found an independent set of size 108 of the graph $C_7^{\boxtimes 4}$ [20]. From Theorems 8.1.5 and 8.1.8 and the facts given earlier, we are able to determine the following lower and upper bounds

$$3.2237 \leq \Theta(C_7) \leq 3.3176. \tag{8.29}$$

Other known values and bounds for cycles are given in Tables 8.2 and 8.3.

$n \setminus p$	**1**	**2**	**3**	**4**	**5**	**6**	$\tilde{p}$	θ
5	2	5	10	25	50	125	**2**	2.2361
7	3	10	33	115 108	402 343	1101	**4**	3.2237
9	4	18	81	363 324	1458	6561	**3**	4.3267
11	5	27	148	814 761	3996	21904	**3**	5.2896
13	6	39	247	1531	9633	61009	**3**	6.2743
15	7	52	390 382	2770	19864	145924	**3**	7.2558
17	8	68	578	4913	39304	334084	**4**	8.3721
19	9	85	807	7666	68994	651610	**4**	9.3571
21	10	105	1092	11441	114660	1201305	**4**	10.3423
23	11	126	1437	16466	181126	2074716	**4**	11.3278
25	12	150	1875	23125	281250	3515625	**4**	12.3316
27	13	175	2362	31522	413350	5579044	**4**	13.3246
29	14	203	2929	42017	594587	8579041	**4**	14.3171

Table 8.2: It presents exact values or bounds on the independence number $\alpha(C_n^{\boxtimes p})$ for $5 \leq n \leq 29$ [18, 2, 20, 7, 4]. Let us observe that in the table, only odd cycles were given. For even cycles, the independence number is known. In the last column, there was given a lover bound θ for $\Theta(C_n)$, obtained from Theorem 8.1.8 from among six exact values or lower bounds, appearing in the same row and indexed by the letter p. The number $\tilde{p}$ indicates the smallest p, for which the most accurate bound is achieved.

At the end, we will present results concerning the complements of odd cycles. Bohman et al. [3] proposed a brilliant construction that let them show the following theorem.

Theorem 8.3.8 (Bohman et al. [3]). *Let $k \geq 0$. Then*

$$\alpha(\overline{C}_{2k+3}^{2^k}) \geq 2^{2^k} + 1. \tag{8.30}$$

Thus, from Theorem 8.1.8 we get $\Theta(\overline{C}_{2k+3}^{2^k}) \geq (2^{2^k} + 1)^{1/2^k}$.

$n \setminus p$	**1**	**2**	**3**	**4**	**5**	**6**	$\tilde{p}$	θ
31	15	232	3580	54934	830560	12816400	**4**	15.3095
33	16	264	4356	71874	1185921	18974736	**5**	16.3988
35	17	297	5197	90947	1591572	27056724	**5**	17.3926
37	18	333	6142	113586	2101333	37824138	**5**	18.3865
39	19	370	7195	140211	2734074	51947406	**5**	19.3804
41	20	410	8405	171462	3510825	70644025	**5**	20.3743
43	21	451	9696	207514	4454896	94012416	**5**	21.3682
45	22	495	11115	249005	5591997	123543225	**5**	22.3621
47	23	540	12666	296439	6950358	160427556	**5**	23.3562
49	24	588	14406	352947	8588377	207532836	**4**	24.3740
51	25	637	16244	414196	10492965	263867536	**4**	25.3689
53	26	689	18232	483091	12721391	332849699	**4**	26.3638
55	27	742	20377	560244	15313309	415701048	**4**	27.3586
57	28	798	22743	646551	18311493	517244049	**4**	28.3564
59	29	855	25222	742247	21761957	636149284	**4**	29.3520
61	30	915	27877	848193	25714075	777127129	**4**	30.3476
63	31	976	30712	965097	30220701	943226944	**4**	31.3432

Table 8.3: It presents exact values or bounds on the independence number $\alpha(C_n^{\boxtimes p})$ for $31 \leq n \leq 63$ [2, 7, 4]. In the last column, there was given a lower bound θ for $\Theta(C_n)$, obtained from Theorem 8.1.8 from among six exact values or lower bounds, appearing in the same row and indexed by the letter p. The number $\tilde{p}$ indicates the smallest p, for which the most accurate bound is achieved.

8.4 Kneser graphs

Another class of graphs, for which the exact value of Shannon capacity is known, is the class of Kneser graphs. Since the 60s, the value of the independence number of these graphs is known.

Theorem 8.4.1 (Erdős et al. [8]). *Let $q \geq 2r$. Then*

$$\alpha(K(q,r)) = \binom{q-1}{r-1}. \tag{8.31}$$

It turns out that it is possible to determine the eigenvalues of the adjacency matrix of the Kneser graph.

Theorem 8.4.2 (Lovász [15]). *Let $q \geq 2r$. The eigenvalues of the Kneser graph $K(q,r)$ are the numbers*

$$(-1)^t\binom{q-r-t}{r-t}, \qquad t=0,1,\ldots,r.$$

The above results let us determine the Shannon capacity for Kneser graphs.

Theorem 8.4.3 (Lovász [15]). *Let $q \geq 2r$. Then*

$$\Theta(K(q,r)) = \binom{q-1}{r-1}. \tag{8.32}$$

What is interesting, the value of the Shannon capacity is also partially known for the complements of the Kneser graphs $K(q,r)$. Some of the values of the capacity we can determine by the following theorem.

Theorem 8.4.4 (Brouwer et al. [6]). *Let $q \geq 2r$ and $p \geq 2$. Then*

$$\alpha(\overline{K(q,r)}^{\boxtimes p}) \leq \left\lfloor \frac{q}{r}\alpha(\overline{K(q,r)}^{\boxtimes (p-1)})\right\rfloor. \tag{8.33}$$

Additionally, when we consider the first strong power of the graph $\overline{K(q,r)}$, that is for $p=1$, then we get $\alpha(\overline{K(q,r)}) = \lfloor \frac{q}{r} \rfloor$. In this case it is the largest number of pairwise disjoint sets in the set of all r-element subsets of the set $[q]$. We now get

$$\alpha(\overline{K(q,r)}^{\boxtimes p}) \leq \left\lfloor \frac{q}{r}\left\lfloor \frac{q}{r}\cdots\left\lfloor \frac{q}{r}\right\rfloor\right\rfloor\right\rfloor, \tag{8.34}$$

where the floor appears p times. The equality is satisfied for $p=1,2$ [6]. Additionally, if $r|q$, then for every $i \geq 1$ we have

$$\frac{q}{r} = \alpha(\overline{K(q,r)}) \leq \sqrt[i]{\alpha(\overline{K(q,r)}^{\boxtimes i})} = \frac{q}{r}. \tag{8.35}$$

Thus, if $r|q$, then $\Theta(\overline{K(q,r)}) = \frac{q}{r}$. In the remaining cases the problem is still open.

Conjecture 8.4.1. *Let $q \geq 2r$. Then*

$$\Theta(\overline{K(q,r)}) = \frac{q}{r}. \tag{8.36}$$

Let us consider a subclass of the complements of the Kneser graphs, the so called class of *triangular graphs* $T_q = \overline{K(q,2)}$, where $q \geq 4$. T_q contains an induced cycle C_q, thus from Theorem 8.1.1 we get

$$\alpha(T_q^{\boxtimes p}) \geq \alpha(C_q^{\boxtimes p}), \tag{8.37}$$

where $p \geq 1$. Another lower bound on the independence number of strong products of triangular graphs is given in the following theorem.

Theorem 8.4.5 (Mathew et al. [13]). *Let $p \geq 1$ and $q \geq 4$. Then*

$$\alpha(\overline{K(q+2,2)}^{\boxtimes p}) \geq 1 + \sum_{i=1}^{p} \binom{p}{i} \alpha(\overline{K(q,2)}^{\boxtimes i}). \tag{8.38}$$

The above results lead to many new exact values and bounds on the Shannon capacity and the independence number $\alpha(\overline{K(q,2)}^{\boxtimes p})$, which are given in Table 8.4.

$q \setminus p$	**1**	**2**	**3**	**4**
5	2	5	12	27
7	3	10	35	122 114
9	4	18	81	364 327
11	5	27	148	814 776
13	6	39	253 248	1644 1551
15	7	52	390 384	2925 2802
17	8	68	578	4913
19	9	85	807	7666
21	10	105	1102 1092	11571 11441

Table 8.4: It shows the values or lower and upper bounds on the independence number $\alpha(\overline{K(q,2)}^{\boxtimes p})$. It is worth to notice that the columns with $p = 1, 2$ are identical with the appropriate columns in Table 8.2.

8.5 Graphs with fixed independence number

The *Ramsey number* $R(l_1, ..., l_k)$ is defined as the smallest integer n such that, no matter how each 2-element subset of an n-element set is colored with k colors, there exists an $i \leq k$ such that there is a subset of size l_i, all of whose 2-element subsets are color i. For instance, $R(3,3) = 6$, $R(4,4) = 18$ [16].

In 1971, Erdős et al. [9] formulated results that can be written in the form of the following theorem joining Ramsey numbers [16] with the independence number of a graph.

Theorem 8.5.1 (Erdős et al. [9]). *For arbitrary graphs $G_1, G_2, \ldots, G_n$,*

$$\alpha(G_1 \boxtimes G_2 \boxtimes \cdots \boxtimes G_n) < R(\alpha(G_1) + 1, \alpha(G_2) + 1, \ldots, \alpha(G_n) + 1), \tag{8.39}$$

and for all $k_1, k_2, \ldots, k_n > 0$ there exist graphs G_i with $\alpha(G_i) = k_i, 1 \leq i \leq n$, such that

$$\alpha(G_1 \boxtimes G_2 \boxtimes \cdots \boxtimes G_n) = R(k_1 + 1, \ldots, k_n + 1) - 1. \tag{8.40}$$

Furthermore, for the diagonal case $k_i = k$*, there exists a single graph* G *with* $\alpha(G) = k$*, such that* $\alpha(G^{\boxtimes n}) = R_n(k+1) - 1$.

It is worth to note that, from the above theorem for any $k \geq 3$,

$$\lim_{n\to\infty} R_n(k)^{1/n} \tag{8.41}$$

is equal to the supremum of the Shannon capacity $\Theta(G)$ for all graphs G with the independence number equal to $k-1$. The best lower bound for this limit in the case when $k = 3$ is $3.199\ldots$ [22, 21].

8.6 Conclusions

This survey of the results concerning the Shannon capacity confirms the interest of the researchers in this topics. It is worth to notice that many works related to the problem of Shannon capacity were published in the last five years, thus the author plans to enrich the current text in the future.

References

[1] S. H. Badalyan and S. E. Markosyan. "On the independence number of the strong product of cycle-powers". In: *Discrete Math.* 313, No. 1 (2013), pp. 105–110.

[2] L. D. Baumert et al. "A combinatorial packing problem". In: *Computers in algebra and number theory (Proc. SIAM-AMS Sympos. Appl. Math., New York, 1970).* Amer. Math. Soc., Providence, R. I., 1971, pp. 97–108.

[3] T. Bohman and R. Holzman. "A nontrivial lower bound on the Shannon capacities of the complements of odd cycles". In: *IEEE Trans. Inf. Theory* 49, No. 3 (2003), pp. 721–722.

[4] T. Bohman, R. Holzman, and V. Natarajan. "On the independence numbers of the cubes of odd cycles". In: *Electron. J. Combin.* 20, No. 3 (2013), Research Paper P10.

[5] V. Brimkov. "Algorithmic and explicit determination of the Lovász number for certain circulant graphs". In: *Discrete Appl. Math.* 155, No. 14 (2007), pp. 1812–1825.

[6] A. E. Brouwer and A. Schrijver. "Uniform hypergraphs". In: *Mathematical Centre Tracts* 106 (1976), pp. 39–73.

[7] B. Codenotti, I. Gerace, and G. Resta. "Some remarks on the Shannon capacity of odd cycles". In: *Ars Comb.* 66 (2003), pp. 243–257.

[8] P. Erdös, K. Chao, and R. Rado. "Intersection theorems for systems of finite sets". In: *Q. J. Math. Ox. II. Ser.* 12 (1961), pp. 313–320.

[9] P. Erd
os, R. J. McEliece, and H. Taylor. "Ramsey bounds for graph products". In: *Pac. J. Math.* 37 (1971), pp. 45–46.

[10] R. S. Hales. "Numerical invariants and the strong product of graphs". In: *J. Comb. Theory, Ser. B* 15 (1973), pp. 146–155.

[11] M. Jurkiewicz and M. Kubale. "Shannon capacity and Ramsey numbers". In: *Proc. XIV International Conference of System Modelling and Control.* 2011.

[12] M. Jurkiewicz, M. Kubale, and K. Ocetkiewicz. "On the Shannon capacity of circulant graphs". preprint.

[13] A. M. Kizhakkepallathu, P. R. Östergård, and A. Popa. "On the Shannon Capacity of Triangular Graphs." In: *Electron. J. Comb.* 20, No. 2 (2013).

[14] J. Körner and A. Orlitsky. "Zero-error information theory". In: *IEEE Trans. Inf. Theory* 44, No. 6 (1998), pp. 2207–2229.

[15] L. Lovász. "On the Shannon capacity of a graph". In: *IEEE Trans. Inf. Theory* 25, No. 1 (1979), pp. 1–7.

[16] S. P. Radziszowski. "Small Ramsey numbers". In: *Electron. J. Comb.* Dynamic Survey 1 (1994).

[17] M. Rosenfeld. "On a problem of C. E. Shannon in graph theory". In: *Proc. Am. Math. Soc.* 18 (1967), pp. 315–319.

[18] C. E. Shannon. "The zero error capacity of a noisy channel". In: *Institute of Radio Engineers, Transactions on Information Theory* IT-2 (1956), pp. 8–19.

[19] E. Sonnemann and O. Krafft. "Independence numbers of product graphs". In: *J. Comb. Theory, Ser. B* 17 (1974), pp. 133–142.

[20] A. Vesel and J. Žerovnik. "Improved lower bound on the Shannon capacity of C_7". In: *Inf. Process. Lett.* 81, No. 5 (2002), pp. 277–282.

[21] Xiaodong Xu and S. P. Radziszowski. "Bounds on Shannon capacity and Ramsey numbers from product of graphs". In: *IEEE Trans. Inf. Theory* 59, No. 8 (2013), pp. 4767–4770.

[22] Xiaodong Xu et al. "Constructive lower bounds on classical multicolor Ramsey numbers". In: *Electron. J. Comb.* 11, No. 1 (2004), Research Paper R35.

Chapter 9

Lipschitz cell decomposition with a parameter in o–minimal structures

BEATA KOCEL–CYNK

Beata Kocel–Cynk
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, POLAND
e-mail: bkocel@pk.edu.pl

Abstract

In this paper we prove that for a finite family of sets definable in an o–minimal structure there exists a finite decomposition compatible with these sets and such that for any member of this decomposition the fibers (after linear change of coordinates) are Lipschitz cells.

Contents

9.1 Introduction

Parusiński [8] and Kurdyka [6] proved that every subanalytic set admits a finite decomposition into Lipschitz cells. In fact Kurdyka's proof shows the result to be true not only for subanalytic sets, but for sets definable in any o–minimal structure (see [2] and [1] for a definition and fundamental properties). Such a decomposition requires linear changes of coordinates (by [9] it is enough to use only permutations of coordinates).

The aim of the present paper is to give the following version with a parameter of the Parusiński–Kurdyka theorem.

Theorem 9.1.1. *For any subset* $A \subset \mathbb{R}^k \times \mathbb{R}^n$ *and any* $t \in \mathbb{R}^k$*, put* $A_t := \{x \in \mathbb{R}^n : (t,x) \in A\}$*. Let* $\mathcal{A}$ *be a finite set of definable subsets of* $\mathbb{R}^k \times \mathbb{R}^n$*.*

Then there exists a finite decomposition $\mathcal{B}$ *of* $\mathbb{R}^k \times \mathbb{R}^n$ *into definable subsets compatible with* $\mathcal{A}$ *(i.e. for each* $A \in \mathcal{A}, B \in \mathcal{B}$*, if* $A \cap B \neq \emptyset$ *then* $B \subset A$*) and such that, for each* $B \in \mathcal{B}$*, there exists a linear isometry* $\Psi_B : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *such that, for each* $t \in \mathbb{R}^k$*, if* $B_t \neq \emptyset$*, then* $\Psi_B(B_t)$ *is a regular* M_n*–cell in* $\mathbb{R}^n$ *with a constant* $M_n \geq 1$ *depending only on* n*.*

Similar reesult was proved in [7] however the following corollary plays an important part in our recent article [5].

Corollary 9.1.2. *Let* T *be a definable subset of* $\mathbb{R}^k$ *and let* A *be a definable subset of* $T \times \mathbb{R}^n$*.*

Then there exists a finite partition $T = \bigcup_{j=1}^s T_j$ *into definable subsets such that for each* $t \in T_j$ *we have a finite decomposition*

$$A_t = C_{1t} \cup \cdots \cup C_{p_j t},$$

such that each of the sets C_{it} *is a regular* M_n*–cell with a constant* M_N *depending only on* n *in some coordinates system depending only on* j *and* i *and moreover the sets* $\{(t,u) : t \in T_j, u \in C_{it}\}$ *are definable.*

Our proof of Theorem 9.1.1 will be a modification of the argument from [4] as well as a modification of the proof of Kurdyka [6].

9.2 Preliminaries on *o*–minimal structures

Main object of this paper are sets and functions definable in *o*–minimal structures on $(\mathbb{R}, +, \cdot)$, for the sake of the reader convenience we recall some basic definitions.

Definition 9.2.1 ([1])**.** *A* structure $\mathcal{S}$ on $\mathbb{R}$ *consists of a collection* $\mathcal{S}_n$ *of subsets of* $\mathbb{R}^n$*, for each* $n \in \mathbb{N}$*, such that*

1. $\mathcal{S}_n$ *is a boolean algebra of subsets of* $\mathbb{R}^n$*,*
2. $\mathcal{S}_n$ *contains the diagonals* $\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_i = x_j\}$ *for* $1 \leq i < j \leq n$*,*

3. *if* $A \in \mathcal{S}_n$*, then* $A \times \mathbb{R}$ *and* $\mathbb{R} \times A$ *belong to* $\mathcal{S}_{n+1}$*,*

4. *if* $A \in \mathcal{S}_{n+1}$*, then* $\pi(A) \in \mathcal{S}_n$*, where* $\pi : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$ *is the projection on the first* n *coordinates.*

We say that a set $A \subset \mathbb{R}^n$ is *definable* iff $A \in \mathcal{S}_n$. A function $f : A \longrightarrow \mathbb{R}^m$ with $A \subset \mathbb{R}^n$ is called *definable* iff its graph is definable.

Definition 9.2.2 ([1]). *A structure* $\mathcal{S}$ *on* $\mathbb{R}$ *is* o-minimal *iff*

1. $\{(x, y) : x < y\} \in \mathcal{S}_2$ *and* $\{a\} \in \mathcal{S}_1$ *for each* $a \in \mathbb{R}$*,*

2. *each set in* $\mathcal{S}_1$ *is a finite union of intervals* (a, b)*,* $-\infty \leq a < b \leq +\infty$*, and points* $\{a\}$*.*

A structure on $(\mathbb{R}, +, \cdot)$ *is a structure on* $\mathbb{R}$ *containing the graphs of both addition and multiplication.*

We end this section with some examples of polynomially bounded *o*–minimal structures on $(\mathbb{R} + \cdot)$.

Example 9.2.3. *We end this section with easiest examples of o–minimal structures*

1. *Semi-algebraic sets.*

2. *Globally subanalytic sets (*see *[3]).*

3. *Let* $K \subset \mathbb{R}$ *be a field. The structure* $\mathbb{R}^K_{an}$ *generated by:*

   - *addition and multiplication,*
   - *all analytic functions* $f : [-1, 1]^m \longrightarrow \mathbb{R}$*, for all* $m \in \mathbb{N}$*,*
   - *the power functions* $x \mapsto x^r : (0, +\infty) \longrightarrow \mathbb{R}$ *for all* $r \in K$

   (see [1] for details).

9.3 Lipschitz cells

The main technical tool used in the studies of geometry of sets definable in o–minimal structures is the cell decomposition. The notions of a cell and that of a cell decomposition are defined inductively.

Definition 9.3.1. *The cells in $\mathbb{R}^1$ are exactly points and open intervals.*

A definable set $C \subset \mathbb{R}^n$, where $n > 1$, is a cell if its image $\pi(C) \subset \mathbb{R}^{n-1}$ by the projection $\pi : \mathbb{R}^n \ni (x_1, \dots, x_{n-1}, x_n) \mapsto (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$ is a cell and C is of one of the following two types:

either

$$C = \Gamma(f) = \{(x', x_n) \in \pi(C) \times \mathbb{R} : \, x_n = f(x')\}$$

(and then C is called a graph*)*

or

$$C = (g_1, g_2) := \{(x', x_n) \in \pi(C) \times \mathbb{R} : g_1(x') < x_n < g_2(x')\}$$

(and then C is called a band*),*

where $f : \pi(C) \longrightarrow \mathbb{R}$ is a continuous definable function (resp. g_1, $g_2 : \pi(C) \longrightarrow \bar{\mathbb{R}}$ are functions such that $g_1 < g_2$ on $\pi(C)$ and, for each $i \in \{1, 2\}$, g_i is either a continuous definable function $g_i : \pi(C) \longrightarrow \mathbb{R}$ or g_i is identically equal to $-\infty$, or else g_i is identically equal to $+\infty$).

*A cell C is called a $\mathcal{C}^k$–*cell *(where $k \in \mathbb{N} \cup \{\infty, \omega\}$), if $\pi(C)$ is a $\mathcal{C}^k$–cell and f (resp. $g_i, i = 1, 2$ if finite) is a $\mathcal{C}^k$–function. Notice that every $\mathcal{C}^k$–cell is a $\mathcal{C}^k$–submanifold of $\mathbb{R}^n$.*

Definition 9.3.2. A cell decomposition of $\mathbb{R}^1$ *is a finite collection of open intervals and points of the following form:*

$$\{(-\infty, a_1), (a_1, a_2), \dots, (a_k, +\infty), \{a_1\}, \dots, \{a_k\}\},$$

where $a_1 < a_2 < \cdots < a_k$ are real numbers.

A cell decomposition *of $\mathbb{R}^n$ $(n > 1)$ is a finite partition $\mathcal{C}$ of $\mathbb{R}^n$ into cells such that the set of all projections $\{\pi(C) : C \in \mathcal{C}\}$ is a cell decomposition of $\mathbb{R}^{n-1}$, where $\pi : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-1}$ is the projection on the first $n-1$ coordinates as in Definition 9.3.1.*

Theorem 9.3.3 (Cell Decomposition Theorem, [2, Chapter 3, §2])**.** *Any o-minimal structure $\mathcal{S}$ on $(\mathbb{R}, +, .)$ admits $\mathcal{C}^1$ cell decompositions, i.e. the following holds:*

(1) If $A_1, \dots, A_k \subset \mathbb{R}^n$ are definable sets then there exists a $\mathcal{C}^1$ cell decomposition of $\mathbb{R}^n$ compatible with sets $A_1, \dots, A_k$.

(2) For each definable function $f : A \longrightarrow \mathbb{R}$, $A \subset \mathbb{R}^n$ there exists a cell decomposition of $\mathbb{R}^n$ partitioning A and such that for every $C \subset A$ in the decomposition the restriction $f|C : C \longrightarrow \mathbb{R}$ is a $\mathcal{C}^1$ function.

To circumvent some pathological behaviour of cells we distinguish their special class: Lipschitz cells.

Definition 9.3.4. *A* Lipschitz cell *in $\mathbb{R}^1$ is any cell in $\mathbb{R}^1$, and, in case $n > 1$, it is such a cell C in $\mathbb{R}^n$ that the projection $\pi(C)$ of C is a Lipschitz cell in $\mathbb{R}^{n-1}$, and the functions f or g_i, $(i = 1, 2)$ (we adopt the convention that a function equal identically $-\infty$ or $+\infty$ is Lipschitz) as in Definition 9.3.1 are Lipschitz.*

In this paper we also use an equivalent in some sense notion of a regular L–cell, which again is defined by induction.

Definition 9.3.5. *Let $L \in \mathbb{R}, L > 0$. Any open interval and any singleton in $\mathbb{R}$ is a* regular L–cell *in $\mathbb{R}$.*

A regular L–cell *in $\mathbb{R}^n$, where $n > 1$, is a $\mathcal{C}^1$ definable cell such that the projection $\pi(C)$ of C is a regular L–cell in $\mathbb{R}^{n-1}$ and the function f (resp. g_i $(i = 1, 2)$, if $g_i \not\equiv \pm\infty$) as in Definition 9.3.1 satisfies*

$$|d_{x'} f| \le L, \quad (\textit{resp. } |d_{x'} g_i| \le L, i = 1, 2), \quad \textit{for } x' \in \pi(C).$$

Since the derivative of a Lipschitz $\mathcal{C}^1$ function is bounded by the Lipschitz constant, we have the following obvious implication.

Lemma 9.3.6. *A Lipschitz $\mathcal{C}^1$–cell is a regular L–cell, for some $L > 0$.*

To see the converse we shall observe first that every regular L–cell satisfies the Whitney arc condition with exponent 1.

Definition 9.3.7. *A subset T of $\mathbb{R}^n$* satisfies the Whitney arc condition with exponent α *(where $\alpha \in \mathbb{R}$, $\alpha > 0$) if there exists a positive constant C such that any two points p and q of T can be joined in T by a rectifiable arc γ of length less then or equal to $C|p - q|^\alpha$ (cf. [10]).*

Lemma 9.3.8 ([6, Prop. 8])**.** *A regular L–cell in $\mathbb{R}^n$ satisfies the Whitney arc condition with exponent 1 and a constant C depending only on n and L.*

Corollary 9.3.9. *A regular L–cell in $\mathbb{R}^n$ is a Lipschitz $\mathcal{C}^1$ cell with a Lipschitz constant depending only on n and L.*

Example 9.3.10. *The semicircle* $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1, x_2 > 0\}$ *is a cell satisfying the Whitney property with exponent 1 (and constant* $\frac{\pi}{2}$*), but it is not a Lipschitz cell.*

We end this section recalling the definition of a stratification.

Definition 9.3.11. *Let A be a definable subset of* $\mathbb{R}^n$. Definable $\mathcal{C}^k$ stratification *of A is a finite partition of A into definable* $\mathcal{C}^k$ *submanifolds (called* strata*) satisfying the following* boundary condition*: for any two strata S, T of the partition if* $S \cap (\bar{T} \setminus T) \neq \emptyset$ *then* $S \subset \bar{T} \setminus T$.

A definable $\mathcal{C}^k$ *stratification* $\mathcal{S}$ *is called* compatible *with subsets* $B_1, \ldots, B_m \subset A$ *if each of the sets* B_ν $(\nu = 1, \ldots, m)$ *is a union of some strata.*

9.4 Proof of Theorem 9.1.1

Let $\pi : \mathbb{R}^k \times \mathbb{R}^n \longrightarrow \mathbb{R}^k$ denote the natural projection. By Cell Decomposition Theorem (Theorem 9.3.3) applied to the family $\mathcal{A}$, Theorem 9.1.1 reduces by induction to the following

Proposition 9.4.1. *Let* $\Lambda \subset \mathbb{R}^k \times \mathbb{R}^n$ *be a definable* $\mathcal{C}^1$*–cell in* $\mathbb{R}^k \times \mathbb{R}^n$ *such that, for each* $t \in \pi(\Lambda)$, Λ_t *is of dimension p. Then* Λ *has a finite decomposition*

$$\Lambda = S_1 \cup \cdots \cup S_m \cup Q$$

into definable subsets such that, for each $\nu \in \{1, \ldots, m\}$ *there exists a linear isometry* $\Psi_\nu : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *such that, for each* $t \in \pi(S_\nu)$ $\Psi_\nu(S_{\nu t})$ *is a p–dimensional* M_n*–cell in* $\mathbb{R}^n$*, where* $M_n \geq 1$ *is a constant depending only on n, and, for each* $t \in \pi(Q)$, $\dim Q_t < p$.

Definition 9.4.2. The angle between a linear subspace X and a line P *in* $\mathbb{R}^n$ *is the number*

$$\delta(P, X) = \inf\{\sin(P, S) : S \text{ line in } X\}$$

where $\sin(P, S)$ *denotes the sine of the angle between the lines P and S.*

The angle between linear subspaces X and Y *in* $\mathbb{R}^n$ *is the number*

$$\delta(Y, X) := \sup\{\delta(P, X) : P \text{ line in } Y\}.$$

If $Y = 0$ *we put* $\delta(0, X) = 0$.

Definition 9.4.3. *Let* $\epsilon > 0$ *and let* $\Gamma \subset \mathbb{R}^k \times \mathbb{R}^n$ *be a* $\mathcal{C}^1$*–cell. We will say that* Γ *is* ϵ–flat with respect to $\mathbb{R}^n$ *if, for any pair* $(t, x), (s, y) \in \Gamma$,

$$\delta(T_x\Gamma_t, T_y\Gamma_s) \leq \epsilon.$$

Lemma 9.4.4. *Let Λ be a definable $\mathcal{C}^1$–cell in $\mathbb{R}^k \times \mathbb{R}^n$ such that $\dim \Lambda_t = p$, for each $t \in \pi(\Lambda)$. Let $\epsilon > 0$.*

Then Λ has a finite decomposition

$$\Lambda = \Lambda_1 \cup \cdots \cup \Lambda_s \cup Q,$$

where Λ_ν ($\nu \in \{1, \ldots, s\}$) are $\mathcal{C}^1$–cells ϵ–flat with respect to $\mathbb{R}^n$, $\dim \Lambda_{\nu t} = p$, for each $t \in \pi(\Lambda_\nu)$, and $\nu \in \{1, \ldots, s\}$, and Q is a definable set such that $\dim Q_t < p$ for each $t \in \mathbb{R}^k$.

Proof. Consider the map $\mathcal{G} : \Lambda \ni (t, x) \mapsto T_x\Lambda_t \in \mathbb{G}(p, n)$, where $\mathbb{G}(p, n)$ is the Grassmannian of p–dimensional linear subspaces of $\mathbb{R}^n$. There exists a finite, semialgebraic covering $(U_i)_{i \in I}$ of $\mathbb{G}(p, n)$ such that if $W_1, W_2 \in U_i$, then $\delta(W_1, W_2) < \epsilon$ (cf. [6]). Let $\mathcal{T}$ be a cell decomposition of $\mathbb{R}^{k+n}$ compatible with Λ and $\mathcal{G}^{-1}(U_i)$. Let $\Lambda_1, \ldots, \Lambda_s$ be the cells in $\mathcal{T}$ contained in Λ and with fibers of dimension p. Put $Q = \Lambda \setminus (\Lambda_1 \cup \cdots \cup \Lambda_s)$. Clearly, $\Lambda_1, \ldots, \Lambda_s, Q$ satisfy the assertion of the lemma. □

Lemma 9.4.5. *Let $C \subset \mathbb{R}^k \times \mathbb{R}^n$ be a $\mathcal{C}^1$–cell such that, for each $t \in \pi(C)$, C_t is an open cell in $\mathbb{R}^n$. Let $\epsilon > 0$.*

Then C has a finite decomposition

$$C = C_1 \cup \cdots \cup C_s \cup Q$$

such that for each $i \in \{1, \ldots, s\}$, C_i is a definable $\mathcal{C}^1$–cell such that, for each $t \in \pi(C_i)$, C_{it} is open in $\mathbb{R}^n$ and for every i there exists a family $\{\Lambda_{ij}\}$ ($j = 1, \ldots, p_i$, where $p_i \leq 2n$) of $\mathcal{C}^1$–cells ϵ–flat with respect to $\mathbb{R}^n$ such that, for each $t \in \pi(C_i)$, $\partial C_{it} \subset \bigcup_{j=1}^{p_i} \Lambda_{ijt}$ and $\dim \Lambda_{ijt} = n - 1$ and $\dim Q_t < n$, for any $t \in \mathbb{R}^k$.

Proof. We shall use induction on n. For $n = 1$ the lemma is obvious.

Let $\tilde{\pi} : \mathbb{R}^k \times \mathbb{R}^{n-1} \ni (t, \tilde{x}) \mapsto \mathbb{R}^k$ be the projection. By $\mathcal{C}^1$–Cell Decomposition Theorem, Lemma 9.4.4 and the induction hypothesis, we can assume, without any loss in generality, that

$$C = \{(t, \tilde{x}, x_n) : (t, \tilde{x}) \in D, g_1(t, \tilde{x}) < x_n < g_2(t, \tilde{x})\},$$

where D is a definable $\mathcal{C}^1$–cell in $\mathbb{R}^k \times \mathbb{R}^{n-1}$ such that, for each $t \in \tilde{\pi}(D)$, D_t is an open cell in $\mathbb{R}^{n-1}$ and there exists a family $\{\Gamma_j\}$ ($j = 1, \ldots, q$, where $q \leq 2n - 2$) of $\mathcal{C}^1$–cells ϵ–flat with respect to $\mathbb{R}^{n-1}$ such that, for each $t \in \tilde{\pi}(D)$,

$$\partial D_t \subset \bigcup_{j=1}^{q} \Gamma_{jt}$$

and, moreover, the cells

$$\Lambda_i = \{(t,\tilde{x},x_n) : (t,\tilde{x}) \in D, x_n = g_i(t,\tilde{x})\} \quad (i = 1,2)$$

are ϵ–flat. We assume here that g_1 and g_2 are finite, the cases $g_1 \equiv -\infty$ and $g_2 = +\infty$ follow by a modification.

Now, it suffices to observe, that for each $t \in \pi(C) = \tilde{\pi}(D)$,

$$\partial C_t \subset \Lambda_{1t} \cup \Lambda_{2t} \cup \bigcup_{j=1}^{q} (\Gamma_j \times \mathbb{R})_t,$$

where $\Lambda_1, \Lambda_2, \Gamma_j \times \mathbb{R}$ $(j = 1,\dots,2(n-1))$ are $\mathcal{C}^1$–cells ϵ–flat with respect to $\mathbb{R}^n$. □

After this preparations we are in a position to give the proof of Proposition 9.4.1

Proof of Proposition 9.4.1. We shall proceed by induction on n. The case $n = 1$ is obvious.

Let $\Lambda \subset \mathbb{R}^n \times \mathbb{R}^k$ be a definable $\mathcal{C}^1$–cell. We shall separately consider the cases $p := \dim \Lambda_t < n$ and $p = n$.

Case I: $p < n$. In this case, by Lemma 9.4.4, we can assume that Λ is ϵ–flat for an arbitrary small $\epsilon > 0$. There exists a linear subspace $L \subset \mathbb{R}^n$ of dimension $n-p$ such that $\delta(T_x\Lambda_t, L) \geq 1-\epsilon$, for each $(x,t) \in \Lambda$.

There exists a linear isometry $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that $\Psi(L) = \{x \in \mathbb{R}^n : x_1 = \dots = x_p = 0\}$. Using $\mathcal{C}^1$–Cell Decomposition Theorem we can assume that, for each $t \in \pi(\Lambda)$, Λ_t is a graph of a $\mathcal{C}^1$–map with bounded derivatives definde on an open cell $\Gamma_t \in \mathbb{R}^p$, where Γ is the projection of Λ to the space $\mathbb{R}^k \times \mathbb{R}^p$. Applying the inductive assumption to Γ end the proof in this case.

Case II: $p = n$. In this case, applying Lemma 9.4.5, we can assume that there exists a family $\Lambda_1,\dots,\Lambda_p$ $(p \leq 2n)$ of $\mathcal{C}^1$–cells ϵ–flat with respect to $\mathbb{R}^n$ such that, for each $t \in \pi(\Lambda)$, $\partial\Lambda_t \subset \bigcup_{j=1}^{p} \Lambda_{jt}$. By [6, Lem. 3], if $\epsilon > 0$ is small enough, there exists a line L in $\mathbb{R}^n$ such that for any $(t,x) \in \Lambda_j$ and $j = 1,\dots,p$ we have $\delta(L, T_x\Lambda_{jt}) > \alpha_n$, where α_n is a positive constant depending only on n. There exists a linear isometry $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that $\Psi(L) = \{x \in \mathbb{R}^n : x_1 = \dots = x_{n-1} = 0\}$. Applying the $\mathcal{C}^1$ Cell Decomposition Theorem and the inductive hypothesis we can find decomposition $\Psi(\Lambda) = S_1 \cup \dots \cup S_m \cup Q$ into definable subsets satisfying

- $\dim Q_t < n$ for each $t \in \pi(Q)$,
- for each $\mu \in \{1,\dots,m\}$ and for each $t \in \pi(S_\mu)$, $S_{\mu t}$ is a $\mathcal{C}^1$–cell such that the projection of $S_{\mu t}$ onto $\mathbb{R}^n$ is a M_{n-1}–cell (M_{n-1} is a constant depending

only on n) and bounded from above and below by cells contained in some $\Psi(\Lambda_{jt})$.

Since every $\Psi(\Lambda_{jt})$ is a graph of a function with derivatives bounded by a constant depending only on n, we conclude the proof.

□

Acknowledgement

The author would like to thank W. Pawłucki for suggesting the problem and help during the preparation of this result.

References

[1] L. van den Dries and C. Miller. "Geometric categories and o-minimal structures". In: *Duke Math. J.* 84, No. 2 (1996), pp. 497–540.

[2] L. van den Dries. *Tame topology and o-minimal structures.* Cambridge University Press, 1998.

[3] T.l Kaya and G. Raby. "Ensembles sous-analytiques: Quelques propriétés globales". In: *C. R. Acad. Sci., Paris, Sér. I* 308, No. 18.18 (1989), pp. 521–523.

[4] B. Kocel-Cynk. "Definable stratification satisfying the Whitney property with exponent 1". In: *Ann. Pol. Math.* 92, No. 2 (2007), pp. 155–162.

[5] B. Kocel-Cynk, W. Pawłucki, and A. Valette. "A short geometric proof that Hausdorff limits are definable in any o-minimal structure". In: *Adv. Geom.* 14, No. 1 (2014), pp. 49–58.

[6] K. Kurdyka. "On a subanalytic stratification satisfying a Whitney property with exponent 1". In: *Real algebraic geometry, Proc. Conf., Rennes/Fr. 1991.* Vol. 1524. Lect. Notes Math. Springer, Berlin, 1992, pp. 316–322.

[7] K. Kurdyka and A. Parusiński. "Quasi-convex decomposition in o-minimal structures. Application to the gradient conjecture". In: *Singularity theory and its applications. Papers from the 12th MSJ International Research Institute of the Mathematical Society of Japan, Sapporo, Japan, September*

16â€“25, 2003. Vol. 43. Advanced Studies in Pure Mathematics. Math. Soc. Japan, Tokyo, 2006, pp. 137–177.

[8] A. Parusiński. “Lipschitz stratification of subanalytic sets”. In: *Ann. Sci. École Norm. Supér. (4)* 27, No. 6 (1994), pp. 661–696.

[9] W. Pawłucki. “Lipschitz cell decomposition in o-minimal structures. I”. In: *Ill. J. Math.* 52, No. 3 (2008), pp. 1045–1063.

[10] J. Stasica. “The Whitney condition for sub-analytic sets”. In: *Zesz. Nauk. Uniw. Jagiellon., Pr. Mat.* 23 (1982), pp. 211–221.

Chapter 10

Derivations on Rings

Kamil Kular

Kamil Kular
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: kamil-kular@wp.pl

Abstract

We give a review of more or less known results concerning the structure of Lie rings of derivations on abstract rings, and the commutativity of rings which admit derivations with some good properties.

Contents

10.1 Preliminaries and Introduction

Throughout the text R stands for an associative ring, possibly without identity. We denote by $\mathrm{Z}(R)$ the center of R. By an ideal of R we mean a two-sided ideal, unless we explicitly say "left ideal" or "right ideal". The ring R is said to be *prime*, if the product of any two nonzero ideals of R is also a nonzero ideal (in other words, for arbitrary $a, b \in R$, if $aRb = \{0\}$, then $a = 0$ or $b = 0$).

The ring R is called *semiprime*, if it does not contain any nonzero nilpotent ideal (equivalently, $\forall\, a \in R:\ aRa = \{0\} \Rightarrow a = 0$). The characteristic of R is defined as the smallest positive integer n with the property that $na = 0$ for all $a \in R$. (If there is no positive integer n such that $na = 0$ for all $a \in R$, then $\mathrm{char}(R) := 0$). The ring R is 2-*torsion free*, if $2a = 0$ implies $a = 0$, for any $a \in R$. Notice that $\mathrm{char}(R) \neq 2$ whenever R is 2-torsion free; the converse of this implication is false.

For arbitrary $a, b \in R$ we define the *Lie product* $[a, b] = ab - ba$ and the *Jordan product* $a \circ b = ab + ba$. An additive subgroup $U \subseteq R$ is said to be a *Lie ideal* (resp. *Jordan ideal*) of R, if $[u, a] \in U$ (resp. $u \circ a \in U$) for any $u \in U$ and any $a \in R$.

A map $f : R \longrightarrow R$ is called *centralizing* on a set $S \subseteq R$, if $[f(x), x] \in \mathrm{Z}(R)$ for all $x \in S$. If $[f(x), x] = 0$ for all $x \in S$, then f is said to be *commuting* on S.

A map $d : R \longrightarrow R$ is said to be a *derivation*, if it is additive and satisfies the Leibniz rule

$$\forall\, x, y \in R:\ d(xy) = d(x)y + xd(y).$$

The usual derivation on algebras of differentiable functions is obviously a derivation in the above sense. It is easy to see that for a fixed $a \in R$, the map $\partial_a : R \ni x \mapsto [a, x] \in R$ is also a derivation. This derivation is referred to as the *inner derivation* generated by a.

The set $\mathrm{Der}(R)$ of all derivations $d : R \longrightarrow R$ is a *Lie ring* under pointwise addition and the Lie multiplication defined by $[d, \delta] = d\delta - \delta d$, where $d\delta$ and δd stand for compositions of the derivations. The primeness and semiprimeness of $\mathrm{Der}(R)$ are defined by means of Lie ideals. It is not difficult to prove that $\mathrm{IDer}(R) = \{\partial_a :\ a \in R\}$ is a Lie ideal of $\mathrm{Der}(R)$.

Refer to [14] for further information concerning noncommutative algebra.

Derivations on abstract rings are very interesting and challenging object of study in algebra. The number of results on the derivations is immense and still growing. In this article we give an overview of theorems and examples concerning derivations on prime and semiprime rings. Our focus is on relationships between the derivations and commutativity.

10.2 Structure of Lie Rings of Derivations

The study of derivations on abstract rings, though initiated long ago, was given its impetus by Posner's paper [20]. In this paper the following result concerning derivations on a prime ring has been proved.

Theorem 10.2.1. *Let R be a prime ring of characteristic different from 2 and let $d, \delta \in \mathrm{Der}(R)$ be such that the composition $d\delta$ is a derivation. Then $d = 0$ or $\delta = 0$.*

The Posner theorem is a valuable tool for studying the structure of Lie rings of derivations. Several authors investigated and generalized this theorem in many ways (see for example [8, 9, 18]).

In 1978 C. R. Jordan and D. A. Jordan published two theorems about primeness and semiprimeness of Lie rings of derivations (see [16]).

Theorem 10.2.2. *Let R be a prime ring of characteristic different from 2. Then* $\mathrm{Der}(R)$ *is a prime Lie ring.*

Theorem 10.2.3. *Let R be a semiprime 2-torsion free ring. Then* $\mathrm{Der}(R)$ *is a semiprime Lie ring.*

For any integer $s \geqslant 3$ and derivations $d_1, \dots, d_s \in \mathrm{Der}(R)$, we inductively define

$$[d_1, \dots, d_s] = [[d_1, \dots, d_{s-1}], d_s].$$

A Lie ring D of derivations on R (i.e., D is a Lie subring of $\mathrm{Der}(R)$) is said to be *nilpotent*, if there exists a positive integer n such that $[d_1, \dots, d_{n+1}] = 0$ for all $d_1, \dots, d_{n+1} \in D$.

Let us present the main results of [19] and [3].

Theorem 10.2.4 ([19]). *Suppose that R is a semiprime ring. Then the following conditions are equivalent:*

(1) the Lie ring $\mathrm{IDer}(R)$ *is nilpotent,*

(2) R is commutative.

Theorem 10.2.5 ([3]). *If R is a semiprime ring, then either* $\mathrm{Der}(R) = \{0\}$ *or* $\mathrm{Der}(R)$ *is not nilpotent.*

Notice that the semiprimeness assumption in the above theorems cannot be removed.

Example 10.2.6. *The ring*

$$R = \left\{ \begin{pmatrix} 0 & x & y \\ 0 & 0 & z \\ 0 & 0 & 0 \end{pmatrix} : x, y, z \in \mathbb{K} \right\}$$

of all strictly upper triangular 3×3 matrices over a field $\mathbb{K}$ is not commutative, although $[\partial_A, \partial_B] = 0$ for any $A, B \in R$.

Example 10.2.7 ([3]). *The quotient ring $\mathbb{Q}[X]/\langle X^2 \rangle$ is isomorphic to $\mathbb{Q}[a] = \{p + qa : p, q \in \mathbb{Q}\}$, where a is such that $a^2 = 0$. If $\lambda \in \mathbb{Q}$, then the map $d_\lambda : \mathbb{Q}[a] \longrightarrow \mathbb{Q}[a]$ defined by $d_\lambda(p + qa) = \lambda qa$ is a derivation. It is not difficult to prove that $\{d_\lambda : \lambda \in \mathbb{Q}\} = \mathrm{Der}(\mathbb{Q}[a])$. Since $[d_\lambda, d_\mu] = 0$ for all $\lambda, \mu \in \mathbb{Q}$, the Lie ring* $\mathrm{Der}(\mathbb{Q}[a])$ *is therefore nonzero and nilpotent.*

In the next section, relationships between derivations and commutativity of semiprime rings will be discussed in detail.

Now, suppose that R is a commutative (and associative) algebra with identity over a field $\mathbb{K}$. We will denote by $\mathrm{Der}_{\mathbb{K}}(R)$ the Lie algebra of all $\mathbb{K}$-derivations on R. (Recall that a derivation $d : R \longrightarrow R$ is said to be a $\mathbb{K}$-derivation, if it is a $\mathbb{K}$-linear map). This Lie algebra has also the R-module structure defined by $(a\delta)(x) = a\delta(x)$, where $a, x \in R$ and $\delta \in \mathrm{Der}_{\mathbb{K}}(R)$.

Definition 10.2.8. *Let L be a Lie subalgebra of $\mathrm{Der}_{\mathbb{K}}(R)$. An ideal I of the ring R is said to be an L-ideal, if $\delta(I) \subseteq I$ for all $\delta \in L$. Moreover, R is called L-simple, if R and $\{0\}$ are the only L-ideals of R.*

In the following theorems due to D. A. Jordan (see [17]), D stands for a nonzero Lie subalgebra of $\mathrm{Der}_{\mathbb{K}}(R)$ being also an R-submodule of $\mathrm{Der}_{\mathbb{K}}(R)$.

Theorem 10.2.9. *If R is D-simple, then D is a simple Lie algebra except possible when $\mathrm{char}(\mathbb{K}) = 2$ and D is cyclic as an R-module.*

Theorem 10.2.10. *Assume that $\mathrm{char}(\mathbb{K}) = 2$ and $D = R\delta$ is cyclic as an R-module. Then D is a simple Lie algebra if and only if $\delta(R) = R$.*

10.3 Commutativity of Prime and Semiprime Rings with Derivations

A ring which admits a derivation with sufficiently good properties must be commutative. This phenomenon was first investigated by Posner.

Theorem 10.3.1 (see [20]). *Let R be a prime ring and $d : R \longrightarrow R$ a derivation. If $[d(a), a] \in \mathrm{Z}(R)$ for all $a \in R$ (i.e., d is centralizing on R), then $d = 0$ or R is commutative.*

It is not difficult to give an example showing that the primeness is a necessary assumption in the above theorem.

Example 10.3.2. *Consider the ring $R = R_1 \times R_2$, where R_1 is a nonzero commutative ring with a nonzero derivation d_1 and R_2 is a noncommutative ring. Then the formula $d(x_1, x_2) = (d_1(x_1), 0)$ defines a nonzero derivation $d : R \longrightarrow R$. This derivation is commuting on R. However, R is not commutative.*

Many authors extended the Posner result by assuming that the derivation is only centralizing on an appropriate subset of the ring. The following two theorems come from [5].

Theorem 10.3.3. *Let R be a prime ring of characteristic different from 2 and 3. If a nonzero derivation $d : R \longrightarrow R$ is centralizing on an ideal U of the ring R, then $U \subseteq \mathrm{Z}(R)$ (in other words, U is a central ideal).*

Theorem 10.3.4. *Let R be a prime ring of characteristic 2. Suppose that U is simultaneously a subring of R and a Lie or Jordan ideal of R. If a nonzero derivation $d : R \longrightarrow R$ is centralizing on U, then U is commutative.*

If we just assumed that U is a subring OR a Lie (Jordan) ideal, then Theorem 10.3.4 would not be true. Consider examples due to Awtar.

Example 10.3.5. *The ring R of all 2×2 matrices over the field $\mathbb{Z}_2$ is prime. Let*

$$U = \left\{ \begin{pmatrix} x & y \\ z & x \end{pmatrix} : x, y, z \in \mathbb{Z}_2 \right\}.$$

It is easy to see that U is a Lie ideal, but not a subring of R. Define the map $d : R \longrightarrow R$ by

$$d \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} w - z & x - w \\ x - w & y - z \end{pmatrix}.$$

Then d is a nonzero derivation centralizing on U. However, U is not commutative.

Example 10.3.6. *Let R be the ring of all 2×2 matrices over a noncommutative prime ring with identity. Then R is prime. The set U of all diagonal matrices in R is obviously a noncommutative subring, but neither a Lie ideal nor a Jordan ideal of the ring R. If $d : R \longrightarrow R$ is defined by*

$$d \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 0 & -y \\ z & 0 \end{pmatrix},$$

then d is a nonzero derivation commuting on U.

Further extensions of Theorem 10.3.1 have been provided by Bell and Martindale in [7].

Theorem 10.3.7. *Let R be a semiprime ring and $U \subseteq R$ a nonzero left ideal. If R admits a derivation which is nonzero on U and centralizing on U, then R contains a nonzero central ideal.*

Theorem 10.3.8. *Let R be a prime ring and $U \subseteq R$ a nonzero left ideal. If R admits a nonzero derivation which is centralizing on U, then R is commutative.*

Now, we will present a few results of a slightly different nature. The first one has been proved by Herstein.

Theorem 10.3.9 (see [13]). *Let R be a prime ring of characteristic not 2 admitting a nonzero derivation d such that $[d(x), d(y)] = 0$ for all $x, y \in R$. Then R is commutative.*

The next two results are due to Daif and Bell (see [11]).

Theorem 10.3.10. *Let R be a prime ring and let $d : R \longrightarrow R$ be a derivation. If there is a nonzero ideal $U \subseteq R$ such that either $d([x, y]) + [x, y] = 0$ for all $x, y \in U$ or $d([x, y]) = [x, y]$ for all $x, y \in U$, then R is commutative.*

Theorem 10.3.11. *Let R be a semiprime ring admitting a derivation d such that either $d([x, y]) + [x, y] = 0$ for all $x, y \in R$ or $d([x, y]) = [x, y]$ for all $x, y \in R$. Then R is commutative.*

In the sequel we will need technical definitions, introduced also by Bell and Daif.

Definition 10.3.12. *Let U be an ideal (left, right or two-sided) of R. A derivation $d : R \longrightarrow R$ is said to be*

- *a U^*-derivation, if $[d(x), d(y)] + d([x, y]) = 0$ for all $x, y \in U$,*
- *a U^{**}-derivation, if $[d(x), d(y)] = d([x, y])$ for all $x, y \in U$,*
- *a U^{***}-derivation, if $d([x, y]) = 0$ for all $x, y \in U$.*

The purpose of introducing the U^*, U^{**} and U^{***}-derivations was to formulate the following theorems.

Theorem 10.3.13 (see [6]). *Let R be a prime ring and $U \subseteq R$ a nonzero right ideal. If R admits a nonzero U^* or U^{**}-derivation d, then either R is commutative or $d(U)d(U) = \{0\}$.*

Theorem 10.3.14 (see [10]). *Let R be a semiprime ring and U a nonzero ideal of R. If R admits a U^*, U^{**} or U^{***}-derivation which is nonzero on U, then R contains a nonzero central ideal.*

In [10] Daif also proved

Theorem 10.3.15. *Let R be a 2-torsion free semiprime ring and let $U \subseteq R$ be a nonzero ideal. If R admits a derivation d which is nonzero on U and such that $[d(x), d(y)] = 0$ for all $x, y \in U$, then R contains a nonzero central ideal.*

The above theorem would not be true, if we assumed that U is only a one-sided ideal.

Example 10.3.16. *The ring R of all 2×2 matrices over a field $\mathbb{K}$ is prime and does not have nonzero central ideals. Nevertheless, if*

$$a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$U \in \{aR, Ra\}$ and $d : R \longrightarrow R$ is the inner derivation generated by a, then d is nonzero on U and $[d(x), d(y)] = 0$ for all $x, y \in U$.

A complete characterization of central ideals of an arbitrary 2-torsion free semiprime ring in terms of derivations was given by Hongan.

Theorem 10.3.17 (see [15])**.** *Let R be a 2-torsion free semiprime ring and let U be a nonzero ideal of R. Then the following conditions are equivalent:*

(1) R admits a derivation d such that $d([x, y]) - [x, y] \in \mathrm{Z}(R)$ for all $x, y \in U$,

(2) R admits a derivation d such that $d([x, y]) + [x, y] \in \mathrm{Z}(R)$ for all $x, y \in U$,

(3) R admits a derivation d such that for all $x, y \in U$ we have $d([x, y]) - [x, y] \in \mathrm{Z}(R)$ or $d([x, y]) + [x, y] \in \mathrm{Z}(R)$,

(4) $U \subseteq \mathrm{Z}(R)$.

This characterization immediately yields

Corollary 10.3.18. *If a 2-torsion free semiprime ring R admits a derivation d such that for all $x, y \in R$ we have $d([x, y]) - [x, y] \in \mathrm{Z}(R)$ or $d([x, y]) + [x, y] \in \mathrm{Z}(R)$, then R is commutative.*

The assumption "2-torsion free" cannot be removed from Corollary 10.3.18.

Example 10.3.19. *Let R be the ring of all 2×2 matrices over $\mathbb{Z}_2$. Then R is noncommutative, prime and $\mathrm{char}(R) = 2$. On the other hand, it is not difficult to see that the inner derivation $\partial_a : R \longrightarrow R$ generated by*

$$a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

satisfies $\partial_a([x, y]) + [x, y] \in \mathrm{Z}(R)$ for all $x, y \in R$.

In [4] Atteya provided sufficient conditions for the existence of nonzero central ideals, involving Jordan multiplication $\circ$.

Theorem 10.3.20. *Let R be a 2-torsion free semiprime ring and let U be a nonzero ideal of R. Suppose that one of the conditions* *(i)* $[x, y] + x \circ y \in \mathrm{Z}(R)$ *for all* $x, y \in U$*,*

(ii) $[x, y] - x \circ y \in \mathrm{Z}(R)$ *for all* $x, y \in U$*,*

(iii) *R admits a derivation d which is nonzero on U and such that* $d([x, y]) = d(x \circ y)$ *for all* $x, y \in U$

is satisfied. Then R contains a nonzero central ideal.

A similar, although more advanced, theorem can be found in [1].

Interesting and elegant commutativity criteria were given by Gupta, Argaç and Inceboz.

Theorem 10.3.21 (see [12])**.** *Let R be a semiprime ring with identity. If for any* $x, y \in R$ *there exists a positive integer n (depending on x and y) such that* $(xy)^k - x^k y^k \in \mathrm{Z}(R)$*, where* $k = n, n+1, n+2$*, then R is commutative.*

In their three criteria, Argaç and Inceboz made use of derivations (see [2]).

Theorem 10.3.22. *Let R be a prime ring, let U be a nonzero ideal of R, and let n be a positive integer. If R admits a derivation d such that*

$$(d(x)y + xd(y) + d(y)x + yd(x))^n = xy + yx$$

for all $x, y \in U$*, then R is commutative.*

Theorem 10.3.23. *Let R be a prime ring with* $\mathrm{char}(R) \neq 2$*, let* $U \subseteq R$ *be a nonzero ideal, and let n be a positive integer. If R admits a derivation d such that*

$$(d(x)y + xd(y) + d(y)x + yd(x))^n - (xy + yx) \in \mathrm{Z}(R)$$

for all $x, y \in U$*, then R is commutative.*

Corollary 10.3.24. *Let R be a prime ring,* $U \subseteq R$ *a nonzero ideal, and* $d : R \longrightarrow R$ *a derivation. If one of the conditions*

(i) $d(x)x + xd(x) = x^2$ *for any* $x \in U$*,*

(ii) $\mathrm{char}(R) \neq 2$ *and* $d(x)x + xd(x) - x^2 \in \mathrm{Z}(R)$ *for any* $x \in U$

is satisfied, then R is commutative.

It seems noteworthy that the proof of the above corollary consists in substituting $x + y$ for x (linearization), and applying Theorems 10.3.22 and 10.3.23 with $n = 1$.

The following example taken from [2] shows that the primeness assumption in Theorems 10.3.22 and 10.3.23 cannot be omitted.

Example 10.3.25. *Let $\mathbb{K}$ be an arbitrary field. Define*

$$R = \left\{ \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix} : x, y \in \mathbb{K} \right\}$$

and notice that R is a ring under the usual matrix operations. The set

$$U = \left\{ \begin{pmatrix} 0 & y \\ 0 & 0 \end{pmatrix} : y \in \mathbb{K} \right\}$$

is an ideal of R. Consider the map $d : R \longrightarrow R$ defined by

$$d \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}.$$

It is easy to see that d is a derivation and

$$(d(A)B + Ad(B) + d(B)A + Bd(A))^n = AB + BA,$$

for all $A, B \in U$ and all positive integers n. However, the ring R is not commutative.

Analogous results hold for semiprime rings.

Theorem 10.3.26 (see [2]). *Let R be a semiprime ring and let n be a positive integer. If R admits a derivation d such that*

$$(d(x)y + xd(y) + d(y)x + yd(x))^n = xy + yx$$

for all $x, y \in R$, then R is commutative.

Theorem 10.3.27 (see [2]). *Let R be a 2-torsion free semiprime ring and let n be a positive integer. If R admits a derivation d such that*

$$(d(x)y + xd(y) + d(y)x + yd(x))^n - (xy + yx) \in \mathrm{Z}(R)$$

for all $x, y \in R$, then R is commutative.

Corollary 10.3.28. *Let R be a semiprime ring and let $d : R \longrightarrow R$ be a nonzero derivation. If one of the conditions*

(i) $d(x)x + xd(x) = x^2$ *for any $x \in R$,*

(ii) R *is 2-torsion free and* $d(x)x + xd(x) - x^2 \in \mathrm{Z}(R)$ *for any $x \in R$*

is satisfied, then R is commutative.

References

[1] S. Ali and S. Huang. "On derivations in semiprime rings". In: *Algebr. Represent. Theory* 15, No. 6 (2012), pp. 1023–1033.

[2] N. Argaç and H. G. Inceboz. "Derivations of prime and semiprime rings". In: *J. Korean Math. Soc.* 46, No. 5 (2009), pp. 997–1005.

[3] O. D. Artemovych and K. Kular. "On the Lie ring of derivations of a semiprime ring". In: *Algebra Discrete Math. (to appear)* ().

[4] M. J. Atteya. "On commutativity of prime and semiprime rings". In: *J. J. Appl. Sci.* 9, No. 2 (2007), pp. 161–169.

[5] R. Awtar. "Lie and Jordan structure in prime rings with derivations". In: *Proc. Am. Math. Soc.* 41 (1973), pp. 67–74.

[6] H. E. Bell and M. N. Daif. "On derivations and commutativity in prime rings". In: *Acta Math. Hung.* 66, No. 4 (1995), pp. 337–343.

[7] H. E. Bell and W. S. III Martindale. "Centralizing mappings of semiprime rings". In: *Can. Math. Bull.* 30, No. 1-3 (1987), pp. 92–101.

[8] J. Bergen. "Lie ideals with regular and nilpotent elements and a result on derivations". In: *Rend. Circ. Mat. Palermo, II Ser.* 33 (1984), pp. 99–108.

[9] M. A. Chebotar. "The composition of derivations of prime rings". In: *Mosc. Univ. Math. Bull.* 50, No. 2 (1995), pp. 20–22.

[10] M. N. Daif. "Commutativity results for semiprime rings with derivations". In: *Int. J. Math. Math. Sci.* 21, No. 3 (1998), pp. 471–474.

[11] M. N. Daif and H. E. Bell. "Remarks on derivations on semiprime rings". In: *Int. J. Math. Math. Sci.* 15, No. 1 (1992), pp. 205–206.

[12] V. Gupta. "A commutativity theorem for semiprime rings". In: *J. Aust. Math. Soc., Ser. A* 30 (1980), pp. 33–36.

[13] I. N. Herstein. "A note on derivations". In: *Can. Math. Bull.* 21 (1978), pp. 369–370.

[14] I. N. Herstein. *Noncommutative rings.* Mathematical Association of America, Washington, DC, 1996.

[15] M. Hongan. "A note on semiprime rings with derivation". In: *Int. J. Math. Math. Sci.* 20, No. 2 (1997), pp. 413–415.

[16] C. R. Jordan and D. A. Jordan. "Lie rings of derivations of associative rings". In: *J. Lond. Math. Soc., II. Ser.* 17 (1978), pp. 33–41.

[17] D. A. Jordan. "On the simplicity of Lie algebras of derivations of commutative algebras". In: *J. Algebra* 228, No. 2 (2000), pp. 580–585.

[18] J. Krempa and J. Matczuk. “On the composition of derivations”. In: *Rend. Circ. Mat. Palermo, II. Ser.* 33 (1984), pp. 441–455.

[19] K. Kular. “Semiprime rings with nilpotent Lie ring of inner derivations”. In: *Ann. Univ. Paedagog. Crac., Stud. Math. (to appear)* (2014).

[20] E. C. Posner. “Derivations in prime rings”. In: *Proc. Am. Math. Soc.* 8 (1958), pp. 1093–1100.

Chapter 11

Simple Chooser Options with Some Risk Reducing Derivatives

Anna Milian

Anna Milian
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: amilian@pk.edu.pl

Abstract

In this paper we discuss two portfolios: one consisting of simple chooser options and second- additionally containing some proposed risk reducing derivatives. We use the Black-Scholes model to describe the price of underlying asset and we derive an explicit formula of the non-arbitrage price of the proposed derivative. We use Monte-Carlo methods to analyze the impact of the proposed instrument on the risk of investing in simple chooser options. We show that the proposed derivative instrument can be used to reduce the risk of large loss from investing in simple chooser options. The presented simulation is performed using MAPLE.

Contents

11.1 Introduction

An option is a contract between a holder of an option and a seller that gives the holder the right - but not the obligation - to buy or to sell the underlying asset at an agreed price at a later date. The agreed price in the contract is called the strike price(strike, exercise price), the date - the expiration date(maturity). In this paper we will have three kinds of options: calls, puts and simple choosers. A call stock option gives the holder of the option the right to buy specified quantity of the stock at the strike price on or before the expiration date. The seller of the call option has the obligation to sell the stock at the agreed price, if the holder of the call option decides to exercise his right to buy. A put option gives the holder the right to sell specified quantity of the underlying asset at the strike price on or before the expiration date. The seller of the put option has the obligation to buy the stock at the strike price if the holder decides to exercise his right to sell. This paper deals with European options which may be exercised only at the expiration date and is devoted to the study of chooser options. They are sometimes named "you-choose" or "as-you-like" options because they are purchased in the present, but are chosen to be either put or call at some specific point in time $t(t < T)$. Chooser options are suitable when strong volatility of the underlying asset is expected but investors are not certain about direction of the change [3]. When the underlying asset decreases, over a period of time, the holder of the option will choose the put option because it will have a higher value than the call option. Otherwise, in the case of rising value of the underlying asset the choice will be the call option. Once this choice was made at time the option stays as either a call or a put to maturity. If the strike prices of the call and the put are the same, just as their expirations, such option is referred to as a simple chooser. Chooser options have been traded since July 1990 with the initial contracts traded by Bankers Trust [6]. In [6] the relationships between choice date and chooser price, and between chooser price and its strike price were examined. Monte-Carlo methods to price simple chooser options were used in [8]. Moreover dependence of the distribution of rate of return from investment in simple chooser options on the strike price was investigated in [8]. To protect a holder of simple chooser options against large loss, in the presented paper we propose a portfolio composed of simple chooser option, a certain amount invested in underlying asset (stocks) and some derivative financial instrument, paying an agreed amount of money when the value of the underlying asset falls below an agreed level. The writer of this instrument could be a large financial institution with high liquidity. An investor investing in risky simple chooser options and risky stocks would be the buyer of this instrument. In this paper, we obtain an analytical closed form formula to price the proposed derivative instrument. Using Monte Carlo simulations we show that the proposed investment portfolio is less risky than the portfolio composed exclusively of simple chooser options. The idea of the financial instrument is based on the idea of catastrophe bonds [10]. We use the Black-Scholes model to describe the price of an underlying asset. We assume that there are no riskless arbitrage opportunities, there are no transaction costs, there are no dividends during the life of an option, security trading is continuous, the risk-free rate of interest and the stock price volatility are constant and the price of an underlying asset follows a geometric Brownian process. Using crude Monte Carlo, we examine distributions of rate of return from considered portfolios. The simulations are performed using MAPLE. In this chapter we present also results obtained in [7].

11.2 Model description

The Black-Scholes model is based on the assumptions that markets are arbitrage free and the price of underlying asset follows a geometric Brownian motion

$$S_t = S_0 exp((r - \frac{1}{2}\sigma^2)t + \sigma W_t), t \in [0, T], \tag{11.1}$$

where $W = W_t, t \in [0, T]$ is a standard Brownian motion under the risk-neutral probability P, r is the risk-free interest rate, $S_0$is the stock price at time 0, T is the expiry date and $\sigma > 0$ is the stock price volatility [2]. Under the assumption of no arbitrage, the price of a generic derivative security can be expressed as the expected value of its discounted payouts, taken with respect to the risk-neutral measure. Then today's price of a derivative, that pays at some time t according to a F_t- measurable payoff function $H(t)$, is

$$E(e^{-rt}H(t)) \tag{11.2}$$

where E denotes the expectation operator under P- measure [11]. Independent replications $H^{(i)}(T), i = 1, \ldots, n$ give us not only the estimation of the price of the derivative [4]

$$c \approx exp(-rT)\frac{1}{n}\sum_{i=10}^{n} H^{(i)}(T) \tag{11.3}$$

at $t = 0$ but also the sample of rate of return R, expressed in percentage:

$$R^{(i)} = \frac{exp(-rT)H^{(i)}(T) - c}{c} 100\%, \quad i = 1, ..., n$$

Let S_t denote the price of the underlying asset at t, T- the expiry date, $T-t$ time to maturity, and let K denote the strike price. For a call option, the pay-off is given by

$$H(T) = (S_T - K)^+$$

and by

$$H(T) = (K - S_T)^+$$

for a put option. Let $T-t$ denote time to maturity and

$$C(S_t, K, T-t) \text{ - premium of European call option,}$$

$$P(S_t, K, T-t) \text{ - premium of European put option.}$$

Hence $C(S_0, K, T) = (S_T - K)^+$ and $P(S_0, K, T) = (K - S_T)^+$.
Let us consider now the situation of a holder of simple chooser option. At time t he will choose the call option if

$$C(S_t, K, T-t) > P(S_t, K, T-t),$$

otherwise he will choose the put option [6].
Using the put-call parity

$$C(S_t, K, T-t) - P(S_t, K, T-t) = S_t - Kexp[-r(T-t)]$$

it is easy to see that above inequality is equivalent to

$$S_t > Kexp(-r(T-t)).$$

Hence the value of chooser option at time t equals

$$ch(t) = max(C(S_t, K, T-t), P(S_t, K, T-t)) = \tag{11.4}$$

$$C(S_t, K, T-t) + max(Kexp(-r(T-t)) - S_t, 0). \tag{11.5}$$

The value of the option at time 0, when the choosing time is t, is equal to

$$CH(t) = exp(-rt)E[C(S_t, K, T-t) + max(Kexp(-r(T-t)) - S_t, 0)]. \tag{11.6}$$

In next section we will define a financial derivative instrument reducing probabilities of large loss in the case of investing in simple chooser options. The payoff function of the derivative is similar to the payoff from catastrophe bonds. Catastrophe bonds are designed to cover mostly costs of natural disasters as hurricanes, windstorms or earthquakes. If X denotes random moment of occurrence of a catastrophe (triggering point), the payoff of the catastrophe bond is

$$H(T) = 1_{(T<X)} S_T.$$

The returns from catastrophe bonds are incorrelated with macroeconomics factors and can be used to diversify investment portfolios [10].

11.3 Two portfolios, pricing

Applying (11.4) we can price the simple option by simulation [5]. For this purpose we express a payoff $H_1(T)$ of the option at T in the following way

$$H_1(T) = 1_A max(S_T - K, 0) + 1_B max(K - S_T, 0) \qquad (11.7)$$

where

$$A = \{S_t > Kexp[-r(T-t)]\}, \quad B = \{S_t < Kexp[-r(T-t)]\}.$$

Let U and V be independent, normal distributed random variables with mean 0 and variance t and $T-t$ respectively,

$$X = S_0 exp[(r - \frac{1}{2}\sigma^2)t + \sigma U]$$

and

$$Y = exp[(r - \frac{1}{2}\sigma^2)(T-t) + \sigma V].$$

By(1) we have $S_T = XY$. We generate independent samples of X and Y. By definition of simple chooser option, if $X > Kexp[-r(T-t)]$ then $H_1(T) = max(S_T - K, 0)$, else $H_1(T) = max(K - S_T, 0)$. In this way , we obtain sample

$$H_1^{(i)}(T), i = 1, \ldots, n,$$

of random variable $H_1(T)$. By (3) we obtain the estimation CH of the simple chooser option price at time 0 and the sample of rate of return from investing in the options, expressed in percentage:

$$R^{(i)}(T) = \frac{exp(-rT)H_1^{(i)}(T) - CH}{CH} 100\%, \quad i = 1, \ldots, n.$$

Figure 11.1 presents a histogram of rate of return , based on $n = 10^4$ simulations, where a maturity T is one year, the underlying asset price S_0 is 50, $r = 10\%, \sigma = 30\%, t = 0.6, K = 50$. Mean is equal 0 and median equals -22.66.

In Table 11.1 we present the probabilities corresponding to different value ranges of return R, obtained by simulation [8].

From Table 11.1 it follows that in the case of simple chooser options, a large loss and a large gain can have high probabilities. The level of risk of investing in these options is high. To reduce the risk we propose a portfolio composed of simple chooser options, an amount invested in underlying asset and some derivative financial instrument, paying an agreed amount of moneywhen the value of the underlying asset falls below an agreed level.

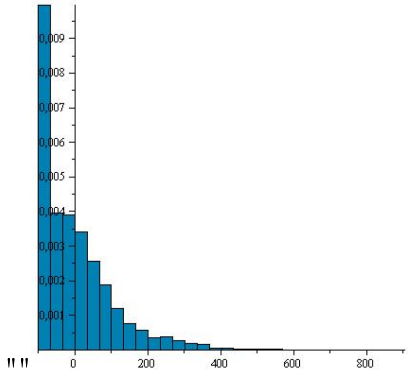


Figure 11.1: Histogram of rates of return for strike price K = 50

Table 11.1: Approximated distributions of rates of return for strike price K = 50

$P(R \leq -75\%)$	0.3
$P(-75\% < R \leq -50\%)$	0.09
$P(-50\% < R \leq -25\%)$	0.1
$P(-25\% < R \leq 0\%)$	0.1
$P(0\% < R \leq 25\%)$	0.09
$P(25\% < R \leq 50\%)$	0.07
$P(50\% < R \leq 75\%)$	0.06
$P(R > 75\%)$	0.19

Precisely, we buy:

1) one simple chooser option on underlying asset with value S_0 at 0 and payoff function given by (11.7)

2) a derivative with the following payoff function

$$H_2(T) = 1_A S_T \tag{11.8}$$

where $A = \{S_T \leq K\}$,

3) an underlying asset with value S_0 at 0, where

$$S_t = S_0 exp[(r - \frac{1}{2}\sigma^2)t + \sigma W_t], \quad t \in [0, T].$$

Thus we obtain the following payoff from the portfolio

$$H(T) = H_1(T) + H_2(T) + S_T \tag{11.9}$$

The payoff $H_2(T)$ is similar to a payoff of catastrophe bonds [9]. Derivative instrument proposed in this paper, with payoff $H_2(T)$ can be considered as an obligation transferring the risk from an individual investor investing in simple chooser options and an underlying instrument (a holder of the derivative) to a writer- a capital market company with high liquidity. Since $H_2(T)$ is positive, $\mathcal{F}_t-$ measurable and square-integrable with respect to measure P, we obtain the today's arbitrage price of the proposed instrument as the expected value of its discounted payouts, according to (11.2):

$$c = E(\exp(-rT)H_2(T)) \tag{11.10}$$

To calculate above expected value let us first note that the random variable

$$X = (r - \frac{1}{2}\sigma^2)T + \sigma W_t$$

has probability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi T}} \exp(-\frac{[x-(r-\frac{1}{2}\sigma^2)T]^2}{2\sigma^2 T}).$$

Since $S_T = S_0 \exp(X)$, it follows that the probability density function g of S_T is expressed as

$$g(x) = \frac{1}{x} f(\ln \frac{x}{S_0}) \text{ for } x > 0 \text{ and } g(x) = 0 \text{ for } x \leq 0.$$

Hence [1]

$$\begin{aligned} c &= \exp(-rT)E(1_{(S_T \leq K)} S_T) \\ &= \exp(-rT)\int_\Omega 1_{(S_T \leq K)} S_T dP = \exp(-rT)\int_0^K xg(x)dx. \end{aligned} \tag{11.11}$$

With the change of variables, we have

$$c = \exp(-rT)\int_0^K f(\ln(\frac{x}{S_0}))dx.$$

Substituting $\ln(\frac{x}{S_0}) = u$ in last integral we have

$$\begin{aligned} c &= S_0 \exp(-rT) \textstyle\int_{-\infty}^{\ln(\frac{K}{S_0})} f(u)\exp(u)du \\ &= \frac{S_0 \exp(-rT)}{\sigma\sqrt{2\pi T}} \int_{-\infty}^{\ln(\frac{K}{S_0})} \exp\{u - \frac{[u-(r-\frac{1}{2}\sigma^2)T]^2}{2\sigma^2 T}\}du. \end{aligned} \tag{11.12}$$

An easy computation shows that

$$u - \frac{[u-(r-\frac{1}{2}\sigma^2)T]^2}{2\sigma^2 T} = \frac{-[u-(r+\frac{1}{2}\sigma^2)T]^2 + 2r\sigma^2 T^2}{2\sigma^2 T}$$

and consequently we obtain the price of considered derivative

$$c = S_0 N(\frac{\ln(\frac{K}{S_0} - (r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}) \tag{11.13}$$

where N is the cumulative probability distribution function for a standarized normal distribution. Moreover,

$$E(\exp(-rT)S_T) = S_0 \tag{11.14}$$

because the process $\exp(-rt)S_t, \quad t \in [0,T]$ is a martingale. Finally, by (11.6), (11.13) and (11.14) we obtain the today's price of the proposed portfolio

$$C(T) = CH(T) + S_0 N(\frac{\ln(\frac{K}{S_0} - (r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}) + S_0. \tag{11.15}$$

11.4 Rates of return, an example

Using Monte Carlo method we can analyse the profit function, expressed by rate of return from a portfolio, at the expiry date.To obtain this goal we simulate values of a payoff function.We present results of the simulations in Tables 1.2-1.4 below.Symbol I in column headers concerns portfolio containing only simple chooser option with payoff function $H_1(T)$ given by (11.7), symbol II denotes portfolio composed of simple chooser option, a proposed risk reducing derivative and an underlying asset, the payoff function of this portfolio $H(T)$ is determined by (11.9). Approximated distributions of rates of return, for given strike prices, are presented in the tables.The results are based on $n = 10^4$ simulations in each case, where a maturity T is one year, the underlying asset price S_0 is 50, $r = 10\%, \sigma = 30\%, t = 0.6$. Mean is equal 0 in each case, with accuracy to fifth decimal place.

As you can see (Tables 1.2, 1.3 1.4), the greatest losses have smaller probabilities in the case of portfolio II than in the case of portfolio I.Moreover probabilities of gain in the interval $(0\%, 25\%]$ corresponding to portfolio II exceed 0.5 for strike price $K \geq 60$.

In the case of $K = 70$ and $K = 80$ probabilities of moderate gains from portfolio II considerably exceed that from portfolio I.

Let us observe that in the cases of strike prices $K = 30, 40, 70, 80$ the probabilities of relative loss from considered portfolios have fairly close values while in

Table 11.2: Approximated distributions of rates of return for strike price K = 30 and K = 40

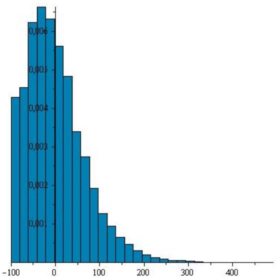
	I, K=30	II, K= 30	I, K = 40	II, K = 40
$P(R \leq -75\%)$	0.1	0	0.24	0
$P(-75\% < R \leq -50\%)$	0.1	0.06	0.13	0
$P(-50\% < R \leq -25\%)$	0.15	0.24	0.13	0.24
$P(-25\% < R \leq 0\%)$	0.18	0.27	0.11	0.38
$P(0\% < R \leq 25\%)$	0.14	0.2	0.08	0.18
$P(25\% < R \leq 50\%)$	0.1	0.12	0.07	0.1
$P(50\% < R \leq 75\%)$	0.1	0.06	0.05	0.05
$P(R > 75\%)$	0.13	0.05	0.19	0.05

Table 11.3: Approximated distributions of rates of return for strike price K = 50 and K = 60

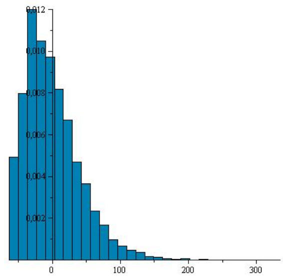
	I, K=50	II, K= 50	I, K = 60	II, K = 60
$P(R \leq -75\%)$	0.3	0	0.26	0
$P(-75\% < R \leq -50\%)$	0.1	0	0.08	0
$P(-50\% < R \leq -25\%)$	0.1	0.19	0.1	0.14
$P(-25\% < R \leq 0\%)$	0.1	0.28	0.1	0.27
$P(0\% < R \leq 25\%)$	0.1	0.43	0.1	0.55
$P(25\% < R \leq 50\%)$	0.08	0.05	0.09	0.02
$P(50\% < R \leq 75\%)$	0.06	0.02	0.08	0.01
$P(R > 75\%)$	0.16	0.03	0.19	0.01

the cases $K = 50$ and $K = 60$ the probabilities are different.If $K = 50$ portfolio I gives loss with probability 0.6, portfolio II with probability 0.47. Respective probabilities are equal to 0.54 and 0.41 in the case $K = 60$. In both last cases, gain from portfolio II is more probable. Large losses from portfolio II, exceeding 50%, have probability 0 in the cases $K \geq 40$ while such losses from portfolio I have probabilities $0.37, 0.4, 0.34, 0.26$ and 0.21 for $K = 40, K = 50, K = 60, K = 70$ and $K = 80$ respectively. In the cases $K \geq 40$ the distributions of rates of return from portfolio II are concentrated around zero, much more than the distributions relating to portfolio I. Let us compare values of $P(-25\% < R \leq 25\%)$ relating to portfolio I and II respectively:0.19 and 0.56 for $K = 40$, 0.2 and 0.71 for $K = 50$, 0.2 and 0.82 for $K = 60$, 0.26 and 0.89 for $K = 70$, 0.32 and 0.95 for $K = 80$. Figure 1.2 presents histograms and medians of rates of return dependent on K, based on performed simulations.

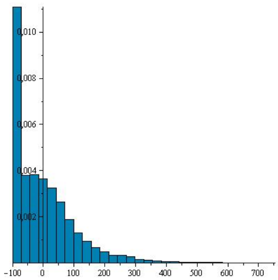
As can be seen the data set of simulated rates of return is unimodal and



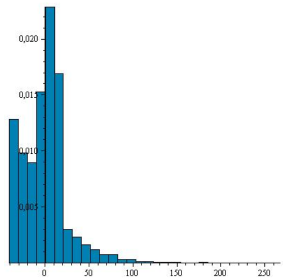
(a) Portfolio I, K=30, Median=-10



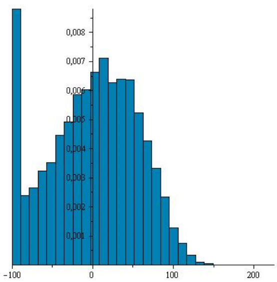
(b) Portfolio II, K=30, Median=-7



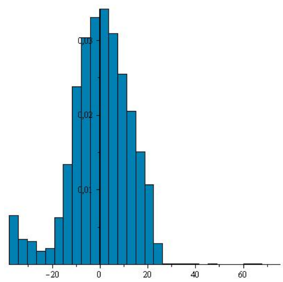
(c) Portfolio I, K=50, Median=-23



(d) Portfolio II, K=50, Median=1



(e) Portfolio I, K=80, Median=4



(f) Portfolio II, K=80, Median=1

Figure 11.2: Histograms and medians of rates of return

Table 11.4: Approximated distributions of rates of return for strike price K = 70 and K = 80

	I, K=70	II, K= 70	I, K = 80	II, K = 80
$P(R \leq -75\%)$	0.18	0	0.13	0
$P(-75\% < R \leq -50\%)$	0.08	0	0.08	0
$P(-50\% < R \leq -25\%)$	0.11	0.09	0.11	0.05
$P(-25\% < R \leq 0\%)$	0.13	0.34	0.15	0.43
$P(0\% < R \leq 25\%)$	0.13	0.55	0.17	0.52
$P(25\% < R \leq 50\%)$	0.12	0.01	0.16	0
$P(50\% < R \leq 75\%)$	0.1	0.01	0.12	0
$P(R > 75\%)$	0.15	0	0.08	0

positively skewed, when $K \leq S_0$. The data set of rate of return is multimodal in the case $K > S_0$.

11.5 Conclusions

Using Monte Carlo simulation we can determine distribution of rate of return from investments in options. Knowledge of this distribution helps to determine an investment risk. As demonstrated in this paper, the risk of incurring a large loss is smaller in case of investing in the proposed portfolio than in the case of investing in simple chooser options only. In both considered cases of portfolios, the distributions of rate of return and consequently level of risk are significantly dependent on the strike price.

References

[1] P. Billingsley. *Prawdopodobieństwo i miara*. Wydawnictwo Naukowe PWN, Warszawa, 2012.

[2] C. J. Corrado and T. W. Miller. “A note on a simple, accurate formula to compute implied standard deviations”. In: *Journal of Banking & Finance* 20, No. 3 (1996), pp. 595–603.

[3] K. Cuthbertson and D. Nitzsche. *Financial Engineering: Derivatives and Risk Management*. Wiley, New York, 2001.

[4] P. Glasserman. *Monte Carlo methods in financial engineering*. Springer, New York, 2004.

[5] E. Korn, R. Korn, and G. Kroisandt. *Monte Carlo methods and models in finance and insurance*. Chapman & Hall/CRC, Boca Raton, FL., 2010.

[6] R. Martinkute-Kauliene. "Exotic Options: A Chooser Option and its Pricing". In: *Business, Management and Education* 10, No. 2 (2012), pp. 289–301.

[7] A. Milian. "On Some Risk-Reducing Derivatives". (to appear).

[8] A. Milian. "Simple Chooser Options with Maple". to appear.

[9] M. Romaniuk. "Pricing the Risk-Transfer financial Instruments via Monte Carlo Methods". In: *Syst. Anal. Modell. Simul.* 43, No. 8 (2003), pp. 1043–1064.

[10] M. Romaniuk and T. Ermolieva. "Wycena obligacji katastroficznych metodami symulacyjnymi". In: *Badania Systemowe – Teoria i Zastosowania. Sesja Sprawozdawcza, Instytut BadaĹ„ Systemowych PAN, Warszawa, luty 2005.* 2005.

[11] A. Weron and R. Weron. *Inżynieria finansowa: wycena instrumentów pochodnych, symulacje komputerowe, statystyka rynku.* Wydawnictwa Naukowo-Techniczne, Warszawa, 1998.

Chapter 12

Finite axiomatization of logical matrices

KATARZYNA PAŁASIŃSKA

Katarzyna Pałasińska
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: kpalasinska@gmail.com

Abstract

This chapter is an overview of the current state of knowledge on the finite axiomatization problem for finite logical matrices.

An axiomatization of a mathematical theory consists of a set of axioms from which all laws true in the theory can be deduced. The finite axiomatization question in algebra, known as the *finite basis problem*, asks whether for a given finite algebra there is a finite set of basic equations such that every identity of the algebra can be derived from them, by means of *equational logic*. We consider a similar question for propositional logics determined by so-called *logical matrices*. The two–element Boolean algebra with the designated value 1 is an example of a matrix. It is well known that its tautologies all follow from a fnite set of axioms. Thus this matrix, used to define the tautologies of the Classical Propositional Logic, is finitely axiomatizable. The question whether for any finite logical matrix there is a finite axiomatization was considered in the literature in three versions. Counterexamples have been found for each of them, but under additional assumptions some positive results have been obtained. In connection with the result of [33], another question was asked: whether the finite axiomatization of a matrix is indepedent of the language, i.e., if it is

preserved under term-equivalence. Some class of logics for which the answer is positive was isolated in ([13]). It has also been recently verified for some special three-element matrices. The general version of the problem if the finite axiomatizability of a finite matrix depends on the language is open.

Contents

12.1 Introduction

This overview is based on the literature, although proofs of some of the results included in section 12.5 are still to be published. In the Introduction we summarize the chapter less formally. More technical details are left for the subsequent sections.

12.1.1 Axiomatizations

Mathematical theories are often presented by a finite set of *axioms* from which all true statements of the theory follow. Set theory or group theory are typical examples. In case of algebraic theories, their axioms take the form of equations and to deduce new equations from them the rules of the so-called *equational logic* are used. These rules express properties of the equality relation: its reflexivity, symmetry, transitivity and the property called *replacement* (see subsection 12.2.3). In general, in other mathematical theories one needs a richer deduction tool – the first order classical logic. This logic can also be presented in an axiomatic form.

In this chapter we will restrict our attention to *propositional logics*. The best known such logic is the classical one, which is the fragment of the first order classical logic not involving quantification – only propositions and connectives are used. Classical Propositional Logic (CPL for short) can be defined as the set of tautologies of the two-element Boolean algebra, i.e., formulas written in symbols used in CPL, such as $\wedge, \vee, \rightarrow, \neg$, that in this algebra always evaluate to 1. But CPL can also be presented as the set of theorems that follow from some axioms by means of a rule of Modus Ponens ((MP) for short). Such axioms and rule form an *axiomatization* of the CPL. Many different finite axiomatizations for CPL have been found. Here is one (due to Frege) that uses the following axioms and rule:

1. $t \rightarrow (s \rightarrow t)$,
2. $(t \rightarrow (s \rightarrow r)) \rightarrow ((t \rightarrow s) \rightarrow (t \rightarrow r))$,
3. $(\neg t \rightarrow \neg s) \rightarrow (s \rightarrow t)$.

(MP) $\dfrac{t,\ t \rightarrow s}{s}$

Other propositional connectives used in CPL such as $\wedge, \vee$ can be defined in terms of $\rightarrow, \neg$ in a standard way.

12.1.2 Finite axiomatizability problems for matrices

The 2-element Boolean algebra with the designated value 1 is an example of a *matrix* (Definition 12). In general, a matrix is an algebraic structure in which some elements have been *designated*. For each matrix, its *tautologies* can be defined as these formulas, written in the language appropriate for the matrix, that always evaluate to a designated value. If a finite martix $\mathfrak{M}$ has an implication connective, denoted by $\rightarrow$, then one may ask if it is finitely axomatizable relatively to the (MP) rule, i.e., if there is a finite set of some “basic" tautologies of $\mathfrak{M}$ from which all of its tautologies can be derived by means of this rule. In his paper [35] from 1977, M. Wajsberg showed that this not always must be the case. For every $n > 2$, he constructed an n-element matrix that is not finitely axiomatizable with the (MP) rule being the only rule allowed in the deduction. See subsection 12.3.2.

Because of the abstract definition of a matrix, the implication connective need not be one of its operations, so not always the (MP) rule can be expressed in its language. Moreover, even if it can there might be some rules other than (MP) that should, in some natural sense, be allowed in the deduction. These are so-called *valid rules* that will be defined later (Definition 18 in subsection 12.3.1). So one may generalize the above question and ask, for a given matrix and a fixed finite set $\mathcal{R}$ of its valid rules, whether there exists a finite set of tautologies of the matrix, from which all of its tautologies can be derived by means of the rules from $\mathcal{R}$ (see [38]). If this is the case, we will say that the matrix is *finitely axiomatizable relatively to* $\mathcal{R}$. Thus we have the following first version of the finite axiomatization problem for matrices:

Question 1. *Given a finite matrix $\mathfrak{M}$ and a finite set $\mathcal{R}$ of its valid rules, is $\mathfrak{M}$ finitely axiomatizable relatively to $\mathcal{R}$?*

It turns out (e.g., [38] or see [32]) that there exists such a finite set $\mathcal{R}$ of rules that all Wajsberg's matrices are finitely axiomatizable relatively to $\mathcal{R}$, see Theorem 23. So it makes sense to ask the following, more general question:

Question 2. *(P. Wojtylak [38]) Given a finite matrix $\mathfrak{M}$, does there exist* any *finite set $\mathcal{R}$ of its valid rules, such that $\mathfrak{M}$ is finitely axiomatizable relatively to $\mathcal{R}$?*

If the answer to Question 1 is "yes" then, obviously, the answer to Question 2 is also positive. But the converse need not be true, as the example of Wajsberg's matrices witnesses. Also in [38], P. Wojtylak presented an example of a finite matrix for which the answer to Question 2 is negative. It was later improved in [37, 12, 26], where smaller counterexamples were found, see section 12.4.

A tautology can be identified with a valid rule of a special form, namely a rule with the empty set of premisses. Therefore, Question 2 can be reformulated as ([38]):

Given a finite matrix $\mathfrak{M}$ does there exist a finite set $\mathcal{R}$ of its valid rules, such that all tautologies of $\mathfrak{M}$ can be deduced from these rules?

One may also ask a more general question:

Question 3. *(R. Suszko, S. L. Bloom, R. Wójcicki) Given a finite matrix, is there some finite set of rules valid in $\mathfrak{M}$, from which all rules valid in $\mathfrak{M}$ can be derived?*

Such a set of rules is called a *basis* for (the consequence operation of) $\mathfrak{M}$ and Question 3 is called the *finite basis problem* for logical matrices. It has been investigated, among others in [8, 7, 40, 34, 39, 33, 9, 11, 10, 6]. Notice that the positive answer to Question 3 obviously implies the positive answer to Question 2 Thus the positive answer to Question 2 is the weakest positive and the negative one is the strongest negative answer. Hence a matrix that is not finitely axiomatizable in Wojtylak's sense neither is finitely based nor is finitely axiomatizable relatively to any fixed set of rules. When we speak about "**a**" finite axiomatization problem, we mean any of the three; while"**the**" finite axomatization problem is the one in Wojtylak's sense.

12.1.3 Connections to Universal Algebra

The finite axiomatization and finite basis problems for matrices are related to the *finite basis problem* in Universal Algebra (UA). This important problem asks if for a given finite algebra **A** there is a finite set of basic equations from which all equations true in the algebra, and only these, can be deduced by means of equational logic. Such a set of equations is called a *basis* of the algebra and if it is finite then, in the terminology of UA, the algebra is called *finitely based*. An equation true in an algebra is also called its *identity*.

Consider, for example, a finite group. The set of its identities is strictly larger than the set of identities satisfied in all groups, so it makes sense to ask if there is a finite set of basic identities of this particular group, from which all of its identities follow. This question was first asked by J. von Neumann in 1937. In the paper [22] of 1964 it has been proved that the answer is "yes" for every finite group.

Similar theorems hold for finite rings ([15, 16]), finite lattices ([19]) and in general for every finite algebra generating a so-called *residually small congruence modular variety* ([20]). This last theorem generalizes an earlier theorem from 1976 due to K. Baker [1]. Baker's theorem says that if a finite algebra generates a so-called *congruence distributive variety* then it has a finite basis for its identities (congruence distributivity is a stronger property than congruence-modularity). An important generalization in another direction, to so-called congruence meet-semidistributive varieties, was given in [36], see also [28, 36, 3, 14, 2, 18], among others. It is also worth mentioning that every two-element algebra is finitely based, [17]. But in general, the answer to the finite basis question in UA is negative: many finite algebras for which no finite basis exists, were discovered.

The finite basis problem in UA and the finite axiomatization relatively to a fixed set of rules can be expressed as one problem in the terminology of so-called *k-deductive systems* introduced in [4], see Definition 7 in subsection 12.2.4

If the set of rules valid in a given matrix forms a so-called *algebraizable deduc-*

tive system defined in [5], then the finite axiomatization problems translate directly to their corresponding problems in UA. For example, CPL is algebraizable, as are some nonclassical logics like intuitionistic propositional logic or Łukasiewicz logic. For nonalgebraizable logics the study of a finite axiomatization does not seem to be reducible to studying finite axiomatization in algebra. For example, it may happen that a matrix is not finitely axiomatizable while its underlying algebra is, as is the case with matrices disscussed in section 12.4.

12.1.4 Some known finite axiomatization results for matrices

Some of the positive finite basis results in universal algebra that were mentioned in subsection 12.1.3 have their analogues for matrices. First of all, all two element matrices are finitely based [13]. Analogues of the Baker's finite basis theorem for propositional logics were proved by J. Czelakowski [9, 10, 11]: if the logic determined by a finite matrix is *filter-distributive* and some additional conditions hold, then this matrix is finitely based and therefore also finitely axiomatizable. The property of filter-distributivity is the analogue of the congruence-distributivity in UA. Both Baker's and Czelakowski's theorems have been generalized to a theorem concerning filter-distributive Universal Horn Logics [23, 24], from which also the main result of [28] follows. A new proof of this theorem can be found in [21]. In the general case, the finite axiomatization of a finite matrix may fail, see Section 12.4.

12.1.5 Term-equivalence

We often define new operations in terms of the basic ones. For example in the presentation of CPL in subsection 12.1.1 it is enough to use only $\rightarrow$ and $\neg$, because other connectives can be defined in their terms. It often happens that also conversely, the basic operations can be expressed in terms of the new ones. Such is the case with the "division" operation in groups: $a \div b = a \cdot b^{-1}$ and then $a^{-1} = e \div a$. Each group can be equivalently presented using the multiplication and division operations rather than multiplication and inverse, as it is usually done. The presentations, using different sets of mutually definable operations, are called *term-equivalent*. It is easy to prove that two term-equivalent algebras are either both finitely based, or none is. When it comes to matrices, however, then it seems far from obvious if finite axiomatization properties are preserved under term-equivalence. This question was asked by professor W. Rautenberg in connection with his work in [33] and so far the answer is not known.

Question 4. *(W. Rautenberg) Given that a matrix $\mathfrak{M}$ is finitely axiomatizable and $\mathfrak{N}$ is term-equivalent to $\mathfrak{M}$, is $\mathfrak{N}$ also finitely axomatizable?*

Here, finite axiomatizability can be understood in any of the three senses, so Question 4 really generates three questions. It would be obviously so, if the deductive system determined by the matrix was algebraizable. The problem was studied in [13], where it was shown, among others, that all two-element matrices are finitely based independently of their language. It has been verified in some cases of three-element matrices, [25, 27], see section 12.5.

12.2 Basic concepts

The concepts that were used informally so far are now going to be precisely defined.

12.2.1 Algebras

An algebra in the sense of universal algebra is a structure consisting of a nonempty set A endowed with some set F of finitary operations. We assume in addition that the set F of operations is finite. For a natural number n, by an n-ary *operation* on a set A we mean a function $f : A^n \to A$, where A^n is the n-th Cartesian power of the set A. If $n = 0$ the operation is called nullary and identified with the constant it picks from the set A.

Definition 1. *Let A be a nonempty set and let F be some finite set of operations on A. The pair $\mathbf{A} = \langle A, F\rangle$ is called an algebra. The set A is then called the universe of $\mathbf{A}$.*

It is customary to denote algebras by boldface capital letters. The universe of an algebra is then denoted by the same capital non-boldface letter. Monoids, groups, rings, modules over finite rings, lattices or Boolean algebras are examples of algebras. Assume that the operations in the set F are indexed by natural numbers, so for example $F = \{f_1, \ldots, f_k\}$. By definition, each operation $f_i \in F$ has some arity, let it be n_i. The function $n : \{1, \ldots, k\} \to \mathbb{N}$ is then a finite sequence $\langle n_i\rangle_i$ of natural numbers, called the *type* of the algebra. Two algebras of the same type are also called *similar*. To illustrate these concepts, let us recall the notion of a Boolean algebra.

Definition 2. *By a Boolean algebra we mean an algebra*

$$\mathbf{B} = \langle B, \{\vee, \wedge, \neg, 0, 1\}\rangle,$$

of type $\langle 2, 2, 1, 0, 0\rangle$ such that the following equalities hold in $\mathbf{B}$:

$x \vee y = y \vee x$;	$x \wedge y = y \wedge x$
$(x \vee y) \vee z = x \vee (y \vee z)$;	$(x \wedge y) \wedge z = x \wedge (y \wedge z)$
$(x \vee y) \wedge x = x$;	$(x \wedge y) \vee x = x$
$(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z)$;	$(x \wedge y) \vee z = (x \vee z) \wedge (y \vee z)$
$(x \vee \neg x) = 1$;	$(x \wedge \neg x) = 0$
$(x \vee 0) = x$;	$(x \wedge 1) = x$

There are many different equivalent definitions of Boolean algebras. In particular, Boolean algebras can be defined using a different set of operations, for example: $\langle \rightarrow, \neg \rangle$ of type $\langle 2, 1 \rangle$, or with just one binary operation $|$ known as the Sheffer stroke. The axioms should then be chosen appropriately. These other presentations would be *term-equivalent* to each other. The 2-element Boolean algebra $\langle \{0, 1\}, \{\vee, \wedge, \neg, 0, 1\} \rangle$ will be denoted by $\mathbf{B}_2$.

Functions between two similar algebras, preserving corresponding operations, are called *h*omomorphisms.

12.2.2 Language and terms

There is a difference between a symbol used to denote an operation and the operation itself. The former is just a sign or a character, while the latter is a specific function. So we distinguish between a set of operations, say F, and a set of symbols, say Λ ($\Lambda = \Lambda_F$), used to denote members of F. With each such symbol $\lambda \in \Lambda$ there is associated its arity, $\rho(\lambda)$, the same as the arity of the corresponding operation.

Definition 3. *By a* language *we mean a pair $\langle \Lambda, \rho \rangle$ such that Λ is a finite set of symbols and $\rho : \Lambda \rightarrow \mathbb{N}$. The function ρ is called the arity function of the language. An* interpretation *of the language $\langle \Lambda, \rho \rangle$ on a set A is a function f assigning to each symbol $\lambda \in \Lambda$ an $\rho(\lambda)$-ary operation f_λ on A, so that $f_\lambda : A^{\rho(\lambda)} \rightarrow A$. This turns A into an algebra $\mathbf{A} = \langle A, \{f_\lambda : \lambda \in \Lambda\} \rangle$ of type $\langle \rho(\lambda) \rangle_{\lambda \in \Lambda}$ indexed by Λ. We then also say that $\mathbf{A}$ is an algebra of the* signature *Λ.*

We often abuse terminology by identifying a language $\langle \Lambda, \rho \rangle$ with its set of symbols Λ, assuming that ρ is implicit. Let Λ be a language and let V be a fixed, countable set of variables, $V = \{x_1, x_2, \ldots\}$. We use the abbreviations: $x := x_1, y := x_2, z := x_3$. The set of all *terms* in variables from V that can be built up using the operation symbols from Λ is denoted by Te_Λ and is defined as follows.

Definition 4. *By Te_Λ we mean the smallest set such that:*

- $V \subseteq Te_\Lambda$

- *for any n, any symbol $\lambda \in \Lambda$ such that $\rho(\lambda) = n$ and for any sequence $t_1, \ldots, t_n \in Te_\Lambda$, the expression $\lambda(t_1, \ldots, t_n) \in Te_\Lambda$.*

The elements of the set Te_Λ are called terms *of the language Λ over V.*

Examples of terms in the language of $\{\to, \neg\}$ are x, $x \to y$, $(\neg x \to \neg y) \to (y \to x)$. In propositional logic, such expressions are usually called *formulas.* The set Te_Λ is turned into an algebra $\mathbf{Te}_\Lambda$ in the standard way: each n-ary operation symbol λ from Λ induces an operation on Te_Λ taking n terms $t_1, \ldots, t_n$ into the term $\lambda(t_1, \ldots, t_n)$. Homomorphisms from this algebra into an algebra $\mathbf{A} = \langle A, F\rangle$ of the signature Λ are called *valuations* in $\mathbf{A}$. Each valuation is uniquely determined by its restriction to variables. If the target algebra of a valuation is the algebra of terms then the valuation is called a *substitution.* For terms $t, s \in \mathrm{Te}_\Lambda$ and a variable $v \in V$, we write $t[s/v]$ for $\sigma(t)$, where σ is the substitution such that $\sigma(v) = s$ and $\sigma(x_i) = x_i$ for all $x_i \neq v$.

12.2.3 Equational logic

Definition 5. *Given a language Λ, by an equation we mean a pair $\varepsilon = \langle s, t\rangle$, written in the form $s \approx t$, where s and t are terms. A quasi-equation is the implication of the form $\varepsilon_1 \wedge \cdots \wedge \varepsilon_m \Rightarrow \varepsilon$, where m is a natural number and for each $i = 1, \ldots, m$, ε_i is an equation. If $m = 0$, we identify the quasi-equation with the equation ε.*

Let us stress that an equation is just a pair of terms and when writing it we use the symbol "$\approx$" in the same sense as a comma. The symbol "$=$" is used to denote the actual identity of objects.

Definition 6. *Let $\mathbf{A}$ be an algebra of type ρ and $\langle \Lambda, \rho\rangle$ a language. We say that an equation $\varepsilon = s \approx t$ in this language is an identity of $\mathbf{A}$ if and only if for every valuation v into $\mathbf{A}$ we have: $v(s) = v(t)$.*

A quasi-equation $s_1 \approx t_1 \wedge \cdots \wedge s_m \approx t_m \Rightarrow s \approx t$ is called a quasi-identity *of $\mathbf{A}$ if for every valuation v into $\mathbf{A}$ we have:*

$$\text{if } v(s_1) = v(t_1), \ldots, v(s_m) = v(t_m) \text{ then also } v(s) = v(t).$$

The following axiom and rules are used to deduce new equations from the old ones. Here $t, s, t_1, \ldots, t_n, s_1, \ldots, s_n$ are arbitrary terms and $\lambda \in \Lambda$ is an n-ary operation symbol of the language.

(refl) $t \approx t$

(sym) If $t \approx s$ then $s \approx t$

(tran) If $t_1 \approx t_2$ and $t_2 \approx t_3$ then $t_1 \approx t_3$

(repl) If $t_1 \approx s_1, \ldots, t_n \approx s_n$ then $\lambda(t_1, \ldots, t_n) \approx \lambda(s_1, \ldots s_n)$.

The axiom (refl) is called reflexivity, and the rules (sym), (trans), (repl) are called symmetry, transitivity and replacement rules, respectively. The system of these four: axiom and rules is called *equational logic*. It is an example of a deductive system, as defined in the next subsection.

12.2.4 Deductive systems

A deductive system is understood as a set of rules of deduction written in some language Λ. Axioms are treated as a special case of rules. Both equational logic and propositional logics are examples of deductive systems, although the formulas considered in them are different in form. The formulas used in propositional logic are just terms over Λ while the formulas of equational logic are equations. A common term that can be used for both types of deductive systems is that of a *k-deductive system*, introduced in [4].

Definition 7. *A k-formula is a k-tuple $\langle t_1, \ldots t_k \rangle$, where $t_1, \ldots, t_k$ are terms.*

The set of all k-formulas of a given language Λ is the Cartesian power $(Te_\Lambda)^k$, which will be written as Te_Λ^k.

Definition 8. *For a finite set $X \cup \{\alpha\} \subseteq Te_\Lambda^k$ the pair $\langle X, \alpha \rangle$, written as $\dfrac{X}{\alpha}$, is called a k-*rule. *The members of the set X are called* premisses *and the k-formula α* the conclusion *of the rule. A rule with the empty set of premisses is identified with its conclusion and is called* axiomatic.

Axiomatic rules are also called *axioms*. For $k = 1$ a k-formula is just a term, which in the context of propositional language is also called a formula. Consider $k = 2$. Since the only 2-deductive systems we consider here are equational logic and its extensions, we write 2-formulas as equations and rules as quasi-equations.

Definition 9. *A k-deductive system is a pair $\mathcal{S} = \langle \Lambda, \mathcal{R} \rangle$, where Λ is a language and $\mathcal{R}$ is a set of k-rules. If the language Λ is fixed, we identify a k-deductive system with the set $\mathcal{R}$ of its k-rules.*

If σ is a substitution and $t = \langle t_1, \ldots, t_k \rangle$ a k-term, then

$$\sigma(t) = \langle \sigma(t_1), \ldots, \sigma(t_k) \rangle.$$

A similar convention is used for valuations. The notion of a *derivation* or *proof* of a k-term in a given k-deductive system $\mathcal{S}$ is defined as follows.

Definition 10. *Let $\mathcal{S} = \langle \Lambda, \mathcal{R} \rangle$ be a k- deductive system. Let X be a set of k-formulas, called the* premisses *of the derivation. Let α be a k-formula and let n be a positive natural number. A sequence $\alpha_1, \ldots, \alpha_n$ of k-terms is called a* derivation *or a* proof *of α from X if and only if*

- $\alpha = \alpha_n$ *and*
- *for each $i = 1, \ldots, n$ either $\alpha_i \in X$ or there exists a rule $\frac{t_1, \ldots, t_m}{t} \in \mathcal{R}$ and a substitution σ such that $\alpha_1 = \sigma(t)$ and for each $j = 1, \ldots, m$, $\sigma(t_j) \in \{\alpha_1, \ldots, \alpha_{i-1}\}$.*

If there is a derivation of a k-formula α from a set of premisses X then we say that the rule $\frac{X}{\alpha}$ is a derived rule *in $\mathcal{S}$ or that it is* derivable *from $\mathcal{R}$. In case $X = \emptyset$ we say that α is a* theorem *of $\mathcal{S}$ and that α is* derivable *from $\mathcal{R}$.*

Definition 11. *Let $\mathcal{S}$ be a k-deductive system, $\mathcal{S} = \langle \Lambda, \mathcal{R} \rangle$. Let $\mathcal{R}'$ be some subset of the set of all k-rules that are derivable in $\mathcal{S}$.*

- *We say that $\mathcal{R}'$ is an* axiomatization *of $\mathcal{S}$ if and only if for every k-formula α, α is a theorem of $\mathcal{S}$ if and only if α is derivable from $\mathcal{R}'$.*
- *The set $\mathcal{R}'$ is called a* basis *for $\mathcal{S}$ if and only if each rule $\frac{X}{\alpha} \in \mathcal{R}$ is derivable from $\mathcal{R}'$.*
- *If the set $\mathcal{R}'$ can be chosen finite, we say that $\mathcal{S}$ is* finitely axiomatizable *or* finitely based, *respectively.*

In other words, a basis is such a set of k-rules that can replace $\mathcal{R}$ in the presentation of the k-deductive system. Notice that every basis is also an axiomatization. It is so, because each theorem is also a derived rule – an axiomatic rule. For example the set of axioms and the (MP) rule for CPL, given in the Introduction (subsection 12.1.1), forms both a basis and an axiomatization for this 1-deductive system.

With each algebra **A** a 2-deductive system can be associated that extends the system of equational logic and this can be done in two ways. The first way is to add all the identities of **A** to the rules of equational logic. Because the set of rules of equational logic is finite, this 2-eductive system is finitely axiomatizable exactly in the case it is finitely based in the sense of Definition 11. But this is equivalent to **A** being finitely based in the sense of UA: all identities of **A** can be deduced in equational logic from a finite set of identities.

The second 2-deductive system associated with **A** consists of all 2-rules, i.e., quasi-equations that are quasi-identities of **A**. This 2-deductive system is finitely based if and only if there is a finite set of quasi-identities of **A** from which all quasi-identities of **A** can be derived. This is sometimes called a finite basis property for quasi-identities of **A**. This system is finitely axiomatizable if and only if all of the identities of **A** can be derived from a finite set of its quasi-identities.

A 1-deductive analogue of an algebra is a *matrix*, the concept which we are now going to introduce.

12.2.5 Matrices and their tautologies

In the Boolean algebra $\mathbf{B}_2$ the element 1 plays a special role: all expressions that always evaluate to 1 are called tautologies of CPL. This is a paradigm for the concept of a matrix: it is an algebra in which one or more special elements have been selected or "designated".

Definition 12. *By a logical matrix (or just matrix, for short), we mean a system $\mathfrak{M} = \langle M, F, D\rangle$, where $\mathbf{M} = \langle M, F\rangle$ is an algebra and D is a nonempty subset of M. The elements of D are called the designated values of $\mathfrak{M}$.*

The notion of a k-matrix, for a natural number $k > 0$, generalizes this.

Definition 13. *By a k-matrix we mean a system $\mathfrak{M} = \langle M, F, D\rangle$, where $\mathbf{M} = \langle M, F\rangle$ is an algebra and D is a nonempty subset of M^k.*

Example 14. *A 2-matrix is an algebra $\mathbf{A}$ with a distinguished binary relation on it. If this relation is the diagonal relation on A then the 2-tautologies of the matrix coincide with the identities of the algebra. We skip the prefix k when it is equal to 1. The two element Boolean matrix is the matrix*

$$\mathfrak{B}_2 = \langle\{0,1\}, \{\vee, \wedge, \neg, 0, 1\}, \{1\}\rangle,$$

where $\langle\{0,1\}, \{\vee, \wedge, \neg, 0, 1\}\rangle$ is the two-element Boolean algebra $\mathbf{B}_2$.

A term-equivalent definition could use $\mathfrak{B}_2 = \langle\{0,1\}, \{\rightarrow, \neg\}, \{1\}\rangle$, with $x \rightarrow y = \neg x \vee y$. Then $x \vee y = \neg x \rightarrow y$ and $x \wedge y = \neg(x \rightarrow \neg y)$.

The following definition generalizes the notion of a classical tautology.

Definition 15. *Let $\mathfrak{M} = \langle M, F, D\rangle$, be a matrix. A k-formula t is called a tautology of $\mathfrak{M}$ if and only if for every valuation into $\mathbf{M}$, $v(t) \in D$. The set of all tautologies of a k-matrix $\mathfrak{M}$ is denoted by $E(\mathfrak{M})$.*

Matrices in the following two examples are associated with the 3-valued Łukasiewicz propositional logic and the 3-valued intuitionistic propositional logic, respectively.

Example 16. *The three-element matrix of Łukasiewicz is the matrix:*

$$\mathfrak{L}_3 = \left\langle \{0, \frac{1}{2}, 1\}, \{\wedge, \vee, \to, \neg\}, \{1\} \right\rangle,$$

with the operations given in the tables below:

$\wedge$	0	$\frac{1}{2}$	1
0	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	1

$\vee$	0	$\frac{1}{2}$	1
0	0	$\frac{1}{2}$	1
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
1	1	1	1

$\to$	0	$\frac{1}{2}$	1
0	1	1	1
$\frac{1}{2}$	$\frac{1}{2}$	1	1
1	0	$\frac{1}{2}$	1

x	$\neg x$
0	1
$\frac{1}{2}$	$\frac{1}{2}$
1	0

For every x in the set $\{0, \frac{1}{2}, 1\}$, the term $\neg\neg x \to x$ is equal to 1, so this formula, known as the law of double negation, is a tautology of $\mathfrak{L}_3$. On the other hand the law of contradiction $\neg(x \wedge \neg x)$ is not. This can be seen by taking $x = \frac{1}{2}$:

$$\neg\left(\frac{1}{2} \wedge \neg\frac{1}{2}\right) = \neg\left(\frac{1}{2} \wedge \frac{1}{2}\right) = \neg\frac{1}{2} = \frac{1}{2} \neq 1.$$

Example 17. *The three-element Heyting matrix is the matrix*

$$\mathfrak{H}_3 = \left\langle \{0, \frac{1}{2} 1\}, \{\wedge, \vee, \to, \neg\}, \{1\} \right\rangle,$$

where the operations are defined in the following tables:

$\wedge$	0	$\frac{1}{2}$	1
0	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	1

$\vee$	0	$\frac{1}{2}$	1
0	0	$\frac{1}{2}$	1
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
1	1	1	1

$\to$	0	$\frac{1}{2}$	1
0	1	1	1
$\frac{1}{2}$	0	1	1
1	0	$\frac{1}{2}$	1

x	$\neg x$
0	1
$\frac{1}{2}$	0
1	0

The law of contradiction, $\neg(x \wedge \neg x)$, is now a tautology of $\mathfrak{H}_3$. On the other hand, the law of double negation, $\neg\neg x \to x$ is not. Again, the witness is $x = \frac{1}{2}$.

Notice that the two matrices above differ only in $\frac{1}{2} \to 0$ and $\neg\frac{1}{2}$. This small difference is enough to make quite an impact on their tautologies.

12.3 Finite axiomatization of logical matrices

12.3.1 Deductive system associated with a matrix

For a given $k > 0$, with every k-matrix $\mathfrak{M}$ there is a k-deductive system $\mathcal{S}_{\mathfrak{M}}$ associated in a natural way. It is determined by all *valid rules* of $\mathfrak{M}$, which we now define.

Definition 18. *Let $\Lambda = \langle \Lambda, \rho \rangle$ be a language and let* $\mathbf{M}$ *be an algebra in the signature* Λ*. Let* $\mathfrak{M} = \langle \mathbf{M}, D \rangle$ *for some subset* D *of* M^k*. Assume that* $X \cup \{\alpha\} \subseteq Te_\Lambda^k$ *and consider a rule* $\dfrac{X}{\alpha}$*. We say that this rule is* valid *in* $\mathfrak{M}$ *iff for every valuation* φ*, we have:* $\varphi(X) \subseteq D \Rightarrow \varphi(\alpha) \in D$*. The set of all rules valid in a* k*-matrix* $\mathfrak{M}$ *is denoted by* $\mathcal{R}_{\mathfrak{M}}$*.*

The Modus Ponens rule is an example of a rule valid in every 1-matrix considered so far. If $\mathbf{A}$ is an algebra and Δ_A denotes the diagonal relation on A, then the 2-rules valid in the 2-matrix $\mathfrak{A}$ correspond to the quasi-identies of $\mathbf{A}$ in the following sense. A 2-rule $\dfrac{\{\epsilon_1, \ldots, \epsilon_n\}}{\alpha}$ is valid in $\mathfrak{A}$ iff $\epsilon_1 \wedge \cdots \wedge \epsilon_n \to \alpha$ is a quasi-identity of $\mathbf{A}$. It is also easy to see that when $\mathfrak{A} = \langle \mathbf{A}, R \rangle$ is a 2-matrix then R is a congruence on $\mathbf{A}$ if and only if all rules of equational logic are valid in $\mathfrak{A}$ (a *congruence* on $\mathbf{A}$ is an equivalence relation that is a subalgebra of the product $\mathbf{A}^2$). It turns out that the 2-tautologies of $\mathfrak{A}$ coincide with the identities of the quotient algebra $\mathfrak{A}/R$. Similarly, valid 2-rules coincide with quasi-identities of this quotient.

Definition 19. *Let $\Lambda = \langle \Lambda, \rho \rangle$ be a language and let* $\mathbf{M}$ *be an algebra in the signature* Λ*. Let* $\mathfrak{M} = \langle \mathbf{M}, D \rangle$ *for some subset* D *of* M^k*. The deductive system associated with* $\mathfrak{M}$ *is the* k*-deductive system* $\mathcal{S}_{\mathfrak{M}} = \langle \Lambda, \mathcal{R}_{\mathfrak{M}} \rangle$*.*

We say that a matrix $\mathfrak{M}$ is *finitely axiomatizable,* or *finiely based*, respectively, if the system $\mathcal{S}_{\mathfrak{M}}$ is such, in the sense of Definition 11. This is equivalent to the following definition.

Definition 20. *A* k*-matrix is*

- finitely axiomatizable *if there is a finite set of its valid* k*-rules from which all tautologies of* $\mathfrak{M}$ *can be derived. It is*
- finitely based *if there is a finite set of its valid* k*-rules from which all valid rules of* $\mathfrak{M}$ *can be derived.*

Suppose there is a fixed set $\mathcal{R}_0$ of k-rules valid in a k-matrix $\mathfrak{M}$. Then one can consider the k-deductive system $\mathcal{S}^0_{\mathfrak{M}} = \langle \Lambda, \mathcal{R}_0 \cup E(\mathfrak{M}) \rangle$ and ask if it is finitely based. It can be shown by a straightforward argument that if the set $\mathcal{R}_0$ is finite then this system is finitely based exactly in the case it is finitely axiomatizable. If this is the case we say that $\mathfrak{M}$ is *finitely axiomatizable relatively to* $\mathcal{R}_0$. According to Definition 11 this is equivalent to the following definition.

Definition 21. *Let* $\mathfrak{M}$ *be a* k*-matrix and let* $\mathcal{R}_0$ *be a finite set of its valid* k*-rules. We say that* $\mathfrak{M}$ *is* finitely axiomatizable relatively to $\mathcal{R}_0$ *if and only if there is a finite set* E_0 *such that all tautologies of* $\mathfrak{M}$ *can be derived from* E_0 *using the rules from* $\mathcal{R}_0$*.*

Let $\mathfrak{M}$ be a finite 1-matrix. Question 1 stated in the Introduction asks if for a given finite set $\mathcal{R}$ of its valid rules, $\mathfrak{M}$ is finitely axiomatizable relatively to $\mathcal{R}$. Question 2 asks if $\mathfrak{M}$ is finitely axiomatizable and Question 3, if it is finitely based. Let $\mathfrak{A} = \langle \mathbf{A}, R\rangle$ be a finite 2-matrix in which the rules of equational logic are valid. Then $\mathfrak{A}$ is finitely axiomatizable as a 2-matrix, relatively to equational logic, if and only if $\mathbf{A}/R$ is finitely based in the sense of UA. As was indicated before, this is equivalent to saying that the 2-deductive system determined by the rules of equational logic and the 2-tautologies of $\mathfrak{A}$ (i.e., identities of the algebra $\mathbf{A}/R$ is finitely based.

12.3.2 Wajsberg's matrices

As mentioned in the Introduction, the problem of finite axiomatization of finite 1-matrices initially focused on tautologies. For a given logical system expressed with standard connectives, such as the implication, conjunction, disjunction and negation, a finite set of axioms was searched for from which all tautologies could be derived by means of Modus Ponens. In 1977 M. Wajsberg provided a sequence of counterexamples, one for each $n > 2$, that we now recall.

Example 22. *([35]) For each $n > 2$ the n-valued Wajsberg matrix is*

$$\mathfrak{M}_n = \Big\langle \{0, 1, \ldots, n-1\}, \neg, \rightarrow, \{0\} \Big\rangle,$$

where for $x, y \in \{0, 1, \ldots, n-1\}$ $\quad x \rightarrow y = y$ and $\neg x = 0$.

It is easy to show that the set of tautologies of this matrix is closed under the Modus Ponens rule. Yet this rule is not strong enough for a finite set of axioms to exist that together with this rule would allow to deduce all tautologies.

Theorem 23. *(M. Wajsberg) For any n, the matrix $\mathfrak{M}_n$ is not finitely axiomatizable relatively to (MP).*

If we allow another rule, however, then a finite axiomatization involvig this rule and a finite set of axioms can be found.

Theorem 24. *The rule $\dfrac{y}{x \rightarrow y}$ and the axiom $\neg x$ form a finite axiomatization of the n-valued Wajsberg matrix, for each $n > 2$.*

Wajsberg's matrices are also finitely based, which is a stronger property than finite axiomatization. The basis is $\left\{ \dfrac{y}{x \rightarrow y}, \dfrac{x \rightarrow y}{y}, \dfrac{\emptyset}{\neg x} \right\}$. [32] contains a detailed explanation of this fact in Polish.

The information given above is based on [37], where P. Wojtylak recalled Wajsberg's result and proposed the notion of finite axiomatizability in the sense of Question 2.

12.3.3 Two-element matrices and filter-distributive protoalgebraic logics

One of the earliest results was the proof in [33] that 2-element matrices in the Post-classification are finitely based and therefore finitely axiomatizable. Together with [13] this gives the following theorem.

Theorem 25. *(Rautenberg, Rautenberg and Herrmann) Every two-element matrix is finitely based.*

This result looks similar to Lyndon's theorem in UA ([17]) stating that every 2-element algebra is finitely based. Yet, as professor Rautenberg pointed out, the obvious fact that an algebra term-equivalent to a finitely based algebra is also finitely based, doesn't have an immediate counterpart in propositional logic. In fact, it is not known to be true that a finite matrix term-equivalent to a finitely based (or axiomatizable) matrix is itself finitely based (or axiomatized, resp.) We return to this problem in section 12.5.

In the theory of logical matrices several analogues of Baker's finite basis theorem in UA were proved, [9, 10, 11, 6]. A result from [23, 24] implies all of them, as well as the Baker's theorem and its generalization to quasivarieties from [28]. Its simplified version for matrices may be stated as follows

Theorem 26. *([23, 24]) Let $\mathfrak{M}$ be a finite k-matrix such that $\mathcal{S}_{\mathfrak{M}}$ is filter-distributive and protoalgebraic. Then $\mathfrak{M}$ is finitely based.*

The term *protoalgebraic* refers to a property of a system that is weaker than *algebraizability*, but still non-trivial. We refer the reader to [6] or [4]for definition. In [21] a new proof of this theorem was given by a technique in the spirit of [3]. The nonfinitely axiomatizable matrices that we will see in the next section are not protoalgebraic.

12.4 Certain three-element matrices

Since every *two*-element matrix is finitely based and finitely axiomatizable it was natural to ask if for some number greater than 2 there exist nonfinitely axiomatizable matrices of this size. First examples that were discovered had six, then five and four elements,[40, 34]. Eventually, as small as a three-element nonfinitely based matrix was found, [39]. This is the matrix $\mathfrak{M}_5$ in the Example 27 below. But $\mathfrak{M}_5$ is finitely axiomatizable, hence it also provides an example that finite axiomatizability does not imply being finitely based, see Theorem 28.

Concerning finite axiomatizability, recall that this notion was introduced in [38]. Also there the first non-finitely axiomatizable matrix was presented. Then a smaller, five-element one was given in [37]. A four-element nonfinitely axiomatizable matrix was subsequently discovered by W. Dziobiak, [12]. Finally, two three-element nonfinitely axiomatizable matrices were found, [26]. These are $\mathfrak{M}_7$ and $\mathfrak{M}_8$ in the example 27 below.

Example 27. *For* $i \in \{1, \dots, 8\}$ *consider a matrix* $\mathfrak{M}_i = \langle\{0, 1, 2\}, \cdot_i, \{2\}\rangle$, *where the operations* $\cdot_i$ *are given by the following tables.*

$\cdot_1$	0	1	2
0	1	2	2
1	1	2	2
2	1	2	2

$\cdot_2$	0	1	2
0	2	2	2
1	2	2	2
2	2	2	2

$\cdot_3$	0	1	2
0	**1**	2	2
1	2	2	2
2	2	2	2

$\cdot_4$	0	1	2
0	2	2	2
1	1	2	2
2	1	2	2

$\cdot_5$	0	1	2
0	2	2	2
1	**1**	2	2
2	2	2	2

$\cdot_6$	0	1	2
0	1	2	2
1	1	2	2
2	2	2	2

$\cdot_7$	0	1	2
0	2	2	2
1	2	2	2
2	**1**	2	2

$\cdot_8$	0	1	2
0	1	2	2
1	2	2	2
2	1	2	2

We will omit the symbol $\cdot$ when writing terms. As it turns out these small and simple matrices, with very similar multiplication tables, differ one from the other significantly when it comes to the finite aximatizability properties. We have the following

Theorem 28. *1.* $\mathfrak{M}_1$*—*$\mathfrak{M}_4$ *are all finitely based and therefore finitely axiomatizable*

2. $\mathfrak{M}_5$ *and* $\mathfrak{M}_6$ *are both finitely axiomatizable but none is finitely based*

3. $\mathfrak{M}_7$ *and* $\mathfrak{M}_8$ *are not finitely axiomatizable.*

The proof that $\mathfrak{M}_5$ is not finitely based can be found in [39] and its details have been worked out also in [31]. The proof for $\mathfrak{M}_6$ is similar to the one for $\mathfrak{M}_5$. The proof that $\mathfrak{M}_7$ and $\mathfrak{M}_8$ are not finitely axiomatizable was given in [26].

It can be observed that

Observation 29. *the term* $\mathbf{2} := x(yz)$ *is a tautology of every* $\mathfrak{M}_i$ *for* $i = 1, \dots, 8$.

In fact, these eight matrices are (up to isomorphism) all possible 3-element matrices of the form $\mathfrak{M} = \langle\{0, 1, 2\}, \cdot, \{2\}\rangle$, in which the formula $x(yz)$ is a tautology. The Observation 29 allows us to conclude:

Proposition 30. *Every non-tautology of a matrix* $\mathfrak{M}_i$ *must be of the form*

$$t = t_1 v_m \cdots v_1, \tag{12.1}$$

for some $m \geq 0$, *where* $v_1, \dots, v_m$ *are variables and* t_1 *is either a variable or a substitution instance of* $\mathbf{2}$.

Due to this proposition nontautologies of $\mathfrak{M}_i$ take a relatively simple form. Also, in view of Theorem 25, three is the smallest number of elements a nonfinitely axiomatizable matrix might have. Therefore, in some sense, the matrices $\mathfrak{M}_7$ and $\mathfrak{M}_8$ are "simplest possible" nonfinitely axiomatizable matrices.

One difference is worthy being pointed out between studying the finite axiomatization in logic and doing this in algebra. With algebras, we have at our disposal the replacement rule. There is no analogue of this rule in propositional logic associated with matrices. Often it turns out that some matrix is not finitely axiomatizable or finitely based, while its underlying algebra is. This fact has also an effect on Question 4 disscussed in Section 12.5.

12.5 Term-equivalence

An elegant and short definition of term-equivalence of algebras involves the notion of the *clone* of operations: two algebras with the same universe are *term-equivalent* if their clones of operations are equal. Here we give a definition omitting the concept of a clone. Let M be a nonempty set and let F and G be two sets of operations on M. Let Λ_F and Λ_G be the corresponding sets of operation symbols and $\mathbf{Te}_F = \mathbf{Te}_{\Lambda_F}$ and $\mathbf{Te}_G = \mathbf{Te}_{\Lambda_G}$ corresponding algebras of terms in variables V. For a valuation of variables $v : V \to M$, let $v_F : \mathbf{Te}_F \to \langle M, F\rangle$ and $v_G : \mathbf{Te}_G \to \langle M, G\rangle$ denote valuations of terms induced by v.

Definition 31. *Let $\mathfrak{M} = \langle M, F, D\rangle$ and $\mathfrak{N} = \langle M, G, D\rangle$ be two matrices with the same underlying set M and the same set of designated values D. We say that $\mathfrak{M}$ and $\mathfrak{N}$ are term-equivalent if*

1. *for every n and every n-ary operation $f \in F$ there is a term $t \in Te_G$ such that for every $v : V \to M$, $f(v(x_1), \ldots, v(x_n)) = v_G(t)$; and*

2. *for every n and every n-ary operation $g \in G$ there is a term $s \in Te_F$ such that for every $v : V \to M$, $g(v(x_1), \ldots, v(x_n)) = v_F(s)$.*

A straightforward proof in UA that an algebra term-equivalent to a finitely based algebra is also finitely based involves the equational logic and the replacement rule in particular. This proof cannot be directly repeated when dealing with matrices, for in propositional logic the replacement rule cannot be, in general, expressed. It was for this reason that the proof given in [33] that every two-element matrix in Post classification is fnitely based was not sufficient to imply the result for all two-element matrices. This lead professor W. Rautenberg to state Question 4. The answer for 2-valued logics was supplied in [13], where it was proved that for matrices with the *finite replacement property* , defined there, the finite basis property is preserved under term-equivalence. It was also proved there that all two-element matrices have the finite replacement property.

All three-element matrices from Example 27 were checked against Question 4. The proof that no matrix term-equivalent to $\mathfrak{M}_5$ is finitely based was given in [25]. The proof for $\mathfrak{M}_6$ is similar. The proofs that all matrices term-equivalent to $\mathfrak{M}_7$ and $\mathfrak{M}_8$ are non-finitely axiomatized can be found in [27], it has also been worked out down to every detail in [31]. Actually it is now known that for any of the eight matrices presented in Section 12.4 the finite axiomatizability property and the finite basis property are preserved under term-equivalence.

Theorem 32. *Let $\mathfrak{M}$ be a three-element matrix with one binary operation $\cdot$ and with one designated element. Assume that the term $x \cdot (y \cdot z)$ is a tautology of $\mathfrak{M}$. Then $\mathfrak{M}$ is finitely based if and only if every matrix term-equivalent to $\mathfrak{M}$ is finitely based. Also, $\mathfrak{M}$ is finitely axiomatizable if and only if every matrix term-equivalent to $\mathfrak{M}$ is finitely axiomatizable.*

Some cases of the proof have been worked out in [30] and [29], others are known to the author and will be published elsewhere. They are done by syntactic arguments.

Rautenberg's problem has been posed over 20 years ago and there is only a small insight into it so far. We do not even know what the answer is in the case of three-element matrices or in case of matrices in one binary operation $\cdot$ in which the term $x \cdot (y \cdot z)$ is a tautology.

References

[1] K. A. Baker. "Finite equational bases for finite algebras in a congruence-distributive equational class". In: *Adv. Math.* 24 (1977), pp. 207–243.

[2] K. A. Baker, G. F. McNulty, and J. Wang. "An extension of Willard's finite basis theorem: Congruence meet-semidistributive varieties of finite critical depth". In: *Algebra Univers.* 52, No. 2-3 (2004), pp. 289–302.

[3] K. A. Baker and J. Wang. "Definable principal subcongruences". In: *Algebra Univers.* 47, No. 2 (2002), pp. 145–152.

[4] W. J. Blok and D. Pigozzi. "Algebraic Semantics for Universal Horn Logic without Equality". In: *Universal algebra and quasigroup theory. Lectures delivered at the conference, held in Jadwisin, Poland, May 23-28, 1989.* Heldermann Verlag, Berlin, 1992, pp. 1–56.

[5] W. J. Blok and D. Pigozzi. "Algebraizable Logics". In: *Mem. Am. Math. Soc.* 369 (1989).

[6] W. J. Blok and D. Pigozzi. "Potoalgebraic Logics". In: *Stud. Log.* 45 (1986), pp. 337–369.

[7] S. L. Bloom. "A representation theorem for the lattice of standard consequence operation". In: *Stud. Log.* 34 (1975), pp. 235–237.

[8] S. L. Bloom. "Some theorems on structural consequence operations". In: *Stud. Log.* 34 (1975), pp. 1–9.

[9] J. Czelakowski. "Equivalential logics. I and II". In: *Stud. Log.* 40 (1981), pp. 227–236, 335–372.

[10] J. Czelakowski. "Filter distributive logics". In: *Stud. Log.* 43 (1984), pp. 353–377.

[11] J. Czelakowski. "Matrices, primitive satisfaction and finitely based logics". In: *Stud. Log.* 42 (1983), pp. 89–104.

[12] W. Dziobiak. "A finite matrix whose set of tautologies is not finitely axiomatizable". In: *Rep. Math. Logic* 25 (1991), pp. 113–117.

[13] B. Herrmann and W. Rautenberg. "Finite Replacement and Finite Hilbert-style Axiomatizability". In: *Z. Math. Logik Grundlagen Math.* 38, No. 4 (1992), pp. 327–344.

[14] K. A. Kearnes and R. Willard. "Residually finite, congruence meet-semidistributive varieties of finite type have a finite residual bound". In: *Proc. Am. Math. Soc.* 127, No. 10 (1999), pp. 2841–2850.

[15] R. L. Kruse. "Identities satisfied by a finite ring". In: *J. Algebra* 26 (1973), pp. 298–318.

[16] I. V. L'vov. "Varieties of associative rings. I, II". In: *Algebra Logika* 12 (1973), pp. 269–297.

[17] R. C. Lyndon. "Identities in two-valued calculi". In: *Trans. Am. Math. Soc.* 71 (1951), pp. 457–465.

[18] M. Maróti and R. McKenzie. "Finite basis problems and results for quasivarieties". In: *Stud. Log.* 78, No 1-2 (2004), pp. 293–320.

[19] R. McKenzie. "Equational bases for lattice theories". In: *Math. Scand.* 27 (1970), pp. 24–38.

[20] R. McKenzie. "Finite equational bases for congruence modular varieties". In: *Algebra Univers.* 24, No. 3 (1987), pp. 224–250.

[21] A. M. Nurakunov and M. M. Stronkowski. "Relation formulas for protoalgebraic equality free quasivarieties; Pałasińska's theorem revisited". In: *Stud. Log.* 101 (2013), pp. 827–847.

[22] S. Oates and M. B. Powell. "Identical relations in finite groups". In: *J. Algebra* 1 (1964), pp. 11–39.

[23] K. Pałasińska. "Deductive systems and finite axiomatization properties". PhD thesis. Iowa State University, 1994.

[24] K. Pałasińska. “Finite basis theorem for filter-distributive protoalgebraic deductive systems and strict universal Horn classes”. In: *Stud. Log.* 74, No. 1-2 (2003), pp. 233–273.

[25] K. Pałasińska. “No Matrix Term-Equivalent to Wroński’s 3-element Matrix is Finitely Based”. In: *Stud. Log.* 77, No. 3 (2004), pp. 413–423.

[26] K. Pałasińska. “Three-element nonfinitely axiomatizable matrices”. In: *Stud. Log.* 53, No. 3 (1994), pp. 361–372.

[27] K. Pałasińska. “Three-element non-finitely axiomatizable matrices and term-equivalence”. In: *Log. Log. Philos.* 23, No. 4 (2014), pp. 481–497.

[28] D. Pigozzi. “Finite basis theorem for relatively congruence-distributive quasivarieties”. In: *Trans. Am. Math. Soc.* 310, No. 2 (1988), pp. 499–533.

[29] R. Polek. “Skończona aksjomatyzowalność i skończona bazowalność matryc termalnie równoważnych pewnym trójelementowym matrycom logicznym”. MA thesis. Politechnika Krakowska, 2012.

[30] R. Polek. “Zagadnienie niezmienniczości własności skończonej aksjomatyzowalności matryc logicznych względem równoważności termalnej (praca licencjacka)”. MA thesis. Politechnika Krakowska, 2010.

[31] M. Porębska. “Brak skończonej aksjomatyzacji dla pewnych małych matryc logicznych i matryc im termalnie równoważnych”. MA thesis. Politechnika Krakowska, 2012.

[32] M. Porębska. “Własność skończonej aksjomatyzowalności dla małych matryc logicznych (praca licencjacka)”. MA thesis. Politechnika Krakowska, 2010.

[33] W. Rautenberg. “2-element Matrices”. In: *Stud. Log.* 40 (1981), pp. 315–353.

[34] A. Urquhart. “A finite matrix whose consequence relation is not finitely axiomatizable”. In: *Rep. Math. Logic* 9 (1977), pp. 71–73.

[35] M. Wajsberg. “On Axiomatizability of Matrices”. In: *Logical Works by Mordechaj Wajsberg.* Polish Academy of Sciences, Institute of Philosophy and Sociology, 1977, pp. 93–106.

[36] R. Willard. “A finite basis theorem for residually finite, congruence meet-semidistributive varieties”. In: *J. Symb. Log.* 65, No. 1 (2000), pp. 187–200.

[37] P. Wojtylak. “An example of a finite though finitely non-axiomatizable matrix”. In: *Rep. Math. Logic* 17 (1984), pp. 39–46.

[38] P. Wojtylak. “Strogly Finite Logics: Finite Axiomatizability and the Problem of Supremum”. In: *Bull. Sect. Logic, Pol. Acad. Sci.* 8 (1979), pp. 99–111.

[39] A. Wroński. “A Three-element Matrix whose Consequence Operation is not Finitely Based”. In: *Bull. Sect. Logic, Pol. Acad. Sci.* 8 (1979), pp. 68–71.

[40] A. Wroński. “On finitely based consequence operations”. In: *Stud. Log.* 35 (1976), pp. 453–458.

Chapter 13

Rank Functions in Algebra, Matrix Theory, and Geometry

Marcin Skrzyński

Marcin Skrzyński
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24, 31-155 Kraków
e-mail: pfskrzyn@cyf-kr.edu.pl

Abstract

A function $\rho : \mathbb{N} \longrightarrow \mathbb{N}$, where $\mathbb{N}$ denotes the set of non-negative integers, is said to be a rank function, if it is weakly decreasing and satisfies the convexity condition

$$\forall\, k \in \mathbb{N} :\ 2\rho(k+1) \leqslant \rho(k) + \rho(k+2).$$

In the article, we discuss some interrelations between the rank functions, semiring theory, matrices, and algebraic sets in matrix spaces.

Contents

13.1 Preliminaries and Introduction

Throughout the text, $\mathbb{F}$ stands for an arbitrary field and m, n, p for non-negative integers. The set of non-negative integers will be denoted by $\mathbb{N}$. We define $\mathcal{M}_n(\mathbb{F})$ to be the $\mathbb{F}$-algebra of all $n \times n$ matrices over $\mathbb{F}$ (notice that $\mathcal{M}_0(\mathbb{F})$ is a singleton set). The $m \times n$ zero matrix and the unit matrix in $\mathcal{M}_n(\mathbb{F})$ will be denoted by $O_{m\times n}$ and I_n, respectively. A permutation matrix is defined to be a square binary matrix that has exactly one 1 in each row and each column. The minors of matrices will be understood to be determinants.

The full linear group $\mathcal{GL}_n(\mathbb{F}) = \{U \in \mathcal{M}_n(\mathbb{F}) : U \text{ is invertible}\}$ acts on $\mathcal{M}_n(\mathbb{F})$ by conjugation. We define $\mathcal{O}(A)$ to be the orbit of a matrix $A \in \mathcal{M}_n(\mathbb{F})$ under this action, i.e.,

$$\mathcal{O}(A) = \{U^{-1}AU : U \in \mathcal{GL}_n(\mathbb{F})\}$$

(in other words, $\mathcal{O}(A)$ is the *conjugacy class* of A). A set $\mathcal{E} \subseteq \mathcal{M}_n(\mathbb{F})$ is called $\mathcal{GL}_n(\mathbb{F})$-*invariant*, if $\mathcal{O}(A) \subseteq \mathcal{E}$ for any matrix $A \in \mathcal{E}$.

All necessary information on matrices can be found in [6] and [7].

Let V be a vector space over $\mathbb{F}$. Suppose that $0 < d := \dim V \neq \infty$. A set $E \subseteq V$ is said to be algebraic, if there exist a linear isomorphism $\varphi : V \longrightarrow \mathbb{F}^d$, a positive integer s, and polynomials $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_d]$ such that

$$E = \{v \in V : f_1(\varphi(v)) = \ldots = f_s(\varphi(v)) = 0\}.$$

For further information on algebraic sets refer to [4]. In this article, an algebraic set $\mathcal{E} \subseteq \mathcal{M}_n(\mathbb{C})$ is called *normal*, if it is irreducible and its coordinate ring $\mathbb{C}[\mathcal{E}]$ is integrally closed in the function field $\mathbb{C}(\mathcal{E})$.

We will use basic concepts of order theory (see [3]).

Let us finally introduce the central notion of the article.

Definition 13.1.1. *A function $\rho : \mathbb{N} \longrightarrow \mathbb{N}$ is said to be a rank function, if it is weakly decreasing and satisfies the convexity condition*

$$\forall\, k \in \mathbb{N} : \; 2\rho(k+1) \leqslant \rho(k) + \rho(k+2).$$

Notice that if $\rho : \mathbb{N} \longrightarrow \mathbb{N}$ is a rank function, then there exists $k_0 \in \mathbb{N}$ such that

$$\begin{cases} \rho(k) = \rho(k_0), & \text{if } k \geqslant k_0, \\ \rho(k) > \rho(k+1), & \text{otherwise.} \end{cases}$$

Moreover, $k_0 \leqslant \rho(0)$. The integer $\rho(k_0)$ will be referred to as the *stable value* of ρ.

The rank functions occur naturally in matrix theory. They have also some applications in algebraic geometry and geometric invariant theory. This article provides a review of more or less known results concerning various applications of rank functions. It is partly based on the talk delivered by the author at the conference "Kangro-100" in Tartu.

13.2 The Semiring of Rank Functions

The main purpose of this section is to provide a quite interesting example of a partially ordered semiring. We begin with a few definitions.

Definition 13.2.1. *Let* $+ : S \times S \ni (a, b) \mapsto a + b \in S$ *and* $\cdot : S \times S \ni (a, b) \mapsto a \cdot b \equiv ab \in S$ *be binary operations on a nonempty set* S*. The triple* $(S, +, \cdot)$ *is said to be a semiring, if the following conditions are satisfied:*

- *addition* $+$ *is associative and commutative,*
- *addition has a neutral element (the neutral element is unique; it is denoted by* 0*),*
- *multiplication* $\cdot$ *is associative,*
- *multiplication is left and right distributive over addition,*
- $0 \cdot a = a \cdot 0 = 0$ *for any* $a \in S$.

Definition 13.2.2. *A nonempty set* $I \subseteq S$ *is said to be an ideal of a semiring* $(S, +, \cdot)$*, if* $a + b, ac, ca \in I$ *for any* $a, b \in I$ *and any* $c \in S$.

Definition 13.2.3. *Let* $(S, +, \cdot)$ *be a semiring and let* $\preceq$ *be a partial order on the set* S*. The quadruple* $(S, +, \cdot, \preceq)$ *is said to be a partially ordered semiring (po-semiring for short), if*

$$\forall\, a, b, c \in S : \begin{cases} a \preceq b \;\Rightarrow\; a + c \preceq b + c, \\ (a \preceq b,\ 0 \preceq c) \;\Rightarrow\; ac \preceq bc. \end{cases}$$

A po-semiring $(S, +, \cdot, \preceq)$ *is said to be positive, if* $0 \preceq a$ *for any* $a \in S$.

Now, consider addition $\cup : 2^X \times 2^X \ni (A, B) \longmapsto A \cup B \in 2^X$, multiplication $\cap : 2^X \times 2^X \ni (A, B) \longmapsto A \cap B \in 2^X$, and the inclusion relation $\subseteq$ on the power set 2^X of a set X. It is easy to see that $(2^X, \cup, \cap, \subseteq)$ is a *commutative* positive po-semiring *with identity* (the identity element coincides with X). Notice also that the set of non-negative integers equipped with the usual addition, the usual multiplication, and the natural order is a *totally ordered* commutative semiring with identity. For more information about semirings we refer to [2].

Let us turn back to rank functions.

Proposition 13.2.4. *If* $\rho, \sigma : \mathbb{N} \longrightarrow \mathbb{N}$ *are rank functions, then the pointwise sum* $\rho + \sigma$ *and the pointwise product* $\rho\sigma$ *are rank functions too.*

Proof. Monotonicity of the sum and the product is obvious as well as the convexity of the sum. We will recall the standard proof of the convexity of the product. Pick an arbitrary $k \in \mathbb{N}$. Since ρ and σ are convex, we have

$$2\rho(k+1)\sigma(k+1) \leq \frac{1}{2}\big(\rho(k)+\rho(k+2)\big)\big(\sigma(k)+\sigma(k+2)\big).$$

One can easily verify that

$$\frac{1}{2}\big(\rho(k)+\rho(k+2)\big)\big(\sigma(k)+\sigma(k+2)\big) =$$

$$= \rho(k)\sigma(k)+\rho(k+2)\sigma(k+2)+\frac{1}{2}\big(\rho(k)-\rho(k+2)\big)\big(\sigma(k+2)-\sigma(k)\big).$$

Monotonicity of ρ and σ yields that $\big(\rho(k)-\rho(k+2)\big)\big(\sigma(k+2)-\sigma(k)\big) \leq 0$. Therefore, $2\rho(k+1)\sigma(k+1) \leq \rho(k)\sigma(k)+\rho(k+2)\sigma(k+2)$. □

The pointwise inequality relation $\leqslant$ defined by

$$\rho \leqslant \sigma \iff \forall\, k \in \mathbb{N}:\ \rho(k) \leqslant \sigma(k)$$

(with the usual inequality on the right hand side) is obviously a partial order on the set $\mathcal{R}$ of all rank functions.

The following two facts, although very easy to prove, form the heart of the section.

Theorem 13.2.5. *The set $\mathcal{R}$ equipped with the pointwise addition, the pointwise multiplication and the partial order defined above is a commutative positive po-semiring with identity. Moreover,*

$$\mathcal{R}_0 = \{\rho \in \mathcal{R} :\ \textit{the stable value of } \rho \textit{ equals } 0\}$$

is an ideal of this semiring.

Proposition 13.2.6. *Let $\mathcal{T}$ be a nonempty finite set of rank functions. Then*

$$\bigvee \mathcal{T} : \mathbb{N} \ni k \mapsto \max_{\rho \in \mathcal{T}} \rho(k) \in \mathbb{N}$$

is a rank function too.

Notice that the pointwise minimum of a nonempty set of rank functions is not necessarily a rank function.

We conclude the section with some natural consequences of Proposition 13.2.6.

Corollary 13.2.7. *Consider the poset $(\mathcal{R}, \leqslant)$, where $\leqslant$ is the partial order defined above, and a set $\mathcal{T} \subseteq \mathcal{R}$. The following properties hold true:*

(i) if $\mathcal{T}$ is bounded from above in $(\mathcal{R}, \leqslant)$, then $\mathcal{T}$ has the supremum in $(\mathcal{R}, \leqslant)$,

(ii) if $\mathcal{T} \neq \emptyset$, then $\mathcal{T}$ has the infimum in $(\mathcal{R}, \leqslant)$,

(iii) $(\mathcal{R}, \leqslant)$ is a lattice.

Proof. It is not difficult to see that $\mathcal{T}$ is bounded from above in $(\mathcal{R}, \leqslant)$ if and only if $\mathcal{T}$ is a finite set. Suppose therefore that $\mathcal{T}$ is bounded from above and nonempty. Then the rank function $\bigvee \mathcal{T}$ defined in Proposition 13.2.6 is evidently the supremum of $\mathcal{T}$. If $\mathcal{T} = \emptyset$, then the supremum of $\mathcal{T}$ coincides with the least element of $(\mathcal{R}, \leqslant)$, i.e., the zero function. Next, if $\mathcal{T} \neq \emptyset$, then the set $\mathcal{L} = \{\rho \in \mathcal{R} : \rho \leqslant \tau \text{ for all } \tau \in \mathcal{T}\}$ is nonempty and bounded from above, and hence $\bigvee \mathcal{L}$ is the infimum of $\mathcal{T}$. Property *(iii)* is now obvious. $\square$

13.3 Rank Functions and Matrices

For a matrix $A \in \mathcal{M}_n(\mathbb{F})$, we define $\mathrm{r}_A : \mathbb{N} \longrightarrow \mathbb{N}$ by $\mathrm{r}_A(k) = \mathrm{rank}(A^k)$. Since, by definition, $A^0 = I_n$, we have $\mathrm{r}_A(0) = n$.

Let us recall that the *direct sum* of matrices $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{F})$ and $B \in \mathcal{M}_m(\mathbb{F})$ is defined to be the block matrix

$$A \oplus B = \begin{bmatrix} A & O_{n\times m} \\ O_{m\times n} & B \end{bmatrix} \in \mathcal{M}_{(n+m)\times(n+m)}(\mathbb{F}).$$

In turn, the *Kronecker* (or *tensor*) *product* of these two matrices is defined to be the block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \dots & a_{nn}B \end{bmatrix} \in \mathcal{M}_{nm\times nm}(\mathbb{F}).$$

It is noteworthy that $\mathrm{rank}(A \otimes B) = \mathrm{rank}(A)\mathrm{rank}(B)$ and

$$\forall\, k \in \mathbb{N} : \ (A \otimes B)^k = A^k \otimes B^k.$$

More information about the Kronecker product can be found in [6].

The following properties are now obvious.

Proposition 13.3.1. *Let $A \in \mathcal{M}_n(\mathbb{F})$, $B \in \mathcal{M}_m(\mathbb{F})$, $U \in \mathcal{GL}_n(\mathbb{F})$, and $\lambda \in \mathbb{F} \setminus \{0\}$. Then*

(i) r_A is a weakly decreasing function,

(ii) $\mathrm{r}_{\lambda A} = \mathrm{r}_A = \mathrm{r}_{U^{-1}AU}$,

(iii) $\mathrm{r}_{A\oplus B} = \mathrm{r}_A + \mathrm{r}_B$ *(pointwise sum),*

(iv) $\mathrm{r}_{A\otimes B} = \mathrm{r}_A\mathrm{r}_B$ *(pointwise product).*

We will also need the core-nilpotent decomposition (cf. [7]).

Theorem 13.3.2. *If $A \in \mathcal{M}_n(\mathbb{F})$, then there exist a matrix $U \in \mathcal{GL}_n(\mathbb{F})$, an integer $s \in \{0,\dots,n\}$, a matrix $C \in \mathcal{GL}_s(\mathbb{F})$, and a nilpotent matrix $\widetilde{A} \in \mathcal{M}_{n-s}(\mathbb{F})$ such that $U^{-1}AU = \widetilde{A} \oplus C$.*

Finally, let us recall some useful formulas related to the Jordan canonical form of a nilpotent matrix.

Proposition 13.3.3. *For a positive integer j we define $\ell_j(A)$ to be the number of blocks of size j contained in the Jordan canonical form of a nilpotent matrix $A \in \mathcal{M}_n(\mathbb{F})$. The following equalities hold true:*

(i) $\forall\, k \in \mathbb{N} :\ r_A(k) = \displaystyle\sum_{j=k+1}^{\infty} (j-k)\ell_j(A),$

(ii) $\ell_j(A) = r_A(j-1) + r_A(j+1) - 2r_A(j).$

We are in a position to state and prove the main theorem of the section (cf. [9]).

Theorem 13.3.4. *For a function $\rho : \mathbb{N} \longrightarrow \mathbb{N}$, the following conditions are equivalent:*

(1) *ρ is a rank function,*

(2) $\exists\, A \in \mathcal{M}_{\rho(0)}(\mathbb{F}) :\ \rho = r_A.$

Moreover, if ρ is a rank function with stable value 0, then the matrix A is unique up to conjugation (i.e., the set $\{A \in \mathcal{M}_{\rho(0)}(\mathbb{F}) :\ r_A = \rho\}$ is a single conjugacy class).

Proof. Suppose that condition *(2)* is satisfied. Then obviously $\rho = \mathrm{r}_A$ is weakly decreasing. Next, by Theorem 13.3.2,

$$\exists\, U \in \mathcal{GL}_{\rho(0)}(\mathbb{F})\, \exists\, s \in \{0,\dots,\rho(0)\} :\ U^{-1}AU = \widetilde{A} \oplus C,$$

where $C \in \mathcal{GL}_s(\mathbb{F})$ and $\widetilde{A} \in \mathcal{M}_{\rho(0)-s}(\mathbb{F})$ is a nilpotent matrix. Since

$$\mathrm{r}_A = \mathrm{r}_{U^{-1}AU} = \mathrm{r}_{\widetilde{A}\oplus C} = \mathrm{r}_{\widetilde{A}} + \mathrm{r}_C$$

and r_C is a constant function with value s, for any $k \in \mathbb{N}$ we have

$$\rho(k) + \rho(k+2) - 2\rho(k+1) = \mathrm{r}_{\widetilde{A}}(k) + \mathrm{r}_{\widetilde{A}}(k+2) - 2\mathrm{r}_{\widetilde{A}}(k+1) = \ell_{k+1}(\widetilde{A}) \geq 0.$$

Therefore, ρ is convex. Condition *(1)* follows.

Suppose that *(1)* is satisfied. Then there exists a non-negative integer $k_0 \leqslant \rho(0)$ with the property that

$$\forall\, k \in \mathbb{N} : \begin{cases} \rho(k) = \rho(k_0), & \text{if } k \geqslant k_0, \\ \rho(k) > \rho(k+1), & \text{otherwise.} \end{cases}$$

Moreover, $\rho(j-1) + \rho(j+1) - 2\rho(j) \geqslant 0$ for any positive integer j. Observe that

$$\sum_{j=1}^{\infty} j\big(\rho(j-1)+\rho(j+1)-2\rho(j)\big) = \sum_{k=0}^{k_0-1}(k+1)\rho(k) + \sum_{k=2}^{k_0+1}(k-1)\rho(k) - 2\sum_{k=1}^{k_0} k\rho(k) =$$

$$= \rho(0) + 2\rho(1) + (k_0-1)\rho(k_0) + k_0\rho(k_0+1) - 2\rho(1) - 2k_0\rho(k_0) =$$

$$= \rho(0) - \rho(k_0).$$

Consequently, there exists a nilpotent matrix $B \in \mathcal{M}_t(\mathbb{F})$, where $t = \rho(0) - \rho(k_0)$, such that $\ell_j(B) = \rho(j-1) + \rho(j+1) - 2\rho(j)$ for any positive $j \in \mathbb{N}$. Pick some $V \in \mathcal{GL}_{\rho(k_0)}(\mathbb{F})$ and define $A = B \oplus V$. Since B is nilpotent and $t \leqslant \rho(0)$, we have $\mathrm{r}_A(k) = \mathrm{r}_B(k) + \rho(k_0) = \rho(k_0) = \rho(k)$ for all integers $k \geqslant \rho(0)$. The equalities

$$\forall\, j \in \{1, \ldots, \rho(0)\} : \rho(j-1) + \rho(j+1) - 2\rho(j) = \ell_j(B) =$$

$$= \mathrm{r}_B(j-1) + \mathrm{r}_B(j+1) - 2\mathrm{r}_B(j) = \mathrm{r}_A(j-1) + \mathrm{r}_A(j+1) - 2\mathrm{r}_A(j)$$

therefore yield $\rho(\rho(0)-1) = \mathrm{r}_A(\rho(0)-1), \ldots, \rho(1) = \mathrm{r}_A(1)$. Condition *(2)* follows.

The "moreover" part is a direct consequence of Proposition 13.3.3 *(ii)*. □

Example 13.3.5. *Consider the rank function $\rho : \mathbb{N} \longrightarrow \mathbb{N}$ defined by*

$$\begin{cases} \rho(0) = 9, \\ \rho(1) = 6, \\ \rho(2) = 4, \\ \rho(k) = 2, \text{ if } k \geqslant 3. \end{cases}$$

The stable value of ρ is equal to 2. Moreover,

$$\rho(0) + \rho(2) - 2\rho(1) = 1, \ \rho(1) + \rho(3) - 2\rho(2) = 0,$$

$$\rho(2) + \rho(4) - 2\rho(3) = 2.$$

It follows that $\rho = r_A$ for

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus [0] \oplus V$$

with any $V \in \mathcal{GL}_2(\mathbb{F})$.

13.4 The Semiring of Conjugacy Classes

Let us define $\mathfrak{M}_n(\mathbb{F}) = \{\mathcal{O}(A) : A \in \mathcal{M}_n(\mathbb{F})\}$ and $\mathfrak{M}(\mathbb{F}) = \bigcup_{n=0}^{\infty} \mathfrak{M}_n(\mathbb{F})$. If $A \in \mathcal{M}_n(\mathbb{F})$, $B \in \mathcal{M}_m(\mathbb{F})$, $U \in \mathcal{GL}_n(\mathbb{F})$, $V \in \mathcal{GL}_m(\mathbb{F})$, and $\bullet \in \{\oplus, \otimes\}$, then $U \bullet V$ is a nonsingular matrix and

$$U^{-1}AU \bullet V^{-1}BV = (U \bullet V)^{-1}(A \bullet B)(U \bullet V).$$

Consequently, $\widetilde{\bullet} : \mathfrak{M}(\mathbb{F}) \times \mathfrak{M}(\mathbb{F}) \ni (\mathcal{O}(A), \mathcal{O}(B)) \longmapsto \mathcal{O}(A \bullet B) \in \mathfrak{M}(\mathbb{F})$ is a well defined binary operation on the set $\mathfrak{M}(\mathbb{F})$. Notice that $\{0\} = \mathcal{M}_0(\mathbb{F})$ and $\{1\} \in \mathfrak{M}_1(\mathbb{F})$ are the neutral elements under $\widetilde{\oplus}$ and $\widetilde{\otimes}$ respectively.

A map between two semirings is said to be a *semiring homomorphism*, if it is additive and multiplicative.

Theorem 13.4.1. *The triple* $(\mathfrak{M}(\mathbb{F}), \widetilde{\oplus}, \widetilde{\otimes})$ *is a commutative semiring with identity. Moreover,*

(i) $\mathfrak{M}_0(\mathbb{F}) = \{\mathcal{O}(A) \in \mathfrak{M}(\mathbb{F}) : A \text{ is nilpotent}\}$ *is an ideal of this semiring,*

(ii) $\Phi : \mathfrak{M}(\mathbb{F}) \ni \mathcal{O}(A) \longmapsto r_A \in \mathcal{R}$ *is a well defined semiring epimorphism,*

(iii) $\Phi|_{\mathfrak{M}_0(\mathbb{F})} : \mathfrak{M}_0(\mathbb{F}) \longrightarrow \mathcal{R}_0$ *is a semiring isomorphism.*

Proof. Let $A \in \mathcal{M}_n(\mathbb{F})$, $B \in \mathcal{M}_m(\mathbb{F})$, $C \in \mathcal{M}_p(\mathbb{F})$, and $\bullet \in \{\oplus, \otimes\}$. Obviously, $(A \bullet B) \bullet C = A \bullet (B \bullet C)$. The associativity of $\widetilde{\oplus}$ and $\widetilde{\otimes}$ follows. Next, if

$$U = \begin{bmatrix} O_{n\times m} & I_n \\ I_m & O_{m\times n} \end{bmatrix},$$

then $U^{-1}(A \oplus B)U = B \oplus A$. This yields the commutativity of $\widetilde{\oplus}$. It is quite easy to see that $B \otimes A = P^{-1}(A \otimes B)P$ for some $nm \times nm$ permutation matrix P. Multiplication $\widetilde{\otimes}$ is therefore commutative, too. Similarly, there exists an $n(m+p) \times n(m+p)$ permutation matrix Q such that

$$A \otimes (B \oplus C) = Q^{-1}(A \otimes B \oplus A \otimes C)Q.$$

The distributive property of $\widetilde{\otimes}$ over $\widetilde{\oplus}$ follows. Recall that for any $k \in \mathbb{N}$ we have $(A \bullet B)^k = A^k \bullet B^k$. This implies property *(i)*. Properties *(ii)* and *(iii)* are straightforward consequences of Proposition 13.3.1 and Theorem 13.3.4. □

To avoid problems with algebraic geometry, from now on we will work over the field $\mathbb{C}$ of complex numbers. Bars will denote closures in the Euclidean topology on $\mathcal{M}_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$.

Let us recall a classical result due to Gerstenhaber (see [2] and [5]).

Theorem 13.4.2. *For nilpotent matrices $A, B \in \mathcal{M}_n(\mathbb{C})$ the following conditions are equivalent:*

(1) $A \in \overline{\mathcal{O}(B)}$,

(2) $r_A \leqslant r_B$ *(the pointwise inequality).*

For $A \in \mathcal{M}_n(\mathbb{C})$ and $B \in \mathcal{M}_m(\mathbb{C})$ we define

$$\mathcal{O}(A) \preceq \mathcal{O}(B) \iff \begin{cases} n \leq m, \\ A \oplus O_{(m-n)\times(m-n)} \in \overline{\mathcal{O}(B)}. \end{cases}$$

Notice that $\overline{\mathcal{O}(B)}$ is a $\mathcal{GL}_m(\mathbb{C})$-invariant set. Therefore, $\preceq$ is a well defined binary relation on $\mathfrak{M}(\mathbb{C})$ (known as the dominance relation).

Theorem 13.4.3. *The following statements are true:*

(i) $\{\mathcal{O}(A) \in \mathfrak{M}(\mathbb{C}) : \{0\} \preceq \mathcal{O}(A)\} = \mathfrak{M}_0(\mathbb{C})$,

(ii) $(\mathfrak{M}(\mathbb{C}), \widetilde{\oplus}, \widetilde{\otimes}, \preceq)$ *is a po-semiring,*

(iii) Φ *defined in Theorem 13.4.1 is an order-preserving map from* $(\mathfrak{M}(\mathbb{C}), \preceq)$ *to* $(\mathcal{R}, \leqslant)$,

(iv) $\Phi|_{\mathfrak{M}_0(\mathbb{C})} : \mathfrak{M}_0(\mathbb{C}) \longrightarrow \mathcal{R}_0$ *is an order isomorphism.*

Proof. Equality *(i)* is a direct consequence of the well known fact that a matrix $A \in \mathcal{M}_n(\mathbb{C})$ is nilpotent if and only if $O_{n\times n} \in \overline{\mathcal{O}(A)}$. Let us turn to property *(ii)*. Relation $\preceq$ is obviously reflexive. Next, it is also well known that if $\overline{\mathcal{O}(A_1)} = \overline{\mathcal{O}(A_2)}$ for some $A_1, A_2 \in \mathcal{M}_n(\mathbb{C})$, then $\mathcal{O}(A_1) = \mathcal{O}(A_2)$. This implies the antisymmetry of $\preceq$.

Pick matrices $A \in \mathcal{M}_n(\mathbb{C})$, $B \in \mathcal{M}_m(\mathbb{C})$, and $C \in \mathcal{M}_p(\mathbb{C})$. Suppose that $\mathcal{O}(A) \preceq \mathcal{O}(B)$. Then $n \leqslant m$ and $A \oplus O_{(m-n)\times(m-n)} \in \overline{\mathcal{O}(B)}$. By the continuity of the map $\Gamma : \mathcal{M}_m(\mathbb{C}) \ni T \mapsto T \otimes C \in \mathcal{M}_{mp}(\mathbb{C})$, we have therefore

$$(A \oplus O_{(m-n)\times(m-n)}) \otimes C \in \overline{\Gamma(\mathcal{O}(B))} \subseteq \overline{\mathcal{O}(B \otimes C)}.$$

Since $\mathcal{O}((A \oplus O_{(m-n)\times(m-n)}) \otimes C) = \mathcal{O}((A \otimes C) \oplus O_{(m-n)p\times(m-n)p})$, we finally obtain $(A \otimes C) \oplus O_{(m-n)p\times(m-n)p} \in \overline{\mathcal{O}(B \otimes C)}$. Thus $\mathcal{O}(A)\widetilde{\otimes}\mathcal{O}(C) \preceq \mathcal{O}(B)\widetilde{\otimes}\mathcal{O}(C)$. The compatibility of $\preceq$ with multiplication follows. The compatibility with addition can be proved analogously. Now, suppose that $\mathcal{O}(A) \preceq \mathcal{O}(B)$ and $\mathcal{O}(B) \preceq \mathcal{O}(C)$. Then $n \leqslant m$, $m \leqslant p$, $A \oplus O_{(m-n)\times(m-n)} \in \overline{\mathcal{O}(B)}$, and $B \oplus O_{(p-m)\times(p-m)} \in \overline{\mathcal{O}(C)}$. Since

$$\Delta : \mathcal{M}_m(\mathbb{C}) \ni T \longmapsto T \oplus O_{(p-m)\times(p-m)} \in \mathcal{M}_p(\mathbb{C})$$

is a continuous map, we have

$$A \oplus O_{(p-n)\times(p-n)} = \Delta(A \oplus O_{(m-n)\times(m-n)}) \in \overline{\Delta(\mathcal{O}(B))} \subseteq$$

$$\subseteq \overline{\mathcal{O}(B \oplus O_{(p-m)\times(p-m)})} \subseteq \overline{\mathcal{O}(C)},$$

and hence $\mathcal{O}(A) \preceq \mathcal{O}(C)$. Consequently, relation $\preceq$ is transitive.

Let us turn to *(iii)*. Suppose that $\mathcal{O}(A) \preceq \mathcal{O}(B)$ for some $A \in \mathcal{M}_n(\mathbb{C})$ and $B \in \mathcal{M}_m(\mathbb{C})$. Pick an arbitrary positive integer k. We need to show that $\mathrm{r}_A(k) \leqslant \mathrm{r}_B(k)$. If $T \in \mathcal{O}(B)$, then $\mathrm{r}_T(k) = \mathrm{r}_B(k)$, and hence each minor of the matrix T^k whose size exceeds $\mathrm{r}_B(k)$ is equal to 0. By the continuity of minors as functions from $\mathcal{M}_m(\mathbb{C})$ to $\mathbb{C}$, it follows that if $T \in \overline{\mathcal{O}(B)}$, then each minor of T^k whose size exceeds $\mathrm{r}_B(k)$ is also equal to 0. Therefore, since $A \oplus O_{(m-n)\times(m-n)} \in \overline{\mathcal{O}(B)}$, we obtain

$$\mathrm{r}_A(k) = \mathrm{rank}((A \oplus O_{(m-n)\times(m-n)})^k) \leqslant \mathrm{r}_B(k).$$

Property *(iv)* follows immediately from Theorems 13.4.1 and 13.4.2. □

13.5 Rank Varieties

Rank varieties were introduced by Eisenbud and Saltman in [1].

Definition 13.5.1. *The rank variety associated to a rank function ρ is defined to be the set $\mathcal{X}_\rho = \{A \in \mathcal{M}_{\rho(0)}(\mathbb{C}) : r_A \leqslant \rho\}$.*

By Theorem 13.3.4, if $\mathcal{X}_\rho = \mathcal{X}_\sigma$, then the rank functions ρ and σ coincide too. Theorem 13.4.2 in turn says that if $A \in \mathcal{M}_n(\mathbb{C})$ is nilpotent, then $\overline{\mathcal{O}(A)}$ is the rank variety associated to r_A.

Fundamental properties of the rank varieties can be summarized as follows (cf. [1]).

Theorem 13.5.2. *For any rank function ρ, the variety $\mathcal{X}_\rho$ is a $\mathcal{GL}_{\rho(0)}(\mathbb{C})$-invariant normal algebraic subset of $\mathcal{M}_{\rho(0)}(\mathbb{C})$. Moreover,*

$$\dim \mathcal{X}_\rho = \big(\rho(0)\big)^2 - \sum_{k=0}^{\infty} \big(\rho(k) - \rho(k+1)\big)^2.$$

For the proof refer also to [8].

Notice that the *determinantal variety* $\mathcal{H}_n^r = \{A \in \mathcal{M}_n(\mathbb{C}) : \mathrm{rank}(A) \leqslant r\}$, where $r \in \{0, \ldots, n\}$, coincides with the rank variety associated to the function ρ given by

$$\begin{cases} \rho(0) = n, \\ \rho(k) = r, \text{ if } k \geqslant 1. \end{cases}$$

This fact can be easily generalized.

Proposition 13.5.3. *Let E be a nonempty set of positive integers. Consider an arbitrary function $\varphi : E \longrightarrow \mathbb{N}$. Then there exists a unique rank function ρ such that*

$$\{A \in \mathcal{M}_n(\mathbb{C}) :\ r_A(k) \leqslant \varphi(k) \text{ for all } k \in E\} = \mathcal{X}_\rho.$$

Proof. The set $\mathcal{T} = \{\tau \in \mathcal{R} :\ \tau(0) = n,\ \tau(k) \leqslant \varphi(k) \text{ for all } k \in E\}$ is nonempty and finite. To complete the proof, it is therefore enough to define $\rho = \bigvee \mathcal{T}$ (see Proposition 13.2.6). □

The above ρ is obviously the greatest element of the set $\mathcal{T}$.

Example 13.5.4. *Suppose that $n \geqslant 7$. Then*

$$\mathcal{Z} = \{A \in \mathcal{M}_n(\mathbb{C}) :\ \operatorname{rank}(A^2) \leqslant n-3,\ \operatorname{rank}(A^3) \leqslant n-7\}$$

coincides with the rank variety associated to the greatest element ρ of the set $\{\tau \in \mathcal{R} :\ \tau(0) = n,\ \tau(2) \leqslant n-3,\ \tau(3) \leqslant n-7\}$. It is not difficult to see that

$$\begin{cases} \rho(1) = n-3, \\ \rho(2) = n-5, \\ \rho(k) = n-7, \text{ if } k \geqslant 3. \end{cases}$$

Therefore, $\mathcal{Z}$ is a normal (hence irreducible) algebraic set and

$$\dim \mathcal{Z} = n^2 - \sum_{k=0}^{\infty} \big(\rho(k) - \rho(k+1)\big)^2 = n^2 - 3^2 - 2 \cdot 2^2 = n^2 - 17.$$

It seems also noteworthy that $\max\limits_{A \in \mathcal{Z}} \operatorname{rank}(A^2) = n-5$.

Finally, notice that the full nilpotent cone

$$\mathcal{N}_n = \{A \in \mathcal{M}_n(\mathbb{C}) :\ A \text{ is nilpotent}\} = \{A \in \mathcal{M}_n(\mathbb{C}) :\ \mathrm{r}_A(n) = 0\} = \overline{\mathcal{O}(J_n)},$$

where J_n stands for the nilpotent Jordan block of size n, is a rank variety.

13.6 Rank Functions and Irreducibility

Let us return for a moment to an arbitrary base field $\mathbb{F}$. For a set $\mathcal{E} \subseteq \mathcal{M}_n(\mathbb{F})$, we define $\mathcal{R}(\mathcal{E}) = \{\mathrm{r}_A :\ A \in \mathcal{E}\}$. Observe that $\mathcal{R}(\mathcal{E})$ is a finite subset of $(\mathcal{R}, \leqslant)$. The following result comes from [10].

Theorem 13.6.1. *The number of irreducible components of an algebraic set $\mathcal{E} \subseteq \mathcal{M}_n(\mathbb{F})$ is greater than or equal to the number of maximal elements of $\mathcal{R}(\mathcal{E})$. In particular, if $\mathcal{E}$ is irreducible, then $\mathcal{R}(\mathcal{E})$ has a greatest element.*

Proof. Suppose that $\mathcal{E} \neq \emptyset$. Let $\mathcal{K}$ be the set of all maximal elements of $\mathcal{R}(\mathcal{E})$. For any $\mu \in \mathcal{K}$ we define $\mathcal{E}_\mu = \{A \in \mathcal{E} : \mathrm{r}_A \leqslant \mu\}$. Each $\mathcal{E}_\mu$ is an algebraic set. Moreover, $\bigcup_{\mu \in \mathcal{K}} \mathcal{E}_\mu = \mathcal{E}$ and none of the sets $\mathcal{E}_\mu$ is contained in the union of the others. Consequently, if $\mu_0 \in \mathcal{K}$, then there exists an irreducible component $\mathcal{C}$ of the set $\mathcal{E}$ such that $\mathcal{C} \subseteq \mathcal{E}_{\mu_0}$ and $\mathcal{C}$ is not contained in $\mathcal{E}_\mu$ with any μ different from μ_0. This yields an injection from $\mathcal{K}$ to the family of irreducible components of $\mathcal{E}$. The assertion follows. □

Notice that the set

$$\mathcal{E} = \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

is reducible, although $\mathcal{R}(\mathcal{E})$ is a singleton.

We conclude the article with two remarks on the geometry of $\mathcal{GL}_n(\mathbb{C})$-invariant sets of nilpotent matrices (cf. [10]).

Proposition 13.6.2. *Let $\mathcal{E} \subseteq \mathcal{M}_n(\mathbb{C})$ be a $\mathcal{GL}_n(\mathbb{C})$-invariant algebraic set of nilpotent matrices. Then the number of maximal elements of $\mathcal{R}(\mathcal{E})$ is equal to the number of irreducible components of $\mathcal{E}$. In particular, $\mathcal{E}$ is irreducible if and only if $\mathcal{R}(\mathcal{E})$ has a greatest element.*

Proof. Consider the sets $\mathcal{K}$ and $\mathcal{E}_\mu$ defined in the proof of Theorem 13.6.1. Since $\mathcal{E}$ consists of nilpotent matrices, it follows from Theorem 13.4.2 and the invariance of $\mathcal{E}$ that for an arbitrary $\mu = \mathrm{r}_A \in \mathcal{K}$ we have

$$\mathcal{E}_\mu = \mathcal{E} \cap \overline{\mathcal{O}(A)} = \overline{\mathcal{O}(A)}.$$

Therefore, every set $\mathcal{E}_\mu$ is irreducible, and hence $\{\mathcal{E}_\mu : \mu \in \mathcal{K}\}$ coincides with the family of irreducible components of $\mathcal{E}$. □

Corollary 13.6.3. *If $n \leqslant 5$ and $\mathcal{E} \subseteq \mathcal{M}_n(\mathbb{C})$ is a nonempty $\mathcal{GL}_n(\mathbb{C})$-invariant algebraic set of nilpotent matrices, then $\mathcal{E}$ is irreducible.*

Proof. It is easy to see that if $n \leqslant 5$, then the pointwise inequality is a total order on the set $\{\rho \in \mathcal{R} : \rho(0) = n, \text{the stable value of } \rho \text{ is equal to } 0\}$. □

Observe also that if $n \geqslant 6$, then in the above set of rank functions there exist elements which are not comparable.

References

[1] D. Eisenbud and D. Saltman. “Rank varieties of matrices”. In: *Commutative algebra. Proc. Microprogram, Berkeley/CA (USA) 1989, Math. Sci. Res. Inst. Publ.* ***15***. 1989, pp. 173–212.

[2] J. S. Golan. *Semirings and their applications.* Dordrecht: Kluwer Academic Publishers, 1999.

[3] B. Grell. *Wstęp do matematyki: zbiory, struktury, modele.* Wydawnictwo Uniwersytetu Jagiellońskiego, 2006.

[4] R. Hartshorne. *Algebraic geometry.* Springer, 1983.

[5] M. Hazewinkel and C. F. Martin. "Representations of the symmetric group, the specialization order, systems and Grassmann manifolds". In: *Enseign. Math., II. Sér.* 29 (1983), pp. 53–87.

[6] A. J. Laub. *Matrix analysis for scientists and engineers.* Philadelphia, PA: SIAM, 2005.

[7] C. D. Meyer. *Matrix analysis and applied linear algebra.* Philadelphia, PA: SIAM, 2000.

[8] M. Skrzyński. "On the dimension of certain $\mathcal{GL}_n$-invariant sets of matrices". In: *Univ. Iagell. Acta Math.* 48 (2010), pp. 137–147.

[9] M. Skrzyński. "Rank functions of matrices". In: *Univ. Iagell. Acta Math.* 37 (1999), pp. 139–149.

[10] M. Skrzyński. "Remarks on applications of rank functions to algebraic sets of matrices". In: *Demonstr. Math.* 32, No. 2 (1999), pp. 263–271.

Chapter 14

Conformal Killing forms and conformal Killing tensors in Riemannian geometry

GRZEGORZ ZBOROWSKI

Grzegorz Zborowski
Cracow University of Technology
Institute of Mathematics
ul. Warszawska 24
31-155 Kraków, Poland
e-mail: gzborowski@pk.edu.pl

Abstract

In this note we want to give some basic results concerning the so-called *conformal Killing tensors* and their connections to another geometric object – *conformal Killing forms*. We are mainly interested in the situation when the Ricci tensor of a Riemannian manifold is a conformal Killing tensor and we want to give a sketch of construction of examples of such manifolds.

Contents

14.1 Conformal Killing tensors

In this section we will give a definition of a conformal Killing tensor and some basic properties of this object. Throughout this chapter let (M, g) be a Riemannian manifold, i.e. a differential manifold M with a symmetric, positive and non-degenerate tensor field g called a Riemannian metric. We denote by ∇ the Levi-Civita connection of g.

Definition 14.1.1. *A symmetric tensor field K on M is called* conformal Killing *if there exists a tensor field P such that*

$$\nabla_X K(X, X) = P(X)g(X, X) \tag{14.1}$$

is satisfied for any vector field X. Moreover, K is called a Killing tensor field *if P is zero in the above identity.*

One can immediately observe that this condition generalizes a so-called *parallel tensor field*, i.e. one that satisfies

$$\nabla K = 0$$

giving immediate trivial examples like the metric tensor g. If a conformal Killing tensor field is not parallel we call it a *strict conformal Killing tensor field.*

It is easy to show that the above condition (14.1) is equivalent to the following cyclic sum condition:

$$\mathcal{C}_{X,Y,Z} \nabla_X K(Y, Z) = \mathcal{C}_{X,Y,Z} P(X)g(Y, Z), \tag{14.2}$$

for any vector fields X, Y and Z, where $\mathcal{C}_{X,Y,Z}$ denotes the cyclic sum over X, Y and Z. A tensor field satisfying (14.2) with $P = 0$ is called *cyclic parallel.*

Killing tensors are important in physics due to the following property.

Proposition 14.1.1. *A tensor field K is a Killing tensor on (M, g) if and only if the function $K(\gamma'(t), \gamma'(t))$ is constant for every geodesic γ.*

It is easily shown, that the 1-form P from Definition 14.1.1 is eqaul to

$$P(X) = \frac{1}{n+2}\left(2\text{div}K(X) + d\text{tr}_g K(X)\right) \tag{14.3}$$

with $\text{div}K(X) = \sum_{i=1}^{n} \nabla_{E_i} K(E_i, X)$ being the divergence of a tensor field K, $\{E_i\}_{i=1}^{n}$ a local orthonormal frame of TM with respect to g, n the dimension of the manifold M, tr_g the trace with respect to g and d denoting the operator of exterior differentiation.

Next we give some properties of conformal Killing tensor fields, after [8] and [5] characterizing those tensor fields due to their eigenvalues. For the remaining part of this section let K be a conformal Killing tensor field with eigenfunctions $\lambda_1, \ldots, \lambda_m$ corresponding to eigendistributions $\mathcal{D}_1, \ldots, \mathcal{D}_m$ giving an orthogonal decomposition of TM with respect to g.

The first step in characterizing conformal Killing tensor fields is

Proposition 14.1.2. *Let K be a conformal Killing tensor field as above. Then for sections X, Y and Z of distributions $\mathcal{D}_i$, $\mathcal{D}_j$ and $\mathcal{D}_k$ respectively, where $i \neq j \neq k$ we have*

$$\mathcal{C}_{X,Y,Z} K(\nabla_X Y + \nabla_Y X, Z) = 0. \tag{14.4}$$

Next, for the 1-form P from the definition 14.1.1 we have

Proposition 14.1.3. *For a conformal Killing tensor field as above, we have*

$$\mathcal{D}_i \subset \ker(P - d\lambda_i). \tag{14.5}$$

To state the last property we need to define a special kind of distribution of TM. We say that a regular distribution $\mathcal{D}$ of TM is *totally umbilical* if for any sections X and Y of $\mathcal{D}$ there exists a vector field ξ such that

$$(\nabla_X Y + \nabla_Y X)|_{\mathcal{D}^\perp} = 2g(X, Y)\xi,$$

where by $\mathcal{D}^\perp$ we mean an orthogonal complement of $\mathcal{D}$. The vector field ξ is then called a *mean curvature normal vector field* of $\mathcal{D}$.

The last property is the following.

Proposition 14.1.4. *Let K be a conformal Killing tensor field as above. For any eigenfunction λ of K corresponding to the eigendistribution $\mathcal{D}_\lambda$ we have that $\mathcal{D}_\lambda$ is totally umbilical. Furthermore, we have*

$$g(\xi_\lambda, Z) = -\frac{1}{2(\mu - \lambda)} \left(d\mu(Z) - d\lambda(Z) \right), \tag{14.6}$$

where Z is a section of some other distribution $\mathcal{D}_\mu$ with $\lambda \neq \mu$ and ξ_λ is a mean curvature vector field of $\mathcal{D}_\lambda$.

It turns out, as showed in [5], that those three properties characterize conformal Killing tensor fields completely.

Theorem 14.1.1. *Let K be a symmetric tensor field and let $TM = \bigoplus_{i=1}^m \mathcal{D}_i$ be an orthogonal decomposition of TM. Let K be defined as follows: $K(X, X) = \lambda_i g(X, X)$ for X a section of $\mathcal{D}_i$. Then K is a conformal Killing tensor field if and only if the conditions* (14.4) *and* (14.6) *are satisfied and there exists a* 1*-form P defined by $P(X) = d\lambda_i(X)$ for X a section of $\mathcal{D}_i$ satisfying* (14.5).

Using this theorem authors of [5] were able to locally characterize metrics admitting conformal Killing tensor fields with two eigenvalues corresponding to eigendistributions of dimension 1 and $n-1$, where n is the dimension of M.

In fact the case of two eigenvalues is the most well-known until now. In addition to the above mentioned local classification we have the following useful property ([8]).

Proposition 14.1.5. *Let K be a Killing tensor with a constant trace and exactly two eigenvalues. Then both eigenvalues are constants. Moreover K is parallel if and only if both eigendistributions are integrable.*

14.2 $\mathcal{AC}^{\perp}$- and $\mathcal{A}$-manifolds - introduction

Of special interest is the situation, when the Ricci tensor is a conformal Killing tensor field. The following definition is due to A. Gray [6].

Definition 14.2.1. *A Riemannian mnaifold (M, g) such that the Ricci tensor* Ric *is a conformal Killing tensor is called $\mathcal{AC}^{\perp}$-*manifold*. If the Ricci tensor is a Killing tensor then we call (M, g) an $\mathcal{A}$-*manifold*.*

Using (14.3) and the fact that $2\mathrm{divRic} = d\mathrm{scal}$, where scal denotes the scalar curvature of g, we have that every $\mathcal{AC}^{\perp}$-manifold satisfies

$$\nabla_X \mathrm{Ric}(X, X) = \frac{2}{n+2} d\mathrm{scal}(X) g(X, X).$$

A trivial example of an $\mathcal{A}$-manifold is an Einstein manifold, as the Ricci tensor is just a constant multiple of the metric tensor.

The problem of finding examples of non-homogeneous and compact $\mathcal{A}$-manifolds was stated in [3], paragraph 16.56, problem (i).

First example of a strict $\mathcal{A}$-manifold is due to Gray ([6]). This is a homogeneous metric on the sphere S^3. The first non-homogeneous example is due to W. Jelonek, [8]. It is a special metric on a S^1-principal bundle over a Kähler-Einstein manifold. Some other examples examples were constructed by Jelonek, including a K-contact and Sasakian $\mathcal{A}$-manifolds, [12], almost Kähler $\mathcal{A}$-manifolds on a twistor bundle, [7]. The author generalized the construction from [8] in [21]. The difference is in taking a principal r-dimensional torus bundle and a base which is a product of Kähler-Einstein manifolds. Furthermore, as a generalization of the K-contact examples the author obtained an example of an $\mathcal{A}$-manifold on a r-torus bundle over a product of almost Kähler $\mathcal{A}$-manifolds [20]. This example generalizes all constructions on S^1-bundles obtained before. In the last section we will give a sketch of construction of this example.

It is worth mentioning that Z. Tang and W. Yan have obtained in [19] new examples of $\mathcal{A}$-manifolds as focal sets of isoparametric hypersurfaces in spheres.

Interest in $\mathcal{AC}^{\perp}$-manifolds is also stated in [3], although the author just says that not much is known about them. Since then, those manifolds where extensively studied by Jelonek. Constructions of such manifolds can be found for example in [13], [9] and [11]. The first is the so-called *warped product metric* g on a product manifold $\mathbb{R} \times M_1 \times M_2$, where M_1 and M_2 are compact Einstein manifolds. The metric g is then defined as

$$g = dt \otimes dt + f g_1 + \frac{1}{f} g_2,$$

where dt is the length element on $\mathbb{R}$, g_1 and g_2 are Einstein metrics on M_1 and M_2 respectively and f is some positive, periodic function on $\mathbb{R}$. This manifold is of course not compact. One can obtain a compact example from it in the following way. Let A_1 and A_2 be isometries of M_1 and M_2 respectively and let T be the period of the mapping f. Then

$$A(t, x_1, x_2) = (t + T, A_1(x_1), A_2(x_2))$$

is an isometry of $\mathbb{R} \times M_1 \times M_2$ and the quotient manifold $\mathbb{R} \times M_1 \times M_2/\mathbb{Z}A$ is compact. Jelonek showed, that when M_1 and M_2 satisfy some assumptions concerning dimensions and Einstein constants, there exists a function f such that both $\mathbb{R} \times M_1 \times M_2$ and the quotient manifold are $\mathcal{AC}^{\perp}$-manifolds.

One can prove a following property of $\mathcal{A}$-manifolds due to Gray ([6]), situating the whole class somewhere between Einstein manifolds and constant scalar curvature manifolds.

Proposition 14.2.1. *Every $\mathcal{A}$-manifold has constant scalar curvature.*

As for non-existence results we have a classic one about $\mathcal{A}$-manifolds, due to Sekigawa and Vanhecke ([16]).

Theorem 14.2.1. *Every Kähler $\mathcal{A}$-manifold has parallel Ricci tensor.*

Moreover, as observed by Jelonek (unpublished) we can prove the following.

Proposition 14.2.2. *Let (M, g) be a conformally Kähler manifold, i.e. suppose that there exists a smooth function f on M such that $(M, e^f g)$ is Kähler. Moreover, suppose that (M, g) is an $\mathcal{A}$-manifold with two eigenvalues, such that one of the eigendistributions is one-dimensional. Then the Ricci tensor of g is parallel.*

14.3 Conformal Killing forms

Let $\wedge$ denote the wedge product of differential forms and let ι_X denote the interior product with a vector field X defined for a differential p-form α by

$$\iota_X \alpha(X_1, \ldots, X_{p-1}) = \alpha(X, X_1, \ldots, X_{p-1}),$$

for any vector fields $X_1, \ldots, X_{p-1}$. In the following we identify a vector field X with its dual one-form defined by $X(Y) = g(X, Y)$. Let d be the operator of exterior differentiation and denote by δ the co-differential. One can prove that d and δ are linked to the Levi-Civita connection ∇ by the following formulae

$$d = \sum_{i=1}^{n} E_i \wedge \nabla_{E_i}, \quad \delta = -\sum_{i=1}^{n} \iota_{E_i} \nabla_{E_i},$$

where $\{E_i\}_{i=1}^n$ is some local orthonormal frame on (M, g) and n is the dimension of M. Furthermore d and δ are dual with respect to the pointwise scalar product g on the space of differential p-forms induced by the Riemannian metric g of M in the sense that

$$g(d\alpha, \beta) = g(\alpha, \delta\beta),$$

where α is a differential $p-1$-form and β is a differential p-form.

Now we can state the following definition, following Semmelmann ([17]).

Definition 14.3.1. *Let α be a differential p-form on M. We call α a* conformal Killing form *or a* twistor form *if the following identity is satisfied:*

$$\nabla_X \alpha = \frac{1}{p+1} \iota_X d\alpha - \frac{1}{n-p+1} X \wedge \delta\alpha, \tag{14.7}$$

for any vector field X. Moreover, if α is in addition co-closed it is called a Killing form*.*

An extensive study of those differential forms (known also by the name of Killing-Yano tensors in physics) is given in the series of articles by U. Semmelmann and A. Moroianu, here we only cite [17] and [18], which will be of relevance to us.

A simple example of a conformal Killing form is given by the dual differential 1-form of a conformal Killing vector field, i.e. a vector field X satisfying

$$L_X g = fg,$$

where L denotes the Lie differential, and f is some smooth function on M. If $f \equiv 0$, then the field X satisfying $L_X g = 0$ is called a Killing field, and its metric dual gives an example of a Killing 1-form.

It is widely known that a symmetrized product of Killing fields is a Killing tensor field. By a symmetrized product of Killing fields $\xi_1, \dots, \xi_k$ we mean a tensor field $K_{\xi_1,\dots,\xi_k}$ defined by

$$K_{\xi_1,\dots,\xi_k}(X,Y) = \sum_{i,j=1}^{k} g(\xi_i, X) g(\xi_j, Y),$$

for any vector fields X and Y.

The following generalization of the above fact can be found in physics literature. As the author could not find any proof and since it is quite simple we give it below.

Proposition 14.3.1. *Let α and β be conformal Killing p-forms. Then the tensor field $K_{\alpha,\beta}$ defined by*

$$K_{\alpha,\beta}(X,Y) = g(\iota_X \alpha, \iota_Y \beta) + g(\iota_X \beta, \iota_Y \alpha)$$

is a conformal Killing tensor field. Here by g we mean the pointwise scalar product induced on differential p-forms by the Riemmanian metric g on M.

Proof. Let X be any vector field and α, β conformal Killing p-forms. We will check that $K_{\alpha,\beta}$ as defined above satisfies (14.1).

$$\begin{aligned} \nabla_X K_{\alpha,\beta}(X,X) &= 2X\left(g(\iota_X\alpha, \iota_X\beta)\right) - 2g(\iota_{\nabla_X X}\alpha, \iota_X\beta) \\ &\quad -2g(\iota_{\nabla_X X}\beta, \iota_X\alpha) \\ &= 2g(\nabla_X(\iota_X\alpha), \iota_X\beta) + 2g(\iota_X\alpha, \nabla_X(\iota_X\beta)) \\ &\quad -2g(\iota_{\nabla_X X}\alpha, \iota_X\beta) - 2g(\iota_{\nabla_X X}\beta, \iota_X\alpha) \\ &= 2g(\iota_X\nabla_X\alpha, \iota_X\beta) + 2g(\iota_X\alpha, \iota_X\nabla_X\beta). \end{aligned}$$

From the fact that α satisfies (14.7) we have

$$\begin{aligned} g(\iota_X\nabla_X\alpha, \iota_X\beta) &= \frac{1}{p+1} g(\iota_X(\iota_X d\alpha), \iota_X\beta) \\ &\quad - \frac{1}{n-p+1} g(\iota_X(X \wedge \delta\alpha), \iota_X\beta) \\ &= -\frac{1}{n-p+1}\left(g(X,X)g(\delta\alpha, \iota_X\beta) - g(X \wedge (\iota_X\delta\alpha), \iota_X\beta)\right) \\ &= -\frac{1}{n-p+1} g(X,X) g(\delta\alpha, \iota_X\beta). \end{aligned}$$

The same is valid for β with

$$g(\iota_X\nabla_X\beta, \iota_X\alpha) = -\frac{1}{n-p+1} g(X,X) g(\delta\beta, \iota_X\alpha).$$

Hence we have

$$\nabla_X K_{\alpha,\beta}(X,X) = -\frac{2}{n-p+1} g(X,X)\left(g(\delta\alpha, \iota_X\beta) + g(\delta\beta, \iota_X\alpha)\right).$$

□

Another link between conformal Killing forms and conformal Killing tensors is present when (M,g,J) is a Kähler manifold, i.e. J is the so-called *complex structure tensor* which is parallel with respect to the Levi-Civita connection of g and is compatible with g, meaning $g(JX,JY) = g(X,Y)$ for any vector fields X and Y. Additionaly on a Kähler manifold we can define the Kähler form ω by $\omega(X,Y) = g(JX,Y)$. This form is closed and co-closed, i.e. $d\omega = 0 = \delta\omega$. We have

Proposition 14.3.2. *Let α be a J-invariant ($\alpha(X,Y) = \alpha(JX,JY)$) conformal Killing 2-form. Then the symmetric tensor field K defined by $K(X,Y) = -\alpha(JX,Y)$ is a conformal Killing tensor field.*

Examples of conformal Killing forms on Kähler manifolds can be obtained using the concept of a Hamiltonian 2-form.

Definition 14.3.2. *Let α be a J-invariant differential two-form on a Kähler manifold (M,g,J). We call α a* Hamiltonian 2-form *if there exists a smooth function σ such that α satisfies*

$$\nabla_X\alpha = \frac{1}{2}\left(d\sigma \wedge JX - Jd\sigma \wedge X\right)$$

for any vector field X.

Hamiltonian 2-forms were studied in [1], where local classification of manifolds admitting such forms was obtained. As observed in this work and also in [18] we have the following.

Proposition 14.3.3. *Let (M,g,J) be a Kähler manifold admitting a Hamiltonian 2-form ψ. Then $\alpha = \psi - \frac{1}{2}g(\omega,\psi)\omega$ is a conformal Killing 2-form. Conversely, if α is a conformal Killing 2-form and the real dimension of M is greater than 4, then $\psi = \alpha - \frac{2}{n-4}g(\omega,\alpha)\omega$ is a Hamiltonian 2-form.*

Thanks to this theorem and examples of Hamiltonian 2-forms given in [1] and its sequels we obtain many examples of conformal Killing tensor fields on Kähler manifolds.

Of particular interest is the case of *weakly Bochner-flat*, i.e. Kähler manifolds with harmonic Bochner tensor (Bochner tensor is the Kähler analogue of the

Riemannian Weyl tensor). Examples of such manifolds can be found in literature, to mention only Bochner-flat manifolds ([4]) and weakly self-dual Kähler surfaces classified in [10] and [2]. In [1] authors proved that the normilzed Ricci form $\tilde{\rho}$ of a weakly Bochner-flat manifold is a hamiltonian 2-form. A normalized Ricci form is obtained from the Ricci form $\rho(X,Y) = \mathrm{Ric}(JX,Y)$ of a Kähler manifold by the formula

$$\tilde{\rho} = \rho - \frac{\mathrm{scal}}{n+2}\omega.$$

By Propositon 14.3.3 the normalized Ricci tensor $\tilde{\mathrm{Ric}}(X,Y) = -\tilde{\rho}(JX,Y)$ is a conformal Killing tensor. Due to the easy observation that if K is a conformal Killing tensor, then also $K + fg$ is a conformal Killing tensor for any smooth function f on M we have that $\mathrm{Ric} = \tilde{\mathrm{Ric}} + \frac{\mathrm{scal}}{n+2}g$ is a conformal Killing tensor. Hence we proved

Proposition 14.3.4. *Every weakly Bochner-flat manifold is an $\mathcal{AC}^{\perp}$-manifold.*

This result was obtained in a different way by Jelonek (see [11] and [9]).

14.4 Bundle construction

In the last part of this note we want to give an example of a construction of an $\mathcal{A}$-manifold on a r-torus principal fibre bundle over a special kind of almost Kähler manifolds following [20]. In the course of the construction we will give a useful theorem about lifting Killing tensors from the base manifold to the total space of the r-torus bundle.

Let (M,h) be a Riemannian manifold and suppose that β_i are closed 2-forms on M for $i = 1,\dots,r$ such that their cohomology classes $[\beta_i]$ are integral. In [14] it was proven that to each such cohomology class there corresponds a principal circle bundle $p_i : P_i \to M$ with a connection form θ_i such that

$$d\theta_i = 2\pi p_i^* \beta_i. \qquad (14.8)$$

Taking the Whitney sum of bundles (p_i, P_i, M) we obtain a principal r-torus bundle $p : P \to M$ classified by cohomology classes of β_i, $i = 1,\dots,r$. The connection form θ is a vector valued 1-form with coefficients θ_i, where θ_i are as before. For each connection form θ_i we define a vector field ξ^i by $\theta_i(\xi^i) = 1$. This vector field is just the fundamental vector field for θ_i corresponding to 1 in the Lie algebra of i-th S^1-factor of the bundle (p, P, M), namely $\mathbb{R}$.

It is easy to check that the tensor field g given by

$$g = \sum_{i,j=1}^{r} b_{ij}\theta_i \otimes \theta_j + p^* h \qquad (14.9)$$

is a Riemannian metric on P if $[b_{ij}]_{i,j=1}^r$ is some symmetric, positive definite $r \times r$ matrix with real coefficients. This Riemannian metric makes the projection $p : (P, g) \to (M, h)$ a Riemannian submersion (see [15]).

It is easy to prove the following lemma.

Lemma 14.4.1. *Each vector field ξ^i for $i = 1, \ldots, r$ is Killing with respect to the metric g. Moreover, define a tensor field T_i by $T_i X = \nabla_X \xi^i$ for a vector field X on P, where ∇ is the Levi-Civita connection of g. Then we have*

$$T_i \xi^j = 0, \quad L_{\xi^i} T_j = 0,$$

for $i \neq j$.

Hence, tensor fields T_i are horizontal, i.e. for each i there exists a tensor field $\tilde{T}_i$ on M such that $p_* \circ T_i = \tilde{T}_i \circ p_*$.

Using formulae for the Ricci tensor of the Riemannian submerson from Chapter 9 of [3] we have

$$\mathrm{Ric}(U, V) = \sum_{i=1}^{m} g(A_{E_i} U, A_{E_i} V), \tag{14.10}$$

$$\mathrm{Ric}(X, U) = -\sum_{i=1}^{m} g\left((\nabla_{E_i} A)_{E_i} X, U\right), \tag{14.11}$$

$$\mathrm{Ric}(X, Y) = \mathrm{Ric}_M(X, Y) - 2\sum_{i=1}^{m} g(A_X E_i, A_Y E_i), \tag{14.12}$$

where $\{E_i\}_{i=1}^m$ is some orthonormal basis of (M, h) lifted to the horizontal distribution of the submersion, Ric_M is the lifted Ricci tensor field of the metric h on M, U, V are sections of the vertical distribution and X, Y of the horizontal distribution. Moreover A is the O'Neill tensor which in our situation is given by

$$A_X Y = \sum_{i,j=1}^{r} b^{ij} \left(g(X, T_i Y)\xi^j + g(\xi^i, X) T_j Y\right),$$

for any vector fiedls X and Y. Observe that we omitted all components containing the O'Neill tensor T. This is because one can prove that the fibres of the Riemannian submersion $p : (P, g) \to (M, h)$ are totally geodesic and this cause the vanishing of the O'Neill tensor T (see for example [3], Chapter 9). Some calculations applied to (14.10) - (14.12) gives the following formulae for components of the Ricci tensor:

$$\mathrm{Ric}(U,V) = \sum_{i=1}^{m} g\left(\sum_{s,t=1}^{r} b^{st} g(\xi^s, U) T_t E_i, \sum_{k,l=1}^{r} b^{kl} g(\xi^k, V) T_l E_i\right), \tag{14.13}$$

$$\mathrm{Ric}(X,U) = \sum_{t=1}^{r} \delta d\theta_t(X) g(\xi^t, U) \tag{14.14}$$

$$\mathrm{Ric}(X,Y) = \mathrm{Ric}_M(X,Y) - \frac{1}{2}\sum_{s,t=1}^{r} b^{st} g(T_s X, T_t Y). \tag{14.15}$$

Now, let (M,g,J) be an almost Hermitian manifold, where J denotes the almost complex structure, i.e. a tensor field such that $J^2 = -\mathrm{id}_{TM}$, and g is any compatible metric satisfying $g(X,Y) = g(JX,JY)$ for some vector fields X and Y on M. We denote by ω the Kähler form. If ω is closed we call (M,g,J) an almost Kähler manifold. Moreover one can prove that in this case the Kähler form is also co-closed. If J is not integrable, then (M,g,J) is a strict almost Kähler manifold.

In [7] Jelonek constructed a strictly almost Kähler $\mathcal{A}$-manifold with non-parallel and J-invariant Ricci tensor. Moreover the Kähler form of such a manifold has a useful property. It is a constant multiple of some differential 2-form that belongs to an integral cohomology class i.e. a differential form in $H^2(M;\mathbb{Z})$. An almost Kähler manifold whose Kähler form satisfies this condition is called an *almost Hodge manifold*.

Returning to our construction suppose that (M_i, g_i, J_i), $i = 1, \ldots, n$ are almost Hodge manifolds such that Kähler forms ω_i are constant multiples of 2-forms α_i and their cohomology classes are integral, i.e. $[\alpha_i] \in H^2(M_i;\mathbb{Z})$. Denote by (M,g,J) the product manifold with the product metric and product almost complex structure and let pr_i be the projection on the i-th factor. From our earlier discussion we know that there exists a principal r-torus bundle classified by cohomology classes of differential forms $\beta_1, \ldots, \beta_r$ given by

$$\beta_j = \sum_{i=1}^{n} a_{ji} pr_i^* \alpha_i,$$

where $[a_{ji}]$ is some $r \times n$ matrix with integer coefficients. By (14.8) the coefficients θ_j of the connection form of (p,P,M) satisfy

$$d\theta_j = 2\pi p^* \beta_j = 2\pi \sum_{i=1}^{n} a_{ji} p^* \left(pr_i^* \alpha_i\right)$$

for every $j = 1, \ldots, r$. Since α_i's and Kähler forms ω_i of (M_i, g_i, J_i) are connected by $\omega_i = c_i\alpha_i$ for some constants c_i, $i = 1, \ldots, n$ we have

$$d\theta_j = 2\pi \sum_{i=1}^{n} \frac{a_{ji}}{c_i}\omega_i^*, \tag{14.16}$$

where by ω_i^* we denote the 2-form obtained from lifting ω_i to P. From the fact that θ_i's are dual to Killing vector fields ξ_i we get a formula for each tensor field $\tilde{T}_i$

$$\tilde{T}_i X = \pi \sum_{j=1}^{r} b_{ij} \sum_{k=1}^{n} \frac{a_{jk}}{c_k} J_k^* X$$

where J_k^* is the almost complex structure tensor of (M_k, g_k, J_k) lifted to the product manifold M.

From (14.14) and from (14.16) we can deduce that $\mathrm{Ric}(U, X) = 0$ for a vertical vector field U and horizontal field X, hence the horizontal and vertical distributions of the submersion (p, P, M) are Ricci-orthogonal.

Again after some computations we get that

$$\mathrm{Ric}(U, V) = \pi^2 \sum_{s,l=1}^{r} g(\xi^s, U)g(\xi^l, V) \sum_{i=1}^{m}\sum_{k=1}^{n} g_k\left(\frac{a_{sk}}{c_k}J_k E_i, \frac{a_{lk}}{c_k}J_k E_i\right). \tag{14.17}$$

Hence the vertical component of the Ricci tensor is a symmetrized product of Killing vector fields and is a Killing tensor by Proposition 14.3.1.

To prove that the horizontal part of Ric is also a Killing tensor field we need the following theorem about lifting Killing tensors from the base of a S^1-principal bundle ([20]).

Theorem 14.4.1. *Let K_i be a Killing tensor on (M_i, g_i, J_i) for $i = 1, \ldots, n$. Then the lift K^* of $K = K_1 + \ldots + K_n$ to P is a Killing tensor iff each K_i is J_i-invariant.*

It is worth noting, that we cannot lift in that way a conformal Killing tensor with non-vanishing P. In fact taking three vertical vector fields we see that P vanishes on vertical distribution. On the other hand for two vertical vector fields U, V and one horizontal vector field X the left-hand side of (14.2) vanish and the right-hand side reads $P(X)g(U, V)$, hence P has to vanish also on the horizontal distribution.

Corollary 14.4.1. *An r-torus bundle with metric defined by* (14.9) *can not be an $\mathcal{AC}^{\perp}$-manifold. Especially there exists no $\mathcal{AC}^{\perp}$ structures on K-contact and Sasakian manifolds.*

Next we show that the second component of the horizontal part of the Ricci tensor (14.15) is just a sum of lifts of metrics g_k, $k = 1, \ldots, n$.

$$\sum_{s,t=1}^{r} b^{st} g(T_s X, T_t Y) = \pi^2 \sum_{s,t=1}^{r} h \left(\sum_{j=1}^{r} b_{sj} \sum_{k=1}^{n} \frac{a_{jk}}{c_k} J_k^* X, \sum_{i=1}^{r} b_{ti} \sum_{l=1}^{n} \frac{a_{il}}{c_l} J_l^* Y \right).$$

Since J_k and J_l are orthogonal for different $k, l = 1, \ldots, n$ we obtain

$$\begin{aligned} &\sum_{s,t=1}^{r} b^{st} g(T_s X, T_t Y) \qquad (14.18) \\ &= \pi^2 \sum_{s,t=1}^{r} \sum_{k=1}^{n} h \left(\sum_{j=1}^{r} b_{sj} \frac{a_{jk}}{c_k} J_k^* X, \sum_{i=1}^{r} b_{ti} \frac{a_{ik}}{c_k} J_k^* Y \right) \\ &= \pi^2 \sum_{j,l=1}^{r} b_{jl} \sum_{k=1}^{r} \frac{a_{jk} a_{lk}}{c_k^2} g_k(X, Y). \end{aligned}$$

From the above Theorem we infer that, since a Riemannian metric is a Killing tensor and each g_k is J_k-invariant, the tensor field $K(X, Y) = \sum_{s,t=1}^{r} b^{st} g(T_s X, T_t Y)$ is a Killing tensor field.

As a corollary from the above considerations we obtain the following theorem.

Theorem 14.4.2. *Let P be a r-torus bundle over a Riemannian product (M, h) of almost Hodge $\mathcal{A}$-manifolds (M_k, g_k, J_k), $k = 1, \ldots n$ with metric g defined by (14.9). Then (P, g) is itself an $\mathcal{A}$-manifold.*

Proof. Since distributions $\mathcal{H}$ and $\mathcal{V}$ are orthogonal with respect to the Ricci tensor Ric of (P, g) we can write it as

$$\begin{aligned} \mathrm{Ric}(E, F) = \pi^2 \sum_{s,l=1}^{r} g(\xi^s, E) g(\xi^l, F) \sum_{i=1}^{m} \sum_{k=1}^{n} g_k \left(\frac{a_{sk}}{c_k} J_k E_i, \frac{a_{lk}}{c_k} J_k E_i \right) \\ + \mathrm{Ric}_M(E, F) - \frac{1}{2} \pi^2 \sum_{j,l=1}^{r} b_{jl} \sum_{k=1}^{r} \frac{a_{jk} a_{lk}}{c_k^2} g_k(E, F) \end{aligned}$$

using (14.18) and (14.17).The first component is a Killing tensor as a symmetrized product of Killing vector fields. The second and third components are Killing tensors by Theorem 14.4.1, since the Ricci tensor of the base is J-invariant. Since a sum of Killing tensors with constant coefficients is again a Killing tensor we have proved the theorem. □

References

[1] V. Apostolov, D. M. J. Calderbank, and P. Gauduchon. "Hamiltonian 2-forms in Kähler geometry, I General Theory". In: *arXiv:math/0202280v2* (2004).

[2] V. Apostolov, D. M. J. Calderbank, and P. Gauduchon. "The geometry of weakly self-dual Kähler surfaces". In: *Compos. Math.* 135, No. 3 (2003), pp. 279–322.

[3] A. L. Besse. *Einstein manifolds.* Springer-Verlag, Berlin etc., 1987.

[4] R. L. Bryant. "Bochner-Kähler metrics". In: *J. Am. Math. Soc.* 14, No. 3 (2001), pp. 623–715.

[5] B. Coll, J. J. Ferrando, and J. A. Sáez. "On the geometry of Killing and conformal tensors". In: *J. Math. Phys.* 47, No. 11 (2006).

[6] A. Gray. "Einstein-like manifolds which are not Einstein". In: *Geom. Dedicata* 7 (1978), pp. 259–280.

[7] W. Jelonek. "Almost Kähler $\mathcal{A}$-structures on twistor bundles". In: *Ann. Global Anal. Geom.* 17, No. 4 (1999), pp. 329–339.

[8] W. Jelonek. "A-tensors in Riemannian geometry". In: *preprint, Institute of Mathematics, Polish Academy of Sciences* 551 (1995).

[9] W. Jelonek. "Compact Kähler surfaces with harmonic anti-self-dual Weyl tensor". In: *Differ. Geom. Appl.* 16, No. 3 (2002), pp. 267–276.

[10] W. Jelonek. "Complex Kähler surfaces with harmonic anti-self-dual Weyl tensor". In: *Differ. Geom. Appl.* 16 (2002), pp. 267–276.

[11] W. Jelonek. "Higher-dimensional Gray Hermitian manifolds". In: *J.Lond. Math. Soc., II Ser.* 80, No. 3 (2009), pp. 729–749.

[12] W. Jelonek. "K-contact $\mathcal{A}$-manifolds". In: *Colloq. Math.* 75, No. 1 (1998), pp. 97–103.

[13] W. Jelonek. "Killing tensors and warped products". manuscript.

[14] S. Kobayashi. "Principal fibre bundles with the 1-dimensional toroidal group". In: *Tohoku Math. J., II. Ser.* 8 (1956), pp. 29–45.

[15] B. O'Neill. "The fundamental equations of a submersion". In: *Mich. Math. J.* 13 (1966), pp. 459–469.

[16] K. Sekigawa and L. Vanhecke. "Symplectic geodesic symmetries on Kähler manifolds". In: *Q. J. Math. Oxf. II Ser.* 37 (1986), pp. 95–103.

[17] U. Semmelmann. "Conformal Killing forms on Riemannian manifolds". In: *arXiv:math/0206117v1* (2002).

[18] U. Semmelmann and A. Moroianu. "Twistor forms on Kähler manifolds". In: *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5)* 2, No. 4 (2003), pp. 823–845.

[19] Z. Tang and W. Yan. "Isoparametric foliation and a problem of Besse on generalizations of Einstein condition". In: *arXiv:math/1307.3807v2* (2013).

[20] G. Zborowski. "$\mathcal{A}$-manifolds on a principal torus bundle over an $\mathcal{A}$-manifold base". In: *arXiv:1406.2465v1* (2014).

[21] G. Zborowski. "Construction of an $\mathcal{A}$-manifold on a principal torus bundle". In: *Ann. Univ. Paedagog. Crac. Stud. Math.* 12 (2013), pp. 5–19.

ISBN 978-83-63663-54-4