CONTEMPORARY MATHEMATICS

proceedings of the conference contemporary mathematics in kielce

sciendo

Proceedings of the Contemporary Mathematics in Kielce 2020

24-27 February 2021, Kielce, Poland

editor
Szymon Walczak

sciendo

sciendo

De Gruyter Poland Sp. z o.o.
Bogumila Zuga 32a, 01-811 Warsaw, Poland

Ministry
of Education
and Science

Co-financed by the "Excellent Science"
program of the Minister of Education and Science
grant No. DNK/SP/465255/2020

ISBN 978-83-66675-36-0

Proceedings of the Contemporary Mathematics in Kielce 2020

24-27 February 2021, Kielce, Poland

This page is intentionally left blank

PREFACE

This volume contains papers written by the participants of the international conference Contemporary Mathematics in Kielce 2020 held at the Department of Mathematics of Jan Kochanowski University in Kielce.

95 registered participants from 13 countries from 4 continents, as well as students of mathematics from Jan Kochanowski University in Kielce, had participated in this conference. The lectures were carried out in several thematic sections: Topology & Algebra (organizer: Taras Banakh), Mathematical Analysis (organizer: Grzegorz Łysik), Geometry (organizer: Szymon Walczak) and Mathematical Education (organizer: Szymon Walczak).

Plenary lectures were given by Antonio Avilés (University of Murcia), Michael Barnsley (Australian National University), Krzysztof Ciesielski (Jagiellonian University), Wiesław Kubiś (Cardinal Stefan Wyszyński University / Academy of Sciences of the Czech Republic), and Michał Wojciechowski (IM PAN). 58 lectures were delivered within the thematic sections. The conference speakers included both experienced mathematicians with recognized achievements and mathematicians starting their research work.

Due to COVID-19 epidemic constraints, the conference was held entirely remotely. The Scientific Committee of the conference was led by Taras Banakh (UJK), while the employees of the Department of Mathematics of the Jan Kochanowski University in Kielce were responsible for the conference organization.

The conference was part of the celebration of the 50th anniversary of Jan Kochanowski University and it was held under the patronage of the Rector of the Jan Kochanowski University in Kielce. Contemporary Mathematics in Kielce 2020 was co-financed by the "Excellent Science" program of the Minister of Education and Science (grant No. DNK/SP/465255/2020).

The papers contained in this volume are closely related to the lectures given at the conference. All the papers contained in this volume were refereed by the experts. Some of them are research papers containing original results, while some contain surveys.

The conference organizers are grateful to both the authors and reviewers as well as everyone involved in the creation of this publication.

The Editor

This page is intentionally left blank

CONTENTS

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON THE FUKS–LODDER–KOTSCHICK CLASS FOR DEFORMATIONS OF FOLIATIONS

TARO ASUKE

ABSTRACT. Given differential families of foliations, we can define characteristic classes for them. The most fundamental ones are derivatives of 'classical' secondary classes such as the Godbillon–Vey class. On the other hand, there are some classes which are not obtained as derivatives. The Fuks–Lodder–Kotschick class, FLK class for short, is the most fundamental one and is discussed in this article. We will present some conditions for the FLK class to be trivial first in terms of cohomology classes associated with deformations, and second in terms of transverse projective structures. It is unknown in the real category if there is a family of foliations of which the FLK class is non-trivial. In this article, we will give an example of a family of transversely holomorphic foliations of which the complex version of the FLK class is non-trivial.

This work was supported by JSPS KAKENHI Grant Number JP21H00980

KEYWORDS: foliations, deformations, characteristic classes

MSC2010: 57R30, 58H10, 37F75

Received 16 April 2021; revised 27 May 2021; accepted 15 June 2021

INTRODUCTION

Given a regular foliation (foliations without singularities), we can define so-called secondary characteristic classes such as the Godbillon–Vey class. One of the most significant properties of such classes is that some of them, especially the Godbillon–Vey class, vary continuously according to deformations of foliations. If the deformation is differentiable, we can consider the derivative of such classes with respect to deformation parameters. It is classical that the derivative of the Godbillon–Vey class is non-trivial, namely, there is a family of foliations with respect to which the derivative is non-trivial. This kind of derivatives can be

 https://doi.org/10.2478/9788366675360-001.

regarded as characteristic classes for deformations of foliations. The theory of derivatives of characteristic classes for foliations is well-developed. Some of basic references are [13], [8], [14]. In the transversely holomorphic case, we refer to [7], [10]. There is also a general construction to obtain characteristic classes for deformations of foliations by means of a certain differential graded algebra [3]. Thus defined classes include classes which are not the derivatives of 'classical' secondary classes as above. The most basic one is the Fuks–Lodder–Kotschick class, the FLK class for short, which we will discuss in this article. The FLK class is defined as a combination of the Godbillon–Vey class and its derivative. It is first found by Fuks [9]. Then, Lodder [22] and Kotschick [19] continue the study. Throughout the study, non-triviality has been always in question. In particular, Kotschick gave some sufficient conditions for triviality. It is known that the derivatives of the Godbillon–Vey class is related with transverse projective structures [23], [2]. We briefly review the relation and introduce a certain analogue for the FLK class after [4]. This leads to some vanishing theorems for the FLK class: the FLK class is trivial if $H^k(M;\Theta_{\mathcal{F}}) = 0$ for $k = 1$ or $k = 2$, or $\mathcal{F}$ is transversely projective (Theorem 3.6). On the other hand, the author gave in [1] an example of a transversely holomorphic foliation of which the complex version of the FLK class is non-trivial. We will give an example of the same kind, as well as a generalization to obtain non-trivial FLK classes in the transversely holomorphic category (Example 3.7 and Proposition 3.8). Unfortunately, it seems difficult to modify our example to find non-trivial FLK classes in the real category. The non-triviality of the FLK class in the real category is still quite unknown.

1. Preliminaries

We first recall some basic notions. A fundamental reference is [6]. Let M be a closed manifold and $\mathcal{F}$ a regular foliation (foliation without singularities) of M. Let TM be the tangent bundle of M and $T\mathcal{F}$ the subbundle which consists of vectors tangent to the leaves of $\mathcal{F}$. We set $Q(\mathcal{F}) = TM/T\mathcal{F}$ and call it the *normal bundle* of $\mathcal{F}$.

In what follows, we assume that foliations are transversely oriented. The codimension of $\mathcal{F}$ is always denoted by q unless otherwise mentioned. Let ω be a trivialization of $\bigwedge^q Q(\mathcal{F})^*$. The trivialization ω is naturally regarded as a q-form on M which defines $\mathcal{F}$. That is, if we set $\ker\omega = \{v \in TM \mid \iota_v\omega = 0\}$, where ι_v denotes the interior product with v, then we have $T\mathcal{F} = \ker\omega$. By the Frobenius theorem, there exists a 1-form, say η, such that $d\omega = -\eta\wedge\omega$. It is known that the $(2q+1)$-form $\eta \wedge (d\eta)^q$ represents a cohomology class independent of the choice of ω and η.

Definition 1.1. The class in $H^{2q+1}(M;\mathbb{R})$ represented by $\eta \wedge (d\eta)^q$ is called the *Godbillon–Vey class* of $\mathcal{F}$ and denoted by $\mathrm{GV}(\mathcal{F})$.

Remark 1.2. The Godbillon–Vey class is often normalized as $(-1/2\pi)^{q+1}\eta \wedge (d\eta)^q$.

Non-triviality of the Godbillon–Vey class has been known since the beginning. Namely, there exists a foliated manifold $(M, \mathcal{F})$ such that $\mathrm{GV}(\mathcal{F}) \in H^{2q+1}(M; \mathbb{R})$ is non-trivial for any q. This is first found by Roussarie for $q = 1$ (see [11]) and examples for $q \geq 2$ were soon found. Much more is known.

Definition 1.3. A *smooth* 1*-parameter family* of codimension-q foliations $\{\mathcal{F}_s\}$ is a family of codimension-q foliations of M such that $\{T\mathcal{F}_s\}$ is a smooth 1-parameter family.

Theorem 1.4 (cf. [27], [15])**.** *For any q, there exists a closed manifold M and smooth family $\{\mathcal{F}_s\}$ of codimension-q foliations of M such that*

$$\mathrm{GV}(\mathcal{F}_s) \in H^{2q+1}(M; \mathbb{R})$$

varies smoothly with respect to s.

In the above theorem, $H^{2q+1}(M; \mathbb{R})$ is regarded as a vector space and $\mathrm{GV}(\mathcal{F}_s)$ is a point in it.

The secondary characteristic classes are divided into two types as follows.

Definition 1.5. Let γ be a characteristic class for foliations. The class γ is said to be *variable* if γ admits a *continuous variation*, namely, if there is a family $\{\mathcal{F}_s\}$ of foliations such that $\gamma(\mathcal{F}_s)$ varies continuously with respect to s. Otherwise, the class γ is said to be *rigid*.

When we discuss characteristic classes for deformations of foliations, so-called transverse structures are relevant. The following ones will play important roles.

Definition 1.6. Let $\mathcal{F}$ be a foliation.

(1) If $\mathcal{F}$ admits a Riemannian metric on $Q(\mathcal{F})$ invariant under holonomy, then $\mathcal{F}$ is said to be *Riemannian*.
(2) If $\mathcal{F}$ admits a projective structure in the transverse direction which is invariant under holonomy, then $\mathcal{F}$ is said to be *transversely projective*. If moreover the projective structure is flat, then $\mathcal{F}$ is said to be *transversely projectively flat*.
(3) If $\mathcal{F}$ admits a holomorphic structure, namely, an integrable almost complex structure, in the transverse direction which is invariant under holonomy, then $\mathcal{F}$ is said to be *transversely holomorphic*. In this case the half of the codimension of $\mathcal{F}$ is called *complex codimension* of $\mathcal{F}$.

Transverse projective structures are usually assumed to be invariant under holonomy, however, we can introduce transverse projective structure not necessarily invariant under holonomy. It is similar to the cases of transverse Riemannian structures. Riemannian foliations are those which admit Riemannian metrics on

$Q(\mathcal{F})$ invariant under holonomy. However, we can always consider Riemannian metrics on $Q(\mathcal{F})$. Transverse projective structures are defined by means of transverse Thomas–Whitehead connections (TW-connections for short) [26], [12], [2]. It is a linear connection on the normal bundle of a foliation associated with $\bigwedge^q Q(\mathcal{F})$, and is a kind of Cartan connections.

Characteristic classes of foliations with transverse structures of certain kinds are well-studied ([21], [16], [25], etc). For example, we have the following

Theorem 1.7 (cf. [21], [24])**.** *If $\mathcal{F}$ is Riemannian, then* $\mathrm{GV}(\mathcal{F})$ *is trivial.*

The triviality of the Godbillon–Vey class in the above theorem can be explained in several ways. One can show that $(d\eta)^q = 0$ (cf. [21]) as well as that η can be chosen to be 0 under suitable choices of connections (cf. [24]).

2. Godbillon–Vey class and its derivative

By Theorem 1.4, we can consider derivatives of the Godbillon–Vey class with respect to deformations. We fix a manifold M. Let 0 be the base point of deformation parameter, and let $\mathcal{F}$ also denote $\mathcal{F}_0$.

Definition 2.1. Let $\{\mathcal{F}_s\}$ be a smooth family of codimension-q foliations of M. We set

$$D_{\{\mathcal{F}_s\}}\mathrm{GV}(\mathcal{F}) = \frac{\partial}{\partial s}\mathrm{GV}(\mathcal{F}_s)\bigg|_{s=0} \in H^{2q+1}(M;\mathbb{R})$$

and call it the *derivative* of $\mathrm{GV}(\mathcal{F})$ with respect to $\{\mathcal{F}_s\}$.

Remark 2.2. A smooth 1-parameter family $\{\mathcal{F}_s\}$ of foliations determines a class in $H^1(M;\Theta_{\mathcal{F}})$, where $\Theta_{\mathcal{F}}$ denotes the 'structural sheaf' of $\mathcal{F}$ which we define by applying the Kodaira–Spencer theory to foliations. In particular, $H^*(M;\Theta_{\mathcal{F}})$ is isomorphic to the cohomology of $Q(\mathcal{F})$-valued differential forms along the leaves of $\mathcal{F}$. It is known that the derivative $D_{\{\mathcal{F}_s\}}\mathrm{GV}(\mathcal{F})$ depends only on the class determined by $\{\mathcal{F}_s\}$, not the family itself. Actually, we can extend $D_\bullet\mathrm{GV}(\mathcal{F})$ to the whole $H^1(M;\Theta_{\mathcal{F}})$. We refer for details to [15] and references therein.

Definition 2.3.

(1) We call elements of $H^1(M;\Theta_{\mathcal{F}})$ *infinitesimal deformations* of $\mathcal{F}$.
(2) We call one-parameter families of deformation of $\mathcal{F}$ *actual deformations* of $\mathcal{F}$.

Under this terminology, an actual deformation determines an infinitesimal deformation.

Let $\{\mathcal{F}_s\}$ be a smooth 1-parameter family of foliations. Then, we can find a smooth 1-parameter family $\{\omega_s\}$ such that ω_s defines $\mathcal{F}_s$ for each s. We can also find a smooth 1-parameter family $\{\eta_s\}$ such that $d\omega_s = -\eta_s \wedge \omega_s$ holds for each s. Let $\mathcal{F}$, ω and η denote $\mathcal{F}_0$, ω_0, and η_0, respectively. We set $\dot\eta = (\partial/\partial s)\eta_s|_{s=0}$. Then, we have the following

Theorem 2.4. *The derivative $D_{\{\mathcal{F}_s\}}\mathrm{GV}(\mathcal{F})$ is represented by $(q+1)\dot{\eta}\wedge(d\eta)^q$.*

Proof. We have

$$\begin{aligned}\frac{\partial}{\partial s}\mathrm{GV}(\mathcal{F}_s) &= \frac{\partial}{\partial s}(\eta_s \wedge (d\eta_s)^q)\\ &= \left(\frac{\partial}{\partial s}\eta_s\right)\wedge(d\eta_s)^q + \eta_s \wedge \left(\frac{\partial}{\partial s}(d\eta_s)^q\right).\end{aligned}$$

On the other hand, we have

$$\frac{\partial}{\partial s}(d\eta_s)^q = qd\left(\frac{\partial}{\partial s}\eta_s\right)\wedge(d\eta_s)^{q-1},$$

and

$$\begin{aligned}&d\left(\eta_s \wedge \left(\frac{\partial}{\partial s}\eta_s\right)\wedge(d\eta_s)^{q-1}\right)\\ &= (d\eta_s)\wedge\left(\frac{\partial}{\partial s}\eta_s\right)\wedge(d\eta_s)^{q-1} - \eta_s\wedge d\left(\frac{\partial}{\partial s}\eta_s\right)\wedge(d\eta_s)^{q-1}\\ &= \left(\frac{\partial}{\partial s}\eta_s\right)\wedge(d\eta_s)^{q} - \eta_s\wedge d\left(\frac{\partial}{\partial s}\eta_s\right)\wedge(d\eta_s)^{q-1}.\end{aligned}$$

Therefore, $(\partial/\partial s)\mathrm{GV}(\mathcal{F}_s)$ is represented by $(q+1)\left((\partial/\partial s)\eta\right)\wedge(d\eta_s)^q$. By setting $s=0$, we obtain the theorem. □

The following is a precise version of Theorem 1.4.

Theorem 2.5. *For any q, there exist a closed manifold M and a smooth family $\{\mathcal{F}_s\}$ of codimension-q foliations of M such that $D_{\{\mathcal{F}_s\}}\mathrm{GV}(\mathcal{F}) \in H^{2q+1}(M;\mathbb{R})$ is non-trivial.*

Much more is known about variations and derivatives of characteristic classes. For example, there is a family of variable characteristic classes, say $\{\gamma_i\}$, for which there exists a family of foliations $\{\mathcal{F}_\lambda\}$ such that there are no linear dependence among $\gamma_i(\mathcal{F}_\lambda)$'s according to variations of λ, where λ is a multi-parameter in general [15]. On the other hand, some characteristic classes are known to be non-trivial but rigid ([17], [18], etc.):

Theorem 2.6 ([1])**.**

(1) *For any q, there exists a manifold M and a transversely holomorphic foliation $\mathcal{F}$ of M, of complex codimension q such that $\mathrm{GV}(\mathcal{F}) \in H^{2q+1}(M;\mathbb{R})$ is non-trivial.*

(2) *Let $\{\mathcal{F}_s\}$ be a smooth one-parameter family of transversely holomorphic foliations, where we assume that the family is not only smooth as real foliations but transverse holomorphic structures form also a smooth family. Then, $D_{\{\mathcal{F}_s\}}\mathrm{GV}(\mathcal{F}_\mathbb{R}) \in H^{4q+1}(M;\mathbb{R})$ is trivial, where q denotes the complex codimension of $\mathcal{F}$ and $\mathcal{F}_\mathbb{R}$ denotes the foliation $\mathcal{F}$ regarded as a real*

foliation of codimension $2q$ by we forgetting the transverse holomorphic structure.

The derivative of the Godbillon–Vey class is related with transverse projective structures. By means of connection matrices of transverse TW-connections, we can define a homomorphism

$$\mathcal{L}_{\mathcal{P}} \colon H^r(M; \Theta_{\mathcal{F}}) \to H^{2q+r}(M; \mathbb{R}),$$

where $\mathcal{P}$ denotes the projective structure, namely, the projective equivalence class determined by the transverse transverse TW-connection. We have the following

Theorem 2.7 ([2])**.** *The homomorphism $\mathcal{L}_{\mathcal{P}}$ depends only on the projective structure $\mathcal{P}$. Moreover, if $\mathcal{P}$ is holonomy invariant, then $\mathcal{L}_{\mathcal{P}}$ is trivial.*

Theorem 2.8 ([23], [2])**.** *Let $\mu \in H^1(M; \Theta_{\mathcal{F}})$ be an infinitesimal deformation. Then, the derivative of the Godbillon–Vey class with respect to μ is represented by $\mathcal{L}_{\mathcal{P}}(\mu)$.*

Theorem 2.8 is first found by Maszczyk [23] for codimension-one foliations. In this case, $\mathcal{L}_{\mathcal{P}}$ is obtained as follows. Let ω_s and η_s be as in Section 1. We set $\dot{\omega} = (\partial/\partial_s)\omega_s|_{s=0}$ and $\dot{\eta} = (\partial/\partial_s)\eta_s|_{s=0}$. We have

$$d\omega + \eta \wedge \omega = 0, \tag{2.9}$$

and

$$d\dot{\omega} + \eta \wedge \dot{\omega} + \dot{\eta} \wedge \omega = 0. \tag{2.10}$$

Indeed, if ω and η satisfy (2.9), then any $\dot{\omega}$ and $\dot{\eta}$ which satisfy (2.10) are called *infinitesimal deformations* of ω and η. We consider a certain equivalence relation among such $\dot{\omega}$'s and the resulting space is known to be isomorphic to $H^1(M; \Theta_{\mathcal{F}})$. Now, let y be the local coordinate in the direction transversal to the leaves of $\mathcal{F}$. As $T\mathcal{F} = \ker \omega$, we can locally find a function f such that $\omega = f dy$, where the equality is understood to locally hold. We have $d\omega = df \wedge dy = d \log f \wedge \omega$ so that we can also locally find a function g such that $\eta = -d \log f + g\omega$. Then, we have $d\eta = (dg - g\eta) \wedge \omega$ so that we have

$$\begin{aligned}
\dot{\eta} \wedge d\eta &= \dot{\eta} \wedge (dg - g\eta) \wedge \omega \\
&= (dg - g\eta) \wedge (d\dot{\omega} + \eta \wedge \dot{\omega}) \\
&= (dg - g\eta) \wedge d\dot{\omega} + dg \wedge \eta \wedge \dot{\omega}.
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
d((dg - g\eta) \wedge \dot{\omega}) &= -(dg \wedge \eta + g d\eta) \wedge \dot{\omega} - (dg - g\eta) \wedge d\dot{\omega} \\
&= -\dot{\eta} \wedge d\eta - g d\eta \wedge \dot{\omega}, \\
g d\eta \wedge \dot{\omega} &= g(dg - g\eta) \wedge \omega \wedge \dot{\omega} \\
&= g(dg + g d\log f) \wedge \omega \wedge \dot{\omega}, \\
g^2 \omega \wedge d\dot{\omega} &= -g^2 \omega \wedge (\dot{\eta} \wedge \omega + \eta \wedge \dot{\omega}) \\
&= -g^2 (d \log f) \wedge \omega \wedge \dot{\omega}, \\
d(g^2 \omega \wedge \dot{\omega}) &= 2g dg \wedge \omega \wedge \dot{\omega} + g^2 (d\log f) \wedge \omega \wedge \dot{\omega} - g^2 \omega \wedge d\dot{\omega}, \\
&= 2g dg \wedge \omega \wedge \dot{\omega} + 2g^2 (d \log f) \wedge \omega \wedge \dot{\omega}.
\end{aligned}$$

Hence we have

$$\dot{\eta} \wedge d\eta = -d\left(\left(dg - g\eta + \frac{1}{2} g^2 \omega\right) \wedge \dot{\omega}\right).$$

Let y' be another coordinate in the transversal direction and suppose that $y' = \gamma(y)$. We define f' and g' in a parallel way to f and g but respect to the coordinate y'. Then, we have

$$\left(dg' - g'\eta + \frac{1}{2} g'^2 \omega\right) - \left(dg - g\eta + \frac{1}{2} g^2 \omega\right) = -S(\gamma) \frac{1}{f^2} \omega,$$

where $S(\gamma) = (\gamma'''/\gamma') - 3/2(\gamma''/\gamma')^2$ denotes the *Schwarzian derivative* of γ. Note that $\omega = f dy$ and that $S(\gamma) dy \otimes dy$ is a cocycle in the sense that we have $S(\gamma_2 \circ \gamma_1) dy_1 \otimes dy_1 = \gamma_1^*(S(\gamma_2) dy_2 \otimes dy_2) + S(\gamma_1) dy_1 \otimes dy_1$, where $y_2 = \gamma_1(y_1)$. The mapping $\mathcal{L}_{\mathcal{P}}$ is given by setting

$$\mathcal{L}_{\mathcal{P}}(\mu) = -d\left(\left(dg - g\eta + \frac{1}{2} g^2 \omega\right) \wedge \mu\right).$$

Note that the right hand side is locally defined so that $\mathcal{L}_{\mathcal{P}}(\mu)$ is closed but not necessarily exact. Although we need more involved calculations, we can generalize the formula in higher codimensional cases to obtain Theorem 2.8. In the higher codimensional cases, the Schwarzian derivative is replaced by components of the curvature of transverse projective connections.

Theorems 2.7 and 2.8 have the following immediate

Corollary 2.11 ([23], [2])**.**

(1) *If $H^1(M; \Theta_{\mathcal{F}}) = 0$, then the derivative of the Godbillon–Vey class with respect to any infinitesimal deformation of $\mathcal{F}$ is trivial.*

(2) *If $\mathcal{F}$ is transversely projective, then the derivative of the Godbillon–Vey class with respect to any infinitesimal deformation of $\mathcal{F}$ is trivial.*

The first assertion is a tautology but we state it for comparison with Theorem 3.6.

Remark 2.12. The Godbillon–Vey class is trivial if $\mathcal{F}$ is Riemannian by theorem 1.7. It is known that the Godbillon–Vey class is also trivial if $\mathcal{F}$ is transversely affine. Hence, roughly speaking, the Godbillon-Vey class is trivial if $\mathcal{F}$ admits transverse geometric structures of first order. On the other hand, Corollary 2.11 says that the Godbillon-Vey class is rigid if $\mathcal{F}$ is transversely projective. This suggests that the Godbillon–Vey class is rigid if $\mathcal{F}$ admits geometric structures of second order. It suffices to establish the rigidity for transversely conformal foliations. This is partly done, indeed, if the codimension of $\mathcal{F}$ is equal to two, then $\mathcal{F}$ is transversely holomorphic. In the category of transversely holomorphic foliations, the Godbillon–Vey class is rigid under deformations (Theorem 2.6). Some other results are also known but the author does not know any unified account.

Remark 2.13. If $\mathcal{F}$ is transversely holomorphic, then we can define a complex version of the Godbillon–Vey class and its derivative by imitating the above constructions. That is, the one-form η is now a $\mathbb{C}$-valued and $\eta\wedge(d\eta)^q \in H^{2q+1}(M;\mathbb{C})$, where q denotes the complex codimension of $\mathcal{F}$. The class represented by $\eta\wedge(d\eta)^q$ or $(-1/2\pi\sqrt{-1})^{q+1}\eta\wedge(d\eta)^q$ when normalized, is called the *Bott class* or complex Godbillon–Vey class. Note that the Bott class [5] was found prior to the Godbillon–Vey class [11].

3. The Fuks–Lodder–Kotschick class

Let $\{\mathcal{F}_s\}$ be a family of foliations. By differentiating the equality $d(\eta_s \wedge (d\eta_s)^q)) = 0$ with respect to s, we obtain $d\dot\eta\wedge(d\eta)^q = d(\dot\eta\wedge(d\eta)^q) = 0$. It follows that the $(2q+2)$-form $\dot\eta\wedge\eta\wedge(d\eta)^q$ is closed. Indeed, we have $d(\dot\eta\wedge\eta\wedge(d\eta)^q) = \eta\wedge d(\dot\eta\wedge(d\eta)^q) - \dot\eta\wedge d(\eta\wedge(d\eta)^q) = 0$.

Definition 3.1 ([9],[22],[19])**.** The class in $H^{2q+2}(M;\mathbb{R})$ represented by $\dot\eta\wedge\eta\wedge(d\eta)^q$ is called the *Fuks–Lodder–Kotschick class* of $\{\mathcal{F}_s\}$ and denoted by $\mathrm{FLK}_\mu(\mathcal{F})$, where μ denotes the element of $H^1(M;\Theta_{\mathcal{F}})$ determined by $\{\mathcal{F}_s\}$.

We refer to the Fuks–Lodder–Kotschick class as the FLK class for short.

Remark 3.2.

(1) The FLK class is normalized as $(-1/2\pi)^{q+2}\dot\eta\wedge\eta\wedge(d\eta)^q$ in [4] in the real case. If transversely holomorphic foliations are considered, then the FLK-class is an element of $H^{2q+2}(M;\mathbb{C})$ and is normalized as $(-1/2\pi\sqrt{-1})^{q+2}\dot\eta\wedge\eta\wedge(d\eta)^q$, where q denotes the complex codimension of $\mathcal{F}$.

(2) It is necessary to show the independence for FLK class of choices of ω, η, $\dot\omega$ and $\dot\eta$. This is done in [22] and [19]. On the other hand, we can introduce a certain differential graded algebra which describes characteristic classes for deformations of foliations. Constructions become involved but we can also show the independence by establishing a characteristic mapping for deformations in this more general setting [3].

(3) The FLK class is actually defined for any infinitesimal deformation of $\mathcal{F}$. That is, μ need not be determined by an actual deformation $\{\mathcal{F}_s\}$ in Definition 3.1 [1].

The FLK class is first introduced by Fuks [9]. Lodder continued the study and discussed the relationship of the FLK class and the Leibniz cohomology in [22]. After that, Kotschick gave some sufficient conditions for triviality [19]. In particular, it is shown that the FLK class is trivial if $\mathcal{F}$ is of codimension one and if $\mathcal{F}$ is transversely projectively flat (cf. Definition 1.6 and Theorem 3.6). The (non-)triviality has been always in question. It is still unknown if the FLK class can be non-trivial in the real category. On the other hand, we will present an example with non-trivial FLK class in the transversely holomorphic category (Example 3.7).

We begin with general properties of the FLK class. First, we have the following

Theorem 3.3.

(1) *In the real category, the FLK class for $\mathcal{F}$ with respect to μ is independent of the choice of trivializations of $\bigwedge^q Q(\mathcal{F})^*$.*
(2) *In the transversely holomorphic category, the FLK class for $\mathcal{F}$ with respect to μ depends on the homotopy types of trivializations of $\bigwedge^q Q(\mathcal{F})^*$.*

Actually, (1) follows from (2) in Theorem 3.3, because the trivialization of $\bigwedge^q Q(\mathcal{F})^*$ is unique up to homotopy once the transverse orientation is fixed. Taking Theorem 3.3 into account, we will denote the FLK class of $\mathcal{F}$ with respect to μ by $\mathrm{FLK}_\mu(\mathcal{F}; e)$ if we emphasize the trivialization e.

The FLK class is further understood by representing the FLK class by $\mathcal{L}_{\mathcal{P}}$ in Theorem 2.7. We have the following

Lemma 3.4 (Lemma 2.1 of [4]). *We have a well-defined mapping*

$$\theta^e \wedge \colon H^r(M; \Theta_{\mathcal{F}}) \to H^{r+1}(M; \Theta_{\mathcal{F}})$$

which depends on the homotopy class of the trivialization e of $\bigwedge^q Q(\mathcal{F})$.

The symbol 'θ^e' stands for the connection form of a Bott connection on $\bigwedge^q Q(\mathcal{F})$ with respect to the trivialization e of $\bigwedge^q Q(\mathcal{F})$. We remark that Lemma 3.4 can be regarded as a version of permanence theorem for elements of $H^*(M; \Theta_{\mathcal{F}})$. We refer to [20] for a permanence theorem for 'classical' secondary characteristic classes. Related calculations can be also found in [13].

Lemma 3.5 (Lemma 2.7 of [4]). *If $\mu \in H^1(M; \Theta_{\mathcal{F}})$, then we have*

$$\mathcal{L}_{\mathcal{P}}(\theta^e \wedge \mu) = \frac{1}{q} \mathrm{FLK}_\mu(\mathcal{F}; e).$$

Lemma 3.5 is shown in a similar way to proving Theorem 2.8. Combining Lemmata 3.4 and 3.5, we obtain the following

Theorem 3.6 ([4])**.**
(1) *If* $H^1(M;\Theta_{\mathcal{F}}) = 0$ *or* $H^2(M;\Theta_{\mathcal{F}}) = 0$*, then the FLK class with respect to any infinitesimal deformation of* $\mathcal{F}$ *is trivial.*
(2) *If* $\mathcal{F}$ *is transversely projective, then the FLK class with respect to any infinitesimal deformation of* $\mathcal{F}$ *is trivial.*

As we have already mentioned, (2) of Theorem 3.6 is known in the most basic case [19]. We now give an example of which the FLK class is non-trivial.

Example 3.7 (cf. Example 5.11 of [1])**.** We begin with a classical example of Bott [5]. Let $\lambda, \mu \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$ and $X = \lambda z \frac{\partial}{\partial z} + \mu w \frac{\partial}{\partial w}$ be a holomorphic vector field on $\mathbb{C}^2$, where (z, w) denotes the standard coordinates on $\mathbb{C}^2$. The integral curves of X form a holomorphic foliation of $\mathbb{C}^2 \setminus \{(0,0)\}$ which is denoted by $\mathcal{F}_\tau$, where $\tau = \lambda/\mu$. Note that $\mathcal{F}_\tau$ actually depends only on the ratio τ, not λ and μ themselves. In what follows, we regard τ as the parameter of deformations. We set $\omega_\tau = w dz - \tau z dw$ and

$$\eta_\tau = -\left(1 + \frac{1}{\tau}\right) \frac{\bar{z} dz}{|z|^2 + |w|^2} - (1 + \tau) \frac{\bar{w} dw}{|z|^2 + |w|^2}.$$

Then, we have $d\omega_\tau = -\eta_\tau \wedge \omega_\tau$. We have

$$\begin{aligned} d\eta_\tau &= -\left(1 + \frac{1}{\tau}\right) \frac{|w|^2 d\bar{z} \wedge dz - \bar{z}\bar{w} dw \wedge dz - \bar{z} w d\bar{w} \wedge dz}{(|z|^2 + |w|^2)^2} \\ &\quad - (1 + \tau) \frac{|z|^2 d\bar{w} \wedge dw - \bar{z}\bar{w} dz \wedge dw - \bar{w} z d\bar{z} \wedge dw}{(|z|^2 + |w|^2)^2}, \\ \eta_\tau \wedge d\eta_\tau &= -\left(\tau + 2 + \frac{1}{\tau}\right) \frac{1}{(|z|^2 + |w|^2)^2} (\bar{z} dz \wedge dw \wedge d\bar{w} + \bar{w} dz \wedge d\bar{z} \wedge dw). \end{aligned}$$

If we assume that the ratio τ is not a negative real number, then, leaves of $\mathcal{F}_\tau$ are transversal to the unit sphere S^3 so that we obtain a transversely holomorphic foliation of S^3 which is denoted by $\mathcal{F}'_\tau$. After normalization by multiplying $(-1/2\pi\sqrt{-1})^2$ and considering the volume of the unit sphere S^3 to be equal to 1, we have $\mathrm{Bott}(\mathcal{F}'_\tau) = \tau + 2 + 1/\tau \in H^3(S^3;\mathbb{C})$ and $D_\tau \mathrm{Bott}(\mathcal{F}'_\tau) = 1 - 1/\tau^2$. This is seen by taking the derivative of $D_\tau \mathrm{Bott}(\mathcal{F}'_\tau)$ with respect to τ, or by direct calculations which make use of the equality

$$\dot{\eta_\tau} = \frac{1}{\tau^2} \frac{\bar{z}}{|z|^2 + |w|^2} dz - \frac{\bar{w}}{|z|^2 + |w|^2} dw,$$

where the symbol '$\cdot$' represents the derivative with respect to τ. Note that we have $\dot{\omega_\tau} = -z dw$ and $d\dot{\omega_\tau} = -\dot{\eta_\tau} \wedge \omega_\tau - \eta_\tau \wedge \dot{\omega_\tau} = -dz \wedge dw$. On the other hand, the FLK class $\mathrm{FLK}(\mathcal{F}'_\tau; e)$ is trivial for any trivialization e of $Q(\mathcal{F}'_\tau)$. The simplest reason is that the FLK class is of degree four while the dimension of S^3 is three. We can give another explanation as follows. The dual of e is a trivialization of $Q(\mathcal{F}'_\tau)^*$ so that it is of the form $\varphi\omega_\tau$, where φ is a $\mathbb{C}^*$-valued function on S^3. We have then $d(\varphi\omega_\tau) = -(\eta_\tau - d\varphi/\varphi) \wedge (\varphi\omega_\tau)$. We also have

$$\begin{aligned} d(\varphi\dot{\omega_\tau}) &= d\varphi \wedge \dot{\omega_\tau} - \varphi\dot{\eta_\tau} \wedge \omega_\tau - \varphi\eta_\tau \wedge \dot{\omega_\tau} \\ &= -\left(\eta_\tau - \frac{d\varphi}{\varphi}\right) \wedge \varphi\dot{\omega_\tau} - \dot{\eta_\tau} \wedge \varphi\omega_\tau. \end{aligned}$$

Hence the FLK class of $\mathcal{F}'_\tau$ with respect to the trivialization e is represented by

$$\dot{\eta_\tau} \wedge \left(\eta_\tau - \frac{d\varphi}{\varphi}\right) \wedge d\eta_\tau = -\dot{\eta_\tau} \wedge \frac{d\varphi}{\varphi} \wedge d\eta_\tau$$

because $\dot{\eta_\tau} \wedge \eta_\tau \wedge d\eta_\tau = 0$. It follows that the FLK class of $\mathcal{F}'_\tau$ is decomposed up to multiplicative constants into the derivative of the Bott class and the class represented by $d\varphi/\varphi$. The latter is trivial because $H^1(S^3;\mathbb{C}) = 0$. Therefore, the FLK class of $\mathcal{F}'_\tau$ is trivial independent of the choice of the trivialization e. We can modify the construction to obtain non-trivial FLK classes as follows. We set $M = S^3 \times S^1 = (\mathbb{C}^2 \setminus \{(0,0)\})/\times 2$. As the vector field X commutes with homothecies, $\mathcal{F}_\tau$ induces a holomorphic foliation of M, which we call $\mathcal{G}_\tau$. We define a function ρ_m, where $m \in \mathbb{Z}$, on $\mathbb{C}^2 \setminus \{(0,0)\}$ by setting

$$\rho_m(z,w) = \exp\left(\frac{m\pi\sqrt{-1}}{\log 2}\log(|z|^2 + |w|^2)\right).$$

Then, ρ_m naturally induces a function on M, which we call ρ_m by abuse of notations. We set

$$e_m = \rho_m \frac{1}{|z|^2 + |w|^2}\left(\frac{\bar{w}}{\mu}\frac{\partial}{\partial z} - \frac{\bar{z}}{\lambda}\frac{\partial}{\partial w}\right).$$

Let ω_m denote the dual of e_m. Then, $\omega_m = (1/\rho_m)\omega_\tau$. Note that e_m and ω_m naturally induce trivializations of $Q(\mathcal{G}_\tau)$ and $Q(\mathcal{G}_\tau)^*$. If we set

$$\begin{aligned} \eta_m &= \eta_\tau + \frac{d\rho_m}{\rho_m} \\ &= \eta_\tau + \frac{m\pi\sqrt{-1}}{\log 2}\frac{1}{|z|^2 + |w|^2}(\bar{z}dz + zd\bar{z} + \bar{w}dw + wd\bar{w}), \end{aligned}$$

then we have $d\omega_m = -\eta_m \wedge \omega_m$. Hence we have

$$\begin{aligned} \dot{\eta_m} \wedge \eta_m \wedge d\eta_m &= \dot{\eta_\tau} \wedge \left(\eta_\tau - \frac{d\rho_m}{\rho_m}\right) \wedge d\eta_\tau \\ &= \frac{d\rho_m}{\rho_m} \wedge \dot{\eta_\tau} \wedge d\eta_\tau \\ &= -\frac{m\pi\sqrt{-1}}{\log 2}\frac{1}{(|z|^2 + |w|^2)^2}\left(1 - \frac{1}{\tau^2}\right) dz \wedge d\bar{z} \wedge dw \wedge d\bar{w}. \end{aligned}$$

If we adopt $\frac{-1}{8\pi^2 \log 2}\frac{1}{(|z|^2+|w|^2)^2} dz \wedge d\bar{z} \wedge dw \wedge d\bar{w}$ as a volume form of M with respect to which the volume of M is equal to 1, and if we normalize the FLK class by multiplying $(-1/2\pi\sqrt{-1})^3$, we have $\mathrm{FLK}(\mathcal{G}_\tau; e_m) = m(1 - 1/\tau^2) \in H^4(M;\mathbb{C})$.

Example 3.7 is generalized as follows.

Proposition 3.8 (Proposition 3.1 of [4]). *Let $\mathcal{F}$ be a transversely holomorphic foliation of a manifold M of which the complex codimension is equal to q. We assume that $\bigwedge^q Q(\mathcal{F})$ is trivial with a trivialization e. Suppose that $\mathrm{FLK}_\mu(\mathcal{F};e)$ is trivial in $H^{2q+2}(M;\mathbb{C})$ and that $D_\mu\mathrm{Bott}(\mathcal{F})$ is non-trivial in $H^{2q+1}(M;\mathbb{C})$. Let $\mu^\circ = \pi^*\mu$ be the pull-back of $\mu \in H^1(M;\Theta_\mathcal{F})$ and e_m the trivialization of $\bigwedge^q Q(\mathcal{F}^\circ)$ defined by $e_m = t^m e$, where e on the right hand side denotes the pull-back of the trivialization e of $\bigwedge^q Q(\mathcal{F})$. Then, $\mathrm{FLK}_{\mu^\circ}(\mathcal{F}^\circ;e_m)$ is non-trivial in $H^{2q+2}(M \times S^1;\mathbb{C})$ if $m \neq 0$. More precisely, we have $\pi_!\mathrm{FLK}_{\mu^\circ}(\mathcal{F}^\circ;e_m) = mD_\mu\mathrm{Bott}(\mathcal{F})$, where $\pi_!$ denotes the integration along the fiber. In particular, $H^2(M \times S^1;\Theta_{\mathcal{F}^\circ}) \neq 0$.*

Remark 3.9. It seems difficult to construct a real version of Example 3.7, because the construction depends on the existence of non-homotopic trivializations of $\bigwedge^q Q(\mathcal{G}_\tau)$.

Geometric or dynamical meanings the FLK class is always in question. We describe some relationship of the FLK class with the Bott class and its derivative for transversely holomorphic foliations. Let $\mathcal{F}$ be a transversely holomorphic foliation of complex codimension q. The foliation $\mathcal{F}$ is denoted by $\mathcal{F}_\mathbb{R}$ when we regard $\mathcal{F}$ as a real foliation by forgetting the transverse holomorphic structure. Note that $\mathcal{F}_\mathbb{R}$ is naturally transversely oriented so that $\mathrm{FLK}^\mathbb{R}(\mathcal{F}_\mathbb{R})$ is naturally defined, where $\mathrm{FLK}^\mathbb{R}$ denotes the FLK class for real foliations for clarification. Recall that the FLK class for real foliations is independent of the choice of trivializations once the transverse orientation is fixed. It is known that we can define a certain 1-form u_1 with the following properties.

(1) If we set $v_1 = du_1$, then v_1 represents the first Chern class of $Q(\mathcal{F})$.
(2) Let $\mu \in H^1(M;\Theta_\mathcal{F})$. If $\dot{u}_1$ denotes the derivative of u_1 with respect to μ, then $D_\mu\mathrm{Bott}(\mathcal{F})$ is represented by $\dot{u}_1 v_1^q$.
(3) If we set $h_1 = \sqrt{-1}(u_1 - \bar{u}_1)$ and $c_1 = dh_1$, then $h_1 c_1^{2q}$ represents $\mathrm{GV}(\mathcal{F}_\mathbb{R})$.
(4) If we set $\dot{h}_1 = \sqrt{-1}(\dot{u}_1 - \dot{\bar{u}}_1)$, then $\dot{h}_1 h_1 c_1^{2q}$ represents $\mathrm{FLK}_{\mu_\mathbb{R}}(\mathcal{F}_\mathbb{R})$, where $\mu_\mathbb{R} \in H^1(M;\Theta_{\mathcal{F}_\mathbb{R}})$ is naturally induced by μ.
(5) If we in addition assume that $\bigwedge^q Q(\mathcal{F})$ is trivial, then $u_1 v_1^q$ represents $\mathrm{Bott}(\mathcal{F})$.

In the above assertions, the representatives might differ up to real multiplicative constants depending on literatures. We also omitted to write '$\wedge$'.

Suppose now that $\bigwedge^q Q(\mathcal{F})$ is trivial. Then, we have

$$\begin{aligned}\mathrm{FLK}^\mathbb{R}_{\mu_\mathbb{R}}(\mathcal{F}_\mathbb{R}) &= \dot{h}_1 h_1 c_1^{2q} \\ &= -(\dot{u}_1 - \dot{\bar{u}}_1)(u_1 - \bar{u}_1)v_1^q \bar{v}_1^q \\ &= -(\dot{u}_1 u_1 v_1^q \bar{v}_1^q + \dot{\bar{u}}_1 \bar{u}_1 \bar{v}_1^q v_1^q) + (\dot{u}_1 v_1^q \bar{u}_1 \bar{v}_1^q + \dot{\bar{u}}_1 \bar{v}_1^q u_1 v_1^q)\end{aligned}$$

up to real multiplicative constants. As v_1 represents the first Chern class, we have $v_1 = \bar{v}_1$ at the cohomology level. On the other hand, we have $v_1^{q+1} = 0$

as differential forms thanks to the Bott vanishing theorem. Hence we have $(\dot{u}_1 u_1 v_1^q)\bar{v}_1^q = (\dot{u}_1 u_1 v_1^q)v_1^q = \dot{u}_1 u_1 v_1^{2q}$ at the cohomology level. The most right hand side is equal to 0. Therefore, we have

$$\begin{aligned}\mathrm{FLK}^{\mathbb{R}}_{\mu_{\mathbb{R}}}(\mathcal{F}_{\mathbb{R}}) &= (\dot{u}_1 v_1^q \bar{u}_1 \bar{v}_1^q + \dot{\bar{u}}_1 \bar{v}_1^q u_1 v_1^q) \\ &= D_\mu \mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})} + \overline{D_\mu \mathrm{Bott}(\mathcal{F})}\,\mathrm{Bott}(\mathcal{F}).\end{aligned}$$

Thus the FLK class of $\mathcal{F}_{\mathbb{R}}$ with respect to $\mu_{\mathbb{R}}$ is related with the real part of $D_\mu \mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})}$. The imaginary part of $D_\mu \mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})}$ appears as the derivative of $\mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})}$. Indeed, we have

$$\begin{aligned}D_\mu(\mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})}) &= D_\mu \mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})} + \mathrm{Bott}(\mathcal{F})\,\overline{D_\mu \mathrm{Bott}(\mathcal{F})} \\ &= D_\mu \mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})} - \overline{D_\mu \mathrm{Bott}(\mathcal{F})}\,\mathrm{Bott}(\mathcal{F}).\end{aligned}$$

Remark 3.10. A typical example of a transversely holomorphic foliation such that $\bigwedge^q Q(\mathcal{F})$ is trivial is obtained as follows. Let $\mathcal{G}$ be a transversely holomorphic foliation of M and E the principal S^1-bundle associated with $\bigwedge^q Q(\mathcal{G})$. Let $\mathcal{F}$ be the foliation of E obtained by pulling-back $\mathcal{F}$. Then $\mathcal{F}$ is such an example. In this case, it is known that the integration of $\mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})}$ along the fiber is equal to $\mathrm{GV}(\mathcal{G})$. Indeed, it can be shown that u_1 and $\bar{u}_1$ correspond to the Thom class, and

$$\int \mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})} = \int u_1 v_1^q \bar{u}_1 \bar{v}_1^q = v_1^q \bar{u}_1 \bar{v}_1^q - u_1 v_1^q \bar{v}_1^q = (u_1 - \bar{u}_1) v_1^q \bar{v}_1^q = \mathrm{GV}(\mathcal{G})$$

holds up to multiplicative constants (cf. Theorem 3.3.10 of [1]). Note that u_1 and $\bar{u}_1$ are not necessarily well-defined on M but $(u_1 - \bar{u}_1)$ is well-defined on M. Taking the derivative, we obtain

$$\int D_\mu(\mathrm{Bott}(\mathcal{F})\,\overline{\mathrm{Bott}(\mathcal{F})}) = (q+1)(\dot{u}_1 - \dot{\bar{u}}_1) v_1^q \bar{v}_1^q.$$

It is known that $\dot{u}_1$ and $\dot{\bar{u}}_1$ are well-defined on M so that we have $\dot{u}_1 v_1^q \bar{v}_1^q = (\dot{u}_1 v_1)^q v_1^q = 0$ and $\dot{\bar{u}}_1 \bar{v}_1^q v_1^q = 0$ for the same reason as above. This is also a consequence of Theorem 2.6. On the other hand, we have

$$\begin{aligned}\int \mathrm{FLK}^{\mathbb{R}}_{\mu_{\mathbb{R}}}(\mathcal{F}_{\mathbb{R}}) &= \int (\dot{u}_1 v_1^q \bar{u}_1 \bar{v}_1^q + \dot{\bar{u}}_1 \bar{v}_1^q u_1 v_1^q) \\ &= -\dot{u}_1 v_1^q \bar{v}_1^q + \dot{\bar{u}}_1 \bar{v}_1^q v_1^q \\ &= 0.\end{aligned}$$

Hence we cannot obtain a non-trivial example of the FLK class in this way.

References

[1] T. Asuke, *Godbillon–Vey class of transversely holomorphic foliations*, MSJ memoirs **24**, June 2010.

[2] T. Asuke, *Transverse projective structures of foliations and infinitesimal derivatives of the Godbillon–Vey class*, Internat. J. Math. **26** (2015), 1540001, 29pp.

[3] T. Asuke, *Derivatives of secondary classes and 2-normal bundles of foliations*, J. Math. Sci. Univ. Tokyo **22** (2015), The special issue for the 20th anniversary, 893–937.

[4] T. Asuke, *On a characteristic class associated with deformations of foliations*, preprint, arXiv:2011.04340.

[5] R. Bott, *On the Lefschetz formula and exotic characteristic classes*, Symposia Mathematica, Vol. X (Convegno di Geometria Differenziale, INDAM, Rome, 1971), Academic Press, London, 1972, 95–105.

[6] A. Candel, L. Conlon, *Foliations I*, Grad. Stud. Math. **23**, Amer. Math. Soc., Providence RI, 2000.

[7] T. Duchamp, M. Kalka, *Deformation theory for holomorphic foliations*, J. Differential Geom. **14** (1979), 317–337.

[8] I. M. Gel'fand, B. L. Feĭgin, D. B. Fuks, *Cohomology of the Lie algebra of formal vector fields with coefficients in its adjoint space and variations of characteristic classes of foliations*, Funct. Anal. and its Appl. **8** (1974), 99–112, translated from Funkcional. Anal. i Priložen. **18** (1974), 13–29.

[9] D. B. Fuks, *Cohomology of infinite-dimensional Lie algebras*, Contemporary Soviet Mathematics, Consultants Bureau, New York, 1986, Translated from the Russian by A. B. Sosinskiĭ.

[10] J. Girbau, A. Haefliger, D. Sundararaman, *On deformations of transversely holomorphic foliations*, Journal für die reine und angewandte Mathematik **345** (1983), 122–147.

[11] C. Godbillon, J. Vey, *Un invariant des feuilletages de codimension* 1, C. R. Acad. Sci. Paris Sér. A-B **273** (1971), A92–A95.

[12] J. Hebda, C. Roberts, *Examples of Thomas–Whitehead projective connections*, Differ. Geom. Appl. **8** (1998), 87–104.

[13] J. L. Heitsch, *Deformations of secondary characteristic classes*, Topology **12** (1973), 381–388.

[14] J. L. Heitsch, *A cohomology for foliated manifolds*, Comment. Math. Helvetici **50** (1975), 197–218.

[15] J. Heitsch, *Independent variation of secondary classes*, Ann. of Math. **108** (1978), 421–460.

[16] J. Heitsch, *Secondary invariants of transversely homogeneous foliations*, Michigan Math. J. **33** (1986), 47–54.

[17] S. Hurder, *Independent rigid secondary classes for holomorphic foliations*, Invent. Math. **66** (1982), 313–323.

[18] F. W. Kamber, P. Tondeur, *On the linear independence of certain cohomology classes of* $B\Gamma_q$, Studies in algebraic topology, Adv. in Math. Suppl. Stud. **5**, Academic Press, New York-London, 1979, pp. 213–263.

[19] D. Kotschick, *Godbillon–Vey invariants for families of foliations*, Symplectic and contact topology: interactions and perspectives (Toronto, ON/Montreal, QC, 2001), Fields Inst. Commun. **35**, Amer. Math. Soc., Providence, RI, 2003, 131–144.

[20] C. Lazarov, *A permanence theorem for exotic classes,* J. Differential Geometry **14** (1979), 475–486.

[21] C. Lazarov, J. Pasternack, *Secondary characteristic classes for riemannian foliations*, J. Differential Geometry **11** (1976), 365–385.

[22] J. M. Lodder, *Rigidity of secondary characteristic classes*, Differential Geom. Appl. **12** (2000), 207–218.

[23] T. Maszczyk, *Foliations with rigid Godbillon–Vey class*, Math. Z. **320** (1999), 329–344.

[24] P. Molino, *Riemannian foliations*, Progress in Math. **73**, Birkhäuser, Boston, 1988.

[25] S. Morita, *On characteristic classes of conformal and projective foliations*, J. Math. Soc. Japan **31** (1979), 693–718.

[26] C. W. Roberts, *The projective connections of T. Y. Thomas and J. H. C. Whitehead applied to invariant connections*, Differ.. Geom. Appl. **5** (1995), 237–255.
[27] W. Thurston, *Noncobordant foliations of S^3*, Bull. Amer. Math. Soc. **78** (1972), 511–514.

University of Tokyo
Komaba 3-8-1, Meguro-ku, Tokyo 153-8914, Japan
asuke@ms.u-tokyo.ac.jp

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

A POLISH GROUP CONTAINING A HAAR NULL F_σ-SUBGROUP THAT CANNOT BE ENLARGED TO A HAAR NULL G_δ-SET

TARAS BANAKH

ABSTRACT. Answering a question of Elekes and Vidnyánszky, we construct a Polish meta-Abelian group H and a subgroup $F \subset H$, which is a Haar null F_σ-set in H that cannot be enlarged to a Haar null G_δ-set in H.

KEYWORDS: Haar null set, Polish group, meta-Abelian group

MSC2010: 22A10; 28C10

Received 5 April 2021; revised 23 June 2021; accepted 10 July 2021

A Borel subset B of a Polish group H is called *Haar null* if there exists a σ-additive Borel probability measure μ on H such that $\mu(xBy) = 0$ for all $x, y \in H$. It is well-known [2] that a Borel subset B of a Polish locally compact group H is Haar null if and only if B has Haar measure zero if and only if B is contained in a Haar null G_δ-subset of H. Mycielski [5, P_1] asked whether every Haar null subset of a Polish group H can be enlarged to a Haar null G_δ-set in H. This question of Mycielski was answered in negative by Nagy [6] who constructed in the Polish Abelian group $\mathbb{Z}^\omega$ a Haar null subset which is a difference of two G_δ-sets but cannot be covered by a Haar null G_δ-subset of $\mathbb{Z}^\omega$. On the other hand, Elekes and Vidnyánszky [4] constructed in every non-locally compact Polish Abelian group H a Haar null Borel subset $B \subset H$ which is not contained in a Haar null G_δ-set in H. However, the construction of the Borel set B exploited in [4] does not allow to evaluate the Borel complexity of B. Because of that, Elekes and Vidnyánszky asked in [4] whether every Haar null F_σ-subset B of a Polish (Abelian) group H can be enlarged to a Haar null G_δ-set. In this paper we partly answer this problem of Elekes and Vidnyánszky presenting an example of a Polish meta-Abelian group

 https://doi.org/10.2478/9788366675360-002.

H and a Haar null F_σ-subgroup $F \subset H$ that cannot be enlarged to a Haar null G_δ-set in H.

A topological group H is *meta-Abelian* if H contains a closed normal Abelian subgroup $A \subseteq H$ such that the quotient group H/A is Abelian.

We define a subset B of a topological group H to be *thick* if for every compact subset $K \subseteq H$ there are points $x, y \in H$ such that $xKy \subseteq B$. It is easy to see that a thick Borel subset of a Polish group cannot be Haar null.

Theorem 1. *There exists a Polish meta-Abelian group H containing a subgroup $F \subset H$ such that F is a Haar null F_σ-set in H but every G_δ-set $G \subseteq H$ containing F is thick and hence is not Haar null in H.*

Proof. Observe that the countable power $R = \mathbb{R}^\omega$ of the real line has the structure of a unital topological ring, and the dense G_δ-set $(\mathbb{R} \setminus \{0\})^\omega$ coincides with the set R^* of invertible elements of the ring R. On the product $H = R \times R^*$ consider the binary operation $\star : H \times H \to H$ defined by the formula

$$(x, a) \star (y, b) = (x + ay, ab) \quad \text{for} \quad (x, a), (y, b) \in H = R \times R^*.$$

The Polish space $H = R \times R^*$ endowed with this binary operation is a Polish group called the *semidirect product* of R and R^*.

In the topological ring R consider the F_σ-subset

$$R_0 = \{(x_n)_{n\in\omega} \in \mathbb{R}^\omega : \exists n \in \omega \;\; \forall m \geq n \;\; x_m = 0\}$$

consisting of all eventually zero sequences. It follows that R_0 is a subring of R and $F = R_0 \times R^* \subset H$ is a subgroup of the Polish group H. We claim that the subgroup F has the desired properties.

Lemma 2. *The subset F is Haar null in H.*

Proof. Taking into account that R_0 is a Borel non-open subgroup of the Polish Abelian group R, we can apply a classical result of Christensen [2, Theorem 2] and conclude that R_0 is Haar null in R. This allows us to find a probability measure μ on R such that $\mu(x + R_0) = 0$ for all $x \in R$. Identifying R with the normal subgroup $R \times \{1\}$ of H, we can consider the measure μ as a measure on H. Then for every elements $(x, a), (y, b) \in H$ we get

$$\big((x,a)\star F\star(y,b)\big)\cap\big(R\times\{1\}\big) = (x,a)\star\big(R_0\times\{a^{-1}b^{-1}\}\big)\star(y,b) = (x+aR_0+b^{-1}y, 1)$$

and hence

$$\mu\big((x,a) \star F \star (y,b)\big) = \mu(x + aR_0 + b^{-1}y) = \mu(x + b^{-1}y + R_0) = 0$$

by the choice of μ. So, the measure μ witnesses that the set $F = R_0 \times R^*$ is Haar null in H. □

Lemma 3. *Every G_δ-set $G \subseteq H$ containing F is thick.*

Proof. Given a G_δ-set $G \subseteq H$ containing F, consider its complement $H \setminus G$ and write it as a countable union $H \setminus G = \bigcup_{k\in\omega} E_k$ of an increasing sequence $(E_k)_{k\in\omega}$ of closed sets in H. To prove that G is thick in H, it suffices to find for every compact subset $K \subset H$ an element $h \in H$ such that $hKh^{-1} \subseteq G$. Given a compact set $K \subset H$ choose compact sets $C \subset R$ and $C^* \subset R^*$ such that $K \subseteq C \times C^*$. Observe that for every element $a \in R^*$ and the element $h = (0, a) \in H$ we get $h^{-1} = (0, a^{-1})$ and hence

$$hKh^{-1} \subseteq (0, a) \star (C \times C^*) \star (0, a^{-1}) \subseteq aC \times C^*.$$

Now it suffices to find an element $a \in R^*$ such that $(aC \times C^*) \cap E_k = \emptyset$ for all $k \in \omega$. The compactness of the set C^* guarantees that the projection $\mathrm{pr} : R \times C^* \to R$, $\mathrm{pr} : (x, y) \mapsto x$, is a closed map, and for every $k \in \omega$ the set $P_k = \mathrm{pr}(E_k \cap (R \times C^*))$ is closed in R and does not intersect the F_σ-subgroup R_0. Observe that the set $R^*_k = \{a \in R^* : aC \cap P_k = \emptyset\}$ is contained in $\{a \in R^* : (aC \times C^*) \cap E_k = \emptyset\}$. The compactness of the set C implies that the set R^*_k is open in R^*.

Claim 4. For every $k \in \omega$ the set R^*_k is dense in R^*.

Proof. Given any element $a = (a_n)_{n\in\omega} \in R = \mathbb{R}^\omega$ and any neighborhood $O_a \subseteq R$, find $m \in \omega$ such that the closed subspace $\{(b_n)_{n\in\omega} \in \mathbb{R}^\omega : \forall n < m \;\; b_n = a_n\}$ is contained in O_a. Find a sequence $(C_n)_{n\in\omega}$ of compact subsets of the real line such that $C \subseteq \prod_{n\in\omega} C_n$. Consider the compact space $\Pi = \prod_{n<k} a_n C_n$ and observe that the projection $\pi : \Pi \times \mathbb{R}^{\omega\setminus k} \to \mathbb{R}^{\omega\setminus k}$, $\pi : (x, y) \mapsto y$, is a closed map (here we identify the natural number k with the set $\{0, \dots, k-1\}$). This implies that the set $\tilde{P}_k = \pi\big((\Pi \times \mathbb{R}^{\omega\setminus k}) \cap P_k\big)$ is closed in $\mathbb{R}^{\omega\setminus k}$ and does not contain the constant zero function $z \in \mathbb{R}^{\omega\setminus k}$. By the compactness of the product $C_{\geq k} = \prod_{n\geq k} C_n \subseteq \mathbb{R}^{\omega\setminus k}$ and the continuity of multiplication in the topological ring $\mathbb{R}^{\omega\setminus k}$, there exists an element $\varepsilon \in (\mathbb{R} \setminus \{0\})^{\omega\setminus k}$ so close to zero that $(\varepsilon \cdot C_{\geq k}) \cap \tilde{P}_k = \emptyset$. Then for the element $b = (b_n)_{n\in\omega} \in O_a$ defined by $b_n = a_n$ for $n < k$ and $b_n = \varepsilon_n$ for $n \geq k$ we get

$$(bC) \cap P_k \subseteq \Big(\prod_{n\in\omega} b_n C_n\Big) \cap P_k \subseteq \Big(\prod_{n<k} a_n C_n\Big) \times \Big(\big(\prod_{n\geq k} \varepsilon_n C_n\big) \cap \tilde{P}_k\Big) = \emptyset$$

and hence $b \in O_a \cap R^*_k$. □

Claim 4 combined with the Baire Theorem guarantees that $\bigcap_{k\in\omega} R^*_k$ is a dense G_δ-set in the Polish space R^*. Then we can take any point $a \in \bigcap_{k\in\omega} R^*_k$ and conclude that $(aC \times C^*) \cap \bigcup_{k\in\omega} E_k = \emptyset$ and hence for the element $h = (0, a)$ we have $hKh^{-1} \subseteq G$. □

□

However, Theorem 1 gives no answer to the following two problems posed by Elekes and Vidnyánszky in [4].

Problem 5. Is each Haar null F_σ-subset of an uncountable Polish Abelian group G contained in a Haar null G_δ-subset of G?

Problem 6. Is each countable subset of an uncountable Polish group G contained in a Haar null G_δ-subset of G?

By [7] (see also [1, §4] and [4, Remark 5.3]), the answer to Problem 6 is affirmative for Polish Abelian groups. An interesting discussion of Problems 5 and 6 can be found in [3, §5.4].

References

[1] T. Banakh, S. Głąb, E. Jabłońska, J. Swaczyna, *Haar-$\mathcal{I}$ sets: looking at small sets in Polish groups through compact glasses*, Dissert. Math. (accepted).

[2] J. Christensen, *On sets of Haar measure zero in Abelian Polish groups*, Israel J. Math. **13** (1972), 255–260.

[3] M. Elekes, D. Nagy, *Haar null and Haar meager sets: a survey and new results*, Bull. Lond. Math. Soc. **52**:4 (2020), 561–619.

[4] M. Elekes, Z. Vidnyánszky, *Haar null sets without G_δ-hulls*, Israel J. Math. **209** (2015) 199–214.

[5] J. Mycielski, *Some unsolved problems on the prevalence of ergodicity, instability, and algebraic independence*, Ulam Quarterly, **1**:3 (1992), 30–37.

[6] D. Nagy, *Low-complexity Haar null sets without G_δ hulls in $\mathbb{Z}^\omega$*, Fund. Math. **246** (2019), 275–287.

[7] S. Solecki, *Haar null and non-dominating sets*, Fund. Math. **170**:1-2 (2001), 197–217.

Jan Kochanowski University in Kielce, Poland
& Ivan Franko National University of Lviv, Ukraine
taras.banakh@ujk.edu.pl, t.o.banakh@gmail.com

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

TEACHERS' PERSPECTIVE ON THE ADVANTAGES AND DISADVANTAGES OF POSTGRADUATE STUDIES ALLOWING TO TEACH MATHEMATICS AS A SECOND SUBJECT

BARBARA BARAŃSKA & MAŁGORZATA ZAMBROWSKA

Abstract. Various researchers in mathematics education have addressed the issue of competencies for teaching mathematics in schools. Much has already been said about the different types of knowledge teachers should possess, as well as about the relationship between teachers' competencies and the achievement of their students. However, there is also a body of research dealing with the issue of qualifications for teaching mathematics as a school subject. It appears that in many countries around the world the phenomenon of 'out of field teaching' is quite common and is gaining increasing attention. In this text, we look at the phenomenon of postgraduate studies that allow Polish teachers of various school subjects to obtain qualifications to teach mathematics as a second or further subject. The results of a survey conducted among 160 teachers who have obtained qualifications for teaching mathematics on such studies allow us to see what positive and negative sides of postgraduate studies the teachers themselves indicate. Analyzing the results of the questionnaire, we tried to determine how the completion of these studies has affected, positively or not, various aspects of teachers' lives and functioning.

Keywords: postgraduate studies qualifying for teaching mathematics as another subject, advantages and disadvantages of postgraduate studies

MSC2010: 97B30, 97B50, 97A40, 97C60

Received 30 April 2021; revised 10 August 2021; accepted 20 August 2021

1. The phenomenon of 'out of field teaching'

A broad body of literature addresses the problem of teaching a subject without necessary qualifications which is called out-of-field teaching (OOFT) or teaching across specialisations (TAS). Surprisingly, this phenomenon is widely spread

 https://doi.org/10.2478/9788366675360-003.

across the globe. Its occurrence has been noticed and described, for example, by researchers from Germany ([2]), Ireland ([11]), UK ([5]), USA ([8],[9],[13]) and Australia ([6]). Several factors that give rise to the OOFT are mentioned in the literature. The most frequently pointed, and seemingly most obvious source of this phenomenon, are shortages of qualified teachers in particular subjects. It is stated that they hit schools depending on, e.g. their location ([6]). For instance, sometimes it is very difficult to find qualified teachers to all school subjects in small towns and villages. Also the type of school can be a critical factor ([14]). It may be more difficult for example to find teachers to work in special education with students having specific learning needs. Our own observations show that graduates of the teaching specialization in mathematics, despite being fully qualified to teach at any stage, prefer to start working as primary school teachers, which they find easier. Regarding school subjects, staff shortages have been observed frequently with respect to mathematics and science teachers ([6]). All these problems, combined with the desire of school principals to maintain continuity in education, contribute to the decisions on assigning teachers to subjects for which they do not have adequate education and preparation (e.g., [9]). However, in ([8]) organizational structure of schools and the occupational conditions and characteristics of teaching are shown to be perhaps even more important sources of the occurrence of out-of-field teaching. The very fact that teaching a subject can be entrusted to any un(der)qualified teacher ([14]) shows that the teaching profession has been deprived of its status and is no longer seen as a profession requiring expert knowledge and competence. Ingersoll says ([8], p. 776):

> Few would require cardiologists to deliver babies, real estate lawyers to defend criminal cases, chemical engineers to design bridges, or sociology professors to teach English. The commonly held assumption is that such traditional professions require a great deal of skill and training and that, hence, specialization is necessary. In contrast, the commonly held assumption is that teaching in elementary and secondary schools requires far less skill, training, and expertise.

According to ([11], p. 292), to perceive that teachings skills can be obtained without formal preparation undermines and underestimates the complexity of the profession. Hobbs & Törner ([7], p.4) add that being a specialist in a subject and its teaching means:

> to become immersed and expert in a defined and bounded body of knowledge and skills such that there is coherence, connectedness and flexibility to what is known and what one can do.

This point of view corresponds with that of Törner & Törner ([14], p. 203) who notice that teachers who teach OOF lack an adequate grasp of the subject they teach:

> It is not primarily missing elements of knowledge that exacerbate the lessons, but rather the (on a meta-level) insufficiently detailed 'maps' of the subject. Though these teachers indeed know viable ways to solve the canonical problems, however, they are afraid of variations and digressions into an in this case frequently unknown territory.

The complexity and multidimensionality of teaching a subject without adequate qualifications and preparation is likely to cause anxiety, tension and discomfort in OOF teachers. When discussing this phenomenon, researchers have already raised questions about teacher's identity, self-efficacy and well-being.

It is worth to mention that mathematics is considered a subject suitable for OOFT ([14]). Research reports show that it is perceived positively by non-qualified teachers who welcome teaching mathematics as a new, interesting challenge ([Ibid.]). Also those who teach out-of-field are found to be "more on the learners' side" ([Ibid.]). It is because not long ago they were learning the material that their students now have to deal with. As a result, these teachers can better empathize with their students' and may understand (but unfortunately also share) the difficulties they are experiencing.

2. Teaching mathematics as a second subject in Poland

It is difficult to say what is the scale of the OOFT phenomenon in Poland today. To our knowledge, neither the Central Statistical Office nor the Ministry of Education and Science have any data on this matter. According to the Eurydice report from 2002 ([4]), the prevalence of OOFT in Poland was negligible at that time. When it is not possible to employ a suitably qualified person as a teacher of a particular subject, the school principal may employ a person without such qualifications or without pedagogical preparation. Although such emergency measures do not solve the problem of staff shortages, they do allow for the continuity of education. There are many reasons to believe that currently in Poland the problem of lack of qualifications has been solved by means of widely available postgraduate studies, where one can obtain pedagogical preparation to teach or qualifications to teach another subject.

According to the Regulation of the Polish Minister For Science And Higher Education (25 July 2019) on the standard of teachers' training:

> Postgraduate studies preparing to the profession of teaching may be provided as preparation for: (...) the substantive and didactic requirements for teaching another subject (...) for graduates of first-cycle and second-cycle programs or long-cycle studies in fields of study whose curricula specified learning outcomes including knowledge and skills corresponding to general requirements of the core curriculum of a given subject (...) having substantial,

> psychological and pedagogical preparation, preparation in the field of basic didactics and voice emission as well as didactic preparation for teaching a subject (...) Postgraduate studies preparing for the teaching professions shall last not less than 3 semesters.

Prior regulations in this area placed no restrictions on the field of previously completed studies. This allowed teachers to obtain qualifications to teach more subjects, even those completely different from their original education. For instance, a physical education teacher or a catechist could have gained qualifications to teach mathematics in just two semesters, like as if it was neither too hard, nor too demanding to master a new area within such a short period of time ([10]). Despite the declarations that the educational results will be similar to those achieved at full-time studies, participants of postgraduate studies have much more difficult and less comfortable conditions for solid work. They have to match professional duties and personal commitments with learning of a new discipline and they undergo intensive preparation for their new profession during weekend classes. Studies that are supposed to last three semesters, in practice may be completed within a single year, e.g., when the summer break period is considered the third semester. Gaining the knowledge and skills to become confident and build a sense of competence will take a lot of time and a lot of own work. Post-graduate studies offer too little time to get a complete preparation for teaching a new subject, but they allow to learn a little about the subject. For these reasons, we regard such teachers who decide to teach mathematics as a second subject as being **neither fully „in", nor completely „out"** of the field of mathematics ([1]).

3. Methodology of the study of teachers' indications of positive and negative sides of postgraduate studies

In this paper we present only part of the data collected in a broader study on the evaluation of the postgraduate studies qualifying for teaching mathematics as a second subject. Our research instrument consisted of: 8 metric questions, 7 questions strictly related to the postgraduate studies, 3 questions on the teaching experience after completing postgraduate studies and 2 supplementary questions (for more details see: [1]). In two questions we asked the teachers to indicate positive and negative sides of the studies qualifying for teaching mathematics as a second subject. We have intentionally been unspecific in these questions and we have made no comments on what we expected the 'positive/negative sides' to mean. We left the respondents freedom to answer according to their own interpretation. From the answers we obtained, we wanted to extract information about the advantages and disadvantages of completing postgraduate studies the teachers themselves indicate, i.e. we wanted to learn what positive or negative impact completing these studies had on various aspects of their lives. The study was conducted with the use of an electronic questionnaire form, the link to which was made available to teachers. The link has been sent to all primary and secondary schools across Poland (email addresses of almost 24 000 schools have been retrieved from the website: https://rspo.men.gov.pl/) and also to the Association of Mathematics Teachers. We also tried to reach teachers through the publishers of school math textbooks and through social networking websites. The data was collected within the period of 8 weeks (February-April 2020).

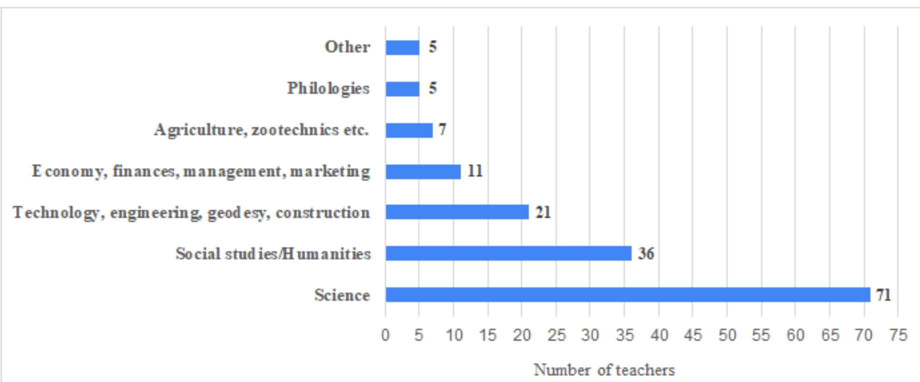
We performed our analysis using data driven, two-cycle coding ([12]) in order to develop the categories of collected answers.

Participants

Participation in the study was voluntary so only the willing teachers took part in it. Therefore, the surveyed group was not randomly selected and thus cannot be treated as representative. The group of teachers participating in the study consisted of 160 respondents who declared that they had completed postgraduate studies qualifying them to teach mathematics as another subject. The vast majority (87,5%) of the participants were women, which is not surprising since in Poland the teaching profession is mostly taken by women ([3]). The table below outlines the age groups of the respondents:

Age (years)	Percentage of the group
25-30	7%
31-40	19%
41-50	42%
51-60	28%
Above 60	4%

We wanted to know the participants' primary field of expertise, hence we asked them what discipline they graduated in. Seven people gave only the name of their university, which did not clearly indicate the discipline. Some people, however, declared having graduated from more than one department. The diagram below shows the total of 156 answers given:



Many of the respondents have graduated in science, technology or economics-related subjects, but the two most represented groups were science (about 44% of the group) and social studies/humanities subjects (about 22%). We have included in science disciplines such as chemistry (30), physics (24), biology or biology with chemistry (6), geography (6) and informatics (5) teachers. Social studies/Humanities category included graduates in pedagogy (33), history (1), psycho-pedagogical counseling (1) and revalidation (1).

We also asked the participants about the number of years they have worked as teachers and the number of years of service as mathematics teachers. At the time when this study was conducted, more than a half of the surveyed teachers had more than 20 years of teaching experience. Detailed data on the number of years of service are presented in the table below:

Years of service as a teacher	**Percentage of the group**
0-5	8%
6-10	10%
11-15	12%
16-20	19%
Above 20	51%

The number of years that the respondents had worked as mathematics teachers were the following:

Years of service as a math teacher	**Percentage of the group**
0-5	42%
6-10	16%
11-15	12%
16-20	18%
Above 20	12%

Unfortunately, our data do not allow us to determine after how many years of teaching the respondents began teaching mathematics.

Research questions

In this study we wanted to find out what positive and negative sides of postgraduate studies qualifying to teach mathematics as a second subject the graduates of these studies acknowledge and indicate. It was of our particular interest to investigate the advantageous or disadvantageous impact that the completion of these studies had on the teachers' lives.

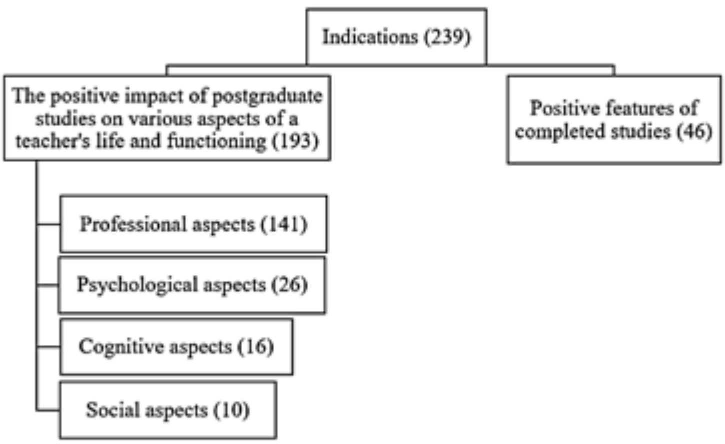
4. The study results

The table below shows the number of teachers who gave answers of different type to the questions on positive and negative sides of the postgraduate studies they have completed:

POSITIVE sides of postgraduate studies	**NEGATIVE sides of postgraduate studies**
Responses • No response - 6 persons, • Answer ”No positive sides” - 2 persons, • Vague, ambiguous answers (excluded from the analysis) - 2 persons, • Answers including indication of positive sides - 150 persons.	Responses • No response - 15 persons, • Answer ”No negative sides” - 76 persons, • Vague, ambiguous answers (excluded from the analysis) - 3 persons, • Answers including indication of negative sides - 66 persons.

In the whole group there were 5 respondents who did not provide their answers to any of these two questions, 1 person who listed no positive sides and gave a non-informative answer to the other question, and 10 teachers who marked some positive sides of the studies and gave no answer to the question about the negative ones.

4.1. **POSITIVE sides of the studies.** Among the 150 teachers who indicated positive sides of the completed studies, 127 referred to the positive impact the studies had on various aspects of their lives and functioning, and 30 teachers referred to the positive features of the completed studies. The teachers gave 239 indications altogether. After initial coding of the raw data, we grouped the answers into the categories shown below:



4.1.1. *Professional aspects.* Within the categories of responses indicating professional aspects, 56 referred to professional stability, 45 were a clear reference to the opportunity of teaching mathematics and 36 addressed general professional development and upskilling. Teachers who found completed postgraduate studies contributing to their professional stabilisation, gave 29 indications related to their work and 27 particularly referring to the teaching hours. The prevailing themes in the statements of respondents who referred to the issue of labour were as follows: getting or keeping a job, job security, ease of getting a job, ability to choose where to work, ability to do a job other than determined by one's primary education, working at school, being attractive on the labor market. Below we give some examples of what the respondents wrote in this category (the numbers in brackets are the individual numbers assigned to the respondents as part of coding the answers):

> (109) *I finally have a job.*
> (1) *Being able to work at school.*
> (108) *Keeping my job.*
> (8) *Due to the reduction of biology hours in high school, I would be left without a job.*
> (126) *Keeping a job at school. In small schools, there is a lack of hours for one subject.*
> (15) *I am attractive in the job market because I can teach chemistry and mathematics.*
> (145) *I am more competitive in the labor market.*
> (117) *More employment opportunities.*
> (43) *Today I can teach mathematics in all classes of primary school even secondary school. Which gives me a wider choice of work.*

A group of teachers who gave 27 responses referring to the teaching hours indicated the possibilities of getting a full-time position at school or filling in missing hours. Examples of statements made by these teachers are the following:

> (2) *I had a full-time job rather than a few hours.*
> (21) *More hours in school.*
> (32) *Getting teaching hours instead of caring, educating hours.*
> (55) *Being able to have more hours in one job.*
> (79) *After liquidation of gymnasium I can teach in one school combining two subjects (physics and mathematics).*
> (92) *I have not been short of full-time hours for 19 years.*
> (102) *I have 1.5 full-time jobs (still).*
> (109) *I finally have a job even with overtime.*
> (127) *There is much more mathematics than science in the hourly grid.*
> (158) *I was teaching two subjects at the same time, when they reduced number of hours in physics (basic level after lower secondary school 1 hour per cycle), mathematics allowed me to work full time in one school.*

> (54) *If there are not enough hours due to the school's work organisation, I can count on supplementing my full-time position with hours from subjects I am qualified to teach.*

The teachers who gave 45 indications related strictly to the possibility of teaching mathematics, wrote about gaining qualifications for teaching mathematics, being given the opportunity to teach a subject they like and have always wanted to teach, having the opportunity to put cross-curricular correlation into practice, having the opportunity to become an examiner in mathematics, teaching a leading subject and having the opportunity to give private lessons in mathematics. Below we give a few examples of what the teachers wrote:

> (43) *Today, I can teach mathematics in all classes of primary school even secondary school.*
> (36) *Currently, while teaching grades 1-3, I can "legally" teach mathematics as a substitute in the older grades.*
> (54) *When maths teachers are absent, the principal can use my qualifications by assigning me maths hours as a substitute.*
> (96) *I can teach the subject I always wanted to teach, but life first wrote a different scenario.*
> (104) *I teach a subject I am passionate about.*
> (137) *I teach what I wanted, I love it.*
> (2) *I prefer teaching mathematics because it is easier.*
> (8) *Postgraduate studies have opened up new opportunities for me (...) to earn extra money from tutoring.*
> (69) *[The studies] help to teach programming and algorithmics.*
> (132) *Opportunity to teach 2 subjects; physics and mathematics come together.*
> (152) *Possibility to correlate maths in chemistry and vice versa (logarithms, concentrations, exponential notation).*
> (15) *I could become an OKE [Polish Regional Examination Board – BP, MZ] examiner in mathematics.*
> (21) *The leading subject on the examination.*

A group of respondents marked the importance of general professional development and obtaining some new qualifications. Here mathematics was not mentioned as the main subject of interest. Few teachers wrote about making their current job more attractive and some wrote about obtaining a diploma or simply getting a "paper". Examples of statements made by these teachers are as follows:

> (5) *I have obtained a qualification.*
> (7) *I have obtained a qualification in a second subject.*
> (25) *I am qualified to teach two subjects at the same time.*
> (34) *A paper that no one will question or challenge.*

> (94) *As a matter of fact, apart from recalling secondary school material (my studies do not qualify me to teach at this level), all I have is a "paper".*
> (139) *I have received a graduation diploma.*
> (77) *I have developed professionally, made my own work more attractive, moved on.*
> (157) *I will be able to help students with special needs who will be entrusted to me next year.*

There were also 4 other indications we could not put easily into any of the above-mentioned categories. We quote these answers below:

> (55) *More hours with a class (fewer children) compared to the same number of hours for e.g. a music teacher who has all classes and I only have 4. And everyone has the same number of hours.*
> (77) *Working with older children I got to know the next stage of education. I certainly know their capabilities better compared to teachers who have never worked with young children.*
> (30) *I work with older children. It is easier.*
> (86) *Eliminating staff shortages at school.*

The first two statements reveal that it is important for their authors to be able to better understand students' abilities and backgrounds and also build relationships with their mentees and have more frequent contact with them. The third statement shows that postgraduate studies qualifying to teach a second subject can also provide an opportunity to start working with a different age group of students. We believe that this is an aspect that is not often raised in discussions regarding such studies.

4.1.2. *Psychological aspects.* Among the answers that we categorized as indications related to psychological aspects, there were responses pointing to: self-improvement, personal growth, development of one's interests and passions, fulfilling dreams, satisfaction, sense of fulfilment, being appreciated and needed, increase in motivation to work and getting sense of security. We illustrate this group of responses with several examples:

> (58) *I have not only consolidated my knowledge during my studies, but also my conviction that I meet the requirements methodologically.*
> (71) *I feel valued and needed at school.*
> (88) *My studies have broadened the scope of my mathematical knowledge and given me more confidence and freedom to teach.*
> (109) *I am constantly learning something new, the satisfaction of completing these difficult studies at a mature age.*
> (116) *Thanks to the diploma, no one will accuse me of incompetence. I feel more confident at work.*

4.1.3. *Cognitive aspects.* Only 16 indications were directly related to the general cognitive aspects of the completion of studies. References to the broadening and acquiring knowledge took a form of statements like, for instance:

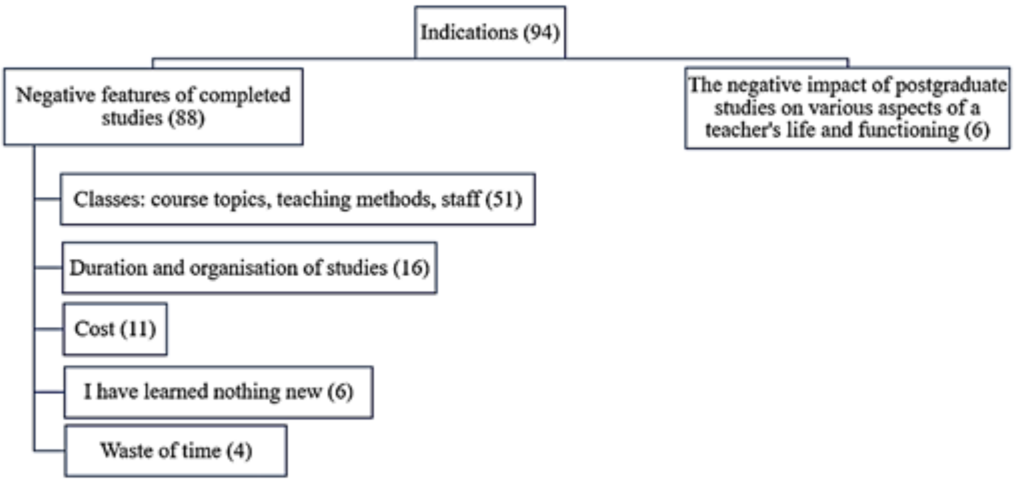
(22) *Reinforcement and extension of existing mathematical knowledge and skills.*
(70) *Additional knowledge about the school system, teacher and working with children.*
(108) *Acquisition of knowledge in advanced mathematical and computing subjects.*
(118) *Advanced mathematical knowledge.*
(128) *Increased knowledge of natural sciences.*

4.1.4. *Social aspects.* The least numerous group were indications of the social aspects of graduation. Teachers reported that it was important for them to have contacts with other teachers. It is worth noting that teachers met during such studies are in the same position. They have in common not only the experience of e.g. not having a full-time position, fear of losing their jobs, but also potential future discrimination and unequal treatment by fellow mathematicians teaching in the same school on the basis of qualifications obtained at full-time university studies. Other important social issues highlighted by the teachers were the opportunity to meet new people and particularly the chance to meet famous mathematicians. Some of the answers in this category are given below:

(13) *Many contacts with people and a different perspective on my profession.*
(53) *People I have met - maths teachers.*
(74) *I met nice people whom I still keep in touch with.*
(135) *Meeting new teachers, sharing experiences from school work. Meeting famous mathematicians and authors of mathematics textbooks.*

4.1.5. *Positive features of completed studies.* Only 46 out of the 239 teacher indications were related to the characteristics and qualities of the studies as such. In some statements, it was emphasized that the great advantage of these studies is their duration (acceptable time span) and the possibility of matching in time studying and working at school at the same time. Another group of responses pointed to the advantages of the substantive preparation in mathematics that one could receive during the studies. Respondents also wrote about the revision of knowledge and substantive preparation for the material covered by the school curriculum. However, a larger number of respondents paid attention to the didactical part of their preparation for teaching mathematics. The teachers indicated various aspects of the methodological knowledge obtained during the studies: knowledge of the didactics of mathematics, methods of teaching mathematics, learning how to introduce concepts, learning about different tasks and exploring different ways of solving them, a lot of training with tasks, gaining knowledge about working with gifted students. In the opinions of the respondents, apprenticeships carried out in schools were also important and beneficial to them. Moreover, there was a positive feedback concerning the classes (interesting and well conducted), the teaching staff (qualified, demanding, with practical experience) and the materials made available to the students.

4.2. **NEGATIVE sides of the studies.** Only 66 teachers indicated negative sides of the completed studies. This time most of the answers, coming from 60 teachers, referred to the negative features of the studies as such. Only 6 teachers mentioned any negative impact the studies had on some aspects of their lives and functioning. The teachers gave 94 indications altogether. After initial coding of the raw data, we grouped the answers into the categories shown below:



4.2.1. *Classes.* In the group of statements indicating negative features of postgraduate studies, the largest number of indications were related to the classes taught during these studies. The surveyed teachers indicated the following problems: lack of specific types of classes or topics, unattractive course offerings, type of delivered knowledge, especially its lack of relevance to school mathematics, as well as the way of teaching and some staff issues. The quotes below give a closer look at the content of the respondents' statements:

> (1) *Lack of interesting classes in didactics - the ones I had were taught by an 80 year old man who didn't have the faintest idea about working with today's youth.*
> (72) *Solving mathematical tasks which I had been solving during my Master's studies in physics.*
> (73) *Uninteresting program.*

> (47) *Several subjects were taught in a very stressful way, e.g. mathematical analysis. Little time was spent on explaining and a lot on demanding. The selection of the person for the blackboard was done by the lecturer going around the room and putting down the chalk in front of the chosen person. This and similar methods are not allowed to be used by us, the teachers, but were applied to us constantly. I didn't feel safe in most of the classes, and I learned less than I could have. In geometry classes, we didn't have time to rewrite everything, let alone think about what we were actually writing and why.*
> (101) *Teachers (some of them) treating us as if we were inexperienced in teaching (despite many years of experience in school).*
> (1) *I was learning number theory, and then I had to learn how to explain word tasks to children in a comprehensible way.*
> (2) *There was no translation into school practice. The methodology teacher showed me how to work with gifted children (because she worked in such a city), but in the rural school where I worked there were no such children or only single cases. But a lot of children have so called specific difficulties and it would be nice to learn already at university how to work with them, activate them and encourage to learn this subject.*
> (57) *I do not really know why I was made to learn, for example, complex numbers and other elements of higher mathematics.*

Statements that point to the subject matter of the courses that were offered, reveal the needs of some postgraduate students that may not have been met in a satisfactory manner. According to the respondents, the following kind of classes or topics were missing or tackled during insufficient number of hours:

- too few hours of didactics of mathematics, methodology: teaching methods, examples of interesting activities, ability to solve tasks, little use of modern technologies,
- too little or no preparation to work in a primary school,
- too little or no preparation to work with gifted students or students with difficulties,
- lack of analysis of the core curriculum,
- lack of psychological preparation for: working with difficult youth, talking to parents.

4.2.2. *Duration and organization.* A total of 16 responses indicating negative sides of postgraduate studies were related to the duration and organization of these studies. The respondents complained about high teaching intensity with a lot of students' own work at home, practice load, difficulty in harmonising studies with daily job, and the duration of studies – to some teachers they lasted too long, to others too short to learn everything.

> (6) *Schedule.*
> (18) *High intensity (3 times a month from Friday to Sunday).*
> (14) *Lots of work at home.*
> (36) *Studies are too short to learn everything you need to learn. I think this applies to postgraduate studies in any other subject.*
> (70) *Too much time consuming.*
> (115) *I completed a two-year postgraduate course, it was a huge time investment.*
> (150) *A lot of people "from nowhere", each planning to work in a different type of school, each had different expectations and needs.*
> (93) *Too many hours of practice - difficult to reconcile with being a teacher who works full time or more.*

4.2.3. *Cost.* There were 11 respondents who wrote that the studies were very expensive. They also complained that they had to pay for them with their own money.

> (2) *I paid for it with my own money (over 3,000 PLN) because the school did not have the money for it.*

'I have learned nothing', waste of time

Ten respondents felt that they did not learn anything new or even that the time spent on these studies, was time wasted.

> (94) *my studies did not really give me anything but took up my time and money.*
> (3) *Lack of knowledge and skills.*
> (34) *Spending a lot of time on passing things I already knew - mathematics was higher at polytechnic studies than at postgraduate studies.*
> (73) *I knew everything.*
> (111) *[The studies] did not add anything new.*
> (17) *Time wasted on getting a "paper".*
> (141) *Waste of time.*

Negative impact on various aspects of life

Only six of the surveyed teachers referred to the negative impact that completing the studies had on their later lives and functioning.

> (50) *My competence is being questioned by other staff.*
> (136) *These will never be a full-time studies - treated worse.*
> (102) *When teaching 4 subjects, I am tired, but I would not be able to give up my first subject (arts) to teach only maths.*

(105) *A lot of work in comparison to other less responsible subjects + external exams for students, and for me more charity work for the same money as for religion, art, physical education.*
(140) *A lot of responsibility for teaching results and stress.*
(116) *Today I would rather be an engineer. Teachers in Poland are despised.*

5. Summary

To the best of our knowledge, our project is the first attempt to address the broadly understood problem of mathematics teacher education on postgraduate studies in Poland. The results presented in this text provide some insight into how the teachers themselves perceive postgraduate education.

In the surveyed group, more teachers were able to give positive sides of the completed studies than negative ones: 150 respondents indicated positive sides and only 66 people reported some negatives. There were 2.5 times more indications of positives (239 contributions) than indications of negative sides (94 contributions). Regarding the positive sides of the completed studies: 127 teachers referred to the impact the studies have had on their lives and only 30 respondents referred to the features of the studies. Regarding the negative sides of the completed studies only 6 teachers referred to the impact the studies have had on their lives and 60 referred to the features of the studies. Among the responses concerning the impact of the studies on different aspects of teachers' lives, there were about 32 times more indications of positive features. Among the responses concerning the characteristics of the studies, there were almost 2 times more indications of negative features.

Given the importance of postgraduate studies for the professional development of teachers, it is worth making every effort to ensure that the university offer is substantively meaningful to them, and that it is not detached from school reality.

Acknowledgements

This work has been supported by the National Science Centre, Poland, grant No. 2018/31/N/HS6/03976.

References

[1] B. Barańska, M. Zambrowska, *Neither fully "in", nor completely "out" of the field: The case of teaching mathematics as a second subject in Poland*, in L. Hobbs, R. Porsch (Eds.), *Out-of-field teaching across teaching disciplines and contexts*, Springer. (In press).

[2] M. Bosse, M., G. Törner, *The practice of out-of-field teaching in mathematics classrooms – a German case study*, in L. Sumpter (Ed.), *Current State of Research on Mathematical Beliefs XX: Proceedings of the MAVI-20 Conference Kultur och Lärande*, Falun: Högskolan Dalarna University Press **2015**:04 (2015), 77–88.

[3] Eurostat (2016). *Women teachers largely over-represented in primary education in the EU. News release 189/2016* (Retrieved from: ec.europa.eu/eurostat).

[4] Eurydice. *The teaching profession in Europe (profile, trends and concerns) Report 2: Supply and demand. General lower secondary education*, Brussels: The Information Network on Education in Europe, European Commission, (2002), retrieved from https://op.europa.eu.

[5] O. Fitzmaurice, R. Walsh, R., K. Burke, *The 'Mathematics Problem' and preservice post primary mathematics teachers – analysing 17 years of diagnostic test data*, International Journal of Mathematical Education in Science and Technology **52**.2 (2021), 259–281.

[6] L. Hobbs, *Teaching 'out-of-field'as a boundary-crossing event: Factors shaping teacher identity*, International Journal of Science and Mathematics Education, **11**.2 (2013), 271–297.

[7] L. Hobbs, G. Törner, *Teaching out-of-field as a phenomenon and research problem*, in *Examining the Phenomenon of "Teaching Out-of-field"*, Springer (2019), 3-20.

[8] R.M. Ingersoll, *The problem of out-of-field teaching*, The Phi Delta Kappan, **79**.10 (1998), 773–776.

[9] R.M. Ingersoll, *The Realities of Out-of-Field Teaching.* Educational leadership, **58**.8 (2001), 42–45.

[10] A. Krause, B. Muchacka, S. Przybyliński, *Kształcenie nauczycieli – analiza na podstawie eksperckich doświadczeń z ocen programowych i instytucjonalnych Polskiej Komi-sji Akredytacyjnej*, Rocznik Pedagogiczny (2017), 89–103.

[11] M.N. Ríordáin, A. Hannigan, *Who teaches mathematics at second level in Ireland?*, Irish Educational Studies (2011) **30**.3, 289–304.

[12] J. Saldaña, *The coding manual for qualitative researchers.* SAGE (2009).

[13] L. Shah, C. Jannuzzo, T. Hassan, B. Gadidov, H.E. Ray, G.T. Rushton, *Diagnosing the current state of out-of-field teaching in high school science and mathematics*, PloS One **14**.9 (2019).

[14] G. Törner, A. Törner, *Underqualified Math Teachers or Out-of-Field-Teaching in Mathematics – A Neglectable Field of Action?*, in W. Blum, R. B. Ferri, K. Maass (Eds.) *Mathematikunterricht im Kontext von Realität, Kultur und Lehrerprofessionalität* Vieweg+Teubner Verlag (2012), 196–206.

Pedagogical University of Cracow, Poland
barbara.baranska@up.krakow.pl

The Maria Grzegorzewska University, Poland
mzambrowska@aps.edu.pl

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

CENTRAL OPEN SETS TILINGS

LOUISA F. BARNSLEY & MICHAEL F. BARNSLEY

ABSTRACT. We introduce a method for constructing collections of subsets of $\mathbb{R}^n$, using an iterated function system, a set T, and a cost function. We refer to these collections as tilings. The special case where T is the central open set of an iterated function system that obeys the open set condition is emphasized. The notion of the central open set associated with an iterated function system of similitudes, introduced in 2005 by Bandt, Hung, and Rao, is reviewed. A practical method for calculating pictures of central open sets is described. Some general properties and examples of the tilings are presented.

KEYWORDS: iterated function systems, fractal geometry, tilings

MSC2010: 28A80, 05B45, 52C22

Received May 13, 2021; accepted 28 July 2021

1. Introduction

The goal of this paper is to describe a simple method for producing a wide range of tilings (in a generalized sense). The method uses an iterated function system (IFS) acting on $\mathbb{R}^n$, together with a set T, and a cost function c. Figure 1 illustrates part of such a tiling where T is the central open set of an IFS. The possible structures of the new tilings are diverse, yet all are handled with the same underlying mathematical device. The formalism yields examples in analysis, geometry, and dynamics; it leads to extensive and rich families of variants of the self-similar tilings introduced in [10, 11]. Our methods are based on addresses associated with IFSs and mappings from these addresses into tilings and tiling spaces.

 https://doi.org/10.2478/9788366675360-004.

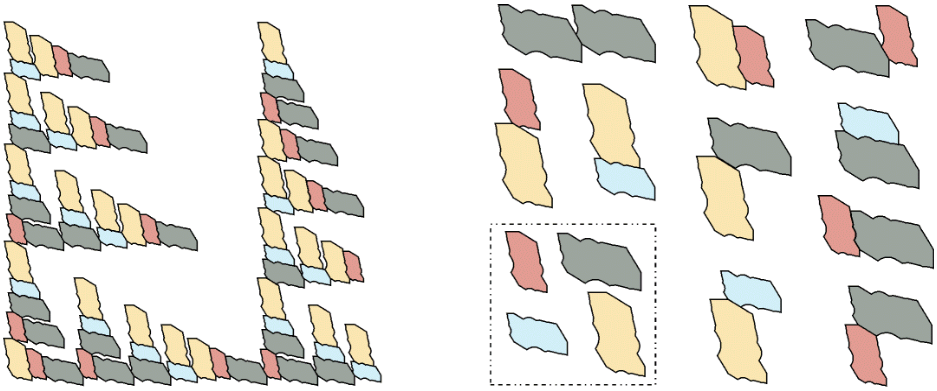


FIGURE 1. See Example 5.1. The left panel illustrates part of an unbounded central open set tiling. The right panel shows some of the ways in which the four prototiles (see inset box) can meet. See also Figures 5 and 6.

Our main examples use the central open set of the attractor of an IFS to provide the shapes of the tiles. In this case the tilings have properties that suggest they may be used to model patterns that arise naturally. For example, such tiles have self-similar features and may touch, and there may also be gaps that repeat at different scales. Type "mudcracks" into a search engine to see illustrations of seemingly related natural patterns.

In this paper we use the following lexicon. We use the word *tiling* to mean a collection of closed subsets of $\mathbb{R}^n$, and a *tile* is a member of the collection. This is more general than the standard definitions, see Grunbaum and Sheppard [15]. We say that two *tiles meet* if their intersection is non-empty, and we say that they *touch* if their intersection is non-empty and contains no interior points. The *support* of a tiling is the union of its tiles. Two *tilings meet* if the union of the tiles in their intersection is a tiling whose support is the intersection of the supports of the two tilings. A set of *prototiles* for a tiling is a set of tiles such that each member of the tiling is related to a prototile by an isometry. The set of isometries may be restricted to translations. We say that two tiles have the same *shape* if they are related by a similitude. If two tiles are related by an isometry we say that they are *copies* of one another. We say that a tiling is *commensurate* if the sizes of all its tiles belong to a geometrical progression; otherwise the tiling is *incommensurate;* see also [22]. A *patch* of a tiling is the set of its tiles that have non-empty intersection with a finite set, typically a disk or rectangle.

In this paper we are concerned with the situation where all tiles have the same shape. But the theory is readily generalized to multiple shapes, by using graph-directed IFS, as in [13].

2. Iterated function systems and central open sets

Here we review the notion of an IFS of similitudes and of a central open set introduced in 2005 by Bandt, Hung, and Rao, [4].

Roughly following [4], let $F = \{\mathbb{R}^n; f_1, f_2 \dots f_m\}$ denote a collection of contractive similitudes, that is $f_i : \mathbb{R}^n \to \mathbb{R}^n$ with

$$|f_i(x) - f_i(y)| = \lambda_i |x - y| \text{ for all } x, y \in \mathbb{R}^n$$

where the $\lambda_i \in (0, 1)$ are the contraction factors and $|\cdot|$ denotes the Euclidean norm. We refer to F as an iterated function system. By slight abuse of notation we use the same symbol F to denote the mapping from sets to sets $F : 2^{\mathbb{R}^n} \to 2^{\mathbb{R}^n}$ defined by $F(S) = \{f_i(x) : x \in S, i = 1, \dots m\}$. It is well-known that there exists a unique non-empty compact set $A \subset \mathbb{R}^n$ such that

$$A = F(A) = \bigcup_{i=1}^{m} f_i(A)$$

where $f_i(A) = \{f_i(x) | x \in A\}$, [16].

The set A is called the *attractor* of the iterated function system F, because

$$\lim_{k\to\infty} F^k(\{x\}) = A \text{ for all } x \in \mathbb{R}^n,$$

where convergence is with respect to the Hausdorff metric on $\mathbb{R}^n$, and F^k is the function F composed with itself k times. We say that the *basin* of A is $\mathbb{R}^n$. The *fast basin* [5, 6] of A is a subset of the basin defined by

$$B = \{x \in \mathbb{R}^n | F^k(\{x\}) \cap A \neq \emptyset, \text{ some } k \in \mathbb{N}\},$$

where $\mathbb{N}$ is the set of positive integers. The fast basin is the set of points such that some finite orbit meets A. Fast basins are related to but distinct from the fractal blow-ups introduced by Strichartz [24] and the macro-fractals introduced by Banakh and Novosad [2]. We will use $B \backslash A$ in calculations in Section 4.

Roughly quoting [4], the attractor A is the union of smaller copies of itself, $A_i = f_i(A)$, where each A_i consists of smaller copies $A_{ij} = f_i(f_j(A))$, and so on. For any positive integer k, we can consider the set Σ^k of words $\mathbf{i} = i_1 \dots i_k$ from the alphabet $\Sigma = \{1, 2, \dots m\}$. Writing $f_{\mathbf{i}} = f_{i_1} f_{i_2} \dots f_{i_k}$ and $A_{\mathbf{i}} = f_{\mathbf{i}}(A)$ we have

$$A = \bigcup \{f_{\mathbf{i}}(A) | \mathbf{i} \in \Sigma^k\}.$$

When k tends to infinity, this induces a continuous map that we call the *address* map, $\pi : \Sigma^\infty \to A$, from the set Σ^∞ of infinite sequences $i_1 i_2 \dots$ onto the attractor.

The IFS F is said to satisfy the open set condition (OSC) if there is a nonempty open set $O \subset \mathbb{R}^n$ such that

$$F(O) \subset O \text{ and } f_i(O) \cap f_j(O) = \emptyset \text{ for } i \neq j.$$

Such a set O is called a *feasible open set* of F. The OSC plays an important role in fractal geometry. For example, if F obeys the OSC, then [19] the Hausdorff dimension of A is the unique positive solution D of

$$\sum_{i=1}^{m} \lambda_i^D = 1.$$

See [4] for a succinct account of the history and mathematical significance of the OSC. Here we are interested in a particular feasible open set, the *central open set* of F, and its relationship to fractal tilings [11].

The second requirement of the OSC may be written

$$O \cap f_{\mathbf{i}}^{-1} f_{\mathbf{j}}(O) = \emptyset$$

for $i_1 \neq j_1$. The maps in

$$\mathcal{N} = \{f_{\mathbf{i}}^{-1} f_{\mathbf{j}} | \mathbf{i}, \mathbf{j} \in \Sigma^*, i_1 \neq j_1\} \text{ where } \Sigma^* = \bigcup_{k=1}^{\infty} \Sigma^k$$

are called *neighbor maps*, [3, 4]. Neighbor maps may be used to provide an algebraic formulation of the OSC: there is a constant $\kappa > 0$ such that $\|h - id\| > \kappa$ for all neighbor maps h. Neighbor maps are related to the fast basin by

$$B \backslash A = H \backslash A \text{ where } H = \bigcup \{h(A) | h \in \mathcal{N}\}.$$

Any feasible open set O must have empty intersection with H.

The *central open set* C for F is defined to be

$$C = \{x \in \mathbb{R}^n | d(x, A) < d(x, H)\} = \{x \in \mathbb{R}^n | d(x, A) < d(x, B \backslash A)\}$$

where $d(x, Y) = \inf\{|x - y| \mid y \in Y\}$. Bandt et al. [4] prove the following theorem and its elegant corollary.

Theorem 2.1. *If the OSC holds, then the central open set C is a feasible open set. If the OSC does not hold then C is empty.*

Corollary 2.2. The OSC holds if and only if A is not contained in $\overline{H}$.

It is an unanswered question as to whether or not it is true that the OSC holds if and only if $A \neq \overline{A \cap H}$. The latter was claimed by M. Moran [19], but his proof contains a gap [4].

Questions relating to the existence of, and the structure of, a feasible open set for a given IFS are very subtle, see for example [25].

3. Diverse tilings derived from an IFS and a cost function

In this Section we describe a general construction of tilings using an IFS F, a set $T \subset \mathbb{R}^n$, and a cost function c defined below. The resulting tilings may have overlapping tiles with non-empty interiors. Such tilings might be used to model fallen leaves carpeting a forest floor, duckweed on the surface of a pond, cracks in dried mud, or to design patterns for wallpaper.

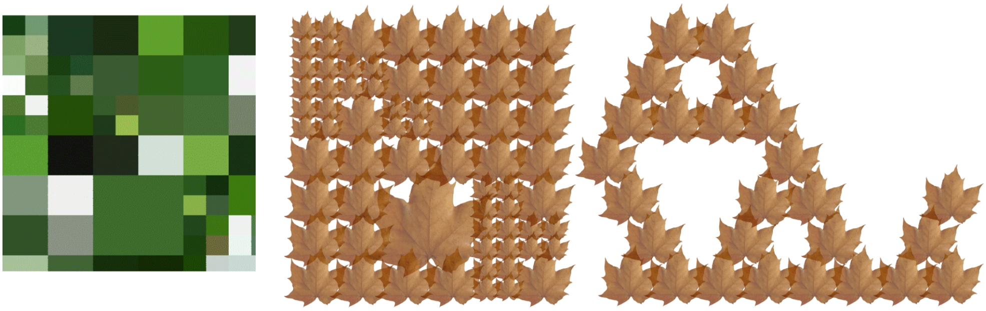


FIGURE 2. Examples of patches of three unbounded tilings. The leftmost two images are related to an IFS whose attractor is a square. The rightmost image is related to an IFS whose attractor is a Sierpinski triangle. Two of the panels use $T = L$, a decorated leaf-shaped set, see Example 3.2.

Let F be a IFS consisting of at least two distinct similitudes, and let T be a closed subset of $\mathbb{R}^n$. For convenience we suppose that $A \cap T \neq \emptyset$, but this is not necessary. We use sets of similitudes, which are subsets of $\mathcal{N}$ determined by the cost function and are applied to T, to form collections of scaled and translated and possibly flipped (i.e. turned upside down, in the two-dimensional case) copies of T, with possible overlaps.

Let $\mathbb{H}(\mathbb{R}^n)$ be the closed bounded subsets of $\mathbb{R}^n$ equipped with the spherical Hausdorff metric, see [12, 13]. This metric d is defined as follows. Let $P(x)$ be the stereographic projection of $x \in \mathbb{H}(\mathbb{R}^n)$ onto the $(n+1)$ dimensional sphere tangent to $\mathbb{R}^n$ at the origin. Then $d(x, y)$ is the Hausdorff distance between $P(x)$ and $P(y)$ using the round metric d_R on the sphere. Let $d_H(X, Y)$ be the Hausdorff distance between sets of subsets X and Y of $\mathbb{R}^n$ calculated using d_R. Let $\mathbb{H}(\mathbb{H}(\mathbb{R}^n))$ be the collections of subsets of $\mathbb{H}(\mathbb{R}^n)$ that are closed with respect to d_H.

For each $i \in \{1, \dots M\}$ write $\lambda_i = s^{a_i}$ where $s = \max\{\lambda_i | i = 1, ...m\}$ and assign a cost $c_i > 0$ to the map f_i. For example we may choose $c_i = a_i$. For $\mathbf{i}$=$i_1 i_2 \cdots \in \Sigma^\infty$, write $\mathbf{i}|k = i_1 \dots i_k \in \Sigma^*$, and $\mathbf{i}|0 = \emptyset$. Define a *cost function* $c : \{\emptyset\} \cup \Sigma^* \to (0, \infty)$ by

$$c(\mathbf{i}|k) = c_{i_1} + c_{i_2} + \cdots + c_{i_k},\ c(\emptyset) = 0.$$

Define a mapping $\Pi_T : \{\emptyset\} \cup \Sigma^* \cup \Sigma^\infty \to \mathbb{H}(\mathbb{H}(\mathbb{R}^n))$ by

$$\begin{aligned}
\Pi_T(\mathbf{i}|k) &= f_{-(\mathbf{i}|k)}(\{f_{(\mathbf{j}|l)}(T) | \mathbf{j} \in \Sigma^\infty, l \in \mathbb{N},\ c(\mathbf{j}|l-1) \leq c(\mathbf{i}|k) < c(\mathbf{j}|l)\}),\\
f_{-(\mathbf{i}|k)} &:= f_{i_1}^{-1} \dots f_{i_k}^{-1}, f_{(\mathbf{j}|l)} = f_{j_1} ... f_{j_l},\\
\Pi_T(\mathbf{i}) &= \bigcup_{k=1}^{\infty} \Pi_T(\mathbf{i}|k) \text{ for all } \mathbf{i} \in \Sigma^\infty, \Pi_T(\emptyset) = \{f_1(T), f_2(T), ... f_m(T)\}.
\end{aligned}$$

Π_T is well-defined because $\{\Pi_T(\mathbf{i}|k)|k = 1, 2, ...\}$ is a nested increasing sequence of collections of sets:

$$\Pi_T(\mathbf{i}|0) \subset \Pi_T(\mathbf{i}|1) \subset \Pi_T(\mathbf{i}|2) \ldots \tag{3.1}$$

for all $\mathbf{i} \in \Sigma^\infty$. This is true because

$$\begin{aligned}\Pi_T(\mathbf{i}|k+1) &= f_{-(\mathbf{i}\,|\,k+1)}(\{f_{(\mathbf{j}\,|\,l)}(T) \mid \mathbf{j} \in \Sigma^\infty, l \in \mathbb{N},\\ &\quad c(\mathbf{j}\,|\,l-1) \leq c(\mathbf{i}\,|\,k+1) < c(\mathbf{j}\,|\,l)\})\\ &\supset f_{-(\mathbf{i}\,|\,k+1)}(\{f_{(\mathbf{j}\,|\,l)}(T) \mid \mathbf{j} \in \Sigma^\infty, l \in \mathbb{N},\\ &\quad c(\mathbf{j},\ |\,l-1) \leq c(\mathbf{i}\,|\,k+1) < c(\mathbf{j}\,|\,l), j_1 = i_{k+1}\})\\ &= f_{i_1}^{-1} \ldots f_{i_{k+1}}^{-1}(\{f_{i_{k+1}} f_{j_2} \ldots f_{j_l}(T) \mid \mathbf{j} \in \Sigma^\infty, l \in \mathbb{N},\\ &\quad c(\mathbf{j}\,|\,l-1) \leq c(\mathbf{i}\,|\,k) < c(\mathbf{j}\,|\,l)\}\\ &= \Pi_T(\mathbf{i}|k).\end{aligned}$$

Equation (3.1) is the key mathematical device in this paper. In general $\Pi_T(\mathbf{i})$ is a collection of subsets of $\mathbb{R}^n$, that we call tiles. These tiles are translated, scaled, maybe flipped and/or rotated, copies of T. They may be overlapping and the support of the tiling $\Pi_T(\mathbf{i})$, namely $\bigcup\{t \in \Pi_T(\mathbf{i})\} \subset \mathbb{R}^n$ may be complicated.

If F obeys the OSC and $T \subset \overline{C}$, then distinct sets of the form $f_{-(\mathbf{i}|k)} f_{(\mathbf{j}|l)}(T)$ in the tiling $\Pi_T(\mathbf{i})$ are non-overlapping; that is, the interiors of the intersections of distinct tiles are empty.

Denote the range of $\Pi_T : \Sigma^\infty \to \mathbb{H}(\mathbb{H}(\mathbb{R}^n))$ by $\mathbb{T}_T = \{\Pi_T(\mathbf{i})|\mathbf{i} \in \Sigma^\infty\}$. As a consequence of properties of, and structures associated with, the shift map $\sigma : \Sigma^\infty \to \Sigma^\infty$, much can be said, along the lines of [12, 13], about continuity properties of $\Pi_T : \Sigma^\infty \to \mathbb{T}_T$ with respect to the metric d_H defined above, dynamics, invariant measures, and ergodic properties associated with mappings that take $\mathbb{T}_T$ into itself, such as certain inflation and deflation operations.

Example 3.1. Let $F = \{\mathbb{R}^1; f_1, f_2\}, f_1(x) = \frac{x}{2}, f_2(x) = \frac{x+1}{2}$. Then $A = [0, 1]$. Choosing $T = [-\frac{1}{3}, \frac{4}{3}]$, we find $\Pi_{[-\frac{1}{3},\frac{4}{3}]}(\overline{1}) = \{[-\frac{1}{6} + \frac{n-1}{2}, \frac{1}{6} + \frac{n}{2}]|\ n = 1, 2, \ldots\}$. That is, $\Pi_{[-\frac{1}{3},\frac{4}{3}]}(\overline{1})$ is a collection of overlapping closed intervals whose union is $[-\frac{1}{6}, \infty)$.

Example 3.2. (i) The leftmost panel in Figure 2 illustrates part of a tiling $\Pi_A(\overline{1})$ where the IFS of four similitudes each with scaling factor 0.5, attractor $A = [0, 1] \times [0, 1]$, and the cost function is defined by $c_1 = 1$, $c_2 = 1.3$, $c_3 = 1.5$, $c_4 = 2$. (ii) The middle panel illustrates the same part, but of $\Pi_L(\overline{1})$ where L is the support of a leaf picture. In this case the tiles have been decorated by a picture of a leaf. (iii) The rightmost panel is related to an IFS of three maps, whose attractor is a Sierpinski triangle, and the cost function specified by $c_1 = c_2 = c_3 = 1$, and the same set L.

For a two-dimensional affine transformation $f : \mathbb{R}^2 \to \mathbb{R}^2$ we write

$$f = \begin{bmatrix} a & b & e \\ c & d & g \end{bmatrix} \text{ for } f(x, y) = (ax + by + e, cx + dy + g)$$

where $a, b, c, d, e, g \in \mathbb{R}$.

Example 3.3. In the special case where $c_i = a_i \in \mathbb{N}$ for all $i = 1, \dots m$, we call the sets

$$T_k = s^{-k}\{f_{(\mathbf{i}|l)}(A)|c(\mathbf{i}|l-1) \leq k < c(\mathbf{i}|l)\}, T_0 = F(A),$$

canonical tilings. They play a natural role in fractal tilings [13], and in connecting them to algebraic geometry [1]. Two sequences of canonical tilings are illustrated in Figure 3. The IFSs may be deduced from $T_0 = \{f_1(A), f_2(A)\}$. The IFS for the top sequence is $F = \{\mathbb{R}^2; f_1, f_2\}$ where

$$f_1 = \begin{bmatrix} 0 & s & 0 \\ -s & 0 & s \end{bmatrix}, f_2 = \begin{bmatrix} -s^2 & 0 & 1 \\ 0 & s^2 & 0 \end{bmatrix}$$

where $s + s^2 = 1$, $s > 0$, and for the lower sequence

$$f_1 = \begin{bmatrix} s & 0 & 1 \\ 0 & s & 0 \end{bmatrix}, f_2 = \begin{bmatrix} 0 & s^2 & 0 \\ s^2 & 0 & 0 \end{bmatrix}$$

with $s + s^4 = 1$, $s > 0$. In these two cases, and others like them, when the cost function is defined by $c_i = i$, there is a simple relationship between the canonical tilings and the tilings $\Pi_A(\mathbf{i})$, namely

$$\Pi_A(\mathbf{i}|k) = f_{-(\mathbf{i}|k)} s^{c(\mathbf{i}|k)} T_{c(\mathbf{i}|k)}, \text{ where } c(\mathbf{i}|k) = i_1 + i_2 + \dots i_k,$$

for all $\mathbf{i}$ and k. That is, $\Pi_A(\mathbf{i}|k)$ is isometric to $T_{c(\mathbf{i}|k)}$, see [13].

4. Central open set tilings

In this Section we assume that F obeys the OSC, and consider the two special cases $T = A$ and $T = \overline{C}$ in the mapping $\Pi_T : \{\emptyset\} \cup \Sigma^* \cup \Sigma^\infty \to \mathbb{H}(\mathbb{H}(\mathbb{R}^n))$. For consistency with [13] we define

$$\Pi = \Pi_A \text{ and } \Xi = \Pi_{\overline{C}}\ .$$

These tilings are of particular interest to us. Here's why. If A has nonempty interior, so that $\overline{C} = A$, these tilings may be examples of conventional self-similar tilings as defined by [15], or tilings with fractal boundaries [17, 18]. But they are more general because they may be tilings with infinitely many incommensurate tile sizes. The case $T = \overline{C}$ is special because it seems to be an extreme case: we conjecture that if T is chosen to be a closed set that contains an open set that contains $\overline{C}$, then $\Pi_T(\mathbf{i})$ contains overlapping tiles.

The mapping $\Pi = \Pi_A$, and the fractal tilings it generates when $c_i = a_i$, were introduced and studied in [9, 10, 11]. We refer to $\Pi(\mathbf{i})$ as a fractal tiling, and we refer to sets of the form $f^{-1}_{(\mathbf{i}|k)} f_{(\mathbf{j}|l)}(A)$ as fractal tiles. Note that fractal tiles may have empty interiors. They have non-empty interiors when A has non-empty

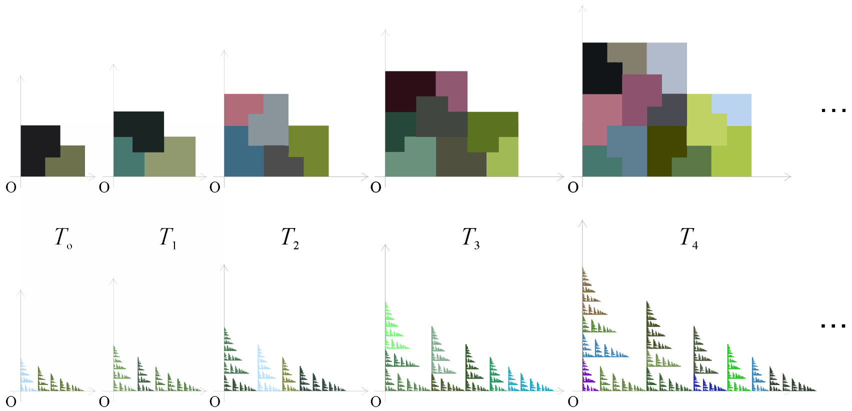


FIGURE 3. Canonical tilings $\{T_k\}$ for two different iterated function systems $\{\mathbb{R}^2, f_1, f_2\}$. The tiling T_{k+1} is derived from T_k by replacing each isometric copy of A in $s^{-1}T_k$ by a copy of $T_0 = \{f_1(A), f_2(A)\}$. See Example 3.3.

interior. The relationship of the address map $\pi : \Sigma^\infty \to A$ to the contractive IFS F has analogies with the relationship of Π to the expansive IFS $F^{-1} = \{\mathbb{R}^2; f_1^{-1}, \dots f_m^{-1}\}$, see also [2, 24].

We call $\Xi(\mathbf{i}) = \Pi_{\overline{C}}(\mathbf{i})$ a *central open set tiling*. We refer to sets of the form $f^{-1}_{(\mathbf{i}|k)} f_{(\mathbf{j}|l)}(\overline{C})$ as central open set tiles. The interiors of the tiles in $\Xi(\mathbf{i})$ are non-empty and disjoint for any fixed $\mathbf{i}$, but the tiles may touch.

Example 4.1. Let $F = \{\mathbb{R}^1; f_1, f_2\}, f_1(x) = \frac{x}{2}, f_2(x) = \frac{x+1}{2}$. Then $A = [0,1]$, $C = (0,1)$ and we find $\Pi(\overline{1}) = \Xi(\overline{1}) = \{[\frac{n-1}{2}, \frac{n}{2}] \subset \mathbb{R}|\ n \in \mathbb{N}\}$, and if the tail of $\mathbf{i} \in \Sigma^*$ is neither $\overline{1} = 11...$ nor $\overline{2} = 22...$, then $\Pi(\mathbf{i}) = \Xi(\mathbf{i}) = \{[\frac{n-1}{2}, \frac{n}{2}]|\ n \in \mathbb{Z}\}$. If the tail of $\mathbf{i} \in \Sigma^*$ is either $\overline{1} = 11...$ or $\overline{2} = 22...$, then $\Pi(\mathbf{i})$ is also tiling by half unit intervals, but the support is either $[k, \infty)$ or $(-\infty, k]$ for some $k \in \mathbb{Z}$.

Example 4.2. Let $F = \{\mathbb{R}^2; f_1, f_2\}, f_1(x,y) = (\frac{x}{2}, \frac{y}{2}), f_2(x,y) = (\frac{x+1}{2}, \frac{y}{2})$. The attractor is $A = [0,1] \times \{0\}$, and the central open set is $C = (0,1) \times (-\infty, \infty)$ is unbounded. If the tail of $\mathbf{i} \in \Sigma^*$ is neither $\overline{1} = 11...$ nor $\overline{2} = 22...$, then $\Pi(\mathbf{i}) = \{[\frac{n}{2}, \frac{n+1}{2}] \times \{0\} \subset \mathbb{R}^2 | n \in \mathbb{Z}\}$, and $\Xi(\mathbf{i}) = \{[\frac{n}{2}, \frac{n+1}{2}] \times (-\infty, \infty) : n \in \mathbb{Z}\}$.

Example 4.3. Let $F = \{\mathbb{R}^2; f_1, f_2, f_3\}$ where

$$f_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}, f_2 = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}, f_3 = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{4} \end{bmatrix}.$$

Then the attractor is a Sierpinski triange and the central open set is a hexagon, see [4]. See Figure 4.

In Theorem 4.4 below we establish some properties of the collection of tilings $\mathbb{T}_T$. We use the following terminology. We say that $\mathbf{i} \in \Sigma^\infty$ is *disjunctive*, when

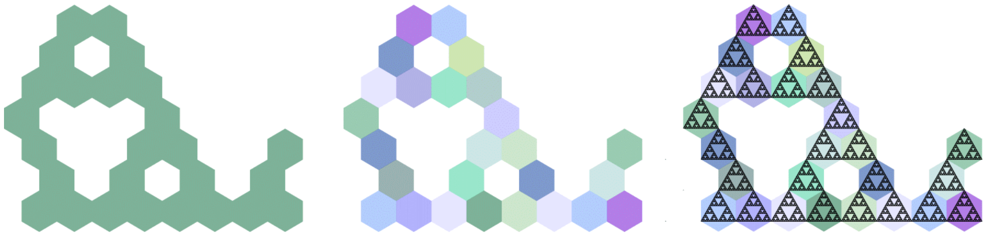


FIGURE 4. From left to right: (i) part of the support of a central open set tiling; (ii) a patch of the same central open set tiling; (iii) same patch, also showing the underlying fractal tiles. See Example 4.3.

given any finite word $\mathbf{j}|p = j_1 j_2 \dots j_p$, there is $k \in \mathbb{N}$ such that $i_{k+1} \dots i_{k+p} = j_1 j_2 \dots j_p$. We say that $\mathbf{i} \in \Sigma^\infty$ is *reversible* when A has non-empty interior A° and there exists $k < l$ such that $f_l f_{l-1} \dots f_k(A) \subset A^\circ$, the interior of A. If A has non-empty interior, then disjunctive is a special case of reversible, [8, 9], but disjunctiveness is much easier to check than reversibility.

Theorem 4.4. *Let $F = \{\mathbb{R}^n; f_1, \dots f_m\}$ be an IFS of contractive similitudes. Let $T \subset \mathbb{R}^n$ be closed and let c be a cost function. Then $\Pi_T(\mathbf{i}|k)$ and $\Pi_T(\mathbf{i})$ are well-defined collections of closed subsets of $\mathbb{R}^n$, (i.e. they are tilings), and Equation (3.1) holds for all $k \in \mathbb{N}$ and all $\mathbf{i} \in \Sigma^\infty$.*

If F obeys the OSC, then the following statements are true.

(i) For all $\mathbf{i} \in \Sigma^\infty$ the interiors of the tiles that comprise $\Xi(\mathbf{i}) = \Pi_{\overline{C}}(\mathbf{i})$ are disjoint.

(ii) The interior of $f^{-1}_{(\mathbf{i}|k)} f_{(\mathbf{j}|l)}(\overline{C})$ is the central open set for the iterated function system SFS^{-1} where $S = f^{-1}_{(\mathbf{i}|k)} f_{(\mathbf{j}|l)}$ is a similitude. In this sense all tiles in $\Xi(\mathbf{i})$ are central open sets.

(iii) In $\mathbb{R}^2$, if A is a polygon and $c_i = a_i \in \mathbb{N}$ for all $i \in \{1, ...m\}$, then $\Xi(\mathbf{i}) = \Pi(\mathbf{i})$. In this case, if also $\mathbf{i} \in \Sigma^\infty$ is reversible, then the support of $\Pi(\mathbf{i})$ is $\mathbb{R}^n$ and $\Pi(\mathbf{i})$ is a self-similar tiling in the sense of standard works such as [1, 15, 23] and many others such as [14, 17, 18, 20] .

(iv) In general the tiling $\Pi_T(\mathbf{i})$ is incommensurate both as defined here and in the sense of [22]. The tilings $\Pi_T(\mathbf{i})$, in particular $\Pi(\mathbf{i})$ and $\Xi(\mathbf{i})$, are commensurate when $c_i = a_i \in \mathbb{N}$ for all $i = 1, 2, \dots m$.

(v) Let $\mathbf{i}, \mathbf{j} \in \Sigma^\infty$, $p, q \in \mathbb{N}$, and the cost function c, be such that $\sigma^p \mathbf{i} = \sigma^q \mathbf{j}$, $c(\mathbf{i}|p) = c(\mathbf{j}|q)$ and $c_l = a_l$ for all $l = \{1, 2, \dots m\}$. Then

$$\Pi_T(\mathbf{i}) = E\Pi_T(\mathbf{j}) \tag{4.1}$$

for all T, where $E = f^{-1}_{i_1} \dots f^{-1}_{i_p} f_{j_q} \dots f_{j_1}$.

(vi) If $\mathbf{i} \in \Sigma^\infty$ is reversible and A has non-empty interior, then the support of $\Pi(\mathbf{i})$, namely $\bigcup\{t \in \Pi_A(\mathbf{i})\}$, is $\mathbb{R}^n$.

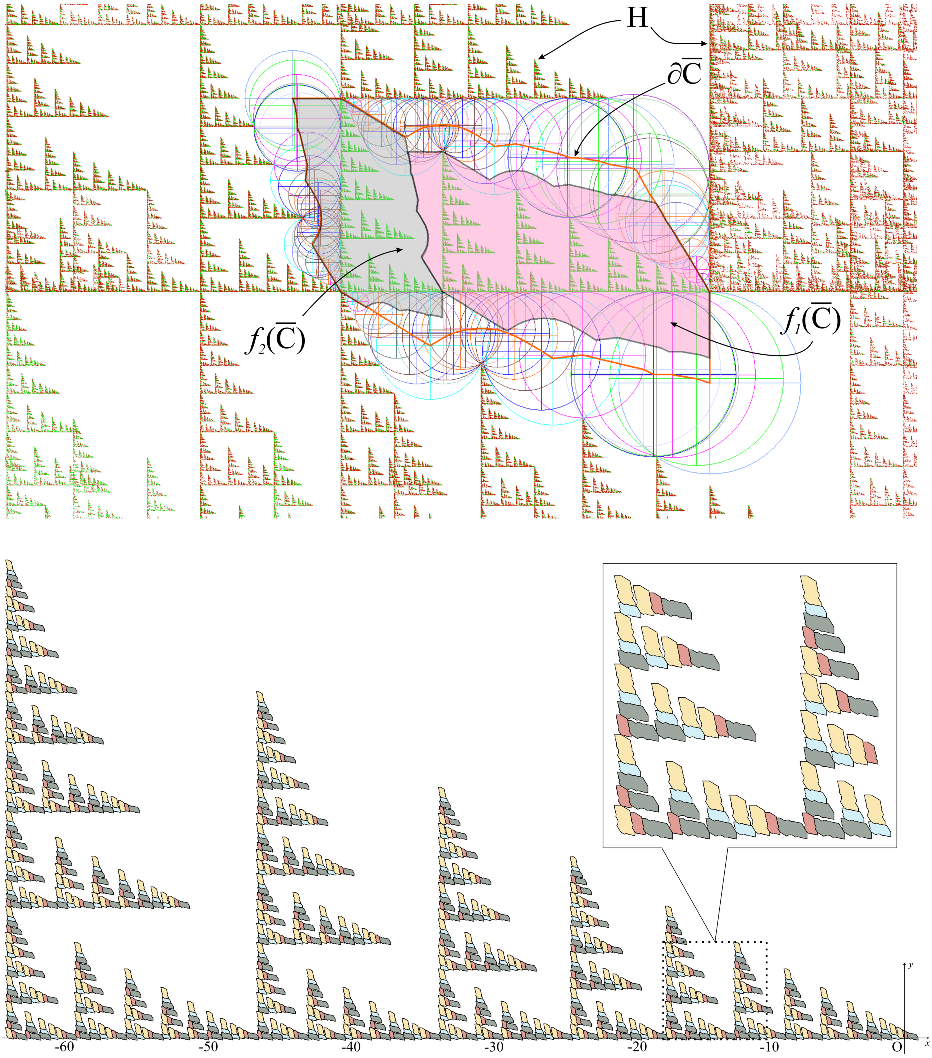


FIGURE 5. The top panel illustrates (i) H for the IFS in Example 5.1; (ii) $\partial\overline{C}$, the boundary of the central open set C; (iii) $f_1(\overline{C})$ and $f_2(\overline{C})$; (iv) circles whose centers approximate points on ∂C. The bottom panel illustrates $\Xi(1111111111111)$, and (inset) a patch of this tiling that also appears in Figure 1.

Proof. The initial assertion follows at once from Equation (3.1). This generalizes, in the case of a single vertex, a core result in [13].

(i) We need to show that for all $\mathbf{i} \in \Sigma^{\infty}$ the interiors of the tiles that comprise $\Xi(\mathbf{i}) = \Pi_{\overline{C}}(\mathbf{i})$ are disjoint. Suppose

$$f_{(\mathbf{i}|k)}^{-1} f_{(\mathbf{j}|l)}(C) \cap f_{(\mathbf{i}|p)}^{-1} f_{(\mathbf{t}|q)}(C) \neq \emptyset$$

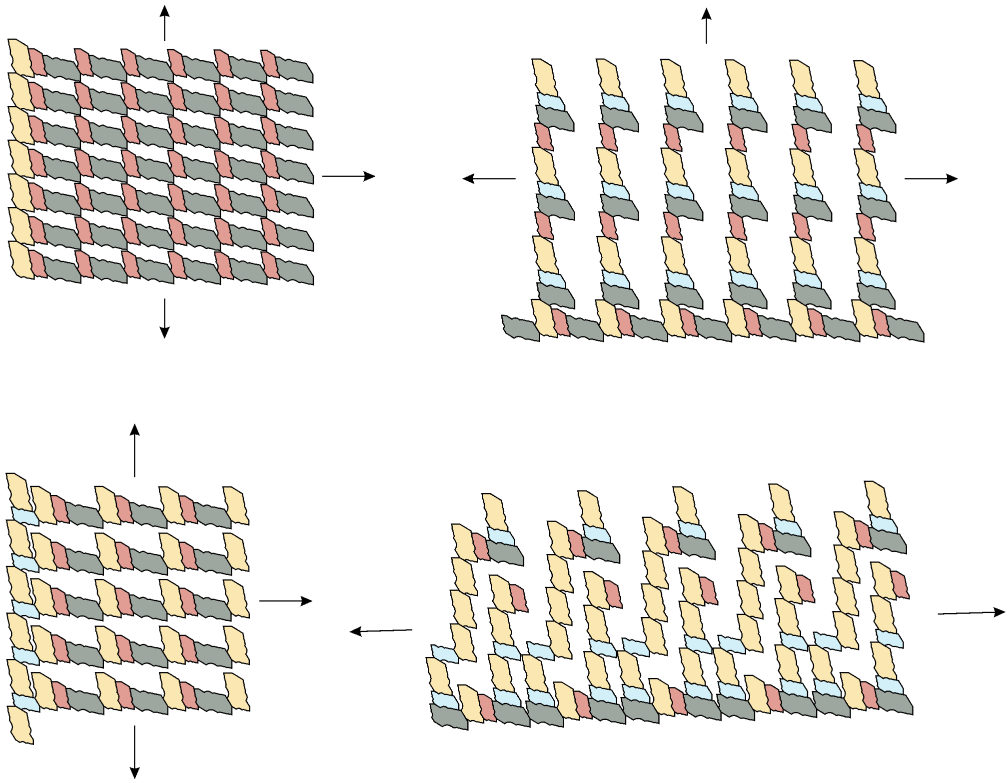


FIGURE 6. Examples of some tiling patterns that can be formed using the prototiles in Figure 7, when the tiles are allowed to touch only as in Figure 7. The arrows show directions in which part of the tiling may be repeated periodically.

for some $k, p, l, q \in \mathbb{N}$, $\mathbf{j}, \mathbf{t} \in \Sigma^\infty$ such that

$$c(\mathbf{j}|l-1) \leq c(\mathbf{i}|k) < c(\mathbf{j}|l) \text{ and } c(\mathbf{t}|q-1) \leq c(\mathbf{i}|p) < c(\mathbf{t}|q).$$

We can assume $k < p$, $i_k \neq j_1$ and $i_p \neq t_1$. It follows that

$$f_{j_1}...f_{j_l}(C) \cap f_{i_{k+1}}^{-1} \dots f_{i_p}^{-1} f_{t_1}...f_{t_q}(C) \neq \emptyset.$$

This implies

$$f_{i_p} \dots f_{i_{k+1}} f_{j_1}...f_{j_l}(C) \cap f_{t_1}...f_{t_q}(C) \neq \emptyset.$$

This implies

$$f_{i_p}(C) \cap f_{t_1}(C) \neq \emptyset,$$

where $i_p \neq t_1$ which contradicts the fact that C obeys the OSC.

(ii) This is an exercise in change of coordinates. We show that the interior of $f_{(\mathbf{i}|k)}^{-1} f_{(\mathbf{j}|l)}(\overline{C})$ is the central open set for the iterated function system SFS^{-1} where $S = f_{(\mathbf{i}|k)}^{-1} f_{(\mathbf{j}|l)}$ is a similitude.

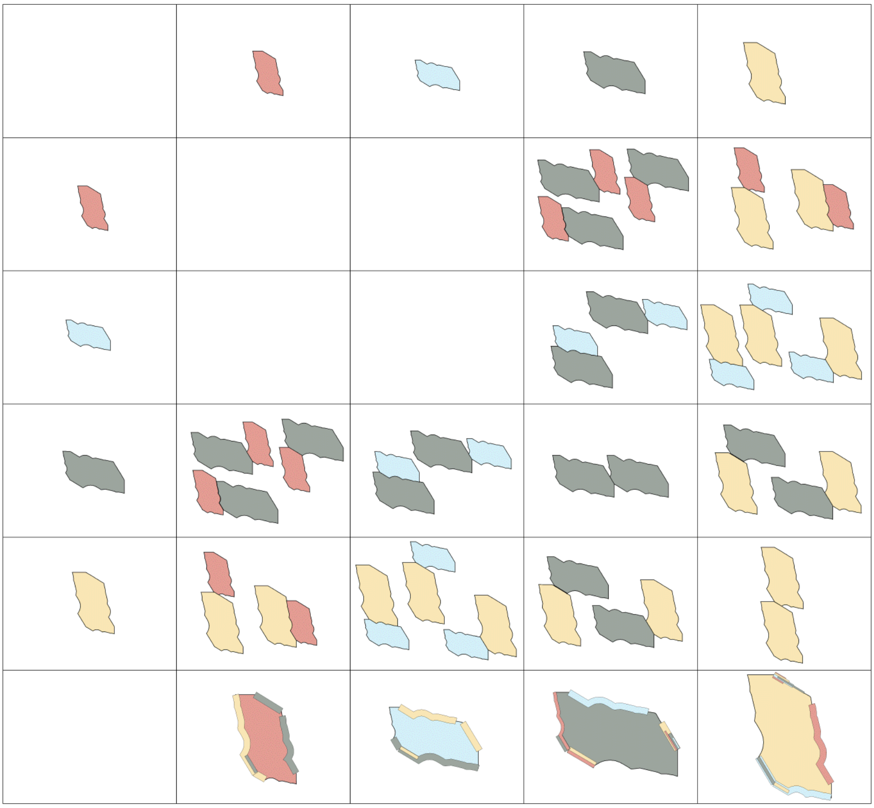


FIGURE 7. Left column: The four prototiles, with respect to translations, in Figures 1, 5, 6. The bottom row illustrates the points of contact in every central open set tiling $\Xi(\mathbf{i})$, $\mathbf{i} \in \Sigma^\infty$. The remaining boxes illustrate all of the allowed combinations between pairs of tiles, indexed by row and column.

Let S be the similitude $f^{-1}_{(\mathbf{i}|k)} f_{(\mathbf{j}|l)}$. Then $A' := f^{-1}_{(\mathbf{i}|k)} f_{(\mathbf{j}|l)}(\overline{C}) = SA$ is the attractor of the IFS $F' = SFS^{-1} := \{\mathbb{R}^n; Sf_i S^{-1} | i = 1, 2, \ldots m\}$. The neighbor maps of F' are $\left\{ Sf^{-1}_{\mathbf{p}} f_{\mathbf{q}} S^{-1} | p_1 \neq q_1, \mathbf{p}, \mathbf{q} \in \Sigma^* \right\}$, so the central open set for F' is

$$\begin{aligned} C' &= \{x' \in \mathbb{R}^m | d(x', A') < d(x, C')\} \\ &= \{x' \in \mathbb{R}^m | d(x', SA) < d(x', SC)\} \\ &= \{Sx \in \mathbb{R}^m | d(Sx, SA) < d(Sx, SC)\} \\ &= S\{x \in \mathbb{R}^m | d(x, A) < d(x, C)\} = SC \end{aligned}$$

where in the penultimate step we have used the fact that S is a similitude.

(iii) This follows from [10], which considers the case of self-similar polygonal tilings, upon noting that if A has non-empty interior and obeys the OSC then $A = \overline{C}$ is a polygon.

(iv) If $c_i = a_i \in \mathbb{N}$ for all i, then it is readily seen that each tile in $\Pi(\mathbf{i})$ is a copy of T scaled by s^a for some $a \in \{1, 2, \ldots a_{\max}\}$ where $a_{\max} = \max\{a_1, a_2, \ldots a_m\}$.

(v) Since $i_{p+1}i_{p+2}\cdots = j_{q+1}j_{q+2}\ldots$ it follows that

$$\Pi_T(\sigma^p \mathbf{i}) = \Pi_T(\sigma^q \mathbf{j})$$

where $\sigma : \Sigma^\infty \to \Sigma^\infty$ is the shift operator. Since $c(\mathbf{i}|p) = c(\mathbf{j}|q)$, we have

$$\{f_{j'_1} \ldots f_{j'_l}(T) | \mathbf{j}' \in \Sigma^\infty, l \in \mathbb{N}, c(\mathbf{j}'|l-1) \leq c(\mathbf{i}|p) < c(\mathbf{j}'|l)\}$$
$$=$$
$$\{f_{j_1} \ldots f_{j_l}(T) | \mathbf{j}' \in \Sigma^\infty, l \in \mathbb{N}, c(\mathbf{j}'|l-1) \leq c(\mathbf{j}|q) < c(\mathbf{j}'|l)\}.$$

The result now follows from

$$f_{i_p} \cdots f_{i_1} \bigcup_{k \geq p} \{f_{i_1}^{-1} \cdots f_{i_k}^{-1}(\{f_{j_1} \ldots f_{j_l}(T) | \mathbf{j}' \in \Sigma^\infty, l \in \mathbb{N}, c(\mathbf{j}|l-1) \leq c(\mathbf{i}|k) < c(\mathbf{j}|l)\})\}$$
$$=$$
$$f_{j_q} \cdots f_{j_1} \bigcup_{k \geq q} \{f_{j_1}^{-1} \cdots f_{j_k}^{-1}(\{f_{j_1} \ldots f_{j_l}(T) | \mathbf{j}' \in \Sigma^\infty, l \in \mathbb{N}, c(\mathbf{j}|l-1) \leq c(\mathbf{i}|k) < c(\mathbf{j}|l)\})\}.$$

(vi) This follows similar lines to [9, 13] and is omitted here. The argument there rests on the observation that, for reversible addresses $\mathbf{i}$, the support of the tiling $\Pi(\mathbf{i}|k)$ is contained in the interior of the tiling $\Pi(\mathbf{i}|k+l)$ for large enough l. □

When the IFS F is rigid and $c_i = a_i \in \mathbb{N}$ for all $i = 1, 2, \ldots m$, a converse of (v) in the Theorem is true. We say that F is *rigid* (with respect to translations) when the statement "T_k meets ET_l" for any $k, l \in \{1, 2, \ldots \max\{a_i\}\}$ and any translation E, implies "T_k is contained in ET_l or vice-versa". Both examples in Figure 2 are rigid. The only way that a rigid tiling $\Pi(\mathbf{i})$ can meet a translation of another rigid tiling $\Pi(\mathbf{j})$ is when Equation 4.1 holds. Such tilings cannot be periodic and have interesting properties, see [13] and references.

5. Calculations and examples

In this Section we present examples, including ones which show how we calculate approximations to central open sets.

Example 5.1. We consider the IFS $\{\mathbb{R}^2; f_1, f_2\}$ where

$$f_1 = \begin{bmatrix} s & 0 & 1 \\ 0 & s & 0 \end{bmatrix}, f_2 = \begin{bmatrix} 0 & s^2 & 0 \\ s^2 & 0 & 0 \end{bmatrix}$$

with $s + s^4 = 1$, $s > 0$. This IFS was mentioned in Example 3.3 and its attractor is illustrated at the bottom left in Figure 3 . In the top image in Figure 5 we illustrate the central open set and suggest how it was approximated. In this example and others the attractor A and parts of the fast basin B were calculated by using random iteration [7]. The portion of $B \backslash A$ closest to A was assumed to be the union of the sets $f_{\mathbf{i}}^{-1} f_{\mathbf{j}}(A)$ for $i_1 \neq j_1$ and $\mathbf{i}, \mathbf{j} \in \Sigma^4$, also computed by random iteration. In order to estimate points on $\partial C = \partial \overline{C}$, circles that appeared to touch both $B\backslash A$ and A were constructed. Calculations and constructions were performed on digital images of resolution 1024×1024. The bottom image in Figure 5 illustrates $\Xi(1111111111111)$. This picture was constructed by starting from a computed image of $\Pi(1111111111111)$ which is a translation of the canonical tiling T_{13}. Figure 6 illustrates four tiling patterns constructed using the tiling rules in Figure 7. The arrows point in directions in which a part of the pattern could be repeated periodically.

Example 5.2. We consider the IFS $F = \{\mathbb{R}^2; f_1, f_2, f_3\}$ defined by

$$f_1 = \begin{bmatrix} .85 & -.05 & .53842 \\ .05 & .85 & -.15789 \end{bmatrix},$$

$$f_2 = \begin{bmatrix} .17 & .22 & .195909 \\ -.22 & .17 & .776364 \end{bmatrix},$$

$$f_3 = \begin{bmatrix} -.17 & -.22 & .805 \\ -.22 & .17 & .776364 \end{bmatrix}.$$

The attractor is illustrated in green in the central zone of the top left panel in Figure 8. The attractor is totally disconnected although the image makes it appear to have connected components, because of digitization effects. The red set in the top left panel is an approximation to the relevant part of $B\backslash A$ and was calculated in the same way as in Example 5.1. A close-up is shown on the right, illustrating how we estimated the central open set. To make pictures of some tilings we chose $c_1 = 1$ and $c_2 = c_3 = 8$. The scaling factor for f_1 is $s_1 = 0.851\ldots$ while the scaling factors for f_2 and f_3 are $s_2 = s_3 = 0.278\ldots$ so $s_1^8 \simeq s_2 = s_3$, but $s_1^8 \neq s_2$. Thus $\Pi(\mathbf{i})$ and $\Xi(\mathbf{i})$ incommensurate tilings for any $\mathbf{i} \in\Sigma^\infty$. The bottom left image illustrates the tiling $\Xi(1111111)$ surrounded by $f_1^{-8}\partial C$. It illustrates the relationship between ∂C and some tiles. We observed that there appeared to be eight different tile sizes in any square patch of $\Pi(\mathbf{i})$ digitized at resolution 2048×2048, when keeping $s_1^{-8}A$ to be roughly the size of the viewing window. This accords with a comment in [4] regarding a result of Schief [21]: "*There exists an integer N such that at most N incomparable pieces $A_{\mathbf{j}}(= f_{\mathbf{j}}(A))$ of size $\geq \varepsilon$ can intersect the ε-neighborhood (sic) of a piece $A_{\mathbf{i}}$ of diameter ε.*" (The sets $A_{j_1\ldots j_n}$ and $A_{i_1\ldots i_m}$, referred to in the quote as "pieces", are said to be incomparable if there exists no $k_1\ldots k_p$ such that $j_1\ldots j_n = i_1\ldots i_m k_1\ldots k_p$ or $i_1\ldots i_m = j_1\ldots j_n k_1\ldots k_p$.) An example of part an unbounded tiling $\Pi(\mathbf{i})$ is illustrated at the bottom right in Figure 8.

Example 5.3. In Figure 9 we illustrate the calculation of the central open set for an IFS involving two maps. The method is the same as described in connection with Figures 8 and 5. Here the IFS is close to the one illustrated in the top row of Figure 3, see Example 3.3. Each map here is slightly more contractive and

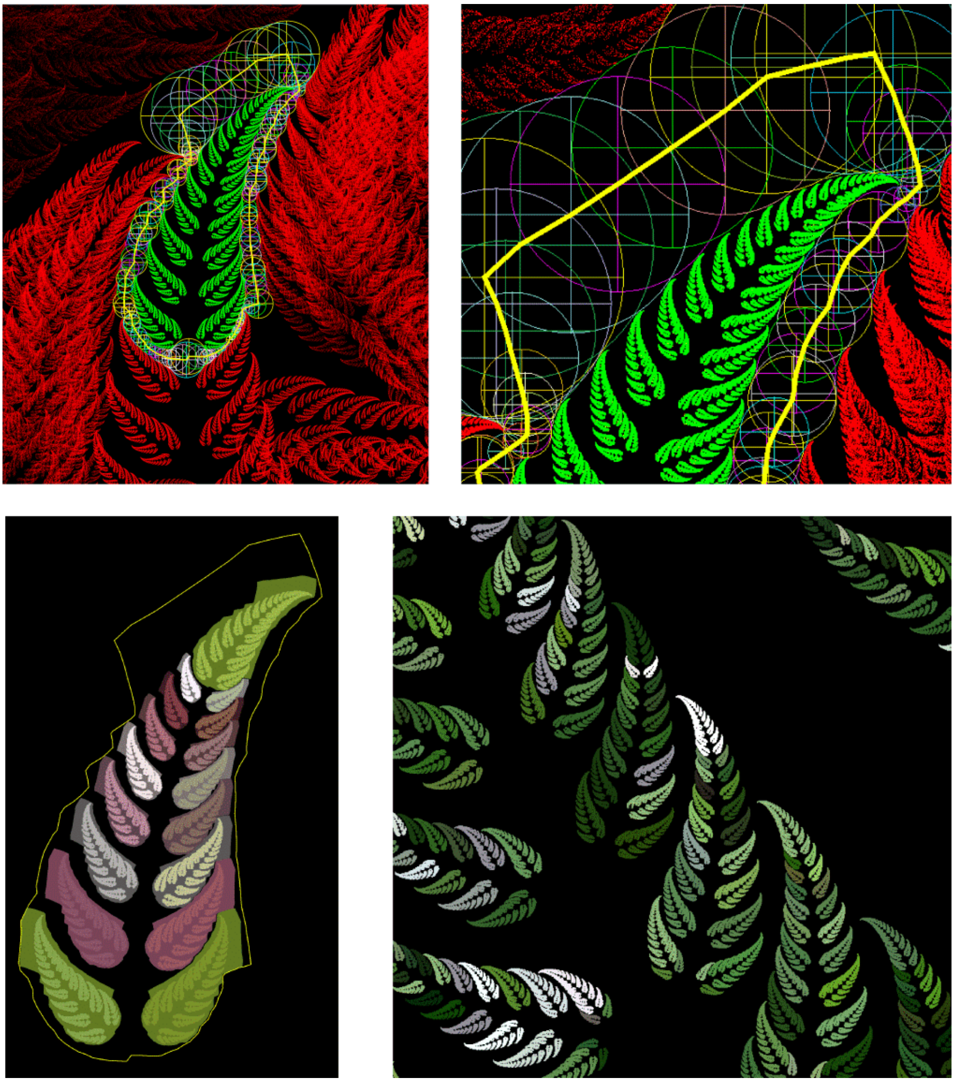


FIGURE 8. The top panels illustrate part of B/A where B is the fast basin and A is the attractor, for the IFS in Example 5.2; ∂C, the boundary (yellow) of the central open set C; and circles whose centers approximate points on ∂C. The bottom left panel shows scaled copies of C in part of a tiling of the form $\Xi(11111111)$ overlayed on $\Pi(11111111)$. The bottom right panel shows part of $\Pi(\mathbf{i})$ for some $\mathbf{i} \in \Sigma^{\infty}$.

rotated by a small amount. The IFS here is $\{\mathbb{R}^2; f_1, f_2\}$ where

$$f_1 = \begin{bmatrix} -.02447 & .777910 & 0 \\ -.77791 & -.02447 & .78615 \end{bmatrix},$$

$$f_2 = \begin{bmatrix} .61156 & -.019221 & -.019221 \\ -.019221 & .61156 & 0 \end{bmatrix}.$$

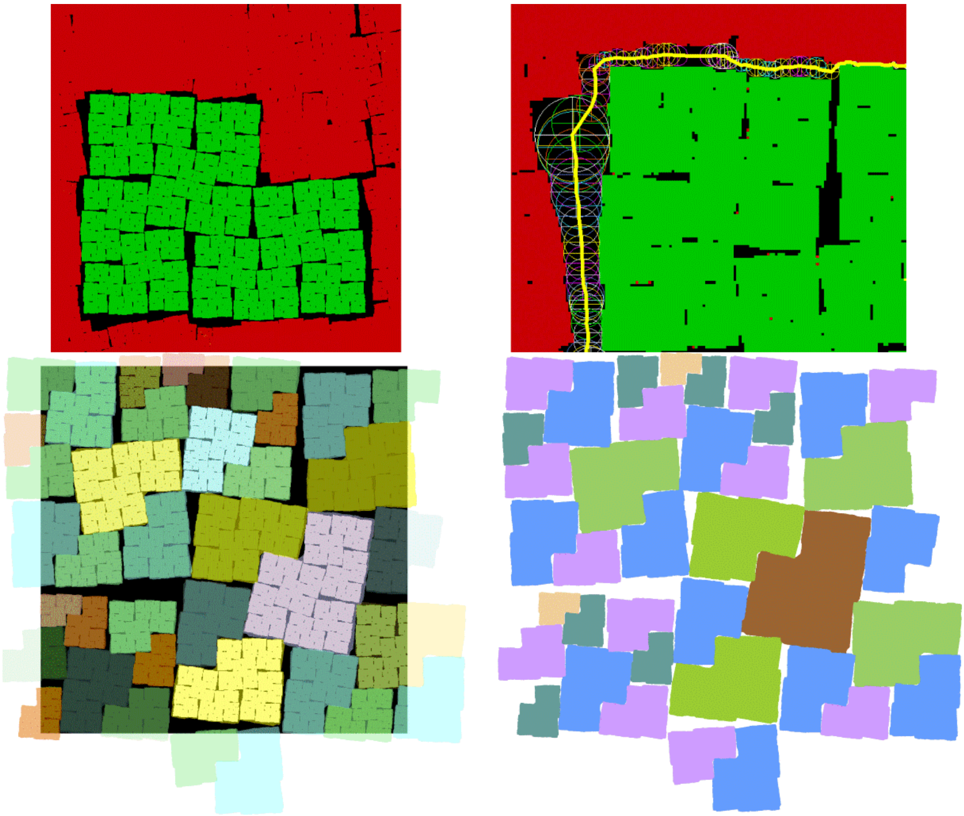


FIGURE 9. See Example 5.3. Top left illustrates the attractor of an IFS of two maps (green), and its fast basin minus the attractor (red). The top right hand image illustrates how the boundary, shown in yellow, of the central open set C was estimated. The lower left image shows a patch of $\Pi(\mathbf{i})$, specifically the tiles which meet a (dark) square. The lower right image shows the corresponding patch of of $\Xi(\mathbf{i})$.

As in Example 5.2 this numerical model is only approximately scaling, but we treat it as though it is by choosing $c_1 = 1$ and $c_2 = 2$. We find this example interesting because it suggests natural situations involving cracks.

Example 5.4. See Figure 10. The IFS here is $\{\mathbb{R}^2; f_1, f_2\}$ where

$$f_1 = \begin{bmatrix} .6413 & -.3283 & .3231 \\ .3283 & .6413 & -.133 \end{bmatrix}, f_2 = \begin{bmatrix} -.2362 & .4620 & .8052 \\ .4620 & .2362 & .5093 \end{bmatrix}$$

One can see how $\overline{C}$ is approximately tiled by $f_1(\overline{C})$ and $f_2(\overline{C})$ in the left image. That is, one can see the relationship between $\overline{C}$ and $\{f_1(\overline{C}), f_2(\overline{C})\}$. Note that

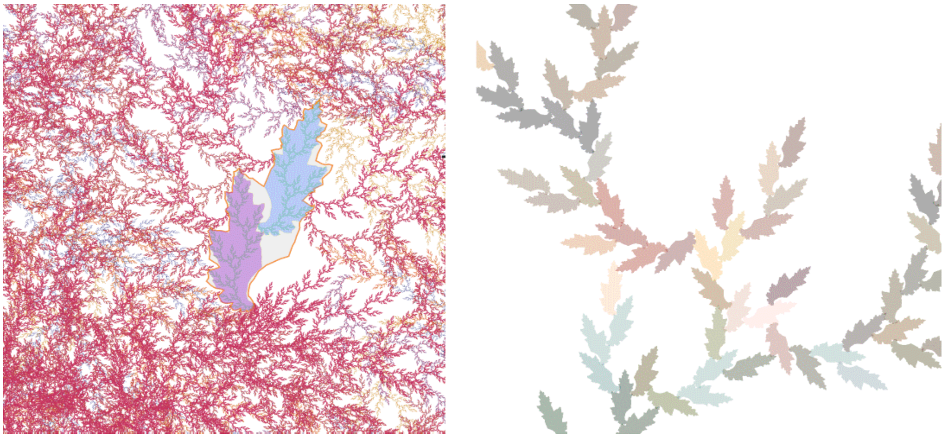


FIGURE 10. See Example 5.4. The left image shows an approximate central open set $\widetilde{C}$ and corresponding fast basin. The sets $f_1(\widetilde{C})$ and $f_2(\widetilde{C})$ are shown to fit neatly inside $\widetilde{C}$, and are surrounded by an approximation to $B\backslash A$. It may well be true that the attractor is overlapping and there is no central open set. Nonetheless, a tiling $\Pi_{\widetilde{C}}(\mathbf{i})$, a patch of which is illustrated on the right, is only slightly overlapping.

the actual attractor may not obey the OSC, our estimated "central open set" $\widetilde{C}$ may not obey the OSC, and the tiling $\Pi_{\widetilde{C}}(\mathbf{i})$ may be overlapping.

We thank to Krzystof Leśniak and Nina Snigreva for many helpful comments and suggestions.

References

[1] J. E. Anderson and I. F. Putnam, *Topological invariants for substitution tilings and their associated C^*-algebras,* Ergod. Th. & Dynam. Sys.**18** (1998), 509-537.

[2] T. Banakh, N. Novosad, *Micro and macro fractals generated by multi-valued dynamical systems*, Fractals **22** (2014), 140012 (https://doi.org/10.1142/S0218348X14500121)

[3] C. Bandt, S. Graf, *Self-similar sets VII. A characterization of self-similar fractals with positive Hausdorff measure*, Proc. Amer. Math. Soc **114** (1992), 995-1001.

[4] C. Bandt, N. V. Hung, H. Rao, *On the open set condition for self-similar fractals with positive Hausdorff measure*, Proc. Amer. Math. Soc **134** (2005),1369-1374.

[5] M. F. Barnsley, K. Lesniak, M. Rypka, *Basic topological structure of fast basins*, Fractals **26** (2018), 1850011. (https://doi.org/10.1142/S0218348X18500111)

[6] M. F. Barnsley, A. Vince, *Fast basins and branched fractal manifolds of attractors of iterated function systems*, SIGMA **11** (2015), 084, 21 pages. (https://doi.org/10.3842/SIGMA.2015.084)

[7] M. F. Barnsley, J. E. Elton, *A new class of Markov processes for image encoding,* Adv. Appl. Prob. **20** (1988), 14-32.

[8] M. F. Barnsley, A. Vince, *Developments in fractal geometry,* Bull. Math. Sci. **3** (2013), 299-348.

[9] M. F. Barnsley, A. Vince, *Fractal tilings from iterated function systems*, Discrete and Computational Geometry **51** (2014), 729-752.
[10] M. F. Barnsley, A. Vince, *Self-similar polygonal tilings*, Amer. Math. Monthly **124** (2017), 905-921.
[11] M. F. Barnsley, A. Vince, *Self-similar tilings of fractal blow-ups*, Contemporary Mathematics **731** (2019), 41-62.
[12] M. F. Barnsley, A. Vince, *Tilings from graph-directed iterated function systems*, Geom. Dedicata **212** (2021), 299-324. (https://doi.org/10.1007/s10711-020-00560-4)
[13] M. F. Barnsley, A. Vince, *Tiling Iterated Function Systems*, arXiv:2002.03538[math.DS]
[14] N. P. Frank, M. F. Whittaker, *A fractal version of the Pinwheel tiling*, Math. Intelligencer **33** (2011), 7-17.
[15] B. Grünbaum, G. S. Shephard, Tilings and Patterns, Freeman, New York (1987).
[16] J. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30** (1981), 713-747.
[17] R. Kenyon, *The construction of self-similar tilings*, Geometric and Functional Analysis **6** (1996), 471-488.
[18] J. Lagarias, Y. Wang, *Self-affine tiles in* R^n, Adv. Math. **121** (1996), 21-49.
[19] M. Morán, *Dynamical boundary of a self-similar set*, Fundamenta Mathematicae **160** (1999), 1-14.
[20] C. Radin, M. Wolff, *Space tilings and local isomorphisms,* Geom. Dedicata **42** (1992), 355-360.
[21] A. Schief, *Separation properties for self-similar sets,* Proc. Am. Math Soc **122** (1994), 111-115.
[22] Y. Smilansky, Y. Solomon, *Multiscale Substitution Tilings,* Proc. Lond. Math. Soc. (2021) (https://doi.org/10.1112/plms12404) (arXiv: 2003.11735v1 [Math.DS] 26 Mar 2020).
[23] B. Solomyak, *Dynamics of self-similar tilings*, Ergodic Theory & Dyn. Syst. **17** (1997), 695-738.
[24] R. S. Strichartz, *Fractals in the large,* Canad. J. Math. **50** (1998), 638-657.
[25] A. Tetenov, K. Kalamutdinov, *On one-point intersection property for self-similar fractals,* Nonlinearity **33** (2020), 408-416.

Australian National University, Canberra, Australia
LouisaBarnsley@gmail.com
Michael.Barnsley@anu.edu.au

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON SOME PARADOXES IN VOTING THEORY

KRZYSZTOF CIESIELSKI

ABSTRACT. We present some paradoxes concerning voting theory – both apportionment methods and elections of a winner. All the examples in the paper were constructed by the author. Mathematical background of some voting methods is explained. Also, the fundamental results on voting theory concerning the nonexistence of fair methods with some historical remarks are described. Finally, a new result on a weak method of n votes is presented.

KEYWORDS: voting method, apportionment method, proportional representation, method, method of n votes, Arrow's Theorem.

MSC2010: 91B12, 91B14, 00A05

Received May 14, 2021; accepted 28 July 2021

1. On some paradoxical phenomena in apportionment methods

In many countries members of the parliament are elected with the use of *apportionment methods*. Number of seats given to a party is determined by the number of votes it receives. It is frequently said that thanks to such a method we get a proportional representation of parties or political groups in the parliament. For example, one may suppose that a party receiving 20% of the national vote would receive approximately 20% of the seats in the parliament. Anyway, although those methods are called "proportional", very frequently an obtained result is far from proportional and those methods lead to many paradoxes.

In such a method, usually a country is divided into wards, i.e. electoral subdivisions. In each ward some members of the parliament are elected. In voting, each voter indicates a preference for one party and for one candidate on the list of this party. As the result, numbers of seats that are given for those parties in the ward are determined.

https://doi.org/10.2478/9788366675360-005.

Let us explain in the example the method currently used in several countries, including Poland (the method is known as the *Jefferson method* or the *D'Hondt method*). The example is shown in Table 1.

	votes (:1)	:2	:3	:4	:5	Representatives
Fees	$\star$ 24000	$\star$ 12000	$\star$ 8000	$\star$ 6000	4800	4
Bees	$\star$ 21600	$\star$ 10800	$\star$ 7200	5400	4320	3
Jees	$\star$ 12000	6000	4000			1
Kees	$\star$ 8000	4000				1
Nees	$\star$ 7000	3500				1
Pees	4200					–
Rees	4000					–

TABLE 1.

The rule is as follows (its mathematical background will be explained in Section 2). We divide numbers of votes by consecutive natural numbers, i.e. by $1, 2, 3, \ldots$ If in the ward there are n seats to be taken, then n greatest quotients give seats to respective parties. In Table 1 we assume that there are 10 representatives to be elected. The quotients that give seats to the parties are indicated by stars.

In this example we already see that the distribution may be far from proportional. In the case of *Bees*:*Jees* we have the votes with the proportion $1.8 : 1$ and the seats with $3 : 1$. If we consider *Fees*:*Jees*, we have $2 : 1$ against $4 : 1$. It is not difficult to construct more such examples. In fact, if a number of seats in a ward is not too large (say: about 10 or smaller), very many natural models lead to some disproportions.

In many countries, the *election threshold* is introduced. This means that to be considered in the distribution of seats, a party must receive a specified minimum percentage of votes (in Poland, it is 5%). What is important, this percentage must be obtained in the whole country, not in a considered ward. Let us come back to the previous example, now with the final result shown in Table 2.

Assume that although in this ward *Bees* obtained a very good score, in the rest of the country the voters did not support *Bees* so strongly and as the result this party did not obtain 5% votes in the whole country. Thus the party is not considered in the distribution of seats. Nevertheless, in the ward 10 seats are to be taken, so three seats that were previously given to *Bees* must be given to other parties. The quotients that give seats are again indicated by stars.

Now, the result is even more strange. We see that *Pees* got a seat and *Bees* not, although the proportion is $5.14 : 1$. Note that *Pees* got a seat in this ward just because of the lack of a suitable number of votes for *Bees* in other wards, perhaps in a region of the country which is very far from this one and where *Pees* do not apply for any seat. We may also see some disproportions if we compare seats for parties which gained seats in this ward.

	votes (:1)	:2	:3	:4	:5	Representatives
Fees	⋆ 24000	⋆ 12000	⋆ 8000	⋆ 6000	⋆ 4800	5
Bees	~~21600~~	~~10800~~	~~7200~~	~~5400~~	~~4320~~	–
Jees	⋆ 12000	⋆ 6000	4000			2
Kees	⋆ 8000	4000				1
Nees	⋆ 7000	3500				1
Pees	⋆ 4200					1
Rees	4000					–

Table 2.

Suppose that in a certain ward there is a perfect candidate and all the voters in this ward vote for this candidate. However, the number of voters in this ward is smaller than 5% of the population in the country, so this candidate will not be elected to the parliament. He (or she) will not get 5% votes in the whole country, although all the voters in this region wants him (or her) as their representative.

We may see other paradoxes. Consider the example presented in Table 3.

In a certain ward, where 5 seats are to be distributed, three parties apply for those seats. *Bees* will get 6000 votes, *Fees* 5700 votes, *Kees* 1950 votes (see upper rows of Table 3). The distribution 3 : 2 : 0 is reasonable. However, assume that just before the day of election a group of 600 voters (possible because anti-*Fees* agitation organized by *Bees*) gave up voting for *Fees*. Then 400 of them changed their opinion and voted for *Bees*, but 200 decided to give their votes to *Kees*.

If the method was logical, there would be only two possibilities. Either nothing would change (as the modification would not be essential enough), or *Fees* would lose a seat (or seats), first for the benefit of *Bees*.

	votes (:1)	:2	:3	:4	Repr.
Bees	⋆ 6000	⋆ 3000	⋆ 2000	1500	3
	⋆ 6400	⋆ 3200	2133	1600	2
Fees	⋆ 5700	⋆ 2850	1900		2
	⋆ 5100	⋆ 2550	1700		2
Kees	1950	975			–
	⋆ 2150	1075			1

Table 3.

However, the result is shown in lower rows of Table 3. Although *Fees* lost votes, the party did not lose any seat. *Bees* got extra votes, moreover – they got more votes than anyone else, but they lost a seat. In this example we see a paradox which shows that this method is far from proportionality and logical rules.

One may presume that when the voting is organized in many wards, then the disproportions would disappear. Indeed, this may happen. However, the story may be completely different, as is shown in the following example.

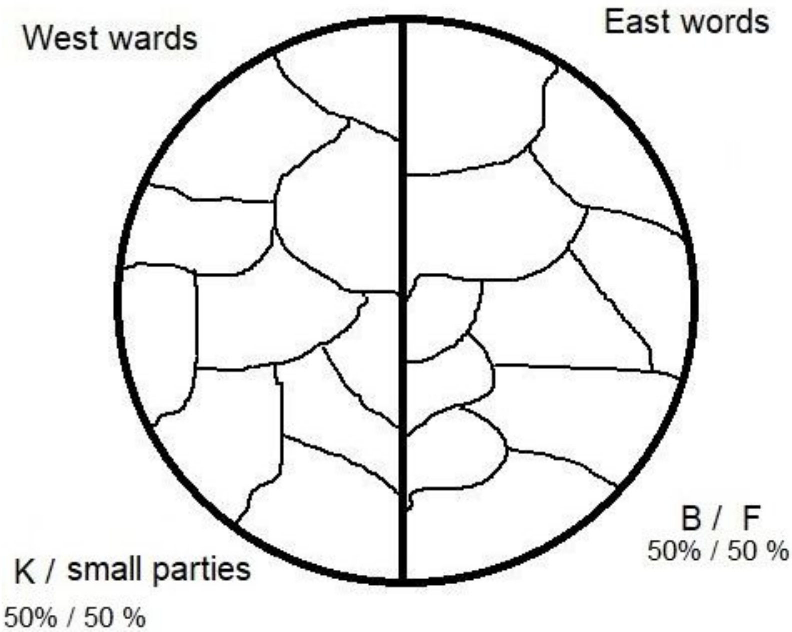


FIGURE 1.

Suppose that in a certain country the opinions of voters substantially differ in the East and in the West. In each East ward 50% voters vote for *Bees* and 50% voters vote for *Fees*. In each West ward 50% voters vote for *Kees* and other voters vote for many small parties which are finally eliminated because of the election threshold (see Figure 1). As the result, *Kees* get 50% seats in the parliament, *Bees* get 25%, and *Fees* get 25%, although all three parties have the same support in the whole country.

This example is artificial, but theoretically (as we consider mathematical models) possible. Nevertheless, generally parties that have representatives in the parliament get (in percent) much smaller support of voters than the number of seats in parliament.

Anyway, as a result of the election we get not anonymous seats, but concrete members of the parliament. Here the rule is also simple. Assume that *Bees* won n seats in a ward. As was mentioned above, each voter indicates one candidate on the list of the chosen party. Now, n candidates with the highest scores from this list are elected.

This also leads to strange phenomena. Consider the following example. Assume that in a ward 6 seats are to be taken and two parties apply for them: Mathematicians and Politicians. The result of voting is presented in Table 4.

	votes (:1)	:2	:3	:4	:5	:6
Mathematicians	$\star$1200	$\star$ 600	$\star$ 400	$\star$ 300	$\star$ 240	200
Politicians	$\star$ 201	101				

TABLE 4.

According to the rule in this method, 5 seats are won by Mathematicians and 1 seat by Politicians. Assume that the votes for the object in Mathematicians' list were as follows: Square 201, Triangle 200, Circle 200, Ball 200, Cube 200, Product 199. In the Politicians' list the greatest score (35) was due to Bureaucrat. Other politicians obtained the following results: one of them 34 and each of the remaining four candidates 33. In this model, Product gained 199 votes against 35 votes for Bureaucrat, but Product will not be elected, although Bureaucrat will become a member of the parliament. Note that Product itself got almost the same number of votes as Politicians altogether.

However, generally such situations rather do not occur in practice. Frequently the majority of votes goes to one candidate on the list (usually the one put in "pole position"). So, assume that Mathematicians gained 100000 votes and because of that have 4 representatives in the parliament. Personally, Triangle got 99995 votes, Tangent Bundle 2 votes, Fibre Bundle 2 votes and Frame Bundle 1 vote. However, in the parliament all the representatives are equal and three Bundles may force the act on the superiority of differential geometry over classical geometry. Not speaking about the case where a member of the parliament changes the party that put him on a list to another one, which happens in some countries. Although in fact three Bundles became members of the parliament thanks to votes for Triangle, being the representatives they may change their party and become members of the party of Physicists.

There are many apportionment methods. Another popular method is the *Webster method* (known also as the *Sainte-Laguë method*). Here the numbers of votes are divided by consecutive odd numbers, i.e. by 1, 3, 5, 7,. . . .

Let's come back to the example concerning Mathematicians and Polititians. If the seats were distributed according to the Sainte-Laguë method, even the paradox worse than previously may occur, see Table 5.

Now Mathematicians' candidates gained the following numbers of votes: Square 316, Triangle 315, Circle 315, Ball 315, Cube 315, Product 314. Product obtained much bigger support than all the politicians together (200), but Product will not become a member of the parliament although one politician will get a seat.

The second paradox will be shown in the example presented in Table 6.

	votes (:1)	:3	:5	:7	:9	:11
Mathematicians	⋆1890	⋆ 630	⋆ 378	⋆ 270	⋆ 210	171.8
Politicians	⋆ 200	66.6				

Table 5.

	votes (:1)	:3	:5	:7	:9	Representatives
Waiters	⋆ 1050	⋆ 350	⋆ 210	⋆ 150	117	4
Sportsmen	⋆ 1008	⋆ 336	⋆ 202	144	112	3
Waiters	⋆ 1050	⋆ 350	⋆ 210	150	117	3
Footballists	⋆ 504	⋆ 168	101	56		2
Basketballists	⋆ 504	⋆ 168	101	56		2

Table 6.

In this example, Waiters gained 1050 votes and got 4 seats of 7 (see upper rows of Table 6), whereas Sportsmen gained 1008 votes and got 3 seats of 7. This seems reasonable. However, if instead of Sportsmen's party two ,,sport parties" take part in the election and the votes initially intended for Sportsmen are equally distributed among two parties: Footballists and Basketballists, then the distribution of seats will be as 3 : 2 : 2 (see lower rows of Table 6). Note that Waiters get more than 50% votes, but do not have majority in the parliament! If Sportsmen predicted a possible result (for example on the base of a precise poll) they could artificially split their party into two parties to get majority in the parliament.

Let us move on to a practical surprising application. In 1997 the author of this paper published in a Polish popular monthly *Wiedza i Życie* (*Knowledge and Life*) a popular article [6] explaining the rules of elections. Then in Poland the so-called the *modified Sainte-Laguë method* was in use – here the first seat for the party was obtained for the division the number of votes by 1.4, the rest was the same as in Sainte-Laguë method. This method also shared paradoxes described above. The paradox on the possible splitting of the party to get more votes was described in this article. The article was noticed by a physicist, Jerzy Przystawa (1939–2012), who actively worked to introduce single-winner voting in Poland. Przystawa spread this property and called it *Ciesielski law* (*prawo Ciesielskiego*, see [9]). Then it was applied in real life. In 2002, in the elections of the City Council in Nysa, 23 seats were to be taken. The Mayor of Nysa, Janusz Sanocki (1954–2020) and other members of his group realized that the group would not get majority in the election and, consequently, would not have majority in the City Council. Thus they split their association into two: *Liga Nyska* and *Komitet Obywatelski Ziemi Nyskiej*. The first one won 9 seats in the City Council, the second won 3 seats. Together they had 12 seats of 23 in the City Council, so they had majority. If they had run in this election as one association, they would not have obtained a suitable results. The story is described in [13] as “a practical application of Ciesielski law”.

2. The mathematical background of apportionment methods

The apportionment methods (which nowadays in public opinion are associated mainly with elections) have a source completely different from party lists. In the House of Representatives in the United States the number of seats given to each state is determined by the population of this state. According to the Constitution of the US, the representatives should be apportioned to the states according to their respective numbers of persons. A natural question arises: how to apportion those seats? The problem appeared at the end of the 18th century. The number of states was changing, so rules must have been analyzed and possibly modified. Then, different methods were used and those methods were changing, as new paradoxes were being discovered and noticed. For the description of the history of those changes, see [10].

Let us introduce the mathematical background of these methods. The following description will be based not on elections to the parliament, but on the distribution of seats in the House of Representatives.

First, fix the notation. Assume that there are k states that should have their representatives. Denote the populations in these states by $p_1, \ldots, p_k$ and assume that $p_1 + \ldots + p_k = p$, which is total population in the country. The numbers of seats given to states are $s_1, \ldots, s_k$ and $s_1 + \ldots + s_k = s$, where s is the number of seats in the House of Representatives. Mathematically, in an apportionment method we have to find a function

$$(k, s, (p_1, \ldots, p_k)) \;\mapsto\; (s_1, \ldots, s_k)$$

satisfying the condition $s_1 + \ldots + s_k = s$. Of course, the number of seats for a state must be an integer.

An important concept in the theory is *quota*. A quota of a state i is defined as $q_i = \frac{p_i}{p} \cdot s$. Roughly speaking, quota represents the appropriate number of seats which should be assigned to a state, if the distribution is done fairly. Unfortunately, usually quota is not an integer. We define a standard divisor $d = \frac{p}{s}$ (roughly speaking, this is „the value of a seat”). Note that $q_i = \frac{p_i}{d}$, i.e a result of the division of the population of state i by the standard divisor.

We define the lower quota as the floor $\lfloor q_i \rfloor$ of the quota q_i, and the upper quota as the ceiling $\lceil q_i \rceil$ of the quota q_i.

It is natural that each stage should get at least its lower quota seats. When we give to each state its lower quota, it is almost certain then some seats will not be distributed. It seems also natural that one may give remaining seats to the states with the highest differences $q_i - \lfloor q_i \rfloor$. This method, called the *Hamilton method* or the *largest remainder method* has been used several times. In fact, it was the first method suggested to be used in the apportionment in the United States, but then it was vetoed. Nevertheless, later it was used from time to time. This method admits some original paradoxes, in particular the Alabama Paradox and the Oklahoma Paradox. In short, the Alabama Paradox is connected with the situation where the number of all seats is enlarged and as a consequence one state loses a seat. The Oklahoma Paradox deals with the situation where an extra state joins the country, so some seats (according to the proportion in populations), say n, have to be added to the House of Representatives, and then this extra state really gets n seats, but the numbers of seats for some other states change. The precise description of those paradoxes may be easily found in the literature, for example in [10].

Let's move to the *Jefferson method*, proposed at the end of the 18th century by Thomas Jefferson (then a Secretary of State, later the third President of the US) and, independently, about a century later, by Belgian lawyer Victor D'Hondt. Assume that we want to give to each state its lower quota $\lfloor q_i \rfloor = \lfloor \frac{p_i}{d} \rfloor$; recall that $d = \frac{p}{s}$ is a real value of a seat. Then the situation in which not all seats will be taken is an event with probality close to one. Let us modify a standard divisor d and take a number $\widetilde{d}$ (in fact, $\widetilde{d} < d$). This gives us modified quotas $\widetilde{q_i} = \frac{p_i}{\widetilde{d}}$ and consequently modified lower quotas $\lfloor \widetilde{q_i} \rfloor$. We change the divisor up to the moment when the numbers of seats obtained by modified lower quotas sum up to s. Almost always seats obtained by lower quotas with standard divisor d result with a number smaller than s, so an appropriate modified divisor $\widetilde{d}$ is slightly smaller than d. It is proved that the distribution of seats by this method gives the same result as the algorithm of dividing p_i by $1, 2, 3, \ldots$ and taking s greatest quotients.

If in this procedure we will consider upper quotas instead of lower quotas, we get the *Adams method* (named by John Quincy Adams, the sixth President of the US). Then an appropriate modified divisor $\widetilde{d}$ is slightly greater than d. We may round the modified quota $\widetilde{q_i} = \frac{p_i}{\widetilde{d}}$ to the nearest integer (round up if $\widetilde{q_i} = k.5$ for some k) – this gives the *Webster method*. This method was later independently introduced in Europe by André Sainte-Laguë and here the distribution of seats boils down to dividing p_i by consecutive odd numbers. Note that in this method a standard divisor may be appropriate.

In the Webster method rounding to the nearest integer may be presented as a comparison of $\widetilde{q_i}$ to the arithmetic mean $\widetilde{A_i} = \frac{\lfloor \widetilde{q_i} \rfloor + \lceil \widetilde{q_i} \rceil}{2}$: if $\widetilde{q_i} \geq \widetilde{A_i}$, we round it up, and if $\widetilde{q_i} < \widetilde{A_i}$, we round it down. However, we may consider here also other means and they are also taken into account in certain methods. In the case of the geometric mean we have the *Hill–Huntington method* (since 1930s used in the distribution of seats in the House of Representatives in the US) and in the case of the harmonic mean we have the *Dean* method.

A good apportionment method does not exist. Below are stated three natural conditions. If in a certain method of apportionment any of them is not fulfilled, then a method cannot be regarded as fair. We use the notation introduced above. The conditions are:

- *quota condition*: $\lfloor q_i \rfloor \leq s_i \leq \lceil q_i \rceil$ for any state i
- *monotonicity property*: $p_i > p_j \;\Rightarrow\; s_i \geq s_j$ for any states i, j
- *population property*: assume that k and s are given, but populations and seats of states change (we denote it by $a \mapsto \overline{a}$); then there are no i, j with $\overline{p_i} > p_i$, $\overline{s_i} < s_i$ and $\overline{p_j} < p_j$, $\overline{s_j} > s_j$.

In other words, quota condition says that each state should not get less seats than its lower quota and more seats than its upper quota (note that this implies that if quota is an integer, then a state gets its quota). Monotonicity property means that one state is allowed to get more seats than another only if it has a greater (or equal) population. Population property says that it is impossible that the population of one state increases and this state loses a seat, but simultaneously the population of another state decreases and this second state gets an extra seat.

In 1982 Michel Balinski and H. Peyton Young proved in [4] a fundamental theorem.

Theorem 2.1 (Balinski-Young Theorem). *If $k \geq 4, s \geq 7$, then there is no apportionment method satisfying all three conditions: quota condition, monotonicity property, and population property.*

To end this chapter note that although apportionment methods are frequently paired off with the elections, they are applicable to many situations – in fact, frequently in much reasonable sense. We need to use apportionment methods in many situations in real life. Say, there are some shareholders that contributed in some venture and their contributions to this venture are different. As the result of their contributions they should get some goods, let's call them shares. Contributions and shares are expressed in integers and shares are expressed in much smaller integers. It is easy to give several examples from real life. The question is: how to distribute shares fairly?

Note that one share is indistinguishable from another one. In the distribution of seats to the House of Representatives in the United States the seats are given to states and there is not indicated who personally would get a seat. Some paradoxes in the application of apportionment methods applied to voting appear because the concrete people are elected simultaneously with determining of numbers seats; in other words, shares (seats) are distinguishable.

3. Single winner voting methods.

Now we turn to methods that lead to the selection of a winner. Of course, such a selection need not be connected with politics, parliament etc. First fix the terminology.

Assume that two sets are given: $\mathbf{V}$ – a *set of voters* and $\mathbf{C}$ – a *set of candidates*. Each voter ranks all candidates (an order in $\mathbf{C}$ fixed by a particular voter is called a *voter's preference order*). A collection of all voters' preference orders is called a profile. A *winner method* is a method that on the base of a profile outputs a set of winners (a subset of $\mathbf{C}$, possibly empty) – of course, we prefer getting a single winner. Formally, if we denote a set of all profiles by $\mathbf{\Sigma}$, then a winner method is a function $f\colon \mathbf{\Sigma} \to 2^{\mathbf{C}}$. If by $Or(\mathbf{C})$ we denote the set of all (linear) orders in $\mathbf{C}$, then $\mathbf{\Sigma} = \{M\colon \mathbf{V} \to Or(\mathbf{C})\}$.

Now consider the following example. Assume that five candidates: Pooh, Tigger, Rabbit, Kanga and Eeyore apply for a title of Milne's Star and it will be decided by 55 voters. Of course, it is possible that two different voters will rank candidates with different orders (as $5! > 55$). We consider much simpler profile: there are only six possible ordered preferences. They are shown in Table 7 (in the first row, numbers of voters that rank candidates this way is written).

	14	11	10	9	6	5
1.	P	T	K	R	E	P
2.	E	K	R	E	K	R
3.	T	R	E	K	R	E
4.	R	E	T	T	T	T
5.	K	P	P	P	P	K

TABLE 7.

Now consider five different voting methods.

In so-called *plurality method* the winner is the candidate who is ranked as the first choice of the most voters. Here Pooh wins with 19 votes.

Now take so-called *president method*. Then the winner is the candidate who gets more than 50% first position. If such a candidate does not exist, we take two candidates with the greatest number of first-place votes and compare them. In our case, the winner is Tigger, who in second round conquers Pooh.

Another method is the *Hare method*. We find the candidate (or candidates) who has the fewer first-place votes and eliminate this candidate. Then voters rearrange their lists (removing that candidate and moving up all the candidates that were ranked lower), and vote again. We proceed up to the moment when there is a candidate with a majority of first-place votes. This method was for many years used in the elections of rector in many Polish universities. In our profile, Kanga wins (the first eliminated candidate is Eeyore, then Rabbit, Tigger and Pooh).

We may use points. If there are n candidates, we may assign points for positions of voters' lists: $n-1$ points for the first position, $n-2$ points for the second position, …, 0 points for the last position. A sum of points obtained by candidates in all lists in the profiles determine the winner. This is called the *Borda Count* *method*. In the case of the profile presented in Table 7 Eeyore is the winner with 134 points (R:129, K:109, T:102, P:76).

Now we need to find a profile where Rabbit wins. We may use pairwise comparison between candidates. For two candidates, say A and B, let's count how many voters rank A above B and conversely. If the number of voters that rank A above B is greater than the number of voters that rank B above A, then A conquers B. If there is a candidate who conquers anyone else, then that candidate seems to be an obvious winner. In our case, Rabbit is such a candidate (we have: R:P – 36:19, R:T – 30:25, R:M – 35:20, R:K – 28:27).

The disadvantage of this metod is that in several cases it will not determine a winner (there is a famous *Condorcet paradox*: A conquers B, B conquers C, C conquers A). But it is easy to repair this method; a modified method is called the *Copeland method*. If A conquers B in pairwise comparison, A gets one point. If the comparison results in a tie, a method gives half a point to A and half a point to B. Finally, a candidate (candidates in the case of a tie) with the greatest number of points is a winner. This method is used, for example, in the Handball World Championship or the Handball European Championship, in preliminary round (with points doubled by 2, i.e. 2, 1, 0). For many years it was used in football championship, now 2 point for a winner are replaced by 3 points.

Note that the profile was not artificial. Moreover, all methods are quite natural and logical. In fact, all of them have been used somewhere. However, each of them gives another winner.

Anyway, note that in each of above methods the winner had a significant support. This is one of advantages of such a method in comparison to the election of representatives on the base of the list of a party in an apportionment method. Here the case where a winner gets a very small number of votes is impossible.

Also in the case of voting methods that select the winner, there is the theorem on the nonexistence of a good method. Before presenting a fundamental theorem we introduce some notation and formulate some conditions.

Consider more detailed voting method. As a result of voting, we do not pick up only winners, but a weak order in the set of candidates. A weak order is defined as follows. Take an equivalence relation $\sim$ in $\mathbf{C}$ and consider a linear order in the quotient space $\mathbf{C}/\sim$. Then for $x, y \in \mathbf{C}$ we say that $x < y$ if and only if $[x]_\sim < [y]_\sim$. In other words, we rank elements of $\mathbf{C}$ and admit ties. If as a result of voting we consider a weak order in $\mathbf{C}$, we call a method an *order voting method*. Of course, a winner method is a particular case of an order voting method – we have two classes in the set $\mathbf{C}$: winners and losers.

Formally, if we denote by $We(\mathbf{C})$ the set of all weak orders in $\mathbf{C}$, the order voting method is a function $f\colon \mathbf{\Sigma} \to We(\mathbf{C})$; recall that $\mathbf{\Sigma} = \{M\colon \mathbf{V} \to Or(\mathbf{C})\}$.

The notation $A \overset{v,M}{<} B$ means that in profile M voter v prefers B to A. The notation $A \underset{M}{<} B$ means that as a result in profile M candidate B is ranked higher than A.

An order voting method

- is *anonymous* if
$$\forall M \in \mathbf{\Sigma}\ \forall w, v \in \mathbf{V}$$
$$w \text{ and } v \text{ exchange their votes} \Rightarrow \text{result does not change;}$$
- satisfies *Pareto rule* if
$$\forall M \in \mathbf{\Sigma}\ \forall A, B \in \mathbf{C}$$
$$(\forall v \in \mathbf{V}:\ A \overset{v,M}{<} B) \ \Rightarrow\ A \underset{M}{<} B;$$
- satisfies *Independence of Irrelevant Alternatives* (IIA) if
$$\forall M, N \in \mathbf{\Sigma} \text{ and } \forall A, B \in \mathbf{C}$$
$$(\forall v \in \mathbf{V}:\ A \overset{v,M}{<} B \Leftrightarrow A \overset{v,N}{<} B) \Rightarrow (A \underset{M}{<} B \Leftrightarrow A \underset{N}{<} B)$$

In other words, a method is anonymous if the votes of all voters are equal. The Pareto rule means that if all voters prefer B to A, then as a result B is ranked above A. The condition IIA means that if voters change their votes but none of them changes the relation between A and B, then the final relation between A and B will not change (that is, the opinion about other candidates should not impact on the final relation of A and B; we may also say that the final preference between two candidates depends only on the individual voter's preferences between those two candidates).

Fundamental Arrow's Impossibility Theorem proved in 1950 by Kenneth Arrow (Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel winner in 1972), see [1] and [2], says:

Theorem 3.1 (Arrow's Impossibility Theorem). *If* $\mathbf{C}$ *contains at least 3 elements and* $\mathbf{V}$ *contains at least 2 elements, then there is no anonymous order voting method satisfying both the Pareto rule and IIA.*

The theorem may be presented in slightly more general way. For this purpose, introduce the next definition. By the *dictatorship* we mean a method where the final results is identical with a vote of one particular voter. Of course, this method is not anonymous. Arrow's Theorem says:

Theorem 3.2 (Arrow's Impossibility Theorem). *If* $\mathbf{C}$ *contains at least 3 elements and* $\mathbf{V}$ *contains at least 2 elements, then the only order voting method satisfying both the Pareto rule and IIA is dictatorship.*

Arrow's Theorem may be formulated in more general mathematical form, without referring to voting. Note that the Pareto rule and IIA may be formulated just for ordered sets. Let us keep the notation introduced above and present a next definition.

A set $\mathbf{T} \subset \mathbf{V}$ is called a *decision set* if

$$\forall M \in \mathbf{\Sigma}\ \forall A, B \in \mathbf{C}\ \ (\forall v \in \mathbf{T} :\ A \overset{v,M}{<} B)\ \Rightarrow\ A \underset{M}{<} B.$$

In the language of voting this means that if all voters in $\mathbf{T}$ prefer B to A, then finally B is ranked above A.

In the following theorem we do not assume that $\mathbf{V}$ and $\mathbf{C}$ are finite. For the basic information of filters, see for example [3].

Now we have

Theorem 3.3 (Arrow's Theorem in form of ultrafilters). *If $\mathbf{C}$ contains at least 3 elements and $\mathbf{V}$ contains at least 2 elements, $f : \mathbf{\Sigma} \to We(\mathbf{C})$, f satisfies both the Pareto rule and IIA, then the family of decision sets is an ultrafilter on $\mathbf{V}$.*

Arrow's Impossibility Theorem for voting theory is an immediate consequence of above theorem, as an ultrafilter on a finite set must be generated by one element, so one voter determines the final result.

Arrow's Theorem in form of ultrafilters was first published by Alan P. Kirman and Dieter Sondermann in [7]. The authors credit this idea to Don Brown and Peter S. Fishburn.

For more information on Arrow's Impossibility Theorem see [8], [10] and [14].

4. Methods of k votes

The topic of the last chapter are methods of k votes. Again, first introduce the terminology.

As previously, assume that each voter ranks candidates (there are n candidates). The method is called a *positional method* and denoted $P(a_1, a_2, \ldots, a_n)$ if a candidate obtains a_1 points for each first-place vote, a_2 points for each second-place vote and so on. We assume that $a_1, \ldots, a_n$ are integers and $a_1 \geq a_2 \geq \ldots \geq a_n \geq 0$.

We have already considered such methods. The plurality method is a positional method $P(1, 0, \ldots, 0)$. The Borda Count method is a positional method $P(n-1, n-2, \ldots, 1, 0)$.

Now we consider a method

$$P(\underbrace{1, 1, \ldots, 1}_{k \text{ times}}, \underbrace{0, 0, \ldots, 0}_{n-k \text{ times}})$$

and denote it by $P_n(k)$. In other words, each voter votes for k candidates, treating them equally, although the voter ranked these k candidates. If as the result we obtain one winner (or more in the case of ties) we call this method *the method of k votes*. If as result we obtain k winners (or more in the case of a tie in the last winning position) we call this method *the weak method of k votes*. Both methods are in use on several occasions. For example, the weak method of k votes is used in many Polish universities in the election of students' representatives or young academic staff representatives in the Faculty Council. Sometimes election threshold 50% for a candidate is required. Then if l candidates ($l < k$) obtain this threshold, they are elected and the voters vote again according to their lists, but now the list of candidates is reduced to $2(k - l)$ candidates with greatest scores in the first round (not elected yet) and so on.

First take into account a method of k votes. Consider profiles presented in Table 8 and Table 9.

	4	3	6
1.	A	A	C
2.	B	C	B
3.	C	B	A

TABLE 8.

In profile presented in Table 8 there are 13 voters and 3 candidates. A wins under the method of 1 vote (with 7 votes), B wins under the method of 2 votes (with 10 votes). Considering here the method of 3 votes is useless, as all the candidates would have equal results, but we may find the winner under the Borda Count method. Here it is C with 15 points (A obtains 14 points, B obtains 10 points).

	4	2	2	1
1.	D	A	A	A
2.	B	B	C	D
3.	C	C	D	C
4.	A	D	B	B

TABLE 9.

In profile presented in Table 9, A wins under the method of 1 vote (with 5 votes), B wins under the method of 2 votes (with 6 votes), C wins under the method of 3 votes (with 9 votes), D wins under the Borda Count method with 16 points (A: 15, B: 12, C: 11).

This paradoxical effect may be generalized for each number of candidates greater than 2. In 1992 Donald G. Saari proved the following theorem.

Theorem 4.1 (Saari's Theorem)**.** *For any $n \geq 3$ and the set of candidates $\{c_1, c_2, \ldots, c_n\}$ there exist a profile so that c_k wins the election under the method of k votes $P_n(k)$ for $k = 1, 2, \ldots, n-1$ and c_n wins under the Borda Count method.*

The theorem is presented and clearly explained (without proof) in [11]. The original proof was published in [12] and was based on advanced geometrical theory constructed in this purpose. However, the theorem may be proved elementary ([5]).

Now turn to weak method of k votes.

The weak method of two votes was used in the election to Polish Senate in 1990–2010. In all (except two) election wards two senators were elected. Consider the example presented in Table 10.

	30%	30%	20%	20%
1.	Pooh	Pooh	Rabbit	Tigger
2.	Rabbit	Tigger	Tigger	Rabbit
3.	Tigger	Rabbit	Pooh	Pooh

TABLE 10.

If only one candidate was elected in one-vote method, Pooh would be the winner. It is obvious that Pooh is the best candidate, as 60% of voters put Pooh in the first place. However, in the weak method of 2 votes Rabbit and Tigger win (with 70% votes each, and Pooh still keeps his 60% votes).

This paradoxical result may be generalized as follows.

Theorem 4.2. *If the set of candidates* $\mathbf{C}$ *contains at least* $2n+1$ *candidates* ($c_1, \ldots, c_{2n+1} \in \mathbf{C}$)*, then for each* $n \geq 1$ *there exists such a profile that in voting with weak method of* n *votes the winners are* $c_1, \ldots, c_n$*, and in voting with weak method of* $n+1$ *votes the winners are* $c_{n+1}, \ldots, c_{2n+1}$*. Moreover, in both cases each winner is supported by more than 50% voters.*

Proof. Consider the following profile (in Table 11 rankings of first $n+1$ positions are presented). Assume that there are v voters.

	k_1	k_2	$\ldots$	k_n	k_{n+1}	k_{n+2}	$\ldots$	k_{2n+1}
1.	c_1	c_2	$\ldots$	c_n	c_{n+1}	c_{n+2}	$\ldots$	c_{2n+1}
2.	c_2	c_3	$\ldots$	c_1	c_{n+2}	c_{n+3}	$\ldots$	c_{n+1}
3.	c_3	c_4	$\ldots$	c_2	c_{n+3}	c_{n+4}	$\ldots$	c_{n+2}
$\ldots$		$\ldots$				$\ldots$		
$n-1.$	c_{n-1}	c_n	$\ldots$	c_{n-2}	c_{2n-1}	c_{2n}	$\ldots$	c_{2n-2}
$n.$	c_n	c_1	$\ldots$	c_{n-1}	c_{2n}	c_{2n+1}	$\ldots$	c_{2n-1}
$n+1.$		$\star$			c_{2n+1}	c_{n+1}	$\ldots$	c_{2n}

TABLE 11.

The numbers of voters are:

$$k_1 = k_2 = \ldots = k_n = v\left(\frac{1}{2n} + \varepsilon\right),$$

$$k_{n+1} = k_{n+2} = \ldots = k_{2n+1} = v\left(\frac{1}{2(n+1)} - \varepsilon + \frac{\varepsilon}{n+1}\right)$$

In the $(n+1)$th positions in the lists of voters $k_1, \ldots, k_n$ (denoted by $\star$) candidates $c_{n+1}, c_{n+2}, \ldots c_{2n+1}$ are posed, each of them $v\frac{1+2n\varepsilon}{2(n+1)}$ times.

In the position $n+2$ we put each of candidates $c_1, c_2, \ldots c_n$ on $v\left(\frac{1}{2n} - \varepsilon\right)$ lists and each of candidates $c_{n+1}, c_{n+2}, \ldots c_{2n+1}$ on $v\frac{2n\varepsilon+1}{2(n+1)}$ lists, of course each candidate is placed on the list that this candidate was not ranked before. The same procedure is applied to the position $n+3$ and so on, up to the position $2n+1$.

If the number of candidates is greater than $2n+1$, then we put candidate c_{2n+2} in the position $2n+2$ in each list, and so on.

Now we need to show the following properties:

- the construction of the profile is correct
- $v\left(\frac{1}{2n} + \varepsilon\right) > v\left(\frac{1}{2(n+1)} - \varepsilon + \frac{\varepsilon}{n+1}\right)$
- $nv\left(\frac{1}{2n} + \varepsilon\right) < nv\left(\frac{1}{2(n+1)} - \varepsilon + \frac{\varepsilon}{n+1}\right) + \frac{v}{n+1}$
- $n\left(\frac{1}{2n} + \varepsilon\right) > \frac{1}{2}$
- $n\left(\frac{1}{2(n+1)} - \varepsilon + \frac{\varepsilon}{n+1}\right) + \frac{1}{n+1} > \frac{1}{2}$

This can be done be simple calculations. Then, for ε and n satisfying those conditions (it turns out that it is enough to assume that $\varepsilon < \frac{1}{2(2n^2+n)}$) and chosen in such a way that a suitable numbers are integers, we get the desired result. □

References

[1] K.J. Arrow, *A difficulty in the concept of social welfare*, Journal of Political Economy **58** (1950), 328-346.

[2] K.J. Arrow, *Social choice and individual values*, John Wiley & Sons, New York 1951.

[3] N. Bourbaki, *Éléments de mathématique. Topologie générale. Chapitres 1 à 4*, Hermann, Paris 1971.

[4] M.L. Balinski, H.P. Young, *Fair Representation: Meeting the Ideal of One Man, One Vote*, Yale University Press, New Haven and London 1982.

[5] K. Ciesielski, *An elementary proof of Saari's Theorem*, in preparation.

[6] K. Ciesielski, *Paradoksy ordynacji, czyli matematyka wyborcza*, Wiedza i Życie 1997, No. 8, 18–23.

[7] A.P. Kirman, D. Sondermann, *Arrow's Theorem, many agents, and invisible dictators*, Journal of Economic Theory **5** (1972), 267–277.

[8] J. Lützen, *How mathematical impossibility changed welfare economics: a history of Arrow's impossibility theorem*, Historia Math. **46** (2019), 56–87.

[9] J. Przystawa, *Konieczność nieuświadomiona*, Wektory, Kąty Wrocławskie 2006.

[10] E.A. Robinson, D.H.Ullman, *A Mathematical Look at Politics*, CRC Press, New York 2011.

[11] D.G. Saari, *Chaotic elections! A Mathematical Looks at Voting*, AMS, Providence, RI 2001.

[12] D.G. Saari, *Millions of election rankings from a single profile*, Soc. Choice & Welfare **9** (1992), 277–306.
[13] J. Sanocki, *Wojownicy*, Nyskie Towarzystwo Społeczno-Kulturalne, Nysa 2005.
[14] A.D. Taylor, A.M. Pacelli, *Mathematics and Politics. Strategy, Voting, Power and Proof*, Springer, New York 2008.

Mathematics Institute, Jagiellonian University, Poland
krzysztof.ciesielski@im.uj.edu.pl

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON HAMILTONIANS RELATED TO THE SECOND PAINLEVÉ EQUATION

ANTON DZHAMAY, GALINA FILIPUK,
ADAM LIGĘZA & ALEXANDER STOKES

Abstract. In this paper we study different Hamiltonian forms of the second Painlevé equation. We find relations between corresponding Hamiltonian systems, and use these to generate new Hamiltonians, including new Hamiltonian forms of the Painlevé XXXIV equation.

Keywords: Painlevé equations, Hamiltonian systems

MSC2010: 34M55

Received 20 April 2021; accepted 20 June 2021

1. Introduction

The second Painlevé equation is a second order nonlinear differential equation, solutions of which have no movable critical points. It is given by

$$x'' = 2x^3 + tx + \alpha, \tag{1.1}$$

where t is an independent complex variable and α is a complex parameter. Equation (1.1) has many remarkable properties [2]. It possesses rational solutions when α is an integer and solutions expressed in terms of the Airy functions when $\alpha+1/2$ is an integer. Equation (1.1) admits Bäcklund transformation symmetries which relate solutions for different parameter values and form an extended affine Weyl group of type $A_1^{(1)}$ [8]. Among these are transformations which relate solutions for parameters α and $\alpha \pm 1$, which correspond to translation elements of the symmetry group. There is also a so-called folding transformation, which is an algebraic change of variables which allows a solution for $\alpha = -1/2$ to be constructed from one with $\alpha = 0$ [11].

Equation (1.1) can be written as a non-autonomous Hamiltonian system, with the following polynomial Hamiltonian as provided by Okamoto [8]:

$$H = H(q,p,t;\alpha) = \frac{1}{2}p(p-2q^2-t) - (\alpha+1/2)q. \tag{1.2}$$

The corresponding Hamiltonian system is given by

$$\begin{aligned} \frac{dq}{dt} &= \frac{\partial H}{\partial p} = p - q^2 - \frac{t}{2}, \\ \frac{dp}{dt} &= -\frac{\partial H}{\partial q} = 2pq + \alpha + \frac{1}{2}. \end{aligned} \tag{1.3}$$

Eliminating the variable p one obtains the second Painlevé equation (1.1) for q, with parameter α. Eliminating the variable q, one obtains

$$p'' = \frac{p'^2}{2p} + 2p^2 - tp - \frac{(2\alpha+1)^2}{8p}, \tag{1.4}$$

which is known as the Painlevé XXXIV equation, due to its appearance as equation number XXXIV in the list of Painlevé and Gambier [6].

Another natural Hamiltonian [5] associated with the second Painlevé equation is given by

$$K = K(x,y,t;\alpha) = \frac{1}{2}y^2 - \frac{1}{2}x^4 - \frac{1}{2}tx^2 - \alpha x. \tag{1.5}$$

The corresponding Hamiltonian system is

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial K}{\partial y} = y, \\ \frac{dy}{dt} &= -\frac{\partial K}{\partial x} = 2x^3 + tx + \alpha. \end{aligned} \tag{1.6}$$

Clearly, eliminating the variable y, we obtain the second Painlevé equation with parameter α.

2. Main results

In what follows, we obtain and study two birational transformations which identify systems (1.3) and (1.6). In both cases the transformations are canonical, i.e. they preserve the 2-forms associated with the Hamiltonian systems:

$$dq \wedge dp + dH \wedge dt = dx \wedge dy + dK \wedge dt.$$

Since we are dealing with t-dependent canonical transformations, the Hamiltonians will not necessarily coincide and there may be a correction term. Remarkably, for our transformations these corrections correspond only to changes in parameters in the Hamiltonians (1.2) and (1.5). We consider the systems associated with these transformed Hamiltonians, and find relations between them. We repeat this procedure several times and find that at each stage the resulting Hamiltonian systems give either the second Painlevé equation or the Painlevé XXXIV equation. Among those we obtain are new Hamiltonian forms of the Painlevé XXXIV equation.

2.1. **The first birational transformation.**

Theorem 2.1. *Systems (1.3) and (1.6) are related by a birational transformation*

$$q = x, \quad p = x^2 + y + \frac{t}{2}. \tag{2.1}$$

The inverse transformation is given by

$$x = q, \quad y = p - q^2 - \frac{t}{2}. \tag{2.2}$$

The transformation is canonical for the Hamiltonians H (1.2) and K (1.5).

Proof. The result can be verified by direct calculation, and was obtained through the geometric approach [10, 9, 7] (see [3, 4] for illustrations of the method in similar examples). This involves the construction of spaces of initial conditions for systems (1.3) and (1.6) using the blowup technique from algebraic geometry, with the change of variables coming from an appropriate identification of these spaces.

The space of initial conditions for system (1.3) can be constructed from $\mathbb{CP}^1 \times \mathbb{CP}^1$ through eight blowups as in [7, Section 8.2.23] and is minimal in the sense of [1]. The same procedure for system (1.6) leads to a space of initial conditions which is not minimal - it is obtained through ten blowups, with two blowdowns required to obtain the identification with that of system (1.3) and the change of variables in the theorem.

We now provide the configuration of points to be blown up to construct the space of initial conditions in each case. For details of the geometric theory and calculation methods we refer the reader to [7]. For system (1.3), the phase space is initially the trivial bundle over the independent variable space $\mathbb{C}$ (with coordinate t) with fibre $\mathbb{C}^2$ (with coordinates (q, p)). We compactify the fibres from $\mathbb{C}^2$ to $\mathbb{CP}^1 \times \mathbb{CP}^1$ by introducing coordinates $Q = 1/q, P = 1/p$, so $\mathbb{CP}^1 \times \mathbb{CP}^1$ is covered by four charts $(q, p), (Q, p), (q, P), (Q, P)$. Each time we blow up a point we introduce coordinate charts according to the following convention: after blowing up a point p_i given in some coordinate chart (x, y) by $(x, y) = (a, b)$, we introduce charts $(u_i, v_i), (U_i, V_i)$ according to $x = a + u_i, y = b + u_i v_i$ and $x = a + U_i V_i, y = b + V_i$, so the exceptional divisor replacing p_i is given by $u_i = 0, V_i = 0$. The space of initial conditions for system (1.3) is obtained by blowing up the following sequence of points:

$$\begin{aligned}
p_1 : (Q, p) = (0,0) &\longleftarrow p_2 : (u_1, v_1) = \left(0, -\alpha - \frac{1}{2}\right) \\
p_3 : (Q, P) = (0,0) &\longleftarrow p_4 : (u_3, v_3) = (0,0) \longleftarrow p_5 : (u_4, v_4) = \left(0, \frac{1}{2}\right) \\
&\longleftarrow p_6 : (u_5, v_5) = (0,0) \longleftarrow p_7 : (u_6, v_6) = \left(0, -\frac{t}{4}\right) \\
&\longleftarrow p_8 : (u_7, v_7) = \left(0, \frac{1-2\alpha}{8}\right).
\end{aligned}$$

For system (1.6), we similarly compactify the fibres of the phase space to $\mathbb{CP}^1 \times \mathbb{CP}^1$ by introducing $X = 1/x, Y = 1/y$, and obtain the space of initial conditions by blowing up the following sequence of points, in which we have used the same convention for blowup charts but denoted points by z_i instead of p_i :

$$\begin{aligned}
z_1 : (X, Y) = (0,0) \longleftarrow z_2 &: (u_1, v_1) = (0,0) \\
z_2 \longleftarrow z_3 &: (u_2, v_2) = (0,-1) \longleftarrow z_4 : (u_3, v_3) = (0,0) \\
&\longleftarrow z_5 : (u_4, v_4) = \left(0, \frac{t}{2}\right) \longleftarrow z_6 : (u_5, v_5) = \left(0, \frac{1}{2} + \alpha\right) \\
z_2 \longleftarrow z_7 &: (u_2, v_2) = (0,1) \longleftarrow z_8 : (u_7, v_7) = (0,0) \\
&\longleftarrow z_9 : (u_8, v_8) = \left(0, -\frac{t}{2}\right) \longleftarrow z_{10} : (u_9, v_9) = \left(0, \frac{1}{2} - \alpha\right).
\end{aligned}$$

The transformation in the theorem is obtained by requiring that it induces a map between the rational surfaces obtained through the blowups of the points above. We omit the details here for brevity but refer the reader interested in this method to [3, 4].

□

In light of this transformation between the Hamiltonian systems (1.3) and (1.6), we now examine how the Hamiltonian functions themselves are related. Substituting the change of variables (2.1) into the Hamiltonian H we obtain a function of x and y, which differs from K by some correction due to the t-dependence of the transformation (2.1). Explicitly, we find

$$H\left(x, x^2 + y + \frac{t}{2}, t; \alpha\right) = K(x, y, t; \alpha) - \frac{x}{2} - \frac{t^2}{8}.$$

We may use this to define a new Hamiltonian function which we denote by $K^{(1)}$. Explicitly, written using the variables x_1, y_1 to emphasise that they do not solve the same system as x, y, this is given by

$$K^{(1)} = K^{(1)}(x_1, y_1, t; \alpha) = \frac{1}{2}(y_1^2 - x_1^4 - tx_1^2 - 2\alpha x_1 - x_1) - \frac{t^2}{8}. \tag{2.3}$$

The corresponding Hamiltonian system is

$$\begin{aligned} x_1' &= \frac{\partial K^{(1)}}{\partial y_1} = y_1, \\ y_1' &= -\frac{\partial K^{(1)}}{\partial x_1} = 2x_1^3 + tx_1 + \alpha + \frac{1}{2}, \end{aligned} \tag{2.4}$$

which is precisely the same as system (1.6) with α replaced with $\alpha + 1/2$, so eliminating the variable y_1 we get the second Painlevé equation for the variable x_1 with parameter $\alpha + 1/2$. The reason for this is that the correction term is (modulo functions of t) precisely the same as that which would arise from shifting the parameter in K:

$$K^{(1)}(x, y, t; \alpha) = K(x, y, t; \alpha + 1/2) - \frac{t^2}{8}$$

Similarly, substituting the inverse change of variables (2.2) into the Hamiltonian K we obtain a function of q and p which also differs from H by some correction:

$$K\left(q, p - q^2 - \frac{t}{2}, t; \alpha\right) = H(q, p, t; \alpha) + \frac{q}{2} + \frac{t^2}{8}.$$

This is the transformed version of the Hamiltonian function H, which we denote by $H^{(1)}$ and again write with q, p relabelled as q_1, p_1:

$$H^{(1)} = H^{(1)}(q_1, p_1, t; \alpha) = \frac{p_1^2}{2} - p_1 q_1^2 - \alpha q_1 - \frac{1}{2} t p_1 + \frac{t^2}{8}. \tag{2.5}$$

The corresponding Hamiltonian system is

$$\begin{aligned} q_1' &= \frac{\partial H^{(1)}}{\partial p_1} = p_1 - q_1^2 - \frac{t}{2}, \\ p_1' &= -\frac{\partial H^{(1)}}{\partial q_1} = 2p_1 q_1 + \alpha. \end{aligned} \tag{2.6}$$

Similarly, we see that this system is the same as (1.3) with a parameter shift, as

$$H^{(1)}(q, p, t; \alpha) = H(q, p, t; \alpha - 1/2) + \frac{t^2}{8}.$$

Eliminating the variable p_1 we get the second Painlevé equation for the variable q_1 with parameter $\alpha - 1/2$. It turns out that we can obtain a birational transformation between systems (2.6) and (2.4) by similar methods as Theorem 2.1, which again can be verified by direct calculation.

Theorem 2.2. *Systems (2.6) and (2.4) are related by a birational transformation*

$$q_1 = -x_1 - \frac{2\alpha}{2x_1^2 - 2y_1 + t}, \quad p_1 = x_1^2 - y_1 + \frac{t}{2}. \tag{2.7}$$

The inverse transformation is given by

$$x_1 = -q_1 - \frac{\alpha}{p_1}, \quad y_1 = q_1^2 - p_1 + \frac{2\alpha q_1}{p_1} + \frac{\alpha^2}{p_1^2} + \frac{t}{2}. \tag{2.8}$$

The transformation is canonical for the Hamiltonians $H^{(1)}$ (2.5) and $K^{(1)}$ (2.3).

Next we repeat the procedure and attempt to define two more new Hamiltonian functions through the substitutions (2.7) and (2.8) in $H^{(1)}$ and $K^{(1)}$, neglecting correction terms. Substitution (2.7) gives us back the Hamiltonian K. We find

$$H^{(1)}\left(-x_1 - \frac{2\alpha}{2x_1^2 - 2y_1 + t}, x_1^2 - y_1 + \frac{t}{2}, t; \alpha\right) = K(x_1, y_1, t; \alpha).$$

Substitution (2.8), on the other hand, gives a new rational function which we denote by $H^{(2)}$. Explicitly,

$$H^{(2)} = H^{(2)}(q_2, p_2, t; \alpha) = \frac{p_2^2}{2} - \frac{t}{2}p_2 + \frac{\alpha}{2p_2} - p_2 q_2^2 - (\alpha - 1/2)q_2. \tag{2.9}$$

The corresponding Hamiltonian system is

$$\begin{aligned} q_2' &= \frac{\partial H^{(2)}}{\partial p_2} = p_2 - q_2^2 - \frac{\alpha}{2p_2^2} - \frac{t}{2}, \\ p_2' &= -\frac{\partial H^{(2)}}{\partial q_2} = 2p_2 q_2 + \alpha - \frac{1}{2}. \end{aligned} \tag{2.10}$$

Unlike the previous case, this system can not be obtained from the original Okamoto Hamiltonian by shifting parameters, and the correction is given according to

$$H^{(2)}(q, p, t; \alpha) = H(q, p, t; \alpha) + q + \frac{\alpha}{2p}.$$

However, this system can still be related to a Painlevé equation: eliminating the variable q_2 we obtain the Painlevé XXXIV equation for p_2 with parameter α as in (1.4). The equation for q_2 is of second order and second degree. It is cumbersome so we omit it.

Theorem 2.3. *Systems (2.10) and (1.6) are related by the birational transformation*

$$q_2 = x + \frac{1}{t + 2x^2 + 2y}, \quad p_2 = \frac{t}{2} + x^2 + y. \tag{2.11}$$

The inverse transformation is given by

$$x = q_2 - \frac{1}{2p_2}, \quad y = -q_2^2 + p_2 + \frac{q_2}{p_2} - \frac{1}{4p_2^2} - \frac{t}{2}. \tag{2.12}$$

The transformation is canonical for the Hamiltonians $H^{(2)}$ (2.9) and K (1.5).

Using the fact that there is no correction between $K^{(2)}$ and K, the Theorem 2.3 leads immediately to the following.

Theorem 2.4. *Systems (2.10) and (1.3) are related by a birational transformation*

$$q_2 = q + \frac{1}{2p}, \quad p_2 = p. \tag{2.13}$$

The inverse transformation is given by

$$q = q_2 - \frac{1}{2p_2}, \quad p = p_2. \tag{2.14}$$

The transformation is canonical for the Hamiltonians $H^{(2)}$ (2.9) and H (1.2).

Again, we repeat the procedure and define two more new Hamiltonian functions from $H^{(2)}$ and $K^{(2)}$ by discarding correction terms. Substitution (2.13) recovers the Hamiltonian $K^{(1)}$:

$$H^{(2)}\left(x + \frac{1}{t+2x^2+2y}, \frac{t}{2} + x^2 + y, t; \alpha\right) = K^{(1)}(x, y, t; \alpha).$$

Again the other substitution (2.31) gives the function $H^{(3)}$:

$$H^{(3)} = H^{(3)}(q_3, p_3, t; \alpha) = \frac{p_3^2}{2} - \frac{t}{2}p_3 + \frac{\alpha - 1/2}{2p_3} - p_3 q_3^2 + (1-\alpha)q_3 + \frac{t^2}{8}, \tag{2.15}$$

This is related (modulo functions of t) to $H^{(2)}$ by a parameter shift:

$$H^{(3)}(q, p, t; \alpha) = H^{(2)}(q, p, t; \alpha - 1/2) + \frac{t^2}{2}.$$

The corresponding Hamiltonian system is

$$\begin{aligned} q_3' &= \frac{\partial H^{(3)}}{\partial p_3} = p_3 - q_3^2 - \frac{\alpha - 1/2}{2p_3^2} - \frac{t}{2}, \\ p_3' &= -\frac{\partial H^{(3)}}{\partial q_3} = 2p_3 q_3 + \alpha - 1. \end{aligned} \tag{2.16}$$

It follows that eliminating the variable q_3 will lead to the Painlevé XXXIV equation for p_3 with parameter $\alpha - 1/2$.

Theorem 2.5. *Systems (2.16) and (2.4) are related by the birational transformation*

$$q_3 = -x_1 - \frac{1-2\alpha}{t + 2x_1^2 - 2y_1 + 1}, \quad p_3 = \frac{t}{2} + x_1^2 - y_1. \tag{2.17}$$

The transformation is canonical for the Hamiltonians $H^{(3)}$ (2.15) and $K^{(1)}$ (2.3).

Again, applying this transformation to the Hamiltonian function $K^{(1)}$ neglecting correction terms leads again to K. Substitution into $H^{(3)}$, however, yields a new Hamiltonian that cannot be obtained from any of the preceding ones by shifting parameters, similarly to $H^{(2)}$:

$$H^{(4)} = H^{(4)}(q_4, p_4, t; \alpha) = \frac{p_4^2}{2} - \frac{t}{2}p_4 + \frac{2\alpha - 1}{2p_4} - p_4 q_4^2 - (\alpha - 3/2)q_4. \tag{2.18}$$

This is related to previously obtained Hamiltonians according to

$$H^{(4)}(q, p, t; \alpha) = H^{(2)}(q, p, t; \alpha) + q + \frac{\alpha - 1}{2p} = H(q, p, t; \alpha) + 2q + \frac{2\alpha - 1}{2p}.$$

The corresponding Hamiltonian system is

$$\begin{aligned} q_4' &= \frac{\partial H^{(4)}}{\partial p_4} = p_4 - q_4^2 - \frac{\alpha - 1/2}{p_4^2} - \frac{t}{2}, \\ p_4' &= -\frac{\partial H^{(4)}}{\partial q_4} = 2p_3 q_3 + \alpha - \frac{3}{2}. \end{aligned} \tag{2.19}$$

The variable p_4 satisfies the Painlevé XXXIV equation with parameter α.

2.2. **The second birational transformation.** Throughout this subsection we assume $\alpha = 0$, and relabel the Hamiltonians H and K with this parameter specialisation by H_0 and K_0. We present some observations along the same lines as in the previous subsection, obtaining relations between Hamiltonian systems, studying the functions obtained by discarding the corrections between Hamiltonians and repeating the process.

Theorem 2.6. *Systems (1.3) and (1.6) with $\alpha = 0$ are related by a birational transformation*

$$q = -x, \quad p = x^2 - y + \frac{t}{2}. \tag{2.20}$$

The inverse transformation is given by

$$x = -q, \quad y = q^2 - p + \frac{t}{2}. \tag{2.21}$$

The transformation is canonical for the Hamiltonians H_0 and K_0.

Remark 2.7. This transformation is related to that of Theorem 2.1. There is a certain Bäcklund transformation for system (1.3) corresponding to an automorphism of the Dynkin diagram $A_1^{(1)}$ which relates solutions for parameters α and $-\alpha$. Specialising to the fixed point $\alpha = 0$ of the action of this symmetry on parameter space allows the two transformations to exist simultaneously, the relation between them being precisely this Bäcklund transformation.

Substituting the change of variables (2.20) into the Hamiltonian H_0 we obtain

$$H_0\left(-x, x^2 - y + \frac{t}{2}, t\right) = K_0(x, y, t) + \frac{x}{2} - \frac{t^2}{8},$$

which we use to define a new Hamiltonian function

$$K_0^{(1)} = K_0^{(1)}(x_1, y_1, t) = \frac{1}{2}(y_1^2 - x_1^4 - tx_1^2 + x_1) - \frac{t^2}{8}. \tag{2.22}$$

The corresponding Hamiltonian system is

$$\begin{aligned} x_1' &= \frac{\partial K_0^{(1)}}{\partial y_1} = y_1, \\ y_1' &= -\frac{\partial K_0^{(1)}}{\partial x_1} = 2x_1^3 + tx_1 - \frac{1}{2}. \end{aligned} \tag{2.23}$$

Eliminating the variable y_1 leads to the second Painlevé equation with parameter $\alpha_1 = -1/2$ for the variable x_1.

Substituting the inverse change of variables (2.21) into the Hamiltonian K_0 we obtain

$$K_0\left(-q, q^2 - p + \frac{t}{2}, t\right) = H_0(q, p, t) + \frac{q}{2} + \frac{t^2}{8}.$$

Again we use this to define another new Hamiltonian function

$$H_0^{(1)} = H_0^{(1)}(q_1, p_1, t) = \frac{p_1^2}{2} - p_1 q_1^2 - \frac{1}{2}tp_1 + \frac{t^2}{8}. \tag{2.24}$$

The corresponding Hamiltonian system is

$$\begin{aligned} q_1' &= \frac{\partial H_0^{(1)}}{\partial p_1} = p_1 - q_1^2 - \frac{t}{2}, \\ p_1' &= -\frac{\partial H_0^{(1)}}{\partial q_1} = 2p_1 q_1. \end{aligned} \tag{2.25}$$

Eliminating the variable p_1 leads to the second Painlevé equation with parameter $\alpha_1 = -1/2$ for the variable q_1.

Theorem 2.8. *Systems (2.25) and (2.23) are related by a birational transformation*

$$q_1 = x_1, \quad p_1 = x_1^2 + y_1 + \frac{t}{2}. \tag{2.26}$$

The inverse transformation is given by

$$x_1 = q_1, \quad y_1 = p = q^2 - \frac{t}{2}. \tag{2.27}$$

The transformation is canonical for the Hamiltonians $H_0^{(1)}$ (2.24) and $K_0^{(1)}$ (2.22).

Next we repeat the procedure and attempt to define two more new Hamiltonian functions using (2.26) and (2.27) and $H_0^{(1)}$ and $K_0^{(1)}$. Substitution (2.26) gives us back the Hamiltonian K_0. We find

$$H_0^{(1)}\left(x_1, x_1^2 + y_1 + \frac{t}{2}, t\right) = K_0(x_1, y_1, t).$$

Substitution (2.27) gives a new function which we denote by $H_0^{(2)}$. Explicitly,

$$H_0^{(2)} = H_0^{(2)}(q_2, p_2, t) = \frac{p_2^2}{2} - \frac{1}{2}tp_2 - p_2 q_2^2 + \frac{q_2}{2} = H_0(q_2, p_2, t) + q_2. \tag{2.28}$$

The corresponding Hamiltonian system is

$$\begin{aligned} q_2' &= \frac{\partial H_0^{(2)}}{\partial p_2} = p_2 - q_2^2 - \frac{t}{2}, \\ p_2' &= -\frac{\partial H_0^{(2)}}{\partial q_2} = 2p_2 q_2 - \frac{1}{2}. \end{aligned} \tag{2.29}$$

Eliminating the variable p_2 we get the second Painlevé equation with parameter $\alpha = -1$.

Theorem 2.9. *Systems (2.29) and (1.6) with $\alpha = 0$ are related by a birational transformation*

$$q_2 = -x - \frac{1}{2y - 2x^2 - t}, \quad p_2 = x^2 - y + \frac{t}{2}. \tag{2.30}$$

The inverse transformation is given by

$$x = -q + \frac{1}{2p_2}, \quad y = q_2^2 - p_2 - \frac{q_2}{p_2} + \frac{1}{4p_2^2} + \frac{t}{2}. \tag{2.31}$$

The transformation is canonical.

3. Conclusion

The primary result of this paper is the relation between the two different Hamiltonian systems associated with the second Painlevé equation, namely systems (1.3) and (1.6). Following this, our aim was to report observations which followed this, beginning with the fact that the correction between these two Hamiltonians corresponds to parameter shifts in the original systems, so applying the transformation to the Hamiltonians themselves still gave systems related to Painlevé equations. These could then be identified and the process repeated, which continued to produce Hamiltonian forms of Painlevé equations, some of which could not be obtained from the original ones by simply shifting parameters. The most notable examples are the systems (2.10) and (2.19), which seem to be new Hamiltonian forms of the Painlevé XXXIV equation.

Our work is very much in the spirit of Okamoto's original studies of symmetries of the second Painlevé equation [8], in which Bäcklund transformation symmetries were studied via an auxiliary Hamiltonian system. We also remark that the correction terms between Hamiltonians observed in this paper, in particular between

$H^{(1)}$ and H, are similar to those arising from the folding transformation [11], which is an algebraic (not birational) canonical transformation that maps the Okamoto Hamiltonian system for (q, p) to the same Hamiltonian system for new variables. This is to be expected, since the folding transformation involves the same shift by a half in parameters. However, the Hamiltonians here are obtained through a different process; firstly, our differential systems for (q, p) and (q_1, p_1) are different. The folding transformation, on the other hand, identified the differential systems themselves and relied on the parameter specialisation $\alpha = 0$ in order to work.

It would be very interesting to continue iterating the procedure to determine whether the resulting Hamiltonians would be related to Painlevé equations and whether they are new. It is natural to ask whether our observations are related to the affine Weyl group symmetries or folding transformation for the second Painlevé equation. An interpretation of our results in terms of the geometric theory of Painlevé equations is also natural to pursue.

Finally, we would also like to remark that there is one more intriguing relation for the Hamiltonian K: it is invariant under the following change

$$K\left(-x - \frac{2\alpha}{2x^2 - 2y + t}, -y - \frac{4\alpha x}{2x^2 - 2y + t} - \frac{4\alpha^2}{(2x^2 - 2y + t)^2}, t\right) = K(x, y, t).$$

This deserves further investigation.

References

[1] A. S. Carstea, T. Takenawa, *A note on minimization of rational surfaces obtained from birational dynamical systems*, J. Nonlinear Math. Phys. **20** (2013), suppl. 1, 17–33.

[2] Digital Library of Mathematical Functions, https://dlmf.nist.gov/32.

[3] A. Dzhamay, G. Filipuk, A. Stokes, *Differential equations for the recurrence coefficients of semi-classical orthogonal polynomials and their relation to the Painlevé equations via the geometric approach*, submitted.

[4] A. Dzhamay, G. Filipuk, A. Ligęza, A. Stokes, *Different Hamiltonians for the Painlevé $P_{\rm IV}$ equation and their identification using a geometric approach*, in preparation.

[5] G. Filipuk, H. Żołądek, *Painlevé equations, elliptic integrals and elementary functions*, J. Differential Equations **258** (2015), no. 4, 1303–1355.

[6] E. L. Ince, *Ordinary Differential Equations*, Dover Publications, New York, 1944.

[7] K. Kajiwara, M. Noumi, Y. Yamada, *Geometric aspects of Painlevé equations*, J. Phys. A **50** (2017), no. 7, 073001, 164 pp.

[8] K. Okamoto, *Studies on the Painlevé equations. III. Second and fourth Painlevé equations, $P_{\rm II}$ and $P_{\rm IV}$*, Math. Ann. **275** (1986), no. 2, 221–255.

[9] K. Okamoto, *Sur les feuilletages associés aux équations du second ordre á points critiques fixes de P. Painlevé* (French) [On foliations associated with second-order Painlevé equations with fixed critical points], Japan. J. Math. (N.S.) **5** (1979), no. 1, 1–79.

[10] H. Sakai, *Rational surfaces associated with affine root systems and geometry of the Painlevé equations*, Comm. Math. Phys. **220** (2001), no. 1, 165–229.

[11] T. Tsuda, K. Okamoto, H. Sakai, *Folding transformations of the Painlevé equation*, Math. Ann. **331** (2005), 713–738.

The University of Northern Colorado, USA
adzham@unco.edu

University of Warsaw, Poland
filipuk@mimuw.edu.pl; akligeza@gmail.com

University College London, United Kingdom
alexander.stokes.14@ucl.ac.uk

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

MORAN TYPE THEOREMS AND IRREDUCIBLE FRACTAL STRUCTURES

MANUEL FERNÁNDEZ-MARTÍNEZ

Abstract. In this paper, we revisit a classical problem in Fractal Geometry which consists of the calculation of the fractal dimension of IFS-attractors. In this way, the open set condition hypothesis has been considered as the right separation property for self-similar sets. In fact, the classical Moran's Theorem guarantees that Hausdorff dimension coincides with the similarity dimension of every attractor whose iterated function system lies under that separation condition. However, it depends on an external open set. Moreover, it has been proved that the open set condition stands as a sufficient (though not necessary) hypothesis leading to that identity.

On the other hand, we provide a separation condition for attractors that becomes necessary to reach the equality among the similarity dimension of an attractor and some fractal dimension models that are defined with respect to a natural fractal structure for self-similar sets. It is worth mentioning that our hypothesis, the level separation property, better describes the self-similar structure of the attractor and becomes weaker than the SOSC in the general setting of complete metric spaces.

Keywords: fractal structure, IFS-attractor, Hausdorff dimension, similarity dimension, open set condition, separation conditions for attractors, Moran's Theorem.

MSC2010: 28A80; 28A78, 54E40, 37F20, 37B10

Received 23 April 2021; accepted 29 July 2021

1. Introduction

A classical problem in Fractal Geometry deals with determining under what conditions on the pieces of a strict self-similar set $\mathcal{K}$, the equality between the

 https://doi.org/10.2478/9788366675360-007.

similarity dimension and the Hausdorff dimension of $\mathcal{K}$ holds. In this way, a well-known result proved by P. A. P. Moran in the forties (c.f. [1, Theorem III]) states that under the open set condition (OSC in the sequel), a property required to the pieces of $\mathcal{K}$ that guarantees that their overlaps are thin enough, such an identity holds. Afterwards, Lalley introduced the strong open set condition (SOSC) by further requiring that the (feasible) open set provided by the OSC intersects the attractor $\mathcal{K}$. The next chain of implications and equivalences stands in the Euclidean setting and is best possible (c.f. [2]):

$$\text{SOSC} \Leftrightarrow \text{OSC} \Leftrightarrow \mathcal{H}^{\alpha}(\mathcal{K}) > 0 \Rightarrow \dim_{\mathrm{H}}(\mathcal{K}) = \alpha, \tag{1.1}$$

where $\mathcal{H}^{\alpha}$ denotes $\alpha-$dimensional Hausdorff measure, $\dim_{\mathrm{H}}$ is Hausdorff dimension, and α is the similarity dimension of the self-similar set. A counterexample due to Mattila (personal communication to A. Schief) guarantees that the last implication in Eq. (1.1) cannot be inverted, in general. Accordingly, the OSC becomes sufficient, but not neccesary, to reach the desired identity.

A further extension of the problem above takes place in the general setting of attractors on complete metric spaces. Schief also proved that (c.f. [3])

$$\mathcal{H}^{\alpha}(\mathcal{K}) > 0 \Rightarrow \text{SOSC} \Rightarrow \dim_{\mathrm{H}}(\mathcal{K}) = \alpha, \tag{1.2}$$

i.e., the SOSC being necessary for a positive Hausdorff measure of $\mathcal{K}$, and only sufficient for $\dim_{\mathrm{H}}(\mathcal{K}) = \alpha$. Once again, Mattila's Counterexample implies that Eq. (1.2) is best possible. As such, both Eqs. (1.1) and (1.2) imply that the SOSC is a sufficient condition on the pre-fractals of $\mathcal{K}$ leading to $\alpha = \dim_{\mathrm{H}}(\mathcal{K})$.

In this paper, we apply the concept of fractal structure (c.f. Subsection 2.5) to define and characterize a novel separation property for self-similar sets in both settings. Interestingly, that separation condition stands weaker than the OSC and becomes necessary to guarantee the equality $\alpha = \dim_{\mathrm{H}}(\mathcal{K})$. Accordingly, we conclude that, in the general setting, it holds that

$$\mathcal{H}^{\alpha}(\mathcal{K}) > 0 \Rightarrow \text{SOSC} \Rightarrow \dim_{\mathrm{H}}(\mathcal{K}) = \alpha \Rightarrow \text{LSP},$$

where LSP denotes the so-called weak separation property for attractors we introduce in Section 4. We also refer the reader to [4] for additional details regarding the proofs of the results stated therein.

2. Preliminaries

This section contains the basics on IFS-attractors (c.f. Subsection 2.2). In this regard, notations from the domain of words will be used (c.f. Subsection 2.1). On the other hand, we also recall some equivalent definitions leading to the open set condition (c.f. Subsection 2.3), usually regarded as the right separation condition for self-similar sets. The definition of Hausdorff dimension appears in Subsection 2.4. The key concept of a fractal structure is introduced in Subsection 2.5, which additionally explains how any self-similar set could be always endowed with a fractal structure from a natural viewpoint. Finally, in Subsection 2.6, we recall the definitions of both fractal dimensions III and IV for a fractal structure, which will play a crucial role to characterize our weaker separation condition for attractors (c.f. Theorem 4.2).

2.1. **Domain of words' notation.** Let $k \in \mathbb{N}$ and $\Sigma = \{1, \ldots, k\}$ be a finite (nonempty) alphabet. Then for each $n \in \mathbb{N}$, the set $\Sigma^n = \{\mathbf{i} = i_1 \ldots i_n : i_j \in \Sigma, j = 1, \ldots, n\}$ contains all the words of (finite) length n from Σ. Also, $\Sigma^\infty = \cup_{n \in \mathbb{N}} \Sigma^n \cup \Sigma^{\mathbb{N}}$ denotes the set consisting of all the finite or infinite words from Σ. The prefix order is defined on Σ^∞ by $x \sqsubseteq y$, if and only if, x is a prefix of y.

2.2. **IFS-attractors.** Let $k \geq 2$. By an iterated function system (IFS, hereafter), we shall understand a finite set of similitudes on a complete metric space (X, ρ), $\mathcal{F} = \{f_1, \ldots, f_k\}$, where each self-map $f_i : X \to X$ satisfies the identity

$$\rho(f_i(x), f_i(y)) = c_i \cdot \rho(x, y) \text{ for all } x, y \in X,$$

where $0 < c_i < 1$ is the similarity ratio of each f_i. In particular, if $X = \mathbb{R}^d$, then $\mathcal{F}$ is said to be a Euclidean IFS (EIFS).

Under the previous assumptions, there exists a unique (nonempty) compact subset $\mathcal{K} \subseteq X$ for which the following Hutchinson's equation holds (c.f. [5]):

$$\mathcal{K} = \cup\{f_i(\mathcal{K}) : i \in \Sigma\}. \tag{2.1}$$

In this way, $\mathcal{K}$ is usually called as the IFS-attractor (attractor, or self-similar set) of $\mathcal{F}$. It is worth noting that each self-similar set $\mathcal{K}$ consists of smaller self-similar copies of itself, say $\mathcal{K}_i = f_i(\mathcal{K})$ for all $i \in \Sigma$, sometimes named pre-fractals of $\mathcal{K}$. We shall also denote $\mathcal{K}_{ij} = f_i(f_j(\mathcal{K}))$, and so on. As such, if $f_\mathbf{i} = f_{i_1} \circ \ldots \circ f_{i_n}, c_\mathbf{i} = c_{i_1} \circ \ldots \circ c_{i_n}$, and $\mathcal{K}_\mathbf{i} = f_\mathbf{i}(\mathcal{K})$, then Eq. (2.1) can be rewritten equivalently as $\mathcal{K} = \cup\{\mathcal{K}_\mathbf{i} : \mathbf{i} \in \Sigma^n\}$. Hence, the so-called address map, π, appears as a continuous function from Σ^∞ onto $\mathcal{K}$. Notice that if the similarity ratios c_i are *small*, then the pre-fractals $\mathcal{K}_i$ are disjoint, π is a homemomorphism, and $\mathcal{K}$ becomes a Cantor set.

2.3. **The open set condition.** In the Euclidean setting, there are, at least, four equivalent descriptions of the open set condition (OSC), a separation property that controlling the overlapping among the pre-fractals of $\mathcal{K}$. They are:

(1) **Moran's open set condition** (introduced by P.A.P. Moran, c.f. [1]). It is said that $\mathcal{F} = \{f_1, \ldots, f_k\}$ (or its attractor $\mathcal{K}$) is under the OSC if there exists a nonempty open subset $\mathcal{V} \subseteq \mathbb{R}^d$ such that $f_i(\mathcal{V}) \cap f_j(\mathcal{V}) = \varnothing$ for all $i \neq j$ with $f_i(\mathcal{V}) \subseteq \mathcal{V}$ for all $i = 1, \ldots, k$. That $\mathcal{V}$ is called as a *feasible open set* of $\mathcal{F}$ (resp., of $\mathcal{K}$).
(2) **The finite clustering property** (provided by Schief, c.f. [2]). First, two pre-fractals of $\mathcal{K}$, $\mathcal{K}_\mathbf{j}$ and $\mathcal{K}_\mathbf{k}$, are said to be incomparable provided that $\mathbf{j} \not\sqsubseteq \mathbf{k}$ and $\mathbf{k} \not\sqsubseteq \mathbf{j}$. The finite clustering property holds whenever there exists

an integer N such that at most N incomparable pieces $\mathcal{K}_{\mathbf{j}}$ of size $\geq \varepsilon$ can intersect the $\varepsilon-$neighborhood of a piece $\mathcal{K}_{\mathbf{i}}$ of diameter equal to ε.

(3) **The neighbor map condition** (contributed by Bandt and Graf, c.f. [6]). Let the next collection of neighbor maps be given:

$$\mathcal{N} = \{h = f_{\mathbf{i}}^{-1} f_{\mathbf{j}} : \mathbf{i}, \mathbf{j} \in \Sigma^*, i_1 \neq j_1\}, \text{ where } \Sigma^* = \cup\{\Sigma^n : n \geq 1\}.$$

Thus, the OSC can be expressed in the following terms: there exists a constant $\lambda > 0$ such that $\|h - \mathrm{id}\| > \lambda$ for all neighbor map $h \in \mathcal{N}$, where $\|g\| = \sup\{g(x) : \|x\| \leq 1\}$ denotes the norm of each affine map g on $\mathbb{R}^d$. I.e., such a condition states that *compared to their size*, two self-similar copies $\mathcal{K}_{\mathbf{i}}$ and $\mathcal{K}_{\mathbf{j}}$ of $\mathcal{K}$ cannot be arbitrarily close to each other.

(4) **Positive $\alpha-$dimensional Hausdorff measure of $\mathcal{K}$** (c.f. [1, 2]): $\mathcal{H}^\alpha(\mathcal{K}) > 0$, where α is the similarity dimension of $\mathcal{K}$, i.e., the (unique) solution of the equation $\sum_{i=1}^k c_i^\alpha = 1$ (c.f. Definition 3.2).

Notice that the classical OSC may be too weak to achieve results concerning the fractal dimension of $\mathcal{K}$ since the feasible open set $\mathcal{V}$ and the attractor $\mathcal{K}$ may be disjoint. In this regard, Lalley introduced the strong open set condition (SOSC) by requiring that, in addition to the OSC, it holds that $\mathcal{V} \cap \mathcal{K} \neq \varnothing$ (c.f. [7]). Interestingly, Schief proved that both the SOSC and the OSC are equivalent in the Euclidean setting (c.f. [2, Theorem 2.2]). Moreover, such a result was further extended to conformal IFSs [8] and self-conformal random fractals [9].

We would like also to mention that Schief explored some conditions leading to the equality between the similarity dimension (c.f. Definition 3.2) and the Hausdorff dimension of IFS$-$attractors in the general setting (complete metric spaces). However, in that case, the OSC no longer yields the equality between such dimensions (c.f. [3]).

2.4. **Hausdorff dimension.** This is the oldest and also the most accurate definition of fractal dimension [10]. Its analytical construction, based on a measure, can be sketched as follows. Let (X, ρ) be a metric space and $A \subseteq X$. Let us denote $\mathrm{diam}\,(A)$ the diameter of A, i.e., $\mathrm{diam}\,(A) = \sup\{\rho(x,y) : x, y \in A\}$, as usual. Further, let $\delta > 0$ and $F \subseteq X$. By a $\delta-$cover of F, we shall understand a countable family of subsets $\{B_i\}_{i \in I}$ such that $F \subseteq \cup_{i \in I} B_i$ with $\mathrm{diam}\,(B_i) \leq \delta$. Let $\mathcal{C}_\delta(F)$ denote the class of all $\delta-$covers of F and define

$$\mathcal{H}_\delta^s(F) = \inf\left\{\sum_{i \in I} \mathrm{diam}\,(B_i)^s : \{B_i\}_{i \in I} \in \mathcal{C}_\delta(F)\right\}.$$

It holds that $\mathcal{H}^s(F) = \lim_{\delta \to 0} \mathcal{H}_\delta^s(F)$ always exists, being known as the ($s-$dimensional) Hausdorff measure of F. As such, the Hausdorff dimension of F is fully described as the (unique) critical point $s \geq 0$ where the Hausdorff measure of F "jumps" from ∞ to 0, i.e.,

$$\dim_{\mathrm{H}}(F) = \sup\{s : \mathcal{H}^s(F) = \infty\} = \inf\{s : \mathcal{H}^s(F) = 0\}.$$

In particular, for $s = \dim_{\mathrm{H}}(F)$, we have $\mathcal{H}^s(F) \in [0, \infty]$.

2.5. **Fractal structures.** Fractal structures were first sketched by Bandt and Retta in [11], and formally introduced and applied afterwards by Arenas and Sánchez-Granero to characterize non-Archimedean quasi-metrization (c.f. [12]).

By a covering of a nonempty set X, we shall understand a family Γ of subsets of X such that $X = \cup\{A : A \in \Gamma\}$. Let Γ_1 and Γ_2 be two coverings of X. Then $\Gamma_2 \prec \Gamma_1$ means that Γ_2 is a refinement of Γ_1, i.e., for all $A \in \Gamma_2$, there exists $B \in \Gamma_1$ such that $A \subseteq B$. In addition, $\Gamma_2 \prec\prec \Gamma_1$ means that $\Gamma_2 \prec \Gamma_1$ and $B = \cup\{A \in \Gamma_2 : A \subseteq B\}$ for all $B \in \Gamma_1$. Hence, a fractal structure on X is defined as a countable family of coverings of X, $\mathbf{\Gamma} = \{\Gamma_n\}_{n\in\mathbb{N}}$, such that $\Gamma_{n+1} \prec\prec \Gamma_n$ for all $n \in \mathbb{N}$. Here, covering Γ_n is named level n of $\mathbf{\Gamma}$. Every fractal structure induces a transitive base of quasi-uniformity (and hence, a topology) given by the transitive family of entourages $U_{\Gamma_n} = \{(x, y) \in X \times X : y \in X \setminus \cup_{A\in\Gamma_n, x\notin A} A\}$. In this work, a set will be allowed to appear twice or more in a level of a fractal structure.

Let $\mathbf{\Gamma}$ be a fractal structure on X and assume that $\mathrm{St}(x, \mathbf{\Gamma}) = \{\mathrm{St}(x, \Gamma_n)\}_{n\in\mathbb{N}}$ is a neighborhood base of x for all $x \in X$, where $\mathrm{St}(x, \Gamma_n) = \cup\{A \in \Gamma_n : x \in A\}$. That $\mathbf{\Gamma}$ is called as a starbase fractal structure. A fractal structure is said to be finite if all its levels are finite coverings. An example of a finite fractal structure is the one which any IFS$-$attractor can be always endowed with naturally. Such a fractal structure plays a key role in this paper.

Definition 2.1 (c.f. [13], Definition 4.4)**.** Let $\mathcal{F}$ be an IFS whose attractor is $\mathcal{K}$. The natural fractal structure on $\mathcal{K}$ (as a self-similar set) is the countable family of coverings $\mathbf{\Gamma} = \{\Gamma_n\}_{n\in\mathbb{N}}$ whose levels are given as $\Gamma_n = \{f_{\mathbf{i}}(\mathcal{K}) : \mathbf{i} \in \Sigma^n\}$.

Alternatively, the levels of the natural fractal structure of any IFS$-$attractor $\mathcal{K}$ could be described as follows: $\Gamma_1 = \{f_i(\mathcal{K}) : i \in \Sigma\}$, and $\Gamma_{n+1} = \{f_i(A) : A \in \Gamma_n, i \in \Sigma\}$ for each $n \in \mathbb{N}$. That natural fractal structure is starbase (c.f. [13, Theorem 4.7]). In addition, the next remark will be useful for upcoming purposes.

Remark 2.2. All the elements in a same level Γ_n are incomparable.

2.6. **Fractal dimensions for fractal structures.** The fractal dimension models for a fractal structure involved in this paper, i.e., fractal dimensions III and IV, were first introduced in [14, 15], and could be understood as subsequent models from those explored in [16]. It is worth pointing out that they allowed generalizing both box dimension (c.f. [14, Theorem 4.15]) and Hausdorff dimension (c.f. [15, Theorem 3.13]) in the context of Euclidean subsets endowed with their natural fractal structures (c.f. [16, Definition 3.1]). Thus, they become ideal candidates to explore the self-similar structure of IFS$-$attractors [17].

Let $\mathbf{\Gamma}$ be a fractal structure on a metric space (X, ρ), F be a subset of X, and $\mathcal{A}_n(F)$ be the set of all the elements in level n of $\mathbf{\Gamma}$ that intersect F, i.e., $\mathcal{A}_n(F) =$

$\{A \in \Gamma_n : A \cap F \neq \varnothing\}$. In addition, we define $\operatorname{diam}(\Gamma_n) = \sup\{\operatorname{diam}(A) : A \in \Gamma_n\}$, and $\operatorname{diam}(F, \Gamma_n) = \sup\{\operatorname{diam}(A) : A \in \mathcal{A}_n(F)\}$.

Definition 2.3. (c.f. [14, Definition 4.2] and [15, Definition 3.2]) Assume that $\operatorname{diam}(F, \Gamma_n) \to 0$ and consider

$$\mathcal{H}^s_{n,k}(F) = \inf\left\{\sum \operatorname{diam}(A_i)^s : \{A_i\}_{i\in I} \in \mathcal{A}_{n,k}(F)\right\} \text{ for } k = 3, 4, \text{ where}$$

- $\mathcal{A}_{n,3}(F) = \{\mathcal{A}_l(F) : l \geq n\}$.
- $\mathcal{A}_{n,4}(F) = \{\{A_i\}_{i\in I} : A_i \in \cup_{l\geq n}\Gamma_l, F \subseteq \cup_{i\in I} A_i, \operatorname{Card}(I) < \infty\}$ (here, $\operatorname{Card}(I)$ denotes the cardinal number of I).

Moreover, let $\mathcal{H}^s_k(F) = \lim_{n\to\infty} \mathcal{H}^s_{n,k}(F)$. The fractal dimension III (resp., IV) of F, $\dim^k_{\mathbf{\Gamma}}(F)$, is the (unique) critical point satisfying the identity

$$\sup\{s \geq 0 : \mathcal{H}^s_k(F) = \infty\} = \inf\{s \geq 0 : \mathcal{H}^s_k(F) = 0\}.$$

It is worth mentioning that fractal dimension III always exists since the sequence $\{\mathcal{H}^s_{n,3}(F)\}_{n\in\mathbb{N}}$ is monotonic in $n \in \mathbb{N}$. On the other hand, the condition $\operatorname{diam}(F, \Gamma_n) \to 0$, though necessary in Definition 2.3, is not too restrictive as the following remark highlights.

Remark 2.4. Let $\mathcal{K}$ be an IFS$-$attractor endowed with its natural fractal structure as a self-similar set, $\mathbf{\Gamma}$. Then $\operatorname{diam}(\mathcal{K}, \Gamma_n) \to 0$ since the sequence of diameters $\{\operatorname{diam}(\Gamma_n)\}_{n\in\mathbb{N}}$ decreases geometrically.

3. Moran's type theorems under the OSC

The main goal in this paper is to provide some (necessary) conditions on the pieces of an attractor to achieve the identity $\alpha = \dim_{\mathrm{H}}(\mathcal{K})$, which would allow easily calculating the Hausdorff dimension of $\mathcal{K}$.

All the results contributed in this paper stand under the following assumptions, which are stated in the general setting.

IFS conditions 3.1. *Let $k \geq 2$ and $(X, \mathcal{F})$ be an* IFS, *where X is a complete metric space, $\mathcal{F} = \{f_1, \ldots, f_k\}$ is a finite collection of similitudes on X, and $\mathcal{K}$ is the* IFS$-$*attractor of $\mathcal{F}$. In addition, let $\mathbf{\Gamma}$ be the natural fractal structure on $\mathcal{K}$ as a self-similar set (c.f. Definition 2.1) and c_i be the similarity ratio of each similitude $f_i \in \mathcal{F}$.*

Next, we recall the concept of similarity dimension for IFS$-$attractors.

Definition 3.2. Let $\mathcal{F}$ be an IFS and $\mathcal{K}$ its attractor. The similarity dimension of $\mathcal{K}$ is the unique positive solution, α, of the equation $p(s) = \sum_{i=1}^k c_i^s - 1 = 0$.

The similarity dimension of $\mathcal{K}$ will be denoted as α, hereafter. It is worth noting that (without any additional assumption) $\mathcal{H}^\alpha(\mathcal{K}) < \infty$ for any IFS-attractor $\mathcal{K}$ (c.f. [5, Proposition 4 (i)]).

The next result states that the fractal dimension III of $\mathcal{K}$ (c.f. Definition 2.3) equals its similarity dimension without requiring any separation property. On the other hand, we also recall the statement of Moran's Theorem, a classical result in Fractal Geometry that guarantees the equality between the Hausdorff dimension of $\mathcal{K}$ and its similarity dimension whenever $\mathcal{F}$ lies under the OSC.

We would like to point out that each result appearing in this paper has been assigned one of the two labels IFS or EIFS. In the first case, it means that the corresponding result holds for attractors on complete metric spaces, whereas EIFS means that the result stands in the Euclidean setting.

Theorem 3.3 (IFS)**.** *(c.f.* [14, Theorem 4.20]*)*

$$\dim_{\mathbf{\Gamma}}^{3}(\mathcal{K}) = \alpha \textit{ with } 0 < \mathcal{H}_3^{\alpha}(\mathcal{K}) < \infty.$$

Theorem 3.4 (EIFS)**.** *(c.f.* [1, Theorem III] *and* [18]*)*

$$\text{OSC} \Rightarrow \dim_H(\mathcal{K}) = \alpha \textit{ with } 0 < \mathcal{H}_H^{\alpha}(\mathcal{K}) < \infty.$$

Following Theorem 3.4, by a Moran's type theorem, we shall understand a result that yields the equality between a fractal dimension of $\mathcal{K}$, $\dim(\mathcal{K})$, and its similarity dimension α, i.e., $\dim(\mathcal{K}) = \alpha$.

The following result follows by combining both Theorems 3.3 and 3.4 regarding the fractal dimension III of $\mathcal{K}$ in the Euclidean setting.

Corollary 3.5 (EIFS)**.** *(c.f.* [14, Corollary 4.22]*)*

$$\text{OSC} \Rightarrow \dim_H(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^{3}(\mathcal{K}) = \alpha.$$

The next result is quite general and holds for finite fractal structures.

Lemma 3.6. *(c.f.* [15, Proposition 3.5 (3)]*) Let $\mathbf{\Gamma}$ be a finite fractal structure on a metric space (X, ρ), F be a subset of X, and assume that* $\operatorname{diam}(F, \Gamma_n) \to 0$. *Then*

$$\dim_H(F) \le \dim_{\mathbf{\Gamma}}^{4}(F) \le \dim_{\mathbf{\Gamma}}^{3}(F).$$

Corollary 3.7 (IFS)**.**

$$\dim_H(\mathcal{K}) \le \dim_{\mathbf{\Gamma}}^{4}(\mathcal{K}) \le \dim_{\mathbf{\Gamma}}^{3}(\mathcal{K}) = \alpha.$$

In addition, the following result, which could be understood as an extension of Corollary 3.5, provides a Moran's type theorem (under the OSC) involving fractal dimension IV as a consequence of the previous corollaries.

Theorem 3.8 (EIFS)**.**

$$\text{OSC} \Rightarrow \dim_H(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^{4}(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^{3}(\mathcal{K}) = \alpha.$$

To conclude this section, we recall two key results showed by Schief (c.f. [2, 3]), which provide sufficient conditions for Moran's type theorems regarding the Hausdorff dimension of $\mathcal{K}$ in both the Euclidean and the general settings. Such conditions consist of appropriate separation properties for IFS−attractors.

Theorem 3.9.

(EIFS) SOSC $\Leftrightarrow$ OSC $\Leftrightarrow \mathcal{H}^\alpha(\mathcal{K}) > 0 \Rightarrow \dim_H(\mathcal{K}) = \alpha.$

(IFS) $\mathcal{H}^\alpha(\mathcal{K}) > 0 \Rightarrow$ SOSC $\Rightarrow \dim_H(\mathcal{K}) = \alpha.$

Interestingly, Mattila provided a counterexample that ensures that Theorem 3.9 is best possible.

Mattila's Counterexample (c.f. [2, 3])**.** Let $\mathcal{F} = \{f_1, f_2, f_3\}$ be an EIFS on $\mathbb{R}^2$, with similitudes given by $f_i(x) = x_i + \frac{1}{3}(x - x_i)$, where $x_1 = (0,0), x_2 = (1,0)$, and $x_3 = (\frac{1}{2}, \frac{1}{2\sqrt{3}})$. The attractor of $\mathcal{F}$, $\mathcal{K}$, is a nonconnected Sierpiński gasket in the plane which lies under the SOSC and whose similarity dimension is $\alpha = 1$. Thus, almost all the Lebesgue projections of $\mathcal{K}$ on $1-$dimensional subspaces of $\mathbb{R}^2$ (that are self-similar sets themselves) possess Hausdorff dimensions equal to 1 (due to Marstrand's Projection Theorem, c.f. [19, Projection theorem 6.1] and [20, 21]) but zero $\mathcal{H}^1$ measure.

From Theorem 3.9, it holds that the OSC provides a sufficient (though not necessary) condition to get the equality between the similarity and the Hausdorff dimensions of Euclidean IFS$-$attractors. It is worth mentioning that the implication $\mathcal{H}^\alpha(\mathcal{K}) > 0 \Rightarrow$ SOSC, that was proved by Schief (c.f. [2, Theorem 2.1]), guarantees the equivalence among OSC, SOSC, and a $\alpha-$dimensional Hausdorff measure of $\mathcal{K}$. However, in the general setting, the OSC no longer leads to a Moran's type theorem (c.f. [3, Example 3.1]). In this regard, the OSC must be replaced by the SOSC (c.f. [3, Theorem 2.6]). On the other hand, [3, Example 3.2] highlights that the SOSC does not suffice to guarantee a positive Hausdorff measure in the general setting. Also, from Mattila's Counterexample it follows that the implication $\dim_{\mathrm{H}}(\mathcal{K}) = \alpha \Rightarrow$ SOSC cannot be guaranteed, in general.

4. The Level Separation Property

Next, we introduce a novel separation property for each level of the natural fractal structure $\mathbf{\Gamma}$ that any IFS$-$attractor can be always endowed with (c.f. Definition 2.1). In upcoming Theorem 4.2, we prove that such a separation property, named the level separation property, is equivalent to $\mathbf{\Gamma}$ being irreducible or $\mathbf{\Gamma}$ being a tiling. Interestingly, that separation property by levels of $\mathbf{\Gamma}$ states how the natural fractal structure of $\mathcal{K}$ should be in order to achieve a Moran's type theorem for both fractal dimensions III and IV of $\mathcal{K}$ (c.f. Theorem 4.5).

Definition 4.1. (c.f. [4, Definition 3.3]) We say that $\mathcal{F}$ is under the level separation property (LSP, hereafter) if the two following conditions hold for each level Γ_n of $\mathbf{\Gamma}$:

LSP1: $A^\circ \cap B^\circ = \varnothing$ for all $A, B \in \Gamma_n : A \neq B$.

LSP2: $A^\circ \neq \varnothing$ for each $A \in \Gamma_n$,

where the interiors are considered in $\mathcal{K}$.

Notice that the LSP does not depend on an external open set, unlike the OSC.

Let Γ be a covering of X. We recall that Γ is a tiling provided that all the elements in Γ have disjoint interiors and are regularly closed, i.e., $\overline{A^\circ} = A$ for every $A \in \Gamma$. In particular, a fractal structure $\mathbf{\Gamma}$ is said to be a tiling provided that each level Γ_n of $\mathbf{\Gamma}$ is a tiling itself.

The key result in this paper is stated next.

Theorem 4.2 (IFS). *(c.f. [4, Theorem 3.6]) The following are equivalent:*

(1) $\mathbf{\Gamma}$ *irreducible.*
(2) $\mathcal{H}_4^\alpha(\mathcal{K}) > 0$.
(3) $\dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^3(\mathcal{K}) = \alpha$.
(4) LSP.
(5) LSP2 *and* $A^\circ \cap B^\circ = \varnothing$ *for each* $A, B \in \Gamma_1 : A \neq B$.
(6) LSP2 *and* $A_{\mathbf{i}} \subseteq A_{\mathbf{j}}$ *implies* $\mathbf{j} \sqsubseteq \mathbf{i}$.
(7) $\mathbf{\Gamma}$ *tiling.*

It is worth pointing out that the LSP is a separation property for attractors weaker than the OSC in the Euclidean setting.

Corollary 4.3 (EIFS). *(c.f. [4, Proposition 5.4])* OSC $\Rightarrow$ LSP.

The reciprocal of Corollary 4.3 is not true (c.f. Mattila's Counterexample). The following result stands in the general setting as a consequence of Corollary 3.7 and Theorems 3.9 and 4.2:

Corollary 4.4 (IFS). *(c.f. [4, Corollary 3.7])* SOSC $\Rightarrow \alpha = \dim_H(\mathcal{K}) \Rightarrow$ LSP,

where the reciprocal to the first implication is not true (Mattila's Counterexample), and the reciprocal to the second one still remains open (c.f. Open question 5.7).

In particular, the following Moran's type theorem holds for both the fractal dimensions III and IV of $\mathcal{K}$ whenever $\mathcal{F}$ lies under the LSP. That theorem is analogous to the SOSC being sufficient to achieve the identity $\dim_{\mathrm{H}}(\mathcal{K}) = \alpha$ in the general setting (c.f. Theorem 3.9).

Theorem 4.5 (IFS). LSP $\Leftrightarrow \dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^3(\mathcal{K}) = \alpha$.

However, Theorem 4.5 does not guarantee the identity $\dim_{\mathrm{H}}(\mathcal{K}) = \alpha$, unlike Theorem 3.9. In return, it provides an equivalence between the LSP, a separation property for IFS−attractors weaker than the OSC, and the identity $\dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^3(\mathcal{K}) = \alpha$.

5. In summary

In this section, all the results contributed in this paper are summarized.

Theorem 5.1. *Consider the next statements:*

(1) $\mathcal{H}^{\alpha}(\mathcal{K}) > 0$.
(2) SOSC.
(3) OSC.
(4) $\dim_H(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^3(\mathcal{K}) = \alpha$.
(5) $\dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^3(\mathcal{K}) = \alpha$.
(6) $\mathbf{\Gamma}$ *irreducible.*
(7) $\mathbf{\Gamma}$ *tiling.*
(8) $\mathcal{H}_4^{\alpha}(\mathcal{K}) > 0$.
(9) LSP.

The following chains of implications and equivalences hold:

(EIFS) (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Leftrightarrow$ (6) $\Leftrightarrow$ (7) $\Leftrightarrow$ (8) $\Leftrightarrow$ (9).
(IFS) (1) $\Rightarrow$ (2) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Leftrightarrow$ (6) $\Leftrightarrow$ (7) $\Leftrightarrow$ (8) $\Leftrightarrow$ (9).

We also highlight some connections among some separation conditions for attractors.

Theorem 5.2. *Consider the following statements:*

(1) SOSC.
(2) OSC.
(3) LSP.

The next chains of implications and equivalences hold and are best possible:

(EIFS) (1) $\Leftrightarrow$ (2) $\Rightarrow$ (3).
(IFS) (1) $\Rightarrow$ (2) *and* (1) $\Rightarrow$ (3).

A counterexample due to Schief (c.f. [3, Example 3.1]) guarantees the existence of (non-Euclidean) IFS-attractors under the OSC that do not lie under the SOSC.

Schief's Counterexample. Let (X, d) be a complete metric space, where $X = \{(x, n) : x \neq 0, x \in \mathbb{R}, n \in \mathbb{Z}\} \cup \{0\}$ and d is defined as

$$\begin{cases} d(0, (x, n)) = |x| \\ d((x, n), (y, n)) = |y - x| \\ d((x, n), (y, m)) = |x| + |y| \text{ provided that } n \neq m. \end{cases}$$

Further, define the similarities $f_i(0) = 0$ and $f_i(x, n) = (\frac{x}{2}, F_i(n))$ for $i = 1, 2$, where

$$F_1(n) = \begin{cases} 2n & \text{if } n \geq 0 \\ -n & \text{if } n < 0 \text{ odd} \\ n/2 & \text{if } n < 0 \text{ even} \end{cases}, \qquad F_2(n) = \begin{cases} 2n - 1 & \text{if } n > 0 \\ \frac{1}{2}(n-1) & \text{if } n < 0 \text{ odd} \\ -n & \text{if } n \leq 0 \text{ even.} \end{cases}$$

Here, the similarity ratios are $c_1 = c_2 = \frac{1}{2}$, $\alpha = 1$, and $\mathcal{K} = \{0\}$, so $\dim_{\mathrm{H}}(\mathcal{K}) = 0$. The SOSC cannot hold since $\alpha \neq \dim_{\mathrm{H}}(\mathcal{K})$. However, the OSC stands. In fact, $\mathcal{V} = \{(x, n) : x > 0, n > 0\}$ is a feasible open set.

Remark 5.3. Regarding Theorem 5.2, the following implications are not true, in general:

(EIFS) (3) $\Rightarrow$ (2) (c.f. Mattila's Counterexample).

(IFS) (2) $\Rightarrow$ (3) by Schief's Counterexample. In fact, the LSP does not hold since $f_1(\mathcal{K}) = f_2(\mathcal{K})$.
(3) $\Rightarrow$ (2) (c.f. Mattila's Counterexample).
(3) $\Rightarrow$ (1) (c.f. either Corollary 4.3 or Mattila's Counterexample).
(2) $\Rightarrow$ (1) (c.f. Schief's Counterexample).

To conclude this paper, we provide a pair of results (one for each context, EIFS or IFS) comparing our results vs. those proved by Schief.

Theorem 5.4 (EIFS, comparative theorem)**.**

- $\mathcal{H}^\alpha(\mathcal{K}) > 0 \Leftrightarrow \text{OSC} \Leftrightarrow \text{SOSC} \Rightarrow \dim_H(\mathcal{K}) = \alpha$.
- $\text{LSP} \Leftrightarrow \mathcal{H}_4^\alpha(\mathcal{K}) > 0 \Leftrightarrow \dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \alpha$.

Theorem 5.5 (IFS, comparative theorem)**.**

- $\mathcal{H}^\alpha(\mathcal{K}) > 0 \Rightarrow \text{SOSC} \Rightarrow \dim_H(\mathcal{K}) = \alpha$.
- $\text{LSP} \Leftrightarrow \mathcal{H}_4^\alpha(\mathcal{K}) > 0 \Leftrightarrow \dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \alpha$.

Both Theorems 5.4 and 5.5 are best possible since no other implications are true, in general. In fact, in the Euclidean setting, the SOSC becomes sufficient but not necessary (c.f. Mattila's Counterexample) to achieve the identity $\dim_{\mathrm{H}}(\mathcal{K}) = \alpha$. On the other hand, we proved the equivalence among the LSP and the conditions $\dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \alpha$ and $\mathcal{H}_4^\alpha(\mathcal{K}) > 0$ (which is calculated by finite coverings).
In the general setting, though, the SOSC remains only sufficient for a Moran's type theorem for Hausdorff dimension. However, our chain of equivalences involving the LSP is valid for complete metric spaces. We also proved that the SOSC is stronger than the LSP (once again, Mattila's Counterexample works), a separation property for IFS−attractors which does not depend on any external open set (unlike the SOSC). Anyway, both results in Theorem 5.5 (Schief's and ours) can be combined into the next one, which stands in the general setting.

Corollary 5.6 (IFS)**.**

$$\mathcal{H}^\alpha(\mathcal{K}) > 0 \Rightarrow \text{SOSC} \Rightarrow \dim_H(\mathcal{K}) = \alpha \Rightarrow \text{LSP}, \ \textit{where}$$
$$\text{LSP} \Leftrightarrow \mathcal{H}_4^\alpha(\mathcal{K}) > 0 \Leftrightarrow \dim_{\mathbf{\Gamma}}^4(\mathcal{K}) = \alpha.$$

Interestingly, Corollary 5.6 highlights that the LSP becomes necessary to reach the identity $\alpha = \dim_{\mathrm{H}}(\mathcal{K})$. In other words, if the natural fractal structure which any attractor can be endowed with is not irreducible, then a Moran's type theorem cannot hold.

To conclude this paper, we consider interesting to pose some natural questions that still remain open.

Open question 5.7 (EIFS/IFS)**.** Is it true that $\text{LSP} \Rightarrow \dim_{\mathrm{H}}(\mathcal{K}) = \alpha$?

Open question 5.8 (EIFS/IFS)**.** Is it true that $\dim_{\mathrm{H}}(\mathcal{K}) = \dim_{\mathbf{\Gamma}}^{4}(\mathcal{K})$?

Notice that an affirmative response to Open question 5.8 would imply Open question 5.7 being true.

Acknowledgements

M. Fernández-Martínez would like to express his gratitude to Prof. Dr. M.A. Sánchez-Granero for a fruitful collaboration (c.f. [4]) on which the current work has been based.

The author appreciates the partial support of Ministerio de Ciencia, Innovación y Universidades, grant number PGC2018-097198-B-I00, and Fundación Séneca of Región de Murcia, grant number 20783/PI/18.

References

[1] P. A. P. Moran, *Additive functions of intervals and Hausdorff measure*, Mathematical Proceedings of the Cambridge Philosophical Society **42** (1946), no. 1, 15–23.
[2] A. Schief, *Separation properties for self-similar sets*, Proceedings of the American Mathematical Society **122** (1994), no. 1, 111–115.
[3] A. Schief, *Self-similar sets in complete metric spaces*, Proceedings of the American Mathematical Society **124** (1996), no. 2, 481–490.
[4] M.A. Sánchez-Granero, M. Fernández-Martínez, *Irreducible fractal structures for Moran's type theorems*, Chaos, Solitons & Fractals **119** (2019), 29–36.
[5] J.E. Hutchinson, *Fractals and self-similarity*, Indiana University Mathematics Journal **30** (1981), no. 5, 713–747.
[6] C. Bandt, S. Graf, *Self-similar sets. VII. A characterization of self-similar fractals with positive Hausdorff measure*, Proceedings of the American Mathematical Society **114** (1992), no. 4, 995–1001.
[7] S.P. Lalley, *The packing and covering functions of some self-similar fractals*, Indiana University Mathematics Journal **37** (1988), no. 3, 699–710.
[8] Y. Peres, M. Rams, K. Simon, B. Solomyak, *Equivalence of positive Hausdorff measure and the open set condition for self-conformal sets*, Proceedings of the American Mathematical Society **129** (2001), no. 9, 2689–2699.
[9] N. Patzschke, *The strong open set condition for self-conformal random fractals*, Proceedings of the American Mathematical Society **131** (2003), no. 8, 2347–2358.
[10] F. Hausdorff, *Dimension und äußeres Maß*, Mathematische Annalen **79** (1918), no. 1-2, 157–179.
[11] C. Bandt, T. Retta, *Topological spaces admitting a unique fractal structure*, Fundamenta Mathematicae **141** (1992), no. 3, 257–268.
[12] F. G. Arenas, M.A. Sánchez-Granero, *A characterization of non-Archimedeanly quasimetrizable spaces*, Rend. Istit. Mat. Univ. Trieste **30** (1999), no. suppl., 21–30.
[13] F. G. Arenas, M.A. Sánchez-Granero, *A characterization of self-similar symbolic spaces*, Mediterranean Journal of Mathematics **9** (2012), no. 4, 709–728.
[14] M. Fernández-Martínez, M.A. Sánchez-Granero, *Fractal dimension for fractal structures: A Hausdorff approach*, Topology and its Applications **159** (2012), no. 7, 1825–1837.
[15] M. Fernández-Martínez, M.A. Sánchez-Granero, *Fractal dimension for fractal structures: A Hausdorff approach revisited*, Journal of Mathematical Analysis and Applications **409** (2014), no. 1, 321–330.
[16] M. Fernández-Martínez, M.A. Sánchez-Granero, *Fractal dimension for fractal structures*, Topology and its Applications **163** (2014), 93–111.

[17] M. Fernández-Martínez, M.A. Sánchez-Granero, J.E. Trinidad Segovia, Juan L.G. Guirao, *Fractal dimensions for fractal structures with Applications to Finance*, Springer Nature International Publishing, Cham, 2019.

[18] M. Fernández-Martínez, Juan L.G. Guirao, G. Rodríguez-Bermúdez, *More than seventy years from a milestone in fractal geometry: Moran's theorem*, Chaos **29** (2019), no. 1, 013106.

[19] K. Falconer, *Fractal geometry. Mathematical Foundations and Applications*, 1st ed., John Wiley & Sons, Ltd., Chichester, 1990.

[20] R. Kaufman, *On Hausdorff dimension of projections*, Mathematika **15** (1968), no. 2, 153–155.

[21] J. M. Marstrand, *Some fundamental geometrical properties of plane sets of fractional dimensions*, Proceedings of the London Mathematical Society **s3-4** (1954), no. 1, 257–302.

University Centre of Defence at the Spanish Air
Force Academy, Spain
manuel.fernandez-martinez@cud.upct.es

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON A REGULARISATION OF A NONLINEAR DIFFERENTIAL EQUATION RELATED TO THE NON-HOMOGENEOUS AIRY EQUATION

GALINA FILIPUK, THOMAS KECKER & FEDERICO ZULLO

Abstract. In this paper we study a nonlinear differential equation related to a non-homogeneous Airy equation. The linear equation has two families of solutions. We apply a procedure of resolution of points of indeterminacy to a system of first order differential equations equivalent to the nonlinear equation and study how the corresponding families of solutions are transformed.

Keywords: the Airy equation, the Painlevé test, nonlinear differential equations, singularities

MSC2010: 34M55, 14E05, 14E15

Received 13 April 2021; revised 5 July 2021; accepted 19 July 2021

1. Introduction

The non-homogeneous Airy equation is given by

$$\frac{d^2y(z)}{dz^2} = zy(z) + c, \tag{1.1}$$

where $c \in \mathbb{C}$ is a constant. In [13] some observations on the distribution of zeros of solutions of the non-homogeneous Airy equation were presented. The existence of a principal family of solutions, with simple zeros, and particular solutions, characterized by a double zero at a given position of the complex plane, was shown. In addition, a recursion describing the distribution of the zeros was introduced. These results generalise previous results on the distribution of zeros of solutions of the corresponding homogeneous equation [14, 15]. As shown in

 https://doi.org/10.2478/9788366675360-008.

[13], a simple extension to the non-homogeneous case (by adding a constant term to the homogeneous equation) considerably changes the distribution of the zeros.

Equation (1.1) possesses entire solutions with order of growth equal to 3/2. Since this number is not an integer, all solutions of (1.1) possess an infinite number of zeros [12]. These zeros are movable, in the sense that if the initial conditions change, the positions of the zeros change.

Entire solutions of a linear second order non-homogeneous differential equation with entire coefficients may have double zeros. Indeed, for any $p \in \mathbb{C}$, equation (1.1) possesses always a solution with a double zero in p. More precisely (see [13]), given $p \in \mathbb{C}$, equation (1.1) possesses just one solution $\tau_d(z,p)$, proportional to c, with a double zero in $z = p$. This solution is defined by the series

$$\tau_d(z,p) = c\sum_{n=2} e_n(z-p)^n, \quad e_2 = \frac{1}{2},\ e_3 = 0,\ e_4 = \frac{p}{24},\ e_{n+2} = \frac{pe_n + e_{n-1}}{(n+1)(n+2)}. \tag{1.2}$$

The set of functions $\tau_d(z,p)$ then represents a family of particular solutions of equation (1.1). Besides the family $\tau_d(z,p)$, equation (1.1) possesses a *principal family* of solutions. Indeed, following [13] one has the following statement. Given $q \in \mathbb{C}$, equation (1.1) possesses a solution $\tau(z,q,\alpha)$ with a simple zero in $z = q$. This solution is defined by the series

$$\tau(z,q,\alpha) = \sum_{n=1} f_n(z-q)^n, \quad f_1 = \alpha,\ f_2 = \frac{c}{2},\ f_3 = \frac{q\alpha}{6},\ f_{n+2} = \frac{qf_n + f_{n-1}}{(n+1)(n+2)}. \tag{1.3}$$

Here q and α are arbitrary parameters and so the series (1.3), that converges everywhere in the complex plane, can be considered as the general solution of equation (1.1).

2. Main results

The main objective of this paper is to study the nonlinear differential equation obtained from the non-homogeneous Airy equation using the logarithmic derivative of $y(z)$, i.e., by taking

$$u(z) = \frac{y'(z)}{y(z)}.$$

The function u solves the following differential equation

$$\frac{d^2u(z)}{dz^2} + 3u(z)\frac{du(z)}{dz} + u(z)^3 = 1 + zu(z). \tag{2.1}$$

Equation (2.1) is a coupled Riccati equation in disguise, that is, it can be written as $w = u' + u^2 - z$, where $w' + uw = 0$. Therefore, it possesses a one parameter family of Riccati solutions when $w = 0$. Clearly, any entire solution y of (1.1) gives a meromorphic solution u of (2.1). There is the so-called Painlevé property, which demands that all movable singularities of the equation be poles. The Painlevé test is a useful criterion of integrability. Inserting a formal Laurent series into the equation, after determining the leading order of a possible solution, one can recursively compute the coefficients of the series. If there is no obstruction in computing the coefficients and a sufficient number of such formal Laurent series solutions exist, the equation is said to pass the test (see e.g. [1, 4]). Applying the Painlevé test to equation (2.1) it is not difficult to see [13] that the dominant balances give singularities of the type $c_0(z-p)^{-1}$, but, seeking the resonances, one finds that there are two families of solutions. One family is characterized by $c_0 = 1$ and with a resonance polynomial given by $(r-1)(r+1)$. The other one is characterized by $c_0 = 2$ and with a resonance polynomial given by $(r+1)(r+2)$. The resonance $r = -1$ corresponds to the arbitrariness of the position of the pole for $u(z)$, i.e. the arbitrariness of the position of the zero $z = p$ for $y(z)$. The coefficients of the Laurent expansion of the logarithmic derivative of $y(z)$ are explicitly connected with the zeros of $y(z)$. In particular, in the first family, $c_0 = 1$ implies that the zero in $z = p$ of $y(z)$ is simple, whereas the resonance $r = 1$ indicates that the constant term in the Laurent expansion of $u(z)$ is the other arbitrary constant describing the solutions of the second order equation (2.1). This is the principal family of solutions [1]. In the second family, $c_0 = 2$ implies that the zero in $z = p$ of $y(z)$ is double, whereas the resonance $r = -2$ is negative: this indicates that the second family is a particular solution of equation (2.1) [1].

As mentioned in [13], the non-homogeneous Airy equation has several applications in mathematical physics [8]. For example, it is related to the second member of the Burger's hierarchy [7]:

$$\psi_t + (\psi_{xx} + \psi^3 + 3\psi\psi_x)_x = 0. \tag{2.2}$$

Under the self-similarity transformation

$$\psi = \frac{1}{\eta(3t)^{1/3}} f(z), \quad z \doteq \frac{x}{\eta(3t)^{1/3}} - \frac{b}{\eta^3}, \quad \eta^3 = a \tag{2.3}$$

equation (2.2) becomes

$$\frac{d^3 f}{dz^3} + 3\left(\frac{df}{dz}\right)^2 + 3f\frac{d^2 f}{dz^2} + 3f^2\frac{df}{dz} = (az+b)\frac{df}{dz} + af. \tag{2.4}$$

Integrating once one gets

$$\frac{d^2 f}{dz^2} + 3\frac{df}{dz}f + f^3 = k + (az+b)f(z), \tag{2.5}$$

where k is the integration constant. Some of the solutions of (2.5) have been considered in [7] for the description of liquids with gas bubbles. Equation (2.1) is a particular case of (2.5) with $k = a = 1$, $b = 0$. In general, in case $a = k$ we can scale solutions of (2.5) by taking $f(z) = a^{1/3}u(a^{1/3}z + b/a^{2/3})$ and get equation (2.1).

Solutions of a given non-linear second-order ordinary differential equation in general have infinitely many singularities in the complex plane, the location of which, apart from a finite number of fixed singularities of the equation, depends on the initial data of the equation. These singularities are therefore called movable. The method of blowing up points of indeterminacy of certain systems of two ordinary differential equations was applied to obtain information about the singularity structure of the solutions of the corresponding nonlinear differential equations in [2, 6].

Equation (2.1) is similar to the Painlevé-Ince equation

$$u'' + 3uu' + u^3 = 0 \tag{2.6}$$

considered in [2]. When the usual Painlevé analysis is applied, the Painlevé-Ince equation also possesses both positive and negative resonances. This equation can also be solved explicitely. The general solution is given by $u = 1/(z-a)+1/(z-b)$ with a and b constants [1]. All solutions are rational functions and, therefore, equation (2.6) possesses the Painlevé property. In [2] three equivalent to equation (2.6) systems were studied. It was shown that for all of them there is an infinite sequence of blow-ups for one of the base points and another one that terminates, which further gives a Laurent expansion of the solution around a movable pole. The aim of this paper is to study equation (2.1) in a similar way.

A blow-up is a construction originating in algebraic geometry to de-singularise an algebraic curve [10]. It can be adapted to the setting of differential equations where it serves to regularise a system of equations at points of indeterminacy of the equations. Okamoto studied all six Painlevé equations in their Hamiltonian form from a geometric point of view in [9], where he introduced the notion of a space of initial conditions, obtained by blowing up the phase space at a finite sequence of points. Points of indeterminacy also known as base points of the system are the points where the vector field is ill-defined. Geometrically the blow-up procedure separates lines through base points according to their slopes and, hence, adds a projective line, which is called an exceptional divisor. The blow-up at a point $(x, y) = (a, b)$, where $a = a(z)$ and $b = b(z)$ can in general be rational functions in z, is defined by the following construction. One introduces new coordinate charts, $x = a + u = a + UV$ and $y = b + uv = b + V$ and re-writes the system in new coordinates (u, v) and (U, V). The exceptional line then corresponds to $u = 0$ or $V = 0$. For more information and examples of application to the Painlevé equations see [5].

Let us consider the following system equivalent to equation (2.1):

$$u' = v - 3/2u^2, \quad v' = -(u^3 - 1 - zu). \tag{2.7}$$

Following the procedure described in [2], we extend the system to study it on $\mathbb{P}^1 \times \mathbb{P}^1$ and find points of indeterminacy of the vector field. Let us rename $u = q$, $v = p$ not to confuse the notation. Since we have a polynomial vector field, there are no points of indeterminacy in the coordinate chart (q, p) (where both the numerator and the denominator on the right hand sides of equations vanish). In the chart $(Q, p) = (1/q,\, p)$ the system is

$$Q' = \frac{3 - 2pQ^2}{2}, \quad p' = \frac{Q^3 + zQ^2 - 1}{Q^3}.$$

We see that there are no points of indeterminancy of the vector field. In the chart $(q,\, P) = (q,\, 1/p)$ we have

$$q' = \frac{2 - 3Pq^2}{2P}, \quad P' = -P^2(1 + zq - q^3).$$

Here we also see that there are no points of indeterminancy of the vector field. In the chart $(Q,\, P) = (1/q,\, 1/p)$ the system is

$$Q' = \frac{3P - 2Q^2}{2P}, \quad P' = -\frac{P^2(Q^3 + zQ^2 - 1)}{Q^3}.$$

We see that when $Q = P = 0$ we have a point of indeterminacy.

To regularise the system (2.7), we need to resolve the base point at $(P,\, Q) = (0,\, 0)$. If the regularisation is successful (i.e., if the base point can be resolved in a finite number of blow-ups), then one can fully describe singularities of solutions of the differential equation [6, 2]. However, in some differential equations it might happen that one or more cascades of base points do not terminate after a reasonable (small) finite number of steps, in such case we call the cascade infinite. The calculations which we did for the case of system (2.7) show that the cascade splits into one finite and one infinite cascade. In particular, in the first step we see that by resolving the first point of indeterminacy $(Q,\, P) = (0,\, 0)$, that is, by taking $Q = u_1$, $P = u_1 v_1$, we can find the second point of indeterminacy $(u_1,\, v_1) = (0,\, 0)$. The corresponding system is

$$u_1' = \frac{3}{2} - \frac{u_1}{v_1}, \quad v_1' = 1 - \frac{3v_1}{2u_1} - zv_1^2 + \frac{v_1^2}{u_1^2} - u_1 v_1^2.$$

Then the cascade splits. We have new coordinates $u_1 = u_2$, $v_1 = u_2 v_2$ and new points of indeterminacy $(u_2,\, v_2) = (0,\, 1)$ and $(u_2,\, v_2) = (0,\, 2)$. The corresponding system is

$$u_2' = \frac{3}{2} - \frac{1}{v_2}, \quad v_2' = \frac{2 + v_2(v_2 - 3 - u^2(z + u_2)v_2)}{u_2}.$$

The first cascade is then infinite in the sense that it does not resolve in a reasonable number of steps. For the second one the system becomes regular after one more blow up. By taking $u_2 = u_3$, $v_2 = 2 + u_3 v_3$ one finds the system for u_3 and v_3 which has no more points of indeterminacy of the vector field. So one of the cascades of points of indeterminacy is resolved. We have the following statement.

Theorem 2.1. *The transformation*

$$u = \frac{1}{u_3}, \quad v = \frac{1}{u_3^2(2 + u_3 v_3)}.$$

transforms system (2.7) into

$$u_3' = \frac{4+3u_3v_3}{2(2+u_3v_3)}, \quad v_3' = -\frac{f(u_3,v_3)}{2(2+u_3v_3)}, \tag{2.8}$$

where

$$f(u_3,v_3) = 16(z+u_3)+24u_3(z+u_3)v_3+3(4u_3^2(z+u_3)-1)v_3^2+2u_3(u_3^2(z+u_3)-1)v_3^3,$$

which has no more points of indeterminacy of the vector field.

Theorem 2.2. *The principal family (1.3) corresponds to the following expansion for the function u:*

$$u(z) = \frac{1}{z-q} + \frac{c}{2\alpha} + \frac{(4q\alpha^2-3c^2)(z-q)}{12\alpha^2} + O((z-q)^2).$$

The function v then has expansion

$$v(z) = \frac{1}{2(z-q)^2} + \frac{3c}{2\alpha(z-q)} + \frac{32q\alpha^2-15c^2}{24\alpha^2} + O((z-q)).$$

Expansions of the functions u_3 and v_3 are, respectively,

$$u_3(z) = (z-q) - \frac{c(z-q)^2}{2\alpha} + \left(\frac{c^2}{2\alpha^2} - \frac{q}{3}\right)(z-q)^3 + O((z-q)^4),$$

$$v_3(z) = -\frac{4c}{\alpha} - \frac{4(q\alpha^2-3c^2)(z-q)}{\alpha^2} + O((z-q)^2).$$

On the other hand, if we use system (2.8) and search for its solutions in the form of the Taylor series with conditions $u_3(q) = 0$, $v_3(q) = -4c/\alpha$, then we recover this solution. Moreover, from the general theorems of regular systems we can find the radius of convergence of such series, so regularising the system is very important in undertanding the behaviour of solutions.

Theorem 2.3. *The second family (1.2) corresponds to*

$$u(z) = \frac{2}{z-p} + \frac{1}{6}p(z-p) + O((z-p)^2),$$

$$v(z) = \frac{4}{(z-p)^2} + \frac{7p}{6} + O((z-p)).$$

We further find that

$$u_3(z) = \frac{(z-p)}{2} - \frac{1}{24}p(z-p)^3 + O((z-p)^3),$$

$$v_3(z) = -\frac{2}{z-p} - \frac{5}{12}p(z-p) + O((z-p)^2),$$

so v_3 has a pole at p.

To reproduce this expansion from system (2.8) we take one further transformation $v_3 = 1/V_3$, so

$$V_3(z) = -\frac{z-p}{2} + \frac{5}{48}p(z-p)^3 + O((z-p)^4),$$

and from the system for u_3 and V_3 we can reproduce this expansion by searching for a solution with $u_3(p) = 0$, $v_3(p) = 0$ and $u_3'(p) = 1/2$, which also gives $v_3'(p) = -1/2$.

However, the system for u_3 and V_3 has further points of indeterminacy. This is refleced in the fact that we had an infinite cascade for the original system and we cannot regularise the system completely. We would also like to remark that we can find other equivalent systems to equation (2.1) and the cascades of points of indeterminacy will be different but still infinite. The computations are essentially the same but cumbersome so we omit them.

Since equation (2.1) has meromorphic solutions, it would be interesting to understand how to regularise the system completely and thus construct a space of initial conditions in the sense of Okamoto [9] for it. As it is well known, for the Painlevé equations, the method of blowing up the space of dependent variables at points of indeterminacy leads to the space of initial conditions. In certain proofs for the Painlevé property of the Painlevé equations (e.g. [3] and [11]) a main part is played by a system of equations in transformed coordinates where the system becomes regular at points where the original variables tend to infinity. As mentioned before, from the relation to the non-homogeneous Airy equation, equation (2.1) has meromorphic solutions. However, the regularisation by blow-ups, in comparison with the Painlevé equations, does not work properly in the sense that one of the cascades of points of indeterminacy seems to be infinite and so a similar formal proof of the Painlevé property using the regularising system is not possible. It would be interesting to understand further the question of connection between regularisation of the system of equations and the Painlevé property of the related nonlinear differential equation. This deserves further study. Other open problems include understanding the distribution of poles of nonlinear equations. In this case the dynamics of the zeros of the non-homogeneous Airy equation (1.1) may help understand the behavior of poles of the related nonlinear equation (2.1).

Acknowledgments

GF acknowledges the support of National Science Center (Narodowe Centrum Nauki NCN) OPUS grant 2017/25/B/BST1/00931 (Poland). FZ acknowledges the support of University of Brescia, INdAM-GNFM and INFN.

References

[1] R. Conte, A.P. Fordy, A. Pickering, *A perturbative Painlevé approach to nonlinear differential equations,* Phys. D **69** (1993), no. 1-2, 33–58.

[2] G. Filipuk, T. Kecker, *On singularities of certain non-linear second-order ordinary differential equations*, in preparation.

[3] A. Hinkkanen and I. Laine, *Solutions of the first and second Painlevé equations are meromorphic*, J. Anal. Math. **79** (1999), 345–77.

[4] A.N.W. Hone, *Painlevé tests, singularity structure and integrability*, in Integrability, A.V. Mikhailov (ed.), Springer Lect. Notes Phys. **767**, 245–277, 2009.

[5] K. Kajiwara, M. Noumi, Y. Yamada, *Geometric aspects of Painlevé equations*, J. Phys. A **50** (2017), no. 7, 073001, 164 pp.

[6] T. Kecker, G. Filipuk, *Desingularising transformations for complex differential equations with algebraic singularities*, submitted, https://arxiv.org/abs/2101.03432.

[7] N. A. Kudryashov, D. I. Sinelshchikov, *Analytical and numerical studying of the perturbed Korteweg-de Vries equation*, Journal of Mathematical Physics **55** (2014), 103504, doi: 10.1063/1.4897445.

[8] A. J. MacLeod, *Computation of inhomogeneous Airy functions*, Journal of Computational and Applied Mathematics **53** (1994), 109–116.

[9] K. Okamoto, *Sur les feuilletages associés aux équations du second ordre à points critiques fixes de P. Painlevé*, Japan. J. Math. (N.S.) **5** (1979), 1–79.

[10] I. R. Shafarevich, *Basic Algebraic Geometry 1. Varieties in projective space*, Third ed., Springer, Heidelberg, 2013.

[11] S. Shimomura, *Proofs of the Painlevé property for all Painlevé equations*, Japan. J. Math. **29** (2003), 159–180.

[12] E.C. Titchmarsh, *The Theory of Functions*, Oxford University Press, London, 1939.

[13] F. Zullo, *Notes on the zeros of the solutions of the non-homogeneous Airy's equation*, accepted for publication in the proceedings of FASnet conference, to be published by World Scientific, 2021.

[14] F. Zullo, *On the solutions of the Airy equation and their zeros*, in Complex Differential and Difference Equations, G. Filipuk, A. Lastra, S. Michalik, Y. Takei and H. Żołądek eds., De Gruyter Proceedings in Mathematics, De Gruyter, Berlin/Boston, 2020.

[15] F. Zullo, *On the dynamics of the zeros of solutions of the Airy equation*, Mathematics and Computers in Simulation **176** (2020), 312–318.

University of Warsaw, Poland
`filipuk@mimuw.edu.pl`

University of Portsmouth, United Kingdom
`thomas.kecker@port.ac.uk`

University of Brescia, Italy
`federico.zullo@unibs.it`

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

A BRIEF PANORAMA OF PERIOD-LIKE MOTIONS IN DYNAMICAL SYSTEMS

KAROL GRYSZKA

Abstract. In the article we draw our attention on the selection of existing notions of non-periodic motions that generalize periodic motions. These are from classic recurrence to slightly less known and recent ones such as near periodicity or G-asymptotic periodicity. We provide mutual relations between all of them and fill the existing gaps that arise in their investigation.

Keywords: Periodicity, Recurrence, Almost periodicity, Asymptotic periodicity, Near periodicity, Stability

MSC2020: 37B20, 37C25, 34D05, 34C99, 34D20

Received 28 April 2021; revised 10 May 2021; accepted 19 June 2021

1. Introduction

Generalizations of periodic motions in flows appeared in 1920's and since then, many variations to the non-periodic case have been provided. It appears that authors usually skip the general comparison between notions, thus motivated by that we provide 10 generalizations of periodicity that are defined using topology or metric of the space, including the classic ones as well as the less known ones. Then we formulate their mutual relations and present examples in case one does not imply the other. Our approach is focused on providing examples showing that some relations fail to occur without additional assumptions. As a result, these notions cannot be freely interchangeable.

 https://doi.org/10.2478/9788366675360-009.

2. Basic definitions and properties

Let us start by introducing fundamental definitions used in the entire paper. Throughout the paper X is a metric space, and the corresponding metric function on X is d.

2.1. **Dynamical systems.** A *dynamical system* (a flow) ϕ is a continuous mapping $\phi\colon \mathbb{R} \times X \to X$ such that $\phi(0, x) = x$ and for any x, s and t, $\phi(t, \phi(s, x)) = \phi(t + s, x)$. We call X a *phase space* of ϕ. A *motion through* x is the mapping $t \mapsto \phi(t, x)$. We will identify properties of the motion through x with properties of the orbit of x or the point itself. Given dynamical system ϕ and $x \in X$, the set $o(x) = \phi(\mathbb{R}, x)$ is the *orbit* of x, $o^+(x) = \phi([0, +\infty), x)$ is the *positive orbit* of x and $o^-(x) = \phi((-\infty, 0], x)$ is the *negative orbit*. A point x is *stationary* if $x = \phi(t, x)$ for any $t \in \mathbb{R}$. If for some $T > 0$ we have $x = \phi(T, x)$ and x is not stationary, then x is *periodic*. If $T > 0$ is the smallest such that $x = \phi(T, x)$, then we say that x is *T-periodic* and call T the *period* of x. A set U is *invariant* (resp., *positively invariant, negatively invariant*) if $\phi(\mathbb{R}, U) \subset U$ (resp., $\phi([0, +\infty), U) \subset U$, $\phi((-\infty, 0], U) \subset U$). A set M is *minimal*, if it is invariant, closed, non-empty and no proper subset of M has these properties.

The following definitions refer to various limit properties of the orbit. The *ω-limit set* $\omega(x)$ consists of all points $y \in X$ such that there exists a strictly increasing and diverging to $+\infty$ sequence $(t_n)_{n\in\mathbb{N}}$ of times with the property: $\phi(t_n, x) \to y$. The set $J^+(x)$, called *first positive prolongation limit set*, consists of all points $y \in X$ such that there exists a strictly increasing and diverging to $+\infty$ sequence $(t_n)_{n\in\mathbb{N}}$ and a sequence $(x_n)_{n\in\mathbb{N}}$ of points from X such that $d(x_n, x) \to 0$ and $\phi(t_n, x_n) \to y$. If in the previous definition the only limitation to t_n is that $t_n \geq 0$, then we describe the set $D^+(x)$ called *first positive prolongation set*. By changing values of time to negative and limits to $-\infty$ we describe the sets $\alpha(x), J^-(x)$ and $D^-(x)$, called *α-limit set*, *first negative prolongation limit set* and *first negative prolongation set*, respectively.

The above definitions play a fundamental role in the theory of stability and topology of dynamical systems. For a very wide preview of the theory see for instance [6, 18, 19, 35, 39, 47].

Note the following result, which is a direct consequence of the definitions.

Lemma 2.1. *We have $\omega(x) \subset J^+(x) \subset D^+(x)$ and $o^+(x) \cup J^+(x) = D^+(x)$. Similarly, we have $\alpha(x) \subset J^-(x) \subset D^-(x)$ and $o^-(x) \cup J^-(x) = D^-(x)$.*

It is worth mentioning that the concepts of limit sets and prolongation sets are topological concepts (see for example [6]).

2.2. **Stability.** Stability theory of dynamical systems and/or their orbits play substantial role in the field of dynamical systems. There are many ways of defining and describing stability-like conditions. Some of the conditions are related to the limit set, while different conditions are related to the behaviour of points and orbits in the vicinity of a given point/orbit.

We say that a point x is *positively Poisson stable* (resp. *negatively Poisson stable, Poisson stable*) (abbreviated: P^+-stable, resp. P^--stable, P-stable) if $x \in \omega(x)$ (resp. $x \in \alpha(x)$, $x \in \omega(x) \cap \alpha(x)$).

A different approach is presented by the Lagrange stability. Here and later we only define positively stable motions, similar definitions can be formulated for negatively stable and stable motions. A point x is *positively Lagrange stable* (abbreviated: L^+-stable) if the set $\overline{o^+(x)}$ is compact. A motion through x is *positively Lyapunov stable* in a subset N of X if for any $\varepsilon > 0$ there is $\delta > 0$ such that $y \in N \cap B(x, \delta)$ implies $d(\phi(t, x), \phi(t, y)) < \varepsilon$ for $t \geq 0$. If in the definition we take N as a neighborhood of x, then we can delete "in a subset N of X". In such a case the definition simplifies to the following condition: for any $\varepsilon > 0$ there is $\delta > 0$ such that if $d(x, y) < \delta$, then $d(\phi(t, x), \phi(t, y)) < \varepsilon$ for $t \geq 0$. In the literature Lyapunov stable motions are simply called stable motions.

There are other, either weaker or stronger stability-like conditions than the ones just presented. One such, called *semi-stability*, was introduced in [37]; in [24] the author defines *positively Lipschitz stable* motions (this concept is stronger than Lyapunov stability). See [41] for further details and applications. See also [6, 35, 39, 47] for an overview of the theory of stability.

2.3. **Recurrence and almost periodicity.** One of the most intuitive generalization of periodicity is to demand that the point remain close to the starting position after a specific amount of time. Such a generalization is described by the notion of recurrence, introduced by Birkhoff in [7]. Birkhoff's original idea was to generalize various notions from the theory of differential equations and to introduce the notion of a minimal flow. The latter was an origin of Birkhoff-recurrent point and recurrent flows.

Definition 2.2 (Generalization **1**.)**.** A point $x \in X$ (an orbit of x) is *recurrent* if for any $\varepsilon > 0$ there is a positive number T (depending on ε) such that for any $t \in \mathbb{R}$ and any $\alpha \in \mathbb{R}$ there is t_0 such that

$$\alpha - T \leq t_0 \leq \alpha + T \quad \text{and} \quad d(\phi(t, x), \phi(t_0, x)) < \varepsilon.$$

The importance of studying recurrent orbits is significant. This is because they have a very direct connection with minimal sets.

Lemma 2.3 (folklore)**.** *(1) Every orbit in a compact minimal set is recurrent. Every compact minimal set is the closure of a recurrent orbit.*

(2) If x is recurrent and the set $\overline{o(x)}$ is compact, then $o(x)$ is a minimal set.

(3) In complete metric spaces the closure of a recurrent orbit is a compact minimal set.

It is also clear that recurrent motions are P-stable.

The characterization of recurrent motions provided in Lemma 2.3 is due to Birkhoff [7].

The more general approach to periodicity is due to Bohr, who introduced almost periodic motions [9, 10, 11]. Recall that a set $D \subset \mathbb{R}$ is *relatively dense* in $\mathbb{R}$ if there exists $L > 0$ such that for any $t \in \mathbb{R}$ we have $D \cap [t - L, t + L] \neq \emptyset$.

Definition 2.4 (Generalization **2.**)**.** A point x is *almost periodic* if for any $\varepsilon > 0$ there exists a relatively dense sequence $(\tau_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$, called the displacements set, such that for any $t \in \mathbb{R}$ and any $n \in \mathbb{N}$ we have $d(\phi(t, x), \phi(t + \tau_n, x)) < \varepsilon$.

An analogous to Lemma 2.3 result also describes almost periodic motions. It can be found in most basic handbooks on dynamical systems (for instance in [6, 35]).

Lemma 2.5 (folklore)**.**

(1) If the set $\overline{o(x)}$ is compact and $x \in M$ is almost periodic, then every point $y \in \overline{o(x)}$ is almost periodic with the same set of displacements for a given ε, but with the strict inequality $<$ replaced by $\leq$.
(2) If M is a compact minimal set and one point of M is almost periodic, then every other point of M is almost periodic.
(3) If x is almost periodic, then it is recurrent.

Further properties of almost periodic motions are connected to the Lyapunov stability.

Theorem 2.6 (Bochner, [8])**.** *If x is almost periodic and $\overline{o(x)}$ is a compact set, then the motion through x is Lyapunov stable in both directions in the set $\overline{o(x)}$.*

Theorem 2.7 (Markov, [30])**.** *If a motion through x is recurrent and positively Lyapunov stable in $o(x)$, then it is almost periodic.*

More properties of almost periodic motions can be found in [6, 14] and in the papers of Bochner [8] and Markov [30].

Notions of recurrence and almost periodicity are based on the global property of the orbit (we demand that inequalities hold for every point from the orbit). However, there are definitions of recurrence that are limited to positive orbits (see [41]). Similarly to the latter one can also provide a definition of almost periodic motion for positive time only. In this article, some of the further definitions require the presence of another orbit in its entirety and rely on the dynamics of the whole orbit as well. One can of course work with positively (or negatively) recurrent motions, but that is not necessary for the purpose of this paper (see [4] and [41]).

Note that there is a notion of weak almost periodicity in flows, investigated by Ellis and Nerurkar [16]. Such a notion is based on the weak topology of maps in the space of continuous function, which is not suitable for our investigation.

The notion of almost periodicity was later generalized by von Neumann [36] in groups, and further generalized by many others (see the research monographs

[3, 17, 29] for the basic theory of almost-periodic functions; for some applications see the research monographs [15, 23, 26, 48]). Note that the theory of almost-periodic functions is still a very active field of investigations of many authors (see for example the recent paper of Kostić and Du [27]). It is also full of open problems and possibilities for further expansions.

2.4. **G-asymptotically periodic motions.** The following generalization relies on the asymptotic behaviour of the orbit outside of a small neighbourhood of a point belonging to the positive orbit of x. This idea was introduced in [20]. We briefly introduce the necessary notation.

Let ϕ be a flow on X. Fix $x \in X$ and $\varepsilon > 0$, and define

$$A(x, \varepsilon) := \{t \geq 0 \mid d(\phi(t, x), x) > \varepsilon\}.$$

This set is the union of at most countably many pairwise disjoint and open intervals denoted by (q_i, r_i). Define

$$w_{x,\varepsilon}(t) := \begin{cases} 0, & t \notin A(x, \varepsilon), \\ \operatorname{diam}(q_i, r_i), & t \in (q_i, r_i). \end{cases}$$

The set $W_{x,\varepsilon} := \{w_{x,\varepsilon}(t) \mid t \geq 0\}$ contains at most countably many different non-negative real numbers, including $+\infty$ if necessary. Set

$$W(x, \varepsilon) := \limsup_{t \to +\infty} w_{x,\varepsilon}(t).$$

Definition 2.8 (Generalization **3.**)**.** The *G-asymptotic period* of x (of the orbit of x) is defined as

$$\text{G-AP}(x) := \lim_{\varepsilon \to 0} \limsup_{t \to +\infty} W(\phi(t, x), \varepsilon).$$

If $\text{G-AP}(x) = 0$, then x is called *G-asymptotically fixed.* If x has a finite asymptotic period, then it is called *G-asymptotically periodic.* If $\text{G-AP}(x) = +\infty$, then a point x is called *G-asymptotically non-periodic.*

Originally, the author of [20] called *G-asymptotically periodic* points as *asymptotically periodic* points. However, in [22] it was fixed to avoid confusion with other notions and the prefix "*G-*" was added to highlight that the notion generalize asymptotically periodic motions in a different way (see Theorem 3.3 for details).

This notion has many interesting properties [20, 21, 22]. For instance, the Poincare-Bendixon-type theorem can be establish.

Theorem 2.9 (Gryszka, [21])**.** *Assume that $X = \mathbb{R}^2$ and ϕ is a dynamical system, generated by the C^1 vector field. Let $x \in X$ be such that* G-AP(x) *is finite and positive. Then $\omega(x)$ is a periodic orbit.*

2.5. **Positively asymptotically periodic motions.** The following definition is by Pelczar [38]. It is a simple generalization of periodicity – we asymptotically search for a periodic-like behaviour.

Definition 2.10 (Generalization **4.**)**.** Assume that $x \in X$ is not a stationary point. We say that x is *positively asymptotically periodic* if there exists $T > 0$ such that for any $\varepsilon > 0$ there is such $s \geq 0$ that if $t \geq s$, then

$$d(\phi(t+T,x),\phi(t,x)) < \varepsilon.$$

There are two interesting properties of the above definition. The first is that it allows characterizing periodic orbits (Theorem 2.11). The second is that in the general case positively asymptotically periodic orbits can have empty limit sets (see Section 4 for the example).

Theorem 2.11 (Pelczar, [38])**.** *Let (X, d) be locally compact, ϕ be a flow on X and $\omega(x) = o(y)$ for some $x, y \in X$. The following conditions are equivalent:*

(1) A point y is stationary or periodic.
(2) A motion through x is positively asymptotically periodic.

The above definition and properties can be reformulated, by changing the time direction, to *negatively asymptotically periodic motions.* More extensive research on asymptotically periodic motions can also be found in [22, 25, 38, 40]. See [40] for further generalizations of asymptotic periodicity.

2.6. **(Weak) near periodicity.** The following definitions of positive near periodicity [1, 24, 41] and positive weak near periodicity [24, 41], introduced in the 90s, received some interest in the past.

Definition 2.12 (Generalizations **5.** and **6.**)**.** A point x is called *positively* (*negatively*) *nearly periodic* if

$$\omega(x) = D^+(x) \qquad (\alpha(x) = D^-(x)).$$

A point x is called *positively* (*negatively*) *weakly nearly periodic* if

$$\overline{o^+(x)} = J^+(x) \qquad (\overline{o^-(x)} = J^-(x)).$$

The above definitions were established for pseudo-dynamical systems, where we do not assume that ϕ is continuous. In such a case these notions generally differ (see [24] for the example), but in the continuous case, they coincide. In any case near periodicity always implies weak near periodicity.

Theorem 2.13 (see Theorem 3.2 in [24])**.** *If x is positively (negatively) nearly periodic, then it is positively (negatively) weakly nearly periodic. Furthermore, positive (negative) weak near periodicity is equivalent to the condition:*

$$\overline{o^+(x)} = J^+(x) = D^+(x) \qquad (\overline{o^-(x)} = J^-(x) = D^-(x)).$$

Theorem 2.14 (Pelczar, [41])**.** *Let x be positively weakly nearly periodic. Then x is P^+-stable and positively nearly periodic. Similarly, if x is negatively weakly nearly periodic, then it is P^--stable and negatively nearly periodic.*

From now on we will refer to positively weakly nearly periodic points and positively nearly periodic points using the common term *positively (weakly) nearly periodic*.

There is one more notion that is slightly weaker than positive (weak) near periodicity. It was mentioned by Pelczar in [41]. Here, we recall the definition with a minor refinement.

Definition 2.15 (Generalization **7**.)**.** A point x has the *property* $\mathcal{P}^+$ (*property* $\mathcal{P}^-$) if $\omega(x) = J^+(x)$ $(\alpha(x) = J^-(x))$ and either of the limit sets is non-empty.

Remark 2.16. Originally, the property $\mathcal{P}^+$ had no "non-empty" assumption to sets. This oversight could lead to some unwanted examples. Take for example the flow $\phi(t, x) = t + x$, where $x \in \mathbb{R}$. In such a case we have $\omega(x) = J^+(x) = \emptyset$ for all x's, but there is no period-like behaviour for any x.

The following straightforward connection is an immediate consequence of definitions.

Corollary 2.17. *If a point x is positively (weakly) nearly periodic, then it has the property $\mathcal{P}^+$. If it is negatively (weakly) nearly periodic, then it has the property $\mathcal{P}^-$.*

The converse of the second fact from the corollary is false as points with the property $\mathcal{P}^+$ $(\mathcal{P}^-)$ need not be P^+-stable (P^--stable). In fact, P^+-stable (P^--stable) motions need not have the property $\mathcal{P}^+$ $(\mathcal{P}^-)$ either.

Many relations between limit sets and corresponding period-like properties are established by Pelczar in [41].

2.7. **Asymptotic period-like motions.** The following concept differs from positive asymptotic periodicity. Here, we assume that the orbit tends to the periodic-like orbit. Such an approach is widely investigated by many authors. We give some references at the end of this section.

Definition 2.18 (Generalizations **8** $-$ **10**.)**.** A motion through x is called *asymptotically stationary* (resp. *asymptotically T-periodic, asymptotically almost periodic, asymptotically recurrent*) if there exists a stationary (resp., T-periodic, almost periodic, recurrent) motion through p such that

$$\lim_{t \to +\infty} d(\phi(t, x), \phi(t, p)) = 0.$$

We also call a motion through x *asymptotically periodic*, if there is $T > 0$ such that x is asymptotically T-periodic.

The following theorem is a combination of many facts from [2]. It summarize necessary and/or sufficient conditions for the asymptotic behaviour. Let P_x denote the set consisting of all points p such that $p \in \omega(x) \cap \omega(p)$ and $d(\phi(t,x), \phi(t,p)) \to 0$ as $t \to +\infty$,

Theorem 2.19.

(1) *If a motion through x is asymptotically almost periodic, then P_x consists of a single point.*
(2) *A motion through x is asymptotically stationary (resp., asymptotically T-periodic, asymptotically almost periodic), if and only if the following conditions hold:*
  (a) *x is L^+-stable,*
  (b) *$o^+(x)$ is uniformly Lyapunov stable in positive direction with respect to the set $o^+(x)$, that is, for any $\varepsilon > 0$ there is $\delta > 0$ such that if $u, v \in o^+(x)$ and $d(u,v) < \delta$, then $d(\phi(t,u), \phi(t,v)) < \varepsilon$ for any $t \geq 0$.*
  (c) *the set $\omega(x)$ is the stationary point (resp., T-periodic orbit, closure of the almost periodic point).*
(3) *A motion through x is asymptotically T-periodic if and only if the sequence $(\phi(T \cdot k, x))_{k \in \mathbb{N}}$ converges.*

Proofs of the above can be found in [12]. Many authors consider asymptotically periodic, stationary or almost periodic motions as well. Seifert [44, 45] (mainly asymptotically almost periodic motions), Bhatia [4] (characterization of asymptotic recurrence), Bhatia and Chow [5], Millionshchikov [31, 32], Nemytskii [33, 34] or Ruess with Summers [42] (asymptotically almost periodic motions and their characterization) are just a selection of many authors who investigated these notions in details. For a much greater list of references see [2].

3. The relations

In the previous section we have already provided basic connections between selected notions. All of them and the remaining ones are summarized below.

Corollary 3.1.

(1) *If x is almost periodic, then it is recurrent.*
(2) *A motion through x is positively (weakly) nearly periodic if and only if it is positively nearly periodic. Both notions imply the property $\mathcal{P}^+$.*
(3) *If x is periodic, then it need not be positive (weak) nearly periodic nor have the property $\mathcal{P}^+$.*
(4) *If x is asymptotically almost periodic, then it is asymptotically recurrent.*
(5) *If x is asymptotically T-periodic (asymptotically stationary), then it is asymptotically almost periodic.*
(6) *If x is almost periodic (resp., recurrent, T-periodic), then it is asymptotically almost periodic (resp., asymptotically recurrent, asymptotically T-periodic).*

There are only two non-trivial cases in which there is a direct implication between the notions and they are both related to asymptotic T-periodicity.

Theorem 3.2 (Proposition 3.1 in [22]). *If a motion through x is asymptotically T-periodic, then it is positively asymptotically periodic.*

Interestingly, the converse to the above theorem does not hold. This is because positively asymptotically periodic orbits need not be bounded.

The final relation is between asymptotic T-periodicity and G-asymptotic periodicity. It was established in [22].

Theorem 3.3 (Theorem 3.3 in [22]). *If x is asymptotically T-periodic, then it is G-asymptotically periodic.*

The summary of the positive relations will be provided in Section 5.

4. Counterexamples

In the following section, we provide several examples of flows that show that the remaining relations between introduced periodic-like properties need not be true. The presentation of each flow is divided into two steps. Firstly, we construct the flow. Secondly, we describe its properties and, if necessary, prove them.

The first example is a continuous approach to the famous discrete irrational rotation system on the circle.

Example 4.1. Let $\alpha \in \left[0, \frac{\pi}{2}\right)$ be such that $\frac{\alpha}{2\pi} \notin \mathbb{Q}$ and denote $I = [0,1]$. Let $T \subset \mathbb{R}^3$ be the torus $I^2/\sim$, where the relation $\sim$ glues opposite sides of the square while preserving their orientation.

We define a dynamical system ϕ on T induced by the solutions of the following differential equations on I^2.

$$\begin{cases} x' = 1, \\ y' = \alpha. \end{cases}$$

Proposition 4.2. *A dynamical system ϕ constructed in Example 4.1 contains a point that is almost periodic and positively (weakly) nearly periodic, but it is neither G-asymptotically periodic nor positively asymptotically periodic.*

Proof. We have $\omega(x) = J^+(x) = \overline{o^+(x)} = D^+(x) = T$ for any $x \in T$, so any x is positively (weakly) nearly periodic and recurrent. Furthermore, since the dynamical system is Lyapunov stable in both directions (in fact, the mapping $x \mapsto \phi(t,x)$ is an isometry for any $t > 0$), it is almost periodic.

To show that ϕ is not G-asymptotically periodic we note that since orbits are placed homogenously and the motion occurs with constant velocity, it is sufficient to consider the limit $\lim_{\varepsilon\to 0} W(x,\varepsilon)$, where x is any point of T. We divide the proof into four steps. Fix $\varepsilon_0 = 33$ and $N_0 = 11$.

Step 1. Consider a function $f\colon \mathbb{S}^1 \to \mathbb{S}^1$ (we think of $\mathbb{S}^1$ as a unit interval with identified endpoints) which is given by the formula $f(x) = x + \alpha \mod 1$. The function f is a continuous isometry, rotating points of the circle by the irrational angle α. Pick any $N_1 > N_0$ and define

$$\varepsilon_{N_1} := \min\{d(0, f^n(0)) \mid n = 1, \ldots, N_1\}.$$

It follows from the definition of ε_{N_1} that the point 0 requires more than N_1 iterations under f, before it returns to the ball $B(0, \varepsilon_{N_1})$. Let $M_1 > N_1$ be the first iteration such that $f^{M_1}(x) \in B(0, \varepsilon_{N_1})$. Define $x_1 := f^{M_1}(0)$ and, considering ε_{N_1} as an element of $\mathbb{S}^1$, let

$$\eta_1 := \frac{1}{2} \min\{\min\{d(f^i(0), B(0, \varepsilon_{N_1})) \mid i = 1, \ldots, M_1 - 1\}, \\ \min\{d(x_1, -\varepsilon_{N_1}), d(x_1, \varepsilon)\}\}.$$

Since f^{M_1} is an isometry it follows that if $d(x, 0) < \eta_1$, then $d(f^{M_1}(x), x_1) < \eta_1$. In particular, we have $f^{M_1}(x) \in B(0, \varepsilon_{N_1})$ and M_1 is the first index such that the M_1-th iteration of x returns to the ball $B(0, \varepsilon_{N_1})$.

Step 2. The set $\{f^n(0) \mid n \in \mathbb{N}\}$ (the orbit of 0) is dense in $\mathbb{S}^1$. Therefore we can find a sequence $(y_{n,1})_{n\in\mathbb{N}}$ of elements from $\mathbb{S}^1$, such that $d(0, y_{n,1}) < \eta_1$ for any $n \in \mathbb{N}$ and each $y_{n,1}$ is some iteration of a point 0. In particular, for countably many elements belonging to the orbit of 0 the return time to the ball $B(0, \varepsilon_{N_1})$ is greater or equal to N_1. Hence the upper limit of the sequence of all return times of 0 is bounded from below by N_1.

Step 3. Pick $N_2 > 2M_1$ and let

$$\varepsilon_{N_2} := \frac{1}{2} \min\{\varepsilon_{N_1}, \min\{d(0, f^n(0)) \mid n = 1, \ldots, N_2\}\}.$$

With the values N_2 and ε_{N_2}, we proceed as in Step 1 and Step 2, substituting N_2 and ε_{N_2} in place of N_1 and ε_{N_1}, respectively. By induction we can therefore obtain two sequences $(N_k)_{k\in\mathbb{N}}$ and $(\varepsilon_{N_k})_{k\in\mathbb{N}}$. For these sequences we have (as $k \to +\infty$), $\varepsilon_{N_k} \to 0$ and $N_k \to +\infty$.

Step 4. We proceed to the continuous case. Let $x = (0, 0)$. We need to prove that

$$\lim_{\varepsilon\to 0} W((0, 0), \varepsilon) = +\infty.$$

Notice that this can be done using the reasoning from Steps 1–3. Indeed, we have the following correspondence:

- iterations of f become a single segment of the orbit from the bottom side to the top side of the square I^2,
- iterations where 0 returns to $B(0, \varepsilon)$ become return times defined by the sequence $(r_i)_{i\in\mathbb{N}}$ with respect to the ball $B((0, 0), \alpha_T(\varepsilon))$.

The function α_T is a linear transformation and it only depends on the angle α.

From Step 3 it follows that $W((0, 0), \varepsilon) > N(\varepsilon)$, where $N(\varepsilon)$ is the number related to ε as described in Step 3. Therefore we obtain

$$\lim_{\varepsilon\to 0} W((0, 0), \varepsilon) \geq \lim_{\varepsilon\to 0} N(\varepsilon) = +\infty.$$

Hence G-AP$(x) = +\infty$.

It is remaining to prove the orbit of $x = (0, 0)$ is not positively asymptotically periodic. Indeed, given $T > 0$ we would have to have $d(\phi(T, x), x) < \varepsilon$, regardless of $\varepsilon > 0$. We can skip $s > 0$ and $t \geq s$ part from the definition because the torus is homogeneously filled with orbits, so any point of reference is the same. Since in the inequality ε is arbitrary, $\phi(T, x) = x$, so x is periodic. A contradiction. □

Remark 4.3. Any almost periodic motion is asymptotically almost periodic and asymptotically recurrent. Thus, any point x from Example 4.1 is asymptotically almost periodic and asymptotically recurrent. Furthermore, it follows that none of these and positive (weak) near periodicity imply G-asymptotic periodicity or positive asymptotic periodicity. Furthermore, there are no points in this example that are periodic or asymptotically periodic.

Let us recall an important example of a system that possesses the ω-limit set equal to two-dimensional torus and the orbit is not contained in that torus. One of the possible constructions of that system has been provided in [20].

Example 4.4. We sketch the construction of the flow; detailed construction is provided in [20, Example 2.5]. We let $\mathbb{T} \subset \mathbb{R}^3$ denote a surface of a torus and we fill $\mathbb{T}$ homogeneously with orbits that intersect each latitudinal circle at constant angle α. The flow on $\mathbb{T}$ is defined to have each orbit "moving" at constant velocity η.

The next step is to construct a family of tori $(\mathbb{T}_k)_{k\in\mathbb{N}}$ such that $\mathbb{T}_{k+1}$ is inside $\mathbb{T}_k$, $\mathbb{T}$ is inside all $\mathbb{T}_k$'s ($\mathbb{T}$ is a "limit" of $\mathbb{T}_k$'s) and each torus is partially filled with a piece of the orbit of the same point x. The entire orbit of x is built in such a way it covers $\mathbb{T}_k$ with some accuracy, namely the condition

$$\forall\, x' \in \mathbb{T}_k \quad d(x', \phi([0, t_k], x_k)) \leq \frac{1}{k}$$

is satisfied for some $x_k \in o(x) \cap \mathbb{T}_k$ and $t_k > 0$. Moreover, the piece of $o(x)$ which is in $\mathbb{T}_k$ is a segment wrapped round $\mathbb{T}_k$ and has an angle $\alpha + \alpha_k$ so that $\tan(\alpha + \alpha_k) \notin \mathbb{Q}$ and $\alpha_k \to 0$. Once $\mathbb{T}_k$ is filled with desired accuracy, the orbit "jumps" to $\mathbb{T}_{k+1}$ and starts filling it from the point x_{k+1}. Finally, the point x moves across its orbit at constant velocity η.

The phase of space X is composed of $\mathbb{T}$, the set $o(x)$ described as above, and $\alpha(x)$. By the construction, $\omega(x) = \mathbb{T}$. Also, depending on the choice of the angle α we can manipulate with the dynamics of the limit set, from "all orbits in $\mathbb{T}$ are dense in $\mathbb{T}$" to "all orbits in $\mathbb{T}$ are periodic".

Proposition 4.5. *A dynamical system ϕ constructed in Example 4.4 contains a point that has the property $\mathcal{P}^+$ and satisfies none of the remaining conditions.*

Proof. The point we chose is any point from the distinguished orbit of x. Let us call it z.

From the construction we obtain $\omega(z) = J^+(z) = \mathbb{T}$, so z has the property $\mathcal{P}^+$.

The proof that x is not G-asymptotically periodic or positively asymptotically periodic is similar to the proof of Proposition 4.2. It is clear that $\overline{o^+(z)}\setminus J^+(z) \neq \emptyset$, so z is not positively (weakly) nearly periodic.

To simplify the reasoning, let us now set α so that all orbits contained in $\mathbb{T}$ are periodic. For such a case we know that P_z (see page 113 for the definition) is a single point which we call w (see Theorem 2.19), therefore w has to be periodic. By the definition, $d(\phi(t,w),\phi(t,z)) \to 0$, so the limit set of z cannot be the entire torus. A contradiction.

Almost periodic and recurrent orbits coincide in $\mathbb{T}$, so the existence of recurrent point satisfying the definition of asymptotically recurrent motion would imply asymptotic almost periodicity. This means z cannot be asymptotically recurrent. □

The next example shows that positive asymptotic periodicity can be detected in a very peculiar system. The flow was presented in [22].

Example 4.6. Consider a dynamical system ϕ on the plane with the point $(0,0)$ removed. It is defined to have its trajectories coinciding with the curves in polar coordinates

$$r(t) = \begin{cases} \frac{1}{t} & \text{for } t \geq e, \\ e^{t-e} & \text{for } t < e, \end{cases} \qquad \vartheta(t) = t + c.$$

For any point $x \neq (0,0)$, the orbit intersects the same half of the line $Ax+By = 0$ (A and B are fixed) in equal time gaps, equal to 2π.

Proposition 4.7 (Proposition 3.5 in [22])**.** *A dynamical system ϕ constructed in Example 4.6 contains a point that is positively asymptotically periodic.*

The latter example shows something very surprising - the system for which every orbit is unbounded, yet at the same time each orbit satisfies some sort of periodicity-like behaviour. It is likely that Pelczar was aware of the property described in Proposition 4.7. However, we do not know any reliable source of a similar example. The presence of the orbit with an empty limit set cannot be eliminated with some minor adjustments to the phase space. For instance, for the case of G-asymptotically periodic motions, if the G-asymptotic period is finite, then it is sufficient to assume that X is a proper metric space. Then the limit set $\omega(x)$ must be a non-empty, compact, and connected set [21]. We can modify Example 4.6 so that $(0,0)$ is a stationary point and the qualitative behaviour of the remaining orbits is maintained. Then the phase space is proper, which is still insufficient to ensure the limit set is not empty.

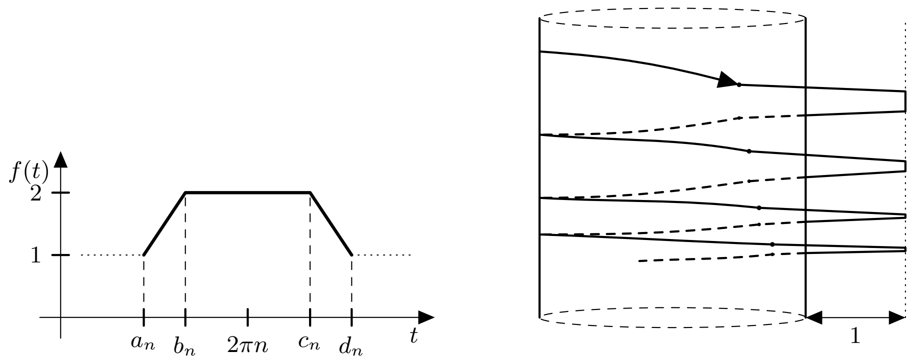
Remark 4.8. By the construction of the dynamical system in Example 4.6 we can also conclude that positively asymptotically periodic motions need not satisfy any of the remaining definitions.

The next example comes from [22] and its purpose was to show that G-asymptotically points need not be positively asymptotically periodic (recall that by Theorem 3.3 the converse is true).

Example 4.9 (Example 3.6 in [22])**.** Let us introduce the following notation:

$$a_n := 2\pi n - \tfrac{1}{n}, \qquad b_n := 2\pi n - \tfrac{1}{n+1},$$
$$c_n := 2\pi n + \tfrac{1}{n+1}, \qquad d_n := 2\pi n + \tfrac{1}{n},$$
$$e_n := \pi(2n+1), \qquad A_n := [a_n, d_n].$$

Define a function $f\colon \mathbb{R} \to \mathbb{R}$ by the following rule. If t is not contained in any interval A_n, set $f(t) = 1$. For any n, define f on A_n to be a piecewise linear function joining points $(a_n, 1), (b_n, 2), (c_n, 2), (d_n, 1)$. We will call such piecewise parts hills (see Figure 1a).



(A) The hill of f on the interval A_n. (B) The sketch of the orbit.

FIGURE 1. The flow

The function f we have just defined has the following key property. For any n we have $f(b_n) = 2$ and $f(b_n + 2\pi) = f(a_{n+1}) = 1$.

Define the curve in $\mathbb{R}^3$ using the following description:

$$\Xi : \mathbb{R} \ni t \mapsto (f(t)\cos t, f(t)\sin t, e^{-t}) \in \mathbb{R}^3.$$

The parametric representation has the following intuition. The graph of f is wrapped around the upper half of the cylinder $x^2+y^2 = 1$ with the coils converging to the plane $z = 0$. The hills stick away from the cylinder (Figure 1b).

We can now define a dynamical system ϕ so that the curve Ξ is its only orbit. The motion occurs towards the plane $z = 0$ with a constant angular velocity equal to 1.

Proposition 4.10 (Proposition 3.7 in [22])**.** *The dynamical system ϕ constructed in Example 4.9 contains a point that is G-asymptotically periodic but neither is positively asymptotically periodic nor asymptotically T-periodic for any value of T.*

Remark 4.11. It is clear that the dynamical system ϕ constructed in Example 4.9 does not satisfy any of the remaining definitions. The ω-limit set is empty. It is unknown whether a similar example with the non-empty limit set can be provided.

The next example is (eventually) used to show that orbits can be periodic but they do not have to satisfy the property $\mathcal{P}^+$.

Example 4.12. Consider a dynamical system ϕ on the set $X := \overline{B}((0,0),2)$ equipped with standard metric inherited from $\mathbb{R}^2$. That system is induced by the system of differential equations:

$$\begin{cases} r' = r(1-r)^2(2-r), \\ \vartheta' = 1, \end{cases}$$

Figure 2 shows the sketch of that system.

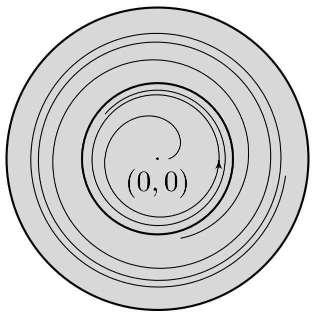


FIGURE 2. A sketch of a system defined in Example 4.12.

Proposition 4.13. *A dynamical system ϕ constructed in Example 4.12 contains a point that is G-asymptotically periodic, positively asymptotically periodic, asymptotically almost periodic, asymptotically recurrent, asymptotically periodic but it is not recurrent, almost periodic, positively (weakly) nearly periodic or periodic.*

Proof. Since the angular velocity of the motion is constant, it follows that

$$\text{G-AP}(x) = 2\pi$$

for any point $x \in X$ such that $0 < \|x\| < 1$ (one can also pick any point x such that $1 < \|x\| < 2$). Such a point is positively asymptotically periodic because the ω-limit set is a periodic orbit (see Theorem 2.11). The latter also implies that x is asymptotically periodic, so it is asymptotically almost periodic and asymptotically recurrent.

On the other hand, such x cannot be recurrent. It is clear that $\omega(x)$ is a proper subset of $D^+(x)$, hence x cannot be positively (weakly) nearly periodic. $\square$

Proposition 4.14. *A dynamical system ϕ defined in Example 4.12 contains a point that is periodic, but it is neither positively (weakly) nearly periodic nor has the property $\mathcal{P}^+$.*

Proof. A point $x = (1, 0)$ is periodic. On the other hand we have:

$$\begin{aligned} \omega(x) &= \{x \mid \|x\| = 1\}, \\ J^+(x) &= \{x \mid \|x\| = 1 \vee \|x\| = 2\}, \\ D^+(x) &= \{x \mid 1 \leq \|x\| \leq 2\}, \end{aligned}$$

hence neither of the remaining properties hold. □

Remark 4.15. Periodic motions are a special case of almost periodic, recurrent, asymptotically almost periodic and asymptotically recurrent motions. Therefore it follows from Proposition 4.14 that such motions need not be positively (weakly) nearly periodic or have the property $\mathcal{P}^+$

Example 4.16. Consider dynamical system defined in Example 4.1. We modify it slightly by setting $(0, 0)$ to be a stationary point. This forces the orbit passing through that point to fall into three different orbits, one of them having its α-limit set equal to $(0, 0)$, the other having its ω-limit set equal to $(0, 0)$, and the stationary point. The presence of the stationary point now forces small changes of the behaviour of all near orbits. Nevertheless, all but one (the stationary one) orbits of the new system are dense in the torus. We have thus defined a new dynamical system ϕ on T^2.

Proposition 4.17. *A dynamical system ϕ defined in Example 4.16 contains a point that is positively (weakly) nearly periodic but not recurrent.*

Proof. Let x be any point that is not stationary and its ω-limit set is not $(0, 0)$. Since the orbit of it is dense, we have $\omega(x) = J^+(x) = \overline{o^+(x)} = D^+(x) = T$. It follows that x is positively (weakly) nearly periodic.

The set $\overline{o(x)}$ is compact, but it is not minimal, thus x cannot be recurrent. □

Remark 4.18. It follows from Example 4.16 that nearly periodic points need not be almost periodic, asymptotically almost periodic or asymptotically recurrent.

The next example is a Denjoy's flow, which we use to show the lack of relation between recurrence and almost periodicity (also in asymptotic case).

Recall that if $f : X \to X$ is a continuous mapping, then we can define

$$X^* := \{(r, x) \colon 0 \leq s \leq 1, x \in X\}/((1, x) \sim (0, f(x))),$$

the *suspension* of X with respect to f. Given that we define

$$\phi^t \colon X^* \to X^*; \quad (s, x) \to (\tau, f^k(x)),$$

where $k = [t + s]$ (the integer part of the number $t + s$) and $\tau = t + s - k \in [0, 1)$. Then ϕ^t is called the *suspension flow* of f. It satisfies the identity $\phi^k(0, x) = (0, f^k(x))$ for $k \in \mathbb{N}$.

Example 4.19 (Denjoy's flow)**.** Denote by $S(\lambda)$ the circle obtained by identifying the endpoints of the interval $[0,\lambda] \subset \mathbb{R}$. Consider the rotation on $S(2\pi)$ by an irrational number α:

$$R\colon x \mapsto x+\alpha \qquad \mod 2\pi.$$

Take arbitrary point $x_0 \in S(2\pi)$ and consider its two sided orbit $\mathcal{O}(x_0) = \{R^n(x_0) \mid n \in \mathbb{Z}\}$.

Take a sequence $(a_n)_{n\in\mathbb{Z}}$ of positive numbers such that $a = \sum_{n\in\mathbb{Z}} a_n > 0$ is finite. With each term of that sequence we associate an open interval I_n of the length a_n. The series of such intervals in now located on the circle $S(2\pi+a)$ in such a way they are pairwise disjoint and their mutual arrangement corresponds with the mutual arrangement of the points x_n on $S(2\pi)$.

The complement Ω of the union of all intervals I_n is a nowhere dense set. Consequently, Ω is homeomorphic to the Cantor set. Denote by Ω_1 the subset of Ω consisting of all endpoints of I_n's. Then take $\Omega_2 = \Omega \setminus \Omega_1$.

Since the arrangement of the intervals agrees with the arrangement of the points from $\mathcal{O}(x_0)$, there exists an orientation preserving and continuous mapping $h\colon S(2\pi+a) \to S(2\pi)$ that transforms the closed interval $\overline{I_n}$ into x_n for all $n \in \mathbb{Z}$, and is a bijection on the set Ω.

We can define the diffeomorphism $f\colon S(2\pi+a) \to S(2\pi+a)$, so that $R\circ h = h\circ f$. We outline the basic idea; rigorous definitions and proof of all properties can be found in [2, 47].

The way we can find f is by constructing the derivative of f; the function $F = f'$ defined on $S(2\pi+a)$ and taking values in $\mathbb{R}$. We assume that $F=1$ on the set Ω and if $x \in I_n = (\alpha_n, \beta_n)$, then we can choose the value of $F(x)$ (F can be a polynomial of degree 2) so that F is continuous at the points of Ω, and eventually F is continuous on $S(2\pi+a)$. Then, F can be used to define a natural lift mapping of F, which in turn becomes a derivative of F

Since F is continuous and $F' = f$, f is a C^1 diffeomorphism of the circle $S(2\pi+a)$. It is called the Denjoy diffeomorphism. Furthermore, for all $x \in S(2\pi+a)$, the ω-limit set (which is the union of all possible sets $\omega(x)$) coincides with Ω. Consequently, the Cantor set Ω is the limit set of the Denjoy diffeomorphism.

The final step is to consider the suspension flow ϕ^t of f. This gives the flow on the torus $\mathbb{T}$ such that each orbit of ϕ^t in the limit set $\mathcal{D}$ (which is the set of all orbits passing through points of Ω) is dense in $\mathcal{D}$. Note that the section of $\mathcal{D}$ is homeomorphic to the product of Ω and a unit interval.

Proposition 4.20. *The Denjoy's flow contains a point x that is recurrent, but it is not almost periodic.*

Proof. The orbit of any $x \in \mathcal{D}$ is minimal and compact, so it is recurrent.

On the other hand, we have, for some integer $m > 0$,

$$\phi(m, [\alpha_k, \beta_k]) = [\alpha_0, \beta_0],$$

where the length of the interval $[\alpha_k, \beta_k]$ can be arbitrary small provided $k < 0$ and $|k|$ is much greater than 1. It follows that the flow ϕ is Lyapunov unstable, and thus no orbit from the set $\mathcal{D}$ can be almost periodic. □

Remark 4.21. The Denjoy's flow is used to provide an example of recurrent but not almost periodic orbits. It is worth noting that the Denjoy diffeomorphism cannot be C^2 (the above example is C^1, the C^2 map does not admit minimal sets of Cantor type (see [13])). Furthermore, the Denjoy's flow contains no periodic orbits, so the C^1 diffeomorphism was later used to solve the Seifert conjecture. It states that every non-singular and continuous vector field on the 3-sphere has a closed orbit (see [46]). The conjecture was proven to be wrong by Paul Schweizer, who provided a C^1 counterexample (see [43]). Later, many authors provided smoother examples, such as C^∞ one in [28].

Example 4.22. Consider the Denjoy's flow $\phi\colon \mathbb{R} \times \mathbb{T} \to \mathbb{T}$ from Example 4.19. We extend the phase space by building one additional orbit so that its ω-limit set is the set $\mathcal{D}$. To do that we can take any orbit from $\mathcal{D}$ (say, the orbit of y) and lift a copy of it radially so that every point of the copy is above the torus. Such an orbit, the orbit of z, can be build so that it extends the ϕ and the following conditions are satisfied:

(1) $d(\phi(t, y), \phi(t, z)) = \frac{1}{t}$ as t is sufficiently large; this distance is not achieved by any other point from $\mathbb{T}$.
(2) $\omega(z) = \mathcal{D}$.

The phase space is $X \cup o(z)$ and the ϕ is well-defined by the condition 1.

Proposition 4.23. *A dynamical system ϕ constructed in Example 4.22 contains a point x that is asymptotically almost recurrent but not asymptotically almost periodic.*

Proof. The orbit of z cannot be asymptotically almost periodic, which follows directly from Theorem 2.19. On the other hand, the choice of z with respect to y guarantees that it is asymptotically recurrent. □

5. The Summary

We summarize results from Section 3 and Section 4 in one theorem.

Theorem 5.1. *In the following diagram, if there is a directed path from one node to the other, then the relation between notions holds. If there is no directed path, the relation does not hold.*

$$\begin{tikzcd}
 & \substack{\text{Almost Periodic}\\ \text{(Gen. 2)}} \arrow[r] \arrow[d] & \substack{\text{Recurrent}\\ \text{(Gen. 1)}} \arrow[d] \\
\text{Periodic} \arrow[ur] \arrow[r] \arrow[dr] & \substack{\text{Asymptotically}\\ \text{Almost Periodic}\\ \text{(Gen. 9)}} \arrow[r] & \substack{\text{Asymptotically}\\ \text{Recurrent (Gen. 10)}} \\
 & \substack{\text{Asymptotically}\\ \text{Periodic (Gen. 8)}} \arrow[u] \arrow[r] \arrow[dr] & \substack{\text{Positively}\\ \text{Asymptotically}\\ \text{Periodic (Gen. 4)}} \\
 & & \substack{\text{G-Asymptotically Periodic}\\ \text{(Gen. 3)}}
\end{tikzcd}$$

$$\begin{tikzcd}
\substack{\text{Positively Nearly}\\ \text{Periodic (Gen. 5)}} \arrow[r, leftrightarrow] & \substack{\text{Positively}\\ \text{Weakly Nearly}\\ \text{Periodic (Gen. 6)}} \arrow[r] & \substack{\text{Property } \mathcal{P}^{+}\\ \text{(Gen. 7)}}
\end{tikzcd}$$

ACKNOWLEDGEMENTS

I thank the referee for many suggestions that helped to improve this article.

References

[1] K. H. Ahmad, S. Elaydi, *Nearly periodic dynamical systems*, Appl. Anal. **27** (1998), 243–251.

[2] S. Aranson et al, *Introduction to the Qualitative Theory of Dynamical Systems on Sufraces*, Transactions of Mathematical Monographs, Vol. 153, 1996.

[3] A. S. Besicovitch, *Almost Periodic Functions*, Dover Publ.: New York, NY, USA, 1954.

[4] N. P. Bhatia, *Asymptotic recurrence and dynamical flow near a compact minimal set*, in Seminar on Differential Equations and Dynamical Systems II (Univ. Maryland, College Park, Md., 1969), vol. 144 of Lecture Notes in Mathematics, Springer, Berlin, Germany, (1970), 22—29.

[5] N. P. Bhatia, S. N. Chow, *Weak attraction, minimality, recurrence, and almost periodicity in semi-systems*, Funkcial. Ekvac. **15** (1972), 39-–59.

[6] N. P. Bhatia, G. P. Szegö, *Stability Theory of Dynamical Systems*, Springer - Verlang, Berling - Heidelberg - New York, 1970.

[7] G. D. Birkhoff, *Dynamical Systems*, Amer. Math. Soc. Colloquium Publications, vol. 9, New York, 1927.

[8] S. Bochner, *Beiträge zur Theorie der fastperiodischen Funktionen*, I. Math. Ann. **96** (1926), 119–147.

[9] H. Bohr, *Zur Theorie der Fastperiodischen Funktioner I*, Acta Math. **45** (1924), 29–127.

[10] H. Bohr, *Zur Theorie der Fastperiodischen Funktioner II*, Acta Math. **46** (1925), 101–204.

[11] H. Bohr, *Zur Theorie der Fastperiodischen Funktioner III*, Acta Math. **47** (1926), 237–281.

[12] D. N. Cheban, *Asymptotically Almost Periodic Solutions of Differential Equations*, Hindawi Publishing Corporation, Cairo, 2009.

[13] E. Coddington, N. Levinson, *Theory of Ordianry Differential Equations*, McGrraw-Hill, New York, 1955.

[14] L. G. Deysach, G. R. Sell, *On the existance of almost periodic motions*, Michigan Math. J.**12** (1965), 87–95.
[15] T. Diagana, *Almost Automorphic Type and Almost Periodic Type Functions in Abstract Spaces*, Springer: New York, NY, USA, 2013
[16] R. Ellis, M. Nerurkar, *Weakly almost periodic flows*, Trans. Amer. Math. Soc. **313** (1989), 103–119.
[17] A. M. Fink, *Almost Periodic Differential Equations*, Springer: Berlin, Germany, 1974.
[18] P. Glendinning, *Stability, Instability and Chaos: An Introduction to the Theory of Nonlinear Differential Equations*, Cambridge University Press, 1994.
[19] W. H. Gottschalk, G. A. Hedlund, *Topological Dynamics*, American Mathematical Society Colloquium Publications, vol. 36. Providence, American Mathematical Society, 1955.
[20] K. Gryszka, *Asymptotic period in dynamical systems in metric spaces*, Colloq. Math. **139** (2015), 245–257.
[21] K. Gryszka, *Lagrange stability and asymptotic perions*, Topology Appl. **204** (2016), 168–174.
[22] K. Gryszka, *On Asymptoically periodic-like motions in flows*, Ann. Univ. Paedagog. Crac. Stud. Math **17** (2018), 45–57.
[23] G. M. N. Guérékata, *Almost Automorphic and Almost Periodic Functions in Abstract Spaces*, Kluwer Acad. Publ.: Dordrecht, The Netherlands, 2001.
[24] L. Keon-Hee, *Weak near periodicity in dynamical systems*, Nonlinear Anal., Theory, Methods and Appl. **15** (1990), 399–404.
[25] J. Kłapyta, *Asymptotically periodic motions in pseudo-processes*, Univ. Iagel. Acta Math. **33** (1996), 73–86.
[26] M. Kostić, *Almost Periodic and Almost Automorphic Type Solutions to Integro-Differential Equations*, W. de Gruyter: Berlin, Germany, 2019
[27] M. Kostić, W.-S. Du, *Generalized almost periodicity in Lebesgue spaces with variable exponents*, Mathematics. 2020; **8** (6):928.
[28] K. Kuperberg, *A smooth counterexample to the Seifert conjecture*, Ann. of Math. **140** (1994), 723–732.
[29] M. Levitan, *Almost Periodic Functions*, G.I.T.T.L.: Moscow, Russia, 1959. (In Russian)
[30] A. A. Markov, *Stabilität im Liapunoffschen Sinne und Fastperiodizität*, Math. Z. **36** (1933) 708–738.
[31] V. M. Millionshchikov, *Recurrent and almost periodic limit trajectories of non-autonomous systems of differential equations* (Russian), Doklady Academy of Sciences of USSR **161**, no. 1 (1965), 43-–44.
[32] V. M. Millionshchikov, *Recurrent and almost periodic limit solutions of non-autonomous systems* (Russian), Differential'nye Uravneniya, **4**, no. 9 (1968), 1555-–1559.
[33] V. V. Nemytskii, *On some methods of qualitative investigation in the large for manydimensional autonomous systems* (Russian), Works of Moscow Mathematical Society **5** (1956), 455—482.
[34] V. V. Nemytskii, *Oscillatory regimes of higher-dimensional dynamical systems*, in: Qualitative Methods in the Theory of Non-Linear Vibrations (Proceedings of the International Symposium on Non-linear Vibrations, Vol. II, 1961), Akademii Nauk Ukrainian SSR, Kiev, Ukraine (1963) 308-–314
[35] V. V. Nemytskii, V. V. Stephanov, *Qualitative Theory of Differential Equations*, Princeton University Press, 1960.
[36] J. von Neumann, *Almost periodic functions in a group*, Trans. Amer. Math. Soc., **34** (1934), 445–492.
[37] A. Pelczar, *Semi-stability of motions and regular dependence of limit sets on points in general semi-systems*, Ann. Polon. Math. **42** (1983), 263–282.

[38] A. Pelczar, *Asymptotically periodic motions*, Bull. Pol. Acad. Sci. Math. **33** (1985), 313–319.

[39] A. Pelczar, *Introduction to the Theory of Differential Equations. Part II: Elements of Qualitative Theory of Differential Equations* (Polish), Biblioteka Matematyczna PWN, vol. 67, Warsaw, 1989.

[40] A. Pelczar, *Remarks on some generalizations of asymptotic periodicity in dynamical systems on metric spaces*, Ann. Polon. Math. **51** (1990), 259–268.

[41] A. Pelczar, *Remarks on some reccurency type and periodicity-like conditions in dynamical systems*, Univ. Iagel. Acta Math. **35** (1997), 95–113.

[42] W. M. Ruess and W. H. Summers, *Positive limit sets consisting of a single periodic motion*, J. Differential Equations **71** (1988), 261—269.

[43] P. A. Schweizer, *Counterexamples to the Seifert conjecture and opening closed leaves of foliations*, Ann. of Math. **100** (1974), 386–400.

[44] G. Seifert, *Almost periodicity in semiflows*, Proc. Amer. Math. Soc. **123** (1995), 2895—2899.

[45] G. Seifert, *Nonlinear evolution equations with almost periodic time dependence*, SIAM J. Math. Anal. **18** (1987), 387—392.

[46] H. Seifert, *Closed integral curves in* 3*-space and isotopic two-dimensional deformations*, Proc. Amer. Math. Soc. **1** (1950), 287–302.

[47] D. Tongren, *Approaches to the Qualitative Theory of Ordinary Differential Equations. Dynamical Systems and Nonlinear Oscillations*, World Scientific Publishing Co. Pte. Ltd., Singapore, 2007.

[48] S. Zaidman, *Almost-Periodic Functions in Abstract Spaces*, Pitman Research Notes in Math; Pitman: Boston, MA, USA, 1985; Volume 126.

Institute of Mathematics, Pedagogical University of Kraków, Poland.
karol.gryszka@up.krakow.pl

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON THE MONOID OF COFINITE PARTIAL ISOMETRIES OF $\mathbb{N}$ WITH A BOUNDED FINITE NOISE

OLEG GUTIK & PAVLO KHYLYNSKYI

Abstract. In the paper we study algebraic properties of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ of cofinite partial isometries of the set of positive integers $\mathbb{N}$ with the bounded finite noise j. For the monoids $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ we prove counterparts of some classical results of Eberhart and Selden describing the closure of the bicyclic semigroup in a locally compact topological inverse semigroup. In particular we show that for any positive integer j every Hausdorff shift-continuous topology τ on $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is discrete and if $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a proper dense subsemigroup of a Hausdorff semitopological semigroup S, then $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a closed ideal of S, and moreover if S is a topological inverse semigroup then $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a topological group. Also we describe the algebraic and topological structure of the closure of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ in a locally compact topological inverse semigroup.

Keywords: partial isometry, inverse semigroup, partial bijection, bicyclic monoid, closure, locally compact, topological inverse semigroup

MSC2020: 20M18, 20M20, 20M30, 22A15, 54A10, 54D45

Received 29 April 2021; revised 19 July 2021; accepted 29 July 2021

1. Introduction and preliminaries

In this paper we shall follow the terminology of [9, 12, 27, 29]. We shall denote the first infinite cardinal by ω and the cardinality of a set A by $|A|$. By $\mathrm{cl}_X(A)$ we denote the closure of subset A in a topological space X.

A semigroup S is called *inverse* if for any element $x \in S$ there exists a unique $x^{-1} \in S$ such that $xx^{-1}x = x$ and $x^{-1}xx^{-1} = x^{-1}$. The element x^{-1} is called the *inverse of* $x \in S$. If S is an inverse semigroup, then the function $\mathrm{inv}\colon S \to S$ which assigns to every element x of S its inverse element x^{-1} is called the *inversion*.

 https://doi.org/10.2478/9788366675360-010.

If S is a semigroup, then we shall denote the subset of all idempotents in S by $E(S)$. If S is an inverse semigroup, then $E(S)$ is closed under multiplication and we shall refer to $E(S)$ as a *band* (or the *band of* S).

If S is an inverse semigroup then the semigroup operation on S determines the following partial order $\preccurlyeq$ on S: $s \preccurlyeq t$ if and only if there exists $e \in E(S)$ such that $s = te$. This order is called the *natural partial order* on S and it induces the *natural partial order* on the semilattice $E(S)$ [30].

An inverse subsemigroup T of an inverse semigroup S is called *full* if $E(T) = E(S)$.

A congruence $\mathfrak{C}$ on a semigroup S is called a *group congruence* if the quotient semigroup $S/\mathfrak{C}$ is a group. Any inverse semigroup S admits the *minimum group congruence* $\mathfrak{C}_{\mathbf{mg}}$:

$$a\mathfrak{C}_{\mathbf{mg}}b \quad \text{if and only if} \quad \text{there exists} \quad e \in E(S) \quad \text{such that} \quad ea = eb.$$

Also, we say that a semigroup homomorphism $\mathfrak{h}\colon S \to T$ is a *group homomorphism* if the image $(S)\mathfrak{h}$ is a group, and $\mathfrak{h}\colon S \to T$ is *trivial* if it is either an isomorphism or annihilating.

The bicyclic monoid $\mathscr{C}(p,q)$ is the semigroup with the identity 1 generated by two elements p and q subjected only to the condition $pq = 1$. The semigroup operation on $\mathscr{C}(p,q)$ is determined as follows:

$$q^k p^l \cdot q^m p^n = q^{k+m-\min\{l,m\}} p^{l+n-\min\{l,m\}}.$$

It is well known that the bicyclic monoid $\mathscr{C}(p,q)$ is a bisimple (and hence simple) combinatorial E-unitary inverse semigroup and every non-trivial congruence on $\mathscr{C}(p,q)$ is a group congruence [12].

If $\alpha\colon X \rightharpoonup Y$ is a partial map, then we shall denote the domain and the range of α by $\operatorname{dom}\alpha$ and $\operatorname{ran}\alpha$, respectively. A partial map $\alpha\colon X \rightharpoonup Y$ is called *cofinite* if both sets $X \setminus \operatorname{dom}\alpha$ and $Y \setminus \operatorname{ran}\alpha$ are finite.

Let $\mathscr{I}_\lambda$ denote the set of all partial one-to-one transformations of a non-zero cardinal λ together with the following semigroup operation:

$$x(\alpha\beta) = (x\alpha)\beta \quad \text{if} \quad x \in \operatorname{dom}(\alpha\beta) = \{y \in \operatorname{dom}\alpha \colon y\alpha \in \operatorname{dom}\beta\}, \quad \text{for } \alpha, \beta \in \mathscr{I}_\lambda.$$

The semigroup $\mathscr{I}_\lambda$ is called the *symmetric inverse* (*monoid*) *semigroup* over the cardinal λ (see [12]). The symmetric inverse semigroup was introduced by Wagner [30] and it plays a major role in the theory of semigroups. By $\mathscr{I}_\lambda^{\mathrm{cf}}$ is denoted a subsemigroup of injective partial selfmaps of λ with cofinite domains and ranges in $\mathscr{I}_\lambda$. Obviously, $\mathscr{I}_\lambda^{\mathrm{cf}}$ is an inverse submonoid of the semigroup $\mathscr{I}_\lambda$. The semigroup $\mathscr{I}_\lambda^{\mathrm{cf}}$ is called the *monoid of injective partial cofinite selfmaps* of λ [20].

A partial transformation $\alpha\colon (X,d) \rightharpoonup (X,d)$ of a metric space (X,d) is called *isometric* or a *partial isometry*, if $d(x\alpha, y\alpha) = d(x,y)$ for all $x, y \in \operatorname{dom}\alpha$. It is obvious that the composition of two partial isometries of a metric space (X,d) is a partial isometry, and the converse partial map to a partial isometry is a partial isometry, too. Hence the set of partial isometries of a metric space (X,d) with the operation of composition of partial isometries is an inverse submonoid of the symmetric inverse monoid over the cardinal $|X|$. Also, it is obvious that the set of partial cofinite isometries of a metric space (X, d) with the operation the composition of partial isometries is an inverse submonoid of the monoid of injective partial cofinite selfmaps of the cardinal $|X|$.

We endow the sets $\mathbb{N}$ and $\mathbb{Z}$ with the standard linear order.

The semigroup $\mathbf{ID}_\infty$ of all partial cofinite isometries of the set of integers $\mathbb{Z}$ with the usual metric $d(n, m) = |n - m|$, $n, m \in \mathbb{Z}$, was studied in the papers [7, 8, 21].

Let $\mathbf{I}\mathbb{N}_\infty$ be the set of all partial cofinite isometries of the set of positive integers $\mathbb{N}$ with the usual metric $d(n, m) = |n - m|$, $n, m \in \mathbb{N}$. Then $\mathbf{I}\mathbb{N}_\infty$ with the operation of composition of partial isometries is an inverse submonoid of $\mathscr{I}_\omega$. The semigroup $\mathbf{I}\mathbb{N}_\infty$ of all partial cofinite isometries of positive integers is studied in [22]. There we described the Green relations on the semigroup $\mathbf{I}\mathbb{N}_\infty$, its band and proved that $\mathbf{I}\mathbb{N}_\infty$ is a simple E-unitary F-inverse semigroup. Also in [22], the least group congruence $\mathfrak{C}_{\mathbf{mg}}$ on $\mathbf{I}\mathbb{N}_\infty$ is described and there it is proved that the quotient-semigroup $\mathbf{I}\mathbb{N}_\infty/\mathfrak{C}_{\mathbf{mg}}$ is isomorphic to the additive group of integers $\mathbb{Z}(+)$. An example of a non-group congruence on the semigroup $\mathbf{I}\mathbb{N}_\infty$ is presented. Also it is proved that a congruence on the semigroup $\mathbf{I}\mathbb{N}_\infty$ is a group congruence if and only if its restriction onto an isomorphic copy of the bicyclic semigroup in $\mathbf{I}\mathbb{N}_\infty$ is a group congruence. In [24] it was shown that the monoid $\mathbf{I}\mathbb{N}_\infty$ does not embed isomorphically into the semigroup $\mathbf{ID}_\infty$. Moreover every non-annihilating homomorphism $\mathfrak{h}\colon \mathbf{I}\mathbb{N}_\infty \to \mathbf{ID}_\infty$ has the following property: the image $(\mathbf{I}\mathbb{N}_\infty)\mathfrak{h}$ is isomorphic either to $\mathbb{Z}_2$ or to $\mathbb{Z}(+)$. Also it is proved that $\mathbf{I}\mathbb{N}_\infty$ does not have a finite set of generators, and moreover it does not contain a minimal generating set.

Later by $\mathbb{I}$ we denote the unit elements of $\mathbf{I}\mathbb{N}_\infty$.

Remark 1.1. We observe that the bicyclic semigroup is isomorphic to the semigroup $\mathscr{C}_\mathbb{N}$ which is generated by partial transformations α and β of the set of positive integers $\mathbb{N}$, defined as follows:

$$\operatorname{dom}\alpha = \mathbb{N}, \qquad \operatorname{ran}\alpha = \mathbb{N}\setminus\{1\}, \qquad (n)\alpha = n+1$$

and

$$\operatorname{dom}\beta = \mathbb{N}\setminus\{1\}, \qquad \operatorname{ran}\beta = \mathbb{N}, \qquad (n)\beta = n-1$$

(see Exercise IV.1.11(ii) in [28]). It is obvious that $\mathbb{I} = \alpha\beta$ and $\mathscr{C}_\mathbb{N}$ is a submonoid of $\mathbf{I}\mathbb{N}_\infty$.

The semigroup of monotone (order preserving) injective partial transformations φ of $\mathbb{N}$ such that the sets $\mathbb{N}\setminus\operatorname{dom}\varphi$ and $\mathbb{N}\setminus\operatorname{ran}\varphi$ are finite was introduced in [18] and there it was denoted by $\mathscr{I}_\infty^{\nearrow}(\mathbb{N})$. Obviously, $\mathscr{I}_\infty^{\nearrow}(\mathbb{N})$ is an inverse subsemigroup of the semigroup $\mathscr{I}_\omega$. The semigroup $\mathscr{I}_\infty^{\nearrow}(\mathbb{N})$ is called *the semigroup of cofinite monotone partial bijections* of $\mathbb{N}$. In [18] Gutik and Repovš studied properties of the semigroup $\mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$. In particular, they showed that $\mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$ is an inverse bisimple semigroup and all of its non-trivial group homomorphisms are either isomorphisms or group homomorphisms. It is obvious that $\mathbf{I}\mathbb{N}_{\infty}$ is an inverse submonoid of $\mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$.

A partial map $\alpha\colon \mathbb{N} \rightharpoonup \mathbb{N}$ is called *almost monotone* if there exists a finite subset A of $\mathbb{N}$ such that the restriction $\alpha\,|_{\mathbb{N}\setminus A}\colon \mathbb{N}\setminus A \rightharpoonup \mathbb{N}$ is a monotone partial map. By $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ we shall denote the semigroup of almost monotone injective partial transformations of $\mathbb{N}$ such that the sets $\mathbb{N}\setminus \operatorname{dom}\varphi$ and $\mathbb{N}\setminus \operatorname{ran}\varphi$ are finite for all $\varphi \in \mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$. Obviously, $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ is an inverse subsemigroup of the semigroup $\mathscr{I}_{\omega}$ and the semigroup $\mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$ is an inverse subsemigroup of $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ too. The semigroup $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ is called *the semigroup of cofinite almost monotone injective partial transformations* of $\mathbb{N}$. In the paper [11] the semigroup $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ is studied. In particular, it was shown that the semigroup $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ is inverse, bisimple and all of its non-trivial group homomorphisms are either isomorphisms or group homomorphisms. In the paper [23] we showed that every automorphism of a full inverse subsemigroup of $\mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$ which contains the semigroup $\mathscr{C}_{\mathbb{N}}$ is the identity map. Also there we constructed a submonoid $\mathbf{I}\mathbb{N}_{\infty}^{[\underline{1}]}$ of $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ with the following property: if S be an inverse subsemigroup of $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ such that S contains $\mathbf{I}\mathbb{N}_{\infty}^{[\underline{1}]}$ as a submonoid, then every non-identity congruence $\mathfrak{C}$ on S is a group congruence. We show that if S is an inverse submonoid of $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ such that S contains $\mathscr{C}_{\mathbb{N}}$ as a submonoid then S is simple and the quotient semigroup $S/\mathfrak{C}_{\mathbf{mg}}$, where $\mathfrak{C}_{\mathbf{mg}}$ is minimum group congruence on S, is isomorphic to the additive group of integers. Also, topologizations of inverse submonoids of $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ and embeddings of such semigroups into compact-like topological semigroups established in [11, 23]. Similar results for semigroups of cofinite almost monotone partial bijections and cofinite almost monotone partial bijections of $\mathbb{Z}$ were obtained in [19].

Next we need some notions defined in [22] and [23]. For an arbitrary positive integer n_0 we denote $[n_0) = \{n \in \mathbb{N}\colon n \geqslant n_0\}$. Since the set of all positive integers is well ordered, the definition of the semigroup $\mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ implies that for every $\gamma \in \mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ there exists the smallest positive integer $n_{\gamma}^{\mathbf{d}} \in \operatorname{dom}\gamma$ such that the restriction $\gamma|_{[n_{\gamma}^{\mathbf{d}})}$ of the partial map $\gamma\colon \mathbb{N} \rightharpoonup \mathbb{N}$ onto the set $[n_{\gamma}^{\mathbf{d}})$ is an element of the semigroup $\mathscr{C}_{\mathbb{N}}$, i.e., $\gamma|_{[n_{\gamma}^{\mathbf{d}})}$ is a some shift of $[n_{\gamma}^{\mathbf{d}})$. For every $\gamma \in \mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N})$ we put $\overrightarrow{\gamma} = \gamma|_{[n_{\gamma}^{\mathbf{d}})}$, i.e.

$$\operatorname{dom}\overrightarrow{\gamma} = [n_{\gamma}^{\mathbf{d}}), \quad (x)\overrightarrow{\gamma} = (x)\gamma \quad \text{for all } x \in \operatorname{dom}\overrightarrow{\gamma} \quad \text{and} \quad \operatorname{ran}\overrightarrow{\gamma} = (\operatorname{dom}\overrightarrow{\gamma})\,\gamma.$$

Also, we put

$$\underline{n}_{\gamma}^{\mathbf{d}} = \min\operatorname{dom}\gamma \qquad \text{for} \quad \gamma \in \mathscr{I}_{\infty}^{\Rsh\!\!\nearrow}(\mathbb{N}).$$

It is obvious that $\underline{n}_\gamma^{\mathbf{d}} = n_\gamma^{\mathbf{d}}$ when $\gamma \in \mathscr{C}_{\mathbb{N}}$, and $\underline{n}_\gamma^{\mathbf{d}} < n_\gamma^{\mathbf{d}}$ when $\gamma \in \mathscr{I}_\infty^{\nearrow}(\mathbb{N}) \setminus \mathscr{C}_{\mathbb{N}}$. Also for any $\gamma \in \mathbf{I}\mathbb{N}_\infty$ we denote

$$\underline{n}_\gamma^{\mathbf{r}} = (\underline{n}_\gamma^{\mathbf{d}})\gamma \qquad \text{and} \qquad n_\gamma^{\mathbf{r}} = (n_\gamma^{\mathbf{d}})\gamma.$$

The results of Section 3 of [24] imply that $n_\gamma^{\mathbf{r}} - \underline{n}_\gamma^{\mathbf{r}} = n_\gamma^{\mathbf{d}} - \underline{n}_\gamma^{\mathbf{d}}$ for any $\gamma \in \mathbf{I}\mathbb{N}_\infty$, and moreover for any non-negative integer j

$$\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]} = \left\{\gamma \in \mathbf{I}\mathbb{N}_\infty \colon n_\gamma^{\mathbf{d}} - \underline{n}_\gamma^{\mathbf{d}} \leqslant j\right\}$$

is a simple inverse subsemigroup of $\mathbf{I}\mathbb{N}_\infty$ such that $\mathbf{I}\mathbb{N}_\infty$ admits the following infinite semigroup series

$$\mathscr{C}_{\mathbb{N}} = \mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[0]} = \mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[1]} \subsetneqq \mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[2]} \subsetneqq \mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[3]} \subsetneqq \cdots \subsetneqq \mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[k]} \subsetneqq \cdots \subset \mathbf{I}\mathbb{N}_\infty.$$

For any positive integer k the semigroup $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[k]}$ is called the *monoid of cofinite isometries of positive integers with the noise k*.

A *(semi)topological semigroup* is a topological space with a (separately) continuous semigroup operation. An inverse topological semigroup with continuous inversion is called a *topological inverse semigroup*.

A topology τ on a semigroup S is called:

- a *semigroup* topology if (S,τ) is a topological semigroup;
- an *inverse semigroup* topology if (S,τ) is a topological inverse semigroup;
- a *shift-continuous* topology if (S,τ) is a semitopological semigroup.

The bicyclic monoid admits only the discrete semigroup Hausdorff topology [13]. Bertman and West in [6] extended this result for the case of Hausdorff semitopological semigroups. Stable and Γ-compact topological semigroups do not contain the bicyclic monoid [1, 25, 26]. The problem of embedding the bicyclic monoid into compact-like topological semigroups was studied in [3, 4, 5, 17].

In this paper we study algebraic properties of the monoid $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ and extend results of the papers [13] and [6] to the semigroups $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$, $j \geqslant 0$. In particular we show that for any positive integer j every Hausdorff shift-continuous topology τ on $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ is discrete and and if $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ is a proper dense subsemigroup of a Hausdorff semitopological semigroup S, then $S \setminus \mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ is a closed ideal of S, and moreover if S is a topological inverse semigroup then $S \setminus \mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ is a topological group. Also we describe the algebraic and topological structure of the closure of the monoid $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ in a locally compact topological inverse semigroup.

Latter in this paper without loss of generality we may assume that j is an arbitrary positive integer $\geqslant 2$.

2. Algebraic properties of the monoid $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$

The following simple proposition describes Green's relations on the monoid $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$.

Proposition 2.1. *For elements γ and δ of the semigroup $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ the following statements hold:*

(*i*) *$\gamma\mathscr{L}\delta$ in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ if and only if $\operatorname{dom}\gamma = \operatorname{dom}\delta$;*

(*ii*) *$\gamma\mathscr{R}\delta$ in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ if and only if $\operatorname{ran}\gamma = \operatorname{ran}\delta$;*

(*iii*) *$\gamma\mathscr{H}\delta$ in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ if and only if $\gamma = \delta$;*

(*iv*) *$\gamma\mathscr{D}\delta$ in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ if and only if $\operatorname{dom}\gamma$ $(\operatorname{ran}\gamma)$ and $\operatorname{dom}\delta$ $(\operatorname{ran}\delta)$ are isometric subsets of $\mathbb{N}$, i.e., there exists an isometry from $\operatorname{dom}\gamma$ $(\operatorname{ran}\gamma)$ onto $\operatorname{dom}\delta$ $(\operatorname{ran}\delta)$;*

(*v*) *$\gamma\mathscr{J}\delta$ in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$, i.e., $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a simple semigroup.*

Proof. Statements (*i*), (*ii*) and (*iii*) immediately follow from Proposition 3.2.11 of [27] and corresponding statements of Proposition 1 of [22].

Statement (*iv*) follows from the definition of the monoid and Proposition 3.2.5 of [27].

Statement (*v*) follows from Theorem 5 of [23]. □

Proposition 2.2 follows from the definition of the natural partial order $\preccurlyeq$ on an inverse semigroup and the statement that every element of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a partial shift of the integers (see [22, Lemma 1]).

Proposition 2.2. *Let γ and δ be elements of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. Then the following conditions are equivalent:*

(*i*) *$\gamma \preccurlyeq \delta$ in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$*

(*ii*) *$n_{\gamma}^{\mathbf{r}} - n_{\gamma}^{\mathbf{d}} = n_{\delta}^{\mathbf{r}} - n_{\delta}^{\mathbf{d}}$ and $\operatorname{dom}\gamma \subseteq \operatorname{dom}\delta$;*

(*iii*) *$n_{\gamma}^{\mathbf{r}} - n_{\gamma}^{\mathbf{d}} = n_{\delta}^{\mathbf{r}} - n_{\delta}^{\mathbf{d}}$ and $\operatorname{ran}\gamma \subseteq \operatorname{ran}\delta$.*

It is obvious that in statements (*ii*) and (*iii*) of Proposition 2.2 we may replace the symbols $n_{\gamma}^{\mathbf{r}}$ and $n_{\gamma}^{\mathbf{d}}$ by $\underline{n}_{\gamma}^{\mathbf{r}}$ and $\underline{n}_{\gamma}^{\mathbf{d}}$, respectively.

The definition of the minimum group congruence $\mathfrak{C}_{\mathbf{mg}}$ on $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ and Proposition 2.2 imply the following proposition.

Proposition 2.3. *Let γ and δ be elements of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. Then $\gamma\mathfrak{C}_{\mathbf{mg}}\delta$ in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ if and only if $n_{\gamma}^{\mathbf{r}} - n_{\gamma}^{\mathbf{d}} = n_{\delta}^{\mathbf{r}} - n_{\delta}^{\mathbf{d}}$. Moreover, the quotient semigroup $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}/\mathfrak{C}_{\mathbf{mg}}$ is isomorphic to the additive group of integers $\mathbb{Z}(+)$ by the map*

$$\pi_{\mathfrak{C}_{\mathbf{mg}}} \colon \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \to \mathbb{Z}(+), \qquad \gamma \mapsto n_{\delta}^{\mathbf{r}} - n_{\delta}^{\mathbf{d}}.$$

Example 2.4. We put $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} = \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \sqcup \mathbb{Z}(+)$ and extend the multiplications from $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ and $\mathbb{Z}(+)$ onto $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ in the following way:

$$k \cdot \gamma = \gamma \cdot k = k + (\gamma)\pi_{\mathfrak{C}_{\mathbf{mg}}} \in \mathbb{Z}(+), \qquad \text{for all} \quad k \in \mathbb{Z}(+) \quad \text{and} \quad \gamma \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}.$$

By Theorem 2.17 from [9, Vol. 1, p. 77] so defined binary operation is a semigroup operation on $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ such that $\mathbb{Z}(+)$ is an ideal in $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. Also, this semigroup operation extends the natural partial order $\preccurlyeq$ from $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ onto $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ in the following way:

(i) all distinct elements of $\mathbb{Z}(+)$ are pair-wise incomparable;

(ii) $k \preccurlyeq \gamma$ if and only if $n_{\gamma}^{\mathbf{r}} - n_{\gamma}^{\mathbf{d}} = k$ for $k \in \mathbb{Z}(+)$ and $\gamma \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$.

For any $x \in \mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ we denote $\uparrow_{\preccurlyeq} x = \left\{y \in \mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} : x \preccurlyeq y\right\}$.

By Proposition 7 of [22] the map $\mathfrak{h}\colon \mathbf{I}\mathbb{N}_{\infty} \to \mathscr{C}_{\mathbb{N}}$, $\gamma \mapsto \overrightarrow{\gamma}$ is a homomorphism. Then its restriction $\mathfrak{h}|_{\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}} \colon \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \to \mathscr{C}_{\mathbb{N}}$ is homomorphism, too.

A *homomorphic retraction* of a semigroup S is a map from S into S which is both a retraction and a homomorphism. The image of the homomorphic retraction is called a *homomorphic retract*. These terms seem to have first appeared in [10].

Since $(\gamma)\mathfrak{h} = \overrightarrow{\gamma} = \gamma$ for any $\gamma \in \mathscr{C}_{\mathbb{N}}$ we get the following proposition.

Proposition 2.5. *The map $\mathfrak{h}\colon \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \to \mathscr{C}_{\mathbb{N}}$, $\gamma \mapsto \overrightarrow{\gamma}$ is a homomorphic retraction, and hence the monoid $\mathscr{C}_{\mathbb{N}}$ is a homomorphic retract of $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$.*

For any subset $M \subseteq \{2, \ldots, j\}$ we denote

$$\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M] = \left\{\gamma \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} : n_{\gamma}^{\mathbf{d}} - x \in M \cup \{0\} \text{ for all } x \in \operatorname{dom}\gamma \text{ such that } x \leqslant n_{\gamma}^{\mathbf{d}}\right\}.$$

For arbitrary $M_1, M_2 \subseteq \{2, \ldots, j\}$ it is obvious that $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M_1] \subseteq \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M_2]$ if and only if $M_1 \subseteq M_2$, and moreover we have that $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M] = \mathscr{C}_{\mathbb{N}}$ when $M = \varnothing$ and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M] = \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ when $M = \{2, \ldots, j\}$.

Remark 2.6. By Lemma 1 of [22] we get that

$$\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M] = \left\{\gamma \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} : n_{\gamma}^{\mathbf{r}} - x \in M \cup \{0\} \text{ for all } x \in \operatorname{ran}\gamma \text{ such that } x \leqslant n_{\gamma}^{\mathbf{r}}\right\}.$$

Proposition 2.7. *$\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M]$ is an inverse semigroup of $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ for any $M \subseteq \{2, \ldots, j\}$.*

Proof. Fix any $\gamma, \delta \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M]$. We consider the following cases.

(1) If $n_{\gamma}^{\mathbf{r}} \leqslant n_{\delta}^{\mathbf{d}}$ then $n_{\gamma\delta}^{\mathbf{r}} = n_{\delta}^{\mathbf{r}}$ and $\operatorname{ran}(\gamma\delta) \subseteq \operatorname{ran}\delta$, because by Lemma 1 from [22] all elements of $\mathbf{I}\mathbb{N}_{\infty}$ are partial shifts of the set $\mathbb{N}$. This and Remark 2.6 imply that $\gamma\delta \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M]$.

(2) If $n_{\gamma}^{\mathbf{r}} > n_{\delta}^{\mathbf{d}}$ then by similar arguments as in the previous case we get that $n_{\gamma\delta}^{\mathbf{d}} = n_{\gamma}^{\mathbf{d}}$ and $\operatorname{dom}(\gamma\delta) \subseteq \operatorname{dom}\delta$. This implies that $\gamma\delta \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M]$.

Remark 2.6 implies that if $\gamma \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}[M]$ then so is γ^{-1}. □

3. On a topologization and a closure of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$

Later in the paper by $\mathbb{I}$ we denote the identity map of $\mathbb{N}$, and assume that α and β are the elements of the submonoid $\mathscr{C}_{\mathbb{N}}$ in $\mathbf{I}\mathbb{N}_{\infty}$ which are defined in Remark 1.1.

It is obvious that $\alpha\beta = \mathbb{I}$ and $\beta\alpha$ is the identity map of $\mathbb{N} \setminus \{1\}$. This implies the following lemma.

Lemma 3.1. *If* $\gamma \in \mathbf{I}\mathbb{N}_{\infty}$, *then*

(*i*) $\beta\alpha \cdot \gamma = \gamma$ *if and only if* $\operatorname{dom}\gamma \subseteq \mathbb{N} \setminus \{1\}$;
(*ii*) $\gamma \cdot \beta\alpha = \gamma$ *if and only if* $\operatorname{ran}\gamma \subseteq \mathbb{N} \setminus \{1\}$.

For any positive integer i let $\varepsilon^{[i]}$ be the identity map of the set $\mathbb{N} \setminus \{i\}$.

The following theorem generalized the results on the topologizabily of the bicyclic monoid obtained in [13] and [6].

Theorem 3.2. *For any positive integer* j *every Hausdorff shift-continuous topology* τ *on* $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ *is discrete.*

Proof. Since τ is Hausdorff, every retract of $\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$ is its closed subset. It is obvious that $\beta\alpha \cdot \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \cdot \beta\alpha$ are retracts of the topological space $\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$, because $\beta\alpha$ is an idempotent of $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. Later we shall show that the set $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \setminus \left(\beta\alpha \cdot \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \cup \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \cdot \beta\alpha\right)$ is finite.

By Lemma 3.1, $\gamma \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \setminus \left(\beta\alpha \cdot \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \cup \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \cdot \beta\alpha\right)$ if and only if $1 \in \operatorname{dom}\gamma$, $1 \in \operatorname{ran}\gamma$, and $\overline{n}_{\gamma}^{\mathbf{d}} - \underline{n}_{\gamma}^{\mathbf{d}} \leqslant j$. Then by Lemma 1 of [22], γ is a partial shift of the set of integers, and hence γ is an idempotent of $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ such that $1 \in \operatorname{dom}\gamma$ and $\varepsilon^{[2]} \cdot \ldots \cdot \varepsilon^{[j-1]} \preccurlyeq \gamma$. It is obvious that such idempotents γ are finitely many in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$, and hence the set $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \setminus \left(\beta\alpha \cdot \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \cup \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \cdot \beta\alpha\right)$ is finite. This implies that the point $\mathbb{I}$ has a finite open neighbourhood and hence $\mathbb{I}$ is an isolated point of the topological space $\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$.

We observe that $\mathbf{I}\mathbb{N}_{\infty}$, and hence $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$, is a submonoid of the semigroup $\mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$ of cofinite monotone partial bijections of $\mathbb{N}$ [22]. By Proposition 2.2 of [18] every right translation and every left translation by an element of the semigroup $\mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$ is a finite-to-one map, and hence such conditions hold for the semigroup $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. Also by Theorem 5 of [23], $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a simple semigroup. This implies that for any $\chi \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ there exist $\alpha, \beta \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ such that $\alpha\chi\beta = \mathbb{I}$, and moreover the equality $\alpha\chi\beta = \mathbb{I}$ has finitely many solutions. Since $\mathbb{I}$ is an isolated point of $\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$, the separate continuity of the semigroup operation in $\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$ and the above arguments imply that $\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$ is the discrete space. □

The following proposition generalized results obtained for the bicyclic monoid in [13] and [16].

Proposition 3.3. *Let* j *be any positive integer and* $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ *be a proper dense subsemigroup of a Hausdorff semitopological semigroup* S. *Then* $I = S \setminus \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ *is a closed ideal of* S.

Proof. By Theorem 3.2, $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a discrete subspace of S, and hence by Lemma 3 of [21], $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is an open subspace of S.

Fix an arbitrary element $y \in I$. If $xy = z \notin I$ for some $x \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ then there exists an open neighbourhood $U(y)$ of the point y in the space S such that $\{x\} \cdot U(y) = \{z\} \subset \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. The neighbourhood $U(y)$ contains infinitely many elements of the semigroup $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. This contradicts Proposition 2.2 of [18], which states that for each $v, w \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ both sets $\left\{u \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \colon vu = w\right\}$ and $\left\{u \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]} \colon uv = w\right\}$ are finite. The obtained contradiction implies that $xy \in I$ for all $x \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ and $y \in I$. The proof of the statement that $yx \in I$ for all $x \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ and $y \in I$ is similar.

Suppose to the contrary that $xy = w \notin I$ for some $x, y \in I$. Then $w \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ and the separate continuity of the semigroup operation in S implies that there exist open neighbourhoods $U(x)$ and $U(y)$ of the points x and y in S, respectively, such that $\{x\} \cdot U(y) = \{w\}$ and $U(x) \cdot \{y\} = \{w\}$. Since both neighbourhoods $U(x)$ and $U(y)$ contain infinitely many elements of the semigroup $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$, both equalities $\{x\} \cdot U(y) = \{w\}$ and $U(x) \cdot \{y\} = \{w\}$ contradict mentioned above Proposition 2.2 from [18]. The obtained contradiction implies that $xy \in I$. $\square$

Lemma 3.4. *Let j be any positive integer $\geqslant 2$. Then the element $\varepsilon \cdot (\beta\varepsilon)^j \cdot \alpha^j$ is an idempotent of the submonoid $\mathscr{C}_{\mathbb{N}}$ for any idempotent ε of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$.*

Proof. Since $\mathbb{I} = \alpha\beta$, we have that

$$\varepsilon \cdot \beta\varepsilon\alpha \cdot \beta^2\varepsilon\alpha^2 \cdot \ldots \cdot \beta^j\varepsilon\alpha^j = \varepsilon \cdot (\mathbb{I}\beta\varepsilon)^j \cdot \alpha^j = \varepsilon \cdot (\beta\varepsilon)^j \cdot \alpha^j$$

and

$$\beta^k\varepsilon\alpha^k \cdot \beta^k\varepsilon\alpha^k = \beta^k\varepsilon\mathbb{I}\varepsilon\alpha^k = \beta^k\varepsilon\varepsilon\alpha^k = \beta^k\varepsilon\alpha^k,$$

for any positive integer k. Also, $\varepsilon(\beta\varepsilon)^j\alpha^j$ is an idempotent of $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$, because $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is an inverse semigroup and the product of idempotents in an inverse semigroup is an idempotent as well.

By definitions of the partial transformations α and β and the above part of the proof we get that

$$n^{\mathbf{d}}_{\beta^k\varepsilon\alpha^k} = n^{\mathbf{d}}_{\varepsilon} + k \qquad \text{and} \qquad \underline{n}^{\mathbf{d}}_{\beta^k\varepsilon\alpha^k} = \underline{n}^{\mathbf{d}}_{\varepsilon} + k, \tag{3.1}$$

and hence

$$n^{\mathbf{d}}_{\beta^k\varepsilon\alpha^k} - \underline{n}^{\mathbf{d}}_{\beta^k\varepsilon\alpha^k} = n^{\mathbf{d}}_{\varepsilon} - \underline{n}^{\mathbf{d}}_{\varepsilon}, \tag{3.2}$$

for any positive integer k. Then equalities (3.1) and (3.2) imply that for any $k = 1, \ldots, j$ the idempotent

$$\varepsilon_k = \varepsilon(\beta\varepsilon)^k\alpha^k$$

has the following properties:

$$n^{\mathbf{d}}_{\varepsilon_k} = n^{\mathbf{d}}_{\beta^k\varepsilon\alpha^k}, \quad \underline{n}^{\mathbf{d}}_{\varepsilon_k} = \underline{n}^{\mathbf{d}}_{\beta^k\varepsilon\alpha^k},$$

and

$$1, \ldots, \underline{n}^{\mathbf{d}}_{\varepsilon}, \ldots, \underline{n}^{\mathbf{d}}_{\varepsilon} + k - 1, n^{\mathbf{d}}_{\varepsilon} - 1, n^{\mathbf{d}}_{\varepsilon}, \ldots, n^{\mathbf{d}}_{\varepsilon} + k - 1 \notin \operatorname{dom}\varepsilon_k.$$

Hence we get that ε_j is the identity map of $\left[n^{\mathbf{d}}_{\varepsilon}+j\right)$, which implies the statement of the lemma. $\square$

Lemma 3.5. *Let j be any positive integer and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ be a proper dense subsemigroup of a Hausdorff topological inverse semigroup S. Then there exists an idempotent $e\in S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ such that $V(e)\cap E(\mathscr{C}_{\mathbb{N}})$ is an infinite subset for any open neighbourhood $V(e)$ of e in S.*

Proof. By Proposition 3.3, $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is an ideal of S. Since S is an inverse semigroup, $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ contains an idempotent.

Put f be an arbitrary idempotent of $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. Since the unit element of a Hausdorff topological monoid is again the unit element of its closure in a topological semigroup, for an arbitrary positive integer k by Proposition 3.3 we have that

$$\beta^k f\alpha^k\cdot\beta^k f\alpha^k=\beta^k f\mathbb{I}f\alpha^k=\beta^k ff\alpha^k=\beta^k f\alpha^k,$$

and hence $\beta^k f\alpha^k\in E(S)\setminus E\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}\right)$. This implies that $e=f\cdot\beta f\alpha\cdot\ldots\cdot\beta^j f\alpha^j$ is an idempotent in S because S is an inverse semigroup. The continuity of the semigroup operation in S implies that for every open neighbourhood $V(e)$ of the point e in S there exists an open neighbourhood $W(f)$ of the point f in S such that

$$W(f)\cdot\beta\cdot W(f)\alpha\cdot\ldots\cdot\beta^j\cdot W(f)\cdot\alpha^j\subseteq V(e).$$

By Proposition II.3 of [13] the set $W(f)\cap E\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}\right)$ is infinite. Since for any positive integer n_0 there exist finitely many idempotents $\varepsilon\in\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ such that $n^{\mathbf{d}}_{\varepsilon}=n_0$, we conclude that the set $\left\{n^{\mathbf{d}}_{\varepsilon}\colon\varepsilon\in W(f)\cap E\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}\right)\right\}$ is infinite, too. Then there exists an infinite sequence $\{\varphi_i\}_{i\in\mathbb{N}}$ of idempotents of $W(f)\cap E\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}\right)$ such that $n^{\mathbf{d}}_{\varphi_{i_1}}\neq n^{\mathbf{d}}_{\varphi_{i_2}}$ for any distinct positive integers i_1 and i_2. Lemma 3.5 implies that $\varphi_i\cdot(\beta\varphi_i)^j\cdot\alpha^j$ is an idempotent of the submonoid $\mathscr{C}_{\mathbb{N}}$ which belongs to $V(e)$ for any positive integer i. Since the set $\left\{n^{\mathbf{d}}_{\varepsilon}\colon\varepsilon\in W(f)\cap E\left(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}\right)\right\}$ is infinite, the set $V(e)\cap E(\mathscr{C}_{\mathbb{N}})$ is infinite, too. $\square$

Theorem 3.6. *Let j be any positive integer and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ be a proper dense subsemigroup of a Hausdorff topological inverse semigroup S. Then $I=S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a topological group.*

Proof. We claim that the ideal I contains a unique idempotent.

Suppose to the contrary that I has at least two distinct idempotent e and f. By Lemma 3.5 without loss of generality we may assume that the set $V(e)\cap E(\mathscr{C}_{\mathbb{N}})$ is infinite for any open neighbourhood $V(e)$ of e in S. Since S is an inverse semigroup $ef=fe=h$ for some $h\in I\cap E(S)$. Fix an arbitrary open neighbourhood $U(h)$ in S. Then there exist disjoint open neighbourhoods $W(e)$ and $W(f)$ of the points e and f in S, respectively, such that $W(e)\cdot W(f)\subseteq U(h)$. Since S is Hausdorff, we can additionally assume that $W(e)\cap U(h)=\varnothing$ if $e\neq h$ and $W(f)\cap U(h)=\varnothing$

if $f \neq h$. Since $e \neq f$ we conclude that $W(e) \cap U(h) = \varnothing$ or $W(f) \cap U(h) = \varnothing$. Since the set $W(f) \cap E\big(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}\big)$ is infinite and for any positive integer n_0 there exist finitely many idempotents $\iota \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ such that $n_\iota^{\mathbf{d}} = n_0$, we conclude that the set $\big\{\underline{n}_\iota^{\mathbf{d}} \colon \iota \in W(f) \cap E\big(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}\big)\big\}$ is infinite as well. Also, the choice of the neighbourhood $W(e)$ implies that the set $\big\{\underline{n}_\iota^{\mathbf{d}} = n_\iota^{\mathbf{d}} \colon \iota \in W(e) \cap E\big(\mathscr{C}_{\mathbb{N}}\big)\big\}$ is infinite, too. Then the semigroup operation in $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ implies that there exist idempotents $\iota_e \in W(e)$ and $\iota_f \in W(f)$ such that $\iota_e \in \iota_e \cdot W(f)$ and $\iota_f \in \iota_f \cdot W(e)$, which implies $W(e) \cap U(h) \neq \varnothing \neq W(f) \cap U(h)$. But this contradicts the choice of the neighbourhoods $W(e)$, $W(f)$, $U(h)$.

Since S is an inverse semigroup, we have that $xx^{-1} = x^{-1}x = e$ for any $x \in I$. This implies that I is a subgroup of S with the unit element e. Also, the continuity of semigroup operation and the inversion in S implies that I is a topological group with the induced topology from S. □

Lemma 3.7 follows from the definition of an element $\overrightarrow{\gamma}$ for an arbitrary $\gamma \in \mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$.

Lemma 3.7. *For any $\gamma \in \mathscr{I}_{\infty}^{\nearrow}(\mathbb{N})$ the following statements hold:*

(i) $\overrightarrow{\gamma} \in \mathscr{C}_{\mathbb{N}}$;
(ii) $\overrightarrow{\gamma}^{-1} = \overrightarrow{\gamma^{-1}}$;
(iii) $\gamma\overrightarrow{\gamma}^{-1} = \overrightarrow{\gamma}\overrightarrow{\gamma}^{-1}$;
(iv) $\overrightarrow{\gamma}^{-1}\gamma = \overrightarrow{\gamma}^{-1}\overrightarrow{\gamma}$.

Proposition 3.8. *Let j be any positive integer and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ be a proper dense subsemigroup of a Hausdorff topological inverse semigroup S. Then the unique idempotent of $S \setminus \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ commutes with all elements of the semigroup $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$.*

Proof. By Theorem 3.6, $S \setminus \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a group. Put e_0 be the unique idempotent of $S \setminus \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. Also, by Lemma 3.5 the set $U(e_0) \cap E(\mathscr{C}_{\mathbb{N}})$ is infinite for any open neighbourhood $U(e_0)$ of the point e_0 in S. This implies that $e_0 \in \operatorname{cl}_S(\mathscr{C}_{\mathbb{N}})$. Then by Proposition III.2 of [13], $e_0 \cdot \gamma = \gamma \cdot e_0$ for any $\gamma \in \mathscr{C}_{\mathbb{N}}$.

Fix an arbitrary $\gamma \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. By Lemma 3.7 we have that

$$\overrightarrow{\gamma} \cdot \overrightarrow{\gamma}^{-1} \cdot \gamma = \gamma \cdot \overrightarrow{\gamma}^{-1} \cdot \overrightarrow{\gamma} = \overrightarrow{\gamma} \in \mathscr{C}_{\mathbb{N}}.$$

Since S is an inverse semigroup and $S \setminus \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is an ideal of S, Lemma 3.7 implies that

$$\begin{aligned} e_0 \cdot \gamma &= \big(e_0 \cdot \overrightarrow{\gamma} \cdot \overrightarrow{\gamma}^{-1}\big) \cdot \gamma = e_0 \cdot \big(\overrightarrow{\gamma} \cdot \overrightarrow{\gamma}^{-1} \cdot \gamma\big) = \\ &= e_0 \cdot \overrightarrow{\gamma} = \overrightarrow{\gamma} \cdot e_0 = \big(\gamma \cdot \overrightarrow{\gamma}^{-1} \cdot \overrightarrow{\gamma}\big) \cdot e_0 = \\ &= \gamma \cdot \big(\overrightarrow{\gamma}^{-1} \cdot \overrightarrow{\gamma} \cdot e_0\big) = \gamma \cdot e_0. \end{aligned}$$

This completes the proof of the proposition. □

Corollary 3.9. *Let j be any positive integer and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ be a proper dense subsemigroup of a Hausdorff topological inverse semigroup S. Then the group $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ contains a dense cyclic subgroup.*

Proof. By Proposition 3.8, the unique idempotent e_0 of $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ commutes with all elements of the semigroup $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ and hence the map $\mathfrak{h}\colon S\to S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$, $(\gamma)\mathfrak{h}=e_0\cdot\gamma$ is a homomorphisms. Since $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a subgroup of S, by Corollary 1.32 of [12] the image $(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]})\mathfrak{h}$ is a cyclic group. Also, since $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a dense subset of a topological semigroup S, Proposition 1.4.1 of [14] implies that the image $(\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]})\mathfrak{h}$ is a dense subset of $S\setminus\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. □

4. On a closure of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ in a locally compact topological inverse semigroup

In [13] Eberhart and Selden described the closure of the bicyclic monoid in a locally compact topological inverse semigroup. We give this description in the terms of the monoid $\mathscr{C}_{\mathbb{N}}$.

Example 4.1. The definition of the bicyclic monoid, its algebraic properties (see [12, Section 1.12]) and Remark 1.1 imply that the following relation

$$\gamma\sim\delta \quad \text{if and only if} \quad n_{\gamma}^{\mathbf{r}}-n_{\gamma}^{\mathbf{d}}=n_{\delta}^{\mathbf{r}}-n_{\delta}^{\mathbf{d}}, \qquad \gamma,\delta\in\mathscr{C}_{\mathbb{N}},$$

coincides with the minimum group congruence $\mathfrak{C}_{\mathbf{mg}}$ on $\mathscr{C}_{\mathbb{N}}$. Moreover, the quotient semigroup $\mathscr{C}_{\mathbb{N}}/\mathfrak{C}_{\mathbf{mg}}$ is isomorphic to the additive group of integers $\mathbb{Z}(+)$ by the map

$$\pi_{\mathfrak{C}_{\mathbf{mg}}}\colon\mathscr{C}_{\mathbb{N}}\to\mathbb{Z}(+), \quad \gamma\mapsto n_{\delta}^{\mathbf{r}}-n_{\delta}^{\mathbf{d}}.$$

The minimum group congruence $\mathfrak{C}_{\mathbf{mg}}$ on $\mathscr{C}_{\mathbb{N}}$ defines the natural partial order $\preccurlyeq$ on the monoid $\mathscr{C}_{\mathbb{N}}$ in the following way:

$$\gamma\preccurlyeq\delta \quad \text{if and only if} \quad n_{\gamma}^{\mathbf{r}}-n_{\gamma}^{\mathbf{d}}=n_{\delta}^{\mathbf{r}}-n_{\delta}^{\mathbf{d}} \quad \text{and} \quad n_{\gamma}^{\mathbf{d}}\geqslant n_{\delta}^{\mathbf{d}}, \qquad \gamma,\delta\in\mathscr{C}_{\mathbb{N}}.$$

We put $\mathcal{C}\mathscr{C}_{\mathbb{N}}=\mathscr{C}_{\mathbb{N}}\sqcup\mathbb{Z}(+)$ and extend the multiplications from the semigroup $\mathscr{C}_{\mathbb{N}}$ and the group $\mathbb{Z}(+)$ onto $\mathcal{C}\mathscr{C}_{\mathbb{N}}$ in the following way:

$$k\cdot\gamma=\gamma\cdot k=k+(\gamma)\pi_{\mathfrak{C}_{\mathbf{mg}}}\in\mathbb{Z}(+), \qquad \text{for all} \quad k\in\mathbb{Z}(+) \quad \text{and} \quad \gamma\in\mathscr{C}_{\mathbb{N}}.$$

Then so defined binary operation is a semigroup operation on $\mathcal{C}\mathscr{C}_{\mathbb{N}}$ such that $\mathbb{Z}(+)$ is an ideal in $\mathcal{C}\mathscr{C}_{\mathbb{N}}$. Also, this semigroup operation extends the natural partial order $\preccurlyeq$ from $\mathscr{C}_{\mathbb{N}}$ onto $\mathcal{C}\mathscr{C}_{\mathbb{N}}$ in the following way:

(i) all distinct elements of $\mathbb{Z}(+)$ are pair-wise incomparable;
(ii) $k\preccurlyeq\gamma$ if and only if $n_{\gamma}^{\mathbf{r}}-n_{\gamma}^{\mathbf{d}}=k$ for $k\in\mathbb{Z}(+)$ and $\gamma\in\mathscr{C}_{\mathbb{N}}$.

For any $x\in\mathcal{C}\mathscr{C}_{\mathbb{N}}$ we denote ${\uparrow_{\preccurlyeq}}x=\{y\in\mathcal{C}\mathscr{C}_{\mathbb{N}}\colon x\preccurlyeq y\}$.

We define the topology τ_{lc} on $\mathcal{C}\mathscr{C}_{\mathbb{N}}$ in the following way:

(i) all elements of the monoid $\mathscr{C}_{\mathbb{N}}$ are isolated points in $(\mathcal{C}\mathscr{C}_{\mathbb{N}},\tau_{\mathsf{lc}})$;

(ii) for any $k \in \mathbb{Z}(+)$ the family $\mathscr{B}_{\mathsf{lc}}(k) = \{U_i(k) \colon i \in \mathbb{N}\}$, where

$$U_i(k) = \{k\} \cup \big\{\gamma \in \mathscr{C}_{\mathbb{N}} \colon k \preccurlyeq \gamma \text{ and } n_\gamma^{\mathbf{d}} \geqslant i\big\},$$

is the base of the topology τ_{lc} at the point $k \in \mathbb{Z}(+)$.

In [13] Eberhart and Selden proved that τ_{lc} is the unique Hausdorff locally compact semigroup inverse topology on $\mathcal{C}\mathscr{C}_{\mathbb{N}}$. Moreover, they shown that if $\mathscr{C}_{\mathbb{N}}$ is a proper dense subsemigroup of a Hausdorff locally compact topological inverse semigroup S, then S is topologically isomorphic to $(\mathcal{C}\mathscr{C}_{\mathbb{N}}, \tau_{\mathsf{lc}})$.

Example 4.2. Let $\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ be a semigroup defined in Example 2.4. Put M be an arbitrary subset of $\{2, \ldots, j\}$.

We define the topology τ_{lc}^M on $\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ in the following way:

(i) all elements of the monoid $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ are isolated points in $\big(\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}, \tau_{\mathsf{lc}}^M\big)$;

(ii) for any $k \in \mathbb{Z}(+)$ the family $\mathscr{B}_{\mathsf{lc}}^M(k) = \big\{U_i^M(k) \colon i \in \mathbb{N}\big\}$, where

$$U_i^M(k) = \{k\} \cup \big\{\gamma \in \mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}[M] \colon k \preccurlyeq \gamma \text{ and } n_\gamma^{\mathbf{d}} \geqslant i\big\},$$

is the base of the topology τ_{lc}^M at the point $k \in \mathbb{Z}(+)$.

Remark 4.3.

1. We observe that a simple verifications show that the following conditions hold:
   (i) if $k = 0$ then $U_i^M(k) = U_i^M(0) = \{0\} \cup \big\{\gamma \in \mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}[M] \colon k \preccurlyeq \gamma \text{ and } \gamma \notin {\uparrow_{\preccurlyeq}}\beta^{i-2}\alpha^{i-2}\big\}$;
   (ii) if $k > 0$ then $U_i^M(k) = \{0\} \cup \big\{\gamma \in \mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}[M] \colon k \preccurlyeq \gamma \text{ and } \gamma \notin {\uparrow_{\preccurlyeq}}\beta^{i-2}\alpha^{i-2+k}\big\}$;
   (iii) if $k < 0$ then $U_i^M(k) = \{0\} \cup \big\{\gamma \in \mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}[M] \colon k \preccurlyeq \gamma \text{ and } \gamma \notin {\uparrow_{\preccurlyeq}}\beta^{i-2-k}\alpha^{i-2}\big\}$.
2. Since all elements of the monoid $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ are isolated points in $\big(\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}, \tau_{\mathsf{lc}}^M\big)$ and all distinct elements of the subgroup $\mathbb{Z}(+)$ are incomporable with the respect to the natural partial order on $\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$, Proposition 2.2 implies that τ_{lc}^M is a Hausdorff topology on $\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$. Also, since for any $\gamma \in \mathscr{C}_{\mathbb{N}}$ the set ${\uparrow_{\preccurlyeq}}\gamma$ is finite we get that $U_i^M(k)$ is compact for any $k \in Z(+)$ and any positive integer i. This implies that the space $\big(\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}, \tau_{\mathsf{lc}}^M\big)$ is locally comapct, and hence by Theorems 3.3.1, 4.2.9 and Corollary 3.3.6 from [14] it is metrizable.

Proposition 4.4. $\big(\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}, \tau_{lc}^M\big)$ *is a topological inverse semigroup.*

Proof. Since all elements of the monoid $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ are isolated points in $\big(\mathcal{C}\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}, \tau_{\mathsf{lc}}^M\big)$ and all distinct elements of the subgroup $\mathbb{Z}(+)$ commute with elements of $\mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$, it is suffices to check the continuity of the semigroup operation at the pairs (γ, k_1) and (k_1, k_2) where $\gamma \in \mathbf{I}\mathbb{N}_\infty^{\boldsymbol{g}[j]}$ and $k_1, k_2 \in \mathbb{Z}(+)$.

Fix any $\gamma \in \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ and $k \in \mathbb{Z}(+)$. Then $\overrightarrow{\gamma} = \beta^p\alpha^r$ for some fixed non-negative integers p and r. Hence

$$\gamma \cdot k = (\gamma)\pi_{\mathfrak{C}_{\mathbf{mg}}} + k = \big(\overrightarrow{\gamma}\big)\pi_{\mathfrak{C}_{\mathbf{mg}}} + k = r - p + k,$$

and for any positive integer $i > \max\{p, r\} + j$ we have that

$$\gamma \cdot U_i^M(k) \subseteq U_i^M(r - p + k).$$

Fix any $k_1, k_2 \in \mathbb{Z}(+)$. Then for any positive integer $i > j$ by Proposition 1.4.7 of [27] and Proposition 2.7 we have that $U_i^M(k_1) \cdot U_i^M(k_2) \subseteq U_i^M(k_1 + k_2)$.

The above arguments and the equality $\big(U_i^M(k)\big)^{-1} = U_i^M(-k)$ complete the proof of the proposition. □

Lemma 4.5. *Let j be any positive integer and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ be a proper dense subsemigroup of a Hausdorff locally compact topological inverse semigroup S. Then $G = S \setminus \mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is topologically isomorphic to the discrete additive group of integers $\mathbb{Z}(+)$.*

Proof. By Corollary 3.9, G is a subgroup of $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ which contains a dense cyclic subgroup. By Theorem 3.2, $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a discrete subspace of S, and hence by Theorem 3.3.9 of [14], G is a closed subspace of S. Then Theorem 3.3.8 of [14] and Theorem 3.6 imply that G with the induced topology from S is a locally compact topological group. By the Weil Theorem (see [31]) the topological group G is either compact or discrete. By Lemma 3.5 the remainder $\operatorname{cl}_S(\mathscr{C}_{\mathbb{N}}) \setminus \mathscr{C}_{\mathbb{N}}$ of the subsemigroup $\mathscr{C}_{\mathbb{N}}$ in S is non-empty. Then by Theorem 3.3.8 of [14], $\operatorname{cl}_S(\mathscr{C}_{\mathbb{N}})$ is a locally compact space. Theorem V.7 of [13] implies that $H = \operatorname{cl}_S(\mathscr{C}_{\mathbb{N}}) \setminus \mathscr{C}_{\mathbb{N}}$ is a group, which is topologically isomorphic to the discrete additive group of integers $\mathbb{Z}(+)$. By Proposition 1.4.19 of [2], H is a closed discrete subgroup of G, and hence by Theorem 1.4.23 of [2] the topological group G is topologically isomorphic to the discrete additive group of integers $\mathbb{Z}(+)$. □

A partial order $\leq$ on a topological space X is called *closed* (or *continuous*) if the relation $\leq$ is a closed subset of $X \times X$ in the product topology [15]. A topological space with a closed partial order is called a *pospace*.

Later we assume that $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a proper dense subsemigroup of a Hausdorff locally compact topological inverse semigroup S and we identify the topological group G with the discrete additive group of integers $\mathbb{Z}(+)$.

We observe that equality ${\uparrow_{\preccurlyeq}}k = \{\gamma \in S \colon \gamma \cdot 0 = k\}$ implies that ${\uparrow_{\preccurlyeq}}k$ is an open-and-closed subset of S for any $k \in \mathbb{Z}(+)$. Since $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a discrete subspace of S the above arguments and Lemma 4.5 imply the following lemma:

Lemma 4.6. *The natural partial order $\preccurlyeq$ on S is closed, and moreover ${\uparrow_{\preccurlyeq}}x$ is open-and-closed subset of S for any $x \in S$.*

Lemma 4.7. *For any $k, l \in \mathbb{Z}(+)$ the subspace ${\uparrow_{\preccurlyeq}}k$ and ${\uparrow_{\preccurlyeq}}l$ of S are homeomorphic. Moreover, the map $P_{\alpha^k} \colon {\uparrow_{\preccurlyeq}}0 \to {\uparrow_{\preccurlyeq}}k$, $x \mapsto x \cdot \alpha^k$ is a homeomorphism for $k > 0$, and the map $\Lambda_{\beta^k} \colon {\uparrow_{\preccurlyeq}}0 \to {\uparrow_{\preccurlyeq}}k$, $x \mapsto \beta^k \cdot x$ is a homeomorphism for $k < 0$.*

Proof. Proposition 1.4.7 from [27] implies that the maps P_{α^k} and Λ_{β^k} are well defined. It is obvious that complete to prove that the second part of the lemma holds. We shall show that the map P_{α^k} determines a homeomorphism from ${\uparrow_{\preccurlyeq}}0$ onto ${\uparrow_{\preccurlyeq}}k$. In the case of the map Λ_{β^k} the proof is similar.

We define a map $P_{\beta^k} \colon {\uparrow_{\preccurlyeq}}k \to {\uparrow_{\preccurlyeq}}0$ by the formula $(x)P_{\beta^k} = x \cdot \beta^k$. Then we have that $(0)P_{\alpha^k} = k$ and $(k)P_{\beta^k} = 0$. Moreover, we have that $(x)P_{\alpha^k}P_{\beta^k} = x$ for any $x \in {\uparrow_{\preccurlyeq}}0$ and $(y)P_{\beta^k}P_{\alpha^k} = y$ for any $y \in {\uparrow_{\preccurlyeq}}k$. Therefore the compositions of maps $P_{\alpha^k}P_{\beta^k} \colon {\uparrow_{\preccurlyeq}}0 \to {\uparrow_{\preccurlyeq}}0$ and $P_{\beta^k}P_{\alpha^k} \colon {\uparrow_{\preccurlyeq}}k \to {\uparrow_{\preccurlyeq}}k$ are identity maps of the sets ${\uparrow_{\preccurlyeq}}0$ and ${\uparrow_{\preccurlyeq}}k$, respectively. Hence the maps P_{α^k} and P_{β^k} are bijections, and hence P_{β^k} is inverse of P_{α^k}. Since right translations in the topological semigroup S are continuous, the maps $P_{\alpha^k} \colon {\uparrow_{\preccurlyeq}}0 \to {\uparrow_{\preccurlyeq}}k$ and $P_{\beta^k} \colon {\uparrow_{\preccurlyeq}}k \to {\uparrow_{\preccurlyeq}}0$ are homeomorphisms. □

By Lemma 3.5 the remainder $\operatorname{cl}_S(\mathscr{C}_{\mathbb{N}}) \setminus \mathscr{C}_{\mathbb{N}}$ of the subsemigroup $\mathscr{C}_{\mathbb{N}}$ in S is non-empty. Also, Theorem V.7 of [13] implies that the remainder $\operatorname{cl}_S(\mathscr{C}_{\mathbb{N}}) \setminus \mathscr{C}_{\mathbb{N}}$ is a group, which is topologically isomorphic to the discrete additive group of integers $\mathbb{Z}(+)$. This and results of [13, Section V] (see Example 4.1) imply the following proposition:

Proposition 4.8. *Let j be any positive integer and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ be a proper dense subsemigroup of a Hausdorff locally compact topological inverse semigroup $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$. Then τ induces the topology τ_{lc} on the semigroup $\mathcal{C}\mathscr{C}_{\mathbb{N}}$.*

If $M = \varnothing$ then we denote the locally compact semigroup inverse topology τ_{lc}^{M} on the monoid $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ by $\tau_{\mathsf{lc}}^{\varnothing}$. Also in the case when $M = \{2, \ldots, j\}$ we denote the topology τ_{lc}^{M} on $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ by $\tau_{\mathsf{lc}}^{[2:j]}$.

Proposition 4.8 implies the following:

Proposition 4.9. *Let j be any positive integer and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ be a proper dense subsemigroup of a Hausdorff locally compact topological inverse semigroup $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$. Then $\tau_{\mathsf{lc}}^{\varnothing} \subseteq \tau \subseteq \tau_{\mathsf{lc}}^{[2:j]}$.*

Theorem 4.10. *Let j be any positive integer and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ be a proper dense subsemigroup of a Hausdorff locally compact topological inverse semigroup (S, τ). Then (S, τ) topologically isomorphic to the topological inverse semigroup $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau_{\mathsf{lc}}^{M}\right)$ for some subset M of $\{2, \ldots, j\}$.*

Proof. Lemma 4.5 implies that the inverse semigroup S is isomorphic to the monoid $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$. Also, by the definition of the monoid $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$, Lemma 4.7 and Proposition 4.9 we get that there exists a maximal subset M_1 of $\{2, \ldots, j\}$ such that the following condition holds:

(∗) for every open neighbourhood V_0 of the point $0 \in \mathbb{Z}(+)$ in $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$ there exists an open neighbourhood $U_i^{M_1}(0)$ of 0 in $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau_{\mathsf{lc}}^{M_1}\right)$ such that $U_i^{M_1}(0) \subseteq V_0$ and $V_0 \setminus U_i^{M_1}(0)$ is infinite.

Since the topology τ is locally compact and $\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$ is a discrete subsemigroup of $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$, without loss of generality we may assume that the open neighbourhood V_0 is compact.

The maximality of M_1 and condition (∗) imply that there exists a subset $M_1^1 \subseteq \{2, \ldots, j\}$ such that $M_1 \subset M_1^1$, $\left|M_1^1 \setminus M_1\right| = 1$ and for every open neighbourhood V_0 of the point $0 \in \mathbb{Z}(+)$ in $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$ the following conditions hold:

$$\left|\left(V_0 \cap U_i^{M_1^1}(0)\right) \setminus U_i^{M_1}(0)\right| = \infty \quad \text{and} \quad \left|U_i^{M_1^1}(0) \setminus \left(V_0 \cap U_i^{M_1^1}(0)\right)\right| = \infty. \tag{4.1}$$

By continuity of the semigroup operation in $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$ there exists a compact-and-open neighbourhood $U_0 \subseteq V_0$ of the point $0 \in \mathbb{Z}(+)$ in the space $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$ such that $\beta \cdot U_0 \cdot \alpha \subseteq V_0$. Then the semigroup operation of $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$, the above inclusion and conditions (4.1) imply that the set $V_0 \setminus U_0$ is infinite, which contradicts the compactness of V_0. This and maximality of M_1 imply that the set $V_0 \setminus U_i^{M_1}(0)$ is finite for every open neighbourhood V_0 of the point $0 \in \mathbb{Z}(+)$ in $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau\right)$ and any open neighbourhood $U_i^{M_1}(0)$ of 0 in $\left(\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}, \tau_{\mathsf{lc}}^{M_1}\right)$. Then the bases of τ and $\tau_{\mathsf{lc}}^{M_1}$ at the point $0 \in \mathbb{Z}(+)$ coincide, and hence by Lemma 4.7 we get that $\tau = \tau_{\mathsf{lc}}^{M_1}$. □

Corollary 4.11. *For any positive integer j there exists exactly 2^{j-1} pairwise topologically non-isomorphic Hausdorff locally compact semigroup inverse topologies on the monoid $\mathcal{C}\mathbf{I}\mathbb{N}_{\infty}^{\boldsymbol{g}[j]}$.*

Acknowledgements

The authors acknowledge Alex Ravsky and the referee for useful important comments and suggestions.

References

[1] L. W. Anderson, R. P. Hunter, R. J. Koch, *Some results on stability in semigroups.* Trans. Amer. Math. Soc. **117** (1965), 521–529.
[2] A. Arhangel'skii, M. Tkachenko, *Topological Groups and Related Structures*, Atlantis, 2008.
[3] T. Banakh, S. Dimitrova, O. Gutik, *The Rees-Suschkiewitsch Theorem for simple topological semigroups*, Mat. Stud. **31** (2009), no. 2, 211–218.
[4] T. Banakh, S. Dimitrova, O. Gutik, *Embedding the bicyclic semigroup into countably compact topological semigroups*, Topology Appl. **157** (2010), no. 18, 2803–2814.
[5] S. Bardyla, A. Ravsky, *Closed subsets of compact-like topological spaces*, Appl. Gen. Topol. **21** (2020), no. 2, 201–214.
[6] M. O. Bertman, T. T. West, *Conditionally compact bicyclic semitopological semigroups*, Proc. Roy. Irish Acad. **A76** (1976), no. 21–23, 219–226.

[7] O. Bezushchak, *On growth of the inverse semigroup of partially defined co-finite automorphisms of integers*, Algebra Discrete Math. (2004), no. 2, 45–55.

[8] O. Bezushchak, *Green's relations of the inverse semigroup of partially defined cofinite isometries of discrete line*, Visn., Ser. Fiz.-Mat. Nauky, Kyiv. Univ. Im. Tarasa Shevchenka (2008), no. 1, 12–16.

[9] J. H. Carruth, J. A. Hildebrant, R. J. Koch, *The Theory of Topological Semigroups*, Vol. I, Marcel Dekker, Inc., New York and Basel, 1983; Vol. II, Marcel Dekker, Inc., New York and Basel, 1986.

[10] D. R. Brown, *Topological semilattices on the two-cell*, Pacific J. Math. **15** (1965), no. 1, 35–46.

[11] I. Ya. Chuchman, O. V. Gutik, *Topological monoids of almost monotone injective co-finite partial selfmaps of the set of positive integers*. Carpathian Math. Publ. **2** (2010), no. 1, 119–132.

[12] A. H. Clifford, G. B. Preston, *The Algebraic Theory of Semigroups*, Vol. I., Amer. Math. Soc. Surveys 7, Providence, R.I., 1961; Vol. II., Amer. Math. Soc. Surveys 7, Providence, R.I., 1967.

[13] C. Eberhart, J. Selden, *On the closure of the bicyclic semigroup*, Trans. Amer. Math. Soc. **144** (1969), 115–126.

[14] R. Engelking, *General Topology*, 2nd ed., Heldermann, Berlin, 1989.

[15] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove, D. S. Scott, *Continuous Lattices and Domains*, Cambridge Univ. Press, 2003.

[16] O. Gutik, *On the dichotomy of a locally compact semitopological bicyclic monoid with adjoined zero*, Visnyk L'viv Univ., Ser. Mech.-Math. **80** (2015), 33–41.

[17] O. Gutik, D. Repovš, *On countably compact* 0*-simple topological inverse semigroups*, Semigroup Forum **75** (2007), no. 2, 464–469.

[18] O. Gutik, D. Repovš, *Topological monoids of monotone, injective partial selfmaps of* ℕ *having cofinite domain and image*, Stud. Sci. Math. Hungar. **48** (2011), no. 3, 342–353.

[19] O. Gutik, D. Repovš, *On monoids of injective partial selfmaps of integers with cofinite domains and images*, Georgian Math. J. **19** (2012), no. 3, 511–532.

[20] O. Gutik, D. Repovš, *On monoids of injective partial cofinite selfmaps*, Math. Slovaca **65** (2015), no. 5, 981–992.

[21] O. Gutik, A. Savchuk, *On the semigroup* $\mathbf{ID}_\infty$*,* Visn. Lviv. Univ., Ser. Mekh.-Mat. **83** (2017), 5–19 (in Ukrainian).

[22] O. Gutik, A. Savchuk, *The semigroup of partial co-finite isometries of positive integers,* Bukovyn. Mat. Zh. **6** (2018), no. 1–2, 42–51 (in Ukrainian).

[23] O. Gutik, A. Savchuk, *On inverse submonoids of the monoid of almost monotone injective co-finite partial selfmaps of positive integers*, Carpathian Math. Publ. **11** (2019), no. 2, 296–310.

[24] O. Gutik, A. Savchuk, *On the monoid of cofinite partial isometries of* ℕ *with the usual metric*, Visn. Lviv. Univ., Ser. Mekh.-Mat. **89** (2020), 17–30.

[25] J. A. Hildebrant, R. J. Koch, *Swelling actions of* Γ*-compact semigroups*, Semigroup Forum **33** (1986), 65–85.

[26] R. J. Koch, A. D. Wallace, *Stability in semigroups*, Duke Math. J. **24** (1957), no. 2, 193–195.

[27] M. Lawson, *Inverse Semigroups. The Theory of Partial Symmetries*, Singapore: World Scientific, 1998.

[28] M. Petrich, *Inverse Semigroups*, John Wiley & Sons, New York, 1984.

[29] W. Ruppert, *Compact Semitopological Semigroups: An Intrinsic Theory*, Lect. Notes Math., **1079**, Springer, Berlin, 1984.

[30] V. V. Wagner, *Generalized groups*, Dokl. Akad. Nauk SSSR **84** (1952), 1119–1122 (in Russian).

[31] A. Weil, *L'intégration dans les groupes topologiques et ses applications*, Actualites Scientifiques No. 869, Hermann, Paris, 1940.

Ivan Franko National University of Lviv, Ukraine
oleg.gutik@lnu.edu.ua, ogutik@gmail.com

Ivan Franko National University of Lviv, Ukraine
pavlo.khylynskyi@lnu.edu.ua

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON THE BEHAVIOR OF INTEGRABLE FUNCTIONS AT INFINITY

JOANNA HORBACZEWSKA

Abstract. The main purpose of the paper is to examine dependence between integrability of function and its behaviour at infinity. Although convergence of series implies convergence to 0 of the sequence under the sum, usually integrability of function does not imply its convergence to 0. In the paper we examine conditions under which every integrable function is convergent to 0 (both in traditional sense and in density). Considered functions can be defined on $\mathbb{N}$ (real sequences), on $\mathbb{R}$ or on any space with an infinite σ-finite measure (then we have to define what we understand by infinity in that space). Achieved results generalize both known dependencies between convergence of series and sequences, and relatively new results of C. P. Niculescu i F. Popovici on convergence in density of functions.

Keywords: integrable functions, convergence, convergence in density

MSC2010: 03G05, 57N01

Received 29 April 2021, revised 28 June 2021, accepted 9 August 2021

1. Introduction

We start with reminding known results on a relation between the convergence of series and the convergence of the sequence of its terms.

One of necessary conditions for convergence of series is Olivier's Theorem published in 1827 [9]:

Theorem 1.1 (Olivier, 1827)**.** *Let (a_n) be a nonnegative and nonincreasing sequence of real numbers. If a series $\sum_{n=1}^{\infty} a_n$ is convergent then $n \cdot a_n \to 0$.*

In 2003 T. Šalát and V. Toma [10] noticed that the monotonocity condition can be omitted if we weaken the thesis to the convergence in density according to the following definition.

https://doi.org/10.2478/9788366675360-011.

Definition 1.2. We define density of a set $A \subset \mathbb{N}$ as a value

$$d(A) = \lim_{n\to\infty} \frac{|A \cap [1,n]|}{n}$$

(if the limit exists) where $|A \cap [1,n]|$ denotes the cardinality of the set $A \cap [1,n]$.

Definition 1.3. We say that a sequence (a_n) of real numbers converges in density to $l \in \mathbb{R}$ if for every $\varepsilon > 0$ the set $A_\varepsilon = \{k \in \mathbb{N} : |a_k - l| \geq \varepsilon\}$ has the density equal to 0 i.e. $d(A_\varepsilon) = 0$. Then we write $a_n \xrightarrow{(d)} l$.

Theorem 1.4 (Šalát, Toma, 2003)**.** *Let a series $\sum_{n=1}^{\infty} a_n$ be absolutely convergent. Then $n \cdot a_n \xrightarrow{(d)} 0$.*

2. Convergence in density in $\mathbb{R}$

The notions of density and convergence in density have their counterparts for real functions defined on the real half-line.

Definition 2.1. Let $\alpha \in \mathbb{R}$ and let $g : (\alpha, +\infty) \to (0, +\infty)$ be a locally integrable function such that $\int_\alpha^\infty g(t)dt = \infty$. Then the density of a set A with respect to g is the value $d_g(A) = \lim_{x\to\infty} \frac{\int_{A\cap(\alpha,x)} g(t)dt}{\int_\alpha^x g(t)dt}$ (if the limit exists).

Definition 2.2. Let $\alpha \in \mathbb{R}$, $f : (\alpha, \infty) \to \mathbb{R}$ be a measurable function and let $g : (\alpha, +\infty) \to (0, +\infty)$ be a locally integrable function such that $\int_\alpha^\infty g(t)dt = \infty$. We denote $A_\varepsilon = \{x \in (\alpha, +\infty) : |f(x) - l| \geq \varepsilon\}$.

We say that the function f converges to $l \in \mathbb{R}$ in density g ($f(x) \xrightarrow{(g)} l$) if for every $\varepsilon > 0$ we have $d_g(A_\varepsilon) = \lim_{x\to\infty} \frac{\int_{A_\varepsilon\cap(\alpha,x)} g(t)dt}{\int_\alpha^x g(t)dt} = 0$.

3. Theorem on convergence of integrable functions defined on the real half-line

The following version of a function counterpart of Theorem 1.4 (Šalát, Toma) was proved by C. P. Niculescu and F. Popovici in [6].

Theorem 3.1 (Niculescu, Popovici, 2012)**.** *Let $n \in \mathbb{N} \cup \{0\}$ and $g(x) = \frac{1}{\prod_{k=0}^{n-1} \ln^k x}$. If f is integrable, then $(\prod_{k=0}^{n} \ln^k x) \cdot f(x) \xrightarrow{(g)} 0$, where $\ln^k$ denotes the k-th composition of the logarithm function.*

For $n = 0$ this theorem is a continuous counterpart of Theorem 1.4. We generalise it for a much wider class of functions g.

Theorem 3.2. *Let $\alpha \in \mathbb{R}$, $f : (\alpha, \infty) \to \mathbb{R}$ be an integrable function, and let $g : (\alpha, +\infty) \to (0, +\infty)$ be a locally integrable function such that $\int_\alpha^\infty g(t)dt = \infty$. Let $h : (\alpha, \infty) \to \mathbb{R}$ be a function such that $|h(x)| \leq \frac{\int_\alpha^x g(t)dt}{g(x)}$. Then $h(x) \cdot f(x) \xrightarrow{(g)} 0$.*

Proof. Let $G(x) = \int_\alpha^x g(t)dt$. Then $G : (\alpha, \infty) \to (0, \infty)$ is positive, increasing and it tends to infinity as $x \to \infty$.

Fix $\varepsilon > 0$, $\delta > 0$ and denote $A_\varepsilon = \{x \in \mathbb{R} : |h(x)f(x)| \geq \varepsilon\}$. Since $\int_\alpha^\infty |f(t)|dt < \infty$, we can choose $y \geq \alpha$ such that $\int_y^\infty |f(t)|dt < \frac{\varepsilon\delta}{2}$. Then there exists $x_0 > y$ such that $G(x_0) > \frac{2}{\varepsilon\delta} G(y) \int_\alpha^y |f(t)|dt$, and as a consequence for every $x > x_0$ we have $G(x) > \frac{2}{\varepsilon\delta} G(y) \int_\alpha^y |f(t)|dt$. Hence for every $x > x_0$ we have (where $\mathbf{1}_{\frac{G(t)}{g(t)}|f(t)|\geq\varepsilon}(t) = \mathbf{1}_{\{t\in[\alpha,\infty):\frac{G(t)}{g(t)}|f(t)|\geq\varepsilon\}}(t)$):

$$\frac{\int_{A_\varepsilon\cap(\alpha,x)} g(t)dt}{G(x)} \leq \frac{\int_\alpha^x g(t)\mathbf{1}_{\frac{G(t)}{g(t)}|f(t)|\geq\varepsilon}(t)dt}{G(x)} \leq \frac{\int_\alpha^x g(t)\mathbf{1}_{\frac{G(t)}{g(t)}|f(t)|\geq\varepsilon}(t)\frac{G(t)}{g(t)}|f(t)|\frac{1}{\varepsilon}dt}{G(x)} \leq$$

$$\leq \frac{\int_\alpha^x g(t)\frac{G(t)}{g(t)}|f(t)|\frac{1}{\varepsilon}dt}{G(x)} = \frac{\int_\alpha^x G(t)|f(t)|dt}{\varepsilon G(x)} = \frac{\int_\alpha^y G(t)|f(t)|dt}{\varepsilon G(x)} + \frac{\int_y^x G(t)|f(t)|dt}{\varepsilon G(x)} \leq$$

$$\leq \frac{\int_\alpha^y G(y)|f(t)|dt}{\varepsilon G(x)} + \frac{\int_y^x G(x)|f(t)|dt}{\varepsilon G(x)} = \frac{G(y)}{\varepsilon G(x)} \int_\alpha^y |f(t)|dt + \frac{1}{\varepsilon} \int_y^x |f(t)|dt <$$

$$< \frac{G(y)}{\varepsilon G(x)} \cdot \frac{\varepsilon\delta G(x)}{2G(y)} + \frac{1}{\varepsilon} \cdot \frac{\varepsilon\delta}{2} = \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Therefore for every $\varepsilon > 0$ we have $d_g(A_\varepsilon) = 0$. Hence $h(x) \cdot f(x) \xrightarrow{(g)} 0$. □

In the next theorem we add an assumption about a monotonocity of a function $\frac{|f(x)|}{g(x)}$, but in exchange we can fortify our thesis to the classical convergence and we achieve a continuous counterpart of Theorem 1.1.

Theorem 3.3. *Let $\alpha \in \mathbb{R}$, $f : (\alpha, \infty) \to \mathbb{R}$ be an integrable function, and let $g : (\alpha, +\infty) \to (0, +\infty)$ be a locally integrable function such that $\int_\alpha^\infty g(t)dt = \infty$. Let $\frac{|f(x)|}{g(x)}$ be nonincreasing. Moreover, let $h : (\alpha, \infty) \to \mathbb{R}$ be a function such that $|h(x)| \leq \frac{\int_\alpha^x g(t)dt}{g(x)}$. Then $h(x) \cdot f(x) \to 0$.*

Proof. Let $G(x) = \int_\alpha^x g(t)dt$. Then $G : (\alpha, \infty) \to (0, \infty)$ is increasing and it converges to infinity when $x \to \infty$.

Fix $\varepsilon > 0$. Since $\int_0^\infty |f(t)|dt < \infty$ we can choose $y > \alpha$ such that $\int_y^\infty |f(t)|dt < \frac{\varepsilon}{2}$. In that case there exists $x_0 > y$ such that $G(x_0) > 2G(y)$ and as a consequence for every $x > x_0$ we have $G(x) > 2G(y)$. Hence $G(x) < 2(G(x) - G(y))$. For every $x > x_0$ we have:

$$|h(x)f(x)| \leq \frac{G(x)}{g(x)}|f(x)| \leq 2\frac{G(x) - G(y)}{g(x)}|f(x)| =$$

$$= 2\int_y^x \frac{|f(x)|}{g(x)} g(t)dt \leq 2\int_y^x \frac{|f(t)|}{g(t)} g(t)dt =$$

$$= 2\int_y^x |f(t)|dt < 2 \cdot \frac{\varepsilon}{2} = \varepsilon$$

We showed that for every $\varepsilon > 0$ there exists $x_0 \in \mathbb{R}$ such that for every $x > x_0$ we have $|h(x)f(x)| < \varepsilon$. Hence$|h(x)f(x)| \to 0$. □

4. Point at infinity and its neighbourhoods in a measure space

To define the point at infinity and its neighbourhood system in a measure space we will mimic the construction of the one-point compactification of a topological locally compact space. In compactification, neighbourhoods of infinity are defined as complements of compact sets. Similarly we would like to define neighbourhoods of infinity in our measure space as complements of sets that are 'small', so the ones with finite measure. Unfortunately if we will take complements of all sets with finite measure we will not get generalisations of above theorems. To achieve our goal we will consider some subfamilies $\{X_t\}_{t\in T}$ of the family of sets with finite measure, indexed by a directed set T.

Definition 4.1. Let $(X, \mathcal{F}, \mu)$ be a space with an infinite σ-finite measure. Let $T \neq \emptyset$ be a directed set, and let $\{X_t\}_{t\in T}$ be a subfamily of family of sets with finite measure in X such that for any $s, t \in T$ we have $s \leq t \Rightarrow X_s \subset X_t$. Moreover we assume that for any $A \in \mathcal{F}$ with finite measure and any $\varepsilon > 0$ there exists $t \in T$ such that $\mu(A \setminus X_t) < \varepsilon$. We will call the family $\{X \setminus X_t : t \in T\}$ the system of neighbourhoods of infinity and denote $U_t = X \setminus X_t$ dla $t \in T$.

The system of neighbourhoods of infinity introduced in Definition 4.1 determines convergence of functions at infinity.

Definition 4.2. Let $f : X \to \mathbb{R}$ and $l \in \mathbb{R}$. We say that $f(x)$ converges to l ($f(x) \to l$) if for every $\varepsilon > 0$ there exists $U \in \{U_t\}_{t\in T}$ such that for any $x \in U$ we have $|f(x) - l| < \varepsilon$.

Remark 4.3. Similarly we can define infinite limits. We say that $f(x)$ tends to $+\infty$ ($f(x) \to +\infty$), if for every $M \in \mathbb{R}$ there exists a neighbourhood of infinity U such that for every $x \in U$ we have $f(x) > M$. Similarly $f(x)$ tends to $-\infty$ ($f(x) \to -\infty$), if for every $M \in \mathbb{R}$ there exists a neighbourhood of infinity U such that for every $x \in U$ we have $f(x) < M$.

It is important to notice that different choices of a subfamily $\{X_t\}_{t\in T}$ for a fixed space with a measure $(X, \mathcal{F}, \mu)$ can lead to different determinations of limits.

Example 4.4. Let $(X, \mathcal{F}, \mu) = ((0, +\infty), Borel((0, \infty)), \lambda)$, where λ stands for the Lebesgues measure. Put $T = (0, +\infty)$, $X_t = (0, t)$ for $t \in T$ and let $\{\widetilde{X}_b\}_{b\in B}$ be a family of all subsets of X with the finite Lebesgues measure. The families $\{X_t\}_{t\in T}$ and $\{\widetilde{X}_b\}_{b\in B}$ determine two systems of neighbourhoods of infinity $\{U_t\}_{t\in T}$ and $\{\widetilde{U}_b\}_{b\in B}$. Of course the convergence in $\{U_t\}_{t\in T}$ is equivalent to classical convergence on $\mathbb{R}$ at $+\infty$.

Let $Y = \bigcup_{n=0}^{\infty} \left(n, n+\frac{1}{2^n}\right)$ and

$$f(x) = \begin{cases} 1 & \text{for } x \in Y, \\ 0 & \text{for } x \notin Y. \end{cases}$$

The function f is not convergent in $\{U_t\}_{t\in T}$ but is is convergent to 0 in $\{\widetilde{U}_b\}_{b\in B}$.

However we can find some dependencies.

Remark 4.5. Let $(X, \mathcal{F}, \mu)$ be any space with an infinite σ-finite measure and let $\{U_t\}_{t\in T}$ and $\{U_b\}_{b\in B}$ be systems of neighbourhoods of infinity such that $\{U_b\}_{b\in B} \subset \{U_t\}_{t\in T}$. Then if $f(x) \to l$ in $\{U_b\}_{b\in B}$ then $f(x) \to l$ in $\{U_t\}_{t\in T}$.

Remark 4.6. Let $(X, \mathcal{F}, \mu)$ be any space with an infinite σ-finite measure and let $\{U_t\}_{t\in T}$ be a system of neighbourhoods of infinity. Moreover, let $B \subset T$ be cofinal in T. Then the convergence $f(x) \to l$ in $\{U_b\}_{b\in B}$ is equivalent to the convergence $f(x) \to l$ in $\{U_t\}_{t\in T}$.

5. Convergence in density at infinity

We will generalise notions of the density and the convergence in density for any space with an infinite σ-finite measure $(X, \mathcal{F}, \mu)$ and a system of neighbourhoods of infinity determined by a family $\{X_t\}_{t\in T}$ (see Definition 4.1).

Definition 5.1. We define a density of a set $A \in \mathcal{F}$ with respect to a measure ν on $(X, \mathcal{F})$ such that $\forall_{t\in T}\ \nu(X_t) < \infty$ and $\sup\{\nu(X_t) : t \in T\} = \infty$ as the value $d_\nu(A) = \lim_t \frac{\nu(X_t \cap A)}{\nu(X_t)}$ (if the limit exists).

Definition 5.2. Let $f : X \to \mathbb{R}$ be an $\mathcal{F}$-measurable function, $l \in \mathbb{R}$ and let ν be a measure on $(X, \mathcal{F})$ such that $\forall_{t\in T}\ \nu(X_t) < \infty$ and $\sup\{\nu(X_t) : t \in T\} = \infty$. We say that $f(x)$ converges to l in density of measure ν, $(f(x) \xrightarrow{(\nu)} l)$ if for any $\varepsilon > 0$ we have $d_\nu(A_\varepsilon) = \lim_t \frac{\nu(X_t \cap A_\varepsilon)}{\nu(X_t)} = 0$, where $A_\varepsilon = \{x \in X : |f(x) - l| \geq \varepsilon\}$.

Remark 5.3. We can similarly define infinite limits. We say that $f(x)$ tends to $+\infty$ in density of a measure ν $(f(x) \xrightarrow{(\nu)} +\infty)$, when for every $M \in \mathbb{R}$ we have $d_\nu(\{x \in X : f(x) \leq M\}) = 0$. Similarly, $f(x)$ tends to $-\infty$ in density of a measure ν $(f(x) \xrightarrow{(\nu)} -\infty)$, if for every $M \in \mathbb{R}$ we have $d_\nu(\{x \in X : f(x) \geq M\}) = 0$.

Let us notice that the convergence of a function f implies its convergence in any density. It is also true that the convergence in density depends on a choice of the system of neighbourhoods of infinity.

We are going to acknowledge some dependencies, begining with describing sets of 0 density in a measure μ.

Theorem 5.4. *Let $(X, \mathcal{F}, \mu)$ be a space with an infinite σ-finite measure and let $A \in \mathcal{F}$.*

i) *If $\mu(A) < \infty$, then for any family $\{X_t\}_{t\in T}$ satisfying conditions from Definition 4.1 we have $d_\mu(A) = 0$.*
ii) *If $d_\mu(A) = 0$ and $\{X_t\}_{t\in T}$ is a family of all finite-measure sets, then $\mu(A) < \infty$.*
iii) *If $\mu(X \setminus A) = \infty$, then there exists a family $\{X_n\}_{n\in\mathbb{N}}$ such that $d_\mu(A) = 0$.*

Proof. **i)** Let us assume that $\mu(A) < \infty$. We chose $\varepsilon > 0$. Since μ is an infinite σ-finite measure, there exists a set $B \in \mathcal{F}$ such that $\frac{\mu(A)}{\varepsilon} < \mu(B) < \infty$. By the properties of the family $\{X_t\}_{t\in T}$ there exists $t_0 \in T$ such that $\mu(B \setminus X_{t_0}) < \mu(B) - \frac{\mu(A)}{\varepsilon}$. Hence $\mu(X_{t_0}) > \frac{\mu(A)}{\varepsilon}$. Then for every $t \geq t_0$ we have

$$\frac{\mu(X_t \cap A)}{\mu(X_t)} \leq \frac{\mu(A)}{\mu(X_{t_0})} < \varepsilon.$$

Hence $d_\mu(A) = 0$.

ii) Let us assume that $\mu(A) = \infty$. We will show that the equality $d_\mu(A) = 0$ does not hold by proving that for every $t_0 \in T$ there exists $t \geq t_0$ such that $\frac{\mu(X_t\cap A)}{\mu(X_t)} \geq \frac{1}{2}$. Since $\mu(A) = \infty$ and $\mu(X_{t_0}) < \infty$ we have $\mu(A \setminus X_{t_0}) = \infty$. Because μ is σ-finite, we can choose sets of finite measure $B_1 \subset B_2 \subset \dots \in \mathcal{F}$, such that $\bigcup_{n=1}^{\infty} B_n = X$. We will consider sets $(A \setminus X_{t_0}) \cap B_n$. They form an ascending sequence of sets of finite measure and the union of this sequence is the set $A \setminus X_{t_0}$. Therefore $\lim_{n\to\infty} \mu((A \setminus X_{t_0}) \cap B_n) = \mu(A \setminus X_{t_0}) = \infty$. Hence for some n_0 we have $\mu((A \setminus X_{t_0}) \cap B_{n_0}) > \mu(X_{t_0})$. We consider set $X_{t_0} \cup ((A \setminus X_{t_0}) \cap B_{n_0})$. Since it has finite measure, and it contains the set X_{t_0}, there exists $t \geq t_0$ such that $X_{t_0} \cup ((A \setminus X_{t_0}) \cap B_{n_0}) = X_t$. We have

$$\begin{aligned}\frac{\mu(X_t \cap A)}{\mu(X_t)} &= \frac{\mu(X_{t_0} \cap A) + \mu((A \setminus X_{t_0}) \cap B_{n_0})}{\mu(X_{t_0}) + \mu((A \setminus X_{t_0}) \cap B_{n_0})} \\ &\geq \frac{\mu((A \setminus X_{t_0}) \cap B_{n_0})}{\mu(X_{t_0}) + \mu((A \setminus X_{t_0}) \cap B_{n_0})} \geq \frac{1}{2}.\end{aligned}$$

Hence $d_\mu(A) = 0$ does not hold.

iii) Since μ is σ-finite there exist sets of finite measure $B_1 \subset B_2 \subset \dots \in \mathcal{F}$, such that $\bigcup_{n=1}^{\infty} B_n = X$. We will define sets $X_1, X_2, \dots$ by induction. Let $X_1 = B_1$. We assume that X_{n-1} is defined. Since $\mu(X \setminus A) = \infty$, there exists B_{k_n} such that $\mu(B_{k_n} \setminus A) > n \cdot \mu(X_{n-1} \cup B_n)$. We define $X_n = X_{n-1} \cup B_n \cup (B_{k_n} \setminus A)$. We have $X_n \cap A = (X_{n-1} \cup B_n) \cap A \subset (X_{n-1} \cup B_n)$. Hence

$$\frac{\mu(X_n \cap A)}{\mu(X_n)} \leq \frac{\mu(X_{n-1} \cup B_n)}{\mu(B_{k_n} \setminus A)} < \frac{1}{n}.$$

Therefore $d_\mu(A) \leq \lim_{n\to\infty} \frac{1}{n}$, hence $d_\mu(A) = 0$. From the definition of X_n we have $X_1 \subset X_2 \subset \dots$ and $B_n \subset X_n$ for any n, hence $\{X_n\}_{n\in\mathbb{N}}$ satisfies conditions from the Definition 4.1. □

The above theorem enables us to present an easy example showing how the choice of the system of neighbourhoods of infinity can affect the convergence in density.

Example 5.5. Let $(X, \mathcal{F}, \mu)$ be a space with an infinite σ-finite measure, $\nu = \mu$ and let $A \in \mathcal{F}$ be such that $\mu(A) = \infty$ and $\mu(X \setminus A) = \infty$.

Let

$$f(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

According to Theorem 5.4 neither set A nor its compliment $X \setminus A$ has density 0 in $\{U_t\}_{t\in T} = \{X \setminus X_t\}_{t\in T}$ where $\{X_t\}_{t\in T}$ is a family of all finite-measure subsets of X. Therefore the function f is not convergent in density.

According to the same theorem there exists a system of neighbourhoods of infinity $\{U_b\}_{b\in B}$ in which A is a set of density 0 and therefore f is convergent in density to 0.

Similarly, there exists a system of neighbourhoods of infinity $\{U_c\}_{c\in C}$, in which $X \setminus A$, is a set of density 0 and therefore f is convergent in density to 1.

However, some dependencies can be found:

Theorem 5.6. *Let $(X, \mathcal{F}, \mu)$ be a space with an infinite σ-finite measure. Let B and T be directed and cofinal sets such that $B \subset T$. Families $\{U_t\}_{t\in T} = \{X \setminus X_t\}_{t\in T}$ and $\{U_b\}_{b\in B} = \{X \setminus X_b\}_{b\in B}$ are cofinal systems of neighbourhoods of infinity and $\{U_b\}_{b\in B} \subset \{U_t\}_{t\in T}$. Let ν be a measure which satisfies conditions from Definition 5.1 (in correspondence to $\{U_t\}_{t\in T}$, which implicates that they are satisfied also for $\{U_b\}_{b\in B}$). Then from $f(x) \xrightarrow{(\nu)} l$ in $\{U_t\}_{t\in T}$, it follows that $f(x) \xrightarrow{(\nu)} l$ in $\{U_b\}_{b\in B}$.*

Proof. We choose $\varepsilon > 0$ and let $A_\varepsilon = \{x \in X : |f(x) - l| \geq \varepsilon\}$. For any $\delta > 0$ there exists $t_0 \in T$, such that for any $t \geq t_0$ we have $\frac{\nu(A_\varepsilon \cap X_t)}{\nu(X_t)} < \delta$. Since B is cofinal with T, there exists $b_0 \in B$, such that $b_0 \geq t_0$. Therefore for any $b \in B$ such that $b \geq b_0$ we have $b \geq t_0$, and $\frac{\nu(A_\varepsilon \cap X_b)}{\nu(X_b)} < \delta$. Hence $f(x) \xrightarrow{(\nu)} l$ in $\{U_b\}_{b\in B}$ □

6. Generalisation of convergence theorems

We want to generalise Theorems 1.1, 1.4, 3.1, 3.2 and 3.3 by substituting natural numbers and a real half-line by any space $(X, \mathcal{F}, \mu)$ with an infinite σ-finite measure. We need two auxiliary lemmas. We omit the simple and standard (but very technical) proofs.

Lemma 6.1. *Let $(X, \mathcal{F}, \mu)$ be a space with an infinite σ-finite measure, $\{U_t\}_{t\in T} = \{X \setminus X_t\}_{t\in T}$ be a system of neighbourhoods of infinity and let $f : X \to \mathbb{R}$ be a μ-integrable function. For any $\varepsilon > 0$ there exists $s \in T$ such that for any $t \geq s$ we have $\int_{X \setminus X_t} |f(x)|\mu(dx) < \varepsilon$.*

Lemma 6.2. *Let $(X, \mathcal{F}, \mu)$ be a space with an infinite σ-finite measure, $\{U_t\}_{t\in T} = \{X \setminus X_t\}_{t\in T}$ be a system of neighbourhoods of infinity and ν an an infinite measure on $(X, \mathcal{F})$ absolutely continuous with respect to μ , such that for any $t \in T$ we have $\nu(X_t) < \infty$. Then there exists $N \in \mathcal{F}$ such that $\nu(N) = 0$ and $X \setminus \bigcup_{t\in T} X_t \subset N$. Moreover,* $\sup\{\nu(X_t) : t \in T\} = \infty$.

Now we prove a generalisation of Theorem 1.4 (Šalát, Toma, [10]), 3.1 (Niculescu, Popovici, [4, 5, 6, 7]) and Theorem 3.2.

Theorem 6.3. *Let $(X, \mathcal{F}, \mu)$ be a space with an infinite σ-finite measure, and let $\{U_t\}_{t\in T} = \{X \setminus X_t\}_{t\in T}$ be a system of neighbourhoods of infinity.*

Let ν be an infinite measure absolutely continuous with respect to μ, with density g, such that for every $t \in T$ we have $\nu(X_t) < \infty$.

Let $h : X \to \mathbb{R}$ be a measurable function such that $|h(x)| \le \frac{\inf\{\nu(X_t):x\in X_t\}}{g(x)}$ when $g(x) > 0$.

If $f : X \to \mathbb{R}$ is μ-integrable, then $h(x) \cdot f(x) \xrightarrow{(\nu)} 0$.

We assume $\inf\emptyset = +\infty$. Therefore the inequality $|h(x)| \le \frac{\inf\{\nu(X_t):x\in X_t\}}{g(x)}$ is meaningfull only for $x \in \bigcup_{t\in T} X_t$.

Proof. From Lemma 6.2 there exists a set $N \in \mathcal{F}$ such that $X \setminus \bigcup_{t\in T} X_t \subset N$ and $\nu(N) = 0$. For any $x \in X \setminus N$ let $G(x) = \inf\{\nu(X_t) : x \in X_t\}$. For any $t \in T$ we denote $X'_t = X_t \setminus (N \cup \{x \in X : g(x) = 0\})$.

Fix $\varepsilon > 0$, $\delta > 0$. From Lemma 6.1 we can choose $s \in T$, such that

$$\int_{X\setminus X_s} |f(x)|\mu(dx) < \frac{\varepsilon\delta}{2}.$$

As ν is an infinite measure, from Lemma 6.2 we know that there exists $t_0 \ge s$ such that for any $t \ge t_0$ we have $\nu(X_t) > \frac{2}{\varepsilon\delta}\nu(X_s)\int_{X_s} |f(x)|\mu(dx)$. Then for $t \ge t_0$ (where $\mathbf{1}_{|h(x)\cdot f(x)|\ge\varepsilon}(x) = \mathbf{1}_{\{x\in X_t:|h(x)\cdot f(x)|\ge\varepsilon\}}(x)$):

$$\begin{aligned}
\frac{\nu(\{|h(x)\cdot f(x)| \ge \varepsilon\} \cap X_t)}{\nu(X_t)} &= \frac{\int_{X_t} \mathbf{1}_{|h(x)\cdot f(x)|\ge\varepsilon}(x)\nu(dx)}{\nu(X_t)} \\
= \frac{\int_{X'_t} \mathbf{1}_{|h(x)\cdot f(x)|\ge\varepsilon}(x)\nu(dx)}{\nu(X'_t)} &\le \frac{\int_{X'_t} g(x)\mathbf{1}_{\frac{G(x)}{g(x)}|f(x)|\ge\varepsilon}(x)\mu(dx)}{\nu(X'_t)} \\
\le \frac{\int_{X'_t} g(x)\frac{G(x)}{g(x)}|f(x))|\mu(dx)}{\varepsilon\nu(X'_t)} &= \frac{\int_{X'_t} G(x)|f(x)|\mu(dx)}{\varepsilon\nu(X'_t)} \\
= \frac{\int_{X'_s} G(x)|f(x)|\mu(dx)}{\varepsilon\nu(X'_t)} &+ \frac{\int_{X'_t\setminus X'_s} G(x)|f(x)|\mu(dx)}{\varepsilon\nu(X'_t)} \\
\le \frac{\nu(X'_s)\int_{X'_s} |f(x)|\mu(dx)}{\varepsilon\nu(X'_t)} &+ \frac{\nu(X'_t)\int_{X'_t\setminus X'_s} |f(x)|\mu(dx)}{\varepsilon\nu(X'_t)}
\end{aligned}$$

$$< \frac{\nu(X_t') \cdot \frac{\varepsilon\delta}{2}}{\varepsilon\nu(X_t')} + \frac{\frac{\varepsilon\delta}{2}}{\varepsilon} = \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

□

In Theorem 1.4 the space X is $\mathbb{N}$ with the counting measure, the family $\{X_t\}_{t\in T}$ consists of the sets $X_t = \{1, 2, ..., t\}$, where $T = \mathbb{N}$ and $g \equiv 1$.

In Theorem 3.2 the space X is $(\alpha, +\infty)$, the family $\{X_t\}_{t\in T}$ is a family of all beginning intervals (α, t), where $T = (\alpha, +\infty)$. If we choose $g(x) = \frac{1}{\prod_{k=0}^{n-1} \ln^k x}$ and $h(x) = \prod_{k=0}^{n} \ln^k x$, where $\ln^k$ means the k-th composition of the the logarithm function, we have Theorem 3.1.

Therefore Theorems 1.4 (Salat and Toma), 3.1 (Niculescu Popovici) and 3.2 are special cases of Theorem 6.3.

In a similar way we want to generalise Theorem 3.3. For this purpose we have to formulate a monotonicity condition from Theorem 3.3 in therms of a measure space. We have to assume that a set T (and, as a consequence, a family $\{X_t\}_{t\in T}$) is linearly ordered. It is intuitive, for any $t \in T$, to treat elements of $X \setminus X_t$ as greater then the ones in X_t. Then a monotonicity can be expressed as follows: for any $t \in T$ and for any $x \in X \setminus X_t$ we have $\frac{|f(x)|}{g(x)} \leq \inf\left\{\frac{|f(y)|}{g(y)} : y \in X_t\right\}$.

Theorem 6.4. *Let $(X, \mathcal{F}, \mu)$ be a space with an infinite σ-finite measure and let $\{U_t\}_{t\in T} = \{X \setminus X_t\}_{t\in T}$ be a system of neighbourhoods of infinity.*

Let T be linearly ordered and $\bigcup_{t\in T} X_t = X$. Let ν be an infinite measure absolutely continues wrt μ, with density $g : X \to (0, +\infty)$, such that for every $t \in T$ we have $\nu(X_t) < \infty$. We assume also that there exists $a \in \mathbb{R}$ such that for every $x \in X$ we have $\frac{\inf\{\nu(X_t):x\in X_t\}}{\sup\{\nu(X_t):x\notin X_t\}} < a$.

Let $h : X \to \mathbb{R}$ be a measurable function such that $|h(x)| \leq \frac{\inf\{\nu(X_t):x\in X_t\}}{g(x)}$ when $g(x) > 0$.

Let $f : X \to \mathbb{R}$ be a μ-integrable function, such that for every $t \in T$ and every $x \in X \setminus X_t$ we have $\frac{|f(x)|}{g(x)} \leq \inf\left\{\frac{|f(y)|}{g(y)} : y \in X_t\right\}$. Then $h(x) \cdot f(x) \to 0$.

Proof. For any $x \in X$ let $G(x) = \inf\{\nu(X_t) : x \in X_t\}$.

Fix any $\varepsilon > 0$. From Lemma 6.1 we can choose $s \in T$, such that

$$\int_{X\setminus X_s} |f(y)|\mu(dy) < \frac{\varepsilon}{2a}.$$

Since ν is an infinite measure, there is $t_0 \geq s \in T$ such that for any $t \geq t_0$ we have $\nu(X_t) > 2\nu(X_s)$. Fix $x \in X \setminus X_t$ and $t', t'' \geq t$ such that $x \in X_{t'}$, $x \notin X_{t''}$ and $\frac{\nu(X_{t'})}{\nu(X_{t''})} < a$. We have:

$$|h(x) \cdot f(x)| \leq \frac{G(x)}{g(x)}|f(x)| \leq \frac{\nu(X_{t'})}{g(x)}|f(x)| \leq a \cdot \frac{\nu(X_{t''})}{g(x)}|f(x)| \leq$$

$$\leq a\cdot\frac{2(\nu(X_{t''})-\nu(X_s))}{g(x)}|f(x)| = 2a\cdot\nu(X_{t''}\setminus X_s)\cdot\frac{|f(x)|}{g(x)} = 2a\int_{X_{t''}\setminus X_s}\frac{|f(x)|}{g(x)}g(y)\mu(dy).$$

Since $x\in X\setminus X_{t''}$, for any $y\in X_{t''}\setminus X_s\subset X_{t''}$ we have $\frac{|f(x)|}{g(x)}\leq\frac{|f(y)|}{g(y)}$. Hence

$$2a\int_{X_{t''}\setminus X_s}\frac{|f(x)|}{g(x)}g(y)\mu(dy)\leq 2a\int_{X_{t''}\setminus X_s}\frac{|f(y)|}{g(y)}g(y)\mu(dy)=$$
$$=2a\int_{X_{t''}\setminus X_s}|f(y)|\mu(dy)\leq 2a\int_{X\setminus X_s}|f(y)|\mu(dy)<2a\cdot\frac{\varepsilon}{2a}=\varepsilon.$$

□

In Theorem 6.4 additionally to the monotonoicity condition we add a condition stating that there exists $a\in\mathbb{R}$ such that for every $x\in X$ we have $\frac{\inf\{\nu(X_t):x\in X_t\}}{\sup\{\nu(X_t):x\notin X_t\}}<a$. Its necessity can be seen in the example below:

Example 6.5. Consider $X=(1,+\infty)$ with the Lebesgues measure. Let $a_n=2^{2^n}$, $T=\{a_n:n\in\mathbb{N}\}$, $X_t=(1,t)$ for $t\in T$. Moreover, let $f:(1,+\infty)\to\mathbb{R}$ be a function given by the formula $f(x)=\frac{1}{x^2}$, and let $g\equiv 1$.

We take $h:(1,+\infty)\to\mathbb{R}$ given by the formula $h(x)=\inf\{\lambda(X_t):x\in X_t\}=a_n-1$, where $n=\lfloor\log_2\log_2 x\rfloor+1$. There is none $a\in\mathbb{R}$ such that for any $x\in X$ we have $\frac{\inf\{\nu(X_t):x\in X_t\}}{\sup\{\nu(X_t):x\notin X_t\}}<a$, because for any $x\in X_{n+1}\setminus X_n$ we have

$$\frac{\inf\{\nu(X_t):x\in X_t\}}{\sup\{\nu(X_t):x\notin X_t\}}=\frac{a_{n+1}-1}{a_n-1}=\frac{2^{2^{n+1}}-1}{2^{2^n}-1}\xrightarrow{n\to+\infty}\infty.$$

Moreover:

$$h(a_n)\cdot f(a_n)=\frac{a_{n+1}-1}{a_n^2}=\frac{2^{2^{n+1}}-1}{(2^{2^n})^2}\xrightarrow{n\to+\infty}1.$$

Extra assumptions are satisfied both in Theorem 1.1 and in Theorem 3.3.

In Theorem 1.1 the space X is $\mathbb{N}$ with the counting measure, the family $\{X_t\}_{t\in T}$ consists of the sets $X_t=\{1,2,...,t\}$, where $T=\mathbb{N}$ and $g\equiv 1$. Then the family $\{X_t\}_{t\in T}$ is linearly ordered and assumption of Theorem 6.4 is satisfied for $a=3>\frac{n+1}{n}$ for any $n\in\mathbb{N}$.

In Theorem 3.3 the space X is $(\alpha,+\infty)$, the family $\{X_t\}_{t\in T}$ is a family of all beginning intervals (α,t), where $T=(\alpha,+\infty)$. The family $\{X_t\}_{t\in T}$ is linearly ordered and as a we can take any real number greater then 1.

Hence Theorems 1.1 (Olivier) and 3.3 are special cases of Theorem 6.4.

References

[1] J.P. Dix, *Existence of the limit at infinity for a function that is integrable on the half line*, Rose-Hulman Undergrad. Math. J. **14** (2013), no. 1, 1–11.

[2] A. Komisarski, *On the behavior of Lebesgue integrable functions at infinity*, Indag. Math. (N.S.) **28** (2017), no. 2, 580–588.

[3] E. Lesigne, *On the behavior at infinity of an integrable function*, Amer. Math. Monthly **117** (2010), no. 2, 175–181.

[4] M.V. Mihai, *A remark on the behavior of integrable functions at infinity*, An. Univ. Craiova Ser. Mat. Inform. **38** (2011), no. 4, 100–101.

[5] C.P. Niculescu, F. Popovici, *A note on the behavior of integrable functions at infinity*, J. Math. Anal. Appl. **381** (2011), no. 2, 742–747.

[6] C.P. Niculescu, F. Popovici, *The behavior at infinity of an integrable function*, Expo. Math. **30** (2012), no. 3, 277–282.

[7] C.P. Niculescu, F. Popovici, *The asymptotic behavior of integrable functions*, Real Anal. Exchange **38** (2012/13), no. 1, 157–167.

[8] C.P. Niculescu, G.T. Prăjitură, *Some open problems concerning the convergence of positive series*, Ann. Acad. Rom. Sci. Ser. Math. Appl. **6** (2014), no. 1, 92–107.

[9] M. Louis Olivier, *Remarques sur les séries infinies et leur convergence*, J. Reine Angew. Math. **2** (1827), 31–44.

[10] T. Šalát, V. Toma, *A classical Olivier's theorem and statistical convergence*, Ann. Math. Blaise Pascal **10** (2003), no. 2, 305–313.

University of Łódź
joanna.horbaczewska@edu.uni.lodz.pl
asiahorb@gmail.com

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

NORMS ON CATEGORIES AND ANALOGS OF THE SCHRÖDER-BERNSTEIN THEOREM

MATT INSALL & DANIEL LUCKHARDT

Abstract. We generalize the concept of a norm on a vector space to one of a norm on a category. This provides a unified perspective on many specific matters in many different areas of mathematics like set theory, functional analysis, measure theory, topology, and metric space theory. We will especially address the two last areas in which the monotone-light factorization and, respectively, the Gromov-Hausdorff distance will naturally appear. In our formalization a Schröder-Bernstein property becomes an axiom of a norm which constitutes interesting properties of the categories in question. The proposed concept provides a convenient framework for metrizations.

Keywords: norm on categories, light map, monotone map, metric space, dilatation, expansion, Gromov-Hausdorff distance, Schröder-Bernstein theorem

MSC2010: 51F99, 18F99

Received 05 May 2021; revised 28 July 2021; accepted 30 July 2021

1. Introduction

The subject of this paper is the generalization of the concept of a norm on a vector space or, more generally, on an abelian group to one on a category. Our attempt is preceded by many approaches using enriched monoidal categories going back to Lawvere [20] and an industry arising out of this [13, 24]. Independently, Ghez, Lima, and Roberts [11] suggested a notion of normed *-category. Recently, interleaving distance has gained some favour as a playing ground [27, 1]. Another recent approach outside of enriched or monoidal category theory is by Kubiś [18].

Our motivation in the present work is threefold, aspiring a framework

(1) to find analogs of the Cantor-Schröder-Bernstein theorem (CSB theorem) from set theory in other categories,

https://doi.org/10.2478/9788366675360-012.

(2) for systematic and convenient metrization of families of equivalence classes of spaces, like the Gromov-Hausdorff space, moduli spaces, and representation spaces,
(3) to work with categories with large classes of morphisms.

The guiding example of our approach is the following norm on the category of sets: To each function $f\colon X \to Y$ assign the non-negative extended real number

$$\|f\|_{\mathrm{inj}} \coloneqq \sup_{x\in X}{}^{0} \log(\#\{\, y \mid f(x) = f(y)\,\}) \tag{1.1}$$

where $\#$ assigns to a set the numbers of its elements (a member of $\{0, 1, \dots, \infty\}$) and $\sup_{x\in X}^{a} f(x) = \sup\{a\} \cup \{\, f(x) \mid x \in X\,\}$. Note that f is injective if and only if $\|f\|_{\mathrm{inj}} = 0$. This is to say that $\|\,.\,\|_{\mathrm{inj}}$ is measuring the deviation from being injective. Hence the idea for the generalization of the CSB theorem is that in a normed category $(\underline{C}, \|\,.\,\|)$ two objects X and Y are isomorphic as soon as there are morphisms $f\colon X \to Y$ and $g\colon Y \to X$ such that $\|X\| = \|Y\| = 0$.

Returning to our list of motivations, we mention with regard to the first one that the CSB theorem is a fundamental theorem in set theory stating that there is a bijection between two sets as soon as there are injective maps between the sets both ways. In conceptual terms it states that if X can be embedded into Y and vice-versa, then X and Y are isomorphic. A direct formalization of this conceptual idea in category theoretic terms would be the property that in a category $\underline{C}$ two objects $X, Y \in \underline{C}_0$ are isomorphic as soon as there are monomorphisms $X \to Y$ and $Y \to X$. But, unfortunately, most categories considered in practice do not have this property, cf. [19]. Though, there are some notable exceptions, which include measure spaces [28, § 3.3] and a noncommutative version thereof for von Neumann algebras [17, Proposition 6.2.4]. Efforts to generalize the CBS theorem in alternative set-ups have recently revived, including results for categories of universal algebras [10] and in homotopy type theory and boolean ∞-topoi [8] including a formalization in Agda [7]. Of course, one can also weaken the property by replacing "monomorphism" by some stronger or related notion of morphism. In our approach it will be a morphism with vanishing norm.

As for the second motivation, the problems with doing metrizations in practise are often that they become very technical, involve arbitrary choices, and basic properties like the triangle inequality or completeness become hard to check. The theory of uniform structure can be seen as an attempt to abstract from these choices, but it lacks a measure of the size of entourages. In many examples we present it will turn out that the norms can be defined in term of a discrepancy function, a quantity measuring the size of a subobject of an object $X \in \underline{C}_0$. A category theoretical approach is natural considering the example of moduli spaces or representation spaces: representatives of the point (i.e. equivalence classes of spaces) are objects of a category and morphisms are comparison maps.

As for the last motivation, note that when working in some area of mathematics using the language of category theory one often has to limit the class of morphisms under consideration. For instance the category of metric spaces $\underline{\text{Met}}$ is normally defined with morphisms to be non-expansive map in order to guarantee nice properties like existence of limits.

This paper is mainly devoted to the first motivation, leaving the others to be explored by future work. It will turn out that our interpretation of Schröder-Bernstein holds normally on a full subcategory of compact spaces.

1.1. **Notation.** Per common practice, we will typically use the notation of Zermelo-Fraenkel Set Theory with the Axiom of Choice (ZFC), but in many cases, especially when we need to work with proper classes, we will actually use the relatively consistent extension referred to generally as GBN (Gödel-Bernaise-von Neumann) Class-Set Theory with the Axiom of Choice. Foundationally, we could use instead category theoretic foundations, but that seems to us to be merely a matter of taste. Accordingly, we leave it to the readers to adjust our recipes and seasonings for the dishes we describe to their preferences.

Given a set X, we denote by $\mathcal{P}(X)$ the power set of X, i.e. the set of all subsets of X. We will find it convenient to have a systematic notation for image and preimage mappings on the power set of a set, but in fact, for more variants than merely the simplest pair of such, whence, we denote by $f_!(A)$, or just $f_!A$ the set $\{ f(x) \mid x \in A \}$, where A is a subset of the domain of f, denote the preimage of any set B under f by $f^*(B)$ or f^*B.

We denote partial orders by $\mathcal{X} = (X, \leq), \mathcal{Y} = (Y, \sqsubseteq)$, and $\mathcal{Z} = (Z, \subseteq)$. Let $\mathcal{L} = (L, \leq)$ be a complete lattice, i.e. a poset that admits all suprema (and consequently, all infima). We define $\sup\colon L \times \mathcal{P}(L) \to L$ by

$$\sup(m, M) \coloneqq \sup\big(M \cap \{ m' \mid m \leq m' \}\big) \cup \{m\},$$

and then for any $m \in L$, set

$$\sup^m M \coloneqq \sup(m, M)$$

for any subset M of L. Moreover, if I is any (indexing) set, then

$$\sup_P^m f \coloneqq \sup^m \{ f(i) \mid i \in P \},$$

for any subset P of I and function f on I. If not specified otherwise, sup is understood as the supremum function on the extended real numbers $[-\infty, \infty]$.

1.2. **Acknowledgements.** The authors thank Paolo Perrone for his hints on the literature, and to thank also anonymous referees for very helpful feedback.

2. Definitions

Definition 2.1. A **seminorm** on a category $\underline{C}$ is a function $\| \,.\, \| \colon \underline{C}_1 \to [0, \infty]$ such that

(N1) $\|\mathrm{id}_X\| = 0$ for every object $X \in \underline{C}_0$, (N2) $\|f\,;g\| \leq \|f\| + \|g\|$ (triangle inequality).

The tuple $(\underline{C}, \|\,.\,\|)$ is called a **seminormed category**.

Note that in the literature a seminorm is often called a norm [20, 13]. Moreover, note that we do not require the obvious strengthening of (N1), namely that the seminorm of every categorical isomorphism vanishes. An explanation how to view this as a generalization of a seminorm on a vector space is found in Subsection 3.4. An isomorphism $f\colon X \to Y$ with inverse $g\colon Y \to X$ is called a **norm isomorphism** if $\|f\| = \|g\| = 0$. By (N2) being norm isomorphic is an equivalence relation. Moreover any morphism with norm zero is called a **modulator**. Often the category with objects $\underline{C}_0$ and all modulators of $\underline{C}_1$ as morphisms has good categorical properties. We will denoted it by $\boxed{M}(\underline{C}, \|\,.\,\|)$. Two seminormed categories are called **isomorphic** if there is a norm preserving categorical isomorphism between them. A seminorm or norm, respectively, induces a seminorm or norm, resp., on the opposite category in the obvious way.

Definition 2.2. A seminorm is called a **norm** if for all objects X, Y the following holds

(N3) if there are modulators $f\colon X \to Y$ and $g\colon Y \to X$, then X and Y are norm isomorphic; and

(N4) if for all $\varepsilon > 0$ there is $f\colon X \to Y$ with $\|f\| \leq \varepsilon$, then there is an modulator $f\colon X \to Y$.

The way to view (N3) is that a form of CSB theorem holds. The moral idea is that $\|f\| = 0$ is a property that is stronger than being monic and $\|\,.\,\|$ measures the deviation from this property. Let $(\underline{C}, \|\,.\,\|)$ be a seminormed category. We define the **left dual seminorm** as

$$\|f\|^{*\mathrm{L}} := \sup_{f'}{}^{0} \left(\|f'\| - \|f'\,;f\|\right), \tag{2.1a}$$

where $X' \xrightarrow{f'} X \xrightarrow{f} Y$, and the **right dual seminorm** as

$$\|f\|^{*\mathrm{R}} := \sup_{f''}{}^{0} \left(\|f''\| - \|f\,;f''\|\right) \tag{2.1b}$$

where $X \xrightarrow{f} Y \xrightarrow{f''} Y'$. The seminorm $\|\,.\,\|$ is called **left reflexive** if $\|\,.\,\|^{*\mathrm{L}*\mathrm{L}} = \|\,.\,\|$ and **right reflexive** if $\|\,.\,\|^{*\mathrm{R}*\mathrm{R}} = \|\,.\,\|$. As opposed to the case of normed spaces, the dual in our case does not define an entirely new category but merely a new norm on the same category.

To check that the left dual and right dual seminorms are actually seminorms observe for (N1) that $\|\mathrm{id}\|^{*\mathrm{L}} = \sup^0\|f'\| - \|f'\,;\mathrm{id}\| = 0 = \sup^0\|f'\| - \|\mathrm{id}\,;f'\| = \|\mathrm{id}\|^{*\mathrm{R}}$. We show that (N2) holds for left duals and then apply that in the opposite category to show that it holds for right duals. To this end, observe that for any diagram $X \xrightarrow{f} Y \xrightarrow{g} Z$

$$\|f\,;g\|^{*\mathrm{L}} = \sup_{h'}{}^{0}\|h'\| - \|h'\,;f\,;g\|$$

$$\begin{aligned}
&= \sup_{h'}{}^{0} \|h' \,;\, f\| - \|h' \,;\, f \,;\, g\| + \|h'\| - \|h' \,;\, f\| \\
&\leq \sup_{h'}{}^{0} \|h'\| - \|h' \,;\, f\| + \sup_{h'}{}^{0} \|h'\| - \|h' \,;\, g\| \\
&= \|f\|^{*\mathrm{L}} + \|g\|^{*\mathrm{L}}.
\end{aligned}$$

These arguments transfer to the right dual by the fact that the seminorm induced by the right dual on the opposite category coincides with the left dual of norm induced on the opposite category by the original norm.

Remark 2.3. Kubiś [18] defines a norm in our terminology as a seminorm $\|\,.\,\|\colon \underline{C}_0 \to [0,\infty]$ such that $\|\,.\,\|^{*\mathrm{L}} \leq \|\,.\,\|$. He defines a completion of a category with respect to such a norm and proves a version of Banach's fixed point theorem in this set-up.

On the class of norm isomorphism classes of objects of $\underline{C}$, $\mathrm{sk}_0(\underline{C}, \|\,.\,\|)$,[1] we define the **pqmetric** or **pseudoquasimetric induced by** $\|\,.\,\|$

$$d_{\|\,.\,\|}(\hat{X}, \hat{Y}) \coloneqq \inf\{\, \|f\| \mid f \in \underline{C}[X,Y] \text{ for } X \in \hat{X}, Y \in \hat{Y} \,\}. \tag{2.2}$$

Note that by (N2) for the computation of $d_{\|\,.\,\|}(\hat{X}, \hat{Y})$ it is sufficient to look at fixed representatives of $\hat{X}$ and $\hat{Y}$. Observe further that this is indeed a pqmetric since $d_{\|\,.\,\|}(X,X) \leq \|\mathrm{id}_X\| = 0$ and the triangle inequality holds by

$$\begin{aligned}
d_{\|\,.\,\|}(X,Z) &= \inf\{\, \|f\| \mid f\colon X \to Z \,\} \\
&\leq \inf\{\, \|f_1 \,;\, f_2\| \mid f_1\colon X \to Y, f_2\colon Y \to Z \,\} \\
&\leq \inf\{\, \|f_1\| + \|f_2\| \mid f_1\colon X \to Y, f_2\colon Y \to Z \,\} \\
&= d_{\|\,.\,\|}(X,Y) + d_{\|\,.\,\|}(Y,Z)
\end{aligned}$$

for all $X, Y, Z \in \underline{C}_0$. Symmetrizing in some way gives a pseudometric, e.g. by

$$d^{\vee}_{\|\,.\,\|}(X,Y) \coloneqq d_{\|\,.\,\|}(X,Y) \vee d_{\|\,.\,\|}(Y,X) \tag{2.3a}$$

$$d^{+}_{\|\,.\,\|}(X,Y) \coloneqq \frac{1}{2}\left(d_{\|\,.\,\|}(X,Y) + d_{\|\,.\,\|}(Y,X)\right) \tag{2.3b}$$

$$d^{p}_{\|\,.\,\|}(X,Y) \coloneqq \sqrt[p]{d_{\|\,.\,\|}(X,Y)^p + d_{\|\,.\,\|}(Y,X)^p} \tag{2.3c}$$

for $p \in [1,\infty)$. If $\|\,.\,\|$ is actually a norm, so $d_{\|\,.\,\|}$ is a metric.

3. Canonical examples

3.1. **Sets.** On the category $\underline{\mathrm{SET}}$ of sets we define for a function $f\colon X \to Y$ the norm $\|\,.\,\|_{\mathrm{set}}$ measuring the deviation of a function from being injective: we set

$$\|f\|_{\mathrm{set}} = \log \sup_{y\in Y}{}^{1} \#f^*(\{y\}). \tag{3.1}$$

[1] Note the foundational remark in the introduction. In many examples $\mathrm{sk}_0(\underline{C}, \|\,.\,\|)$ admits a set of representatives.

We check that $\|\,.\,\|_{\text{set}}$ is a norm. For the seminorm properties observe that $\|\mathrm{id}_X\|_{\text{set}} = 0$ for any set X. Moreover the triangle inequality is satisfied as it holds trivially whenever $\|f\|_{\text{set}} = \infty$ or $\|g\|_{\text{set}} = \infty$ and otherwise—using (3.1)—

$$\begin{aligned}\|f\,;g\|_{\text{set}} &= \log \sup_{z\in Z}{}^{1} \#(f\,;g)^*\{z\} \\ &= \log \sup_{z\in Z}{}^{1} \# \bigcup \{\, f^*(\{y\}) \mid y \in g^*\{z\}\,\} \\ &\leq \log \left(\sup_{y\in Y}{}^{1} \# f^*(\{y\}) \cdot \sup_{z\in Z}{}^{1} \# g^*\{z\} \right) \\ &= \|f\|_{\text{set}} + \|g\|_{\text{set}}.\end{aligned}$$

Hence $\|\,.\,\|_{\text{set}}$ is a seminorm. We continue with the norm properties: $\|f\|_{\text{set}} = 0$ is to say that f is injective. Property (N3) is exactly the Schröder-Bernstein theorem. The property (N4) is trivial since the image of $\|\,.\,\|_{\text{set}}$ is discrete.

3.2. **Simplicial complexes.** We can use the same norm $\|\,.\,\|_{\text{set}}$, from the previous subsection, on simplicial complexes. Recall that an (unoriented) simplicial complex on a set V is a pair (V, X) where X is a subset of $\mathcal{P}(V)$ such that

(1) each $s \in X$ is finite and nonempty,
(2) $\{x\} \in X$ for each $x \in V$, and
(3) $s' \in X$ for all $s' \subseteq s \in X$ and $s' \neq \emptyset$.

The elements of V are called vertices. A set $s \in X$ is called simplex, and n-simplex if s has n elements. Further let X_n denote the set of all n-simplices in X. A simplicial complex is called finite whenever it is finite as a set. A morphism of simplicial complexes is a function $f\colon V \to W$ such that $(f_!)_!(X) \subseteq Y$. Let <u>SimpCplx</u> denote the category of simplicial complexes and morphisms of simplicial complexes. Let <u>FinSimpCplx</u> denote the full subcategory of finite simplicial complexes, i.e. simplicial complexes with a finite set of vertices (or equivalently a finite set of simplices).

Proposition 3.1. *The seminormed category* (<u>FinSimpCplx</u>, $\|\,.\,\|_{\text{set}}$) *is normed.*

Proof. The axiom (N4) is trivial since the image of $\|\,.\,\|_{\text{set}}$ is discrete. For (N3) assume that (V, X), (W, Y) are two simplicial complexes and $f\colon V \to W$, $g\colon W \to V$ morphisms of simplicial complexes with $\|f\|_{\text{set}} = \|g\|_{\text{set}} = 0$. Thus f and g are injective. Hence $f_!$ and $g_!$ are injective. This is to say that both functions map n-simplices to n-simplices for every n. Hence both complexes have the same number of n-simplices for every n. Since for each n the set of simplices X_n is finite and $f_!$ is injective, the map $f_!$ is actually bijective. Hence both simplicial complexes are norm isomorphic. □

3.3. **Cost functions and polarization.** Let c be a **cost function** on M, i.e., a map $c\colon \operatorname{Conf}_2 M \to [0, \infty]$ where the configuration space $\operatorname{Conf}_2(M)$ is defined as $\{\, (x, y) \in M^{\times 2} \mid x \neq y \,\}$. Cost functions form the corner stone of transportation theory [29]. Remember that a square-free word is a word in which no pattern of the form xx occurs. Define the normed category $(\underline{\mathcal{M}}, \|\,.\,\|_c)$ by

$$\begin{aligned}
\underline{\mathcal{M}}_0 &:= M,\\
\underline{\mathcal{M}}_1 &:= \{\text{square-free words over the alphabet } M\},\\
\underline{\mathrm{M}}[x,y] &:= \{\, w \in \underline{\mathcal{M}}_1 \mid w \text{ starts with } x \text{ and ends with } y \,\}\\
(x\xi_1 \dots \xi_n y)\,;(y\eta_1 \dots \eta_m z) &:= (x\xi_1 \dots \xi_n y \eta_1 \dots \eta_m z)\\
\mathrm{id}_x &:= (x),\\
\|(\xi_1 \dots \xi_n)\|_c &:= c(\xi_1,\xi_2) + \dots + c(\xi_{n-1},\xi_n)
\end{aligned}$$

for $(\xi_1 \dots \xi_n)\colon x \to y$. This obviously defines a seminorm since the triangle inequality (N2) is even an equality.

Moreover this construction induces a qpmetric on M, namely $d_c := d_{\|\,.\,\|_c}$. This is automatically a pmetric since any path from x to y can be transformed into a morphism from y to x of the same length by reversing. We have $d_c \le c$ and equality holds if and only if the extension of c to $M \times M$ by 0 is a metric. In other words, d_c is the largest pseudometric bounded by c.

Exercise 3.2a. Assume that $\mathcal{M} = (M, d_{\mathcal{M}})$ is a metric space. Show that $(\underline{M}, \|\,.\,\|_c)$ with $c(x,y) = d_{\mathcal{M}}(x,y)$ is a normed category

Exercise 3.2b. Let c be a cost function. Show that $\|\,.\,\|_c^{*\mathrm{L}} = \|\,.\,\|_c^{*\mathrm{R}} = 0$.

3.4. **Grothendieck norm for monoids.** Let $M = (M, +, 0)$ be a (not necessarily commutative) monoid. Further let $\|\,.\,\|_M$ be a seminorm on M, i.e. a map $M \to [0,\infty]$ such that $\|a+b\| \le \|a\| + \|b\|$. In the spirit of the previous example we define the category $\underline{M}$ and the **Grothendieck seminorm** $\|\,.\,\|$ on $\underline{C}$ by

$$\begin{aligned}
\underline{M}_0 &:= M\\
\underline{M}[a,b] &:= \{\, (f_+, f_-) \in M^{\times 2} \mid f_+ + a = b + f_- \,\}\\
(f_+, f_-)\,;(g_+, g_-) &:= (g_+ + f_+, g_- + f_-)
\end{aligned}$$

for $(f_+, f_-)\colon a \to b$ and $(g_+, g_-)\colon b \to c$ (well-defined since $g_+ + f_+ + a = g_+ + b + f_- = c + g_- + f_-$)

$$\begin{aligned}
\mathrm{id}_a &:= (0,0) \qquad \text{for all } a \in \underline{M}_0\\
\|(f_+, f_-)\| &:= \|f_+\|_M + \|f_-\|_M.
\end{aligned}$$

This defines indeed a seminorm since $\|(f_+, f_-);(g_+, g_-)\| = \|(g_+ + f_+, g_- + f_-)\| \le \|g_+\|_M + \|f_+\|_M + \|g_-\|_M + \|f_-\|_M = \|(f_+, f_-)\| + \|(g_+, g_-)\|$.

Exercise 3.3. Assume that $\|\,.\,\|_M$ is a norm (i.e. $\|a\|_M = 0$ if and only if $a = 0$). Prove that $(\underline{M}, \|\,.\,\|)$ is a normed category.

Proposition 3.4. *Assume that G is a group with a seminorm $\|\,.\,\|_G$. Then*

(i) *a normed category canonically isomorphic to* $(\underline{M}, \| \,.\, \|)$ *is defined by*

$$\begin{aligned} \underline{G}'_0 &:= G, \\ \underline{G}'[a,b] &:= G, \\ f \,;' g &:= g - b + f, && \text{for } f\colon a \to b \text{ and } g\colon b \to c, \\ \|f\|' &:= \|f\|_G + \|-b + f + a\|_G && \text{for } f\colon b \to c; \end{aligned}$$

(ii) *assuming that* $\|a\|_M = \|-a\|_M$, *we have* $d_{\| \,.\, \|'}(a,b) = \|-b + a\|_G$;

(iii) *under the same assumption we have* $\| \,.\, \|^{*\mathrm{L}} = \| \,.\, \|^{*\mathrm{R}} = \| \,.\, \|$.

Exercise 3.5. Prove this theorem.

Claim (ii) of Proposition 3.4 is the motivation to consider a seminorm on a category a generalization of a seminorm on a vector space where the underlying abelian group takes the role of G.

Example 3.6 (Word metric on a group). As an example of a norm on a monoid take generating set $\{g_i\}_{i \in I}$ of a group $G = (G, +, -(\,.\,), 0)$. Then the free monoid M generated by $\bigcup_{i \in I}\{g_i, -g_i\}$ has the evaluation function $\mathrm{ev}\colon M \to G$. On M there is a norm given on $m \in M$ by the minimal word length of some word w over the alphabet M such that $\mathrm{ev}(w) = m$. Word metrics are a fundamental tool in geometric group theory [21].

3.5. **Operators and normed vector spaces.** Let $\underline{\mathrm{NVect}}_{\mathbb{R}}$ denote the category of normed vector spaces over the reals with linear maps as morphisms. Define the norm of a linear map $A\colon V \to W$ as

$$\|A\|_{\mathrm{op}} := \log \sup_{v \in V}{}^1 \frac{\|v\|_V}{\|Av\|_W}.$$

The pair $(\underline{\mathrm{NVect}}_{\mathbb{R}}, \|.\|)$ is a seminormed category. Being norm isomorphic is equivalent to being isometric as linear normed spaces.

It is not a normed category and especially does not satisfy the CSB axiom (N3).

Exercise 3.7. Find an example that violates (N3). Hint: take the vector space $\mathrm{C}^0([0,1])$ of real valued continuous functions on the unit interval $[0,1]$ vanishing at 0 and 1 and endow this space with the supremum norm $\| \,.\, \|$. Define $V := \mathrm{C}^0([0,1])$ and $W := \mathrm{C}^0([0,1]) \oplus \{\, f \in \mathrm{C}^0([0,1]) \mid f \text{ smooth} \,\}$.

It's even not possible to ensure (N3) by restricting to the category of Banach space, i.e. complete normed vector spaces. Corresponding examples are more complicated but one was found in a celebrated result by Gowers [12]. On the other hand the fully faithful subcategory ${}^{\mathrm{Hilb}}\underline{\mathrm{NVect}}_{\mathbb{R}}$ of $\underline{\mathrm{NVect}}_{\mathbb{R}}$ consisting of Banach spaces that admit an inner product (i.e. admit the structure of a Hilbert space). Recall that the Hilbert dimension of a Hilbert space is defined as the cardinality of a basis. A basis is by definition a maximal orthonormal set $E \subset V$, i.e. $\langle e, e' \rangle = 0$ for all $e, e' \in E$ with $e \neq e'$ and $\|e\| = 1$ for all $e \in E$. Especially, E is linearly independent. Hilbert spaces are norm isomorphic if and only if they have the same Hilbert dimension [4, Theorem 5.4].

If there is an expansive operator $A\colon V \to W$ than the dimension of W is not smaller than the dimension of V: if E is a maximal orthonormal set in V then $A(E)$ is still linearly independent due to linearity and injectivity of A. Since W is a Hilbert space there is a decomposition $W \simeq \overline{A(V)} \oplus W'$. By the Gram-Schmidt process we can find a basis for $\overline{A(V)}$. Extending this basis to a basis of W be see that the dimension of W is not smaller than the dimension of V. Thus if we have expansive operators in both direction, V and W are of the same dimension and hence there is a norm isomorphism.

Exercise 3.8. Show that the left dual of $\|\,.\,\|_{\mathrm{op}}$ is the re-scaled operator norm $\log \sup^1{}_{v \in V} \frac{\|Av\|}{\|v\|}$ and that $\|\,.\,\|_{\mathrm{op}}$ is left reflexive.

4. Norms from precapacities

For concrete categories $\underline{C}$ most seminorms arise from a function on the subobjects of objects in $\underline{C}_0$—or an extension of this concept—valued in the extended real numbers, called precapacity. We will define precapacities as a function on subobjects of objects in $\underline{C}$, the category to be given a seminorm. Given an object X in a category $\underline{C}$ the slice category $\underline{C}/X$ is defined as the category with morphisms $B \in \underline{C}[Y, X]$ (for any $Y \in \underline{C}_0$) as objects and commuting diagrams

$$\begin{tikzcd} \operatorname{source} B \arrow[rr, "\varphi"] \arrow[dr, "B"'] & & \operatorname{source} C \arrow[dl, "C"] \\ & X & \end{tikzcd}$$

as morphisms. Composition is defined by composition of morphisms $\varphi\,;\psi$:

$$\begin{tikzcd} \operatorname{source} B \arrow[r, "\varphi"] \arrow[dr, "B"'] & \operatorname{source} C \arrow[r, "\psi"] \arrow[d, "C"] & \operatorname{source} D \arrow[dl, "D"] \\ & X & \end{tikzcd} .$$

We repeat the standard notion of subobjects from category theory [23, p. 11, 16, A.1.3]. A subobject of an object X in a category $\underline{C}$ is an equivalence class of monomorphisms to X, where equivalent means isomorphic in $\underline{C}/X$. We denote the set of such equivalence classes by $\mathrm{Sub}_0(X)$. Note that the composition of two monomorphisms is a monomorphism.

Exercise 4.1a. Prove that two objects $B, C \in (\underline{C}/X)_0$ are isomorphic if B, C are monomorphisms in $\underline{C}$ and there are morphisms $\varphi \in \underline{C}/X[B, C]$ and $\psi \in \underline{C}/X[C, B]$ such that φ and ψ are monomorphisms in $\underline{C}$.

Exercise 4.1b. Prove that $\mathrm{Sub}_0(X)$ becomes a partially ordered set by the relation $\subseteq$ where $B \subseteq C$ if and only if there is a morphism from B to C in $\underline{C}/X$ that is a monomorphism in $\underline{C}$.

We will denote by $\mathrm{Sub}(X)$ the set $\mathrm{Sub}_0(X)$ endowed with the partial order $\subseteq$. From category theory, the pullback[2] of a monomorphism C along a morphism f is a monomorphism again; it is denoted by f^*C:

$$\begin{tikzcd} \mathrm{source}(f^*C) \arrow[r] \arrow[d, "f^*C"] \arrow[dr, phantom, "\lrcorner", very near start] & \mathrm{source}\, C \arrow[d, "C"] \\ X \arrow[r, "f"] & Y \end{tikzcd} .$$

Definition 4.2. A **concrete category** $(\underline{C}, F)$ is a category $\underline{C}$ together a faithful functor $\underline{C} \to \underline{\textsc{Set}}$. The functor F is called **forgetful functor**.

Definition 4.3. By a **concrete category with generalized subobjects** $(\underline{C}, F; S, F_S, \mathrm{GS})$, which by abuse of notation we will often denote by $(\underline{C}; \mathrm{GS})$, we understand a concrete category $(\underline{C}, F)$ additionally endowed with an extension of the concrete category $(\underline{C}, F)$, meaning a commutative triangle of functors

$$\begin{tikzcd} \underline{SC} \arrow[dr, "F_S"] & \\ \underline{C} \arrow[u, "S"] \arrow[r, "F"] & \underline{\textsc{Set}} \end{tikzcd} ,$$

and a selection function GS that assigns to each object X a family of subobjects in $\mathrm{Sub}(X)$, called **(generalized) subobjects**, such that

(1) for each $X \in \underline{C}_0$ the order preserving induced functor

$$\mathrm{GS}(F_S, X)\colon \mathrm{Sub}(X) \to \mathrm{Sub}(F_S(X)), \quad C \mapsto (F_S)_1(C)$$

is well-defined, i.e. $F_S(C)$ is a monomorphism again, and full, i.e. $|B| \subseteq |C|$ implies $B \subseteq C$ where

$$|C| := (F_S C)(\mathrm{source}\, C) \subseteq F(X) \tag{4.1}$$

for any $C \in \mathrm{GS}(X)$ with $X \in \underline{C}_0$ (note that (4.1) is independent of the representative in $\underline{\textsc{Set}}_1$).

(2) if $f\colon X \to Y$ and $C \in \mathrm{GS}(Y)$, then there is a $B \in \mathrm{GS}(X)$ with $|B| = (Ff)^*(|C|)$, that is maximal in $\mathrm{GS}(Y)$ with this property. This generalized subobject is called the **preimage of** C **under** f and written $B = f^*(C)$.

The last axiom implies that each GS-subobject of $S(X)$ is mapped to a monomorphism. Often we will encounter the case $\underline{SC} = \underline{C}$ and $S = \mathrm{id}_{\underline{C}}$. By abuse of notation we will often write $B \in \mathrm{GS}(X)$ for a representative of B. Note that a concrete category with generalized subobjects has some similarity to a Grothendieck topology: think of a representative of an element of $\mathrm{GS}(X)$ as singleton set. Then one may compare the assignment GS to a Grothendieck topology, that assigns to each object a family of sieves.

[2]We denote pullback diagrams by $\begin{tikzcd} \cdot \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & \cdot \arrow[d] \\ \cdot \arrow[r] & \cdot \end{tikzcd}$.

An example of a concrete category with subobjects is given by

$$(\underline{\text{Top}}, F; \mathrm{id}_{\underline{\text{Top}}}, F, \mathcal{C}\ell) \tag{4.2}$$

where $\mathcal{C}\ell(\mathcal{X})$ is the collection of equivalence classes of homeomorphisms onto closed subspaces of $\mathcal{X}$ and F is the canonical forgetful functor $\underline{\text{Top}} \to \underline{\text{Set}}$: the property of having enough subobjects follows immediately from the fact that preimages of closed sets under continuous maps are closed.

Definition 4.4. A **precapacity** w on a concrete category with subobjects $(\underline{C}; \mathrm{GS})$ is a function

$$c \colon \bigsqcup_{X \in \underline{SC}_0} \mathrm{GS}(X) \to [-\infty, \infty]$$

and it is called a **capacity** if it is monotone or antimonotone, i.e. for any two subobjects $B, C \in \mathrm{GS}(X)$ with $B \subseteq C$ we have that $c(\operatorname{source} B) \leq c(\operatorname{source} C)$.

In practise, capacities are often non-negative. Each precapacity c gives rise to an assignment

$$\|f\|_c \coloneqq \sup{}^0 \left\{ c f^* C - c(C) \;\middle|\; \substack{C \in \mathrm{GS}(Y),\\ c(C) < \infty} \right\}. \tag{4.3}$$

For a concrete category with enough subobjects it is called the **seminorm induced by** w. The seminorm properties are checked immediately by

$$\|\mathrm{id}_X\|_c = \sup_C{}^0 |c(\mathrm{id}_X^*(C))| - |c(C)| = \sup_C{}^0 |c(C)| - |c(C)| = 0$$

and for any diagram $X \xrightarrow{f} Y \xrightarrow{g} Z$ in $\underline{C}$

$$\begin{aligned}
\|f\,;g\|_c &= \sup_C{}^0 \left(|c(f\,;g)^* C| - |c(C)|\right) \\
&= \sup_C{}^0 \left(|c(f\,;g)^* C| - |c(g^* C)| + |c(g^* C)| - |c(C)|\right) \\
&\leq \sup_C{}^0 \left(|c(f\,;g)^* C| - |c g^* C|\right) + \sup_C{}^0 \left(|c g^* C| - |c(C)|\right) \\
&= \|f\|_c + \|g\|_c.
\end{aligned}$$

4.1. **Example: normed category of vector spaces.** We will explain how to formulate the seminormed category $(\underline{\text{NVect}_{\mathbb{R}}}, \|\,.\,\|_{\mathrm{op}})$ from Subsection 3.5 in the framework of capacities explained above. Set

$$\begin{aligned}
(\underline{\text{SNVect}_{\mathbb{R}}})_0 &\coloneqq \{\, (F, V) \mid F \subseteq V, V \in \underline{\text{NVect}_{\mathbb{R}}} \,\} \\
\underline{\text{SNVect}_{\mathbb{R}}}[(F, V), (G, W)] &\coloneqq \{\, f \in \underline{\text{Set}}[F, G] \mid \exists A \in (\underline{\text{NVect}_{\mathbb{R}}})_1 \colon A|_F = f \,\} \\
S \colon \underline{\text{NVect}_{\mathbb{R}}} &\to \underline{\text{SNVect}_{\mathbb{R}}} \\
S(V) &\coloneqq (V, V) \\
F_S(F, V) &\coloneqq F \\
\mathrm{GS}(V) &\coloneqq \left\{ (F, V) \xrightarrow{(v,w) \mapsto (v,w)} (V, V) \;\middle|\; F \subseteq V \right\}
\end{aligned}$$

$$c(F,V) := \sup_{v\in F} \log\|v\|_V$$

where $V = (V, \|v\|_V)$. This actually defines a category with subobject as both properties 1 and 2 are obvious.

Exercise 4.5. Show that $\|A\|_c = \|A\|_{\mathrm{op}}$.

4.2. **Dual seminorms from capacities.** Let c be a precapacity. In view of (4.3) one is of course tempted to look at the quantity

$$\|f\|_{-c} := \sup{}^0\Big\{\, c(C) - cf^*C \;\Big|\; \substack{C\in\mathrm{GS}(Y),\\ c(C)<\infty} \Big\}. \tag{4.4}$$

Exercise 4.6a. Show that for any precapacity c it holds that $\|f\|_c^{*\mathrm{R}} \le \|f\|_{-c}$ and therefore $\|f\|_{-c}^{*\mathrm{L}} \ge \|f\|_c \ge \|f\|_{-c}^{*\mathrm{R}}$.

Exercise 4.6b. Moreover for the **left** and **right biduals**, i.e. the quantities $\|\,.\,\|^{*\mathrm{L}*\mathrm{L}}$ and $\|\,.\,\|^{*\mathrm{R}*\mathrm{R}}$, show that $\|\,.\,\|^{*\mathrm{L}*\mathrm{L}}, \|\,.\,\|^{*\mathrm{R}*\mathrm{R}} \le \|\,.\,\|$.

4.3. **Overview.** In this article we will study the following examples of precapacities and the corresponding induced norms. Note the conventions $\log 0 = -\infty$ and $|-\infty| = \infty$ and that topological spaces are denoted by $\mathcal{X} = (X, \tau_{\mathcal{X}})$.

(1) Topological dimension $C \mapsto |\log(1+\dim C)|$ for $\mathrm{GS}(X) = \mathcal{P}(X)$ in Section 5 giving rise to a seminorm measuring deviation of a map from being light.
(2) Logarithmic number of connected components of topological spaces $\mathcal{X}$

$$\mathrm{GS}(\mathcal{X}) = \mathcal{C}\ell(X) \quad \text{and} \quad |\log\#\mathrm{I}|\colon \mathcal{X} \mapsto |\log(\#(\mathrm{I}\,\mathcal{X}))|$$

where $\mathrm{I}\,\mathcal{X}$ is the set of connected components of $\mathcal{X}$. In Section 5 this gives rise to a seminorm measuring deviation of a map from being monotone.
(3) Diameter $\operatorname{diam}\mathcal{M}$ of a metric space $\mathcal{M}$ for $\mathrm{GS}(\mathcal{M}) = \mathcal{P}(\mathcal{M})$ giving rise to the dilatation norm, that quasi-metrizes Gromov-Hausdorff convergence.

There are other capacities that may be part of future research:

(4) Negative logarithmic diameter $-\log\operatorname{diam}\mathcal{M} = -c_{\mathrm{Lip}}$ giving rise to the Lipschitz norm $\|\,.\,\|_{-c_{\mathrm{Lip}}}$, that quasi-metrizes Lipschitz convergence of compact metric spaces.
(5) The Krull dimension $\dim_{\mathrm{Krull}}$ giving rise to $c_{\mathrm{Krull}}(C) := |\dim_{\mathrm{Krull}} C|$. Recall that the Krull dimension $\dim_{\mathrm{Krull}}\mathcal{X}$ of a topological space $\mathcal{X}$ is the maximal length of a chain of irreducible closed sets, i.e. if there are closed irreducible sets $C_0 \subset C_1 \subset \ldots \subset C_m \subset X$, then $\dim_{\mathrm{Krull}}\mathcal{X}$ is at least m. If there is no proper closed irreducible subset, the dimension is $-\infty$.

5. Topological spaces

The subject of this section are two seminorms both measuring the increase of complexity when passing from a subset of the domain to its preimage. In the first case complexity is measured by topological dimension in the second one by number of connected components. These norms measure the deviation from being light and monotone respectively. The well-known monotone-light factorization implies that in the case of compact spaces the sum of both norms, which we call the topological norm $\|\,.\,\|_{\mathrm{top}}$, is a norm. Actually, the monotone-light factorization is a strengthening of the norm property (N3). The idea of the monotone-light factorization goes back to Eilenberg [6] and Whyburn [30] independently. For a historical overview about the monotone-light factorization and its variations consult [22].

5.1. **Dimension seminorm.** Recall that a map $f\colon \mathcal{X} \to \mathcal{Y}$ is **light** if the fiber $f^*y := f^*(\{y\})$ is totally disconnected for every $y \in Y$. Further we define the **dimension seminorm** using the precapacity $A \mapsto |\log(1 + \dim(A))|$ by

$$\begin{aligned}\|f\|_{\dim} &:= \|f\|_{|\log(1+\dim)|}\\ &= \sup_{\substack{A\in\mathcal{P}(Y),\\ \dim A<\infty}}{}^{0}\; |\log(1 + \dim f^*A)| - |\log(1 + \dim A)|\end{aligned}$$

where we used 2 of Definition 4.3 in the last step.

To further simplify this expression we use the Hurewicz formula which states

$$\dim \mathcal{X} \le \dim \mathcal{Y} + \sup_{y\in Y}|\dim(f^*\{y\})| \tag{5.1}$$

[25, Ch. 9, Prop. 2.6, 15, § VI.4] for any continuous closed surjection f from a T_4-space $\mathcal{X}$ to a metrizable space $\mathcal{Y}$. Use this for:

Exercise 5.1. Prove that for a map f from a T_4-space $\mathcal{X}$ to a metrizable space $\mathcal{Y}$

$$\|f\|_{\dim} = \sup_{y\in Y}|\log(1 + \dim(f^*\{y\}))|, \quad \text{and} \tag{5.2a}$$

$$\dim \mathcal{X} \le \mathcal{Y} + \exp\|f\|_{\dim}. \tag{5.2b}$$

5.2. **Component seminorm.** Following Whyburn [31] and Carboni et al. [3] we call a map $f\colon \mathcal{X} \to \mathcal{Y}$ **monotone** if the preimage of every singleton $\{y\} \subset \mathcal{Y}$ is nonempty and connected. Note that this property implies surjectivity since the empty set consists of zero connected components. For any topological space $\mathcal{X}$ let $\mathcal{I}(\mathcal{X}) = (\mathrm{I}(\mathcal{X}), \tau_{\mathcal{I}(\mathcal{X})})$ denote the collection of connected components of $\mathcal{X}$ endowed with the quotient topology. Define the **component seminorm** as

$$\begin{aligned}\|f\|_{\mathrm{comp}} &:= \|f\|_{|\log\#I|}\\ &= \sup{}^{0}\Big\{\, |\log(\#(\mathrm{I}\, f^*C))| - \log(\#(\mathrm{I}\, C)) \;\Big|\; {}_{|\log\#(\mathrm{I}\,C)|<\infty}^{C\subseteq Y,} \Big\}\\ &= \sup{}^{0}\Big\{\, |\log(\#(\mathrm{I}\, f^*C))| - \log(\#(\mathrm{I}\, C)) \;\Big|\; {}_{0<\#(\mathrm{I}\,C)<\infty}^{C\subseteq Y,} \Big\}\end{aligned}$$

where we used the convention $\log(0) = -\infty$. Note that we can express the number of connected components $\#I\mathcal{X}$ of a nonempty space $\mathcal{X}$ as $\exp\|\mathcal{X} \to \{*\}\|_{\mathrm{comp}}$ where $\mathcal{X} \to \{*\}$ is the canonical map to the singleton space. This norm relates to the dimension of fibers by the obvious inequality

$$\|f\|_{\text{comp}} \geq \sup_{p\in Y}{}^{0} |\log(\#(\mathrm{I}\, f^*\{p\})| =: \mathrm{mon}(f). \tag{5.3}$$

Lemma 5.2. *Let $f\colon \mathcal{X} \to \mathcal{Y}$ be a continuous function. Then*

(i) *We have* $\|f\|_{\text{comp}} = \sup{}^0 \left\{ |\log(\#(\mathrm{I}(f^*C)))| \;\middle|\; {\scriptstyle \emptyset \neq C \subseteq Y, \atop \# \mathrm{I}\, C = 1} \right\}$.

(ii) *The map f is monotone if and only if* $\mathrm{mon}(f) = 0$.

(iii) *If f is closed and monotone then* $\|f\|_{\text{comp}} = 0$.

(iv) *Assume that $\mathcal{X}$ is compact and that $\mathcal{Y}$ is Hausdorff. Then f is monotone, if and only if* $\|f\|_{\text{comp}} = 0$.

Exercise 5.3. Prove claims (i) and (ii) of Lemma 5.2.

Proof. Claims (i) and (ii) are already proved by the exercise. For claim (iii) assume that f is monotone and closed. Note that the restriction $f|_{f^*C}\colon f^*C \to C$ is closed as well: take any relatively closed $A \subseteq f^*C$. Any point x in the closure of $f_!(A)$ relative to C must be in the image $f_!(\overline{A}^{\mathcal{X}})$ of the closure of C in X, but then already $y \in A$ for any $y \in f^*x$. Thus $f_!(A) = f_!(\overline{A}^{\mathcal{X}})$.

Let C be an arbitrary closed connected subset of Y. The preimage f^*C must not be empty because otherwise monotonicity of f would be violated. Since C is closed, so is f^*C. Assume that f^*C is a disjoint union of two sets K and L that are clopen in the relative topology on f^*C. For any $y \in C$ the preimage $f^*\{y\}$ is connected in the relative topology. Hence either $f^*\{y\} \subseteq K$ or $f^*\{y\} \subseteq L$. Since this holds for all $y \in C$, the set C is actually the disjoint union of $f_!K$ and $f_!L$. Due to our observation on the closedness of $f|_{f^*C}\colon f^*C \to C$ both $f_!K$ and $f_!L$ are open in the relative topology. Thus by connectedness of C either $f_!K = \emptyset$ or $f_!L = \emptyset$, a contradiction. Consequently, f^*C is connected. Hence $\|f\|_{\text{comp}} = 0$.

In claim (iv) the direction monotone"$\ldots \implies \|f\|_{\text{comp}} = 0$" follows from claim (iii) and the fact that a continuous function from a compact space to a Hausdorff space is closed. The other direction is implied by (5.3) and claim (ii). ☐

Exercise 5.4. Find an example of a monotone map f from a compact space into a non-Hausdorff space such that $\|f\|_{\text{comp}} > 0$.

Theorem 5.5. *A map $f\colon \mathcal{X} \to \mathcal{Y}$ between totally disconnected compact Hausdorff spaces having $\|f\|_{\text{comp}} = 0$ is a homeomorphism.*

Proof. Assume that $\|f\|_{\text{comp}} = 0$. Then this map is surjective. Since every fiber is totally disconnected, $\|f\|_{\text{comp}} = 0$ implies that each fiber is a singleton. Hence f is bijective. As for compact Hausdorff spaces the notions of closed and compact subsets coincide, and compact subsets are mapped to compact subsets under continuous function, the inverse of f is continuous. Thus f is a homeomorphism. ☐

5.3. **The topological norm.** Define the **topological norm** as

$$\|f\|_{\text{top}} := \|f\|_{\text{comp}} + \|f\|_{\dim}.$$

Proposition 5.6. *Let $\mathcal{X}$ be a compact* T_4 *space and $\mathcal{Y}$ be metrizable. If $\|f\|_{\mathrm{top}} = 0$ for a continuous function $\mathcal{X} \to \mathcal{Y}$, then f is monotone and light.*

Proof. Assume that $\|f\|_{\mathrm{top}} = 0$. Then $\|f\|_{\mathrm{comp}} = \|f\|_{\mathrm{dim}} = 0$. The fact $\|f\|_{\mathrm{comp}} = 0$ implies that f is monotone because points in $\mathcal{Y}$ are closed. The other fact $\|f\|_{\mathrm{dim}} = 0$ implies that f is light by claim (iv) of Lemma 5.2. □

Theorem 5.7. *Let $\mathcal{X}$ be a compact* T_4 *space and $\mathcal{Y}$ be metrizable. If there is map with $\|f\|_{\mathrm{top}} = 0$, then f is a homeomorphism. Especially, the category of compact metrizable spaces is a normed category with respect to $\|\,.\,\|_{\mathrm{top}}$.*

Proof. By Proposition 5.6 the map f is monotone and light. Since the identity on a topological space is monotone and light as well, we have two factorizations of f in a monotone and a light map:

$$\begin{tikzcd}
 & \mathcal{X} \arrow[dr, "f"] \arrow[dd, dashed, "\varphi"] & \\
\mathcal{X} \arrow[ur, "\mathrm{id}"] \arrow[dr, "f"'] & & \mathcal{Y} \\
 & \mathcal{Y} \arrow[ur, "\mathrm{id}"'] &
\end{tikzcd}$$

By the classical uniqueness of the monotone-light factorization [3, 2.8, 7.3] there is a homeomorphism $\varphi\colon \mathcal{X} \to \mathcal{Y}$. □

6. Metric spaces

Let $\underline{\mathrm{Met}}$ denote the category of metric spaces with multi-valued functions among them as morphisms, i.e. functions $f\colon M \to \mathcal{P}(N) \setminus \{\emptyset\}$, $x \mapsto f[x]$. By the notation $\varphi(f(x))$ for some formula φ we mean $\exists y \in f[x]\colon \varphi(y)$, e.g. $y = f(x)$ means $y \in f[x]$. Set builder notation has to be understood as $\{\, f(x) \mid x \in M \,\} = \{\, y \mid y = f(x), x \in M \,\}$. This also explains how suprema and infima are to be understood. The image and preimage functions as defined as $f_!(A) = \bigcup_{x \in A} f(x)$ and $f^*(A) = \{\, x \mid f(x) \cap A \neq \emptyset \,\}$, respectively. Objects of $\underline{\mathrm{Met}}$ we denote by curly letters, e.g. $\mathcal{M} = (M, d_{\mathcal{M}})$, and the metric is abbreviated by $|\,.\ \,.\,| = d_{\mathcal{M}}(\,.\,,\,.\,)$ if no confusion can arise. Moreover let $\underline{\mathrm{Met}}_{\mathrm{cpt}}$ denote the full subcategory of compact metric spaces.

If a function is only densely defined, i.e. $f\colon M \to \mathcal{P}(N)$ and points with $f[x] \neq \emptyset$ are dense, then it is transformed into a morphism of $\underline{\mathrm{Met}}$ by the **closure**

$$\overline{f}(x) := \left\{ y \;\middle|\; y = \lim_{n\to\infty} y_n \text{ for } y_n = f(x_n) \text{ with } x_n \xrightarrow{n\to\infty} x \right\}. \tag{6.1a}$$

Using the notion of precapacity from Section 4, we define the dilatation seminorm of any densely defined multi-valued function $f\colon M \to N$ via the precapacity

$$\operatorname{diam}(A) := \sup_{x,y\in A}{}^{0} |x\,y|, \tag{6.1b}$$

for $A \subseteq M$, as

$$\|f\|_{\mathrm{dil}} := \|f\|_{\mathrm{diam}} = \sup_{A \subseteq N}^{0} \left[\mathrm{diam}(f^* A) - \mathrm{diam}(A)\right]. \tag{6.1c}$$

Exercise 6.1. Let $f\colon M \to N$ be any densely defined multi-valued function. Show that

$$\|f\|_{\mathrm{dil}} = \inf\{\, \varepsilon > 0 \mid \text{for all } x, y \in M\colon\ |x\, y| \leq |f(x)\, f(y)| + \varepsilon \,\} \tag{6.2a}$$

$$= \sup_{x,y \in M}^{0} \left(|x\, y| - |f(x)\, f(y)|\right), \tag{6.2b}$$

$$\|\bar{f}\|_{\mathrm{dil}} = \|f\|_{\mathrm{dil}}. \tag{6.2c}$$

For $r > 0$ let $\mathcal{M}_r = (\{x, y\}, d_r)$ be the two point metric space such that $d_r(x, y) = r$. This definition may serve as a hint for the following exercises.

Exercise 6.2. Show that the seminorm $\|\,.\,\|_{\mathrm{dil}}$ is left reflexive and has the left dual

$$\|f\|_{\mathrm{dil}}^{*\mathrm{L}} = \sup_{x,y \in M}^{0} \left(|f(x)\, f(y)| - |x\, y|\right). \tag{6.3}$$

Exercise 6.3a. Show that $\|f\|_{\mathrm{dil}}^{*\mathrm{L}} \geq \|f\|_{-\,\mathrm{diam}}$. Hint: Use Exercise 6.2.

Exercise 6.3b. Find a counterexample to the reverse inequality $\|f\|_{\mathrm{dil}}^{*\mathrm{L}} \leq \|f\|_{-\,\mathrm{diam}}$.

Note that a function f with $\|d\|_{\mathrm{dil}}^{*\mathrm{L}}$ bounded is single-valued. Moreover let T be the one point metric space. Observe that

$$\mathrm{diam}\, M = \|M \to T\|_{\mathrm{dil}}. \tag{6.4}$$

Further set $\mathrm{d}_{\mathrm{dil}}(\mathcal{M}, \mathcal{N}) := d_{\|\,.\,\|_{\mathrm{dil}}}(\mathcal{M}, \mathcal{N}) = \inf\{\, \|f\|_{\mathrm{dil}} \mid f\colon \mathcal{M} \to \mathcal{N} \,\}$.

We recall the well-known notion of Gromov-Hausdorff distance [26, § 11.1.1, 14, § 3A] employing the following shorthand notations for any $A \subseteq M$

$$A^{r)} := \{\, x \in M \mid |x\, A| < r \,\} \qquad \text{for } r > 0 \text{ and}$$
$$A^{r]} := \bigcap_{r' > r} A^{r')} \qquad \text{for } r \geq 0.$$

A subset $X \subseteq M$ is said to be ***l*-dense in** M if $X^{l]} = M$. Let $A, B \subset M$ be subsets of a metric space $\mathcal{M}$. The **Hausdorff distance** between A and B is given by

$$\mathrm{d}_{\mathrm{H}}(A, B) := \inf\{\, r \in [0, \infty] \mid A \subseteq B^{r]} \text{ and } B \subseteq A^{r]} \,\}.$$

Let M, N be metric spaces. Their **Gromov-Hausdorff distance** is

$$\mathrm{d}_{\mathrm{GH}}(\mathcal{M}, \mathcal{N}) := \inf\{\, \mathrm{d}_{\mathrm{H}}(f_! M, g_! N) \mid \mathcal{M} \xrightarrow{f} \mathcal{L} \xleftarrow{g} \mathcal{N} \,\}$$

where $\mathcal{L}$ ranges over all metric spaces and f, g are metric embeddings. Recall that a function is **Cauchy continuous** if it preserves Cauchy sequences.

Theorem 6.4. *The identity map on* $\mathrm{sk}_0(\underline{\mathrm{MET}}_{\mathrm{cpt}}, \|\,.\,\|_{\mathrm{dil}})$ *with the Gromov-Hausdorff metric* d_{GH} *on the domain and the distance* $\mathrm{d}_{\mathrm{dil}}^{+}$ *on the codomain is 2-Lipschitz with Cauchy continuous inverse.*

The next lemma is a quantitative version of [9], cf. also [2][3]

Lemma 6.5. *Let $\mathcal{M}, \mathcal{M}'$ be compact metric spaces. For all L, l with $l > L \geq 0$ there is a $\delta > 0$ such that for every $h\colon \mathcal{M}' \to \mathcal{M}'$ with $\|h\|_{\mathrm{dil}} < \delta$ and $\mathrm{d}_{\mathrm{dil}}(\mathcal{M}', \mathcal{M}) \leq L$ we have that*

(i) *$h_!(M')$ is l-dense, and*

(ii) *$\|h\|_{\mathrm{dil}}^{*\mathrm{L}} \leq 4l + C\delta$ where $C = C(l-L, \mathcal{M})$.*

Proof. Recall that an **l-packing of $\mathcal{M}$** is a collection $P := (p_1, \dots, p_n) \subset M$ such that $|p\,q| > l$ for $(p,q) \in \operatorname{Conf} P := \{\,(p,q) \in X^{\times 2} \mid p \neq q\,\}$. Let $\#_l^{\mathrm{pack}}(\mathcal{M})$ be the metric l-packing number of $\mathcal{M}$, i.e.

$$\#_l^{\mathrm{pack}}(\mathcal{M}) := \sup{}^0\{\, n \mid \exists l\text{-packing } (p_1, \dots, p_n) \text{ of } \mathcal{M}\,\}.$$

Since $\mathcal{M}$ is compact, $N := \#_l^{\mathrm{pack}}(\mathcal{M})$ is finite. A collection $P := (p_1, \dots, p_N)$ is a **maximal l-packing of $\mathcal{M}$**. Further define the **total distances** for a finite $P \subseteq M$ and of $\mathcal{M}$ itself by

$$\|P\|^{\mathrm{tot}} := \sum_{(p,q)\in \operatorname{Conf} P} |p\,q|, \quad \text{and}$$

$$\|\mathcal{M}\|_l^{\mathrm{tot}} := \sup\{\, \|P\|^{\mathrm{tot}} \mid P \text{ is an } l\text{-packing of } \mathcal{M}\,\}.$$

Let $h\colon \mathcal{M}' \to \mathcal{M}'$ be a map between compact metric spaces. The integer $\#_l^{\mathrm{pack}}(\mathcal{M}')$ is continuous from the left and monotonically decreasing in l. Thus we can find some small $\varepsilon > 0$ such that for all $\delta \in (0, \varepsilon]$, we have $\#_{l-\delta}^{\mathrm{pack}}(\mathcal{M}') = \#_l^{\mathrm{pack}}(\mathcal{M}')$. For any l-packing $P = (p_1, \dots, p_N)$ with $N := \#_l^{\mathrm{pack}}(\mathcal{M}')$ and h with $\|h\|_{\mathrm{dil}} < \delta \leq \varepsilon$ the collection $h(P)$ is an $(l-\delta)$-packing. Since $\#_{l-\delta}^{\mathrm{pack}}(\mathcal{M}') = \#_l^{\mathrm{pack}}(\mathcal{M}') = \#_{l-\varepsilon}^{\mathrm{pack}}(\mathcal{M}')$ this implies that $h(P)$ is even a maximal $(l-\delta)$-packing. Hence $h(M)$ is l-dense; actually even $(h(P))^{l)} = M$. Thus claim (i) holds.

For claim (ii), i.e. $\|h\|_{\mathrm{dil}}^{*\mathrm{L}} \leq 4l + C\delta$, assume further that $\mathrm{d}_{\mathrm{dil}}(\mathcal{M}', \mathcal{M}) \leq L$. Observe that it is possible to find an l-packing P in h^*M' such that

$$\|P\|^{\mathrm{tot}} > \|\mathcal{M}'\|_l^{\mathrm{tot}} - \big(\#_{l-L}^{\mathrm{pack}}(\mathcal{M})\big)^2\delta.$$

We still assume $\delta \leq \varepsilon$. Further observe that $\|h(P)\|^{\mathrm{tot}} \leq \|\mathcal{M}'\|_{l-\delta}^{\mathrm{tot}} = \|\mathcal{M}'\|_l^{\mathrm{tot}} < \|P\|^{\mathrm{tot}} + \big(\#_l^{\mathrm{pack}}(\mathcal{M}')\big)^2\delta \leq \|P\|^{\mathrm{tot}} + \big(\#_{l-L}^{\mathrm{pack}}(\mathcal{M})\big)^2\delta$ and hence,

$$\sum_{(p,q)\in \operatorname{Conf} h(P)} |p\,q| \leq \sum_{(p,q)\in \operatorname{Conf} P} |p\,q| + \big(\#_{l-L}^{\mathrm{pack}}(\mathcal{M})\big)^2\delta.$$

From $|h(p)\,h(q)| \geq |p\,q| - \delta$ and a summand-wise comparison we get that for all $p, q \in P$ (using the notation $\tilde{p} = h(p), \tilde{q} = h(q)$),

$$\begin{aligned}|\tilde{p}\,\tilde{q}| &\leq |p\,q| + \big(\#_l^{\mathrm{pack}}(\mathcal{M}') - 1\big)^2\delta + \big(\#_{l-L}^{\mathrm{pack}}(\mathcal{M})\big)^2\delta \\ &\leq |p\,q| + \big(\#_{l-L}^{\mathrm{pack}}(\mathcal{M}) - 1\big)^2\delta + \big(\#_{l-L}^{\mathrm{pack}}(\mathcal{M})\big)^2\delta\end{aligned}$$

[3]Note that the result of the latter is implied by the former using the closure (6.1a).

$$\leq |p\,q| + C'\delta \tag{6.5}$$

where the parameter C' depends upon $\mathcal{M}$ and $l - L$.

To conclude the argument for claim (ii), let $x, y \in M'$. Set $\tilde{x} = h(x)$ and $\tilde{y} = h(y)$. We derive an estimate for $\tilde{p} = h(p)$ and $\tilde{q} = h(p)$ with $p, q \in P$ such that $|\tilde{x}\,\tilde{p}|, |\tilde{y}\,\tilde{q}| < l$. Observe that $|x\,p|, |y\,q| \leq l + \delta$, so we have

$$|p\,q| \leq |p\,x| + |x\,y| + |y\,q| \leq |x\,y| + 2(l + \delta).$$

Now we can apply (6.5)

$$|x\,y| \geq |p\,q| - 2(l + \delta) \geq |\tilde{p}\,\tilde{q}| - C'\delta - 2(l + \delta).$$

Finally, we obtain by setting $C := C' + 2$

$$|\tilde{x}\,\tilde{y}| \leq |\tilde{x}\,\tilde{p}| + |\tilde{p}\,\tilde{q}| + |\tilde{q}\,\tilde{y}| \leq l + \big(|x\,y| + C'\delta + 2(l + \delta)\big) + l = 4l + C\delta. \qquad \square$$

This theorem in particular implies that in the category $\boxed{M}(\underline{\mathrm{MET}}_{\mathrm{cpt}}, \|\,.\,\|_{\mathrm{dil}})$ every endomorphism is an isomorphism. Such categories are called **EI-categories** and have been studied for several decades [5].

Exercise 6.6. Find a counterexample showing that claim (i) in Lemma 6.5 does not hold when $\|\,.\,\|_{\mathrm{dil}}$ is replaced by $\|\,.\,\|_{-\,\mathrm{diam}}$.

Proof of Theorem 6.4. Set $\mathrm{sk}_0 := \mathrm{sk}_0(\underline{\mathrm{MET}}_{\mathrm{cpt}}, \|\,.\,\|_{\mathrm{dil}})$. First, we prove the 2-Lipschitz property of $\mathrm{id}\colon (\mathrm{sk}_0, \mathrm{d}_{\mathrm{GH}}) \to (\mathrm{sk}_0, \mathrm{d}^+_{\mathrm{dil}})$. Set $l := \mathrm{d}_{\mathrm{GH}}(\mathcal{M}, \mathcal{N})$. For every $\varepsilon > 0$ we have embeddings $\mathcal{M} \xrightarrow{f} \mathcal{L} \xleftarrow{g} \mathcal{N}$ such that $f_! M \subset (g_! N)^{l+\varepsilon]}$ and $g_! N \subset (f_! M)^{l+\varepsilon]}$. Set $h[x] := \mathrm{B}[x, l + \varepsilon] \cap N$, where the ball and the intersection are in $\mathcal{L}$. Observe that $\|h\|_{\mathrm{dil}} = \sup_{x,y} |x\,y| - |f(x)\,f(y)| \geq \sup_{x,y} |x\,f(x)| + |y\,f(y)| \geq 2(l + \varepsilon)$. Since $\varepsilon > 0$ can be chosen arbitrarily small, $\mathrm{d}_{\mathrm{dil}}(\mathcal{M}, \mathcal{N}) \leq 2l$. The analog argument with $\mathcal{M}$ and $\mathcal{N}$ interchanged gives $\mathrm{d}_{\mathrm{dil}}(\mathcal{M}, \mathcal{N}) \leq 2l$. This implies $\mathrm{d}^+_{\mathrm{dil}}(\mathcal{M}, \mathcal{N}) \leq 2l$.

To show that $\mathrm{id}\colon (\mathrm{sk}_0, \mathrm{d}^+_{\mathrm{dil}}) \to (\mathrm{sk}_0, \mathrm{d}_{\mathrm{GH}})$ is Cauchy continuous it suffices to show that for any Cauchy sequence $\mathcal{M}_n$ with respect to $\mathrm{d}^+_{\mathrm{dil}}$ the following holds: for all $N \in \mathbb{N}, L > 0$ we have that if $\forall n > N\colon \mathrm{d}^+_{\mathrm{dil}}(\mathcal{M}_N, \mathcal{M}_n) < L/2$, then $\exists M \geq N\colon \forall n, m \geq M\colon \mathrm{d}_{\mathrm{GH}}(\mathcal{M}_n, \mathcal{M}_m) \leq 5L$.

Take such N and $L > 0$ so that for all $n > N$ we have $\mathrm{d}^+_{\mathrm{dil}}(\mathcal{M}_N, \mathcal{M}_n) < L/2$. We know that $\mathrm{d}_{\mathrm{dil}}(\mathcal{M}_n, \mathcal{M}_N) < L$ for all $n \geq N$. Let $C = C(\mathcal{M}_N, L)$ be the parameter from Lemma 6.5. Choose $M \geq N$ so large that for all $n, m > M$

there are maps $\mathcal{M}_n \xrightarrow{f_{nm}} \mathcal{M}_m \xrightarrow{g_{nm}} \mathcal{M}_m$ such that $\|f_{nm}\,;g_{nm}\|_{\mathrm{dil}} < {}^{L}\!/C \wedge L$ and $\|f_{nm}\|_{\mathrm{dil}}, \|g_{nm}\|_{\mathrm{dil}} < L$. Set $h_{nm} := f_{nm}\,; g_{nm}$. Hence by Lemma 6.5 for sufficiently large n we have that $h_{nm!}(M_n)$ is L-dense in $\mathcal{M}_n$ and $\|h_{nm}\|^{*\mathrm{L}}_{\mathrm{dil}} \leq 5L$. Thus $g_{nm!}(M_m)$ is L-dense in $\mathcal{M}_n$. Therefore $f_{nm!}(M_n)$ must be $2L$-dense in $\mathcal{M}_m$.

On $M_n \sqcup M_m$ consider the symmetric function determined by the assignment

$$d_{nm}(x,y) := \begin{cases} |x\,y|_n & \text{if } x,y \in M_n \\ |x\,y|_m & \text{if } x,y \in M_m \\ 3L + \inf\limits_{x'\in M_n} |x\,h_{nm}(x')| + |f_{nm}(x')\,y|_n & \text{if } x \in M_n, y \in M_m. \end{cases} \tag{6.6}$$

Obviously, d_{nm} distinguishes points. The triangle inequality is left to Exercise 6.7. Within $(M_n \sqcup M_m, d_{nm})$ we have $M_m^{5L]} \supseteq M_n$ since $f_{nm}(M_n)$ is $2L$-dense in $\mathcal{M}_m$. By the same fact, $M_n^{5L]} \supseteq ((f_{nm})_! M)^{2L]} \supseteq M_m$. Hence $\mathrm{d_{GH}}(\mathcal{M}_n, \mathcal{M}_m) \leq 5L$. $\square$

Exercise 6.7. Check the triangle inequality for the symmetric function (6.6).

The fact that the Gromov-Hausdorff space is complete [26, § 11.1.1] implies:

Corollary 6.8. *The space* $(\mathrm{sk}_0(\underline{\mathrm{MET}}_{\mathrm{cpt}}, \|\,.\,\|_{\mathrm{dil}}), \mathrm{d}^{+}_{\mathrm{dil}})$ *is a complete metric space.*

Exercise 6.9. Show that the category $(\underline{\mathrm{MET}}_{\mathrm{cpt}}, \|\,.\,\|_{\mathrm{dil}})$ is normed. Hint: use Lemma 6.5 for (N3) and (6.2c) for (N4).

References

[1] P. Bubenik, V. de Silva, and J. Scott. *Interleaving and Gromov-Hausdorff distance.* April 26, 2018. arXiv: 1707.06288 [math.CT].

[2] O. Bucicovschi and D. Meyer. *Expansive maps are isometries.* July 8, 2015. arXiv: 1507.05044 [math.MG].

[3] A. Carboni et al. "On localization and stabilization for factorization systems". In: *Appl. Categ. Structures* 5 (1 March 1997), pp. 1–58.

[4] J. Conway. *A Course in Functional Analysis.* Grad. Texts in Math. Springer New York, 1994.

[5] T. tom Dieck. *Transformation groups.* Vol. 8. De Gruyter Stud. Math. de Gruyter, 1987.

[6] S. Eilenberg. "Sur les transformations continues d'espaces métriques compacts". In: *Fund. Math.* 1.22 (1934), pp. 292–296.

[7] M. H. Escardó. *The Cantor-Schröder-Bernstein for homotopy types, or ∞-groupoids, in Agda.* 2020. URL: https://www.cs.bham.ac.uk/~mhe/agda-new/CantorSchroederBernstein.html.

[8] M. H. Escardó. *The Cantor-Schröder-Bernstein Theorem for ∞-groupoids.* Version 2. August 26, 2020. arXiv: 2002.07079 [math.AG].

[9] H. Freudenthal and W. Hurewicz. "Dehnungen, Verkürzungen, Isometrien". German. In: *Fund. Math.* 26.1 (1936), pp. 120–122.

[10] H. Freytes. "The Cantor-Bernstein-Schröder theorem via universal algebra". In: *Algebra universalis* 80.2 (2019), pp. 1–28.

[11] P. Ghez, R. Lima, and J. E. Roberts. "W^*-categories." In: *Pacific J. Math.* 120.1 (1985), pp. 79–109.

[12] W. T. Gowers. "A solution to the Schroeder-Bernstein problem for Banach spaces". In: *Bull. Lond. Math. Soc.* 28.3 (1996), pp. 297–304.

[13] M. Grandis. "Categories, norms and weights". In: *J. Homotopy Relat. Struct.* 2.2 (2007), pp. 171–186.

[14] M. Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces.* Ed. by J. Lafontaine and P. Pansu. Progr. Math. Birkhäuser, 1999.

[15] W. Hurewicz and H. Wallman. *Dimension Theory.* Princeton Mathematical series. Princeton University Press, 1941.

[16] P. T. Johnstone. *Sketches of an Elephant. A Topos Theory Compendium.* 2 vols. Oxford University Press, 2002.

[17] R. Kadison and J. Ringrose. *Fundamentals of the Theory of Operator Algebras.* Vol. 2: *Advanced Theory.* Grad. Stud. Math. 16. American Mathematical Society, 1997.

[18] W. Kubiś. *Categories with norms.* May 29, 2017. arXiv: 1705.10189 [math.CT].

[19] D. Laackman. "The Cantor-Schroeder-Bernstein Property in Categories". In: 2010 Summer VIGRE Program for Undergraduates. The University of Chicago. 2010. URL: https://math.uchicago.edu/~may/VIGRE/VIGRE2010/REUPapers/Laackman.pdf.

[20] F. W. Lawvere. "Metric spaces, generalized logic, and closed categories". In: *Rend. Semin. Mat. Fis. Milano* 43 (1973), pp. 135–166.

[21] C. Löh. *Geometric group theory. An Introduction.* Universitext. Springer, 2017.

[22] H. Lord. "Monotone-light factorizations: A brief history". In: *Mathematik-Arbeitspapiere* (1997). URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.123.6329&rep=rep1&type=pdf.

[23] S. MacLane and I. Moerdijk. *Sheaves in Geometry and Logic. A First Introduction to Topos Theory.* Universitext. Springer, 1994.

[24] A. Neeman. "Metrics on triangulated categories". In: *J. Pure Appl. Algebra* 224 (4 2020).

[25] A. R. Pears. *Dimension theory of general spaces.* 1st ed. Cambridge University Press, 1975.

[26] P. Petersen. *Riemannian Geometry.* 3rd ed. Grad. Texts in Math. 171. Springer, 2016.

[27] L. N. Scoccola. "Locally Persistent Categories And Metric Properties Of Interleaving Distances". PhD thesis. The University of Western Ontario, July 15, 2020. URL: https://ir.lib.uwo.ca/etd/7119.

[28] S. Srivastava. *A Course on Borel Sets.* Grad. Texts in Math. 180. Springer, 1998.

[29] C. Villani. *Optimal Transport. Old and New.* Grundlehren Math. Wiss. 338. Springer, 2008.

[30] G. T. Whyburn. "Non-alternating transformations". In: *Amer. J. Math.* 56.1/4 (1934), pp. 294–302.

[31] G. T. Whyburn. "Open and closed mappings". In: *Duke Math. J.* 17.1 (1950), pp. 69–74.

Missouri University of Science and Technology, USA
insall@mst.edu

Ben-Gurion University of the Negev, Israel
luckhard@post.bgu.ac.il

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

A FRACTAL TRIANGLE ARISING IN THE AIMD DYNAMICS

KRZYSZTOF LEŚNIAK, NINA SNIGIREVA & FILIP STROBIN

ABSTRACT. We discuss a necessary and sufficient condition for the contractivity of an IFS consisting of stochastic matrices. This is based on the recent criterion due to R. Kantrowitz and M.M. Neumann (2014) concerning the ergodicity of Markov matrices via metric fixed point theory. The IFSs under consideration appear in the study of the dynamics of the AIMD (additive increase multiplicative decrease) algorithm for resource allocation. We consider a particular example of an AIMD IFS due to M. Corless, C. King, R. Shorten and F. Wirth (2016). We also investigate IFSs consisting of eventually contractive stochastic matrices (a condition equivalent to their ergodicity).

KEYWORDS: IFS, AIMD, column-stochastic matrix, 1-norm, contraction

MSC2020: 28A80, 68W40

Received 29 April 2021; revised 28 June 2021; accepted 19 July 2021

1. INTRODUCTION

Recent investigations of iterated function systems (IFS) concentrate on non-contractive systems, e.g., [4, 5, 8]. In certain cases one can apply techniques from the theory of contractive IFSs to non-contractive systems. This is the case for IFSs which arise in the study of the AIMD (additive increase multiplicative decrease) algorithm for network resources allocation, cf. [7]. In this paper, we will show that, although the IFSs of interest are not globally contractive, they are contractive with respect to the ℓ^1-norm on all c-hyperplanes, H_c, that are parallel to the hyperplane which contains a unit simplex (of probability vectors). Hence, the standard theory with respect to attractors and invariant measures applies to such IFSs. Moreover, we will show that instead of AIMD matrices, one can take a certain class of very general stochastic matrices to form an IFS. We prove this

 https://doi.org/10.2478/9788366675360-013.

by strengthening the recent result of Kantrowitz and Neumann[10]. Furthermore, we will discuss a special IFS considered in [7]. In particular, we will show that its triangle-like attractor is of the Lebesgue measure zero on every hyperplane. In the last section, we will consider examples of IFSs whose matrices are only eventually contractive on H_c (equivalently, ergodic matrices).

2. Notation and definitions

In this paper the underlying space is the Euclidean space of dimension d denoted by $\mathbb{R}^d$. The null vector is written as θ and $u = (u_i)_{i=1}^d$, $v = (v_i)_{i=1}^d$, denote (column) vectors in $\mathbb{R}^d$. We will use the standard notion for the orthonormal vectors, namely $e_j, j = 1, \ldots, d$. Given $c \in \mathbb{R}$, the *c-hyperplane* is

$$H_c = \{u \in \mathbb{R}^d : \sum_{i=1}^{d} u_i = c\}. \tag{2.1}$$

The simplex of probability vectors is

$$\Delta = H_1 \cap \{u \in \mathbb{R}^d : u_i \geq 0 \text{ for } i = 1, \ldots, d\}.$$

We equip $\mathbb{R}^d$ with the 1-norm, $\|u\|_1 = \sum_{i=1}^d |u_i|$, instead of the standard norm. For a closed subset $X \subseteq \mathbb{R}^d$ with the topology induced from $\mathbb{R}^d$ and its subset $S \subseteq X$, we denote by $\mathrm{cl}_X\, S$ the topological closure of S relative to X, and by $\mathrm{int}_X\, S$ the topological interior of S relative to X. The convex hull of S is written as $\mathrm{conv}\, S$.

A real matrix $A = (a_{ij})_{i,j \in \{1,\ldots,d\}}$ is a *column-stochastic matrix* provided $a_{ij} \in [0,1]$ and $\sum_{i=1}^d a_{ij} = 1$ for $i, j \in \{1, \ldots, d\}$. We note that a stochastic matrix is often called a Markov matrix. However, some authors impose additional conditions on a stochastic matrix so that it can be called a Markov matrix, cf. [12, chap. 4.4, definition 4.7].

Observe that $A(H_c) \subseteq H_c$ when A is column-stochastic. An *AIMD matrix* is a $d \times d$-matrix A of the form

$$A = (\beta_i \cdot \delta_{ij} + \gamma_i \cdot (1 - \beta_j))_{i,j \in \{1,\ldots,d\}}, \tag{2.2}$$

where $\delta_{i,j} = 1$ if $i = j$, and 0 otherwise (Kronecker's delta), and parameters form two vectors: $(\beta_i)_{i=1}^d \in [0,1]^d$ and $(\gamma_i)_{i=1}^d \in (0,1)^d$, $\sum_{i=1}^d \gamma_i = 1$ (positive probability vector), cf. [7] eq. (1.12) p.5, eq. (3.15) p.33. One should note that an AIMD matrix is always column-stochastic. The matrix transpose of A will be denoted by A^T.

Let $X \subseteq \mathbb{R}^d$ be a nonempty closed set. The *iterated function system (IFS)* acting on X is a finite collection $\mathcal{F} = (X; f_i : i = 1, \ldots, N)$, $N \geq 1$, of continuous maps $f_i : X \to X$. The *Hutchinson operator* $F : \mathcal{H} \to \mathcal{H}$ acts on the hyperspace

$\mathcal{H}$ of nonempty closed subsets S of X according to the formula

$$F(S) = \mathrm{cl}_X \bigcup_{i=1}^{N} f_i(S).$$

If all maps f_i are Banach contractions, then F can be restricted to the hyperspace $\mathcal{H}_b$ of nonempty closed and bounded subsets of X, and $F : \mathcal{H}_b \to \mathcal{H}_b$ is Banach contractive with respect to the Hausdorff metric. Hence there exists the Hutchinson *attractor* of $\mathcal{F}$, that is a set $\mathbb{A} \in \mathcal{H}_b$ such that for all $S \in \mathcal{H}_b$ the iterates converge $F^n(S) \to \mathbb{A}$ with respect to the Hausdorff metric. One should note that $\mathbb{A}$ is compact then, and invariant in the sense that $F(\mathbb{A}) = \mathbb{A}$.

Given a positive probability vector $(p_i)_{i=1}^N$, $p_i > 0$, $\sum_{i=1}^N p_i = 1$, and an IFS $\mathcal{F}$ on closed $X \subseteq \mathbb{R}^d$ we define the *Markov operator* $M : \mathcal{M} \to \mathcal{M}$ acting on the space $\mathcal{M}$ of Borel probability measures μ on X according to the formula

$$M(\mu)\,(B) = \sum_{i=1}^{N} p_i \cdot \mu(f_i^{-1}(B))$$

for every Borel set $B \subseteq X$. (In other words: $M\mu = \sum_{i=1}^N p_i \cdot \mu \circ f_i^{-1}$ by slight abuse of notation.) If $\mathcal{F}$ consists of Banach contractions, then there exists an *attractive invariant probability measure* for the probabilistic IFS $(\mathcal{F}, (p_i)_{i=1}^N)$, that is a measure $\mu_* \in \mathcal{M}$ such that

$$\text{(invariance)} \qquad \mu_* = M(\mu_*),$$
$$\text{(attractivity)} \qquad M^n(\mu) \underset{n\to\infty}{\rightharpoonup} \mu_*;$$

$\rightharpoonup$ stands for the weak convergence of measures. One should note that the support of μ_* is precisely the attractor $\mathbb{A}$.

Contractive IFSs have good convergence properties not only under iteration of sets via Hutchinson operator and under iteration of measures via Markov operator, but also under iteration of orbits. For $\mathcal{F}$ comprising Banach contractions on X there holds

$$\text{(chaos game representation)} \qquad \mathbb{A} = \bigcap_{n=1}^{\infty} \mathrm{cl}_X \{u^{(m)} : m \geq n\},$$
$$\text{where} \qquad u^{(n)} = f_{i_n}(u^{(n-1)}), n \geq 1,$$

and $(i_n)_{n=1}^\infty \in \{1, \ldots, N\}^\infty$ is a disjunctive sequence, cf. [15]. In particular, if $(i_n)_{n=1}^\infty$ is generated by means of a Bernoulli process, then the equality in the chaos game representation of the attractor holds with probability 1 (classic chaos game algorithm).

For more information on contractive IFSs we refer to [1] and [11].

The considerations in Section 5, the last one, involve non-contractive IFSs. For these kind of systems local attractors, rather than global Hutchinson attractors, are often more appropriate. Following [3, 8] we say that a nonempty compact set $\mathbb{A}_\sharp \subseteq X \subseteq \mathbb{R}^d$ is a *strict attractor* (respectively, *pointwise attractor*) of an IFS $\mathcal{F} = (X; f_i : i = 1, \ldots, N)$, $N \geq 1$, provided there exists an open neighbourhood $U \supseteq \mathbb{A}_\sharp$ such that $F^n(S) \to \mathbb{A}_\sharp$ (the convergence taking place in the Hausdorff metric) for all nonempty compact $S \subseteq U$ (respectively, for all singletons $S = \{x\}$, $x \in U$). The maximal neighbourhood U is called the *basin* of attraction of $\mathbb{A}_\sharp$. Evidently, any strict attractor is pointwise, and the Hutchinson attractor is a strict attractor with full basin $U = X$, but not the other way round.

3. Kantrowitz–Neumann criterion and IFSs of stochastic matrices

The authors in [7] show that the AIMD matrix (2.2) with parameters $(\beta_i)_{i=1}^d \in [0,1]^d$, $(\gamma_i)_{i=1}^d \in (0,1)^d$, is Banach contractive on the hyperplane H_0 provided $\max_{i=1,\ldots,d} \beta_i < 1$, cf. [7, Lemma 3.5 p.33]. We will show that the case when $\beta_{i_*} = 1$ for some $i_* = 1, \ldots, d$ and $\beta_i < 1$ for $i \neq i_*$, important for the analyses in [7], can be treated similarly. In fact we will show that IFSs comprising of very general stochastic matrices, not only AIMD matrices, fall within the farmework of contractive IFSs. We start by recalling the following result from [10] which we will extend and use throughout the paper.

Proposition 3.1 (Kantrowitz–Neumann criterion, [10], Proposition 3.2)**.** For each column-stochastic $d \times d$ matrix $A = (a_{ij})$, the following assertions are equivalent:

(a) $\|Au - Av\|_1 < \|u - v\|_1$ for all $u, v \in \Delta$ with $u \neq v$;
(b) $\|Ae_j - Ae_k\|_1 < \|e_j - e_k\|_1$ for all distinct $j, k \in \{1, \ldots, d\}$ (where e_j is the jth orthonormal vector);
(c) for all $j, k \in \{1, \ldots, d\}$, there exists some $l \in \{1, \ldots, d\}$ for which $a_{lj}, a_{lk} > 0$;
(d) all entries of $A^T \cdot A$ are (strictly) positive.

In particular, these equivalent conditions hold when A contains a row with (strictly) positive entries.

The next theorem is an extension of the above proposition.

Theorem 3.2 (Extended Kantrowitz–Neumann criterion)**.** *Let A be a column-stochastic matrix. Then the following are equivalent:*

(i) *A is a Banach contraction with respect to $\|.\|_1$ on some c-hyperplane H_c;*
(ii) *(contractivity) A is a Banach contraction with respect to $\|.\|_1$ on every c-hyperplane H_c;*
(iii) *(positivity) $A^T \cdot A$ has positive all entries.*

Proof. (i) $\Rightarrow$ (ii). Let $0 \leq \kappa < 1$ be the Lipschitz constant of A on H_c. Fix $c' \neq c$ and $u', v' \in H_{c'}$. Let $w = (w_i)_{i=1}^d \in \mathbb{R}^d$ be such that

$$w_i = d^{-1} \text{ for all } i = 1, \ldots, d. \tag{3.1}$$

Then $u = u' + (c - c') \cdot w, v = v' + (c - c') \cdot w \in H_c$ and $u' - v' = u - v$. Therefore

$$\|Au' - Av'\|_1 = \|Au - Av\|_1 \leq \kappa \|u - v\|_1 = \|u' - v'\|_1.$$

(ii) $\Rightarrow$ (iii). Since A is contractive on H_1, it is contractive on the unit simplex $\Delta \subseteq H_1$. Thus condition (a) in Proposition 3.1 is satisfied. Hence condition (d) of that proposition is satisfied, which is precisely our condition (iii).

(iii) $\Rightarrow$ (i). Condition (d) in Proposition 3.1 is satisfied. Hence condition (a) of that proposition is satisfied. Similarly, as in part (i)$\Rightarrow$ (ii), we see that

$$\|Au' - Av'\|_1 < \|u' - v'\|_1 \text{ for } u', v' \in (\Delta - w), u' \neq v', \tag{3.2}$$

where w is given by (3.1). Since $w \in \mathrm{int}_{H_1} \Delta$, we have that

$$H_0 \cap \overline{S}(\theta, r) \subseteq \mathrm{int}_{H_0}(\Delta - w) \subseteq H_0 \tag{3.3}$$

for all sufficiently small $r > 0$, where $\overline{S}(\theta, r) = \{u \in \mathbb{R}^d : \|u\|_1 = r\}$ stands for the r-sphere at θ. Recall that H_0 is a linear subspace of $\mathbb{R}^d$ and equip H_0 with the norm $\|.\|_1$ induced from $\mathbb{R}^d$. Observe that $H_0 \cap \overline{S}(\theta, r)$ is the r-sphere at $\theta \in H_0$ for the induced norm. Now, we consider the restriction of A to H_0, denoted $A' : H_0 \to H_0$. Its operator norm in H_0 is given by

$$\|A'\| = \max\left\{ \frac{\|A'u'\|_1}{\|u'\|_1} : u' \in H_0 \cap \overline{S}(\theta, r) \right\}.$$

Using (3.2), (3.3) and observing that $H_0 \cap \overline{S}(\theta, r)$ is compact we obtain $\|A'\| < 1$. Thus A is a Banach contraction on H_0 with Lipschitz constant $\|A'\|$. This settles (i) and the whole proof. $\square$

Remark 3.3. As observed in Proposition 3.1, a simple sufficient condition for positivity in Proposition 3.1, part (d), and in Theorem 3.2, part (iii), is

(iv) (row-positivity) A has at least one row with positive entries.

Indeed, the kth row of A corresponds to the kth column of A^T. Hence, on multiplying ith row r_i of A^T by jth column c_j of A we get that a positive kth element of r_i meets a positive k-th element of c_j, which results in (i, j) element of $A^T \cdot A$ being positive.

Remark 3.4. The maps that satisfy condition (a) in Proposition 3.1 and condition (3.2) in the proof of Theorem 3.2 are known in the literature as Edelstein contractions (e.g., [11, Definition 2.2 (iii)]).

We can now single out the following strengthened version of [7, Lemma 3.5 p.33].

Theorem 3.5. *If A is an AIMD matrix (2.2) with parameters $(\beta_i)_{i=1}^d$, $(\gamma_i)_{i=1}^d$ such that it contains at most one trivial column (being unit vector), equivalently, $\beta_i = 1$ for at most one i, then at least one row of A is positive and A is a Banach contraction with respect to $\|.\|_1$ on every c-hyperplane.*

Proof. Case (a). If $\beta_i < 1$ for all i, then $\gamma_i \cdot (1 - \beta_j) > 0$ for all $i, j \in \{1, \ldots, d\}$, so A has all rows positive.

Case (b). If $\beta_{i_*} = 1$ for some i_* and $\beta_i < 1$ for $i \neq i_*$, then i_*th row of A is its only positive row.

Be it Case (a) or (b), condition (iv) from Remark 3.3 is satisfied, so the extended Kantrowitz–Neumann criterion applies to A. □

Remark 3.6. One can easily see that when, in Theorem 3.5, $\beta_i = 1$ for more than one index i, then A is not a Banach contraction on c-hyperplanes.

The following constitutes a far generalization of considerations from [7].

Theorem 3.7 (On AIMD-like IFSs)**.** *Let $\mathcal{F} = (\mathbb{R}^d, f_i : i = 1, \ldots, N)$, $N \geq 1$, be an affine IFS given by column-stochastic matrices A_i, i.e., $f_i(u) = A_i \cdot u$, $u \in \mathbb{R}^d$. Suppose that each matrix $A_i^T \cdot A_i$ has positive all entries. Then each restricted IFS $\mathcal{F}_c = (H_c, f_i : i = 1, \ldots, N)$ is contractive in the norm $\|.\|_1$. Thus for every $c \in \mathbb{R}$ there exist:*

(a) *an attractor $\mathbb{A}(c)$ of $\mathcal{F}_c$, and*
(b) *a unique attractive invariant measure $\mu_*(c)$ of $(\mathcal{F}_c, (p_i)_{i=1}^N)$ that supports $\mathbb{A}(c)$; $(p_i)_{i=1}^N$ is a positive probability vector.*

Moreover,

(c) *$\mathbb{A}(c)$ can be recovered by means of the (disjunctive and probabilistic) chaos game played on the c-hyperplane.*

In particular, all of the above holds when each matrix A_i is an AIMD matrix with at most one trivial column (unit vector).

Proof. $\mathcal{F}_c$ consists of Banach contractions thanks to the extended Kantrowitz–Neumann criterion. Now classical theory applies to Part (a) (the Hutchinson theorem on the attractor of hyperbolic IFS) and Part (b) (the Hutchinson theorem on the invariant measure of hyperbolic IFS); cf. [1, 11]. Part (c) recently has been supplied with simple proofs, e.g., [4, 11]. At the end, one calls Theorem 3.5 to finish the proof. □

It should be stressed that the sufficient condition for contractivity in Theorems 3.5 and 3.7, which demands that the matrix has "at most one trivial column", is only applicable to matrices of special form, like an AIMD matrix.

Example 3.8. Let

$$A = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

Then A is a column stochastic matrix with only one unit column. However, $A^T \cdot A = A$ contains null elements, so A is not contractive with respect to $\|.\|_1$. In fact, A is not contractive on any hyperplane H_c under any metrization of H_c, because $A \cdot u = u$ for all $u = (t, c - 2t, t)^T$, $t \in \mathbb{R}$.

We finish this section by showing that all attractors $\mathbb{A}(c)$ are geometrically similar when $c \neq 0$, and things bear analogously for invariant measures $\mu_*(c)$.

Theorem 3.9. *Let $\mathbb{A}(c)$ be an attractor of $\mathcal{F}_c$, and let $\mu_*(c)$ be a unique invariant measure of $(\mathcal{F}_c, (p_i)_{i=1}^N)$, $c \neq 0$, as defined in Theorem 3.7. Then*

(a) *$\mathbb{A}(c)$ is a c-scaled image of $\mathbb{A}(1)$, that is, $\mathbb{A}(c) = c \cdot \mathbb{A}(1)$;*
(b) *$\mu_*(c)(B) = \mu_*(1)\,(c^{-1} \cdot B)$ for all Borel $B \subseteq H_c$.*

Proof. Define $h : \mathbb{R}^d \to \mathbb{R}^d$ by $h(u) := c \cdot u$ for $u \in \mathbb{R}^d$. Let us restrict $h : H_1 \to H_c$ and $h^{-1} : H_c \to H_1$, where h^{-1} is an inverse of h. (The restrictions are well defined because $h(H_1) = H_c$.) Furthermore denote by $\widetilde{f_i} := f_i|H_c$ and by $\widehat{f_i} := f_i|H_1$ the respective restrictions of f_i to H_c and H_1.

The IFSs $\mathcal{F}_1$ and $\mathcal{F}_c$ are conjugated by h, i.e.,

$$\widetilde{f_i} = h \circ \widehat{f_i} \circ h^{-1}, \qquad 1 \leq i \leq N.$$

Therefore

$$\begin{aligned}
F_c(\,h(\mathbb{A}(1))\,) &= \bigcup_{i=1}^N \widetilde{f_i}(\,h(\mathbb{A}(1))\,) = \\
&= \bigcup_{i=1}^N h \circ \widehat{f_i} \circ h^{-1}(\,h(\mathbb{A}(1))\,) = \\
&= \bigcup_{i=1}^N h \circ \widehat{f_i}(\,\mathbb{A}(1)\,) = h\left(\bigcup_{i=1}^N \widehat{f_i}(\mathbb{A}(1))\right) = \\
&= h(\,F_1(\mathbb{A}(1))\,) = h(\mathbb{A}(1)),
\end{aligned}$$

where F_c denotes the Hutchinson operator induced by $\mathcal{F}_c$. This establishes (a).

Let us write μ_* instead of $\mu_*(1)$, to simplify notation. We are going to check that the Borel measure ν on H_c, defined by $\nu := \mu_* \circ h^{-1}$, is invariant for $(\mathcal{F}_c, (p_i)_{i=1}^N)$. (That is, the transport of μ_* via h coincides with $\mu_*(c)$.) We have

$$\begin{aligned}
M\nu &= \sum_{i=1}^N p_i \cdot \nu \circ \widetilde{f_i}^{-1} = \\
&= \sum_{i=1}^N p_i \cdot \mu_* \circ h^{-1} \circ \widetilde{f_i}^{-1} = \\
&= \sum_{i=1}^N p_i \cdot \mu_* \circ h^{-1} \circ (h \circ \widehat{f_i}^{-1} \circ h^{-1}) = \\
&= \sum_{i=1}^N p_i \cdot \mu_* \circ \widehat{f_i}^{-1} \circ h^{-1} = \mu_* \circ h^{-1} = \nu, \qquad (3.4)
\end{aligned}$$

where M stands for the Markov operator induced by $(\mathcal{F}_c, (p_i)_{i=1}^N)$,

$$\widetilde{f_i}^{-1} = (h \circ \widehat{f_i} \circ h^{-1})^{-1} = (h^{-1})^{-1} \circ \widehat{f_i}^{-1} \circ h^{-1}$$

(in the sense of preimages) and (3.4) follows from the invariance of μ_* for $(\mathcal{F}_1, (p_i)_{i=1}^N)$. This establishes (b). □

Remark 3.10. One should be aware that (a) follows from (b) in Theorem 3.9. Indeed,

$$A_*(c) = \operatorname{supp} \mu_*(c) = \operatorname{supp} \mu_*(1) \circ h^{-1} = h(\operatorname{supp} \mu_*(1)) = h(A_*(1)),$$

where supp denotes the support of a measure.

4. AIMD triangle

Investigations of AIMD dynamics in Example 6.1 in [7] lead to the following special IFS.

Let $\mathcal{F} = (\mathbb{R}^3; f_i : i = 1, 2, 3)$ be a linear IFS given by

$$f_1(x,y,z) = \begin{pmatrix} 1 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{4.1}$$

$$f_2(x,y,z) = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{4} & 1 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{4.2}$$

$$f_3(x,y,z) = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{4.3}$$

It is easy to see that $\mathcal{F}$ is not contractive. However, we know from Theorem 3.7, $\mathcal{F}$ is contractive when restricted to c-hyperplanes. (Note that such restriction is always possible: $f_i(H_c) \subseteq H_c$ for $i = 1, 2, 3$.)

Below we give a direct computation which establishes contractivity of f_3 on H_c:

$$\|f_3(u) - f_3(v)\|_1 \le (3/4) \cdot \|u - v\|_1 \text{ for } u, v \in H_c. \tag{4.4}$$

(By symmetry, f_1 and f_2 have the same Lipschitz constant $3/4$ on H_c.)

Let $u, v \in H_c$. Then $u - v = (a, b, -(a+b))^T \in H_0$, and $f_3(u - v) = (a/2 + b/4, a/4 + b/2, -3a/4 - 3b/4)^T$ for $a = u_1 - v_1$, $b = u_2 - v_2$. Therefore

$$\|f_3(u) - f_3(v)\|_1 = \left|\frac{a}{2} + \frac{b}{4}\right| + \left|\frac{a}{4} + \frac{b}{2}\right| + \frac{3}{4} \cdot |a + b| \le$$
$$\left(\frac{1}{2} + \frac{1}{4}\right) \cdot |a| + \left(\frac{1}{4} + \frac{1}{2}\right) \cdot |b| + \frac{3}{4} \cdot |a + b| = \frac{3}{4}\|u - v\|_1.$$

This yields (4.4).

We can apply to $\mathcal{F}_c = (H_c; f_i : i = 1, 2, 3)$ the standard theory of IFSs and its recent advances delineated in Section 2. The IFS $\mathcal{F}_c$ admits

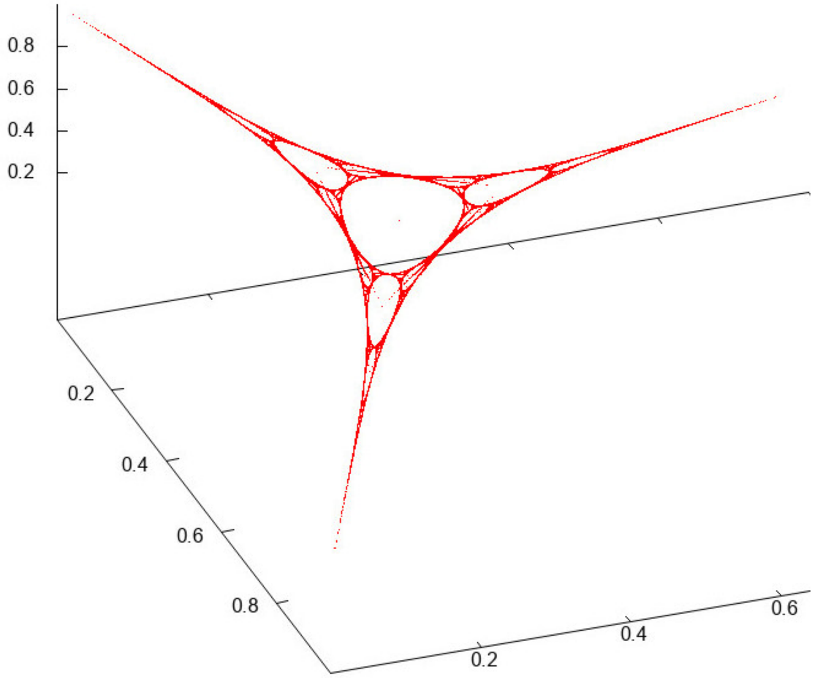


FIGURE 1. A visualization of the attractor $\mathbb{A}(1) \subseteq H_1$ of the IFS given by (4.1), (4.2), (4.3). It is generated as an orbit starting at the point $(1/3, 1/3, 1/3) \notin \mathbb{A}(1)$ (visible inside the largest hole) and driven by a de Bruijn sequence of length 59058 for words of length 10 over 3-letter alphabet. See [6, Section 4] for de Bruijn sequences and [14] for their use as drivers.

(i) the Hutchinson attractor, denoted $\mathbb{A}(c)$;
(ii) an attractive invariant measure on H_c uniquely determined by a positive probability vector on $\{1, 2, 3\}$;
(iii) chaos game representation of its attractor.

The attractor $\mathbb{A}(c)$ is depicted in Figure 1 for $c = 1$ and in Figures 6.4-6.5 in the book [7] for $c = 3$. At first glance it looks similar to the Kigami triangle, i.e., the Sierpiński gasket in harmonic coordinates, cf. [9], [2]. However, a closer inspection of it on Figure 2 reveals "line segments" sticking out of smooth parts in $\mathbb{A}(1)$. Thus, $\mathbb{A}(1)$ is not an instance of the Kigami triangle.

It should be noted that the IFS $\mathcal{F}_c$ is overlapping, the type of systems only very recently shown to be amenable for rigorous analysis in very specific cases, e.g., [13]. Let us inspect the overlapping for $\mathcal{F}_c$. Fix $c \in \mathbb{R}$. Denote $\Delta_c = \operatorname{conv}\{ce_1, ce_2, ce_3\}$. Since $f_i(ce_j) \in \Delta_c$ for $i, j \in \{1, 2, 3\}$ and f_i's are affine, we have $\mathbb{A}(c) \subseteq \Delta_c$. In particular, $\mathbb{A}(0) = \{(0, 0, 0)\}$. For $c \neq 0$ the image $F(\Delta_c)$ splits into three

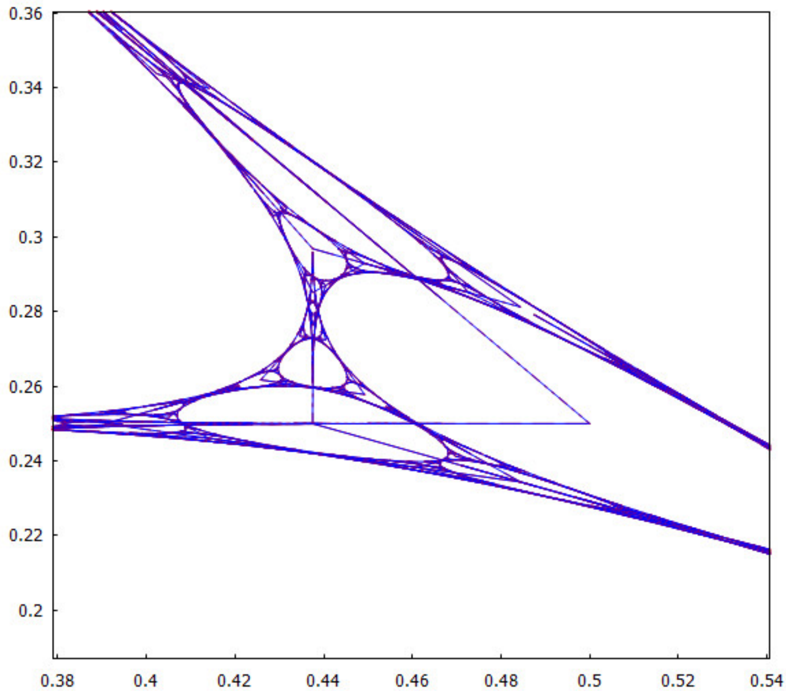


FIGURE 2. Zoomed version of the attractor from Figure 1 when projected onto the xy-plane.

overlapping triangles as shown in Figure 3. The image $F^2(\Delta_c)$ consists of several overlapping triangles positioned in such a way that there is a bounded hole in the exterior of $F^2(\Delta_c)$; see Figure 4.

There are many questions which can be asked regarding the nature of $\mathbb{A}(c)$ for $c \neq 0$. We address only the basic question concerning the invariant measure arising in the IFS $\mathcal{F}_c$ with constant probabilities.

Theorem 4.1. *The attractor of $\mathcal{F}_c$, given by* (4.1), (4.2), (4.3), *has null two dimensional Lebesgue measure $\mathcal{L}$ on the c-hyperplane. In particular, the invariant measure for $\mathcal{F}_c$ with probabilities is singular with respect to $\mathcal{L}$.*

Proof. Observe that

$$F^n(\Delta_c) = \bigcup_{(i_1,\ldots,i_n)\in\{1,2,3\}^n} f_{i_1}\circ\ldots\circ f_{i_n}(\Delta_c).$$

Roughly speaking, the n-th approximation of $\mathbb{A}(c)$ consists of 3^n collaged triangles. Each f_i transforms triangles onto triangles and shrinks the area of a triangle by factor $\frac{3}{16}$. Indeed, let $T \subseteq H_c$ be any triangle. Let $C(B) = \mathrm{conv}(B \cup \{\theta\}) \subseteq \mathbb{R}^3$ be the pyramid (cone) with base $B \subseteq H_c$, apex $\theta \in \mathbb{R}^3$ and height h from the apex

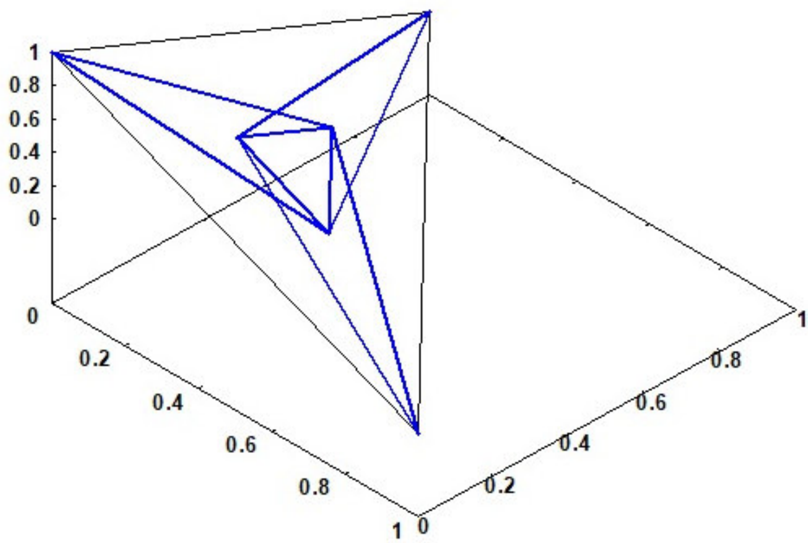


FIGURE 3. The image of the black triangle Δ_c under the action of f_i's given by (4.1), (4.2), (4.3). First iteration $F(\Delta_c)$ consists of overlapping blue triangles $f_i(\Delta_c)$. It is set $i = 1, 2, 3$, $c = 1$.

to the base (i.e., the length of $(c/3, c/3, c/3)^T$). Then $f_i(C(T)) = C(f_i(T))$ and

$$\frac{1}{3}h \cdot \mathcal{L}(f_i(T)) = \mathcal{L}^3(C(f_i(T))) =$$
$$\mathcal{L}^3(f_i(C(T))) = \frac{3}{16} \cdot \mathcal{L}^3(C(T)) = \frac{1}{3}h \cdot \frac{3}{16}\mathcal{L}(T),$$

where $\mathcal{L}^3$ denotes the three dimensional Lebesgue measure, and $3/16$ is the determinant of the matrix generating f_i. Therefore we have

$$\mathcal{L}(F^n(\Delta_c)) \leq 3^n \cdot \left(\frac{3}{16}\right)^n \mathcal{L}(\Delta_c) \to 0 \text{ as } n \to \infty.$$

□

5. Eventual contractivity

Eventual contractivity appears in the criterion for ergodicity of Markov matrices due to Kantrowitz and Neumann, cf. Theorem 3.3 in [10]. Namely, a Markov matrix is ergodic precisely when some of its powers is ℓ^1-contractive on the unit simplex of probabilities Δ_1 (equivalently, on any hyperplane H_c). Hence the question arises whether an IFS consisting of column-stochastic matrices that are eventually contractive (on hyperplane) always admits an attractor (on hyperplane). In the case of general IFSs that are eventually contractive, the answer depends

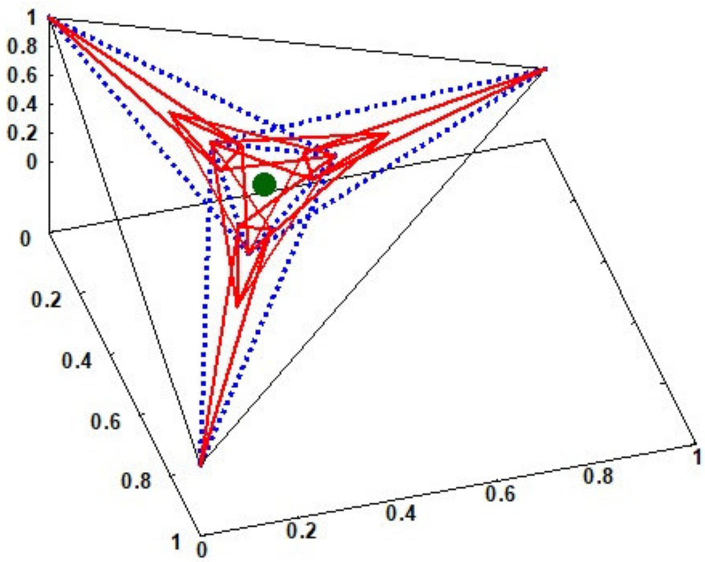


FIGURE 4. The image of the black triangle Δ_c under the action of f_i's given by (4.1), (4.2), (4.3). First iteration $F(\Delta_c)$ consists of dotted blue triangles $f_i(\Delta_c)$. Second iteration $F^2(\Delta_c)$ consists of small overlapping red triangles $f_i \circ f_j(\Delta_c)$. The green point inside one of the polygons marks a hole that is the only bounded connected component of $H_c \setminus F^2(\Delta_c)$. It is set $i, j \in \{1,2,3\}$, $c = 1$.

on compositions of distinct maps from IFS, not only upon higher iterates of individual maps defining IFS, cf. [11, Example 6.3, Remark 6.5]. We illustrate below that this is the case for column stochastic matrices as well.

Example 5.1. Let $\mathcal{F} = (\mathbb{R}^3; f_i : i = 1,2)$, where $f_i(u) = A_i \cdot u$, $u \in \mathbb{R}^3$, are linear maps given by Wieland-type matrices

$$A_i = \begin{pmatrix} 0 & \frac{i}{3} & 0 \\ 0 & 0 & 1 \\ 1 & 1-\frac{i}{3} & 0 \end{pmatrix}$$

(cf. [10, p.1249]). It turns out that although f_i's are not contractive on H_c with respect to $\|\cdot\|_1$, they are eventually contractive. In fact, the IFS $\mathcal{F}^3 = (\mathbb{R}^3; f_i \circ f_j \circ f_k : i,j,k \in \{1,2\})$ is ℓ^1-contractive on H_c since the product $A_iA_jA_k$ has third row with positive entries for all $i,j,k \in \{1,2\}$ and so we can apply Remark 3.3. Note that $\mathcal{F}^2|H_c$ is not ℓ^1-contractive. The attractor of $\mathcal{F}^3|H_c$ exists and, in consequence, it is also an attractor of $\mathcal{F}|H_c$; visualised in Figure 5.

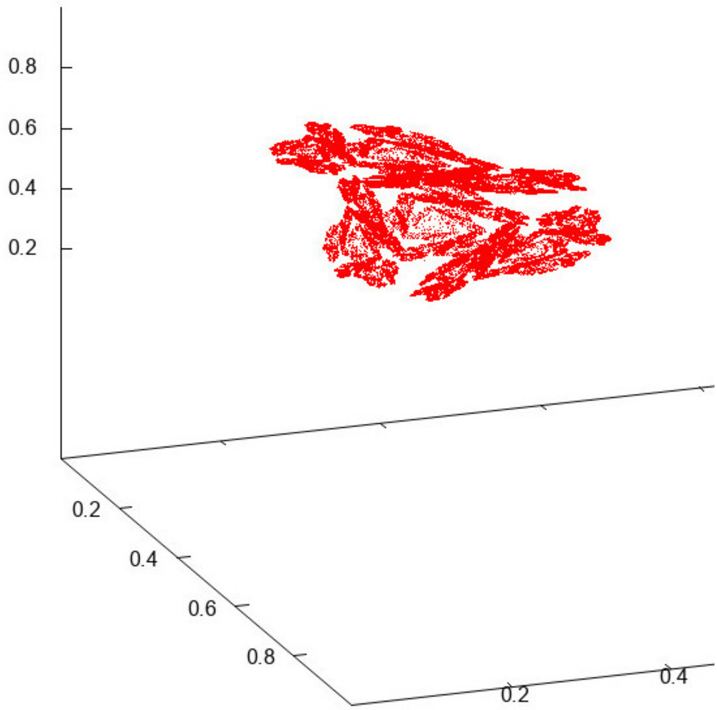


FIGURE 5. A visualization of the attractor of $\mathcal{F}$ from Example 5.1. It is generated as an orbit starting at the point $(1/3, 1/3, 1/3)$ and driven by a de Bruijn sequence of length 65551 for words of length 16 over 2-letter alphabet (cf. caption to Figure 1).

Example 5.2. Let $\mathcal{F} = (\mathbb{R}^3; f_i : i = 1, 2)$, where $f_i(u) = A_i \cdot u$, $u \in \mathbb{R}^3$, are linear maps given by (doubly stochastic) matrices

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}, \qquad A_2 = A_1^T = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

One can easily check that A_i^2 satisfy the Kantrowitz–Neumann positivity criterion, so A_i are eventually contractive on H_c. On the other hand, for instance $A_2 \cdot A_1$ is not eventually contractive (actually, it is idempotent, i.e., $(A_2A_1)^2 = A_2A_1$).

We will show that $\mathcal{F}$ does not admit a pointwise attractor on H_c. In particular $\mathcal{F}|H_c$ does not have a strict attractor. For simplicity $c := 1$. Suppose that $\mathbb{A}_\sharp$ is a pointwise attractor of $\mathcal{F}|H_1$ with basin $U \subseteq H_1$. Define

$$v_t := (1/2 - t/2, t, 1/2 - t/2)^T \in H_1, \qquad t \in \mathbb{R}.$$

Observe that $f_2 \circ f_1(v_t) = v_t$ and $v_{1/3} = f_1^2(v_{1/3})$. Since $A_\sharp$ is the pointwise attractor of the IFS $\mathcal{F}^2 = (H_1; f_i \circ f_j : i, j \in \{1, 2\})$ and f_1^2 is contractive on H_1, we have that $v_{1/3} \in \mathbb{A}_\sharp$. Hence the unbounded line $\{v_t : t \in \mathbb{R}\}$ intersects the set

$\mathbb{A}_\sharp$. Since $\mathbb{A}_\sharp$ is compact and U is open, there exists t_0 with $v_{t_0} \in U$, $v_{t_0} \notin \mathbb{A}_\sharp$. Now consider an orbit $(u^{(n)})_{n=0}^{\infty}$, where $u^{(n)} = f_{i_n}(u^{(n-1)})$, starting at $u^{(0)} := v_{t_0}$ and driven by a periodic sequence $i_n = 1$ for odd n and $i_n = 2$ for even n. Since $\mathbb{A}_\sharp$ is a pointwise attractor, $u^{(n)}$ must be attracted by $\mathbb{A}_\sharp$, that is

$$\inf_{a \in \mathbb{A}_\sharp} \|u^{(n)} - a\| \to 0 \qquad \text{as } n \to \infty, \tag{5.1}$$

where $\|\cdot\|$ is any norm (for instance ℓ^1). However, it turns out that $u^{(n)} = u^{(0)} \notin \mathbb{A}_\sharp$ for even n against (5.1). Overall, $\mathcal{F}$ does not admit a pointwise attractor on H_c.

Remark 5.3. Matrices A_1, A_2 in Example 5.2 are ergodic, while their product $A_2 \cdot A_1$ is not ergodic. This gives the solution to the first part of Exercise 4.21 on p. 148 of [12].

Acknowledgement

We would like to thank Michael Barnsley for his valuable suggestions which improved our paper. We would also like to thank the unknown referee for careful reading of our paper and pointing out to us a mistake in the initial version of Example 5.1.

References

[1] M.F. Barnsley, *Fractals Everywhere. New Edition*, Dover, Mineola, 2012.
[2] M.F. Barnsley, U. Freiberg, *Fractal transformations of harmonic functions*, Proc. SPIE 6417, Complexity and Nonlinear Dynamics, 64170C (3 January 2007); doi.org/10.1117/12.696052.
[3] M.F. Barnsley, K. Leśniak, M. Rypka, *Chaos game for IFSs on topological spaces*, J. Math. Anal. Appl. **435** (2016), no. 2, 1458-1466.
[4] M. Barnsley, A. Vince, *Developments in fractal geometry*, Bull. Math. Sci. **3**, no. 2, 299-348.
[5] P.G. Barrientos, M. Fitzsimmons, F.H. Ghane, D. Malicet, A. Sarizadeh, *Addendum and corrigendum to: "On the chaos game of iterated function systems"*, Topol. Methods Nonlinear Anal. **55** (2020), no. 2, 601-616.
[6] J. Berstel, D. Perrin, *The origins of combinatorics on words*, Eur. J. Combinatorics **28** (2007), 996-1022.
[7] M. Corless, C. King, R. Shorten, F. Wirth, *AIMD Dynamics and Distributed Resource Allocation*, SIAM, Philadelphia, 2016.
[8] M. Fitzsimmons, H. Kunze, *Small and minimal attractors of an IFS*, Commun. Nonlinear Sci. Numer. Simul. **85** (2020), Article ID 105227, 14 p.
[9] N. Kajino, *Analysis and geometry of the measurable Riemannian structure on the Sierpiński gasket*, Contemp. Math. **600** (2013), 91-133.
[10] R. Kantrowitz, M.M. Neumann, *A fixed point approach to the steady state for stochastic matrices*, Rocky Mt. J. Math. **44** (2014), no. 4, 1243-1250.
[11] K. Leśniak, N. Snigireva, F. Strobin, *Weakly contractive iterated function systems and beyond: a manual*, J. Difference Equ. Appl. **26** (2020), no. 8, 1114-1173.
[12] E. Seneta, *Non-negative Matrices and Markov Chains. Revised Printing*, Springer, New York, 2006.
[13] I. Kolossváry, K. Simon, *Triangular Gatzouras-Lalley-type planar carpets with overlaps*, Nonlinearity **32** (2019), no. 9, 3294–3341.

[14] I. McFarlane, S.G. Hoggar, *Optimal drivers for 'Random' Iteration Algorithm*, Comput. J. **37** (1994), 629-640.
[15] C.S. Calude, L. Priese, L. Staiger, *Disjunctive sequences: an overview*, CDMTCS Research Report **63** (1997).

Nicolaus Copernicus University in Toruń, Poland
much@mat.umk.pl

University College Dublin, Ireland
nina@maths.ucd.ie

Łódź University of Technology, Poland
filip.strobin@p.lodz.pl

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

DIRECTIONAL DENSITY OF POLYNOMIAL HULLS AT SINGULARITIES

ANDREAS LIND & EGMONT PORTEN

ABSTRACT. We study the thickening problem on a 2-dimensional Stein variety X with isolated irreducible singularities, i.e. the problem whether the assumption that a compact set K is contained in the interior of another compact set L implies that the same inclusion holds for their holomorphic hulls. The problem is still open except for a positive answer in the special case of quotient singularities. The main result of the present article is the partial result that the holomorphic hull L is has a directional density property at every singular point contained in the hull of K. The proof is based on removability results on pseudoconcave closed sets, which may be of some independent interest.

KEYWORDS: Holomorphic hulls, thickening property, envelopes of holomorphy

MSC2010: 32C20, 32E10, 32E20, 32D10

Received 4 May 2021; accepted 22 June 2021

1. INTRODUCTION

The present article is motivated by the following problem, to which we will refer to as the *thickening problem* for polynomial hulls:

Let X be a pure-dimensional analytic subset of the open unit ball $\mathbb{B}^n$ in $\mathbb{C}^n$ which is regular at all points except for a normal singularity at 0. *Does it hold true that, for every pair of compact sets K and L such that $0 \in \widehat{K}$ and $K \subset \operatorname{int}_X(L)$, we have $0 \in \operatorname{int}_X(\widehat{L})$?*

Here $\widehat{K}$ denotes the polynomial hull of the compact K and $\operatorname{int}_X(\cdot)$ the *relative* interior with respect to X. The problem is stated in [6], together with an indication how it is related to global analysis on Stein spaces. In what follows, we will

 https://doi.org/10.2478/9788366675360-014.

also say that L is a *thickening of* K if $K \subset \mathrm{int}_X(L)$ and refer to any affirmative (partial) answer to the problem above as a *thickening phenomenon.*

For $X = \mathbb{B}^n$ the positive answer is elementary since polynomial hulls follow rigid translations of K. In case that X is a submanifold of $\mathbb{B}^n$, this argument can be generalised by using holomorphic vector fields tangent to X, see [10] and the first step of the proof of Theorem 5.1, where the further extension to abstract Stein manifolds is achieved. Note however that this kind of reasoning does not solve the thickening problem in the general case since holomorphic vector fields vanish at isolated singularities. By now only partial knowledge is available for.the singular case. The aim of the present article is to contribute to a better understanding of the 2-dimensional case.

Let us start by reviewing the few known results, accompanied by some elementary remarks. At first it seems appropriate to place the problem into the global setting of pure-dimensional reduced normal Stein spaces with isolated singularities (with $\mathcal{O}(X)$-hulls instead of polynomial hulls). However, it is not hard to show that the thickening phenomenon is valid in the more general setting if and only if it is locally valid at every singularity of X, see [10] and also Section 4. We will see that Rossi's local maximum modulus principle even permits to reduce the problem to the case where K is contained in the intersection of X with some sphere $\partial B_\epsilon(0)$ in ambient space (after selecting some embedding).

If $\dim X = 1$, the answer is always positive, see the end of Section 3. For $\dim X \geq 2$, the problem seems to become much harder. However there is still the special but important class of quotient singularities where the thickening phenomenon is established in [10]. In this context, one should mention the dichotomy for isolated normal singularities in dimension 2 found in [3], which establishes a link between nonquotient singularities and Stein theory. This and the fact that resolving of singularities is quite explicit in dimension 2 may suggest that the problem is more accessible in dimension 2. Compare this also to the link between the thickening problem and envelopes of holomorphy [10, Thm.5.1], which is also restricted to dimension 2. Finally we mention that the normality assumption can be weakened to *irreducibility* at the isolated singularities without changing the problem in essence, see Lemma 3.1.

In the present paper, we approach the problem by showing that the hull of a thickening has a certain density property called *directional density* at singular points. This property is formulated in terms of the geometry of lifted hulls near the exceptional divisor of arbitrary resolutions, see Section 4 for the definition and our main result in Theorem 5.1. The intuitive idea is that the hull of a thickening comes close to any complex curve passing through the singularity up to arbitrarily high order, and the theorem may be viewed as a result on a weakened notion of thickening.

The article is organised as follows. In the proof of Theorem 5.1, we need some properties of pseudoconcave sets in Stein manifolds, see Section 2. A consequence of a certain independent interest is a Riemann removability theorem for holomorphic set-valued mappings. Analysing the proof of Theorem 5.1 also shows that thickening of the hull of a given compact K occurs if parts of the hull are localised with respect to the exceptional divisor of some resolution, or intuitively with respect to directions of approach to the singularity. This revisits a theme from our earlier work [10], namely thickening of hulls with analytic structure. In Section 6.1, we refine our earlier result to hulls containing a Duval-Sibony current whose support is *not* directionally dense at the singularity.

2. Pseudoconcave sets and set-valued functions

After seminal work by Hartogs [8], Oka [13] and Nishino [12], pseudoconcavity of sets in $\mathbb{C}^n$ has been studied from various points of view, corresponding to various ways of looking at pseudoconvexity properties of the complement. For our purpose, it is convenient to follow the approach from [4], which is motivated by the classical continuity principle.

Definition 2.1. A closed subset A of an n-dimensional complex manifold X is said to be *p-pseudoconcave near* $x_0 \in A$ (for a positive integer p) if there is a neighborhood U_{x_0} of x_0 such that for every domain $U \Subset U_{x_0}$ and every holomorphic mapping $F : V \to \mathbb{C}^p$ defined on $V^{\text{open}} \supset \overline{U}$ the following property holds: For any $x \in A \cap U$, the image $F(x)$ does not lie in the unbounded component of $\mathbb{C}^p \backslash F(A \cap \partial U)$. Moreover, we say that A is *p-pseudoconcave* if A is p-pseudoconcave near every point.

It is immediate that p-pseudoconcavity implies q-pseudoconcavity if $p > q$. Moreover, classical results on Stein manfolds imply that $(n-1)$-pseudoconcavity is equivalent to local Steinness of the complement. For 1-pseudoconcavity, the property we are most interested in, we also shortly write pseudoconcavity.

By definition, local p-pseudoconcavity is an *open* property. First examples of p-pseudoconcave sets are complex subvarieties of codimension p and more generally Levi flat CR submanifolds of CR dimension p (see [11] for more on the role of pseudoconcavity in CR geometry). This indicates that p-pseudoconcavity as defined above should be viewed as a condition expressing *weak* Levi curvature. Note also that p-pseudoconcavity is the same as local *Hartogs convexity* of the complement with respect to analytic sets which are fibers of a holomorphic mapping with p-dimensional target space.

Note that we have slightly modified the definition from [4] by adding the localisation embodied by U_{x_0}. In general complex manifolds, this seems reasonable in order to ensure that sets like the Levi flat $A = \{|\zeta| = 1\} \times \mathbb{P}^1 \subset \mathbb{C} \times \mathbb{P}^1$ are pseudoconcave. Because of this slight difference, it is advisable to check briefly that some essentially well-known results like Lemma 2.2 and Proposition 2.4 fit into our terminology.

We start with a global maximum principle for subsets of Stein manifolds.

Lemma 2.2. *Let A be a closed, pseudoconcave subset of a Stein manifold X. Then for every domain $U \Subset X$, the estimate*

$$\max_{A\cap\overline{U}} \varphi \leq \max_{A\cap\partial U} \varphi \tag{2.1}$$

holds for any function φ which is plurisubharmonic near $\overline{U}$.

The proof will show that it suffices to assume that the intersection of A with some open neighbourhood V of $\overline{U}$ is closed in V.

Proof. The lemma and variants of it are well-known, see for example [4]. For completeness, we provide a short argument. Let us assume that there is a function φ violating the estimate (2.1). Adding a small multiple of a global strictly plurisubharmonic function (which exists since X is Stein), we can make φ strictly plurisubharmonic. Then we can assume φ to be $\mathcal{C}^\infty$-smooth, by Richberg's theorem [15]. Note that the maximum set

$$M = \{x \in A\cap\overline{U} : \varphi(x) = \max_{A\cap\overline{U}} \varphi\}$$

is a compact subset of $A \cap U$. We will conclude by deriving that A is not pseudoconcave near any $x_0 \in M$.

Since the remaining argument is local, we can work in local coordinates $z = (z_1, \ldots, z_n)$ centered at x_0. For a sufficiently small $\epsilon > 0$, we choose first a smooth function $\chi(t) \geq 0$ which equals 1 for $t \leq \epsilon/2$ and 0 for $t \geq \epsilon$ and then $\eta > 0$ so small that $\psi = \varphi + \eta\chi(|z|)$ still is strictly plurisubharmonic on the 2ϵ-ball $B_{2\epsilon}(0)$. It follows that $K = \{\psi \leq \varphi(0)\} \cap \overline{B}_\epsilon(0)$ is holomorphically convex in $B_{2\epsilon}(0)$. In particular, there is $f \in \mathcal{O}(B_{2\epsilon}(0))$ with $f(0) = 1$ and $|f| \leq 1/2$ on K, which clearly rules out pseudoconcavity of A near x_0. □

We localise the maximum property.

Definition 2.3. A closed subset A of an arbitrary complex manifold X satisfies the *local maximumprinciple (for plurisubharmonic functions) near* $x_0 \in A$ if there is a neighborhood U_{x_0} of x_0 such that for every domain $U \Subset U_{x_0}$ the estimate

$$\max_{A\cap\overline{U}} \varphi \leq \max_{A\cap\partial U} \varphi \tag{2.2}$$

holds for any function φ which is plurisubharmonic near $\overline{U}$.

The set $A = \{|\zeta| = 1\} \times \mathbb{P}^1 \subset \mathbb{C} \times \mathbb{P}^1$ mentioned above is also an example of a pseudoconcave set, where the local maximumprinciple is valid, but not the global. In fact the two local properties introduced above are equivalent.

Proposition 2.4. *Let A be a closed subset of a complex manifold X and $x_0 \in A$. Then A is pseudoconcave near x_0 if and only if it satisfies the local maximumprinciple for plurisubharmonic functions at x_0.*

Proof. For both directions it suffices to choose $U_{x_0} \Subset X$ Stein. Then sufficiency immediately follows from Lemma 2.2.

To show necessity, we consider $f : V \to \mathbb{C}$ holomorphic on $V^{\text{open}} \supset \overline{U}$ for some $U \Subset U_{x_0}$. Assume that there is a point $x_1 \in A \cap U$ such that $w_1 = f(x_1)$ is contained in the unbounded component of the complement of the compact $K = f(A \cap \partial U)$. This means that $f(x_1)$ is not contained in $\widehat{K}$ (since $\widehat{K}$ is the union of the bounded components of $\mathbb{C}\backslash K$), allowing us to find a polynomial $g \in \mathbb{C}[w]$ which is close to 0 Hence we have found the continuous plurisubharmonic function $\varphi = |g \circ f|$ defined near $\overline{U}$ which fulfils $\varphi(x_1) > \max_{A\cap\partial U} \varphi$. The proof is complete. □

The following removability theorem is our main tool for the next sections.

Theorem 2.5. *Let X be a Stein manifold and Y be an embedded smooth complex hypersurface. If A be a closed pseudoconcave subset of $X\backslash Y$ such that $\overline{A} \cap Y$ is compact. Then $\overline{A}$ is a pseudoconcave subset of X.*

Proof. It only remains to prove pseudoconcavity near points $x_0 \in \overline{A} \cap Y$. We will choose a Stein neighbourhood $U_{x_0} \Subset X$ of x_0 and verify the maximum principle (2.1) for $U \Subset U_{x_0}$ for functions φ which are plurisubharmonic near $\overline{U}$. By upper semicontinuity, it suffices to show that

$$\varphi(x) \leq \max_{A\cap\partial U} \varphi$$

holds for all $x \in A \cap U$.

It is convenient to work in local coordinates $z_1, \ldots, z_n$ centered at x_0 such that Y is locally defined by $z_1 = 0$. After an appropriate dilation, we may take U_{x_0} as the polydisc $\{|z_1| < 1, \ldots, |z_n| < 1\}$.

For $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_n) \in A \cap U$ fixed, we consider $U_\eta = U\backslash\{|z_1| \leq \eta\}$, see Figure 1. By a compactness argument, we find for every $\epsilon > 0$ sufficiently small numbers $\delta > 0$ and $\eta \in (0, |\hat{z}_1|/2)$ such that $-\epsilon < \delta \log|\hat{z}_1| < 0$ and the plurisubharmonic function $\varphi_\delta = \varphi + \delta \log|z_1|$ satisfies

$$\max_{A\cap\partial U_\eta \cap\{|z_1|\leq 2\eta\}} \varphi_\delta < \varphi(\hat{z}) - \epsilon.$$

We deduce

$$\varphi(\hat{z}) - \epsilon < \varphi_\delta(\hat{z}) \leq \max_{A\cap\partial U_\eta} \varphi_\delta = \max_{A\cap\partial U_\eta\cap\{|z_1|\geq 2\eta\}} \varphi_\delta \leq \max_{A\cap\partial U} \varphi,$$

where the second inequality follows from Lemma 2.2 since $U_\eta \Subset X\backslash Y$, and the equality holds since φ_δ is small on $A \cap \partial U_\eta \cap \{|z_1| \geq 2\eta\}$. Finally we obtain the desired estimate after letting $\epsilon \downarrow 0$. The proof of Theorem 2.5 is complete. □

As a little digression, we point out that Theorem 2.5 implies a Riemann removability theorem for holomorphic set-valued mappings. A *set-valued mapping* between complex manifolds X_1 and X_2 is a closed set $\Gamma_F \subset X_1 \times X_2$ such that the image $F(x_1) = \Gamma_F \cap (\{x_1\} \times X_2)$ of every point $x_1 \in X_1$ is compact. Such

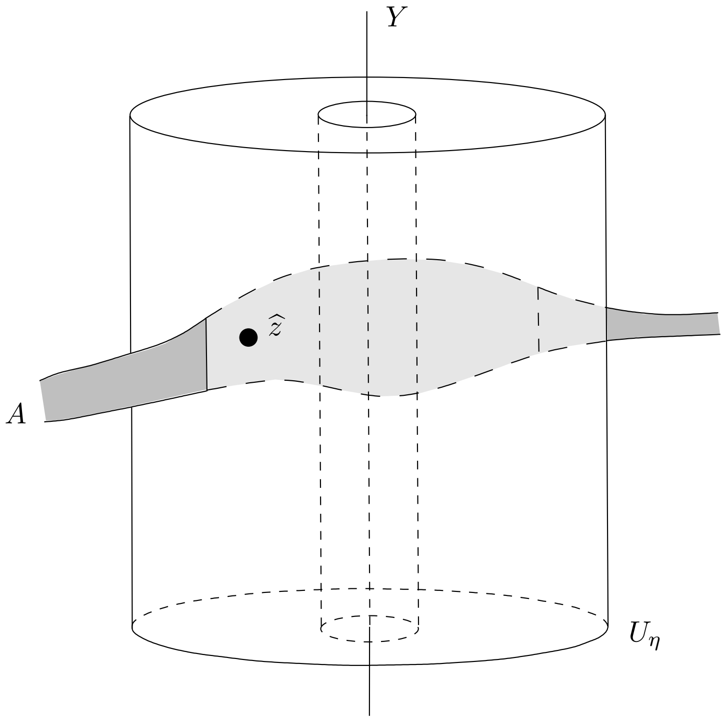


FIGURE 1. Domain U_η with tunnel drilled along Y.

F is upper semicontinuous if for every $\hat{x}_1 \in X_1$, $F(\hat{x}_1)$ equals the limit set of $F(x_1)$ for $x_1 \to \hat{x}_1$, $x_1 \neq \hat{x}_1$. The mapping F is called *holomorphic* if it is upper semicontinuous and the image is n_1-pseudoconcave with $n_1 = \dim X_1$. Holomorphic set-valued mappings are a.o. useful for dealing with meromorphic mappings, correspondences, holomorphic hulls and the spectrum of operators depending on a parameter. For more information on holomorphic set-valued mappings, we refer to [12], [19].

For set-valued mappings with 1-dimensional source manifold we obtain the following generalisation of the classical Riemann removability theorem.

Corollary 2.6. *Let $F : \mathbb{D}^* \to X$ be a holomorphic set-valued mapping from the pointed unit disc $\mathbb{D}^* = \mathbb{D}\backslash\{0\}$ to a Stein manifold X. Assume that the limit set $\overline{\Gamma}_F \cap \{\zeta = 0\}$ is compact. Then F has a unique holomorphic extension to $\mathbb{D}$ defined by setting $\widetilde{F}(0) = \overline{\Gamma}_F \cap \{\zeta = 0\}$.*

As another consequence we derive that the empty set has a special role as value holomorphic set-valued mappings.

Corollary 2.7. *Let $F : \mathbb{D} \to X$ be a holomorphic set-valued mapping. If $F(\zeta) = \emptyset$ holds for some $\zeta \in \mathbb{D}$, then $F \equiv \emptyset$.*

Proof. Note that the set $F^{-1}(\emptyset)$ of all $\zeta \in \mathbb{D}$ where $F(\zeta) = \emptyset$ is open. If F is a function violating the statement, we can select a round open disc $G \Subset \mathbb{D}$ centered at some point $\zeta_0 \in F^{-1}(\emptyset)$ but not contained in $F^{-1}(\emptyset)$. The function $g(\zeta) = \frac{1}{\zeta - \zeta_0}$ is holomorphic near the closure of $U = G \times X$ and fulfils

$$\max_{\Gamma_F \cap \overline{U}} |g| > \max_{\Gamma_F \cap \partial U} |g|,$$

in contradiction to Lemma 2.2. $\square$

Let us finally mention a strengthening of Theorem 2.5, which follows from a straightforward extension of the proof.

Theorem 2.8. *Let $F : R \to X$ be a set-valued upper semicontinuous mapping from a noncompact Riemann surface R to a Stein manifold X, and let E be a polar subset of R. If F restricted to $R \backslash E$ is holomorphic, then it is holomorphic on R.*

3. Polynomial hulls and positive currents

Let K be a compact subset of a reduced complex space X. The set

$$\widehat{K}_{\mathcal{O}(X)} = \{z \in X : |f(z)| \leq \max_K |f| \text{ for every } f \in \mathcal{O}(X)\},$$

is called the *holomorphic hull of K in X* or $\mathcal{O}(X)$-hull. Specifically, $\widehat{K}_{\mathcal{O}(\mathbb{C}^n)}$ is called *polynomial hull* and abbreviated by $\widehat{K}$, and K is *polynomially convex* if $K = \widehat{K}$. A reason for the particular interest of polynomial convexity is that holomorphic convexity in Stein spaces can be reduced to polynomial convexity in $\mathbb{C}^N$, using a proper holomorphic embedding $X \hookrightarrow \mathbb{C}^N$. We will often use that on Stein spaces, the holomorphic hull coincides with the hull defined with respect to plurisubharmonic functions $PSH(X)$. The theory of polynomial convexity has been thoroughly studied, see [18] for an excellent presentation.

Let us elaborate on the observation made in the introduction that one can the replace normality assumption by irreducibility without essentially changing the problem.

Lemma 3.1. *Let X be a pure-dimensional reduced Stein space. Assume that X_{sing} only contains one point at which X is locally irreducible. Then the thickening phenomenon is valid on X if and only if it is valid on its normalisation $\widetilde{X}$.*

For simplicity, we restrict the statement to the case of one singular point. Anyway, the first step in the proof of Theorem 5.1 will localise the general problem at the individual singular points.

Proof. Note that it is not immediate to show the lemma by directly working with holomorphic functions on the two levels. However, there is a short argument based on the fact that $\mathcal{O}(X)$-hulls coincide with plurisubharmonic hulls on Stein spaces. At a singularity, plurisubharmonic functions are those upper semicontinuous functions $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ for which $\varphi \circ \iota$ is subharmonic for every parametrised holomorphic curve $\iota : \mathbb{D} \to X$. The canonical mapping $\widetilde{X} \to X$ is a homeorphism which is biholomorphic on $\widetilde{X}_{\text{reg}}$. Since parametrised holomorphic curves can be lifted and pushed down between X and $\widetilde{X}$, we get a canonic identification between the respective plurisubharmonic functions, which immediately implies the lemma. $\square$

In Section 6.1, we will build on a characterisation of $\widehat{K}$ obtained by Duval and Sibony [5]. Recall some basic terminology on currents: For a complex manifold X of dimension n, we denote by $\mathcal{D}^{p,q}(X)$ the space of (p,q)-forms with compact support and by $\mathcal{D}_{n-p,n-q}(X)$ its topological dual space, the currents of bidimension (p,q) and bidegree $(n-p,n-q)$. If $X = \mathbb{C}^n$ elements in $\mathcal{D}_{n-p,n-q}(X)$ can be identified with $(n-p,n-q)$-forms whose coefficients are distributions. A current T of bidimension (p,p) is *positive* if

$$T[f\ i\alpha_1 \wedge \bar{\alpha}_1 \wedge \ldots \wedge i\alpha_p \wedge \bar{\alpha}_p] \geq 0$$

holds whenever f is a smooth nonnegative function and $\alpha_j \in \mathcal{D}^{1,0}(X)$.

The standard differential operators d, d^c, ∂ and $\overline{\partial}$ extend naturally to currents: T is called *closed* if $dT = 0$, *pluriharmonic* if $dd^cT = 0$, *plurisubharmonic* if dd^cT is a positive current, and *plurisuperharmonic* if $-dd^cT$ is a positive current. We refer to [9] for detailed information on pluripotential theory.

Historically the first strong evidence for the relevance of currents in several complex variables was Lelong's discovery that integration over analytic subsets yields closed positive currents. Certainly analytic sets with boundary in K are contained in $\widehat{K}$, but they do not cover $\widehat{K}$ in general. Whereas the traces of closed positive currents on $\mathbb{C}^n \backslash \widehat{K}$ are still not enough, positive plurisuperharmonic currents of bidimension $(1,1)$ are sufficient. More precisely, Duval and Sibony proved in [5] that for $z_0 \in \widehat{K}$ there is a positive $(1,1)$-current T with bounded support such that $dd^cT = \mu - \delta_{z_0}$ where μ is a representing measure of z_0 with support contained in K. In the sequel, we will refer to these currents as *Duval-Sibony currents*. We mention that there is a related characterisation of polynomials discovered independently by Poletsky [14] and Bu and Schachermacher [1], see also [7] for the relation with Duval-Sibony currents.

The equation $dd^cT = \mu - \delta_{z_0}$ implies that the restriction of T to $\mathbb{C}^n \backslash K$ is a plurisuperhamonic positive $(1,1)$-current. In [4] it is proved that the support of such currents is pseudoconcave. Hence, Lemma 2.2 (applied to $|p|$, $p \in \mathbb{C}[z_1, \ldots, z_n]$) shows that $\operatorname{supp}(T) \subset \widehat{K}$. If K is contained in some analytic subset X of $\mathbb{C}^n$, then the same holds for $\operatorname{supp}(T)$.

We close the section by looking at the thickening phenomenon in complex dimension 1.

Proposition 3.2. *Let X be a noncompact purely 1-dimensional reduced and locally irreducible complex space. If K and L are compact subsets of X such that $K \subset \operatorname{int}(L)$, then $\widehat{K} \subset \operatorname{int}(\widehat{L})$.*

Proof. Since thickening happens at regular points, see the first step in the proof of Theorem 5.1, we only have to prove $0 \in \text{int}(L)$ in case $0 \in \widehat{K}$. This follows from Lemma 3.1 since the normalisation of holomorphic curves is smooth and the thickening phenomenon is valid for Stein manifolds.

An alternative option is to apply the following lemma, which gives slightly more insight.

Lemma 3.3. *Let X be as in Proposition 3.2 and $K \subset X$ compact. Then $\widehat{K}_{\mathcal{O}(X)}$ is the union of the relatively compact connected components of $X \backslash K$.*

Indeed, the lemma implies that every point in $K \backslash \widehat{K}_{\mathcal{O}(X)}$ is already an interior point of $\widehat{K}_{\mathcal{O}(X)}$, which is stronger than the thickening property.

Proof of Lemma 3.3. The lemma is well known for $X = \mathbb{C}$. To prove it in the general case, we first embed X properly in $\mathbb{C}^N$. Then we observe that the lemma immediately follows from Stolzenberg's theorem on the polynomial hull of smooth curves (see [2] for detailed information) in the case K that is a disjoint union Γ of smoothly embedded circles $\Gamma_1, \ldots, \Gamma_k$ in X_{reg}.

For K general we denote by V the union of the relatively compact components of $X \backslash K$. The maximum modulus principle on X implies that $V \subset \widehat{K}$. It remains to show that $x_0 \notin \widehat{K}$ for a given $x_0 \in X \backslash (K \cup V)$. There is a relatively compact subdomain U_K of X such that $x_0 \notin \overline{U}_K$ and ∂U_K is a disjoint union of smooth circles in X_{reg}. The first part of the proof implies that

$$x_0 \notin \widehat{\partial U_K} = \widehat{\overline{U}}_K \supset \widehat{K},$$

completing the proof of Lemma 3.2 and Proposition 3.2. □

4. Approach directions of hulls

Let X be a complex space with an isolated singularity at $x \in X$. Recall that a *resolution* of the singularity x is a complex manifold X' equipped with a proper holomorphic mapping $\pi : X' \to X$ onto a neighbourhood $U^{\text{open}} \subset X$ of x such that $\pi|_{X' \backslash \pi^{-1}(x)} : X' \backslash \pi^{-1}(x) \to U \backslash \{x\}$ is biholomorphic. The set $E = \pi^{-1}(x)$ is called the *exceptional divisor* of the resolution.

Definition 4.1. A set $S \subset X$ is said to be *directionally dense at x* if for every resolution $\pi : X' \to X$ of the singularity x the closure of $\pi^{-1}(S \backslash \{x\})$ contains the exceptional divisor E.

Of course, S is directionally dense at x if $x \in \text{int}(S)$. The terminology is motivated by the situation for conical singularities, where directional density implies nontrivial intersection with every open local cone with vertex at the singularity. However, the definition gives much more since it is formulated with respect to *all* resolutions at x. Note also that the concept is already reasonable (and nontrivial) if $x \in X_{\text{reg}}$.

Example 4.2. a) A complex curve C through x is *never* directionally dense since its strict transform[1] $\widetilde{C} \subset X'$ intersects E only in isolated points.

b) A real hyperplane H in $\mathbb{C}^2$ is both directionally dense and locally polynomially convex at each point. To prove directional density at some point, say $0 \in H$, we fix a resolution $\pi : X' \to \mathbb{C}^2$ of $\mathbb{C}^2$ at 0 and pick, for a given $e \in E \subset X'$, a local parametrised complex curve $\gamma : \mathbb{D} \to X'$ which meets E precisely at e. By the open mapping principle for holomorphic functions, the image of $\pi \circ \gamma$ is either contained in H or intersects both of its sides (the connected components of $\mathbb{C}^2 \backslash H$). In the first case $\pi \circ \gamma$ has image contained in one of the complex lines foliating H; in the second case the image intersects H in a 1-dimensional real analytic set. Hence $e \in \overline{\pi^{-1}(H \backslash \{0\})}$.

c) The closure of a smoothly bounded strictly pseudoconvex domain $D \Subset \mathbb{C}^2$ is not directionally dense at any of its boundary points. Indeed, after a quadratic ccordinate change, we can assume that $D = \{\mathrm{Im}(z_2) > h(z_1, \mathrm{Re}(z_2))\}$ holds near 0, with a strictly convex function h vanishing up to first order at 0. Now we construct $\pi : X' \to \mathbb{C}^2$ by two successive modifications. First we blow up the origin, yielding

$$\pi_1 : X_1 = \{((z_1, z_2), (\zeta_1 : \zeta_2)) \in \mathbb{C}^2 \times \mathbb{P} : z_1\zeta_2 = z_2\zeta_1\} \to \mathbb{C}^2$$

with the projection π_1 onto the first factor, and then the point $(0, (1 : 0))$ corresponding to the z_1-axis. If we express the projections locally by

$$z_1 = u_1, z_2 = u_1v_1, \quad \text{and} \quad u_1 = u_2, v_1 = u_2v_2,$$

the curves $v_2 = cu_2$ through $(u_2, v_2) = (0, 0)$ are mapped to $z_2 = cz_2^2$, which have 0 as an isolated intersection with $\overline{D}$ if $|c|$ is small enough. Hence a neighbourhood of $(u_2, v_2) = (0, 0) \in E$ is disjoint from the closure of $\pi^{-1}(\overline{D} \backslash \{0\})$.

d) It is straightforward to obtain examples in *singular* spaces corresponding (b) and (c), by taking suitable intersections with real hyperplanes and strictly pseudoconvex domains in a local model.

5. Weak thickening at singularities

The following theorem is the main result of the article.

Theorem 5.1. *Let X be a reduced normal Stein space of pure dimension 2 with an isolated irreducible singularity at $\hat{x}$. If K and L are compact subsets of X such that $\hat{x} \in \widehat{K}_{\mathcal{O}(X)}$ and $K \subset \mathrm{int}(L)$, then $\widehat{L}_{\mathcal{O}(X)}$ is directionally dense at $\hat{x}$.*

Proof. If $\hat{x} \in K$ then $\hat{x} \in \mathrm{int}(L)$, and there is nothing to prove. From now on, we assume that $\hat{x} \in \widehat{K}_{\mathcal{O}(X)} \backslash K$.

Step 1: Localisation at the singular point. Our first goal is to localise to the case where $\hat{x} = 0 \in \mathbb{C}^N$ and X is an analytic subset of some neighbourhood of the origin. We sketch how this can be achieved and refer to [10] for the details.

[1]The strict transform of a local complex curve C is the closure of $\pi^{-1}(C \backslash \{x\})$ in X'.

Following [16], we consider the coherent sheaf $\mathcal{T}_X$ of germs of holomorphic vector fields on X. By Theorem A, the stalks $\mathcal{T}_{X,(x)}$ are generated by global holomorphic vector fields. In particular, this permits to find global holomorphic vector fields v_1, v_2 such that $v_1(x), v_2(x)$ span the holomorphic tangent space $T_x^{1,0}X$ at a given $x \in X_{\text{reg}}$. Using the dynamic flows of sufficiently many global vector fields (more precisely of their real parts), this allows us to show that

$$\widehat{K}_{\mathcal{O}(X)} \cap X_{\text{reg}} \subset \text{int}\left(\widehat{L}_{\mathcal{O}(X)}\right),$$

meaning that the thickening phenomenon is valid at regular points.

After choosing a proper holomorphic embedding, we can assume that $X \subset \mathbb{C}^N$ and $\hat{x} = 0$. In particular, $\mathcal{O}(X)$-hulls of compacts in X coincide with polynomial hulls in the ambient space. For $0 < \epsilon \ll 1$ we have that

$$\{0\} = X_{\text{sing}} \cap B_\epsilon(0), \ \overline{B}_\epsilon(0) \cap K = \emptyset,$$

and that $\partial B_\epsilon(0)$ intersects X transversally. By Rossi's local maximum modulus theorem we have $\widehat{K} \cap \overline{B}_\epsilon(0) = (\widehat{K} \cap \partial B_\epsilon(0))\widehat{\ }$ and the corresponding equality for $\widehat{L}$. From the thickening property at regular points, we deduce that $\widehat{K} \cap \partial B_\epsilon(0) \subset \text{int}_X(\widehat{L})$, where $\text{int}_X(\cdot)$ denotes the set of relatively interior points with respect to X. Hence it suffices to prove Theorem 5.1 for $\widehat{K} \cap \partial B_\epsilon(0)$ at the place of K.

Step 2: Microlocalisation at the exceptional divisor. By Step 1 we can assume that $X \subset \mathbb{C}^N$ is a subvariety of a ball $B_{2\epsilon_0}(0) \subset \mathbb{C}^N$, has a unique singularity at $\hat{x} = 0$ and that $K \subset \partial B_{\epsilon_0}(0)$. Taking ϵ_0 small enough, we can in addition assume that the spheres $\partial B_r(0)$, $0 < r < 2\epsilon_0$, are all transverse to X.

Note that we have to argue with respect to an *arbitrary* resolution $\pi : X' \to X$. A fortiori, it is enough to prove the theorem for manifolds which are obtained from the given X' by finitely many successive blow-ups at points of the exceptional divisor. By Zariski's connectedness theorem the exceptional divisor $E = \pi^{-1}(0)$ is a connected union of irreducible complex curves. Since complex curves in a smooth complex surface can be desingularised by successive blow-ups of the ambient space, we can and will assume that $E = \pi^{-1}(0)$ is a *normal crossing divisor*, i.e. a finite union $E_1 \cup \ldots \cup E_k$ of smooth complex curves E_j such that two of them transversally intersect in at most one point.

To argue by contradiction, we assume that the theorem fails for $\pi : X' \to X$ as arranged above and some L with $K \subset \text{int}_X L$, meaning that

$$E^* = E \setminus E_L \neq \emptyset, \text{ where } E_L = \overline{\pi^{-1}(\widehat{L} \setminus \{0\})} \cap E.$$

On the other hand, we know that $E_L \neq \emptyset$ since $0 \in \widehat{K} \subset \widehat{L}$ and E is compact. We deduce that there is an irreducible component of E, which we may assume to be E_1, such that both $E_1 \cap E^*$ and $E_1 \cap E_L$ are nonempty. Otherwise, the disjoint sets E^* and E_L would both be unions of irreducible components of E, which is impossible since E is connected.

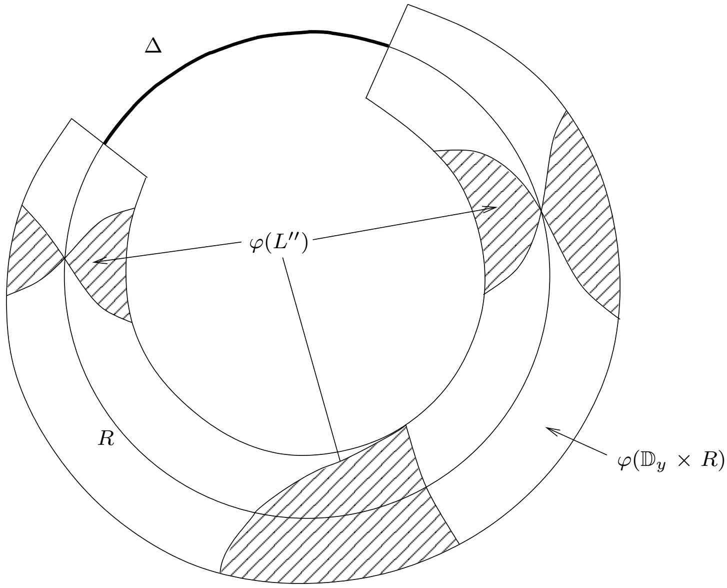


FIGURE 2. Neighborhood of $R = E_1 \backslash \overline{\Delta}$

Step 3: Removability along E_1. Let $\Delta \subset E_1$ be an open disc with $\overline{\Delta} \subset E^*$. Using the triviality of holomorphic vector bundles over noncompact Riemann surfaces combined with [17], we construct a biholomorphism

$$\varphi : \mathbb{D}_\zeta \times R \to U,$$

where $R = E_1 \backslash \overline{\Delta}$ and $U^{\text{open}} \subset X'$, such that

(i) $\varphi(0, x) = x$, for all $x \in R$,

(ii) there is a compact $F_R \subset R$ such that

$$L' = (\pi \circ \varphi)^{-1}(\widehat{L} \backslash \{0\}) \subset \mathbb{D}_\zeta \times F_R\,.$$

This can be arranged for every compact $F_R \subset R$ which contains E_L in its interior (see Figure 2).

Observe that L' is pseudoconcave in $\mathbb{D}^*_\zeta \times R$ since pseudoconcavity is a local property and stable under biholomorphisms. Hence Theorem 2.5, with L' instead of A, $\mathbb{D}_\zeta \times R$ instead to X, and $\{\zeta = 0\}$ instead of Y, yields that $L'' = \overline{L'} = L' \cup E_L$ is pseudoconcave in $\mathbb{D}_\zeta \times R$.

From Lemma 2.2, we derive that

$$L'' \cap \{|\zeta| \leq 1/2\} \subset (L' \cap \{|\zeta| = 1/2\})\widehat{}_{\mathcal{O}(\mathbb{D}_\zeta \times R)},$$

meaning in particular that the $\mathcal{O}(\mathbb{D}_\zeta \times R)$-hull of $L' \cap \{|\zeta| \leq 1/2\}$ has nontrivial intersection with $\{\zeta = 0\}$.

Remark 5.2. The crucial point of Step 3 is the transition from $\mathcal{O}(X)$-hulls to hulls with respect to the *strictly richer* space $\mathcal{O}(\mathbb{D}_\zeta \times R) \supsetneq (\varphi \circ \pi)^* \mathcal{O}(X)$. Theorem

2.5 rules out the a priori possibility that $\widehat{(L'\cap\{|\zeta|=1/2\})}_{\mathcal{O}(\mathbb{D}_\zeta\times R)}$ is smaller than the closure of $L'\cap\{|\zeta|\le 1/2\}$.

Step 4: Thickening the hull in $\mathbb{D}_\zeta\times R$. Now we are in a position to apply thickening results for holomorphic hulls on the Stein manifold $\mathbb{D}_\zeta\times R$, which is a byproduct of Step 1 for the case $X_{\text{sing}}=\emptyset$.

Choose a second compact H such that $K\subset\operatorname{int}(H)\subset H\subset\operatorname{int}(L)$. Applying the reasoning in Step 3 to H instead of L, we obtain that the $\mathcal{O}(\mathbb{D}_\zeta\times R)$-hull of

$$H_{1/2}=(\varphi\circ\pi)^{-1}(\widehat{H})\cap\{|\zeta|=1/2\}$$

has nonvoid intersection with $\{\zeta=0\}$. The thickening phenomenon at points of X_{reg} implies that $H_{1/2}$ is contained in the interior of

$$L_{1/4,\,3/4}=(\varphi\circ\pi)^{-1}(\widehat{L})\cap\{1/4\le|\zeta|\le 3/4\}.$$

Thus thickening of hulls in the Stein manifold $\mathbb{D}_\zeta\times R$ gives that

$$\widehat{(H_{1/2})}_{\mathcal{O}(\mathbb{D}_\zeta\times R)}\subset\operatorname{int}\left(\widehat{(L_{1/4,\,3/4})}_{\mathcal{O}(\mathbb{D}_\zeta\times R)}\right).$$

In particular, we have shown that the relative interior of $\widetilde{L}=\overline{\pi^{-1}(\widehat{L}\backslash\{0\})}$ has nontrivial intersection with E_1.

Step 5: Propagation along E_1. For completing the proof it suffices to show that

$$E_1\subset\operatorname{int}(\widetilde{L}),\tag{5.1}$$

which is in contradiction to the choice of E_1. Since $\operatorname{int}(\widetilde{L})\cap E_1\neq\emptyset$ holds by the previous step, (5.1) follows from [10, Lem. 4.2]. Similarly as in Step 2, we construct domains $R_1\Subset R_2\subsetneq E_1$ with smooth boundaries contained in $E_1\cap\operatorname{int}(\widetilde{L})$ and with $E_1\subset\operatorname{int}(\widetilde{L})\cup R_1$, together with a biholomorphism

$$\psi:\mathbb{D}\times R_2\to U_2^{\text{open}}\subset X'$$

such that

- **(iii)** $\psi(0,x)=x$, for all $x\in R_2$,
- **(iv)** $\psi(\mathbb{D}\times\partial R_1)\subset\operatorname{int}(\widetilde{L})$.

Applying the maximum modulus principle along the bordered complex curves $\psi(\{\zeta\}\times R_1)$, we deduce that $\psi(\mathbb{D}^*\times R_1)\subset\pi^{-1}(\widehat{L}\backslash\{0\})$. This yields (5.1), and the proof of Theorem 5.1 is complete. □

By applying the above reasoning directly to $\widehat{K}$ instead of $\widehat{L}$, we obtain the following corollary as a byproduct.

Corollary 5.3. *Let X be a reduced normal Stein space of pure dimension 2 with an isolated irreducible singularity at $\hat{x}$. Assume that K is a compact subset of X such that $\hat{x}\in\widehat{K}_{\mathcal{O}(X)}$ and $\widehat{K}_{\mathcal{O}(X)}$ is* not *directionally dense at $\hat{x}$. Then for every compact $L\subset X$ with $K\subset\operatorname{int}(L)$, we have $\hat{x}\in\operatorname{int}(\widehat{L}_{\mathcal{O}(X)})$.*

Hence we obtain a partial positive answer to the thickening problem for compacts K whose hull is in a suitable sense thin at the singularities. In the subsequent section, we will strengthen this observation.

6. Hulls with tame parts

Let X be a reduced pure-dimensional Stein space with an isolated irreducible singularity at $\hat{x}$. We say that the hull of a compact K with $\hat{x} \in \widehat{K}_{\mathcal{O}(X)}$ has the *thickening property at* $\hat{x}$ if for every compact L with $K \subset \mathrm{int}_X(L)$ we have that $\hat{x} \in \mathrm{int}_X(\widehat{L}_{\mathcal{O}(X)})$. In earlier work [10] we proved the thickening property for the hulls of compacts which have analytic structure at $\hat{x}$, meaning that $\widehat{K}_{\mathcal{O}(X)}$ contains a piece of a holomorphic curve passing through $\hat{x}$.

Whereas holomorphic hulls do not have holomorphic structure in general, they are the union of supports of Duval-Sibony currents, as summarised in Section 3. We say that $\widehat{K}_{\mathcal{O}(X)}$ has a *tame part at* $\hat{x}$ if there is a Duval-Sibony current such that the measure $\lambda = dd^cT + \delta_{\hat{x}}$ is contained in $\widehat{K}_{\mathcal{O}(X)} \backslash \{\hat{x}\}$ and is not directionally dense at $\hat{x}$. Obviously, hulls with analytic structure at $\hat{x}$ have a tame part there.

The following is a strengthening in dimension 2of our earlier result [10, Prop.4.1] on hulls with analytic structure.

Theorem 6.1. *Let X be a reduced Stein space of pure dimension 2 with an isolated irreducible singularity at $\hat{x}$. Then the hull of every compact $K \subset X$ which has a tame part at $\hat{x}$ has the thickening property at $\hat{x}$.*

Proof. The reasoning follows similar lines like in the proof of Theorem 5.1, with $\mathrm{supp}(T)$ playing the role of $\widehat{K}$. We briefly sketch how to adapt the arguments. Note that it is sufficient this time to argue with a fixed resolution $\pi : X' \to X$ with normal crossing exceptional divisor E.

Localisation can be achieved like in the Steps 1 and 2 in the proof of Theorem 5.1, allowing us to assume that X is a subvariety of $B_{2\epsilon_0} \subset \mathbb{C}^N$ with an isolated singularity at 0 and that $K \subset \partial B_{\epsilon_0}$. Hence we can work with and a plurisuperharmonic $(1,1)$-current T obtained by restricting the original current to B_{ϵ_0}. Finally we set $S^* = \pi^{-1}(\mathrm{supp}(T)\backslash\{0\})$, $S = \overline{S^*}$ and select an irreducible component $E_1 \subset E = \pi^{-1}(0)$ such that $\emptyset \neq S \cap E_1 \subsetneq E_1$.

Next we construct a holomorphic embedding

$$\varphi : \mathbb{D} \times R \hookrightarrow X'$$

such that R is a proper subdomain of E_1, we have $\varphi(0,x) = x$ for $x \in R$, and such that the projection to the second factor in $\mathbb{D} \times R$ maps $\varphi^{-1}(S)$ onto a relatively compact subset of R. As before we derive from Theorem 2.5 that $\varphi^{-1}(S)$ is pseudoconcave and obtain that $E_1 \subset \mathrm{int}\left(\overline{\pi^{-1}(\widehat{L}\backslash\{0\})}\right)$ for any compact L with $K \subset \mathrm{int}(L)$ by the argument in Step 5. Iterated use of this argument yields that

$E \subset \operatorname{int}\left(\overline{\pi^{-1}(\widehat{L}\setminus\{0\})}\right)$ for any compact L with $K \subset \operatorname{int}(L)$ since E is connected. The proof Theorem 6.1 is complete. □

Acknowledgements

The second author is grateful to Jean Ruppenthal for enlighting discussions and help with informations on singularities.

References

[1] S. Q. Bu, W. Schachermacher, *Approximation of Jensen measures by image measures under holomorphic functions and applications*, Trans. Amer. Math. Soc. **331** (1992) 585–608.
[2] E. M. Chirka, *Complex Analytic Sets*, Mathematics and its applications **46**. Springer, Netherlands 2011.
[3] M. Colţoiu, M. Tibar, *Steiness of the universal covering of the complement of a 2-dimensional complex singularity*, Math. Ann. **326** (2003), 95–104.
[4] T.-C. Dinh, M. G. Lawrence, *Polynomial hulls and positive currents*, Ann. Toulouse **12** (2003), 317–334.
[5] J. Duval, N. Sibony, *Polynomial convexity, rational convexity, and currents*, Duke Math. J. **79** (1995), 487–513.
[6] J. E. Fornæss, N. Sibony, *Some open problems in higher-dimensional complex analysis and* Publ. Mat. **45** (2001), 529–547.
[7] E. Fornæss-Wold, *A note on polynomial convexity: Poletsky disks, Jensen measures and positive currents*. J. Geom. Anal. **21** (2011), 252–255.
[8] F. Hartogs, *Über die aus den singulären Stellen einer analytischen Funktion mehrerer Veränderlicher bestehenden Gebilde*, Acta Math. **32** (1909), 57–79.
[9] M. Klimek, *Pluripotential Theory*, Clarendon Press 1991.
[10] A. Lind, E. Porten, *On thickening of holomorphic hulls and envelopes of holomorphy on Stein spaces*, Intern. J. Math. **27** (2016), 1650051.
[11] M. Nacinovich, E. Porten, *$\mathcal{C}^\infty$-hypoellipticity and extension of CR functions*, Ann. Sc. Norm. Sup. Pisa **14** (2015), 677–703.
[12] T. Nishino, *Sur les ensembles pseudoconcaves*, J. Math. Kyoto Univ. **1** (1962), 225–245.
[13] K. Oka, *Note sur les familles de fonctions analytiques multiformes etc.*, J. Sci. Hiroshima Univ. **4** (1934), 93–98.
[14] E. A. Poletsky, *Holomorphic currents*, Indiana Univ. Math. J. **42** (1993), 85–144.
[15] R. Richberg, *Stetige streng pseudokonvexe Funktionen*, Math. Ann. **175** (1968), 251–286.
[16] H. Rossi, *The local maximum modulus principle*, Ann. Math. **72** (1960), 1–11.
[17] Y.-T. Siu, *Every Stein subvariety admits a Stein neighborhood*, Invent. Math. **38** (1976), 89–100.
[18] E. L. Stout, *Polynomial Convexity*, Birkhäuser 2007.
[19] Z. Złodkowski, *Analytic set valued functions and spectra*, Math. Ann. **256** (1981), 363–386.

Mid Sweden University, Sweden
andreas.lind@miun.se

Mid Sweden University, Sweden
egmont.porten@miun.se

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

NEWTON DIAGRAMS AND THE GEOMETRIC DEGREE OF A POLYNOMIAL MAPPING OF $\mathbf{C}^2$

MATEUSZ MASTERNAK

ABSTRACT. Let $H = (f, g) : \mathbf{C}^2 \to \mathbf{C}^2$ be a polynomial mapping. We give a formula for the geometric degree of H in terms of the Newton diagrams of f and g. We say that the mapping H is proper if for any compact set $K \subset \mathbf{C}^2$ its preimage $H^{-1}(K)$ is compact. In the paper a criterion of properness of the mapping H is also given.

KEYWORDS: polynomial mapping, geometric degree, Newton diagram

MSC2010: 11R09, 14E05, 52B20, 14M25, 33C70,

Received 04 May 2021; revised 14 July 2021; accepted 16 July 2021

1. INTRODUCTION

Let $f(X,Y) = \sum c_{\alpha\beta}X^\alpha Y^\beta \in \mathbf{C}[X,Y]$ be a nonzero polynomial of positive degree. We say that the polynomial f is *quasi–convenient (convenient)* if $c_{\alpha 0} \neq 0$ and $c_{0\beta} \neq 0$ for some integers $\alpha, \beta \geq 0$ $(\alpha, \beta > 0)$. Otherwise $f(X,Y) = X^s Y^t \tilde{f}(X,Y)$ for some nonnegative integers s and t, where $\tilde{f}$ is a quasi-convenient polynomial or it is a nonzero constant.

Let $\mathrm{supp} f := \{(\alpha, \beta) \in \mathbf{N}^2 : \ c_{\alpha\beta} \neq 0\}$. We define

$$\Delta_\infty(f) := \mathrm{convex}(\{(0,0)\} \cup \mathrm{supp} f).$$

The polygon $\Delta_\infty(f)$ is called *the Newton diagram at infinity* of the polynomial f.

For any nonzero vector $\vec{w} = [p, q]$ of the real plane $\mathbf{R}^2$ we put

$$\mathrm{in}(f, \vec{w})(X,Y) := \sum_{p\alpha + q\beta = d_{\vec{w}}(f)} c_{\alpha\beta}X^\alpha Y^\beta,$$

https://doi.org/10.2478/9788366675360-015.

where

$$d_{\vec{w}}(f) = \max\{p\alpha + q\beta : (\alpha,\beta) \in \mathrm{supp} f\}.$$

A pair (f,g) of quasi–convenient polynomials is *nondegenerate at infinity* if for any real vector $\vec{w} = [p,q]$ such that $p > 0$ or $q > 0$ the system of equations $\mathrm{in}(f,\vec{w})(X,Y) = \mathrm{in}(g,\vec{w})(X,Y) = 0$ has no solutions in $\mathbf{C}^* \times \mathbf{C}^*$, where $\mathbf{C}^* = \mathbf{C} \setminus \{0\}$.

Let $f(X,Y)$, $g(X,Y) \in \mathbf{C}[X,Y]$ be polynomials of positive degrees. If $P = (a,b) \in \mathbf{C}^2$ is a solution of the system

$$f(X,Y) = 0, \quad g(X,Y) = 0$$

then the symbol $(f,g)_P$ denotes the intersection multiplicity of f and g at P. We use the definition of the intersection multiplicity as in [4]. In particular

$$(f,g)_{(0,0)} := \dim_{\mathbf{C}} {}^{\mathbf{C}\{X,Y\}}/_{\mathrm{I}(f,g)},$$

where $\mathbf{C}\{X,Y\}$ denotes the ring of convergent power series in two complex varaibles and $\mathrm{I}(f,g)$ is the ideal generated by f and g in $\mathbf{C}\{X,Y\}$. We have $(f,g)_P < +\infty$ if and only if P is an isolated solution of the given system.

For a pair of quasi–convenient polynomials $f(X,Y)$, $g(X,Y) \in \mathbf{C}[X,Y]$ we denote by $\nu_\infty(f,g)$ the double Minkowski mixed area (see [13]) of the diagrams $\Delta_\infty(f)$ and $\Delta_\infty(g)$, i.e.

$$\nu_\infty(f,g) := \mathrm{Area}(\Delta_\infty(f) + \Delta_\infty(g)) - \mathrm{Area}\Delta_\infty(f) - \mathrm{Area}\Delta_\infty(g),$$

where $+$ denotes the Minkowski sum. Let us note that if polynomials f, $g \in \mathbf{C}[X,Y]$ are quasi–convenient then $\Delta_\infty(fg) = \Delta_\infty(f) + \Delta_\infty(g)$ (see [11], Proposition 2.1), so $\nu_\infty(f,g) = \mathrm{Area}(\Delta_\infty(fg)) - \mathrm{Area}\Delta_\infty(f) - \mathrm{Area}\Delta_\infty(g)$.

Let us cite the following

Lemma 1.1. [11, Lemma 3.2] *If f, $g \in \mathbf{C}[X,Y]$ are quasi–convenient polynomials of positive degrees then*

(1) $\nu_\infty(f,g) \geq 0$,
(2) $\nu_\infty(f,g) = 0$ *if and only if the diagrams $\Delta_\infty(f)$ and $\Delta_\infty(g)$ are segments included in the same straight line passing through the origin.*

2. Formula for the geometric degree and criterion of the properness of a polynomial mapping

Let $H = (f,g) : \mathbf{C}^2 \to \mathbf{C}^2$ be a polynomial mapping. If the fiber $H^{-1}(z)$, where $z = (u,v) \in \mathbf{C}^2$ is finite then we put

$$\deg H^{-1}(z) := \sum_{P \in H^{-1}(z)} (f-u, g-v)_P.$$

By convention $\deg H^{-1}(z) = 0$ if $H^{-1}(z) = \emptyset$. According to Sard's theorem $\deg H^{-1}(z) = \#H^{-1}(z)$ for almost all $z \in \mathbf{C}^2$. We define *the geometric degree* of the mapping H:

$$d(H) := \sup\{\,\deg H^{-1}(z) : z \in \mathbf{C}^2 \text{ and } \#H^{-1}(z) < +\infty\,\}.$$

In [12] there are proofs of the following propositions:

Proposition 2.1. [12, Proposition 2.1] *Let $H = (f, g) : \mathbf{C}^2 \to \mathbf{C}^2$ be a polynomial mapping. Then the set $\{\, z \in \mathbf{C}^2 : \#H^{-1}(z) = +\infty \,\}$ is a proper algebraic subset of $\mathbf{C}^2$ and $d(H) < +\infty$. Moreover, the set $\{\, z \in \mathbf{C}^2 : \deg H^{-1}(z) = d(H) \,\}$ is a Zariski open, non-empty subset of $\mathbf{C}^2$. The mapping H is dominating (i.e., f and g are algebraically independent) if and only if $d(H) > 0$. Moreover, $d(H) = 0$ if and only if $H(\mathbf{C}^2) = \{\, z \in \mathbf{C}^2 : \#H^{-1}(z) = +\infty \,\}$.*

If $\#H^{-1}(z) < +\infty$, we put $\delta_z(H) := d(H) - \deg H^{-1}(z)$. The number $\delta_z(H)$ is called *the defect of the fiber* $H^{-1}(z)$. By Proposition 2.1 $\delta_z(H) \geq 0$ for each z such that $\#H^{-1}(z) < +\infty$ and $\delta_z(H) = 0$ for almost all $z \in \mathbf{C}^2$.

By definition, a mapping $H : \mathbf{C}^2 \to \mathbf{C}^2$ is *proper* if for any compact set $K \subset \mathbf{C}^2$ its preimage $H^{-1}(K)$ is compact. We say, that a mapping H is *proper over zero* if there exists a compact neighbourhood of the zero V such that its preimage $H^{-1}(V)$ is compact.

Proposition 2.2. [12, Proposition 2.2] *Keeping the above notations we have*

(i) *The mapping H is proper if and only if H has finite fibers and $\delta_z(H) = 0$ for all $z \in \mathbf{C}^2$,*

(ii) *The mapping H is proper over zero if and only if the fiber $H^{-1}(0)$ is finite and $\delta_0(H) = 0$.*

Now, we can state the main result.

Theorem 2.3 (Formula for the geometric degree)**.** *Let (f, g) be a pair of quasi–convenient polynomials. If for any real vector $\vec{w} = [p, q]$ such that $p > 0$ or $q > 0$ and that $d_{\vec{w}}(f)d_{\vec{w}}(g) \neq 0$ the system of equations*

$$\mathrm{in}(f, \vec{w})(X, Y) = \mathrm{in}(g, \vec{w})(X, Y) = 0$$

has no solutions in $\mathbf{C}^* \times \mathbf{C}^*$ then

$$d(H) = \nu_\infty(f, g),$$

where $H = (f, g) : \mathbf{C}^2 \to \mathbf{C}^2$.

Proof. We will use Kouchnirenko–Bernstein Theorem (see [7, 8, 9, 2, 1, 3, 5, 6]). Let us cite its two-dimensional version.

Theorem 2.4. [11, Theorem 2.2, Kouchnirenko–Bernstein] *Let* $f(X,Y)$, $g(X,Y) \in \mathbf{C}[X,Y]$ *be quasi–convenient polynomials. It holds*

(1) *if f and g are coprime then* $\sum_{P\in\mathbf{C}^2}(f,g)_P \leq \nu_\infty(f,g)$,

(2) $\sum_{P\in\mathbf{C}^2}(f,g)_P = \nu_\infty(f,g)$ *if and only if the pair* (f,g) *is nondegenerate at infinity.*

We have to check that the equality

$$\sum_{P\in\mathbf{C}^2}(f-u,g-v)_P = \nu_\infty(f,g)$$

holds for almost all $(u,v) \in \mathbf{C}^2$.

Let us fix $(u,v) \in \mathbf{C}^2$ such that $\tilde{f}(0,0) \neq 0$ and $\tilde{g}(0,0) \neq 0$ where $\tilde{f}(X,Y) = f(X,Y)-u$ and $\tilde{g}(X,Y) = g(X,Y)-v$. Obviously $\Delta_\infty(\tilde{f}) = \Delta_\infty(f)$ and $\Delta_\infty(\tilde{g}) = \Delta_\infty(g)$. The polynomials $\tilde{f}$ and $\tilde{g}$, similarly to f and g, are quasi–convenient and $\Delta_\infty(\tilde{f}\tilde{g}) = \Delta_\infty(fg)$, hence we have $\nu_\infty(\tilde{f},\tilde{g}) = \nu_\infty(f,g)$. Moreover, by our assumptions we state that the pair $(\tilde{f},\tilde{g})$ is nondegenerate at infinity for almost all $(u,v) \in \mathbf{C}^2$. To finish we apply Kouchnirenko-Bernstein Theorem for the pair of the polynomials $\tilde{f}$ and $\tilde{g}$. □

Remark 2.5. Notice that Theorem 2.3 allows to determine the geometric degree of the mapping $H = (f,g)$ without assumption of quasi–convenience of f and g. Namely, let u and v be complex numbers such that the polynomials $\tilde{f} = f-u$ and $\tilde{g} = g-v$ are quasi–convenient. Obviously $d(H) = d(\tilde{H})$ where $\tilde{H} = (\tilde{f},\tilde{g})$ and we can use Theorem 2.3 to compute $d(\tilde{H})$. Hence $d(H) = \nu_\infty(\tilde{f},\tilde{g})$. Let us notice that it does not have to be $\nu_\infty(\tilde{f},\tilde{g}) = \nu_\infty(f,g)$. Consider the following example: if $f(X,Y) = X$ and $g(X,Y) = Y$ then $f(X,Y)\,g(X,Y) = XY$ and $\Delta_\infty(fg)$ is the segment, while $\Delta_\infty(\tilde{f}\tilde{g})$ is the square, so $1 = \nu_\infty(\tilde{f},\tilde{g}) \neq \nu_\infty(f,g) = 0$.

Corollary 2.6. *If the polynomials f and g satisfy the assumptions of Theorem 2.3 and if the diagrams $\Delta_\infty(f)$, $\Delta_\infty(g)$ are not included at the same line passing through the origin then the mapping $H = (f,g)$ is dominating.*

Proof. It suffices to show that $d(H) > 0$ (see Proposition 2.1). According to Theorem 2.3 we have $d(H) = \nu_\infty(f,g)$. Therefore, it is enough to state that if the diagrams $\Delta_\infty(f)$ and $\Delta_\infty(g)$ are not included at the same line passing through the origin, then $\nu_\infty(f,g) > 0$, but it follows directly from Lemma 1.1. □

Another consequence of the formula for geometric degree is

Theorem 2.7.

(1) *If a pair of convenient polynomials (f,g) is nondegenerate at infinity then the mapping $H = (f,g) : \mathbf{C}^2 \to \mathbf{C}^2$ is proper.*

(2) *If a pair of quasi–convenient polynomials* (f, g) *is nondegenerate at infinity then the mapping* $H = (f, g) : \mathbf{C}^2 \to \mathbf{C}^2$ *is proper over zero.*

Proof. Let us note, for the proof of (1), that for any $(u, v) \in \mathbf{C}^2$ the polynomials $\tilde{f}(X, Y) = f(X, Y) - u$ and $\tilde{g}(X, Y) = g(X, Y) - v$ are conveniet, $\Delta_\infty(\tilde{f}) = \Delta_\infty(f)$, $\Delta_\infty(\tilde{g}) = \Delta_\infty(g)$ and that the pair $(\tilde{f}, \tilde{g})$ is nondegenerate at infinity. By Kouchnirenko–Bernstein Theorem (Theorem 2.4) we have $\deg H^{-1}(u, v) = \nu_\infty(f, g)$ for each $(u, v) \in \mathbf{C}^2$. Hence, $\deg H^{-1}(u, v) = d(H)$ for each $(u, v) \in \mathbf{C}^2$. It implies that the mapping H is proper (see Proposition 2.2, (i)).

Similarly, (2) follows from Kouchnirenko–Bernstein Theorem and from the formula for the geometric degree (Theorem 2.3). Namely, if a pair of quasi–convenient polynomials is nondegenerate at infinity, then $\deg H^{-1}(0, 0) = \sum_{P \in \mathbf{C}^2} (f, g)_P = \nu_\infty(f, g)$ and $d(H) = \nu_\infty(f, g)$. Therefore $\deg H^{-1}(0, 0) = d(H)$ but then the mapping H is proper over zero (see Proposition 2.2, (ii)). □

Using different methods, in [10], author gave the formula (Theorem 3.2.) for the Łojasiewicz exponent of a polynomial mapping $H = (f, g) : \mathbf{C}^2 \to \mathbf{C}^2$ in terms of Newton diagrams of its components assuming that the pair (f, g) is nondegenerate at infinity. Mentioned formula implies that if the polynomials f and g are convenient (quasi–convenient) then the Łojasiewicz exponent of the mapping H is positive (non-negative) number and then the mapping H is proper (proper over zero). Therefore Theorem 2.7 may be derived from the formula given in mentioned paper.

References

[1] M.F. Atiyah, *Angular momentum, convex polyhedra and algebraic geometry*, Proceedings of the Edinburgh Mathematical Society **26** (1983), 121–138.

[2] D.N. Bernstein, *The number of roots of a system of equations*, Funk. Anal. Appl. **9**.3 (1975)),1–4 (in Russian).

[3] D.N. Bernstein, A.G. Kouchnirenko, A.G. Khovanski, *Newton polyhedra*, Uspekhi Mat. Nauk XXXI **3** (1976), 201-202 (in Russian).

[4] W. Fulton, *Algebraic curves*, W.A.Benjamin, Inc., (1969).

[5] A.G. Khovanski, *Newton plyhedra and toroidal varieties* , Funk. Anal. Appl., **11**.4 (1977), 56–64.

[6] A.G. Khovanski, *Newton Polyhedra and the genus of complete intersections*, Functional Anal. Appl. **12** (1978), 38–46.

[7] A.G. Kouchnirenko, *Newton polyhedra and a number of roots of a system of k equations in k variables*, Uspekhi Mat. Nauk XXX **2** (1975), 266–267 (in Russian).

[8] A.G. Kouchnirenko, *Newton polyhedra and Bézout theorem*, Funk. Anal. Appl., **10**.3 (1976), 82–83 (in Russian).

[9] A.G. Kouchnirenko, *Polyèdres de Newton et nombres de Milnor*, Inventiones Mathematicae, **32** (1976), 1–31.

[10] A. Lenarcik, *On the Łojasiewicz exponent of the gradient of a polynomial function*, Ann. Pol. Math. **LXXI**.3 (1999), 211–239.

[11] M. Masternak, *Nearly irreducibility of polynomials and the Newton diagrams*, Ann. Univ. Paedagog. Crac. Stud. Math. **19** (2020), 65–77.

[12] A. Płoski, *A Note On The Łojasiewicz Exponent at infinity*, Bull. Soc. Letters Łódź **XLIV** (1994), 11–15.
[13] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, Encyclopedia of Mathematics and its Applications, Vol. 44, Cambridge University Press, Cambridge, 1993.

Jan Kochanowski University in Kielce, Poland
mateusz.masternak@ujk.edu.pl

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

APPLICATION OF STOCHASTIC EQUATIONS UNDER NONCLASSICAL APPROXIMATION SCHEMES TO CONSTRUCTION AND ANALYSIS OF INFORMATION STRUGGLE MODEL

ANATOLII NIKITIN

ABSTRACT. The information struggle model described by a stochastic differential equation with Markov switchings and an impulse perturbation process under Levy and Poisson approximation scheme, has constructed and analyzed. The new model is interpreted as the impact of rare events that quickly change a certain perception of a large number of people. As a result, the number of adherents of different ideas makes stochastic jumps, which can be seen using non-classical approximation schemes.

KEYWORDS: random evolution, Markov switchings, information struggle model, Levy and Poisson approximation conditions

MSC2010: 03G05, 57N01

Received 24 April 2021; revised 21 June 2021; accepted 16 July 2021

1. INTRODUCTION

Random evolution in the form of a stochastic differential equation is used to describe a wide class of natural processes in various branches of science. An important case is the study of the behavior of similar evolutionary systems in a random environment, which is well modeled using Markov processes. A large number of works by famous scientists, including A.V. Skorokhod, J.I. Gikhman, M.M. Bogolyubov, and others, are devoted to the study of such systems. A detailed bibliography on this issue can be found, for example, in the monographs of V.S. Korolyuk [2], in particular, in [7] the approaches used in this article to study the asymptotic behavior of an evolutionary system with diffusion perturbation were initiated. In the first two parts of this work, we consider the case when the perturbations of the system are determined by an impulse process in nonclassical approximation schemes (for details, see [10], [9], [11],[15], [14], [16], [12]). First of all, it is important the question of the asymptotic behavior for the limit generator. Similar problems have been considered before with the use of qualitatively different methods (for example, see [17]). The methods proposed in our work allow to investigate a model that contains Markov switches corresponding to a random environment, as well as to single out an additional diffusion component and large jumps of the perturbing process in the boundary equation, which are described by rare in applied problems.

In the third part of paper costructed an evolutionary model describes two complex systems' conflicting interactions with nontrivial internal structures. The model is determined by a stochastic differential equation with Markov switchings and an impulse perturbation process under Levy and Poisson approximation scheme. The Lotka-Volterra model of prey-predator interaction is one of the main models for simulation of many processes in applied science [1], [4], [8], [6], [5], [18], [3]. Application of this approach to information warfare model proposed in [13]. Authors regard some social community of quantity N_0, potentially exposed some information threat (InfT) of two types, that is, for example, the threat of a negative change in its state by transmitting some information relevant to this group by information two different channels. The values $N_1(t)$, $N_2(t)$ are the numbers of "adherents" depending on time who accepted the new information, ideas, norms, etc. of the type 1 and 2 respectively. These are the main current characteristics of the degree of prevalence of InfT. At present paper more essential model generalization has provided, the generator for the limit process has constructed in explicit form and some interpretation for a model has proposed.

2. Stochastic evolutionary equations under Levy approximation scheme

In this section we investigate the case where system's disturbances are defined by an impulse process is under nonclassical approximation scheme Levy and we pay special attention to the asymptotic behavior of the generator for this system.

Stochastic evolutionary system in an ergodic Markov medium is defined by the stochastic diffenential equation

$$du^\varepsilon(t) = C(u^\varepsilon, x(t/\varepsilon^2))dt + d\eta^\varepsilon(t),\ \ u^\varepsilon(t) \in \mathbb{R}. \tag{1}$$

where uniformly ergodic Markov process $x(t)$ in standart phase space $(X, \mathbf{X})$, is defined by the generator

$$\mathbf{Q}\varphi(x) = q(x) \int_X P(x, dy)[\varphi(y) - \varphi(x)]$$

on the Banach space $B(X)$ of real-valued bounded functions $\varphi(x)$ with the supremum norm $\|\varphi\| = \max_{x\in X} |\varphi(x)|$.

The stochastic kernel $P(x, B)$, $x \in X$, $B \in \mathbf{X}$ defines uniformly ergodic embedded Markov chain $x_n = x(\tau_n)$, $n \geq 0$, with stationary distribution $\rho(B)$, $B \in \mathbf{X}$. Stationary distribution $\pi(B)$, $B \in \mathbf{X}$ of the Markov process $x(t)$, $t \geq 0$ is defined by the relation

$$\pi(dx)q(x) = q\rho(dx), \ q = \int_X \pi(dx)q(x).$$

Let R_0 be denoted a potential operator of the generator $\mathbf{Q}$, which is defined by the equality [2]

$$R_0 = \Pi - (\Pi + \mathbf{Q})^{-1},$$

where $\Pi\varphi(x) = \int_X \pi(dy)\varphi(y)1(x)$ is the projector of zeroes of generator $\mathbf{Q}$ onto subspace $N_Q = \{\varphi : \ \mathbf{Q}\varphi = 0\}$.

The impulse process of perturbations (IPP) $\eta^\varepsilon(t)$, $t \geq 0$, under the Levy approximation scheme is defined by the relation

$$\eta^\varepsilon(t) = \int_0^t \eta^\varepsilon(ds, x(s/\varepsilon^2)), \tag{2}$$

where the set of processes with independent increments $\eta^\varepsilon(t, x)$, $t \geq 0$, $x \in X$, is defined by the generators

$$\Gamma^\epsilon(x)\varphi(w) = \varepsilon^{-2} \int_R (\varphi(w + v) - \varphi(w))\Gamma^\varepsilon(dv, x), \ x \in X \tag{3}$$

and satisfies the properties of Levy approximation [2]

L1. The approximation of averages

$$\int_R v\Gamma^\varepsilon(dv, x) = \varepsilon a_1(x) + \varepsilon^2(a_2(x) + \theta_a(x)), \ \theta_a(x) \to 0, \ \varepsilon \to 0,$$

and

$$\int_R v^2\Gamma^\varepsilon(dv, x) = \varepsilon(b(x) + \theta_b(x)), \ \theta_b(x) \to 0, \ \varepsilon \to 0,$$

L2. The condition imposed on the distribution function

$$\int_R g(v)\Gamma^\varepsilon(dv, x) = \varepsilon^2(\Gamma_g(x) + \theta_g(x)), \ \theta_g(x) \to 0, \ \varepsilon \to 0,$$

for all $g(v) \in C^2(\mathbb{R})$ (the space of real-valued bounded functions such that $g(v)/|v|^2 \to 0$, $|v| \to 0$), where measure $\Gamma_g(x)$ is bounded for all $g(v) \in C^2(\mathbb{R})$ and is defined by the relation (functions from the space $C^2(\mathbb{R})$ separate the measures):

$$\Gamma_g(x) = \int_R g(v)\Gamma_0(dv, x), \ g(v) \in C^3(\mathbb{R});$$

L3. The uniform quadratic integrability

$$\sup \lim_{c\to\infty} \int\limits_{|v|>c} v^2\Gamma_0(dv,x) = 0;$$

Assuming that the balance condition is fulfilled

$$\hat{a}_1 := \int\limits_X \pi(dx)a_1(x) = 0, \tag{4}$$

let the asymptotic properties of the perturbation process be considered.

Theorem 2.1. *If balance condition (4) and conditions L1 – L3 are satisfied, weak convergence*

$$\eta^\varepsilon(t) \to \eta^0(t),\ \varepsilon \to 0$$

takes place for impulse perturbation process (3).

The limit process $\eta^0(t)$ is defined by the generator

$$\Gamma\varphi(w) = \hat{a}_2\varphi'(w) + \frac{1}{2}\sigma^2\varphi''(w) + \int\limits_R [\varphi(w+v) - \varphi(w)]\hat{\Gamma}_0(dv),$$

where

$$\hat{a} = \int\limits_X \pi(dx)(a_2(x) - a_0(x)),$$

$$\sigma^2 = \int\limits_X \pi(dx)(b(x) - b_0(x)) + 2\int\limits_X \pi(dx)a_1(x)R_0a_1(x),$$

$$a_0(x) = \int\limits_R v\Gamma_0(dv,x),$$

$$b_0(x) = \int\limits_R v^2\Gamma_0(dv,x),$$

$$\hat{\Gamma}_0(v) = \int\limits_X \pi(dx)\Gamma_0(v,x),$$

and it is a Levy process with three components – derministic shift, diffusion component and Poisson jump component.

Proof. Firstly, let some additional propositions be justified.

Lemma 2.2. *Generators of processes with independed increments $\eta^\varepsilon(t,x)$, $t \ge 0$, $x \in X$, can be asymptotically representated by relation*

$$\mathbf{\Gamma}^\varepsilon(x)\varphi(w) = \varepsilon^{-1}\Gamma_1(x)\varphi(w) + \Gamma_2(x)\varphi(w), \tag{5}$$

on test functions $\varphi(w) \in C^2(\mathbf{R})$ under conditions L1 – L3, where

$$\Gamma_1(x)\varphi(w) = a_1(x)\varphi'(w),$$

$$\Gamma_2(x)\varphi(w) = (a_2(x) - a_0(x))\varphi'(x) + \frac{1}{2}(b(x) - b_0(x))\varphi''(x) +$$
$$+ \int_R [\varphi(w+v) - \varphi(v)]\Gamma_0(dv, x).$$

Proof. Using the expansion of function $\varphi(w)$ into the Taylor series, let generator (3) be transformed:

$$\mathbf{\Gamma}^\varepsilon(x)\varphi(w) = \varepsilon^{-2}\int_R (\varphi(w+v) - \varphi(v))\Gamma^\varepsilon(dv, x) =$$
$$= \varepsilon^{-2}\int_R (\varphi(w+v) - \varphi(v) - v\varphi'(v) - \frac{1}{2}v^2\varphi''(w))\Gamma^\varepsilon(dv, x) +$$
$$+\varepsilon^{-2}\int_R (v\varphi'(w)\Gamma^\varepsilon(dv, x) + \frac{1}{2}v^2\varepsilon^{-2}\int_R v^2\varphi''(w)\Gamma^\varepsilon(dv, x) =$$
$$= \int_R (\varphi(u+v) - \varphi(v) - v\varphi'(w) - \frac{1}{2}v^2\varphi''(w))\Gamma_0(dv, x) +$$
$$+\varepsilon^{-1}a_1(x)\varphi'(w) + a_2(x)\varphi'(w) +$$
$$+\frac{1}{2}(b(x) - b_0(x))\varphi''(w) +$$
$$+ \int_R (\varphi(u+v) - \varphi(v))\Gamma_0(dv, x) + \gamma^\varepsilon(w)\varphi(w),$$

where the penultimate equality follows from conditions L1 and L2 (note that function $\varphi(w+v) - \varphi(w) - v\varphi'(w) - \frac{1}{2}v^2\varphi''(w) \in C^2(\mathbf{R})$, since it is bounded due to the boundedness of $\varphi(w)$ and of its derivatives and

$$[\varphi(w+v) - \varphi(w) - v\varphi'(w) - \frac{1}{2}v^2\varphi''(w)]/|v^2| \to 0$$

when $v \to 0$.

Taking into account that $\gamma^\varepsilon(w)\varphi(w) = o(\varepsilon^2)$, $\varphi(w) \in C^2(\mathbf{R})$, relation (5) has obtained. □

Lemma 2.3. *Generator of two-component Markov process* $(\eta^\varepsilon, x(t/\varepsilon^2))$, $t \geq 0$ *has the form*

$$\hat{\Gamma}^\varepsilon(x)\varphi(w, x) = \varepsilon^{-2}\mathbf{Q}\varphi(w, x) + \varepsilon^{-1}\Gamma_1(x)\varphi(w, x)$$
$$+\Gamma_2(x)\varphi(w, x) + \gamma^\varepsilon(x)\varphi(w, x), \tag{6}$$

where operators $\Gamma_1(x)$ *and* $\Gamma_2(x)$ *are defined in Lemma 2.2 and the remainder term* $\|\gamma^\varepsilon(x)\varphi(w, x)\| \to 0$ *as* $\varepsilon \to 0$, $\varphi(w, \cdot) \in C^2(\mathbf{R})$.

Proof. Using the definition of generator of Markov process and the form of respective generators of processes $\eta^\varepsilon(t,x)$ and $x(t/\varepsilon^2)$. the statement of Lemma 2.3 becomes obvious. □

The truncated operator has the form [15]

$$\Gamma_0^\varepsilon(x)\varphi(w) = \varepsilon^{-2}\mathbf{Q}\varphi(w,x) + \varepsilon^{-1}\Gamma_1(x)\varphi(w,x) + \\ + \Gamma_2(x)\varphi(w,x). \tag{7}$$

Lemma 2.4. *Under balance condition (4), let the solution of the singular perturbation problem for the truncated operator (7) be founded by the relation*

$$\Gamma_0^\varepsilon(x)\varphi^\varepsilon(w,x) = \Gamma\varphi(w) + \varepsilon\theta_\eta^\varepsilon(x)\varphi(w), \tag{8}$$

on test functions

$$\varphi^\varepsilon(w,x) = \varphi(w) + \varepsilon\varphi_1(w,x) + \varepsilon^2\varphi_2(w,x)$$

where the remainder term $\theta_\eta^\varepsilon(x)\varphi(w)$ is uniformly bounded in x.

The limit operator is defined by the formula

$$\Gamma = \Pi\Gamma_1(X)R_0\Gamma_1(x)\Pi + \Pi\Gamma_2(x)\Pi. \tag{9}$$

Proof. For equality (8) to hold, the coefficients of identical degrees ε on the left and on the right should coincide. Let us calculate

$$\Gamma_0^\varepsilon(x)\varphi^\varepsilon(w,x) = \varepsilon^{-2}\mathbf{Q}\varphi(w) + \\ + \varepsilon^{-1}[\mathbf{Q}\varphi_1(w,x) + \Gamma_1(x)\varphi(w)] + \\ + [\mathbf{Q}\varphi_2(w,x) + \Gamma_1(x)\varphi_1(w,x) + \Gamma_2(x)\varphi_2(w,x)] + \\ + \varepsilon[\Gamma_1(x)\varphi_2(w,x) + \Gamma_2(x)\varphi_1(w,x)] + \\ + \varepsilon^2\Gamma_2(x)\varphi_2(w,x).$$

From the first summand we obtain

$$\mathbf{Q}\varphi(w) = 0 \Leftrightarrow \varphi(w) \in N_{\mathbf{Q}}.$$

From here we can see that $\varphi(w)$ does not depend on x.

Tha balance condition (4) is a condition of solvability for the equation

$$\mathbf{Q}\varphi_1(w,x) + \Gamma_1(x)\varphi(w) = 0.$$

Then

$$\varphi_1(w,x) = R_0\Gamma_1(x)\varphi(w). \tag{10}$$

The equation

$$\mathbf{Q}\varphi_2(w,x) + \Gamma_1(x)\varphi_1(w,x) + \Gamma_2(x)\varphi(w) = \Gamma\varphi(w)$$

acorrding to (10), can be rewrited in a form

$$\mathbf{Q}\varphi_2(w,x) + \Gamma_1(x)R_0\Gamma_1(x)\varphi(w,x) + \Gamma_2(x)\varphi(w) = \Gamma\varphi(w).$$

The resolvability condition for the last equation gives the limit operator in the form (9). Then

$$\varphi_2(w,x) = R_0[\Gamma_1(x)R_0\Gamma_1(x) + \Gamma_2(x) - \Gamma]\varphi(w). \tag{11}$$

Using (10) and (11), we can reduce the other terms of the expansion to the form

$$\begin{gathered}\varepsilon[\Gamma_1(x)\varphi_2(w,x) + \Gamma_2(x)\varphi_1(w,x)] + \varepsilon^2\Gamma_2(x)\varphi(w,x) = \\ = \varepsilon[[\Gamma_1(x)R_0[\Gamma_1(x)R_0\Gamma_1(x) + \Gamma_2(x) - \Gamma] + \Gamma_2(x)R_0\Gamma_1(x)] + \\ +\varepsilon\Gamma_2(x)R_0[\Gamma_1(x)R_0\Gamma_1(x) + \Gamma_2(x) - \\ -\Gamma]]\varphi(w) = \varepsilon\theta_\eta^\varepsilon(x)\varphi(w).\end{gathered}$$

The boundedness of $\theta_\eta^\varepsilon(x)\varphi(w)$ follows from the form of operators Γ_1, Γ_2 and R_0.

The proof of Lemma 2.4 has completed.

The proof of the Theorem 2.1 can be completed similarly to the proof of Theorem 6.4 from [2].

Further let the asymptotic properties of the original evolutionary system (1) be investigated.

Theorem 2.5. *If balance condition (4) and conditions L1 – L3 are satisfied, the weak convergence*

$$u^\varepsilon(t) \to \hat{u}(t),\ \varepsilon \to 0.$$

holds thrue.

The limit process $\hat{u}(t)$ is defined by the generator

$$\mathbf{L}\varphi(w) = \hat{C}(u)\varphi'(w) + \Gamma\varphi(w), \tag{12}$$

where $\hat{C}(u) = \int\limits_X \pi(dx)C(u,x)$.

Remark 2.6. The weak convergence of processes $u^\varepsilon(t) \to \hat{u}(t)$, $\varepsilon \to 0$, follows from the convergence of respective generators when compactness of prelimiting set of processes $u^\varepsilon(t)$ holds true.

Remark 2.7. The limit process $\hat{u}(t)$ can be given by stochastic differential equation

$$d\hat{u}_d(t) = [\hat{C}(\hat{u}(t)) + \hat{a}_2]dt + \sigma dW(t) + \int_{\mathbb{R}} v\tilde{v}(dt,dv),$$

where $\mathbf{E}\tilde{\nu}(dt,dv) = dt\tilde{\Gamma}_0(dv)$.

Remark 2.8. The limit process $\hat{u}(t)$ has three components. The deterministic shift is defined by the solution of the differential equation

$$d\hat{u}_d(t) = [\hat{C}(\hat{u}_d(t)) + \hat{a}_2]dt, \tag{13}$$

where the additional term $\hat{a}_2$ appears due to accumulation with the normalized time t/ε^2, $\varepsilon \to 0$ of small jumps of the impulse process that happen with probability, close to one. The second, diffusion component,is defined by a parameter σ and it arises due to accumulation with growth of the normalized time t/ε^2, $\varepsilon \to 0$ of small jumps of degree ε, that happen with probability, close to one too.

The third component is rare big jumps that take place with nearly zero probability and are defined in terms of averaged measure of jumps $\tilde{\Gamma}_0(dv)$ by the generator

$$\Gamma_j \varphi(w) = \int_R [\varphi(w+v) - \varphi(w)]\tilde{\Gamma}_0(dv).$$

Lemma 2.9. *Generator of two-component Markov process $(u^\varepsilon(t), x(t/\varepsilon^2))$, $t \geq 0$, has the representation*

$$\mathbf{L}^\varepsilon(x)\varphi(w,x) = \varepsilon^{-2}\mathbf{Q}\varphi(w,x) + \Gamma^\varepsilon(x)\varphi(w,x) + \\ + \mathbf{C}(x)\varphi(w,x) + \theta_w^\varepsilon \varphi(w,x), \tag{14}$$

where $\Gamma^\varepsilon(x)$ is generator of the family of processes with independent increments,

$$\mathbf{C}(x)\varphi(w,x) = C(u,x)\varphi'_w(w,x).$$

The remainder term $\|\theta_w^\varepsilon \varphi(w,x)\| \to 0$ as $\varepsilon \to 0$.

The proof of lemma can be provided using scheme in [7].

Lemma 2.10. *Generator $\mathbf{L}^\varepsilon(x)$ has an asymptotic representation*

$$\mathbf{L}^\varepsilon(x)\varphi(w,x) = \varepsilon^{-2}\mathbf{Q}\varphi(w,x) + \varepsilon^{-1}\Gamma_1(x)\varphi(w,x) + \\ + \Gamma_2(x)\varphi(w,x) + \mathbf{C}(x)\varphi(w,x) + \hat{\theta}_w^\varepsilon \varphi(w,x), \tag{15}$$

where

$$\hat{\theta}_w^\varepsilon(x) = \gamma^\varepsilon + \theta_w^\varepsilon(x),$$

$\Gamma_1(x)$ and $\Gamma_2(x)$ are determined in Lemma 1.

The remainder term

$$\|\hat{\theta}_w^\varepsilon \varphi(w,x)\| \to 0$$

when $\varepsilon \to 0$.

The proof uses the representation of representation of operator (15) and results of Lemma 2.9.

The truncated operator has a form

$$\mathbf{L}_0^\varepsilon(x)\varphi = \varepsilon^2\mathbf{Q}\varphi + \varepsilon^{-1}\Gamma_1(x)\varphi + \Gamma_2(x)\varphi + \mathbf{C}(x)\varphi. \tag{16}$$

Lemma 2.11. *Under the balance condition (4), the solution of a singular perturbation problem for the truncated operator (6) can be defined by the relation*

$$\mathbf{L}_0^\varepsilon(x)\varphi^\varepsilon(w,x) = \mathbf{L}\varphi(w) + \varepsilon^2\theta_w^\varepsilon(x)\varphi(w), \tag{17}$$

on test functions

$$\varphi^{\varepsilon}(w,x) = \varphi(w) + \varepsilon\varphi_1(w,x) + \varepsilon^2\varphi(w,x)$$

where the remainder term $\theta_w^{\varepsilon}(x)$ is uniformly bounded with respect to x

The limit operator $\mathbf{L}$ *is defined by the formula*

$$\mathbf{L} = \Pi[\mathbf{C}(x) + \Gamma_1(x)R_0\Gamma_1(x) + \Gamma_2(x)]\Pi. \tag{18}$$

Proof. For equality (17) to hold, the coefficients of identical degrees ε on the left and on the right should coincide. To this end, we calculate

$$\mathbf{L}_0^{\varepsilon}(x)\varphi^{\varepsilon}(w,x) = \varepsilon^{-2}\mathbf{Q}(x)\varphi(w)+$$

$$+\varepsilon^{-1}[\mathbf{Q}\varphi_1(w,x) + \Gamma_1(x)\varphi(w)]+$$

$$+[\mathbf{Q}\varphi_2(w,x) + \Gamma_1(x)\varphi_1(w,x)]+$$

$$+\varepsilon[\Gamma_1(x)\varphi_2(w,x) + \Gamma_2(x)\varphi_1(w,x)) + \mathbf{C}(x)\varphi_1(w,x)]+$$

$$+\varepsilon^2[\Gamma_2(x)\varphi_2(w,x) + \mathbf{C}(x)\varphi_2(w,x)].$$

Since

$$\mathbf{Q}\varphi(w) \Leftrightarrow \varphi(w) \in N_Q,$$

it is obvious that $\varphi(w)$ does not depend on x.

The condition of resolvability to equation has a form

$$\mathbf{Q}\varphi_1(w,x) + \Gamma_1(x)\varphi(w) = 0.$$

Then

$$\varphi_1(w,x) = R_0\Gamma_1(x)\varphi(w).$$

We have the last equation

$$\mathbf{Q}\varphi_2(w,x) + \Gamma_1(x)\varphi(w,x)+$$

$$+\Gamma_2(x)\varphi(w) + \mathbf{C}(x)\varphi(w) = \mathbf{L}\varphi(w).$$

It can be rewrited in form as follows

$$\mathbf{Q}\varphi_2(w,x) = [\mathbf{L} - \Gamma_1(x)R_0\Gamma_1(x) - \Gamma_2(x) - \mathbf{C}(x)]\varphi(w).$$

The condition of resolvability for the last equation gives the form (18) of limit operator $\mathbf{L}$. $\square$

The proof of Theorem 2.5 can be completed similarly to the proof of Theorem 6.4 from [2]

3. Stochastic evolutionary equations under Poisson approximation schemes

In this section let the case where systems perturbations are defined by an impulse process in another nonclassical approximation scheme [2] named by Poisson be considered. Poisson approximation scheme is a generalization of the classical averaging scheme, which in turn is based on the idea of the law of large numbers.

Stochastic evolutionary system in ergodic Markov environment is defined by the stochastic differential equation

$$du^{\varepsilon}(t) = C(u^{\varepsilon}, x(t/\varepsilon))dt + d\eta^{\varepsilon}(t),\ u^{\varepsilon}(t) \in \mathbf{R}, \tag{19}$$

where and the conditions imposed on the Markov process are the same as in the previous section.

The impulse perturbation process $\eta^{\varepsilon}(t)$, $t \geq 0$, under the Poisson approximation scheme is defined by the relation

$$\eta^{\varepsilon}(t) = \int_0^t \eta^{\varepsilon}(ds, x(s/\varepsilon)), \tag{20}$$

where the set of processes with independent increments $\eta^{\varepsilon}(t, x)$, $t \geq 0$, $x \in X$, is defined by the generators

$$\Gamma^{\epsilon}(x)\varphi(w) = \varepsilon^{-1} \int_R (\varphi(w + v) - \varphi(w))\Gamma^{\varepsilon}(dv, x),\ x \in X$$

and satisfies the properties of Poisson approximation [14]

P1 is approximation in averages

$$\int_R v\Gamma^{\varepsilon}(dv, x) = \varepsilon(a(x) + \theta_a(x)),\ \theta_a(x) \to 0,\ \varepsilon \to 0,$$

$$\int_R v^2\Gamma^{\varepsilon}(dv, x) = \varepsilon(b(x) + \theta_b(x)),\ \theta_b(x) \to 0,\ \varepsilon \to 0,$$

P2 is the condition imposed on the distributing function

$$\int_R g(v)\Gamma^{\varepsilon}(dv, x) = \varepsilon(\Gamma_g(x) + \theta_g(x)),\ \theta_g(x) \to 0,\ \varepsilon \to 0,$$

for all $g(v) \in C^2(\mathbb{R})$ (the space of real-valued bounded functions such that $g(v)/|v|^2 \to 0$, $|v| \to 0$), where measure $\Gamma_g(x)$ is bounded for all $g(v) \in C^2(\mathbb{R})$ (functions from the space $C^2(\mathbb{R})$ separate the measures):

$$\Gamma_g(x) = \int_R g(v)\Gamma_0(dv, x),\ g(v) \in C^2(\mathbb{R});$$

P3 is uniform quadratic integrability of

$$\sup \lim_{c\to\infty} \int\limits_{|v|>c} v^2\Gamma_0(dv,x) = 0;$$

P4 is the abcence of diffusion component

$$b(x) = \int\limits_{R} v^2\Gamma_0(dv,x).$$

Let us introduce the notation

$$\Gamma_1(x) = a(x)\varphi'(w) + \int\limits_{R} [\varphi(w+v) - \varphi(v) - v\varphi'(w)]\Gamma_0(dv,x).$$

Let's give a simple example of random variable satisfying properties under Levy and Poisson schemas We cosider α:

$$P\{\alpha = b\} = \varepsilon^2 p,$$
$$P\{\alpha = \varepsilon a_1 + \varepsilon^2 b_1\} = 1 - \varepsilon^2 p.$$

Then

$$\mathbf{E}\alpha = \varepsilon a_1 + \varepsilon^2(bp + b_1) + o(\varepsilon^2),$$
$$\mathbf{E}\alpha^2 = \varepsilon^2(b^2 p + a_1^2) + o(\varepsilon^2).$$

These moment conditions characterize Levy approximation

When $a_1 = 0$, we obtain

$$\mathbf{E}\alpha = \varepsilon^2(bp + b_1) + o(\varepsilon^2),$$
$$\mathbf{E}\alpha^2 = \varepsilon^2 b^2 p + o(\varepsilon^2),$$

then assuming $\widetilde{\varepsilon} = \varepsilon^2$ we obtain

$$\mathbf{E}\alpha = \widetilde{\varepsilon}(bp + b_1) + o(\widetilde{\varepsilon}),$$
$$\mathbf{E}\alpha^2 = \widetilde{\varepsilon} b^2 p + o(\widetilde{\varepsilon}).$$

These moment conditions characterize Poisson approximation.

Following the research scheme given in the previous section, we obtain statements that reveal the form of the limit generator for the impulse process and the evolutionary system itself

Theorem 3.1. *If conditions P1 – P4 are satisfied, weak convergence*

$$\eta^\varepsilon(t) \to \eta^0(t),\ \varepsilon \to 0$$

take place for impulse perturbation process (20). The limit process $\eta^0(t)$ is defined by the generator

$$\Gamma\varphi(w) = \Pi\Gamma_1(x)\varphi(w) =$$
$$= \tilde{a}\varphi'(w) + \int\limits_{R} [\varphi(w+v) - v\varphi'(w)]\tilde{\Gamma}_0(dv),$$

where $\tilde{a} = \int\limits_X \pi(dx)a(x)$, $\tilde{\Gamma}_0(v) = \int\limits_X \pi(dx)\Gamma_0(v,x)$, *and it is a process with independend increments, which has both Poisson component and deterministic shift.*

Theorem 3.2. *If conditions P1 – P4 are satisfied, the weak convergence*

$$u^\varepsilon(t) \to \hat{u}(t),\ \varepsilon \to 0$$

takes place. The limit process $\hat{u}(t)$ *is defined by the generator*

$$\mathbf{L}\varphi(w) = \hat{C}(u)\varphi'(w) + \Gamma\varphi(w),$$

where $\hat{C}(u) = \int\limits_X \pi(dx)C(u,x)$.

Remark 3.3. The limit process $\hat{u}(t)$ has two components. The deterministic drift is defined be the solution of the differential equation

$$d\hat{u}_d(t) = [\hat{C}(\hat{u}_d(t)) + \tilde{a}]dt, \tag{21}$$

where the additional term $\tilde{a}$ appears due to accumulation (with the normalized time t/ε, $t \to 0$) of small jumps of the impulse process that happen with probability close to one. The second component is rare big jumps that take place with nearly zero probability and are defined in terms of averaged measure of jumps $\tilde{\Gamma}_0(dv)$ by the generator

$$\Gamma_j\varphi(w) = \int\limits_R [\varphi(w+v) - \varphi(v) - v\varphi'(w)]\tilde{\Gamma}_0(dv).$$

Remark 3.4. The limit process $\hat{u}(t)$ is purely deterministic and defined by equation (21) in case of zero average measure of jumps $\tilde{\Gamma}_0(dv)$. For example, if all the moments of order three and above for the set of processes with independent increments $\eta^\varepsilon(t,x)$, $t \geq 0$, $x \in X$, are equal to zero, or if the balance condition is satisfied

$$\tilde{\Gamma}_0(v) = \Pi\Gamma_0(v,x) = \int\limits_X \pi(dx)\Gamma_0(v,x) = 0.$$

4. Construction features and analysis of information struggle model under nonclassical approximation schemes

We construct and investigate the information struggle model in a form

$$dN^\varepsilon(t) = C(N^\varepsilon, x(t/\varepsilon^2))dt + d\eta^\varepsilon(t),\ N^\varepsilon(t) \in \mathbb{R}. \tag{22}$$

where

$$C(N^\varepsilon, x(t/\varepsilon^2)) =$$

$$= \begin{pmatrix} -\alpha_1(x) + \beta_1(x)N_0 - \beta_1(x)N_1^\varepsilon(t) & -\alpha_1(x) - \beta_1(x)N_1^\varepsilon(t) \\ -\alpha_2(x) - \beta_2(x)N_2^\varepsilon(t) & -\alpha_2(x) + \beta_2(x)N_0 - \beta_2(x)N_2^\varepsilon(t) \end{pmatrix} \times$$

$$\times \begin{pmatrix} N_1^\varepsilon(t) \\ N_2^\varepsilon(t) \end{pmatrix} + \begin{pmatrix} \alpha_1 N_0 \\ \alpha_2 N_0 \end{pmatrix}$$

where some social community with a constant number of people N_0 is considered, which is potentially exposed to information threat of two types, for example, the threat of negative change of views of community members, transmitting some information, which is of two types. It should be understood that each of the two types of information can have both positive and negative colors, the most interesting is the case of two antagonistic thoughts, the spread of which leads to the polarization of society and the question of the winner in the information war. The values of $N_1(t)$ and $N_2(t)$ are the number of "supporters" depending on the time who perceived the new information, ideas, norms, etc. of type 1 and 2, respectively. Basic assumptions about the classical model [13]: 1. Both ideas are spread among the community through two information channels: - the first is "external" to the community, for example, an advertising media campaign. Its intensity is characterized by the parameters $\alpha_1 > 0$ and $\alpha_2 > 0$ respectively, both are considered to be independent of time;

- the second, "internal" channel - interpersonal communication between members of the social community. (its intensity, that is, the number of equivalent informational contacts, characterized by the parameters $\beta_1 > 0$ and $\beta_2 > 0$ respectively, that are also independent of time). As a result, supporters of the first idea, who have already been "recruited" (their number is equal to $N_1(t)$), make their personal contribution to the process of spreading the idea among the community (their number is equal to $N_0 - N_1(t) - N_2(t)$), influencing its "non-recruited" members. The same applies to the supporters of the second idea. 2. The rate of change of the number of supporters $N_1(t)$ and $N_2(t)$ (ie the number of supporters of the idea, "recruited" per unit time) consists of: - external speed of recruitment; - the internal speed of recruitment.

Thus, the part of the community that has not yet been recruited by the time t (its hypothetical "middle" representative, initially neutral with respect to both N_1, and to N_2) receives information the faster, than larger the values α_1, α_2, β_1, β_2. Moreover, even if the impact N_1 is obviously stronger than the impact N_2, some members of the community still accept N_2 (that is, there is no complete monopoly of one type of information over another). Thus, the model is described by equations of the Lotka-Volterra type (for details on possible solutions and characteristics of a dynamical system, see [13]):

$$\frac{dN_1(t)}{dt} = (\alpha_1 + \beta_1 N_1(t))(N_0 - N_1(t) - N_2(t)),$$

$$\frac{dN_2(t)}{dt} = (\alpha_2 + \beta_2 N_2(t))(N_0 - N_1(t) - N_2(t)),\ t > 0.$$

The main disadvantage of the classical model is, firstly, the constancy of the characteristics (intensities) of information impact, and secondly - the lack of ability to take into account sudden unforeseen events that affect the consciousness of consumers of information. Obviously, in today's world where information is spread instantly and at the same time covers a wide audience, such rare but very influential factors must be taken into account. We offer a model in the form of a dynamic system that takes into account both the random influence of the environment on the intensity of information dissemination and rare random jumps that in the short term, instead significantly change the number of "followers" of relevant ideas.

The main result is that the influence of big jumps is preserved in the limit process. These big jumps, in particular, simulate resonant events that instantly and significantly affect people's thoughts. They occur very rarely, but change opinions significantly. This is not taken into account in any other known model. In our problem statement, the averaged limit model of information struggle has the form

$$\mathbf{L}\varphi(w) = \hat{C}(u)\varphi'(w) + \Gamma\varphi(w),$$

where $\hat{C}(u) = \int\limits_X \pi(dx)C(u,x)$.

In the case of the Ornstein-Uhlenbeck switching process, the above formulas can be explicitly calculated by stationary integration, taking into account the type of potential operator.

5. Conclusion

A new form of information struggle model with an additional influence of chance has proposed. We suppose that such a model could be more essential, as soon as now breaking news produce quick and significant influence on the audience through TV and Internet. The behavior of our model could not be analyzed obviously for any fixed moment of time as it was done in a classical case. But, as it is usual for stochastic models, we may obtain functional limit theorems that present the behavior on large time intervals. Thus, we have averaged limit characteristics of the process and may use them to construct obvious solutions. That is, all the functions that depend on Markov process should be averaged by the stationary measure of the switching Markov process.

References

[1] A. J. Lotka, *Relation between birth rates and death rates*, Science **26** (1907), 21–22.

[2] V.S. Korolyuk, N. Limnios, *Stochastic Systems in Merging Phase Space* , World Scientific, 2005.

[3] J. Tufto, *Effects of releasing maladapted individuals: a demographic evolutionary model*, The American Naturalist **158**.4 (2001), 331–340.

[4] V. Volterra, *Sui tentativi di applicazione della matematiche alle scienze biologiche e sociali*, Giornale degli Economisti **23** (1901), 436–458.

[5] L. Stone, R. Olinky, *Phenomena in ecological systems*, Experimental Chaos: 6th Experimental Chaos Conference (2003), 476–487.

[6] P.P. Verhulst, *Notice sur la loi que la population suit dans son accroissement*, Correspondence mathematique et physique publiee par A. Quetelet **10** (1838), 113–121.

[7] V.S. Korolyuk, N. Limnios, I.V. Samoilenko, *Levy and Poisson approximations of switched stochastic systems by a semimartingale approach*, Comptes Rendus Mathematique **354** (2016), 723–728.

[8] K. I. Takahashi, K.Md.M. Salam, *Mathematical model of conflict with non-annihilating multi-opponent*, J. Interdisciplinary Math. **9**.3 (2006), 459–473.

[9] A.V. Nikitin, *Asymptotic Properties of a Stochastic Diffusion Transfer Process with an Equilibrium Point of a Quality Criterion*, Cybernetics and Systems Analysis, **51**.4 (2015), 650–656.

[10] I.V. Samoilenko, A.V. Nikitin, *Differential equations with small stochastic terms under the Levy approximation conditions*, Ukrainian Mathematical Journal **69**.9 (2018), 1445–1454.

[11] A.V. Nikitin, U.T. Khimka, *Asymptotics of Normalized Control with Markov Switchings*, Ukrainian Mathematical Journal **68**.8 (2017), 1252–1262.

[12] A.V. Nikitin, *Asymptotic dissipativity of stochastic processes with impulsive perturbation in the Levy approximation scheme*, Journal of Automation and Information Sciences **68**.8 (2018), 1252–1262.

[13] A.P. Mikhailov, N.A. Marevtseva, *Models of information warfare*, Math. Models Comput. Simul. **4**.3 (2012), 251–259.

[14] I.V. Samoilenko, Y,M. Chabanyuk, A.V. Nikitin, *Asymptotic Dissipativity of Random Processes with Impulse Perturbation in the Poisson Approximation Scheme*, Cybernetics and Systems Analysis **54**.2 (2018), 205–211.

[15] Y.M. Chabanyuk, A.V. Nikitin, U.T. Khimka, *Asymptotic properties of the impulse perturbation process under Levy approximation conditions with the point of equilibrium of the quality criterion*, Mathematychni Studii **52**.1 (2019), 96–104.

[16] I.V. Samoilenko, A.V. Nikitin, *Double Merging of the Phase Space for Stochastic Differential Equations with Small Additions in Poisson Approximation Conditions*, Cybernetics and Systems Analysis **55**.2 (2019), 265–273.

[17] G. Papanicolaou, D. Stroock, S.R.S. Varadhan, *Martingale approach to some limit theorems*, Duke turbulence conference, Durham, NC, April 23-25, 1976, Duke University Mathematics Series III, New York: Duke University, 1977.

[18] V.V. Anisimov, *Switching processes in queuing models*, London:Wiley-ISTE, 2008.

Jan Kochanowski University in Kielce, Poland
anatolii.nikitin@ujk.edu.pl

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

SOLVING MATH PROBLEMS USING THE AREA METHOD

ANNA PETIURENKO

Abstract. In this paper, we continue the research ushered in [2]. We apply the area method introduced in [4] to reconstruct Euclid's theory of similar figures. In this theory, the area of a triangle is a primitive concept, and Euclid's proposition VI.1 an axiom. We show that the fundamental theorem of the area method, the so-called Co-side Theorem follows from Euclid's VI.1. Our proof builds on four geometric patterns that simplify solving problems in school geometry concerning similar figures.

Keywords: area method, Co-side Theorem, Euclid, proportion, school geometry

MSC2010: 97G40, 03A05

Received 05 May 2021; revised 28 July 2021; accepted 30 July 2021

1. Introduction

The theory of similar figures, presented in school mathematics, on the one hand, mimics results of Book VI of Euclid's *Elements*, on the other, is developed in a metric space with segments having lengths, fractions emulating proportions, and the similarity scale being a real number. However, real numbers and fractions are alien to ancient mathematics. As a result, school mathematics rephrases Euclid's propositions and does not apply his proof technique, i.e., proportions.

Line segments, triangles, rectangles, and generally, figures are terms in ancient proportions, while in school mathematics, these are measures of lines and figures, i.e., lengths and areas. Furthermore, in school mathematics, proportions are fractions processed in the arithmetic of real numbers. Usually, secondary school curricula do not explain how to introduce real numbers into Euclidean geometry. Indeed, it is an intricate endeavor of the foundations of geometry. Few have taken

https://doi.org/10.2478/9788366675360-017.

up that challenge, to mention [5], [3], [9], and regarding school geometry – Zofia Krygowska's textbooks.

From an educational perspective, Euclid's original system has clear merits. Let us consider, for example, the Thales theorem, that is proposition 2 of Book VI. In the *Elements*, it is a two-lines straightforward proof based on VI.1. In metric geometry, it is a long and tiresome derivation.[1] In the Hilbert system, it depends on the arithmetic of line segments, therefore requires some extra proofs and definitions.[2] That is why we propose introducing an axiomatic development of Euclidean theory of similar figures with no reference to real numbers or arithmetic of line segments. Our proposal builds on the axiomatic development of the area method. It provides foundations for rigorous development of the basics of similar figures. Moreover, this system provides a new technique for solving problems in school geometry.

2. Axioms for the area method

The area method, pioneered in [4], is a technique of proving theorems and constructing solutions in Euclidean geometry. [6] provides its axiomatic description. In this section, we introduce that system. In [2], we present a model for these axioms.

From the perspective of formal systems, the language of the area method includes one kind of variables, and symbols of a binary, $\overline{}$, and a ternary function, S. We also need the language of an ordered field, that is symbols of binary functions, $+$, $\cdot$ (sum and product), and unary functions $-, ^{-1}$ (an opposite and inverse element), as well as constants $0, 1$, and finitely many constants and r. Less formally, there are three primitive notions in the area method: point, length of a directed segment, and a signed area of a triangle. An ordered pair of points is called a directed segment, an ordered triple – a triangle. In what follows, capital letters A, B, C, etc., stand for points. The length of a directed segment, $\overline{AB}$, in short, is an element of an ordered field $(\mathbb{F}, +, \cdot, 0, 1, <)$.[3] Similarly, the signed area of a triangle, S_{ABC}, in short, is an element of the ordered field. $\overline{AB}$ and S_{ABC} can be positive, negative, or zero and they are processed in the arithmetic of an ordered field.

To model Euclidean geometry, we need some definitions that we apply in axioms.[4]

Definition 2.1. Points A, B, C are co-linear iff $S_{ABC} = 0$.

[1] [9], chapter 9, or [7], ch. III.

[2] [5], ch. 54, § 20.

[3] Adopting a formal perspective, it can be a commutative field of characteristics other than 2. Yet standards models of elementary geometry are Cartesian planes over an ordered field.

[4] [5] relates these definitions with Euclid's original system

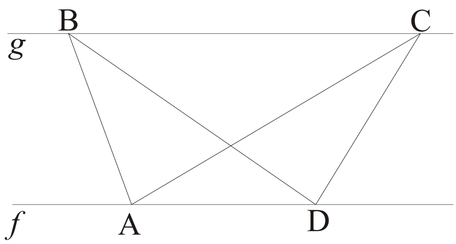


FIGURE 1. Definition of parallel line segments

Definition 2.2. Two segments AD and BC, where $A \neq D$ and $B \neq C$, are parallel, iff $S_{ABC} = S_{DBC}$. For this relation, we adopt the standard symbol $AD \parallel BC$.

Definition 2.3. For three points A, B and C, the Pythagorean difference, denoted by P_{ABC}, is defined by

$$P_{ABC} = \overline{AB}^2 + \overline{BC}^2 - \overline{AC}^2.$$

Definition 2.4. Two segments DB and CA, where $D \neq B$ and $C \neq A$, are perpendicular iff $P_{DCA} = P_{BCA}$. This relation is denoted by $DB \perp CA$.

Here is the first proposition of the *Elements*, Book VI.

Theorem 2.5 (*Elements*, VI.1)**.** *In triangles, ABC and ACD, having the same height AC, as base BC is to base CD, so triangle ABC is to triangle ACD.*

In school geometry, where the product of base and height determines the area of a triangle, that theorem follows from algebraic identity $\frac{a_1 h}{a_2 h} = \frac{a_1}{a_2}$. However, Euclid's proof proceeds within proportion theory; it involves the non-defined addition of triangles and requires comparing triangles as greater–lesser [1]. Both concepts seem foreign to a modern reader. In this paper, we adopt VI.1 as the axiom A10. Thus, on the one hand, we do not need to refer to Euclid's definition of proportion. On the other, do not need to apply any formula for an area of a triangle. Specifically, our account does not involve products of line segments or their measures (i.e., real numbers). Moreover, although objects $\overline{AB}$ and S_{ABC} are elements of an ordered field, we process only terms such as

$$\frac{\overline{AB}}{\overline{CD}} = \frac{\overline{EF}}{\overline{GH}}, \quad \frac{\overline{AB}}{\overline{CD}} = \frac{S_{EFG}}{S_{GIJ}}.$$

To this end, we apply Euclid's propositions regarding proportions, namely V.12, V.17–19, that are also laws of fractions in an ordered field. In an algebraic stylization, these propositions are as follows:[5]

[5]For a discussion of reconstruction of Euclid's Book V in an ordered field, see [2].

V.12 $\dfrac{a}{b} = \dfrac{c}{d},\ \dfrac{a}{b} = \dfrac{e}{f} \Rightarrow \dfrac{a}{b} = \dfrac{(a+c+e)}{(b+d+f)}$.
V.17 $\dfrac{(a+b)}{b} = \dfrac{(c+d)}{d} \Rightarrow \dfrac{a}{b} = \dfrac{c}{d}$.
V.18 $\dfrac{a}{b} = \dfrac{c}{d} \Rightarrow \dfrac{(a+b)}{b} = \dfrac{(c+d)}{d}$.
V.19 $\dfrac{(a+b)}{(c+d)} = \dfrac{a}{c} \Rightarrow \dfrac{b}{d} = \dfrac{(a+b)}{(c+d)}$.

In Euclid's theory, $a, b, c, \ldots$ stand for the so-called magnitudes, that is line segments, triangles, figures and solids, and angles. In the area method, we apply propositions V.12, 17–19 to directed line segments and signed areas.

Combining V.12 and V.19 we get

$$\frac{a}{b} = \frac{c}{d} \Rightarrow \frac{a}{b} = \frac{a+c}{b+d}.$$

We will refer to Euclid's propositions through that version. Furthermore, we will apply V.19 in the following form

$$\frac{(a+b)}{(c+d)} = \frac{a}{c} \Rightarrow \frac{a}{c} = \frac{b}{d}.$$

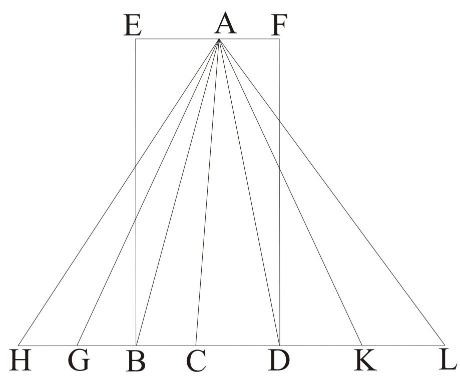


FIGURE 2. *Elements*, VI.1

Here are the axioms for the area method.

A1. $\overline{AB} = 0$ iff A and B are identical.

A2. $S_{ABC} = S_{CAB}$.

A3. $S_{ABC} = -S_{BAC}$.

A4. If $S_{ABC} = 0$, then $\overline{AB} + \overline{BC} = \overline{AC}$ (Chasles' axiom).

A5. There are points A, B and C such that $S_{ABC} \neq 0$ (there are non-collinear points).

A6. $S_{ABC} = S_{DBC} + S_{ADC} + S_{ABD}$ (all points are in the same plane).

A7. For each element r of F, there exists a point P, such that $S_{ABP} = 0$ and $\overline{AP} = r\overline{AB}$ (construction of a point on a line).

A8. If $A \neq B$, $S_{ABP} = 0, \overline{AP} = r\overline{AB}$, $S_{ABP'} = 0$ and $\overline{AP'} = r\overline{AB}$, then $P = P'$.

A9. If $PQ \parallel CD$ and $\frac{\overline{PQ}}{\overline{CD}} = 1$, then $DQ \parallel PC$ (parallelogram).

A10. If $S_{PAC} \neq 0$ and $S_{ABC} = 0$, then $\frac{\overline{AB}}{\overline{AC}} = \frac{S_{PAB}}{S_{PAC}}$ (Euclid's proposition VI.1).

A11. If $C \neq D$ and $AB \perp CD$ and $EF \perp CD$, then $AB \parallel EF$.

A12. If $A \neq B$, $AB \perp CD$ and $AB \parallel EF$, then $EF \perp CD$.

3. Co-side Theorem

The so-called Co-side Theorem is the fundamental tool of the area method.

Theorem 3.1. *For four distinct points A, B, P, Q, let M be the intersection of the lines AB and PQ such that $Q \neq M$. Then the following equality obtains*

$$\frac{S_{PAB}}{S_{QAB}} = \frac{\overline{PM}}{\overline{QM}}.$$

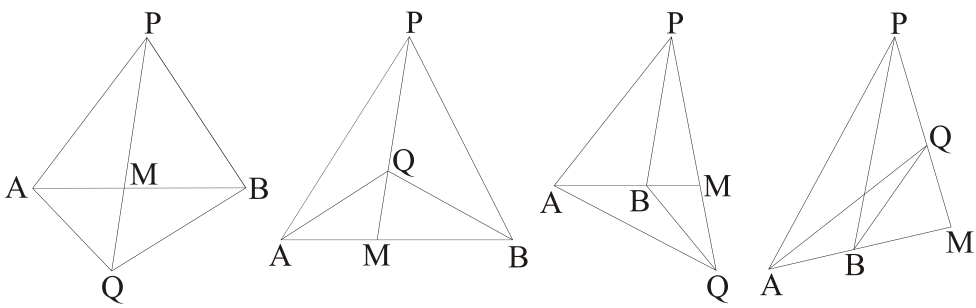


FIGURE 3. Co-side Theorem

Below we present two proofs. The first comes from ([4], pp. 8–17). It builds on an arithmetic trick to represent fraction $\frac{S_{PAB}}{S_{QAB}}$ as the product of three other fractions, namely

$$\frac{S_{PAB}}{S_{PAM}}, \quad \frac{S_{PAM}}{S_{QAM}}, \quad \frac{S_{QAM}}{S_{QAB}}.$$

Then, due to axiom A10, these ratios of triangles are reduced to ratios of line segments, namely

$$\frac{S_{PAB}}{S_{PAM}} = \frac{\overline{AB}}{\overline{AM}}, \quad \frac{S_{PAM}}{S_{QAM}} = \frac{\overline{PM}}{\overline{QM}}, \quad \text{and} \quad \frac{S_{QAM}}{S_{QAB}} = \frac{\overline{PM}}{\overline{QM}}.$$

Although short and simple, it does not provide any geometric insight.

Proof 1.

$$\frac{S_{PAB}}{S_{QAB}} = \frac{S_{PAB}}{S_{PAM}} \cdot \frac{S_{PAM}}{S_{QAM}} \cdot \frac{S_{QAM}}{S_{QAB}} = \frac{\overline{AB}}{\overline{AM}} \cdot \frac{\overline{PM}}{\overline{QM}} \cdot \frac{\overline{AM}}{\overline{AB}} = \frac{\overline{PM}}{\overline{QM}}.$$

□

In the second proof, we consider four cases represented in Fig. 3. They depend on whether M lies between P and Q, or A and B. We base it on Theorem 2.5 (*Elements*, VI.1).

In figures Fig. 4–8 below, we use symbols $S_1, ..., S_4$ to represent triangles. They aim to simplify formulas applied in the proof.

Proof 2. **Case 1.** Point M lies between P, Q and between A, B; see Fig. 4.

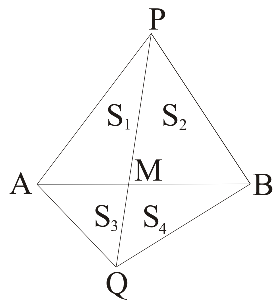


FIGURE 4. Co-side Theorem, case 1.

By Theorem 2.5, we have the following equalities (proportions)

$$\frac{S_1}{S_3} = \frac{\overline{PM}}{\overline{QM}}, \quad \frac{S_2}{S_4} = \frac{\overline{PM}}{\overline{QM}}.$$

The following equalities emulate Euclid's proportions (propositions V.12, 19), yet, they can also be justified in the arithmetic of fractions:

$$\frac{S_1}{S_3} = \frac{S_2}{S_4} = \frac{S_1 + S_2}{S_3 + S_4} = \frac{\overline{PM}}{\overline{QM}}.$$

Since $S_1 + S_2 = S_{PAB}$, and $S_3 + S_4 = S_{QAB}$, finally, the required equality obtains, namely

$$\frac{S_{PAB}}{S_{QAB}} = \frac{\overline{PM}}{\overline{QM}}.$$

Note, when $B = M$, the Co-side Theorem is the same as Euclid's VI.1 (see Fig. 5), namely

$$\frac{S_{PAB}}{S_{QAB}} = \frac{\overline{PB}}{\overline{BQ}} = \frac{\overline{PM}}{\overline{MQ}}.$$

Case 2. Point M lies between A, B, but not between P, Q; see Fig. 6. By Theorem 2.5

$$\frac{S_1 + S_3}{S_3} = \frac{\overline{PM}}{\overline{QM}} = \frac{S_2 + S_4}{S_4}.$$

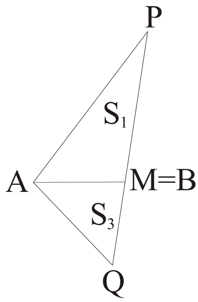


FIGURE 5. Co-side Theorem turns into VI.1.

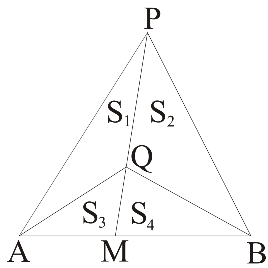


FIGURE 6. Co-side Theorem, case 2.

Similarly, Euclid's theory of proportions, as well as the arithmetic of fractions justifies the following case

$$\frac{S_1 + S_2 + S_3 + S_4}{S_3 + S_4} = \frac{\overline{PM}}{\overline{QM}}.$$

Hence,

$$\frac{S_{PAB}}{S_{QAB}} = \frac{\overline{PM}}{\overline{QM}}.$$

Case 3. Point M lies between P, Q and not between A, B; see Fig. 7. By Theorem 2.5

$$\frac{S_1 + S_2}{S_3 + S_4} = \frac{\overline{PM}}{\overline{QM}} = \frac{S_2}{S_4}.$$

The following equality emulates another proposition concerning Euclid's proportions (proposition V.19), yet, it also can be justified in the arithmetic of fractions.

$$\frac{S_1 + S_2 - S_2}{S_3 + S_4 - S_4} = \frac{S_1}{S_3} = \frac{\overline{PM}}{\overline{QM}}.$$

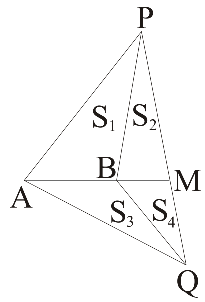


FIGURE 7. Co-side Theorem, case 3.

Hence,

$$\frac{S_{PAB}}{S_{QAB}} = \frac{\overline{PM}}{\overline{QM}}.$$

Case 4. Point M lies neither between P and Q, nor A and B; see Fig. 8. By

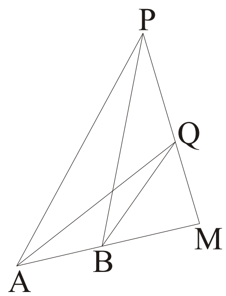


FIGURE 8. Co-side, case 4

Theorem 2.5

$$\frac{S_{PAB} + S_{PBM}}{S_{QAB} + S_{QBM}} = \frac{\overline{PM}}{\overline{QM}} = \frac{S_{PBM}}{S_{QBM}}.$$

As in the previous cases,

$$\frac{S_{PAB}}{S_{QAB}} = \frac{\overline{PM}}{\overline{QM}}.$$

□

To sum up, in Euclid's VI.1, the Co-side requirement guarantees that triangles are of the same height, while their bases can differ. In the Co-side Theorem, triangles share the base, while their heights differ. Euclid's VI.1 enables one to reduce the geometric pattern represented in Fig. 2 to the proportion of lines given by formula $\frac{S_{ABC}}{S_{ADC}} = \frac{\overline{BC}}{\overline{DC}}$. The Co-side Theorem provides us with four more geometric patterns that reduce the proportions of triangles to the proportions of lines. In the next section, we show how to exploit these new patterns in school geometry.

4. Co-side Theorem in solving school geometry problems

In this section, we show how to apply the Co-side Theorem in solving school problems. For each of the four cases discussed in the previous section, we select a separate task. In these cases, we process fractions more liberally, in a way familiar to students.

Exercise 4.1 (Co-side Theorem, case 1)**.** *In trapezium $ABCD$, where $AB \parallel DC$, O is the intersection point of diagonals AC and BD. Area of triangle ADB is 50. Diagonals AC and BD intersect such that $\overline{AO} : \overline{OC} = 5 : 1$. Find the area of the trapezium.*

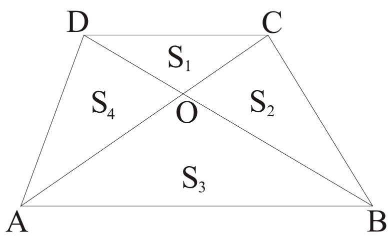


Figure 9. Exercises 1, Co-side Theorem case 1

To simplify our argument, we set (see Fig. 9):

$$S_{DCO} = S_1, S_{OCB} = S_2, S_{OBA} = S_3, S_{ADO} = S_4.$$

By Co-side Theorem,

$$\frac{S_4 + S_3}{S_1 + S_2} = \frac{\overline{AO}}{\overline{OC}} = \frac{5}{1}.$$

Since $S_{ADB} = S_3 + S_4 = 50$, it follows that $S_1 + S_2 = 10$. Hence, $S_{ABCD} = S_4 + S_3 + S_1 + S_2 = 60$. □

Exercise 4.2 (Co-side Theorem, case 2; [11], p. 19)**.** *Let ABC be a triangle with a point P inside. Let lines AP and CB intersect at D, lines BP and CA – at E, and lines CP and AB – at F. Show that*

$$\frac{\overline{PD}}{\overline{AD}} + \frac{\overline{PE}}{\overline{BE}} + \frac{\overline{PF}}{\overline{CF}} = 1.$$

We set (see Fig. 10):

$$S_{PCE} = S_1, S_{PDC} = S_2, S_{PBD} = S_3, S_{PEA} = S_4, S_{PFB} = S_5, S_{PFA} = S_6.$$

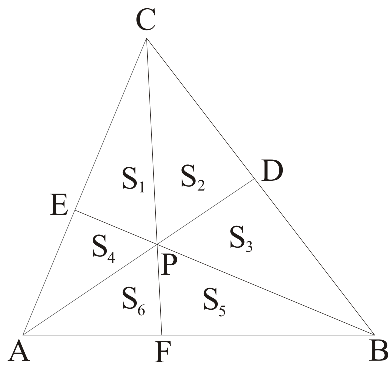


FIGURE 10. Exercises 2, Co-side Theorem case 2

By Co-side Theorem, we have the following equalities (proportions)

$$\frac{\overline{PD}}{\overline{AD}} = \frac{S_2 + S_3}{S_{ABC}},$$

$$\frac{\overline{PE}}{\overline{BE}} = \frac{S_1 + S_6}{S_{ABC}},$$

$$\frac{\overline{PF}}{\overline{CF}} = \frac{S_5 + S_4}{S_{ABC}}.$$

By adding their left and right sides, we obtain

$$\frac{\overline{PD}}{\overline{AD}} + \frac{\overline{PE}}{\overline{BE}} + \frac{\overline{PF}}{\overline{CF}} = \frac{S_{ABC}}{S_{ABC}} = 1.$$

□

Exercise 4.3 (Co-side Theorem, case 3; [8], p. 241)**.** *Show that the three medians of a triangle divide it into six smaller triangles of equal area.*

Let set (see Fig. 11):

$$S_{PBF} = S_1, S_{PEB} = S_2, S_{PCE} = S_3, S_{PFA} = S_4, S_{PDC} = S_5, S_{PDA} = S_6.$$

The medians assumption means:

$$\overline{BF} = \overline{FA}, \quad \overline{AD} = \overline{DC}, \quad \overline{CE} = \overline{EB}$$

By Theorem 2.5 $S_1 = S_4, \quad S_6 = S_5, \quad S_3 = S_2$. Applying Co-side Theorem, we get

$$\frac{S_1 + S_4}{S_2 + S_3} = \frac{\overline{AD}}{\overline{DC}} = 1,$$

which means $S_1 = S_2$. Similarly, we show that $S_3 = S_5$ and $S_4 = S_6$. Hence, $S_1 = S_2 = S_3 = S_4 = S_5 = S_6$. □

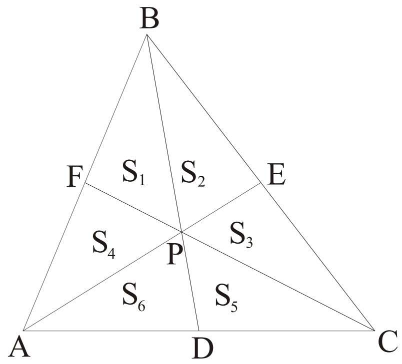


FIGURE 11. Exercises 3, Co-side Theorem case 3

Exercise 4.4 (Co-side Theorem, case 4; [10], p.14). *Let ABC be a triangle with $\overline{BC} = a$. Let D be the mid point of the side $\overline{AC}$, and O of line segment $\overline{BD}$. Let BC and AO intersect at P. Show that $\overline{CP} = \frac{2}{3}a$.*

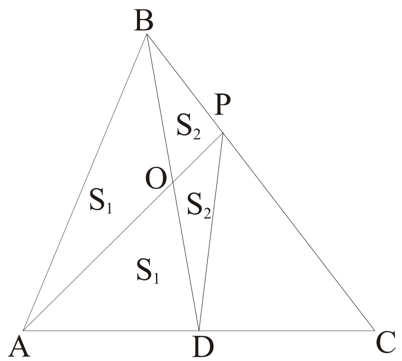


FIGURE 12. Exercises 4, Co-side Theorem case 4

Since $\overline{BO} = \overline{OD}$, by Theorem 2.5 we get

$$S_{BAO} = S_{AOD}, \quad S_{BPO} = S_{POD}.$$

Let set (see Fig. 12) $S_1 = S_{BAO} = S_{AOD}$ and $S_2 = S_{BPO} = S_{POD}$. Since $\overline{AD} = \overline{DC}$, applaing Co-side Theorem we get:

$$\frac{S_{ABP}}{S_{DBP}} = \frac{S_1 + S_2}{2 \cdot S_2} = \frac{\overline{AC}}{\overline{DC}} = \frac{2}{1}.$$

It follows that $S_1 = 3 \cdot S_2$. Again, by Co-side Theorem, we get

$$\frac{\overline{BC}}{\overline{CP}} = \frac{2 \cdot S_1}{S_1 + S_2} = \frac{3}{2}.$$

Since $\overline{BC} = a$, we finally obtain $\overline{CP} = \frac{2}{3}a$. □

5. Advanced problems

Exercise 5.1 ([5], p. 212)**.** *Let D, E, F divide the sides of the triangle ABC into thirds. Draw $\overline{AE}$, $\overline{BD}$, $\overline{CF}$. Show that the area of the triangle inside, S_5, equals $\frac{1}{7}$ the area of the whole triangle.*

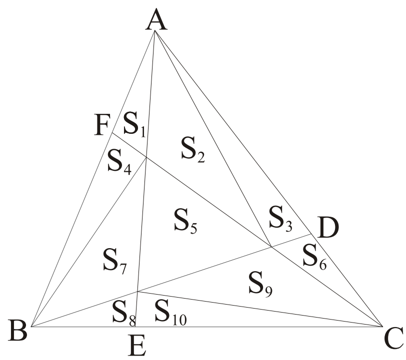


Figure 13. Exercise 5.1

To simplify our argument, we apply the notation as shown in the figure Fig. 13. Thus

$$\overline{FB} = 2\overline{AF}, \ \ \overline{DA} = 2\overline{CD} \ \ \overline{EC} = 2\overline{BE}.$$

By Theorem 2.5, we have $S_4 = 2S_1$, $S_3 = 2S_6$, and $S_{10} = 2S_8$. Co-side Theorem, case 3, gives

$$\frac{S_5 + S_9}{S_7} = \frac{2}{1}, \quad \frac{S_5 + S_2}{S_9} = \frac{2}{1}, \quad \frac{S_5 + S_7}{S_2} = \frac{2}{1}.$$

It follows from the above that $S_5 + S_9 = 2S_7$, $S_5 + S_2 = 2S_9$, and $S_5 + S_7 = 2S_2$. Adding up the above equations we have

$$3S_5 = S_7 + S_9 + S_2 \tag{5.1}$$

Again, by Co-side Theorem

$$\frac{S_9 + S_8 + S_{10}}{S_3 + S_6} = \frac{2}{1}.$$

Hence

$$\begin{aligned} S_9 + S_8 + S_{10} &= 2(S_3 + S_6), \\ S_9 + S_8 + 2S_8 &= 2(2S_6 + S_6), \\ 3S_8 + S_9 &= 6S_6. \end{aligned}$$

Similarly, we show

$$\begin{aligned} 3S_6 + S_2 &= 6S_1, \\ 3S_1 + S_7 &= 6S_8. \end{aligned}$$

Adding up the above equations, we have

$$S_7 + S_9 + S_2 = 3(S_6 + S_1 + S_8)$$

By (5.1), we have

$$S_5 = S_6 + S_1 + S_8. \tag{5.2}$$

Note that $S_{ABC} = 3S_1 + 3S_8 + 3S_6 + S_5 + S_7 + S_9 + S_2$. By the (5.1) and (5.2) we have

$$\begin{aligned} S_{ABC} &= 3S_1 + 3S_8 + 3S_6 + S_6 + S_1 + S_8 + 3(S_6 + S_1 + S_8) \\ &= 7(S_6 + S_1 + S_8) = 7S_5. \end{aligned}$$

Hence $S_5 = \frac{S_{ABC}}{7}$. □

Exercise 5.2. *Let ABC be a triangle, with D lying on line AB and E on line BC, in such a way that $DE \parallel AC$. The diagonals DC and AE intersect at point F. Let lines BF and AC intersect at point G. Show that $\overline{AG} = \overline{GC}$ ([4], p.18).*

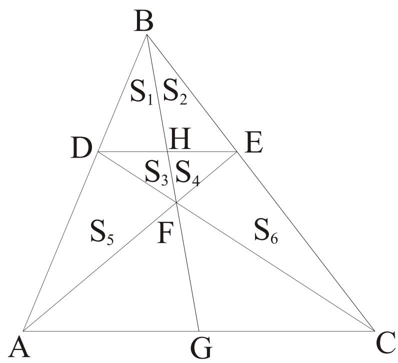


FIGURE 14. Exercise 5.2

We apply the notation as shown in the figure Fig. 14. From the assumption $DE \parallel AC$. By the definition of parallel segments we have

$$S_4 + S_3 + S_5 = S_3 + S_4 + S_6,$$

and

$$S_5 = S_6. \tag{5.3}$$

It follows from *Elements*, VI.2 (Thales's Theorem)[6] that

$$\frac{\overline{BD}}{\overline{DA}} = \frac{\overline{BE}}{\overline{EC}}. \tag{5.4}$$

[6] ([2], p. 52) If some straight-line is drawn parallel to one of the sides of a triangle then it will cut the sides of the triangle proportionally. And if the sides of a triangle are cut proportionally then the straight-line joining the cutting will be parallel to the remaining side of the triangle.

By Theorem 2.5

$$\frac{S_1+S_3}{S_5}=\frac{\overline{BD}}{\overline{DA}},\quad \frac{S_2+S_4}{S_6}=\frac{\overline{BE}}{\overline{EC}}. \tag{5.5}$$

By (5.4) and (5.5) we have

$$\frac{S_1+S_3}{S_5}=\frac{S_2+S_4}{S_6}.$$

By (5.3),

$$S_1+S_3=S_2+S_4. \tag{5.6}$$

Co-side Theorem (case 1) gives

$$\frac{S_1+S_3}{S_2+S_4}=\frac{\overline{DH}}{\overline{HE}},$$

Hence, $\overline{DH}=\overline{HE}$. By Co-side Theorem (case 3)

$$\frac{S_1+S_3+S_5}{S_2+S_4+S_6}=\frac{\overline{AG}}{\overline{GC}}.$$

By (5.6) and (5.3) $S_1+S_3+S_5=S_2+S_4+S_6$. Hence $\overline{AG}=\overline{GC}$. □

References

[1] P. Błaszczyk, *From Euclid's Elements to the methodology of mathematics. Two ways of viewing mathematical theory*, Annales Universitatis Paedagogicae Cracoviensis. Studia ad Didacticam Mathematicae Pertinentia **10** (2018), 5–15.

[2] P. Błaszczyk, A. Petiurenko, *Euclid's theory of proportion revised*, Annales Universitatis Paedagogicae Cracoviensis. Studia ad Didacticam Mathematicae Pertinentia **11** (2018), 37–61.

[3] K.Borsuk, W.Szmielew, *Podstawy geometrii*, PWN, Warszawa (1972).

[4] S. Chou, X. Gao, J. Zhang, *Machine Proofs in Geometry*, World Scientific, Singapore (1994).

[5] R. Hartshorne, *Geometry: Euclid and Beyond*, Springer, New York (2000).

[6] P. Janičić, J. Narboux, P. Quaresma, *The area method: a Recapitulation*, Journal of Automated Reasoning, Springer Verlag **48**.4 (2012), 489–532.

[7] Z. Krygowska, *Geometria*, WSziP, Warszawa (1979).

[8] M. Kurczab, E. Kurczab, *Matematyka 2. Zakres podstawowy. Reforma 2019*, OE Pazdro (2020)

[9] R. Millman, G.Parker, *Geometry A Metric Approach with Models*, Springer, Berlin (1991).

[10] *Teraz Matura. Matematyka. Poziom rozszerzony. Zbiór zadań i zestawów maturalnych*, Nowa Era, Warszawa (2019).

[11] Z. Stankowa, T. Rike *A Decade of the Berkeley Math Circle: The American Experience*, American Mathematical Soc. (2008).

Institute of Mathematics, Pedagogical University of Cracow, Poland
anna.petiurenko@up.krakow.pl

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

THE BAIRE CATEGORY OF THE HYPERSPACE OF NONTRIVIAL CONVERGENT SEQUENCES

MICHAŁ POPŁAWSKI

Abstract. Assume that X is a regular space. We study topological properties of the family $S_c(X)$ of nontrivial convergent sequences in X equipped with the Vietoris topology. In particular, we show that if X has no isolated points, then $S_c(X)$ is a space of the first category which answers the question posed by S. García-Ferreira and Y.F. Ortiz-Castillo in [1].

Keywords:Baire category, Nontrivial convergent sequences, Vietoris topology MSC2010: 54A20, 54B20, 54E52

Received 30 April 2021; revised 15 July 2021; accepted 28 July 2021

1. Introduction

Let X be a topological Hausdorff space. The Vietoris topology on the family $K(X)$ of all compact subsets of X is generated by a base consisting of sets

$$(1.1)\quad \langle V_1, \dots, V_n \rangle := \left\{ K \in K(X) \colon K \subset \bigcup_{i=1}^{n} V_i \text{ and } K \cap V_i \neq \emptyset \text{ for } 1 \leq i \leq n \right\},$$

where n runs over $\mathbb{N}$ and $V_1, \dots, V_n$ are nonempty open sets in X.

Our notation is consistent with that used in [1]. We say that $S \subset X$ is a nontrivial convergent sequence in X if $S = \{x_n\}_{n\in\mathbb{N}} \cup \{\lim S\}$ for some injective sequence $(x_n)_{n\in\mathbb{N}}$ in X which is convergent to some $\lim S \in X \setminus \{x_n\}_{n\in\mathbb{N}}$. In other words, S is a set of terms of an injective convergent sequence with its limit. Obviously, S is compact for any space X, and $S_c(X)$ is empty for a discrete space X. Hence each closed subset F of S is compact. Consequently, the spaces $K(X)$ and the family $CL(X)$ of all closed subsets of X with the topology generated by an analogous base, given by (1.1), introduce the same topology in their subspace $S_c(X)$.

https://doi.org/10.2478/9788366675360-018.

In general, separation axioms of the spaces $CL(X)$ and $K(X)$ depend on X. More precisely, $CL(X)$ is normal if and only if X is compact (see [5]), and $K(X)$ is metrizable if and only if X is metrizable ([3]).

The main aim of this paper is to show that $S_c(X)$ is of first category in itself under the assumption that X is regular and crowded (i.e. has no isolated points). This result gives a positive answer to Question 3.4 in [1]. The authors of [1] asked whether $S_c(X)$ is of the first category in itself if X is a metric crowded space. Note that in [2] authors used other methods to prove this result. From now on, sets of the first category will be called also meager. In our considerations we will use the Banach Category Theorem (see [4]) which states that in any topological space the union of a family of open meager sets is meager, too. Thanks to this fact, it suffices to construct a meager open neighbourhood of any $S \in S_c(X)$. Such a construction is possible in the case of regular crowded space X. For reader's convenience we repeat or modify some arguments from [1].

2. Main result

We will follow some ideas taken from the paper [1] while considering some specific subsets of $S_c(X)$. For given $k, m \in \mathbb{N}, 1 \leq i \leq k$ and pairwise disjoint nonempty closed sets $C_1, \ldots, C_k$, we denote

$$\left\{ S \in S_c(X) \colon S \subset \bigcup_{j=1}^{k} C_j \text{ and } 1 \leq \operatorname{card}(S \cap C_j) \leq m \text{ for all } j \leq k,\ j \neq i \right\}$$

by $N_k^i(m, \{C_j \colon j \leq k\})$ and $\bigcup_{i=1}^{k} N_k^i(m, \{C_j \colon j \leq k\})$ by $N_k(m, \{C_j \colon j \leq k\})$.

It can be immediately seen that, if $S \in N_k^i(m, \{C_j \colon j \leq k\})$, then $\operatorname{card}(S \cap C_i) = \omega$.

The following lemma is from [1, Lemma 3.1].

Lemma 2.1. *Let X be a crowded topological space. Then the set*

$$N_k^i(m, \{C_j \colon j \leq k\})$$

is nowhere dense, closed in $S_c(X)$ whenever $k, m \in \mathbb{N}, 1 \leq i \leq k$ and $C_1, \ldots, C_k$ are pairwise disjoint, nonempty and closed sets.

Note that the assertion of the above fact can be lost if X is not crowded.

Example 2.2. Consider sets $X := [0,1],\ Y := [0,1] \cup \{2\},\ Z := [0,1] \cup [2,3]$ with the Euclidean topology. Note that X is a closed subspace of Y, Y is a closed subspace of Z and both spaces X and Z are crowded but Y has one isolated point. Take $k := 2,\ m := 1,\ C_1 := [0,1],\ C_2 := \{2\},\ i := 2$. By Lemma 2.1, the set $N_2^2(1, \{C_1, C_2\})$ is nowhere dense in the space $S_c(Z)$. Nevertheless, this set is not nowhere dense in $S_c(Y)$. To check it, set $V_1 := [0,1],\ V_2 := \{2\}$, and consider the open set $\langle V_1, V_2 \rangle$. Obviously, each $S \in \langle V_1, V_2 \rangle$ has exactly one point in $S \cap V_2$. Thus $S \in N_2^2(1, \{C_1, C_2\})$ and consequently, $\langle V_1, V_2 \rangle \subset N_2^2(1, \{C_1, C_2\})$.

Moreover, all sets $N_k^i(m, \{C_j : j \le k\})$ as in Lemma 2.1 are nowhere dense in $S_c(X)$.

The next result (see [1, Theorem 3.2]) is an application of the previous lemma.

Lemma 2.3. *Suppose that U_1, U_2 are nonempty, closed and disjoint subsets of a crowded space X. Then $\langle \mathrm{Int}(U_1), \mathrm{Int}(U_2)\rangle$ is meager in $S_c(X)$.*

Proof. Thanks to Lemma 2.1 it suffices to observe that

$$\langle \mathrm{Int}(U_1), \mathrm{Int}(U_2)\rangle \cap S_c(X) \subset \bigcup_{m\in\mathbb{N}} N_2(m, \{U_1, U_2\}).$$

But this follows from the disjointness of closed sets U_1, U_2. □

Now, we will construct the respective neighbourhoods of nontrivial convergent sequences.

Lemma 2.4. *Suppose X is a Hausdorff space and $S = \{x_n\}_{n\in\mathbb{N}} \cup \{\lim S\} \in S_c(X)$. There are neighbourhoods $V_n, n \in \mathbb{N}$ of x_n's and a neighbourhood V_S of $\lim S$ such that*

$$V_1 \cap V_n = \emptyset \textit{ for all } n \ge 2 \quad \textit{ and } V_1 \cap V_S = \emptyset.$$

Proof. Use the Hausdorff axiom to find disjoint neighbourhoods V_1' of x_1, and V_S of $\lim S$. Since $(x_n)_{n\in\mathbb{N}}$ is convergent to $\lim S$, the set

$$M := \{m \ge 2 : x_m \notin V_S\}$$

is finite. Again, for each $m \in M$ use the Hausdorff axiom to find disjoint neighbourhoods $V_{1,m}'$ of x_1, and V_m of x_m. The intersection $V_1 = V_1' \cap \bigcap_{m\in M} V_{1,m}$ satisfies

$$V_1 \cap V_m = \emptyset \text{ for each } m \in M.$$

Then it suffices to define $V_k := V_S$ for all $k \notin M \cup \{1\}$. □

Proposition 2.5. Suppose X is a regular space and

$$S = \{x_n\}_{n\in\mathbb{N}} \cup \{\lim S\} \in S_c(X).$$

Then there are nonempty, closed and disjoint sets U_1, U_2 with $S \in \langle \mathrm{Int}(U_1), \mathrm{Int}(U_2)\rangle$.

Proof. Let V_S, V_n, $n \in \mathbb{N}$, be the respective neighbourhoods of $\lim S$, x_n, $n \in \mathbb{N}$ considered in Lemma 2.4. Since $V_1 \cap (V_S \cup \bigcup_{n\ge 2} V_n) = \emptyset$, we have

$$V_1 \cap \mathrm{cl}(V_S \cup \bigcup_{n\ge 2} V_n) = \emptyset.$$

Then we use the regularity of X to find a neighbourhood W_1 of x_1 such that $\mathrm{cl}(W_1) \subset V_1$. Put $U_1 := \mathrm{cl}(W_1)$, $U_2 := \mathrm{cl}(V_S \cup \bigcup_{n\ge 2} V_n)$. Obviously, these sets are nonempty, closed and disjoint. We need to show that $S \in \langle \mathrm{Int}(U_1), \mathrm{Int}(U_2)\rangle$. Indeed, $x_1 \in W_1 \subset \mathrm{Int}(U_1)$ and $\{\lim S\} \cup \bigcup_{n\ge 2}\{x_n\} \subset V_S \cup \bigcup_{n\ge 2} V_n \subset \mathrm{Int}(U_2)$. □

Theorem 2.6. *Suppose that X is a regular crowded space. Then the space $S_c(X)$ is of the first category in itself.*

Proof. Take $S \in S_c(X)$ and a neighbourhood of S of the form $\langle \mathrm{Int}(U_1), \mathrm{Int}(U_2)\rangle$ considered in Proposition 2.5. Then by Lemma 2.3, this neighbourhood is meager. Therefore, by the Banach Category Theorem, $S_c(X)$ is of the first category as a union of meager neighbourhoods of its points. $\square$

References

[1] S. García-Ferreira, Y.F. Ortiz-Castillo, *The hyperspace of convergent sequences*, Topology Appl. **196** (2015), 795–804.
[2] S. García-Ferreira, Y.F. Ortiz-Castillo, R. Rojas-Hernández, *Categorical properties on the hyperspace of nontrivial convergent sequences*, Topol. P. **52** (2018), 265–279
[3] E. Michael, *Topologies on spaces of subsets*, Trans. Amer. Math. Soc. **71**.1 (1951), 152–182.
[4] J.C. Oxtoby, *Measure and category*, Springer, New York 1980.
[5] N.V. Velichko, *On spaces of closed subsets*, Sib. Math. J. **16** (1975), 484–486.

Jan Kochanowski University in Kielce, Poland
mpoplawski@ujk.edu.pl

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON A RIEMANNIAN MANIFOLD WITH TWO ORTHOGONAL DISTRIBUTIONS

VLADIMIR ROVENSKI

Abstract. In this article, we introduce and study a Riemannian manifold equipped with a distribution represented as the sum of two distributions. We define the mixed scalar curvature of this structure and prove integral formulas generalizing classical results.

Keywords: Riemannian metric, distribution, foliation, second fundamental form, mean curvature vector, mixed scalar curvature

MSC2010: 53C15, 57R25

Received: 11 February 2021; accepted 15 June 2021

1. Introduction

Distributions on a manifold (i.e., vector subbundles of the tangent bundle) are used to build up notions of integrability, and specifically of foliations of arbitrary dimension, and arise in such topics of differential geometry as submersions, Lie groups actions and sub-Riemannian geometry [2, 3]. We do not discuss here the existence of distributions or foliations on a manifold, but only mention that a compact manifold with Euler number equal to zero admits a one-dimensional distribution. Integral formulae are useful for solving many problems in the geometry of manifolds (e.g., the Gauss-Bonnet formula prevents the specification of the Gaussian curvature of a surface) and foliations, see [1, 4, 7, 8, 10]:

- characterizing of foliations, whose leaves have a given geometric property;
- prescribing the higher mean curvatures of the leaves of a foliation;
- minimizing functionals like volume defined for tensor fields on a foliation.

The first known integral formula for a codimension one foliation of a closed Riemannian manifold shows us that the integral mean curvature of the leaves is equal

https://doi.org/10.2478/9788366675360-019.

to zero, see [6]. There is a series of integral formulas for codimension one foliations (starting with the formula of G. Reeb), see survey in [1, 8]. The second formula in the series of total σ_k's (elementary symmetric functions of principal curvatures of the leaves with a unit normal vector field N) is

$$\int_M (2\,\sigma_2 - \mathrm{Ric}_{N,N})\,\mathrm{d}\,\mathrm{vol}_g = 0. \tag{1.1}$$

These integral formulas (for codimension one foliations or distributions) were obtained by applying the Divergence Theorem to suitable vector fields. There is also a series of integral formulas for two complementary orthogonal distributions $\mathcal{D}$ and $\mathcal{D}^\perp$ (or foliations) of arbitrary dimension on a closed Riemannian manifold (M,g), see [7]. The following famous integral formula of P. Walczak [10] generalizes (1.1) and has many applications:

$$\int_M \big(\mathrm{S}_{\,\mathrm{mix}} + \|h\|^2 + \|h^\perp\|^2 - \|H\|^2 - \|H^\perp\|^2 - \|T\|^2 - \|T^\perp\|^2\big)\,\mathrm{d}\,\mathrm{vol}_g = 0. \tag{1.2}$$

Here, $h, T : \mathcal{D}\times\mathcal{D}\to\mathcal{D}^\perp$ and $H = \mathrm{Tr}_g\, h$ are the second fundamental form, the integrability tensor and the mean curvature vector field of the distribution $\mathcal{D}$ (and similarly for $\mathcal{D}^\perp$). The mixed scalar curvature $\mathrm{S}_{\,\mathrm{mix}}$ – the simplest curvature invariant of a Riemannian almost-product manifold – is defined as an averaged sum of sectional curvatures of 2-planes that non-trivially intersect with both distributions; for an adapted orthonormal frame we have

$$\mathrm{S}_{\,\mathrm{mix}} = \sum_{1\le a\le \dim\mathcal{D} < b \le \dim M} K(e_a, e_b).$$

In this article, we represent a distribution on a manifold as the sum of two distributions, and generalize some of the results of [6, 10]. In Section 2 we define the modified divergence operator for a distribution and generalize the integral formula by G. Reeb for a distribution represented as the sum of two distributions, one of which is a line field. In Section 3 we define the mixed scalar curvature of a distribution represented as the sum of two orthogonal distributions on a Riemannian manifold, find the modified divergence of a geometrically significant vector field and prove a new integral formula (with this kind of scalar curvature) that generalizes (1.2). Examples with one-dimensional distributions and paracontact metric manifolds illustrate the results.

2. The Reeb type integral formula

A *distribution* $\mathcal{D}$ on a smooth m-dimensional manifold M is a subbundle of TM of rank n, to each point $x\in M$ it assigns an n-dimensional subspace $\mathcal{D}_x$ of the tangent space T_xM smoothly depending on x. A Riemannian manifold (M,g) equipped with a self-adjoint (1,1)-tensor P such that $P^2=P$ is called a Riemannian almost-product manifold, e.g., [3]. The tensors P and $\mathrm{id}_{TM} - P$ are orthoprojectors onto complementary orthogonal distributions $\mathcal{D}$ and $\mathcal{D}^\perp$.

In the beginning of the section, we define the divergence type operator for distributions on a Riemannian manifold (M, g) and prove the Divergence type theorem. Recall that the divergence of a $(1, s)$-tensor field S is a $(0, s)$-tensor field $\operatorname{div} S = \operatorname{trace}(Y \to \nabla_Y S)$, or, in coordinates, $(\operatorname{div} S)_{i_1,\dots i_s} = \nabla_j S^j_{\cdot\, i_1,\dots i_s}$. For $s = 0$, we get the divergence of a vector field.

Definition 2.1. Given orthoprojector $P : TM \to \mathcal{D}$ on the distribution $\mathcal{D}$, the *P-divergence of a vector field* $X \in \mathcal{X}_M$ is defined by

$$\operatorname{div}_P X = \operatorname{trace}(Y \to P\,\nabla_{PY} X) = \sum\nolimits_s \langle P\,\nabla_{Pe_s} X, e_s\rangle, \tag{2.1}$$

where $\{e_s\}$ is any local orthonormal frame in M.

A distribution is said to be harmonic if its mean curvature vector field vanishes. There are topological restrictions on the existence of a Riemannian metric, for which a foliation (i.e., an integrable distribution) is harmonic, see [9].

The following proposition generalizes the classical Divergence Theorem.

Proposition 2.2. The distributions $\mathcal{D} = P(TM)$ and its orthogonal complement are harmonic on a Riemannian manifold (M, g) if and only if

$$\operatorname{div} P = 0, \tag{2.2}$$

where $P : TM \to \mathcal{D}$ is the orthoprojector. In this case, for a compact manifold M with an inner unit normal ν on the boundary and any vector field X we get

$$\int_M (\operatorname{div}_P X)\, d\operatorname{vol}_g = \int_{\partial M} \langle P(X), \nu\rangle\, d\omega.$$

Proof. The following equality holds for any vector field X:

$$(\operatorname{div} P)(X) = \sum\nolimits_s \langle \nabla_{e_s}(P(X)) - P(\nabla_{e_s} X),\ e_s\rangle.$$

If $X \perp \mathcal{D}$, then $P(X) = 0$ and $(\operatorname{div} P)(X) = \langle X,\, H\rangle$, and if $X \in \mathcal{D}$, then $P(X) = X$ and $(\operatorname{div} P)(X) = -\langle X,\, H^\perp\rangle$. By the above, (2.2) is equivalent to conditions $H = 0 = H^\perp$. Using $P^2 = P$ and the definition of $\operatorname{div}_P X$, see (2.1), we obtain

$$\begin{aligned}
\operatorname{div}_P X &= \sum\nolimits_s \langle P\,\nabla_{Pe_s} X, e_s\rangle = \sum\nolimits_{s,t} \langle Pe_s, e_t\rangle\langle \nabla_{e_t} X, Pe_s\rangle\\
&= \sum\nolimits_{s,t} \langle e_s, Pe_t\rangle\langle P\,\nabla_{e_t} X, e_s\rangle = \sum\nolimits_t \langle P\,\nabla_{e_t} X, Pe_t\rangle\\
&= \sum\nolimits_t \langle P\,\nabla_{e_t} X, e_t\rangle = \operatorname{div}(P(X)) - (\operatorname{div} P)(X).
\end{aligned}$$

Thus, (2.2) is equivalent to the equality

$$\operatorname{div}_P X = \operatorname{div}(P(X)),\ X \in \mathcal{X}_M. \qquad \square \tag{2.3}$$

Suppose that $\mathcal{D}$ is the sum of two orthogonal distributions, that is $\mathcal{D} = \mathcal{D}_1 \oplus \mathcal{D}_2$, and let $P_1 : TM \to \mathcal{D}_1$ and $P_2 : TM \to \mathcal{D}_2$ be the orthoprojectors, hence $P =$

P_1+P_2. The following tensors (called the *P-second fundamental form* and the *P-integrability tensor*) generalize the second fundamental form and the integrability tensor of distributions $\mathcal{D}_1$ and $\mathcal{D}_2$:

$$\begin{aligned}\hbar_1(X,Y) &= P_2(\nabla_{P_1X}P_1Y+\nabla_{P_1Y}P_1X)/2,\\ \hbar_2(X,Y) &= P_1(\nabla_{P_2X}P_2Y+\nabla_{P_2Y}P_2X)/2,\\ \mathcal{T}_1(X,Y) &= P_2\big(\nabla_{P_1X}P_1Y-\nabla_{P_1Y}P_1X\big)/2,\\ \mathcal{T}_2(X,Y) &= P_1\big(\nabla_{P_2X}P_2Y-\nabla_{P_2Y}P_2X\big)/2. \end{aligned}\tag{2.4}$$

The *P-mean curvature vectors* $\mathcal{H}_i=\mathrm{Trace}_g\,\hbar_i$ of $\mathcal{D}_i$ $(i=1,2)$ are given by

$$\mathcal{H}_1=\sum\nolimits_s P_2(\nabla_{P_1e_s}P_1e_s),\quad \mathcal{H}_2=\sum\nolimits_s P_1(\nabla_{P_2e_s}P_2e_s).\tag{2.5}$$

The definition of $\mathcal{H}_i$ is correct since the distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ are orthogonal. We will also use a local orthonormal *adapted* basis, i.e., $\{e_j\}$ in $\mathcal{D}_1$ and $\{e_a\}$ in $\mathcal{D}_2$. Then, $Pe_j=P_1e_j$, $Pe_a=P_2e_a$, $\mathcal{H}_1=\sum_j P_2(\nabla_{e_j}e_j)$, etc. Indices j,k will be used for base vectors in $\mathcal{D}_1$, and indices a,b – for base vectors in $\mathcal{D}_2$. We define the norms of the above tensors as usual by

$$\|\hbar_1\|^2=\sum\nolimits_{s,t}\|\,\hbar_1(e_s,e_t)\|^2,\quad \|\mathcal{T}_1\|^2=\sum\nolimits_{s,t}\|\,\mathcal{T}_1(e_s,e_t)\|^2.$$

A distribution is called *P-harmonic* if the P-mean curvature vector vanishes, and the distribution is *P-integrable* if the P-integrability tensor vanishes, see (2.4).

Remark 2.3. The second fundamental forms h_i, the integrability tensors T_i and the mean curvature vectors $H_i=\mathrm{Trace}_g\,h_i$ of $\mathcal{D}_i$ (of $\mathcal{D}_1$ and $\mathcal{D}_2$ as distributions in TM) are related with the above tensors as $\hbar_i=P\circ h_i$, $\mathcal{T}_i=P\circ T_i$ and $\mathcal{H}_i=P(H_i)$.

We can consider distributions defined outside a "singularity set" Σ, a finite union of pairwise disjoint closed submanifolds of codimension greater than 2. If X is a vector field on $M\setminus\Sigma$ such that $\|X\|\in L^2(M,g)$, then, see [4],

$$\int_M \mathrm{div}\,X\,\mathrm{d}V_g=0.\tag{2.6}$$

The following theorem generalizes the classical result of G. Reeb [6].

Theorem 2.4. *Let (M,g) be a closed Riemannian manifold, and a distribution $\mathcal{D}=P(TM)$ be the sum of a codimension one distribution $\mathcal{D}_1=P_1(TM)$ and a line field $\mathcal{D}_2=P_2(TM)$, spanned by a unit vector field N defined on $M\setminus\Sigma$ (outside the singularity set). Set $\mathcal{H}_{1,sc}=\langle\mathcal{H}_1,N\rangle$ – the mean P-curvature function of $\mathcal{D}_1$. If P satisfies condition (2.2) and $\|N\|\in L^2(M,g)$, then*

$$\int_M \mathcal{H}_{1,sc}\,\mathrm{d}\,\mathrm{vol}_g=0;\tag{2.7}$$

thus, either $\mathcal{H}_{1,sc}\equiv 0$ or $\mathcal{H}_{1,sc}(x)\,\mathcal{H}_{1,sc}(x')<0$ for some points $x\neq x'$ on $M\setminus\Sigma$.

Proof. By (2.1) and (2.5), $\mathcal{H}_{1,sc}=-\,\mathrm{div}_P\,N$. By conditions and (2.3), $\mathrm{div}_P\,N=\mathrm{div}\,N$. By conditions and (2.6), $\int_M \mathrm{div}\,N\,\mathrm{d}V_g=0$. Hence, (2.7) holds. □

3. The integral formula with the mixed scalar curvature

First, we adapt the construction of a curvature type tensor from [5].

Proposition 3.1. The mappings $R_i^P : \mathcal{X}_M \times \mathcal{X}_M \times \mathcal{X}_M \to \mathcal{X}_M$ $(1 \le i \le k)$

$$R_i^P(X,V)W = P_i^\perp(\nabla_{P_i^\perp X} P\,\nabla_{P_i V} - \nabla_{P_i V} P\,\nabla_{P_i^\perp X} - \nabla_{P\,[P_i^\perp X,\ P_i V]})\,P_i W$$

are tensor fields on (M,g).

Define the curvature type (0,4)-tensor $R^P : \mathcal{X}_M \times \mathcal{X}_M \times \mathcal{X}_M \times \mathcal{X}_M \to C^\infty(M)$ by

$$R^P(X,V,W,Y) = R_1^P(X,V,W,Y) + R_2^P(X,V,W,Y), \quad X,Y,V,W \in \mathfrak{X}_M. \tag{3.1}$$

If X and V are unit vectors of different distributions, then $R^P(X,V,V,X)$ will be called the mixed sectional P-curvature.

Definition 3.2. The following function on a Riemannian manifold (M,g) with two orthogonal distributions $\mathcal{D}_i = P_i(TM)$ $(i = 1,2)$ will be called the *mixed scalar P-curvature*:

$$\mathrm{S}^{\,P}_{\mathcal{D}_1,\mathcal{D}_2} = \sum\nolimits_{s<t} R^P(e_t,e_s,e_s,e_t), \tag{3.2}$$

where $\{e_s\}$ is a local orthonormal frame in M and the tensor R^P is given by (3.1).

Remark 3.3. Our definition (3.2) generalizes the notion of the mixed scalar curvature $\mathrm{S}_{\mathcal{D}_1,\mathcal{D}_2}$ of complementary orthogonal distributions, see [10]. Let $\mathcal{D}_1 \oplus \mathcal{D}_2 = TM$ and $\{e_i\}$ be a local adapted orthonormal frame on M: $\{e_1,\ldots,e_n\} \subset \mathcal{D}_1$, $\{e_{n+1},\ldots,e_m\} \subset \mathcal{D}_2$. If $X \in \mathcal{D}_1$ and $Y \in \mathcal{D}_2$ are unit vectors, then $K(X,Y) = R^P(X,Y,Y,X)$ is the mixed sectional curvature. Thus,

$$\mathrm{S}^{\,P}_{\mathcal{D}_1,\mathcal{D}_2} = \sum_{s<t} K(P_2 e_t, P_1 e_s) = \sum_{1\le a\le n<b\le m} K(e_b,e_a) = \mathrm{S}_{\mathcal{D}_1,\mathcal{D}_2}.$$

The tensors $\hbar_i, \mathcal{T}_i$ and $\mathcal{H}_i$, see (2.4) and (2.5), control the extrinsic geometry of the distribution $\mathcal{D}_i$ and are implemented in the formulas below generalizing [10, Proposition 1] on complementary distributions.

Proposition 3.4. For orthogonal distributions $\mathcal{D}_i = P_i(TM)$, $i = 1,2$, and $P = P_1 + P_2$ the following formula is valid:

$$\mathrm{div}_P(\mathcal{H}_1+\mathcal{H}_2) = \mathrm{S}^{\,P}_{\mathcal{D}_1,\mathcal{D}_2} + \|\hbar_1\|^2 + \|\hbar_2\|^2 - \|\mathcal{H}_1\|^2 - \|\mathcal{H}_2\|^2 - \|\mathcal{T}_1\|^2 - \|\mathcal{T}_2\|^2. \tag{3.3}$$

Proof. This consists of stages similar to those of [10, Proposition 1] (see also [5, Proposition 3.5]). $\square$

Remark 3.5. Define the endomorphisms $\mathcal{A}_i$ of distributions $\mathcal{D}_i$ on (M,g) by

$$\langle \mathcal{A}_i X, Y\rangle = \langle \hbar_i(X,Y), \mathcal{H}_i\rangle, \quad X,Y \in \mathcal{D}_i.$$

We can compose $(\mathcal{A}_i)^k = \mathcal{A}_i \circ \ldots \circ \mathcal{A}_i$ k times and consider the vector fields $\mathcal{Z}_k = (\mathcal{A}_1)^k\,\mathcal{H}_2 + (\mathcal{A}_2)^k\,\mathcal{H}_1$. For $k=0$, we have $\mathcal{Z}_0 = \mathcal{H}_1 + \mathcal{H}_2$ – its P-divergence is given in (3.3). The following problem (see [4] for two complementary distributions) seem to be interesting: *find the P-divergence of vector fields $\mathcal{Z}_k$ $(k>0)$ defined by means of extrinsic geometry of orthogonal distributions $\mathcal{D}_1$ and $\mathcal{D}_2$.* The formulae obtained in this way will be more complicated than (3.3), see e.g., calculating for $\mathcal{Z}_1$ in [4] when $\mathcal{D}_1\oplus\mathcal{D}_2=TM$.

The following theorem generalizes [10, Theorem 1] and [4, Theorem 2(i)].

Theorem 3.6. *Let a closed Riemannian manifold (M,g) be equipped with orthogonal distributions $\mathcal{D}_i=P_i(TM)$, $i=1,2$, defined on $M\setminus\Sigma$ (outside the singularity set). Set $P=P_1+P_2$ and assume (2.2) and $\|\mathcal{H}_1+\mathcal{H}_2\|\in L^2(M,g)$. Then*

$$\int_M \big(\mathrm{S}^{\,P}_{\mathcal{D}_1,\mathcal{D}_2}+\|\hbar_1\|^2+\|\hbar_2\|^2-\|\mathcal{H}_1\|^2-\|\mathcal{H}_2\|^2-\|\mathcal{T}_1\|^2-\|\mathcal{T}_2\|^2\big)\,\mathrm{d}\,\mathrm{vol}_g=0. \tag{3.4}$$

Proof. Using Proposition 2.2 and (2.6) to (3.3), gives (3.4). □

Next, we present the following.

Lemma 3.7. *For orthogonal distributions $\mathcal{D}_i=P_i(TM)$, $i=1,2$, and $P=P_1+P_2$, the following equalities are valid:*

$$\mathrm{div}_{P_2}\,\mathcal{H}_1=\mathrm{div}_P\,\mathcal{H}_1+\|\mathcal{H}_1\|^2,\quad \mathrm{div}_{P_1}\,\mathcal{H}_2=\mathrm{div}_P\,\mathcal{H}_2+\|\mathcal{H}_2\|^2. \tag{3.5}$$

Proof. Using definition (2.1), we obtain

$$\mathrm{div}_P\,\mathcal{H}_1=\sum\nolimits_a\langle P_2\nabla_{P_2e_a}\mathcal{H}_1,e_a\rangle+\sum\nolimits_j\langle P_1\nabla_{P_1e_j}\mathcal{H}_1,e_j\rangle=\mathrm{div}_{P_2}\,\mathcal{H}_1-\|\mathcal{H}_1\|^2,$$

– the equality $(3.5)_1$ with $\mathrm{div}_{P_2}\,\mathcal{H}_1$. The proof for $\mathrm{div}_{P_1}\,\mathcal{H}_2$ is similar. □

The following theorem generalizes [10, Theorem 2].

Theorem 3.8. *Let a Riemannian manifold (M,g) be equipped with orthogonal distributions $\mathcal{D}_i=P_i(TM)$ $(i=1,2)$. Set $P=P_1+P_2$ and suppose that $\mathcal{D}_1$ is P-harmonic, $\mathcal{D}_2$ is P-integrable and $\mathrm{S}^{\,P}_{\mathcal{D}_1,\mathcal{D}_2}>0$. Then $\mathcal{D}_1$ has no compact leaves.*

Proof. Assume that L is a compact leaf of $\mathcal{D}_1$, that is $\mathcal{D}_1|_L=TL$. Thus, $\mathcal{T}_1|_L=0$, and by (3.5), (3.3), we obtain

$$\mathrm{div}_{P_1}\,\mathcal{H}_2=\mathrm{S}^{\,P}_{\mathcal{D}_1,\mathcal{D}_2}+\|\hbar_1\|^2+\|\hbar_2\|^2. \tag{3.6}$$

Since $\mathcal{H}_2$ belongs to $\mathcal{D}_1$, then the function $\mathrm{div}_{P_1}\mathcal{H}_2$ along L is equal to $\mathrm{div}_L\,\mathcal{H}_2$. Applying the Divergence Theorem to (3.6) on a compact manifold L, gives a contradiction:

$$0=\int_L\mathrm{div}_L\,\mathcal{H}_2\,d\,\mathrm{vol}_L=\int_L(\mathrm{S}^{\,P}_{\mathcal{D}_1,\mathcal{D}_2}+\|\hbar_1\|^2+\|\hbar_2\|^2)\,d\,\mathrm{vol}_L>0.\quad\square$$

The mixed Ricci P-curvature in the N-direction has the form, see (3.2),

$$\mathrm{Ric}^P_{N,N}=\sum\nolimits_s R^P(N,e_s,e_s,N),$$

where $\{e_s\}$ is any local orthonormal frame in M.

The following corollary of Theorem 3.8 generalizes [10, Corollary 4].

Corollary 3.9. Let a line field $\mathcal{D}_2 = P_2(TM)$ spanned by a unit vector field N on a Riemannian manifold (M, g) be orthogonal to a distribution $\mathcal{D}_1 = P_1(TM)$. Set $P = P_1 + P_2$ and suppose that $\mathcal{D}_1$ is P-harmonic. If $\mathrm{Ric}^P_{N,N} > 0$, then $\mathcal{D}_1$ has no compact leaves.

The following corollary of Theorem 3.8 generalizes (1.1).

Corollary 3.10. Let a line field $\mathcal{D}_2 = P_2(TM)$ spanned by a unit vector field N defined on $M \setminus \Sigma$ (outside the singularity set) be orthogonal to a distribution $\mathcal{D}_1 = P_1(TM)$ on a Riemannian manifold (M, g). Set $P = P_1 + P_2$, $\hbar_{1,sc} = \langle \hbar_1, N\rangle$ – the scalar P-second fundamental form of $\mathcal{D}_1$, and assume conditions (2.2) and $\|P\,\nabla_N N + \sigma_1(\hbar_{1,sc})N\| \in L^2(M, g)$. Then the following integral formula is valid:

$$\int_M (2\,\sigma_2(\hbar_{1,sc}) - \mathrm{Ric}^P_{N,N} + \|\mathcal{T}_1\|^2)\,\mathrm{d}\,\mathrm{vol}_g = 0. \tag{3.7}$$

Proof. Note that $\mathcal{D} = \mathcal{D}_1 \oplus \mathcal{D}_2$ on $M \setminus \Sigma$. We find

$$\|\hbar_1\|^2 - \|\mathcal{H}_1\|^2 = -2\,\sigma_2(\hbar_{1,sc}), \quad \|\hbar_2\|^2 = \|\mathcal{H}_2\|^2, \quad \mathcal{T}_2 = 0.$$

Also, $\mathrm{S}^{\,P}_{\mathcal{D}_1,\mathcal{D}_2} = \mathrm{Ric}^P_{N,N}$. Thus, (3.4) simplifies to (3.7). □

Example 3.11. a) Note that $\sigma_2(\hbar_{1,sc}) \ge 0$ for a P-totally umbilical distribution $\mathcal{D}_1$, and $\sigma_2(\hbar_{1,sc}) \le 0$ for a P-harmonic $\mathcal{D}_1$, see Corollary 3.10. Thus, (3.7) implies nonexistence of P-totally umbilical codimension-one distributions $\mathcal{D}_1 \subset \mathcal{D}$ with $\mathrm{Ric}^P < 0$ on a closed manifold. Also, there are no P-harmonic codimension-one foliations tangent to $\mathcal{D}_1 \subset \mathcal{D}$ with $\mathrm{Ric}^P > 0$ on a closed manifold.

b) This example illustrates Theorem 3.6 and generalizes a special case of the Gauss-Bonnet Theorem. Let a distribution $\mathcal{D}$ and its orthogonal complement on a closed Riemannian manifold (M, g) be harmonic, see (2.2). Consider the special case when $\dim \mathcal{D} = 2$, and the Euler characteristic of $\mathcal{D}$ is equal to zero. Thus, $\mathcal{D} = \mathcal{D}_1 \oplus \mathcal{D}_2$, where $\mathcal{D}_i$ are line fields on M spanned by unit orthogonal vector fields N_1 and N_2, respectively. In view of equalities $\|\hbar_i\|^2 = \|\mathcal{H}_i\|^2$ and $\mathcal{T}_i = 0$ for $i = 1, 2$, (3.3) simplifies to $\mathrm{div}_P(\mathcal{H}_1 + \mathcal{H}_2) = K^P(N_1, N_2)$. Thus,

$$\int_M K^P(N_1, N_2)\,\mathrm{d}\,\mathrm{vol}_g = 0.$$

c) An *almost paracontact structure* on a manifold M^{2n+1} generalizes the almost product and the almost contact structures, see [11]. This is defined by a $(1,1)$-tensor field ϕ of rank $2n$, a vector field ξ and a 1-form η, satisfying

$$\phi^2 = \mathrm{id}_{TM} - \eta \otimes \xi, \quad \eta(\xi) = 1.$$

An almost paracontact structure admits a *compatible* Riemannian metric, i.e., $g(\phi X, \phi Y) = -g(X, Y) + \eta(X)\,\eta(Y)$. Set $\mathcal{D} := \phi(TM) = \mathcal{D}^+ \oplus \mathcal{D}^-$, where $\mathcal{D}^\pm$ are ± 1 eigen-distributions of ϕ. The line field $\widetilde{\mathcal{D}} = \ker \phi$ is spanned by ξ. Then $TM = \mathcal{D}^+ \oplus \mathcal{D}^- \oplus \widetilde{\mathcal{D}}$ and

$$P = \phi^2, \quad P^+ = (\phi^2 + \phi)/2, \quad P^- = (\phi^2 - \phi)/2$$

are the orthoprojectors onto $\mathcal{D}$, $\mathcal{D}^+$ and $\mathcal{D}^-$, respectively. The formula (3.3) is fulfilled, but to apply (3.4) the distributions $\mathcal{D}$ and $\widetilde{\mathcal{D}}$ should be harmonic (see Proposition 2.2). If

$$g(X, \phi Y) = d\eta(X, Y),$$

where $d\eta(X,Y) = (1/2)(X\eta(Y) - Y\eta(X) - \eta([X,Y])$, the almost paracontact metric manifold is said to be a *paracontact metric manifold*. In this case,

$$\nabla_X\, \xi = -\varphi\, X - \varphi\, hX, \tag{3.8}$$

see [11, Lemma 2.5], where the symmetric tensor $h = \frac{1}{2}\,\mathcal{L}_\xi\, \varphi$ anti-commutes with ϕ. Thus, $\mathrm{Tr}(\varphi\, h) = 0$. Since ϕ is skew-symmetric, we also have $\mathrm{Tr}\ \varphi = 0$. By the above, $\nabla_\xi\, \xi = 0$, that is $\widetilde{\mathcal{D}}$ is harmonic. Using a local orthonormal frame (e_i) of $\mathcal{D}$ and applying (3.8), we conclude that $\mathcal{D}$ is harmonic:

$$g(\nabla_{e_i} e_i,\, \xi) = g(\nabla_{e_i}\, \xi,\, e_i) = \mathrm{Tr}\ \varphi + \mathrm{Tr}(\varphi\, h) = 0.$$

Thus, (2.2) is fulfilled for a paracontact metric manifold, and we can apply (3.4).

References

[1] K. Andrzejewski, V. Rovenski, P. Walczak, *Integral formulas in foliations theory*, 73–82, in *Geometry and Its Applications*, Springer Proc. in Math. and Statistics, **72**, Springer, 2014.
[2] A. Bejancu, H. Farran, *Foliations and geometric structures*. Springer, 2006.
[3] A. Gray, *Pseudo-Riemannian almost product manifolds and submersions*, J. Math. Mech. **16**, no. 7 (1967), 715–737.
[4] M. Lużyńczyk, P. Walczak, *New integral formulae for two complementary orthogonal distributions on Riemannian manifolds*, Ann. Glob. Anal. Geom. **48** (2015), 195–209.
[5] P. Popescu and V. Rovenski, *An integral formula for singular distributions*, Results in Mathematics **75**, Article number: 18 (2019), DOI: 10.1007/s00025-019-1145-1
[6] G. Reeb, *Sur la courboure moyenne des variétés intégrales d'une équation de Pfaff* $\omega = 0$, C. R. Acad. Sci. Paris **231** (1950), 101–102.
[7] V. Rovenski, *Integral formulae for a Riemannian manifold with two orthogonal distributions*, Central European J. Math. **9**, No. 3 (2011), 558–577.
[8] V. Rovenski, P. Walczak, *Topics in Extrinsic Geometry of Codimension-One Foliations*, Springer Briefs in Mathematics, Springer-Verlag, 2011.
[9] D. Sullivan, *A homological characterization of foliations consisting of minimal surfaces*, Comm. Math. Helv. **54** (1979), 218–223
[10] P.G. Walczak, *An integral formula for a Riemannian manifold with two orthogonal complementary distributions*, Colloq. Math. **58**, No. 2 (1990) 243–252.
[11] S. Zamkovoy, *Canonical connections on paracontact manifolds*, Ann. Global Anal. Geom. **36**, no. 1 (2009), 37–60.

University of Haifa, Israel
vrovenski@univ.haifa.ac.il

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

ON THE 2-SOLITON ASYMPTOTICS FOR THE $d \times d$-MATRIX KORTEWEG-DE VRIES EQUATION

CORNELIA SCHIEBOLD

ABSTRACT. The present article is concerned with the interaction of solitary wave solutions of the matrix Korteweg-de Vries equation. The picture is essentially richer than in the classical scalar case since collisions may be less elastic in the sense that they do not only cause a position shift but also a change of shape.

Our construction of solutions is based on a general solution formula with matrix parameters. After a discussion which parameters yield solutions consisting of one localized wave, an asymptotic description is obtained for the interaction of two such waves, the crucial point being explicit formulas for the change of shape. The main result extends previous work by Goncharenko, who studied waves coming from matrices of rank 1.

As usual, it is to be expected that nonlinear superpositions of finitely many waves behave like combinations of 2-soliton interactions.

KEYWORDS: matrix KdV equation, soliton solutions, asymptotic analysis

MSC2010: 37K40, 35Q53, 35C08

Received 10 May 2021; accepted 19 July 2021

1. INTRODUCTION

For the classical Korteweg-de Vries (KdV) equation $u_t = u_{xxx} + 6uu_x$, the N-soliton solution is well understood: Asymptotically, we see an *ordinary* superposition of N single solitons, ordered according to their (negative) velocities (in $t \approx -\infty$ the fastest appears rightmost, in $t \approx \infty$ leftmost). In between, pairs of solitons collide successively, locally resembling a 2-soliton solution, at least in generic cases.

In terms of the Inverse Scattering Method (ISM), they belong to reflectionless potentials $u_0(x) = u(x,0)$ in the Schrödinger operator $L = -(d/dx)^2 + u_0(x)$, with

https://doi.org/10.2478/9788366675360-020.

N simple discrete eigenvalues $\lambda_1, \ldots, \lambda_N$ (and norming constants $r_1, \ldots, r_N$ of the corresponding eigenfunctions) as scattering data. Note that, asymptotically, the eigenvalue λ_j determines speed and height of the jth soliton, whereas r_j encodes its position for $t \to \pm\infty$ (up to position-shifts due to collisions).

The $\mathsf{d} \times \mathsf{d}$-matrix KdV equation

$$U_t = U_{xxx} + 3\{U, U_x\}, \tag{1.1}$$

was introduced by Lax in [12], see [1, 3, 11] and references therein for background on noncommutative integrable systems. An explicit formula for the N-soliton solution was derived by Goncharenko [10] using ISM. Roughly speaking, the norming constants $r_1, \ldots, r_N$ in the spectral data are replaced by $\mathsf{d}\times\mathsf{d}$-matrices $R_1, \ldots, R_N$, called spectral matrices in [10]. It should be mentioned that this solution formula is formulated in terms of quasi-determinants (as introduced by Gelfand and Retakh [9]). In [6] a rather general solution formula for the $\mathsf{d} \times \mathsf{d}$-matrix KdV equation containing operator parameters is given, which is expressed in terms of determinants on quasi-Banach ideals (in the sense of Pietsch [13]), and it is shown how Goncharenko's N-soliton solutions fit into this picture.

However, for the $\mathsf{d} \times \mathsf{d}$-matrix KdV equation the asymptotics of the N-soliton solution is not as well understood. In [10], the 1-soliton solution ($N = 1$) is described in the case that the spectral matrix $R := R_1$ is diagonisable, see [6] for a discussion of the general case. Furthermore, in [10] the collisions of two solitons ($N = 2$) is discussed in the case that both spectral matrices R_1, R_2 are multiples of one-dimensional projections: A regularity condition is given, and the solution is described asymptotically. In contrast to the classical scalar case, collision does not only lead to position shifts, but also to changes of shape. More precisely, the shape of every soliton is encoded in some corresponding "local" spectral matrix (which need not be a part of the spectral data[1]). In general, an incoming soliton with spectral matrix $R_{-\infty}$ reemerges from collision as a soliton with a new spectral matrix R_∞.

As already mentioned, there are substantially more localized solutions[2] than those appearing in the interaction analysis of the 2-soliton solution in [10]. The goal of the present article is to complete the interaction analysis by admitting spectral matrices of rank larger than 1.

Our point of departure in Section 2 is the class of solutions generated by a solution formula with matrix parameters, see Proposition 2.1. In contrast to the formula in [10] involving quasi determinants, our solution formula is expressed purely in terms of matrix multiplication, resulting in formulas particularly suitable for asymptotic analysis. To make the article self-contained, a direct proof of

[1]The question which parts of the spectral data reflect behaviour in $t \approx -\infty$ or $t \approx \infty$ is slightly more delicate, but will become transparent from detailed formulas later on.

[2]For parallel research on the matrix-valued modified KdV equation, see [4, 7].

Proposition 2.1 is provided in Appendix A. We also explain how the multisoliton solutions in [10] can be reobtained in our framework.

Before investigating interaction, we have to determine which of the 1-soliton solutions provided by Proposition 2.1 are localized. We address this question in Section 3, together with a short discussion of qualitative properties of 1-soliton solutions in the matrix case (for more information we refer to [6]). As a conclusion, we will see that there is only one class of (degenerate) spectral matrices producing non-localised 1-soliton solutions, namely matrices whose Jordan form contains a block with eigenvalue 0 and size at least 2×2.

We then proceed to a complete study of the asymptotics for 2-soliton solutions corresponding to non-degenerate spectral matrices in Section 4. Our main result is Theorem 4.4, and its addendum in Remark 4.8, which give explicit expressions for the change caused by interaction, both in position and shape. In Corollary 4.7 we observe a striking consequence, which seems to call for a conceptual explanation: If a 1-soliton collides with another 1-soliton coming from a *regular* spectral matrix, then it only suffers a position shift, like solitons in the scalar setting. In Section 5, we illustrate this observation by a worked example with MATHEMATICA plots.

Acknowledgement. I would like to thank Sandra Carillo and Mauro Lo Schiavo for a long and fruitful collaboration on noncommutative integrable systems, which has had a necessary influence on the work presented in the article.

2. AN EXPLICIT FORMULA FOR THE N-SOLITON SOLUTIONS OF THE $\mathsf{d} \times \mathsf{d}$-MATRIX KDV

We start with stating the formula for the N-soliton solutions on which our analysis relies. For a proof of the following proposition, we refer to Appendix A, see also [15, 16, 17] for earlier work on matrix solution formulas in a more general context. Below, I_{d} denotes the unit matrix of size $\mathsf{d} \times \mathsf{d}$.

Proposition 2.1. *Let $k_1, \dots, k_N$ be complex numbers such that $k_i + k_j \neq 0$ for all i, j, and let $C_1, \dots C_N$ be arbitrary $\mathsf{d} \times \mathsf{d}$-matrices.*

Define the $N\mathsf{d} \times N\mathsf{d}$-matrix function $L = L(x,t)$ as the block matrix $L = \big(L_{ij}\big)_{i,j=1}^{N}$ with the $\mathsf{d} \times \mathsf{d}$-blocks

$$L_{ij} = \frac{\ell_j}{k_i + k_j} C_j,$$

where $\ell_j = \ell_j(x,t) = \exp(k_j x + k_j^3 t)$.

Then

$$U(x,t) = \frac{\partial}{\partial x}\left(\Big(\ell_1 C_1 \dots \ell_N C_N\Big) \left(I_{N\mathsf{d}} + L\right)^{-1} \begin{pmatrix} I_{\mathsf{d}} \\ \vdots \\ I_{\mathsf{d}} \end{pmatrix} \right) \tag{2.1}$$

solves the $\mathsf{d} \times \mathsf{d}$-matrix KdV equation (1.1) on the open set where $\det(I_{N\mathsf{d}} + L) \neq 0$.

In the remainder of the section we explain how the N-soliton solutions in [10] can be realised in terms of Proposition 2.1. Part of the argument is taken from [6, pp. 16–17].

Let us first summarise the construction of the N-soliton solution in [10], in the framework of ISM. Note that we apply the transformation $t \to -t$, $U \to -U$ to adjust to the slightly different form of the matrix KdV considered in [10].

The starting point is the scattering data

$$\{\, \mathrm{i}\lambda_1, ..., \mathrm{i}\lambda_N,\ R_1, ..., R_N,\ R(k)\,\}$$

in the reflectionless case, i.e. $R(k) \equiv 0$. Here $\lambda_N > \ldots > \lambda_1 > 0$ and $R_1, \ldots, R_N$ are $\mathsf{d} \times \mathsf{d}$-matrices, called *spectral matrices* in [10].

The first step is to find a solution $K = K(x, y)$ of the matrix Gelfand-Levitan-Marchenko (GLM) equation

$$K(x, y) + H(x + y) + \int_x^{\infty} K(x, z)H(y + z)\, dz = 0 \tag{2.2}$$

with kernel $H(\xi) = \sum_{j=1}^{N} R_j e^{-\lambda_j \xi}$.

Next the t-dependence is incorporated into this solution by adding it to the spectral matrices according to $R_j(t) = R_j e^{-8\lambda_j^3 t}$.

Finally, the N-soliton solution is given by

$$U_{\text{sol}}(x, t) = 2\frac{\partial}{\partial x} K(x, x; t).$$

We now give the details of the above construction, following [10]. In contrast to the presentation in [10], we here use a formulation in terms of matrix multiplication. Besides leading to compact formulas, this will facilitate comparison with our proposition considerably. In fact, the sketch below is precisely what one gets lifting the standard argument for the construction of the *scalar* 1-soliton solution via the GLM equation to the non-commutative level.

Rewriting

$$H(\xi) = R\, e^{-\xi\Lambda}\, J,$$

where Λ is the block diagonal matrix of size $N\mathsf{d} \times N\mathsf{d}$ with the blocks $\lambda_j I_{\mathsf{d}}$, $j = 1, \ldots, N$, on the diagonal, and R the $\mathsf{d} \times N\mathsf{d}$-matrix, J the $N\mathsf{d} \times \mathsf{d}$-matrix given by

$$R = \Big(R_1 \ldots R_N\Big), \qquad J = \begin{pmatrix} I_{\mathsf{d}} \\ \vdots \\ I_{\mathsf{d}} \end{pmatrix},$$

the ansatz to solve (2.2) used in [10] translates to

$$K(x, y) = R\, \widetilde{K}(x) e^{-y\Lambda}\, J.$$

Inserting this into (2.2), one gets

$$R\left(\widetilde{K}(x)+e^{-x\Lambda}+\widetilde{K}(x)\int_x^\infty e^{-z\Lambda}\,JR\,e^{-z\Lambda}\,dz\right)e^{-y\Lambda}\,J=0,$$

for which a solution can be obtained by setting the expression inside the parentheses equal to 0. This yields

$$\widetilde{K}(x)=-e^{-x\Lambda}\Big(I+\int_x^\infty e^{-z\Lambda}\,JR\,e^{-z\Lambda}\,dz\Big)^{-1}.$$

In order to evaluate the integral explicitly, we look for a primitive function for the integrand. Consider the matrix function $f(z)=e^{-z\Lambda}De^{-z\Lambda}$ as a candidate. By the non-commutative product rule, $f'(z)=-\Lambda e^{-z\Lambda}De^{-z\Lambda}-e^{-z\Lambda}De^{-z\Lambda}\Lambda=-e^{-z\Lambda}(\Lambda D+D\Lambda)e^{-z\Lambda}$. Hence, $f(z)$ is primitive function provided that D is a solution of the Sylvester equation $\Lambda D+D\Lambda=-JR$. In fact, it is straightforward to compute

$$D=-\Big(\frac{1}{\lambda_i+\lambda_j}R_j\Big)_{i,j=1}^N.$$

As a result,

$$\widetilde{K}(x)=-e^{-x\Lambda}\Big(I-e^{-x\Lambda}\,D\,e^{-x\Lambda}\Big)^{-1},$$

and, incorporating now also the t-dependence, using $R(t)=R\,e^{-8t\Lambda^3}$ and $D(t)=D\,e^{-8t\Lambda^3}$, we get

$$\begin{aligned}K(x,y;t)&=-R\,e^{-x\Lambda-8t\Lambda^3}\Big(I-e^{-x\Lambda}\,D\,e^{-x\Lambda-8t\Lambda^3}\Big)^{-1}e^{-y\Lambda}\,J\\&=-Re^{-(x+y)\Lambda-8t\Lambda^3}\Big(I-e^{(y-x)\Lambda}\,D\,e^{-(x+y)\Lambda-8t\Lambda^3}\Big)^{-1}J.\end{aligned}$$

In summary,

$$U_{\text{sol}}(x,t)=\frac{\partial}{\partial x}\left(-2Re^{-2x\Lambda-8t\Lambda^3}\Big(I-D\,e^{-2x\Lambda-8t\Lambda^3}\Big)^{-1}J\right). \tag{2.3}$$

To conclude the argument we observe that (2.3) can be realised in Proposition 2.1 with the parameter settings

$$k_j=-2\lambda_j,\qquad C_j=-2R_j, \tag{2.4}$$

establishing the translation to the spectral data in the framework of ISM.

3. 1-soliton solutions

In the case $N=1$, the solution formula in Proposition 2.1, omitting the superfluous index, reads

$$\begin{aligned}U(x,t)&=\frac{\partial}{\partial x}\Big(\ell C(I_{\mathsf{d}}+\frac{1}{2k}\ell C)^{-1}I_{\mathsf{d}}\Big)=2k\frac{\partial}{\partial x}\Big(\ell D\Big((I_{\mathsf{d}}+\ell D)^{-1}\Big)\Big)\\&=2k^2\ell D(I_{\mathsf{d}}+\ell D)^{-2},\end{aligned} \tag{3.1}$$

where $\ell(x,t)=:\exp(kx+k^3t)$ and $D=(1/(2k))\,C$.

Lemma 3.1. *The solution $U(x,t)$ is real for k, D real.*

For the KdV equation, one usually is interested in *real* solutions. In the sequel we assume the parameters k, D to be real. This is a reasonable choice according to the previous corollary. Furthermore we can restrict to D in Jordan form:

Lemma 3.2. *Up to a (real) coordinate transformation, D can be chosen in real Jordan form.*

Note that a consequence of the above lemma is that we can discuss the (real) Jordan blocks of D independently.

Remark 3.3. In [6] a reduction of (3.1) is given using gauge transformations.

Real eigenvalues with 1×1-Jordan blocks. For a 1×1-Jordan block of D with real eigenvalue d, the corresponding entry $u(x,t)$ of $U(x,t)$ is the classical, scalar 1-soliton solution of the KdV equation,

$$u(x,t) = \begin{cases} \frac{k^2}{2}\cosh^{-2}\Big(\frac{1}{2}\big(kx+k^3t+\log|d|\big)\Big), & d>0,\\ 0, & d=0,\\ -\frac{k^2}{2}\sinh^{-2}\Big(\frac{1}{2}\big(kx+k^3t)+\log|d|\big)\Big), & d<0, \end{cases}, \tag{3.2}$$

which is regular iff $d \geq 0$. Hence, several 1×1-Jordan blocks with eigenvalues $d_j>0$, lead to 1-soliton solutions $u_j(x,t)$ on the diagonal of $U(x,t)$, the velocity and height characterised by the same eigenvalue k. They only differ in their initial positions, which are specified by d_j.

Real eigenvalues of higher multiplicity. For a general Jordan block of D with real eigenvalue d, the corresponding block of the solution $U(x,t)$ has band structure. Explicit formulas for the entries $u_j(x,t)$ on the jth off-diagonal ($j\geq 0$) are derived in [6]. In the prototypical case $\mathsf{d}=2$, one can use

Lemma 3.4. *The 1-soliton solution (3.1) of the 2×2-matrix KdV can be written as*

$$U = 2k^2\ell\,\frac{1}{\Delta^2}\Big(D + 2\det(D)\,\ell\, I_2 + \det(D)\,\ell^2\,\widehat{D}\,\Big)$$

with $\Delta = 1+\operatorname{tr}(D)\,\ell + \det(D)\,\ell^2$.

Proof. Let

$$D = \begin{pmatrix} d_{11} & d_{12}\\ d_{21} & d_{22}\end{pmatrix}, \qquad \widehat{D} = \begin{pmatrix} d_{22} & -d_{12}\\ -d_{21} & d_{11}\end{pmatrix}$$

such that $D\widehat{D} = \det(D)I_2$ and $D+\widehat{D} = \operatorname{tr}(D)I_2$. Then $(I_2+\ell D)^{-1} = \frac{1}{\Delta}(I_2+\ell\widehat{D})$ where $\Delta = \det(I_2+\ell D)$. Expansion of the determinant shows $\Delta = 1+\ell\operatorname{tr}(D) + \ell^2\det(D)$. □

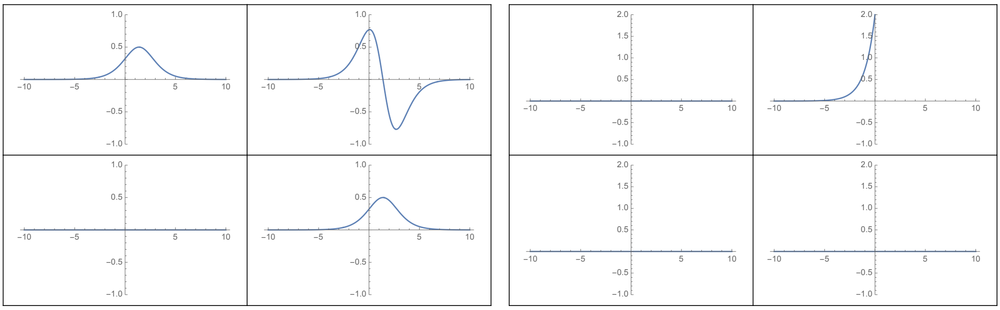


FIGURE 1. The 1-soliton solution in the case that D is a 2×2 Jordan block with eigenvalue d. In both cases $k = 1$. For $d = 1/4$ the wave is localized (to the left), for $d = 0$ not (to the right). The plots show snapshots for $t = 0$.

For a 2×2-Jordan block D, we have $\Delta = (1+d\ell)^2$, the 1-soliton solution $U(x,t)$ being regular iff $d \geq 0$. In addition, we see that $U(x,t)$ is an upper band matrix with the elements

$$u = 2k^2 d\ell/(1+d\ell)^2, \qquad v = 2k^2\ell(1-d\ell)/(1+d\ell)^3$$

on the diagonal respective the first off-diagonal. Hence, for $d > 0$ the diagonal element is the classical (localized) soliton as described in (3.2) and, due to the Jordan structure of D, we have $v = \partial u/\partial d$.

Remark 3.5. Note in particular that the solution for $d = 0$ is *not* localized, see Figure 1 for an illustration. This phenomenon continues to be valid for larger Jordan blocks with eigenvalue 0, as shown in [6]. We will therefore exclude this case from the asymptotic analysis in the next section.

Complex conjugated eigenvalues. To give an impression of this case, we will elaborate the details for $\mathsf{d} = 2$. Applying Lemma 3.4 for D with a pair of complex conjugated eigenvalues $d_1+\mathrm{i}d_2$, $d_1-\mathrm{i}d_2$, where $d_2 \neq 0$, shows $\Delta = (1+d_1\ell)^2+(d_2\ell)^2$ implying regularity of the 1-soliton solution $U(x,t)$. We use

$$D = \begin{pmatrix} d_1 & d_2 \\ -d_2 & d_1 \end{pmatrix}, \; T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix} \implies T^{-1}DT = \begin{pmatrix} d_1+\mathrm{i}d_2 & 0 \\ 0 & d_1-\mathrm{i}d_2 \end{pmatrix}.$$

Similarly, $T^{-1}U(x,t)T$ is diagonal with $u(x,t) = 2k^2\big(d_1+\mathrm{i}d_2)\ell/(1+(d_1+\mathrm{i}d_2)\ell\big)^2$ and its complex conjugate on the diagonal. Transforming back,

$$U = \begin{pmatrix} \mathrm{Re}(u) & \mathrm{Im}(u) \\ -\mathrm{Im}(u) & \mathrm{Re}(u) \end{pmatrix},$$

see Figure 2 for an illustration.

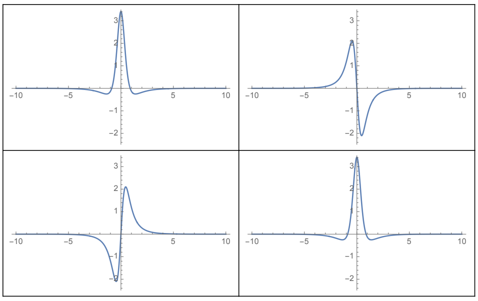


FIGURE 2. The 1-soliton solution in the case that the 2×2-spectral matrix D is the real Jordan block corresponding to the complex conjugated eigenvalues $d = (-1 \pm \mathrm{i})/\sqrt{2}$. Here $k = 1$. The plot shows a snapshot for $t = 0$.

4. Asymptotics of 2-soliton solutions

In this section we study the asymptotics of the 2-soliton solution for the $\mathsf{d}\times\mathsf{d}$-matrix KdV equation in Proposition 2.1,

$$(4.1)\quad U(x,t) = \frac{\partial}{\partial x}\left(\begin{pmatrix} 2k_1\ell_1 D_1 & 2k_2\ell_2 D_2 \end{pmatrix} \begin{pmatrix} I_{\mathsf{d}} + \ell_1 D_1 & \mu_2\ell_2 D_2 \\ \mu_1\ell_1 D_1 & I_{\mathsf{d}} + \ell_2 D_2 \end{pmatrix}^{-1} \begin{pmatrix} I_{\mathsf{d}} \\ I_{\mathsf{d}} \end{pmatrix} \right),$$

describing the collision of the two 1-soliton solutions

$$(4.2)\qquad U_j(x,t) = \frac{\partial}{\partial x}\Big(2k_j\ell_j D_j (I_{\mathsf{d}} + \ell_j D_j)^{-1} \Big),$$

where $D_j = (1/(2k_j))C_j$ and $\mu_j = 2k_j/(k_1+k_2)$, $j = 1, 2$.

To start with, we recall the basic assumption for our asymptotic analysis.

Assumption 4.1. *The Jordan forms of the spectral matrices D_1, D_2 do not contain Jordan blocks with eigenvalue 0 and of size at least 2×2.*

In the sequel, we will focus on what happens to $U_1(x,t)$, and therefore look at the one-parameter family of parallel straight lines $\mathcal{L}_c : k_1 x + k_1^3 t = c$ in (x,t)-space. This will lead to a partially formal discussion, which can be extended to statements on uniform convergence by using techniques from [14, 18], even in the presence of poles.

Assumption 4.2. *Let $0 < k_1 < k_2$. The other cases can be treated similarly.*

Since the arguments for $U_2(x,t)$ are symmetric, we will content ourselves with providing the outcome at the end of the section. It should be stressed that this symmetry includes time reversal. Whereas the soliton $U_1(x,t)$ appears (asymptotically) undisturbed for $t \approx -\infty$, its companion $U_2(x,t)$ appears undisturbed for $t \approx \infty$.

Part 1: **Asymptotic behaviour for $t \to -\infty$.** The first result is that, for $t \to -\infty$, the asymptotic behaviour of $U(x,t)$ along the straight lines $\mathcal{L}_c$ does not depend on $U_2(x,t)$.

Theorem 4.3. *For $t \to -\infty$, the soliton $U_1(x,t)$ appears undisturbed.*

Proof. For $t \to -\infty$, we have $k_2 x + k_2^3 t = k_2(x + k_1^2 t) + k_2(k_2^2 - k_1^2)t = (k_2/k_1)c + k_2(k_2^2 - k_1^2)t \to -\infty$, which means that $\ell_2 \to 0$. Hence,

$$U(x,t) \approx \frac{\partial}{\partial x}\left(\begin{pmatrix} 2k_1\ell_1 D_1 & 0 \end{pmatrix}\begin{pmatrix} I_{\mathsf{d}} + \ell_1 D_1 & 0 \\ \mu_1\ell_1 D_1 & I_{\mathsf{d}} \end{pmatrix}^{-1}\begin{pmatrix} I_{\mathsf{d}} \\ I_{\mathsf{d}} \end{pmatrix}\right).$$

We will determine the asymptotics along lines $\mathcal{L}_c$ on which the (constant) matrix $I_{\mathsf{d}} + \ell_1 D_1$ is invertible (this holds for c generic). In this case we have

$$\begin{pmatrix} I_{\mathsf{d}} + \ell_1 D_1 & 0 \\ \mu_1\ell_1 D_1 & I_{\mathsf{d}} \end{pmatrix}^{-1} = \begin{pmatrix} (I_{\mathsf{d}} + \ell_1 D_1)^{-1} & 0 \\ -\mu_1\ell_1 D_1 (I_{\mathsf{d}} + \ell_1 D_1)^{-1} & I_{\mathsf{d}} \end{pmatrix},$$

showing

$$\begin{aligned} U(x,t) &\approx \frac{\partial}{\partial x}\left(\begin{pmatrix} 2k_1\ell_1 D_1 & 0 \end{pmatrix}\begin{pmatrix} (I_{\mathsf{d}} + \ell_1 D_1)^{-1} & 0 \\ -\mu_1\ell_1 D_1 (I_{\mathsf{d}} + \ell_1 D_1)^{-1} & I_{\mathsf{d}} \end{pmatrix}\begin{pmatrix} I_{\mathsf{d}} \\ I_{\mathsf{d}} \end{pmatrix}\right) \\ &= \frac{\partial}{\partial x}\Big(2k_1\ell_1 D_1 (I_{\mathsf{d}} + \ell_1 D_1)^{-1}\Big), \end{aligned}$$

and the result follows. $\square$

Part 2: **Asymptotic behaviour for $t \to \infty$.** Secondly we investigate the asymptotic behaviour of $U(x,t)$ for $t \to \infty$ along the straight lines $\mathcal{L}_c$. Our main result in Theorem 4.4 describes how the soliton $U_1(x,t)$ changes its spectral matrix as an effect of the collision with $U_2(x,t)$. The resulting spectral matrix $\widetilde{D}_1$ will depend on the spectral matrix D_2 of $U_2(x,t)$.

To give an explicit formula for $\widetilde{D}_1$, we observe that D_2 can be transformed to real Jordan form such that the blocks corresponding to the non-zero eigenvalues are positioned in the upper left corner. More precisely, there is a transformation matrix τ such that

$$\tau^{-1} D_2 \tau = \begin{pmatrix} J & 0 \\ 0 & 0 \end{pmatrix},$$

where J is regular and we have the zero-matrix in the lower right corner because of Assumption 4.1. For the subsequent analysis, we write

$$\tau^{-1} D_2 \tau = DP, \text{ where } D = \begin{pmatrix} J & 0 \\ 0 & I \end{pmatrix}, P = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.3}$$

Note that D is invertible and P is a projection. Moreover D and P commute.

Theorem 4.4. *For $t \to \infty$, we have*

$$U(x,t) \approx \frac{\partial}{\partial x}\Big(2k_1\ell_1 \widetilde{D}_1 \big(I_{\mathsf{d}} + \ell_1 \widetilde{D}_1\big)^{-1}\Big)$$

with $\widetilde{D}_1 = (I_{\mathsf{d}} - \mu_2 \tau P \tau^{-1}) D_1 (I_{\mathsf{d}} - \mu_2 \tau P \tau^{-1})$.

In preparation for our analysis, we explicitly evaluate the formula for 2-soliton solution (4.1). We will use that $I_{\mathsf{d}} + \ell_2 D_2$ is invertible along any line $\mathcal{L}_c$ provided t is large enough. This holds since we have $\ell_2 \to \infty$ on $\mathcal{L}_c$ for $t \to \infty$.

Proposition 4.5. *For the 2-soliton solution* (4.1)*, we have*

$$U(x,t) = U_2(x,t) + \widetilde{U}_1(x,t)$$

with

$$\widetilde{U}_1(x,t) = \frac{\partial}{\partial x}\Big(2k_1(I_{\mathsf{d}}-\mu_2 S_2)\ell_1 D_1(I_{\mathsf{d}}-\mu_2 S_2)\Big(I_{\mathsf{d}}+(I_{\mathsf{d}}-\mu_2 S_2)\ell_1 D_1(I_{\mathsf{d}}-\mu_2 S_2)\Big)^{-1}\Big),$$

where $S_2 = \ell_2 D_2 (I_{\mathsf{d}} + \ell_2 D_2)^{-1}$.

Note that $U_2(x,t) = 2k_2(\partial S_2/\partial x)$, see (4.2).

Proof. For short, we write $B_1 = \ell_1 D_1$, $B_2 = \ell_2 D_2$. In particular, we have $S_2 = B_2(I_{\mathsf{d}} + B_2)^{-1}$. To start with, we observe that

$$S = \begin{pmatrix} I_{\mathsf{d}} + B_1 & \mu_2 B_2 \\ \mu_1 B_1 & I_{\mathsf{d}} + B_2 \end{pmatrix}$$

is invertible iff $T = I_{\mathsf{d}} + B_1 - \mu_1\mu_2 B_2(I_{\mathsf{d}} + B_2)^{-1}B_1$ is invertible, and

$$S^{-1} = \begin{pmatrix} I_{\mathsf{d}} & 0 \\ 0 & (I_{\mathsf{d}} + B_2)^{-1} \end{pmatrix} \begin{pmatrix} T^{-1} & -\mu_2 T^{-1} S_2 \\ -\mu_1 B_1 T^{-1} & I_{\mathsf{d}} + \mu_1\mu_2 B_1 T^{-1} S_2 \end{pmatrix}.$$

This can be verified by a straightforward but lengthy computation.

Therefore,

$$\begin{aligned}
&\begin{pmatrix} 2k_1 B_1 & 2k_2 B_2 \end{pmatrix} S^{-1} \begin{pmatrix} I_{\mathsf{d}} \\ I_{\mathsf{d}} \end{pmatrix} \\
&= \begin{pmatrix} 2k_1 B_1 & 2k_2 B_2 (I_{\mathsf{d}} + B_2)^{-1} \end{pmatrix} \begin{pmatrix} T^{-1}(I_{\mathsf{d}} - \mu_2 S_2) \\ I_{\mathsf{d}} - \mu_1 B_1 T^{-1}(I_{\mathsf{d}} - \mu_2 S_2) \end{pmatrix} \\
&= 2k_2 B_2 (I_{\mathsf{d}} + B_2)^{-1} + \begin{pmatrix} 2k_1 I_{\mathsf{d}} - 2k_2\mu_1 B_2 (I_{\mathsf{d}} + B_2)^{-1} \end{pmatrix} B_1 T^{-1}(I_{\mathsf{d}} - \mu_2 S_2) \\
&= 2k_2 B_2 (I_{\mathsf{d}} + B_2)^{-1} + 2k_1 \begin{pmatrix} I_{\mathsf{d}} - \mu_2 B_2 (I_{\mathsf{d}} + B_2)^{-1} \end{pmatrix} B_1 T^{-1}(I_{\mathsf{d}} - \mu_2 S_2) \\
&= 2k_2 S_2 + 2k_1 (I_{\mathsf{d}} - \mu_2 S_2) B_1 T^{-1} (I_{\mathsf{d}} - \mu_2 S_2),
\end{aligned}$$

where we have used $k_2\mu_1 = k_1\mu_2$. Note finally that $T = I_{\mathsf{d}} + (I_{\mathsf{d}} - \mu_1\mu_2 S_2)B_1$ and, since $\mu_2 - 1 = 1 - \mu_1$, that

$$I_{\mathsf{d}} - \mu_1\mu_2 P = I_{\mathsf{d}} + \mu_2(\mu_2 - 2)P = (I_{\mathsf{d}} - \mu_2 P)^2.$$

So far we have shown

$$U(x,t) = \frac{\partial}{\partial x}\Big(2k_2 S_2 \ + \ 2k_1(I_{\mathsf{d}} - \mu_2 S_2)B_1\Big(I_{\mathsf{d}} + (I_{\mathsf{d}} - \mu_2 S_2)^2 B_1\Big)^{-1}(I_{\mathsf{d}} - \mu_2 S_2)\Big).$$

Finally

$$\Big(I_{\mathsf{d}} + (I_{\mathsf{d}} - \mu_2 S_2)^2 B_1\Big)(I_{\mathsf{d}} - \mu_2 S_2)$$

$$= \quad (I_{\mathsf{d}} - \mu_2 S_2)\Big(I_{\mathsf{d}} + (I_{\mathsf{d}} - \mu_2 S_2)B_1(I_{\mathsf{d}} - \mu_2 S_2)\Big)$$

implies

$$\Big(I_{\mathsf{d}} + (I_{\mathsf{d}} - \mu_2 S_2)^2 B_1\Big)^{-1}(I_{\mathsf{d}} - \mu_2 S_2)$$
$$= \quad (I_{\mathsf{d}} - \mu_2 S_2)\Big(I_{\mathsf{d}} + (I_{\mathsf{d}} - \mu_2 S_2)B_1(I_{\mathsf{d}} - \mu_2 S_2)\Big)^{-1},$$

which yields the formula claimed in the proposition. □

The above proposition tells that it is crucial to determine the limit of $S_2 = \ell_2 D_2(I_{\mathsf{d}} + \ell_2 D_2)^{-1}$ as $\ell_2 \to \infty$. To this end we use

Lemma 4.6. *Let P be a projection and D a matrix commuting with P.*

a) *If $I_{\mathsf{d}} + D$ is invertible, then $I_{\mathsf{d}} + DP$ is invertible with $(I_{\mathsf{d}} + DP)^{-1} = I_{\mathsf{d}} - (I_{\mathsf{d}} + D)^{-1}DP$.*

b) *Assume in addition that D is invertible. Then,*

$$cDP(I_{\mathsf{d}} + cDP)^{-1} \to P$$

for $c \to \infty$.

Proof. a) follows by direct verification. As a consequence, we get

$$\begin{aligned} cDP(I_{\mathsf{d}} + cDP)^{-1} &= cDP\ \big(\ I_{\mathsf{d}} - (I_{\mathsf{d}} + cD)^{-1}cDP\ \big) \\ &= cD\Big(P - (I_{\mathsf{d}} + cD)^{-1}cDP^2\Big) \ = \ cD\Big(I_{\mathsf{d}} - (I_{\mathsf{d}} + cD)^{-1}cD\Big)P \\ &= cD(I_{\mathsf{d}} + cD)^{-1}P = \Big(I_{\mathsf{d}} + \frac{1}{c}D^{-1}\Big)^{-1}P \ \stackrel{c\to\infty}{\longrightarrow}\ P \end{aligned}$$

which shows b). □

Proof of Theorem 4.4. Starting from the formula for the 2-soliton solution in Proposition 4.5, we can use Lemma 4.6 to approximately replace S_2 by P. Observe that the summand $2k_2P$ is constant and vanishes after derivation with respect to x. Finally, after applying the coordinate transformation τ, we arrive at the claimed result. □

It is worth mentioning two special cases of the above result (which do not include the transformation τ in the asymptotic formulas).

Corollary 4.7. *The following asymptotics holds for $t \to \infty$:*

a) *If D_2 is regular, then*

$$U(x,t) \approx \frac{\partial}{\partial x}\Big(2k_1(\Phi^2\ell_1)D_1\Big(I_{\mathsf{d}} + (\Phi^2\ell_1)D_1\Big)^{-1}\Big)$$

with $\Phi = \dfrac{k_1 - k_2}{k_1 + k_1}$.

b) If $D_2 \neq 0$ is a multiple of a projection P, then

$$U(x,t) \approx \frac{\partial}{\partial x}\Big(2k_1\ell_1\widehat{D}_1\Big(I_{\mathsf{d}} + \ell_1\widehat{D}_1\Big)^{-1}\Big)$$

with $\widehat{D}_1 = (I_{\mathsf{d}} - \mu_2 P)D_1(I_{\mathsf{d}} - \mu_2 P)$.

In the case that D_2 is regular, the spectral matrix of $U_1(x,t)$ appears undisturbed up to the *scalar* phase-shift factor Φ^2. Hence the appearance of the soliton $U_1(x,t)$ is not changed but its path experiences a position shift. In fact, Φ^2 is precisely the well-known phase-shift factor appearing in the 2-soliton solution formula for the scalar KdV equation.

Observe also that in the case that D_2 is a multiple of a projection P, the spectral matrix of $U_1(x,t)$ is given by an explicit formula in terms of D_1 and D_2.

Proof. a) For D_2 regular, we choose $T = P = I_{\mathsf{d}}$, $D = D_2$ in (4.3), with which we get $\widetilde{D}_1 = (1-\mu_2)^2 D_1 = \Phi^2 D_1$.

b) For $D_2 = cP$ with $c \neq 0$ and P a projection, we choose $T = I_{\mathsf{d}}$, $P = D_2$, $D = cI_{\mathsf{d}}$. □

Remark 4.8. As already mentioned, the situation for the soliton $U_2(x,t)$ is symmetric (up to time reversal). Hence, along straight lines $k_2x + k_2^3t = c$ we get with analogous arguments as above

a) For $t \to -\infty$: $U(x,t) \approx U_{2,-\infty}(x,t)$, where $U_{2,-\infty}(x,t)$ is the 1-soliton solution with spectral matrix

$$\widetilde{D}_2 = (I_{\mathsf{d}} - \mu_1\tau P\tau^{-1})D_2(I_{\mathsf{d}} - \mu_1\tau P\tau^{-1})$$

and τ such that $\tau^{-1}D_1\tau = DP$ with P, D commuting, P a projection, and D invertible.

b) For $t \to \infty$: $U(x,t) \approx U_2(x,t)$.

5. Illustration for the 2×2-matrix KdV

In the case of the 2×2-matrix KdV, the two particular cases covered by Corollary 4.7 give a rather complete picture. In the sequel we give an example for a 2-soliton solution with one regular spectral matrix and the other one a multiple of a projection.

For $k_1 = 1$, $k_2 = 2$, and

$$D_1 = \begin{pmatrix} 10 & 0 \\ 0 & \frac{1}{100} \end{pmatrix}, \quad D_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

the resulting 2-soliton solution $U(x,t)$ is depicted in Figure 4. Let us first focus on the asymptotics of the first soliton, along the path $x + t = c$. According to our analysis we get

- For $t \to -\infty$, the soliton $U_1(x,t)$ appears undisturbed. Observe that U_1 is diagonal, and that both diagonal entries are 1-solitons of height 1/2 and velocity -1. However, for t frozen, they are positioned on different places on the x-axis.
- For $t \to \infty$, the spectral matrix of $U_1(x,t)$ has changed due to collision with the other soliton to

$$D_{1,\infty} = \left(I_2 - \frac{4}{3}P \right) D_1 \left(I_2 - \frac{4}{3}P \right),$$

where $P = \frac{1}{2}D_2$.

To geometrically understand the emerging 1-soliton solution $U_{1,\infty}$, note first that $D_{1,\infty}$ is diagonisable. After the corresponding coordinate change $U_{1,\infty}$ becomes diagonal. Changing coordinates back, as matrix entries of $U_{1,\infty}$ appear *linear* superpositions of the diagonal entries, i.e. scalar 1-soliton solutions with the same velocity, but with differently located maxima. To facilitate comparison, we provide plots for $U(x,t)$ and for the 1-soliton solution $U_{1,\infty}(x,t)$ generated from the parameters k_1, $D_{1,\infty}$ in Figure 3.

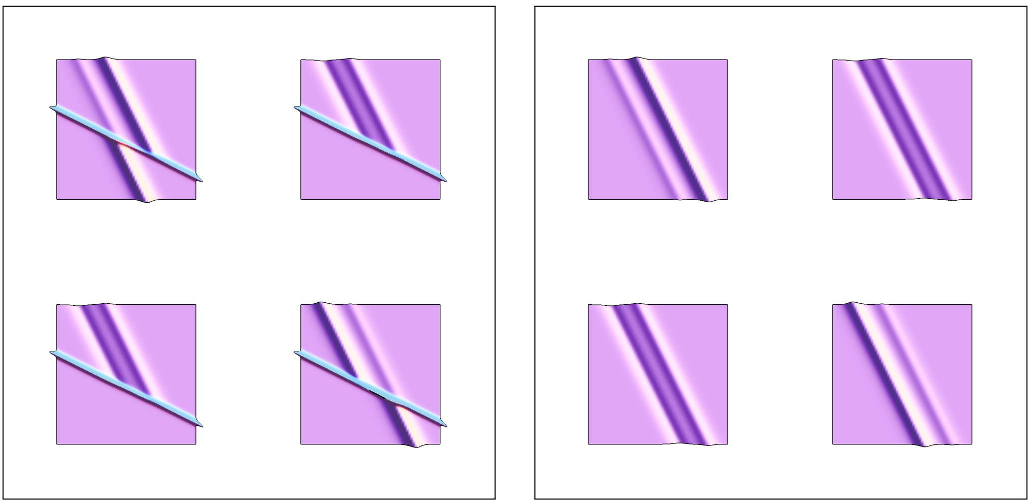


FIGURE 3. Top views of the 2-soliton solution from Section 5 (to the left) and the asymptotic 1-soliton solution $U_{1,\infty}$ appearing along the paths $k_1x + k_1^3t = c$ for $t \to \infty$ (to the right).

For the second soliton, a similar argument shows that

- For $t \to -\infty$, the soliton $U_2(x,t)$ appears undisturbed up to a position-shift, which is given by the same formula as in the scalar case.
- For $t \to \infty$, the soliton $U_2(x,t)$ appears undisturbed.

Note that the spectral matrix of U_2 is diagonisable with 0 being one of the eigenvalues. Hence, by a similar argument as above, the entries of U_2 are multiples of

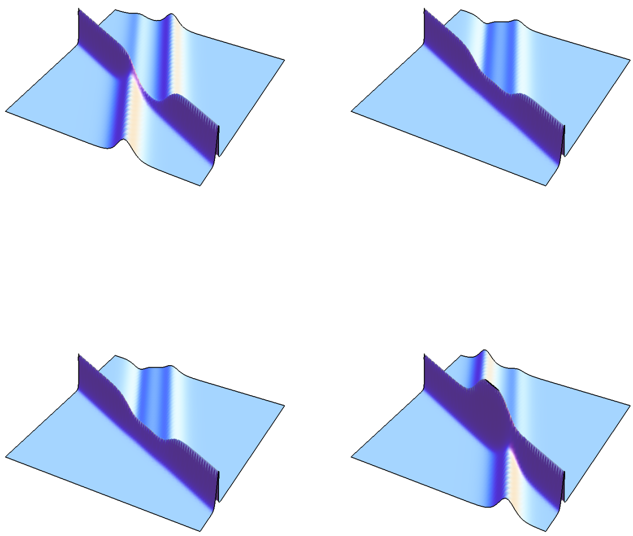


FIGURE 4. The 2-soliton solution discussed in Section 5, where the entries are depicted for $-20 \le x \le 20$, $-8 \le t \le 12$ and the range is $[-0.3, 1.7]$.

one and the same scalar 1-soliton solution. In fact, an explicit computation shows that the entries of U_2 are identical.

APPENDIX A. PROOF OF PROPOSITION 2.1.

We start from an explicit solution formula for the non-commutative KdV (1.1) viewed as an equation with the unknown $U = U(x,t)$ taking its values in the (bounded) operators on some Banach space E. The following result from [8] is essentially proven in [2], see also [5] for a generalisation to the non-commutative KdV hierarchy.

Proposition A.1. ([8, Theorem 7.2]). *Given a bounded operator A on E, and an operator-function $L = L(x,t)$ satisfying the linear differential equations $L_x = AL$ and $L_t = A^3L$. Then*

$$U = \frac{\partial}{\partial x}\Big((I_E + L)^{-1}(AL + LA)\Big)$$

solves the non-commutative KdV equation (1.1) on any open set where $I_E + L$ is invertible.

In the sequel the differential equations imposed on the operator-function L will be met by choosing

$$L(x,t) = e^{Ax+A^3t}B,$$

where B can be any bounded operator on E.

From now on we work under the assumptions of Proposition 2.1. We specialise Proposition A.1 to the finite-dimensional case $E = \mathbb{C}^{N\mathsf{d}}$, and choose the parameters

$$A = \begin{pmatrix} k_1 I_{\mathsf{d}} & & 0 \\ & \ddots & \\ 0 & & k_N I_{\mathsf{d}} \end{pmatrix}, \qquad B = \Big(\frac{1}{k_i + k_j} C_i^T \Big)_{i,j=1}^N.$$

Here T denotes transposition of matrices. Let C, J be the $\mathsf{d} \times N\mathsf{d}$-matrices given by $C = \begin{pmatrix} C_1 & \cdots & C_N \end{pmatrix}$, $J = \begin{pmatrix} I_{\mathsf{d}} & \cdots & I_{\mathsf{d}} \end{pmatrix}$. Then $AB + BA = C^T J$, and

$$\begin{aligned} U(x,t) &= \frac{\partial}{\partial x}\Big(\big(I_{N\mathsf{d}} + \widehat{L}(x,t)B\big)^{-1} \widehat{L}(x,t) \, \big(AB+BA\big)\Big) \\ &= \Big(\frac{\partial}{\partial x}\Big((I_{N\mathsf{d}} + \widehat{L}B)^{-1}\widehat{L}\Big)\Big) \, C^T J \\ &=: \widehat{U}(x,t) \, C^T J, \end{aligned}$$

where $\widehat{L}$ is the block diagonal matrix with the blocks $\ell_j I_{\mathsf{d}}$, $j = 1, \dots, N$, on the diagonal.

So far we have obtained a solution of the $N\mathsf{d} \times N\mathsf{d}$-matrix KdV. To reduce dimension, we use that J has full rank, and therefore admits the right-inverse $K := J^T (JJ^T)^{-1}$.

Claim: The $\mathsf{d} \times \mathsf{d}$-matrix function $u(x,t) = J\,\widehat{U}(x,t)\,C$ solves (1.1).

Proof. Since $u = J\,\widehat{U}C$, we get $u_t = J\,\widehat{U}_t\,C = J\widehat{U}_t\,C(JK) = J(\widehat{U}_t\,CJ)K = JU_tK$ and, similarly, $u_{xxx} = JU_{xxx}K$. Furthermore,

$$uu_x = (J\,\widehat{U}\,C)(J\,\widehat{U}_x\,C) = J\Big(\widehat{U}(CJ)\,\widehat{U}_x\,CJ\Big)K = J(UU_x)K,$$

and analogously $u_x u = J(U_x U)K$.

As a consequence,

$$u_t - \big(\, u_{xxx} + 3\{u, u_x\} \,\big) = J\Big(U_t \; + 3\{U, U_x\}\Big)K = 0$$

because of the solution property of U. □

Note that we now have derived the solution formula

$$u(x,t) = \frac{\partial}{\partial x}\left(\Big(I_{\mathsf{d}} \dots I_{\mathsf{d}} \Big) \left(I_{N\mathsf{d}} + \Big(\frac{1}{k_i+k_j} \ell_i C_i^T \Big)_{i,j=1}^N \right)^{-1} \widehat{L} \begin{pmatrix} \ell_1 C_1^T \\ \vdots \\ \ell_N C_N^T \end{pmatrix} \right).$$

Finally, with u also u^T is a solution of (1.1), and the formula in Proposition 2.1 is established. □

References

[1] M.J. Ablowitz, B. Prinari, A.D. Trubatch, *Discrete and Continuous Nonlinear Schrödinger Systems*, Cambridge University Press 2004.

[2] H. Aden, B. Carl, *On realizations of solutions of the KdV equation by determinants on operator ideals,* J. Math. Phys. **37** (1996), 1833–1857.

[3] F. Calogero, A. Degasperis, *Spectral Transform and Solitons I: Tools to Solve and Investigate Nonlinear Evolution Equations*, Studies in Mathematics and its Applications Vol. 13 (North-Holland Publishing Company 1982).

[4] S. Carillo, M. Lo Schiavo, C. Schiebold, *N-soliton matrix mKdV solutions: a step towards their classification,* Preprint 2020.

[5] S. Carillo, C. Schiebold, *Noncommutative Korteweg-de Vries and modified Korteweg-de Vries hierarchies via recursion methods,* J. Math. Phys. **50** (2009), 073510, 14 pp.

[6] S. Carillo, C. Schiebold, *Matrix Korteweg-de Vries and modified Korteweg-de Vries hierarchies: noncommutative soliton solutions,* J. Math. Phys. **52** (2011), 053507, 21 pp.

[7] S. Carillo, C. Schiebold, *Construction of soliton solutions of the matrix modified Korteweg-de Vries equation,* To appear in: "Nonlinear Dynamics of Structures, Systems and Devices", Ed: W. Lacarbonara et al., Springer 2021.

[8] B. Carl, C. Schiebold, *Nonlinear equations in soliton physics and operator ideals,* Nonlinearity **12** (1999), 333–364.

[9] I.M. Gelfand, V.S. Retakh, *Determinants of matrices over noncommutative rings,* Funct. Anal. Appl. **25** (1991), 91–102.

[10] V.M. Goncharenko, *Multisoliton solutions of the matrix KdV equation,* Theor. Math. Phys. **126** (2001), 81–91.

[11] B.A. Kupershmidt, *KP or mKP: Noncommutative Mathematics of Lagrangian, Hamiltonian, and Integrable Systems*, Mathematical Surveys and Monographs vol. 78, AMS 2000.

[12] P.D. Lax, *Integrals of nonlinear equations of evolution and solitary waves,* Commun. Pure Appl. Math. **21** (1968), 467–490.

[13] A. Pietsch, *Eigenvalues and s-Numbers*, Cambridge University Press 1987.

[14] C. Schiebold, *On negatons of the Toda lattice,* J. Nonlinear Math. Phys. **10** (2003), 181–193.

[15] C. Schiebold, *Noncommutative AKNS systems and multisolitons solutions to the matrix sine-Gordon equation,* Discrete Contin. Dyn. Syst. Suppl. **2009** (2009), 678–690.

[16] C. Schiebold, *The noncommutative AKNS system: projection to matrix systems, countable superposition and soliton-like solutions,* J. Phys. A **43** (2010), 434030, 18 pp.

[17] C. Schiebold, *Matrix solutions for equations of the AKNS system.* In: "Nonlinear Systems and their Remarkable Mathematical Structures", Chapter B.5, p. 256–293, Ed: N. Euler, CRC Press 2018.

[18] C. Schiebold. *Asymptotics for the multiple pole solutions of the nonlinear Schrödinger equation,* Nonlinearity **30** (2017), 2930–2981.

Mid Sweden University, Sweden
Cornelia.Schiebold@miun.se

sciendo

Proceedings of the
Contemporary Mathematics in Kielce 2020
Jan Kochanowski University in Kielce, Poland
February 24-27, 2021

CANAL SURFACES AND FOLIATIONS – A SURVEY

PAWEŁ WALCZAK

KEYWORDS: canal surface, foliation

MSC2010: 53A23

Received 19 April 2021, accepted 3 August 2021

1. Introduction

Here, *conformal geometry* is seen as the theory of invariants of Möbius transformations of space forms $\mathbb{S}^3$, $\mathbb{R}^3$ or $\mathbb{H}^3$. Since the Euclidean space $\mathbb{R}^3$ and the hyperbolic space $\mathbb{H}^3$ are conformally equivalent to open subspaces of the sphere $\mathbb{S}^3$, the case of $\mathbb{S}^3$ is of greatest interest.

All the Möbius transformations of one of the spaces under consideration can be presented as a composition of spherical inversions and spherical inversions map spheres (or planes) onto spheres (or planes), the notion of sphere is conformally invariant. Since all these transformations are smooth, they map spheres osculating a surface onto spheres osculating its images. Also, they transform envelopes of families of spheres onto envelopes of corresponding families of their images.

Canal surfaces considered in this article are defined as envelopes of one parameter families of spheres. Surfaces of this sort can be found easily in nature: water pipes, hoses for vacuum cleaners and blood vessels being some of them. Their analytic models play an important role in 3D computer graphics. The simplest examples of canal surfaces are provided by surfaces of revolution and their images by Möbius transformations, spherical inversions in particular.

Foliations (here, of 3-dimensional manifolds) are particular families of connected and pairwise disjoint surfaces (called *leaves*) filling the manifolds. In differential geometry, several authors discuss existence, properties and classification of foliations by leaves satisfying particular geometric conditions, for example being

https://doi.org/10.2478/9788366675360-021.

totally geodesic, minimal, umbilical, of either constant, or positive, or negative curvatures (sectional, Ricci, scalar, mean etc.) and so on (see [30, 31] and the bibliographies therein).

In this article, we provide a survey of results concerning canal surfaces and *canal foliations*, that is foliations of 3-dimensional manifolds of constant curvature by leaves being canal surfaces, obtained recently (2000 – 2020) by the author and his collaborators. In Section 2, we describe (after [10] and [19]) a useful and well known interpretation of 2-dimensional oriented spheres $S \subset \mathbb{S}^3$ as points of the 4-dimensional de Sitter space. Section 3 is devoted to a description of local conformal invariants of surfaces used in the description of canal surfaces. Section 4 contains a survey of results on canal surfaces while Section 5 those on canal foliations.

Acknowledgments. The author expresses his gratitude to his collaborators for several years of cooperation and friendship, and to the organizers of the conference *Contemporary Mathematics in Kielce 2020* at Jan Kochanowski University for giving him an opportunity to present the survey.

2. Space of spheres

Consider the 5-dimensional *Lorenz space* $\mathbb{L}^5$, that is the vector space $\mathbb{R}^5$ equipped with the *Lorentz quadratic form* $\mathcal{L}$ and the associated *Lorentz bilinear form* $\mathcal{L}(\cdot,\cdot)$ given by

$$\mathcal{L}(x_0, \cdots, x_4) = -x_0^2 + x_1^2 + \cdots + x_4^2$$

and

$$\mathcal{L}(u, v) = -u_0 v_0 + u_1 v_1 + \cdots + u_4 v_4$$

when $u = (u_0, y_1, \ldots, u_4)$ and $v = (v_0, v_1, \ldots, v_4)$.

The isotropy cone $\mathcal{L}_{iso} = \{v \in \mathbb{R}^5\, ; \mathcal{L}(v) = 0\}$ of $\mathcal{L}$ is called the *light cone*. Its non-zero vectors are also called *light-like*. The light cone splits the set of nonzero vectors $v \in \mathbb{L}^5$ into two classes: A vector v in $\mathbb{R}^5$ is called *space-like* if $\mathcal{L}(v) > 0$ and *time-like* if $\mathcal{L}(v) < 0$. A straight line is called space-like (respectively, time-like) if it contains a space-like (respectively, time-like) vector.

The *de Sitter space* Λ^4 is defined as the set of all the points $x = (x_0, x_1, \ldots, x_5)$ of $\mathbb{R}^5$ for which $\mathbb{L}(x) = 1$.

The points at infinity of the light cone in the upper half space $\{x_0 > 0\}$ form a 3-dimensional sphere. Let us denote it by S^3_∞. Since it can be considered as the set of lines through the origin in the light cone, it is identified with the intersection S^3_1 of the upper half light cone and the hyperplane $\{x_0 = 1\}$, which is given by $S^3_1 = \{(x_1, \cdots, x_4) \,|\, x_1^2 + \cdots + x_4^2 = 1\}$.

To each point $\sigma \in \Lambda^4$ there corresponds a sphere $\Sigma = \sigma^\perp \cap S^3_\infty$ (or, $\Sigma = \sigma^\perp \cap S^3_1$), where $\sigma^\perp$ is the 4-dimensional vector subspace of $\mathbb{R}^5$ orthogonal (with respect to the form $\mathbb{L}$) to the straight line passing through σ and the origin (see Figure 1,

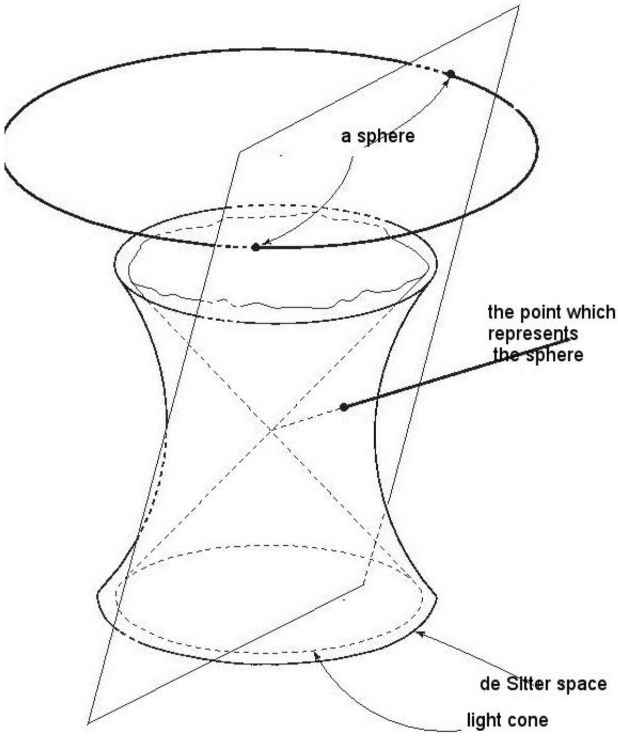


FIGURE 1. Spheres as points of de Sitter space.

where – by obvious technical reasons – the dimensions of all the geometric objects are reduced by 2: $5 \mapsto 3$, $4 \mapsto 2$, $3 \mapsto 1$ and $2 \mapsto 0$).

Since $(-\sigma)^{\perp} = \sigma^{\perp}$, we get two copies of the same sphere $\Sigma \subset \mathbb{S}^3$ equipped with two opposite orientations. This is why de Sitter space Λ^4 can be identified with the space of oriented 2-spheres contained in $\mathbb{S}^3$.

A regular curve $\gamma : I \to \mathbb{R}^5$ is called space-like if, at each point t ot the inerval I, its tangent vector $\dot{\gamma}(t)$ is space-like, that is $\mathbb{L}(\dot{\gamma}(t)) > 0$. If $\gamma(I) \subset \Lambda^4$, γ can be considered as the one-parameter family of the corresponding spheres $\Sigma(t)$. If both these conditions are satisfied, the family of spheres $\Sigma(t)$ associated to the points $\gamma(t)$, $t \in I$, defines an envelope. An extra condition is necessary to guarantee that the envelope is immersed: all the geodesic acceleration vectors $\vec{k_g}(t) = \ddot{\gamma}(t) + \gamma(t)$, $t \in I$, should be time-like.

Certain motivation for considering the correspondence $\sigma \longleftrightarrow \Sigma$ between points σ of de Sitter space Λ^4 and spheres $\Sigma \subset \mathbb{S}^3$ arises from the following.

Given two two-dimensional spheres S and Σ in $\mathbb{R}^3$ (or, $\mathbb{S}^3$), the spherical inversion of Σ with respect to S is another sphere (or, a plane considered as a sphere of infinite radius) $\tilde{\Sigma} = \iota_S(\Sigma)$. The corresponding points s, σ and $\tilde{\sigma}$ of Λ^4 are related by

$$\tilde{\sigma} = 2\mathbb{L}(s, \sigma)s - \sigma. \tag{2.1}$$

From (2.1), several properties of spherical inversion follow easily. For example,

$$\iota_S(S) = S,\ \iota_S^2(\Sigma) = \Sigma,\ \mathbb{L}(\iota_s(\sigma_1), \iota_s(\sigma_2)) = \mathbb{L}(\sigma_1, \sigma_2)$$

for all the spheres $S, \Sigma, \Sigma_1, \Sigma_2 \subset \mathbb{S}^3$ (and the corresponding points σ_1, σ_2 of Λ^4). Therefore, spherical inversions considered in Λ^4 appear to be isometries with respect to the Lorenz scalar product. Since, all the conformal transformations of $\mathbb{S}^3$ are compositions of such inversions, we arrive at

Proposition 2.1. *All the conformal transformations of the sphere $\mathbb{S}^3$ become Lorenzian isometries of Λ^4 under the correspondence between 2-spheres $\Sigma \subset \mathbb{S}^3$ and points of de Sitter space Λ^4 we discussed here.*

Remark 2.2. Note that the analytic description of the inversive image $\tilde{\Sigma}$ in Euclidean coordinates is significantly more complicated than that of $\tilde{\sigma}$ in (2.1):

$$\tilde{\Sigma} = S\Big(o + \frac{\rho^2(c-o)}{\|c-o\}^2 - r}, \frac{\rho^2 r}{(\|c-o\|^2 - r)^2)^2}\Big)$$

when $\Sigma = S(c, r)$ and $S = S(o, \rho)$ is the sphere of inversion.

3. Conformal invariants

Assume now that S is a surface which is *umbilic free*, that is, that the principal curvatures $k_1(x)$ and $k_2(x)$ of S are different at any point x of S. Let X_1 and X_2 be unit vector fields tangent to the curvature lines corresponding to, respectively, k_1 and k_2. Throughout the paper, we assume that $k_1 > k_2$. Put $\mu = (k_1 - k_2)/2$. For a long time, it is known ([34], see also [7]) that the vector fields $\xi_i = X_i/\mu$ and the coefficients θ_i $(i = 1, 2)$ in

$$[\xi_1, \xi_2] = -\frac{1}{2}(\theta_2\xi_1 + \theta_1\xi_2)$$

are invariant under arbitrary (orientation preserving) conformal transformation of the Euclidean space $\mathbb{R}^3$. (In fact, they are invariant under arbitrary conformal change of the Riemannian metric on $\mathbb{R}^3$.) Elementary calculations involving Codazzi equations show that

$$\theta_1 = \frac{1}{\mu^2}\cdot X_1(k_1) \quad \text{and} \quad \theta_2 = \frac{1}{\mu^2}\cdot X_2(k_2).$$

The quantities θ_i $(i = 1, 2)$ are called *conformal principal curvatures* of S.

Another conformally invariant scalar quantity Ψ can be derived from the derivation of the Bryant's *conformal Gauss map* β which maps a point x of S to the sphere tangent at x to S that has the same mean curvature as S at x (see [4]).

The sphere $\beta(x)$ can be seen as a point of $\Lambda^4 \subset \mathbb{L}^5$. We get (with all the scalar products $\langle\cdot,\cdot\rangle$ below denoting the Lorentz bilinear form $\mathcal{L}$ in $\mathbb{L}^5$)

$$\begin{aligned} &\frac{1}{2}\big(\langle\xi_1(\xi_1(\beta)),\xi_1(\xi_1(\beta))\rangle - \langle\xi_2(\xi_2(\beta)),\xi_2(\xi_2(\beta))\rangle \\ &\qquad -\langle\xi_1(\xi_1(\beta)),\xi_2(\beta)\rangle^2 + \langle\xi_2(\xi_2(\beta)),\xi_1(\beta)\rangle\big) \\ &\qquad = \Psi + \frac{1}{2}\left(\theta_1^2 - \theta_2^2 + \xi_1(\theta_1) + \xi_2(\theta_2)\right). \end{aligned} \tag{3.1}$$

The quantities on both sides of the above equality are equal to

$$\frac{1}{\mu^3}\left(\triangle H + 2\mu^2 H\right), \tag{3.2}$$

where H is the mean curvature of S and $\triangle$ is the Laplace operator on S equipped with the Riemannian metric induced from the ambient space. Moreover, this quantity appears in the Euler-Lagrange equation for *Willmore functional*

$$S \mapsto \int_S \mu^2 d\,area.$$

The vector fields ξ_1, ξ_2 (or, the dual 1-forms ω_1, ω_2) together with quantities θ_1, θ_2 and Ψ generate all the local conformal invariants for surfaces and determine a surface locally up to conformal transformations of $\mathbb{R}^3$ ([16], see also [7] again).

Define (5×5) matrices A_1 and A_2 by

$$A_1 = \begin{pmatrix} \theta_1/2 & -(1+\Psi)/2 & b/2 & \theta_1/2 & 0 \\ 1 & 0 & 0 & -1 & (1+\Psi)/2 \\ 0 & 0 & 0 & 0 & -b/2 \\ 0 & 1 & 0 & 0 & -\theta_1/2 \\ 0 & -1 & 0 & 0 & -\theta_1/2 \end{pmatrix} \tag{3.3}$$

and

$$A_2 = \begin{pmatrix} -\theta_2/2 & -c/2 & -(1-\Psi)/2 & \theta_2/2 & 0 \\ 0 & 0 & 0 & 0 & c/2 \\ 1 & 0 & 0 & 1 & (1-\Psi)/2 \\ 0 & 0 & -1 & 0 & -\theta_2/2 \\ 0 & 0 & -1 & 0 & \theta_2/2 \end{pmatrix}, \tag{3.4}$$

where $b = -\theta_1\theta_2 + \xi_2(\theta_1)$ and $c = \theta_1\theta_2 + \xi_1(\theta_2)$.

Proposition 3.1 (Fialkov, [16])**.** *Given a simply connected domain $U \subset \mathbb{R}^2$, linearly independent 1-forms ω_1 and ω_2 and smooth functions θ_1, θ_2 and Ψ defined on U for which the matrix valued 1-form ω,*

$$\omega = A_1\omega_1 + A_2\omega_2, \tag{3.5}$$

satisfies the structural equation

$$d\omega + (1/2)[\omega,\omega] = 0, \tag{3.6}$$

there exists an immersion $\iota : U \to \mathbb{R}^3$ for which $S = \iota(U)$ realizes these forms and functions as local conformal invariants.

The integrability condition (3.6) implies the following.

Proposition 3.2 ([2]). *Any surface $S \subset \mathbb{R}^3$ with constant conformal principal curvatures has at least one of these curvatures equal to zero.*

Moreover, for arbitrary constant c, the family of all the immersed in $\mathbb{R}^3$ surfaces $S = \iota(\mathbb{R}^2)$ with constant conformal principal curvatures 0 and c is nonempty and parametrized by triples (f_2, C_1, C_2), where f_2 and C_1 are smooth real functions of one variable while C_2 is a real constant. The corresponding surface S admits the conformally invariant 1-forms $\omega_i = f_i dx_i$ with f_1 given by

$$f_1(x_1, x_2) = C_1(x_1) \cdot e^{-\frac{1}{2}c\int_0^{x_2} f_2(t)dt}, \tag{3.7}$$

where $C_1 : \mathbb{R} \to \mathbb{R}$ is arbitrary smooth function while the Bryant conformal invariant Ψ of S is given by

$$\Psi(x_1, x_2) = C_2 \cdot \exp\Big(- c \cdot \int_0^{x_2} f_2(t)dt\Big) - 2. \tag{3.8}$$

Example 3.3. Torus, cylinder and cone of revolution, and their images under arbitrary Möbius transformations are the only surfaces with both conformal principal curvatures equal to zero. All these surfaces are called *Dupin cyclides.*

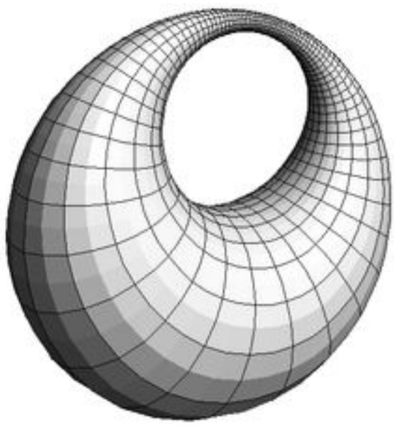


FIGURE 2. A Dupin cyclide.

In [11, 12, 13], Darboux mentioned several results concerning Dupin cyclides. Among them, one can find the following.

Proposition 3.4 (Darboux). *Dupin cyclides are the only surfaces that are in two different ways envelopes of one-parameter families of spheres as well as the only surfaces that have circles as all the lines of curvature.*

4. Canal surfaces

As mentioned in Introduction, *canal surfaces* in space forms are defined as the envelopes of one-parameter families of spheres. Therefore, they can be seen as space-like curves in de Sitter space Λ^4 (of time-like geodesic curvature vectors when regularly immersed).

Since all the Möbius transformations map spheres to spheres, the notion of canal belongs to conformal geometry: conformal image of a canal surface is a canal surface. The simplest examples of canal surfaces are provided by Dupin cyclides (see Section 3), surfaces of revolution and their images under Möbius transformations.

By definition, any canal surface S in $\mathbb{R}^3$ is the solution of a system of equations

$$\begin{cases} (x - x(t))^2 + (y - y(t))^2 + (z - z(t))^2 = r(t)^2, \\ x'(t)(x - x(t)) + y'(t)(y - y(t)) + z'(t)(z - z(t)) = r'(t)r(t), \end{cases} \tag{4.1}$$

where – obviously – the first of these equations defines a sphere enveloped by S while the second one – a plane. The intersection of the osculating sphere with the corresponding plane is – in general – a circle contained in S and is called a *characteristic circle* of S.

Proposition 4.1. *Characteristic circles of canal surfaces appear to be their lines of curvature corresponding to the principal conformal curvature, say θ_1, vanishing identically along the surface: $\theta_1(p) = 0$ for any point p of the canal surface under consideration.*

In some sense, canal surfaces admit osculating Dupin cyclides called in [1] *necklaces.* (For a discussion of osculation of Dupin cyclides and arbitary surfaces see [3].)

Theorem 4.2 ([1])**.** *The osculating spheres $\sigma_2(\phi)$ for the principal curvature k_2 along a characteristic circle Γ (being a line of principal curvature for k_1 parametrized by ϕ) have an envelope which is a Dupin cyclide $\mathcal{D}$; in other terms the corresponding points $\sigma_2(\phi) \in \Lambda^4$ form a circle.*

Since osculating spheres have order of tangency 2 with the corresponding canal, the Bryant invariants Ψ_S and Ψ_N of a canal surface S and its necklace N are equal along the characteristic circle of their tangency. Since Ψ_N is constant, we get the following.

Corollary 4.3 ([1])**.** *The Bryant invariant of a canal surface is constant along its characteristic circles.*

The above Corollary and the condition $\theta_1 \equiv 0$ satisfied for all the canals motivates the interest in canal surfaces for which θ_2 is constant along the characteristic circles. Such canals are called in [1] *special.*

Special canals. Let us classify special canals, that is canal surfaces satisfying the conditions

$$\theta_1 = 0, \quad \theta_2 \text{ is constant along characteristic circles.} \tag{4.2}$$

Theorem 4.4 ([1]). *A surface S in $\mathbb{R}^3$ is a special canal if and only if it is a conformal image of either a surface of revolution or a cylinder over a planar curve, or a cone.*

Note that the three classes of surfaces listed in the above theorem can be characetrized in terms of de Sitter space Λ^4:

Proposition 4.5 ([1]).

(i) *Special canal surfaces are envelopes of spheres which form a curve γ contained in $\Lambda^4 \cap V$, V being a 3-dimensional vector subspace of $\mathbb{R}^5$.*
(ii) *V is of mixed type (resp., degenerate, resp. space-like) whenever S is a conformal image of a surface of revolution (resp., a cylinder, resp., a cone).*
(iii) *in this case, the intersection $\Lambda^4 \cap V$ looks like a 2-dimensional de Sitter subspace Λ^2 of Λ^4, resp., a cylinder generated by parallel light-rays, resp., a sphere.*

Now, let us recall that a surface S in $\mathbb{R}^3$ is said to be *isothermic* whenever it is obtained locally by charts (x_1, x_2) such that the curves $x_i = \text{const.}, i = 1, 2,$ coincide with the curvature lines on S. In other words, S is isothermic whenever there exists a positive function f such that the Lie bracket $[f\xi_1, f\xi_2]$ vanishes identically. From the definition of conformal principal curvatures in Section 3 it follows that this condition is equivalent to the following one:

$$\xi_1(f) = (1/2)f\xi_2, \quad \xi_2(f) = (-1/2)f\xi_1.$$

Note that this is a conformally invariant notion: the image of an isothermic surface by a Möbius transformation in again isothermic. For more information about such surfaces (and, more general, submanifolds of higher dimension) see, for example, [5], [6] and the bibliographies therein.

Finally, if S is an isothermic canal, then, say, $\theta_1 = 0$, $\xi_2(f) = 0$, f is constant along characteristic circles, the same holds for $\xi_2(f)$ and for $\theta_2 = -2\xi_2(f)/f$. This yields the following.

Theorem 4.6 ([1]). *Any isothermic canal surface is special.*

In [28], it has been shown that any *Willmore canal surface*, that is a canal surface which is a critical point of the *Willmore functional*

$$S \mapsto \int_S \mu_2 dA,$$

is isothermic. Therefore,

Corollary 4.7. *Any Willmore canal is special.*

5. Canal foliations

Recall (see, for example, [9] and[18]) that a p-dimensional C^r-*foliation* $\mathcal{F}$ ($r = 0, 1, \dots, \infty$) on an n-dimensional manifold M is a decomposition of M into connected submanifolds – called *leaves* – such that for any $x \in M$ there exists a C^r-differentiable chart $\phi = (\phi', \phi'') : U \to \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$ defined on a neighbourhood U of x and satisfying the condition

($*$) for any L of $\mathcal{F}$, the connected components – called *plaques* – of $L \cap U$ are given by the equation $\phi'' =$ const.

The simplest examples of foliations are provided by submersions $F : M \to N$ with leaves being the connected components of the fibers $F^{-1}(y)$, $y \in N$, in particular by products $M = M_1 \times M_2$ of connected manifolds with leaves $M_1 \times \{y\}$, $y \in M_2$. The first non-trivial example of a foliation of the sphere $\mathbb{S}^3 = \{(w, z) \in \mathbb{C}^2; |w|^2 + |z|2 = 1\}$ has been provided by Reeb [29]: the *Reeb foliation* is obtained by gluing along the common toral boundary $T^2 = \{(w, z); |w|^2 = |z|^2 = 1/2\}$ the two *Reeb components* (see Figure 3) which can be produced from the strip $[-1, 1] \times \mathbb{R}$ foliated by the boundary lines and the graphs of smooth functions $f + c$, $c \in \mathbb{R}$, where $f : (-1, 1) \to \mathbb{R}$ satisfies $f(-t) = f(t)$, $f''(t) > 0$ and $\lim_{t\to\pm1} f(t) + \infty$, by: first, the rotation around the axis $t = 0$ of symmetry of the strip, then passing to the quotient $D^2 \times \mathbb{R}/\mathbb{Z}$, $\mathbb{Z}$ being the group of translations generated by $(w, s) \mapsto (w, s + 1)$, $|w| \le 1$, $s \in \mathbb{R}$.

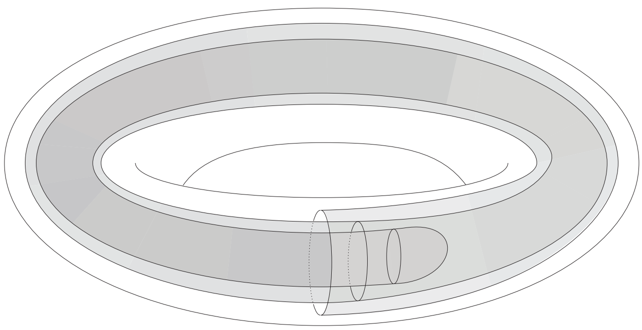


Figure 3. A Reeb component.

Observe that Reeb components can be foliated by canal surfaces, therefore the same happens to Reeb foliations of $\mathbb{S}^3$. Foliations of 3-manifolds of constant curvature by canal surfaces are called here *canal foliations*.

One of the procedures which allows to produce more complicated foliations from given ones is called *turbulization*.

Let us begin with a 2-dimensional foliation $\mathcal{F}$ of a manifold M, $\dim M = 3$. Find a loop Γ transverse to $\mathcal{F}$ and its tubular neighbourhood $N(\Gamma) \approx D^2 \times S^1$. Replace $\mathcal{F}$ outside $N(\Gamma)$ by the foliation shown in Figure 4 and fill the interior of $N(\Gamma)$ with a Reeb component. The resulting foliation $\mathcal{F}'$ is the turbulized $\mathcal{F}$.

In Figure 4, L is a piece of a leaf of $\mathcal{F}$ and γ a curve on L while L' is a piece of a leaf of $\mathcal{F}'$ and γ' a curve on L'.

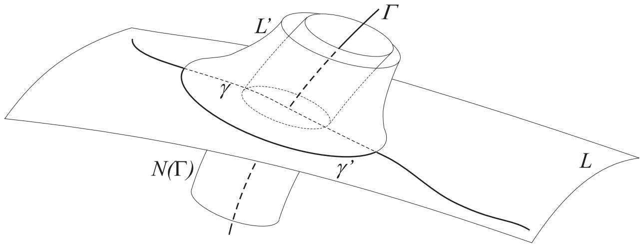


FIGURE 4. Turbulization.

Certainly, this procedure can be repeated as long as we can find new loops transverse to the foliation under consideration and turbulizations of canal foliations can sometimes provide new canal foliations.

Similarly to the construction of Reeb components, one can foliate the zone $Z = T^2 \times [0,1]$ by a family of tori $T^2 \times \{t\}$, $t \in A$, $A \subset [0,1]$ being closed, and filling the zones between two consecutive tori with cylinders spiralling from one boundary component of such zone towards the other one, Figure 5.

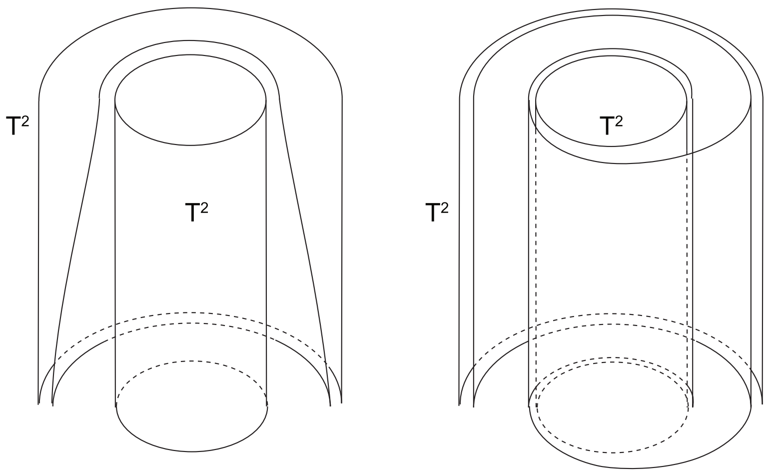


FIGURE 5. Two ways of spiraling.

Canal foliations of the sphere $\mathbb{S}^3$ have been classified:

Theorem 5.1 ([26]). *Any foliation $\mathcal{F}$ of $\mathbb{S}^3$ by canal surfaces is either*

- *a Reeb foliation with the toral leaf being a Dupin cyclide or*
- *is obtained from such a Reeb foliation inserting a zone $Z \simeq \mathbf{T}^2 \times [0,1]$ foliated by toral and cylindrical leaves (as described above).*

A question about the existence of canal foliations on other 3-manifolds of constant curvature arises in a natural way.

The first step is to ask about existence of foliations by the simplest canals: Dupin cyclides. Such foliations are called *Dupin foliations* here.

Theorem 5.2 ([25])**.** (i) *Dupin foliations of $\mathbb{S}^3$ do not exist.*
(ii) *The only Dupin foliations of $\mathbb{R}^3$ are these by parallel planes.*
(iii) *Closed hyperbolic manifolds admit no Dupin foliations.*

Certainly, one can imagine several different examples of canal foliations of the Euclidean 3-space $\mathbb{R}^3$ and of the hyperbolic 3-space $\mathbb{H}^3$. The most interesting case is that of closed hyperbolic 3-manifolds. It occurs, that the best way towards this goal is to find a topological version of the "canalicity".

Following [17], we shall say that a *griddled structure* on a surface L is a 1-dimensional foliation $\mathcal{C}$ (with singularities) such that any singularity of $\mathcal{C}$ is isolated and any regular leaf of $\mathcal{C}$ is homeomorphic to $\mathbb{S}^1$. Similarly, a *griddled structure* on a foliated 3-manifold $(M, \mathcal{F})$ will be an orientable subfoliation $\mathcal{C}$ of the codimension 1 foliation $\mathcal{F}$ which induces by restriction a griddled structure on each leaf L of $\mathcal{F}$. In both cases, we will say that L and $(M, \mathcal{F})$ are *griddled*.

First, we are going to classify griddled surfaces. To this end, recall that an action of $\mathbb{S}^1$ is called *semi-free* if any non trivial isotropy subgroup coincides with $\mathbb{S}^1$. Now a semi-free action of $\mathbb{S}^1$ on a surface defines a griddled structure provided that its singularities are isolated; we call such a structure *canonical*.

Example 5.3. Natural examples of canonical griddled structures on surfaces are provided by the following semi-free actions:

(1) the action of the group of rotations around the origin of the plane $\mathbb{R}^2$ or the unit disk $\mathbb{D}^2$; the action of the group of rotations of the unit sphere S^2 around the vertical axis $\overrightarrow{Oz}$ of $\mathbb{R}^3$; these structures have either one or two singularities,

(2) the natural free action of the first factor on the annulus $\mathbb{A} = \mathbb{S}^1 \times [0, 1]$, the cylinders $\mathbb{S}^1 \times \mathbb{R}$ or $\mathbb{S}^1 \times [0, 1]$, the torus $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$, all the four structures being *regular*, that is without singularities.

Lemma 5.4. *Any singularity of a connected griddled surface $(L, \mathcal{C})$ is a center and $(L, \mathcal{C})$ is topologically conjugate to one of the canonical griddled surfaces of Example 5.3.*

The above lemma together with several technical steps allows to classify griddled foliated 3-manifolds (either with boundary or without):

Theorem 5.5 ([17])**.** *A compact 3-manifold M supports a griddled foliation if and only if M is one of the following manifolds:*
i) $D^2 \times [0, 1]$, $D^2 \times S^1$, $S^2 \times [0, 1]$ or $T^2 \times [0, 1]$ if $\partial M \neq \emptyset$,
ii) $S^2 \times S^1$, T^3 or an S^1-bundle over T^2 if $\partial M = \emptyset$.

Certainly, canal surfaces and many canal foliations are griddled with the griddled structure consisting of all the characteristic circles of the surface/leaves.

However, the standard Reeb foliation on $\mathbb{S}^3$ is not griddled: the griddled structures of the two Reeb components induce two transverse griddlings of the toral leaf. This is why we need to accept the following definitions:

Consider a finite family $\mathcal{D} = \{L_1, L_2, \ldots, L_k\}$ of compact leaves of a codimension 1 compact foliated manifold $(M, \mathcal{F})$. Cutting M along these leaves, we produce a finite family $\{N_1, N_2, \ldots, N_l\}$ of compact foliated manifolds and a foliation preserving submersive map

$$\psi : \coprod_j N_j \to M.$$

The restriction of ψ either to the interior or to any boundary component of any N_j is injective. We call ψ a *foliated decomposition* of $(M, \mathcal{F})$ defined by $\mathcal{D}$. Now, we introduce our last concept:

A codimension 1 foliation $\mathcal{F}$ on a compact connected 3-manifold M (possibly with boundary) is a *topological canal foliation* if there exists a foliated decomposition $\psi : \coprod_j N_j \to M$ of $(M, \mathcal{F})$ defined by a finite family $\mathcal{D}$ of compact leaves verifying the following two conditions:

(i) for each j, the foliation $\mathcal{F}_j$ induced by $\mathcal{F}$ on N_j is tangent to the boundary ∂N_j, admits a griddled structure $\mathcal{C}_j$ and any component of ∂N_j is a regularly griddled torus,

(ii) any torus $L_j \in \mathcal{D}$ being the image by ψ of two boundary components of $\coprod_j N_j$ is endowed with two griddled structures which are mutually transverse.

The elements of $\mathcal{D}$ are called the *turning leaves* of $\mathcal{F}$ and the manifolds N_j are its *griddled components*.

Theorem 5.6 ([17]). *A compact* 3*-manifold M supports a topological canal foliation if and only if M is one of the following:*

(i) $D^2 \times [0,1]$, $D^2 \times S^1$, $S^2 \times [0,1]$ *or* $T^2 \times [0,1]$ *if* $\partial M \neq \emptyset$,

(ii) $S^2 \times S^1$, S^3 *or any Lens space,* T^3 *or any* S^1*-bundle over* T^2 *or any* T^2*-bundle over* S^1 *if* $\partial M = \emptyset$.

None of the manifolds M listed in the theorem above admits a hyperbolic structure. Indeed, the fundamental group of each of them (and its doubling $M\#M$ in the case $\partial M \neq \emptyset$) has the growth of polynomial type while – as shown in [27] and [33] – the fundamental groups of all closed hyperbolic manifolds have exponential type of growth. (For more about types of growth, see – for instance – [35], Chapter 2.) As mentioned before, many geometric canal foliations admit griddled structures and, in fact, all of them admit structures of topological canal foliations. Therefore, we can conclude with the following.

Corollary 5.7. *Closed hyperbolic* 3*-manifolds do not admit (neither geometric nor topological) canal foliations.*

6. Epilogue

Certainly, this article does not exhaust the list of recent result on canals. Here, we rewiev briefly some of those which were not mentioned in previous sections.

1. In [24], the authors find the minimal value of the length (in de Sitter space) of closed space-like curves with non-vanishing non-space-like geodesic curvature vector. These curves are in correspondence with closed almost regular canal surfaces, and their length is a natural quantity in conformal geometry.

2. As shown in Section 2, the space of 2-dimensional spheres in the 3-dimensional space form has dimension 4. Similarly, the space of Dupin cyclides can be shown to be of dimension 9. Given two *contact conditions*, that is two planes equipped with two points (supposed to be points of tangency), in general there is no sphere satisfying them (that is, tangent to the planes at given points). Similarly (see [20, 23]), given three such tangency conditions, in general there is no Dupin cyclide satisfying them. But, one can find a codimension-one subspace of triples of contact conditions such that for any its element there exists a one-parameter family of Dupin cyclides satisfying the three contacts.

3. In [15], the authors provide an algorithm to compute in de Sitter space a characteristic circle of a Dupin cyclide given a point and the tangent line. They provide also iterative algorithms (in the space of spheres) to compute (in 3D space) some characteristic circles of a Dupin cyclide which blends two particular canal surfaces. In [14], tools from geometric algebra are used to study Dupin cyclides.

4. As mentioned in Theorem 5.2, regular foliations by Dupin cyclides do not exist neither on $\mathbb{S}^3$ nor on closed hyperbolic manifolds. However, singular Dupin foliations, like that on $\mathbb{S}^3 \subset \mathbb{C}^2$ by tori $|w| = a,\ |z| = b$ with $a, b > 0$ satisfying $a^2 + b^2 = 1$, may be considered. In [21], the authors constructed several examples of such foliations on the unite sphere and illustrated them with very interesting pictures produced by computer-drawing.

Also, cyclides different than these called Dupin can be considered: A general *cyclide* is a closed surface in $\mathbb{S}^3$ which is spanned by two transverse families of circles such that exactly one circle of each family passes through each point of the surface. Singular foliations by general cyclides on $\mathbb{S}^3$ are studied in [22].

References

[1] A. Bartoszek, R. Langevin, P. Walczak, *Special canal surfaces of* $\mathbb{S}^3$, Bull. Braz. Math. Soc. **42** (2011), 301–319.

[2] A. Bartoszek, P. Walczak, *Foliations by surfaces of a peculiar class*, Ann. Polon. Math. **94** (2008), 89–95.

[3] A. Bartoszek, P. Walczak, Sz. Walczak, *Dupin cyclides osculating surfaces*, Bull. Braz. Math. Soc. **45** (2014), 179–195.
[4] R. Bryant, *A duality theorem for Willmore surfaces*, J. Diff. Geom. **20** (1984), 23–53.
[5] F. E. Burstall, *Isothermic surfaces: conformal geometry, Clifford algebras and integrable systems*, in Integrable systems, Geometry and Topology (ed. C.-L. Terng) AMS/IP Studies in Adv. Math. **36** (2006), 1–82.
[6] F.E. Burstall, D.M.J. Calderbank, *Conformal submanifold geometry I–III*, arXiv:1006.5700[math.DG], 2010.
[7] G. Cairns, R. W. Sharpe, L. Webb, *Conformal invariants for curves and surfaces in three dimensional space forms*, Rocky Mountain J. Math., **24** (1994), 933–959.
[8] C. Camacho, A. Lins Neto, *Geometric Theory of Foliations*, Birkhäuser, 1985.
[9] A. Candel, L. Conlon, *Foliations I & II*, Amer. Math. Soc., 2000 & 2003.
[10] T. Cecil, *Lie sphere geometry*, Universitext, Springer1992.
[11] G. Darboux, *Leçons sur la théorie générale des surfaces*, Gauthier-Villars, Paris 1887.
[12] G. Darboux, *Leçons sur les systèmes orthogonaux et les coordonnées curvilignes*, Gauthier Villars, Paris 1910.
[13] G. Darboux, *Principes de géométrie analytique*, Gauthier-Villars, Paris 1946.
[14] L. Druton, L. Fuchs, L. Garnier, R. Langevin, *The non-degenerate Dupin cyclides in the space of spheres using geometric algebra*, Adv. Appl. Clifford Algebras **24** (2014), 515–532.
[15] L. Druton, L. Garnier, R. Langevin, *Iterative construction of Dupin cyclide characteristic circles using non-stationary iterated functon systems (IFS)*, Comput.-Aided Des. **45** (2013), 568–573.
[16] A. Fialkov, *Conformal differential geometry of a subspace*, Trans. Amer. Math. Soc. **56** (1944), 309–433.
[17] G. Hector, R. Langevin, P. Walczak, *Topological canal foliations*, J. Math. Soc. Japan **71** (2019), 43–63.
[18] G. Hector, U. Hirsch, *Introduction to the Geometry of Foliations*, Parts A & B, Vieweg, 1981 & 1983.
[19] U. Hertrisch-Jeromin, *Introduction to Möbius Differential Geometry*, Cambridge Univ. Press, 2003.
[20] R. Krasauskas, C. Mäurer, *Studying cyclides with Laguerre geometry*, Comp. Aided Geom. Design **17** (2000), 101–126.
[21] R. Langevin, J.-C. Sifre, *Foliations of* $\mathbb{S}^3$ *by Dupin cyclides*, in Foliations 2012, World. Sci. 2013, 67–102.
[22] R. Langevin, J.-C. Sifre, *Foliations of* $\mathbb{S}^3$ *by cyclides*, in Geometry, dynamics and foliations 2013, Adv. Stud. Pure Math. **72** (2017), 395–427.
[23] R. Langevin, J.-C. Sifre, L. Druoton, L. Garnier, M. Paluszny, *Finding a cyclide given three contact conditions*, Comput. Appl. Math. **34** (2014), 275–292.
[24] R. Langevin, G. Solanes, *The geometry of canal surfaces and the length of curves in de Sitter space*, Adv. Geom. **11** (2011), 585–601.
[25] R. Langevin, P. Walczak, *Conformal geometry of foliations*, Geom. Dedicata **132** (2008), 135–178.
[26] R. Langevin, P. Walczak, *Canal foliations on* $\mathbb{S}^3$, J. Math. Soc. Japan **64** (2012), 657–680.
[27] J. Milnor, *A note on curvature and volume*, J. Diff. Geom. **2** (1968), 1–7.
[28] E. Musso, L. Nicoldi, *Willmore canal surfaces in Euclidean space*, Rend. Istit. Mat. Univ. Trieste **31** (1999), 177–202.
[29] G. Reeb, *Certaines Proprieétés Topologiques des Variétés Feuilletées*, Hermann, Paris 1952.
[30] V. Rovenski, P. Walczak, *Topics in Extrinsic Geometry of Codimension-One Foliations*, Springer, 2011.
[31] V. Rovenski, P. Walczak, *Extrinsic Geometry of Foliations*, Birkhäuser, 2021.
[32] R. W. Sharp, *Differential Geometry*, Springer, 1997.

[33] A. S. Svarc, *A volume invariant of coverings*, Dokl. Akad. Nauk SSSR **105** (1955), 32–34 (in Russian).
[34] A. Tresse, *Sur les invariants différentiels d'une surface par rapport aux transformations conformes de l'espace*, C.R. Acad. Sci. Paris **114** (1892), 948–950.
[35] P. Walczak, *Dynamics of Foliations, Groups and Pseudogroups*, Birkhäuser, 2004.

Uniwersytet Łódzki, Poland
pawel.walczak@wmii.uni.lodz.pl