B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

SÉRIE: RECHERCHES SUR LES DÉFORMATIONS

Volume LXI, no. 1

ŁÓDŹ 2011

ŁÓDZKIE TOWARZYSTWO NAUKOWE
PL-90-505 Łódź, ul. M. Cuie-Skłodowskiej 11
tel. (42) 665 54 59, fax (42) 665 54 64
sprzedaż wydawnictw: tel. (42) 665 54 48
e-mail: ltn@ltn.lodz.pl

Wydano z pomocą finansową Ministerstwa Nauki i Szkolnictwa Wyższego

PL ISSN 0459–6854

Wydanie 1.
Nakład 250 egz.

Skład komputerowy: Zofia Fijarczyk

Druk i oprawa: *AnnGraf* s.c.
92-637 Łódź, ul. Rataja 54
tel. +48 601 24 10 83

INSTRUCTION AUX AUTEURS

1. La présente Série du Bulletin de la Société des Sciences et des Lettres de Łódź comprend des communications du domaine des mathématiques, de la physique ainsi que de leurs applications liées aux déformations au sense large.
2. Toute communications est présentée à la séance d'une Commission de la Société par un des membres (avec deux opinions de spécialistes designés par la Rédaction). Elle doit lui être adressée directement par l'auteur.
3. L'article doit être écrit en anglais, français, allemand ou russe et débuté par un résumé en anglais ou en langue de la communication présentée. Dans tous les travaux écrits par des auteurs étrangers le titre et le résumé en polonais seront préparés par la rédaction. Il faut fournir le texte original qui ne peut contenir plus de 15 pages (plus 2 copies).
4. Comme des articles seront reproduits par un procédé photographique, les auteurs sont priés de les préparer avec soin. Le texte tapé sur un ordinateur de la classe IBM PC avec l'utilisation d'une imprimante de laser, est absolument indispensable. Il doit être tapé préférablement en *AMS-TEX* ou, exceptionnellement, en *Plain-TEX* ou *LATEX*. Après l'acceptation de texte les auteurs sont priés d'envoyer les disquettes (PC). Quelle que soient les dimensions des feuilles de papier utilisées, le texte ne doit pas dépasser un cadre de frappe de 12.3×18.7 cm (0.9 cm pour la page courante y compris). Les deux marges doivent être de la même largeur.
5. Le nom de l'auteur (avec de prénom complet), écrit en italique sera placé à la 1ère page, 5.6 cm au dessous du bord supérieur du cadre de frappe; le titre de l'acticle, en majuscules d'orateur 14 points, 7.1 cm au dessous de même bord.
6. Le texte doit être tapé avec les caractères Times 10 points typographiques et l'interligne de 14 points hors de formules longues. Les résumés, les rénvois, la bibliographie et l'adresse de l'auteurs doivent être tapés avec les petites caractères 8 points typographiques et l'interligne de 12 points. Ne laissez pas de "blancs" inutiles pour respecter la densité du texte. En commençant le texte ou une formule par l'alinéa il faut taper 6 mm ou 2 cm de la marge gauche, respectivement.
7. Les texte des théorèmes, propositions, lemmes et corollaires doivent être écrits en italique.
8. Les articles cités seront rangés dans l'ordre alphabétique et précédés de leurs numéros placés entre crochets. Après les références, l'auteur indiquera son adress complète.
9. Envoi par la poste: protégez le manuscript à l'aide de cartons.
10. Les auteurs recevront 20 tirés à part à titre gratuit.

Adresse de la Rédaction de la Série:
Département d'Analyse complexe et Géométrie différentielle
de l'Institut de Mathématiques de l'Académie polonaise des Sciences
BANACHA 22, PL-90-238 ŁÓDŹ, POLOGNE

Name and surname of the authors

TITLE – INSTRUCTION FOR AUTHORS SUBMITTING THE PAPERS FOR BULLETIN

Summary

Abstract should be written in clear and concise way, and should present all the main points of the paper. In particular, new results obtained, new approaches or methods applied, scientific significance of the paper and conclusions should be emphasized.

1. General information

The paper for BULLETIN DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ should be written in LaTeX, preferably in LaTeX 2e, using the style (the file **bull.cls**).

2. How to prepare a manuscript

To prepare the LaTeX 2e source file of your paper, copy the template file **instr.tex** with **Fig1.eps**, give the title of the paper, the authors with their affiliations/addresses, and go on with the body of the paper using all other means and commands of the standard class/style 'bull.cls'.

2.1. Example of a figure

Figures (including graphs and images) should be carefully prepared and submitted in electronic form (as separate files) in Encapsulated PostScript (EPS) format.

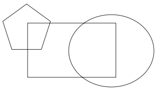

Fig. 1: The figure caption is located below the figure itself; it is automatically centered and should be typeset in small letters.

2.2. Example of a table

Tab. 1: The table caption is located above the table itself; it is automatically centered and should be typeset in small letters.

Description 1	Description 2	Description 3	Description 4
Row 1, Col 1	Row 1, Col 2	Row 1, Col 3	Row 1, Col 4
Row 2, Col 1	Row 2, Col 2	Row 2, Col 3	Row 2, Col 4

3. How to submit a manuscript

Manuscripts have to be submitted in electronic form, preferably via e-mail as attachment files sent to the address **zofija@mvii.uni.lodz.pl**. If a whole manuscript exceeds 2 MB composed of more than one file, all parts of the manuscript, i.e. the text (including equations, tables, acknowledgements and references) and figures, should be ZIP-compressed to one file prior to transfer. If authors are unable to send their manuscript electronically, it should be provided on a disk (DOS format floppy or CD-ROM), containing the text and all electronic figures, and may be sent by regular mail to the address: **Department of Solid State Physics, University of Lodz, Bulletin de la Société des Sciences et des Lettres de Łódź, Pomorska 149/153, 90-236 Lodz, Poland**.

References

[1]

Affiliation/Address

TABLE DES MATIÈRES

1. **A. Jamiołkowski** and **M. Michalski**, Quantum information theory – Toruń school 9–21
2. **M. Vaccaro**, Basics of linear para-quaternionic geometry I. Hermitian para-type structures on a real vector space 23–36
3. **J. Rutkowski, L. Wojtczak,** and **C. Surry**, Singular perturbation problem. How to approximate Dirac function 37–46
4. **M. Skwarczyński**, De Branges theorem and generalized hypergeometric functions I. Bieberbach conjecture and Milin functional 47–67
5. **M. Skwarczyński**, De Branges theorem and generalized hypergeometric functions II. De Branges functional and hypergeometric equation 69–88
6. **M. Skwarczyński**, De Branges theorem and generalized hypergeometric functions III. Basic properties of de Branges functions 89–103
7. **J. Garecki**, Teleparallel equivalent of general relativity: a critical review 105–118
8. **A. K. Kwaśniewski**, Graded posets inverse zeta matrix formula IIA. The formula of inverse ζ-matrix for graded posets with the finite set of minimal elements via natural join of matrices and digraph techniques – A. Relabeling and exercises 119–128
9. **A. K. Kwaśniewski**, Graded posets inverse zeta matrix formula IIB. The formula of inverse ζ-matrix for graded posets with the finite set of minimal elements via natural join of matrices and digraph techniques – B. Weighted reflexive reachability relation 129–141
10. **S. Bednarek** and **T. Bednarek**, The substitute resistances, capacitances and inductances of some fractal networks 143–156

The issues **60***, no.* 3 *and* **61***, no.* 1 *of the journal are dedicated to Professor* ROMAN STANISŁAW INGARDEN*, an outstanding physicist, good friend and teacher of many of us, cheerful and warmhearted person, on the occasion of his ninetieth birthday* (October 1, 2010)

Right to left: Osamu Suzuki (Tokyo), Roman Stanisław Ingarden (Toruń), and Julian Ławrynowicz (Łódź) in the front of shogun's palace (Kyoto)

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 9–21

Dedicated to Professor Roman Stanisław Ingarden
on the occasion of his ninetieth birthday

Andrzej Jamiołkowski and Miłosz Michalski

QUANTUM INFORMATION THEORY — TORUŃ SCHOOL

Summary

Roman S. Ingarden is one of the founders of quantum information theory (QIT). On the one hand QIT can be regarded as a branch of modern quantum mechanics, while on the other it is a generalization of classical (Shannon) information theory founded on noncommutative probability theory. The aim of this paper is to present the role of R. S. Ingarden and a group of his collaborators in the creation and subsequent development of fundamental ideas of QIT.

1. Beginnings

Almost forty years ago Professor R. S. Ingarden published a seminal paper entitled “Quantum Information Theory”, [5], in the introduction of which he wrote:

> *A conceptual analysis of the classical information theory of Shannon* (1948) *shows that this theory cannot be directly generalized to the usual quantum case. The reason is that in the usual quantum mechanics of closed systems there is no general concept of joint and conditional probability. Using, however, the generalized quantum mechanics of open systems* (A. Kossakowski, 1972) *and the generalized concept of observable* (“semiobservable” by E. B. Davies and J. T. Lewis, 1970) *it is possible to construct a quantum information theory being then a straightforward generalization of Shannon's theory.*

In fact, the foundations of such a new approach to QIT and to quantum mechanics of open systems were laid from two independent sides. In 1961, E. C. G.Sudarshan with collaborators announced a paper [20] where for the first time they formulated the ideas of the so-called "quantum stochastic dynamics". Somewhat later, the idea of a non-Hamiltonian quantum statistical theory, whose central notions are an open quantum system and its evolution governed by a "dynamical semigroup", was born and developed in Toruń by the members of Professor Ingarden's group (cf. A. Kossakowski [11]). Immediately emerging in this context is the fundamental problem of the preservation of basic properties of quantum states by such a semigroup. Formally, the defining properties of density matrices representing quantum states, that is positive semidefiniteness and trace being equal to 1, should be preserved during time evolution. This problem was also studied around that time by the Toruń group (cf. A. Jamiołkowski [7]). Let us sketch the above ideas in a bit more detail.

The central object in the theory of classical statistical mechanics is the N-body distribution function $\varrho(t, x)$, where x represents the "phase point" for the entire N-body system. Similarly, in the quantum case an appropriate density matrix $\varrho(t)$ represents the state of the system at time t.

The dynamics of a finite isolated quantum system is usually described by a one-parameter group of unitary transformations in a complex Hilbert space and its action on the system initial state $\varrho(0)$. However, in less idealized physical situations, it is necessary to consider a quantum system as an *open* one, taking into account its interaction with surroundings. Namely, the respective Hilbert space is assumed to be composed of two parts — system and environment, $\mathcal{H} = \mathcal{H}_{\rm s} \otimes \mathcal{H}_{\rm env}$, and the Hamiltonian is taken to consists of 3 terms,

$$H \;=\; H_{\rm s} + H_{\rm int} + H_{\rm env}\,,$$

where $H_{\rm s}$ describes the system itself, $H_{\rm env}$ rules the free evolution of the environment and $H_{\rm int}$ introduces interaction between the two parts. As a whole, such a system can be considered isolated, and hence its dynamics is fully described by a unitary group, $\varrho(t) = U(t)\varrho(0)U^{\dagger}(t)$. This picture, however, is totally impractical due to the enormous number of environmental degrees of freedom entering the model. The usual way of handling this problem is to resort to the *reduced* dynamics, i.e. to average out the environment in $\varrho(t)$, $\varrho_{\rm s}(t) = {\rm Tr}_{\rm env}(\varrho(t))$. Then the notion of a *quantum dynamical semigroup* proves to be useful in more direct description of the reduced dynamics of an open system or, at least, of some of its aspects (cf. e.g. [6, 11]). Thus in contrast to the idealized case of a strictly isolated system with Hamiltonian time evolution prescribed in the form of a unitary group generated by H, the evolution of an open quantum system is governed by a properly selected dynamical semigroup $\Lambda(t)$ characterized by its generator L. We will look closer at dynamical semigroups and their generators in Section 2.

Obviously, an open system needs no longer be conservative: energy can flow back and forth between the system and its environment. So if one considers the Cauchy problem in such context, then some part of the dissipation is due to the reduced description of $\varrho(0)$ (in the sense of Ingarden, see e.g. [18]), and some other part is due to the element of randomness in the dynamics described by the generator L.

Ingarden's "level of description" is fixed by choosing a list of random variables (observables) $Q_1, \dots, Q_n$ among which are the "sure" function I and the energy H. The larger n is, the more detailed is our description. If pure states of our system are described by elements of an N-dimensional Hilbert space (e.g. if we discuss quantum networks), then a set of independent operators I, $Q_1, \dots, Q_n$, where $n = N^2 - 1$ is said to be a complete set of observables. If $n < N^2 - 1$ then one speaks about an incomplete set of observables. It is obvious that a measurement of mean values of a complete set of observables fixes uniquely the state $\varrho(0)$ of our system (assuming that the operators I, $Q_1, \dots, Q_n$ are linearly independent). Note however that it is not possible to use, for instance, I, Q, $Q^2, \dots, Q^{N^2-1}$ for a fixed Q, as a complete set of observables ([6] p. 124) because in this case, according to the Hamilton–Cayley theorem, these operators must be linearly dependent.

Let us turn now to the information-theoretic aspect of open systems. If we adapt the view that statistical mechanics is the study of mechanical systems (classical or quantal) in terms of incomplete information, we are led to regard ϱ as an object representing the available knowledge about the system in question. Loss or gain of information is recognized physically as the gain or loss of entropy — a special function of ϱ. For a given statistical operator ϱ we determine its entropy by the von Neumann formula $s(\varrho) = -\mathrm{Tr}(\varrho \ln \varrho)$. Now, if we want to reverse this procedure and use certain physical properties of the system (e.g. mean values $E(Q_i)$ of a set of observables $Q_1, \dots, Q_n$) to construct a *unique* statistical operator describing the system state, we have to introduce the so-called principle of maximum entropy which, in quantum statistical physics, is an information-theoretical estimation principle (decision rule). This principle was formulated independently by E. T. Jaynes and R. S. Ingarden (cf. e.g. [6, 18]).

In the next section we discuss the main ideas underlying the semigroup description of open quantum system dynamics.

2. Dynamical semigroups

Time evolution of an open quantum system of finitely many degrees of freedom coupled to an infinite system, usually called a reservoir, can be described by a one-parameter semigroup of transformations [6, 19]. We shall define the semigroup accordingly.

Let $\mathcal{H}$ be the Hilbert space of the system in question ($\dim \mathcal{H} = N < \infty$). Let us denote by $\mathcal{T}(\mathcal{H})$ the real Banach space of self-adjoint operators on $\mathcal{H}$ under the trace norm $\|\varrho\| := \mathrm{Tr}(\varrho^* \varrho)^{1/2}$. In this finite-dimensional setting $\mathcal{T}(\mathcal{H})$ contains in fact all self-adjoint operators acting on $\mathcal{H}$. States of the system are described by density operators $\varrho \in \mathcal{P}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H})$, where the set $\mathcal{P}(\mathcal{H})$ is defined as

$$\mathcal{P}(\mathcal{H}) := \{\varrho \in \mathcal{T}(\mathcal{H}); \quad \varrho \geq 0,\ \mathrm{Tr}\,\varrho = 1\}.$$

Let us note that the smallest linear space in which $\mathcal{P}(\mathcal{H})$ can be embedded is just the real Banach space $\mathcal{T}(\mathcal{H})$.

The set of semipositive operators $\varrho \in \mathcal{T}(\mathcal{H})$ constitutes a positive cone $V^+(\mathcal{H})$ in $\mathcal{T}(\mathcal{H})$. Throughout the paper we shall use the terms "positive" and "semipositive" for brevity in place of formally more appropriate "positive definite" and "positive semidefinite" referring to linear operators on $\mathcal{H}$. This cone can be also defined as:

$$V^+(\mathcal{H}) := \{\varrho \in \mathcal{T}(\mathcal{H});\quad \|\varrho\| = \mathrm{Tr}\varrho\},$$

because $\varrho \in V^+(\mathcal{H})$ if and only if the equality $\|\varrho\| = \mathrm{Tr}\varrho$ is fulfilled.

Definition 1. A family $\{\Lambda(t),\ t \in \mathbb{R}^1_+\}$ of linear mappings

$$\Lambda(t) : \mathcal{T}(\mathcal{H}) \longrightarrow \mathcal{T}(\mathcal{H})$$

constitutes a *dynamical semigroup of a quantum system* $\mathcal{S}$ iff

1) $\Lambda(t) : V^+(\mathcal{H}) \longrightarrow V^+(\mathcal{H})$ for all $t \in \mathbb{R}^1_+$,
2) $\|\Lambda(t)\varrho\| = \|\varrho\|$ for all $\varrho \in V^+(\mathcal{H})$,
3) $\Lambda(t)\Lambda(s) = \Lambda(t+s)$ for all $t, s \in \mathbb{R}^1_+$,
4) $\lim_{t\to 0} \Lambda(t) = I$ (I – the identity operator in $\mathcal{T}(\mathcal{H})$).

The limit in the latter equality should be understood as the limit in the norm $\|\cdot\|$ in $\mathcal{T}(\mathcal{H})$.

The meaning of conditions 1) and 2) in the above definition is that for all $t \in \mathbb{R}^1_+$, $\Lambda(t) : \mathcal{P}(\mathcal{H}) \longrightarrow \mathcal{P}(\mathcal{H})$. Condition 3) ensures that the family $\{\Lambda(t),\ t \in \mathbb{R}^1_+\}$ constitutes a semigroup, whereas condition 4) is dictated by the requirement that for all observables $Q \in \mathcal{T}^*(\mathcal{H})$ and $\varrho \in \mathcal{T}(\mathcal{H})$ the equality $\lim_{t\to 0} \mathrm{Tr}(Q(\Lambda(t)\varrho - \varrho)) = 0$ is fulfilled.

It can be shown (cf. [6, 11]) that the family $\{\Lambda(t),\ t \in \mathbb{R}^1_+\}$ of linear mappings $\Lambda(t) : \mathcal{T}(\mathcal{H}) \longrightarrow \mathcal{T}(\mathcal{H})$ constitutes the dynamical semigroup of a quantum system $\mathcal{S}$ if and only if for all $t, s \in \mathbb{R}^1_+$

1° $\mathrm{Tr}(\Lambda(t)\varrho) = \mathrm{Tr}\varrho$ for all $\varrho \in \mathcal{T}(\mathcal{H})$,
2° $\|\Lambda(t)\varrho\| \leq \|\varrho\|$ for all $\varrho \in \mathcal{T}(\mathcal{H})$,
3° $\Lambda(t) \circ \Lambda(s) = \Lambda(t \cdot s)$,
4° $\lim_{t\to 0} \Lambda(t) = I$.

The above theorem essentially states the equivalence of the first two conditions of the definition with the requirements 1° and 2°. These conditions in the form 1° and 2° refer to the whole space $\mathcal{T}(\mathcal{H})$ and not only to the positive cone $V^+(\mathcal{H})$. As we shall see below, this allows us to introduce the notion of the generator of a dynamical semigroup. On the other hand, the form of these conditions, as given in the definition, enables a straightforward physical interpretation — they simply mean that the mappings $\Lambda(t)$, for $t \in \mathbb{R}^1_+$, transform states into states.

By applying the Hille–Yosida theorem (see e.g. [22]) to the dynamical semigroup $\{\Lambda(t),\ t \in \mathbb{R}^1_+\}$ we infer that there exists a linear operator L acting on the space $\mathcal{T}(\mathcal{H})$, called the generator of $\{\Lambda(t)\}$, such that

$$\frac{d}{dt}(\Lambda(t)\varrho) = L(\Lambda(t)\varrho),$$

for all $\varrho \in \mathcal{T}(\mathcal{H})$. If $\varrho(t)$ denotes the operator $\Lambda(t)\varrho_0$, where $\Lambda(t)$ is an element of the dynamical semigroup, then

$$\mathbb{R}^1_+ \ni t \longmapsto \varrho(t) \in \mathcal{T}(\mathcal{H})$$

is a solution of the differential equation $\dot{\varrho}(t) = L\varrho(t)$ with initial condition $\varrho(0) = \varrho_0 \in \mathcal{T}(\mathcal{H})$. In other words, if $\dim \mathcal{H} = N < \infty$, then every family of stochastic mappings $\Lambda(t) : \mathcal{P}(\mathcal{H}) \longrightarrow \mathcal{P}(\mathcal{H})$, satisfying the conditions given in Definition 1, can be represented in the form $\Lambda(t) = \exp Lt,\ t \in \mathbb{R}^1_+$.

Let us note that if the inequality 2° above is replaced by equality

$$\|\Lambda(t)\varrho\| = \|\varrho\|$$

for all $\varrho \in \mathcal{T}(\mathcal{H})$ and $t \in \mathbb{R}^1_+$ or, which amounts for the same, if the equality 2) in Definition 1 holding on $V^+(\mathcal{H})$ is extended to the whole space $\mathcal{T}(\mathcal{H})$, then one can already infer the existence of a continuous one-parameter unitary group $\{U(t),\ t \in \mathbb{R}^1\}$ on $\mathcal{H}$ such that

$$\Lambda(t)\varrho = U(t)\varrho U^\dagger(t)$$

for all $\varrho \in \mathcal{T}(\mathcal{H})$ and $t \geq 0$. In other words, the strengthening of 2) allows one technically to extend the semigroup $\{\Lambda(t),\ t \in \mathbb{R}^1_+\}$ to a full group of mappings $\{\Lambda(t),\ t \in \mathbb{R}^1\}$.

Physical intuitions behind the two situations are that the semigroup case corresponds to the presence of dissipation in the system which gives rise to an irreversible dynamics: one cannot trace back into the past of a process whose "current" state is $\varrho(t_0)$. On the other hand, the group structure of the dynamics allowing one to propagate states backwards in time is rooted in conservative character of the system.

In the latter case, Stone's theorem guarantees a spectral representation of the group $\{U(t),\ t \in \mathbb{R}^1\}$,

$$U(t) = \int_{-\infty}^{\infty} \exp(-it\lambda)E(d\lambda).$$

The differential equation for $\varrho(t) = \Lambda(t)\varrho$ assumes then the form

$$\frac{d}{dt}\varrho(t) = L_0\varrho(t) = -i[H, \varrho(t)],$$

where the self-adjoint operator H acting on $\mathcal{H}$ is given by the spectral resolution formula

$$H = \int_{-\infty}^{\infty} \lambda E(d\lambda).$$

In this manner we arrive at the conditions which determine the Hamiltonian description of a quantum system and the evolution equation in the form of the von Neumann equation.

Arguments of physical nature indicate that a semigroup describing the time evolution of an open quantum system should not only be positive and trace preserving but also *completely positive*. Formally, a linear mapping $\Theta : B(\mathcal{H}) \longrightarrow B(\mathcal{H})$ is completely positive iff the tensor product $\mathbb{I}_n \otimes \Theta$ is a positive map on $\mathcal{M}_n \otimes B(\mathcal{H})$ for any natural n. Here $B(\mathcal{H})$ is the set of bounded operators on $\mathcal{H}$, $\mathcal{M}_n$ is the space of complex $n \times n$ matrices and $\mathbb{I}_n$ denotes the identity mapping on this space. Every such map $\mathbb{I}_n \otimes \Theta$ is called an *amplification* of Θ.

To put it in more intuitive terms, the complete positivity of a semigroup $\{\Lambda(t)\}$ guarantees that it will act consistently on the states (preserving their semipositivity) when our system is treated as an autonomous part of a larger one. The amplification $(\mathbb{I}_n \otimes \Lambda)(t)$ of the semigroup advances the state of our system as before while leaving the supplementary part at rest (see also (6–7) below).

Arguments in favour of completely positive semigroups as the foundation of non-Hamiltonian dynamics along with a thorough study of their properties can be found in the papers of Kraus [12], Lindblad [13] and Gorini, Kossakowski, and Sudardshan [3]. In particular, in [3,13] general form of the generator of a completely positive dynamical semigroup was derived. Namely, a linear operator $L : B(\mathcal{H}) \longrightarrow B(\mathcal{H})$ preserving $\mathcal{T}(\mathcal{H})$ proves to be the generator of a proper completely positive dynamical semigroup if and only if it can be represented in the form

$$L\varrho = -i[H, \varrho] + \frac{1}{2} \sum_j \big([V_j \varrho, V_j^*] + [V_j, \varrho V_j^*] \big) \tag{1}$$

where $V_j \in B(\mathcal{H})$ for $j = 1, 2, \ldots$, and $H \in B(\mathcal{H})$ is self-adjoint.

It should be emphasized though that determining the generator of a semigroup alone is insufficient to describe the evolution of the system in question. Equally essential is the ability to decide its initial state at an arbitrary time instant t_0 (we usually assume that $t_0 = 0$). In the presently considered finite-dimensional setting knowledge of the initial state $\varrho(t_0) \in \mathcal{P}(\mathcal{H})$ it suffices to determine the system state at time $t > t_0$ according to the formula

$$\mathbb{R}^1_+ \ni t \longmapsto \varrho(t) = \exp\big(L(t - t_0)\big)\,\varrho(t_0).$$

The initial state $\varrho(t_0)$ can be determined by measuring expectation values or correlation functions of some observables belonging to the space $\mathcal{T}^*(\mathcal{H})$. It is essential thereby whether the accessible information about the system is sufficient to determine its state uniquely. It is worth emphasizing that even if the space $\mathcal{T}(\mathcal{H})$ is finite-dimensional, the two spaces $\mathcal{T}(\mathcal{H})$ and $\mathcal{T}^*(\mathcal{H})$ should not be identified, although they are algebraically isomorphic as spaces of the same dimension, the reason being that the norms in these spaces are introduced differently. If the expectation values of a relevant set of observables $Q_1, \ldots, Q_n$, where $n < N^2 - 1$, are measured at a finite number of time instants $t_1, \ldots, t_s$, then such procedure of determining the state ϱ_0 can be effective only for N-level systems, because in general infinitely many measurements are necessary to describe ϱ_0 entirely. Nonetheless, in this case it is also possible to establish certain conditions which must be fulfilled (and which are sufficient) in order that the state of an open quantum system be determined uniquely. The branch of physics which is concerned with identification of quantum states is called "quantum tomography" and has also been developed in Toruń since 1980's, [8–10].

3. Superoperators which preserve semipositivity

Another important problem formulated by R. S. Ingarden and discussed in Toruń in 1970's was the preservation of semipositivity and trace of states represented by density matrices by the action of a dynamical semigroup. As we have already mentioned it in the previous section, time evolution of a non-isolated (open) quantum system is described by the differential equation in the general form

$$\frac{d\,\varrho(t)}{dt} \;=\; L\,\varrho(t)\,,$$

with the generator L given by (1). Solutions of the above equation can be rewritten in the form

$$\varrho(t) = \Lambda(t)\varrho(0)\,,$$

where $\Lambda(t)$ is a superoperator with respect to the operators $\varrho \in B(\mathcal{H})$.

This representation of time evolution leads directly to the following question: what are the condintions that a superoperator $\Lambda(t)$ must obey in order to preserve semipositivity of density operators ϱ. An answer to this problem was given in 1973 by A. Jamiołkowski [7]. This answer can be formulated as follows.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two finite-dimensional Hilbert spaces with $\dim\mathcal{H}_1 = n$ and $\dim\mathcal{H}_2 = m$. By $(\cdot,\cdot)_i$ we denote respectively the inner product in $\mathcal{H}_i$, $i = 1, 2$. Let $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ be the vector space of linear transformations between $\mathcal{H}_1$ and $\mathcal{H}_2$. We also write simply $\mathcal{L}(\mathcal{H})$ for $\mathcal{L}(\mathcal{H}, \mathcal{H})$. Let moreover $\mathcal{A}_i = \mathcal{A}(\mathcal{H}_i)$ be the full algebra of linear operators on $\mathcal{H}_i$ with inner product $[A, B]_i = \mathrm{Tr}(B^*A)$, $i = 1, 2$. Note that $\mathcal{L}(\mathcal{H})$, $B(\mathcal{H})$ and $\mathcal{A}(\mathcal{H})$ refer to *the same set* of objects, but with *different structure* in mind: an ordinary vector space, a Banach space and an algebra, respectively.

Let $\mathcal{H}_1 \otimes \mathcal{H}_2$ denote the tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$ which, when endowed with inner product of the form

$$((x_1 \otimes y_1, x_2 \otimes y_2)) \;=\; (x_1, x_2)_1 \cdot (y_1, y_2)_2$$

for any $x_i \in \mathcal{H}_1$ and $y_i \in \mathcal{H}_2$, becomes a Hilbert space of its own. Analogously, the tensor product of algebras $\mathcal{A}_1 \otimes \mathcal{A}_2$ is naturally equipped with a unitary space structure by

$$[[A_1 \otimes B_1, A_2 \otimes B_2]] := [A_1, A_2]_1 \cdot [B_1, B_2]_2$$

for all $A_1, A_2 \in \mathcal{A}_1$ and $B_1, B_2 \in \mathcal{A}_2$.

Let us recall the standard fact that the algebras $\mathcal{A}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ and $\mathcal{A}_1 \otimes \mathcal{A}_2$ are isomorphic.

Now following [7], let $\mathcal{J}$ denote the linear transformation which maps the space $\mathcal{L}(\mathcal{A}_1, \mathcal{A}_2)$ to the space $\mathcal{A}_1 \otimes \mathcal{A}_2$,

$$\mathcal{J} : \mathcal{L}(\mathcal{A}_1, \mathcal{A}_2) \longrightarrow \mathcal{A}_1 \otimes \mathcal{A}_2 \,,$$

whose value for arbitrary $\Lambda \in \mathcal{L}(\mathcal{A}_1, \mathcal{A}_2)$ is defined by the equality

$$[[\mathcal{J}(\Lambda), A^* \otimes B]] \; = \; [\Lambda(A), B]_2 \tag{2}$$

for any $A \in \mathcal{A}_1$ and $B \in \mathcal{A}_2$. It can easily be verified that

$$\mathcal{J}(\Lambda) \; = \; \sum_i E_i^* \otimes \Lambda(E_i) \,, \tag{3}$$

for any orthonormal basis $\{E_i\}$ in $\mathcal{A}_1$. In other words, $\mathcal{J}$ is an isomorphism between linear maps from $\mathcal{A}_1$ to $\mathcal{A}_2$ and operators on $\mathcal{H}_1 \otimes \mathcal{H}_2$. In literature this isomorphism if often referred to as the *Jamiołkowski isomorphism* or *Choi-Jamiołkowski isomorphism*. The isomorphism was originally introduced in 1972 by one of the present authors (A. J.) [7] in the physically motivated context of conservation of density matrix properties by open quantum dynamics, while it also emerged a bit later, in 1975, in a purely mathematical study by M. D. Choi, [2]. It easily follows from (2) that $\mathcal{J}$ is, in fact, an isometry. Moreover, in [7] the following relevant properties of $\mathcal{J}$ are demonstrated:

(i) $\Lambda : \mathcal{A}_1 \to \mathcal{A}_2$ preserves hermiticity if and only if its image by $\mathcal{J}$ is Hermitian in $\mathcal{A}_1 \otimes \mathcal{A}_2 = \mathcal{A}(\mathcal{H}_1 \otimes \mathcal{H}_2)$;

(ii) Λ preserves strict positivity of operators if and only if its image by $\mathcal{J}$ is Hermitian and

$$((\mathcal{J}(\Lambda) x \otimes y, x \otimes y)) \; > \; 0 \tag{4}$$

for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$. Observe that the last condition is weaker than ordinary positive definiteness of $\mathcal{J}(\Lambda)$, since it is required to hold only for *product vectors* $x \otimes y$ which form a proper subset of $\mathcal{H}_1 \otimes \mathcal{H}_2$.

(iii) In the same spirit, Λ preserves semipositivity of operators iff its image by $\mathcal{J}$ is Hermitian and (4) holds with "$\geq$" replacing the strict inequality "$>$".

Let us mention that there is an important application of the $\mathcal{J}$ mapping in the theory of entanglement in bipartite quantum systems, namely it establishes an equivalence between positive but not completely positive maps and so-called entanglement witnesses. The two notions are of fundamental importance for the surprisingly involved problem of distinction between separable and entangled states in contemporary quantum information science. We will draw a perspective of these and related issues in the next section.

4. Entanglement, entanglement witnesses and positive maps

If $\mathcal{H}$ is a Hilbert space of a composite quantum system, e.g. $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, then a large part of its elements cannot be represented in the product form $z \neq x \otimes y$

regardless of the basis chosen. Such vectors are called *entangled* and there exists a standard algebraic procedure, called Schmidt decomposition, by means of which one can check whether a given vector is a product or an entangled one. Entanglement is understood as the manifestation of *quantum correlations* between the constituent parts in the system as opposed to much weaker *classical correlations* which can well be encoded in a product or — in alternative terminology — a *separable* state.

These notions of separability and entanglemet are immediately extended to mixed states of a compound quantum system.

Definition 2. A mixed state $\varrho \in \mathcal{P}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is *separable* iff it can be represented as a convex combination of projections onto product vectors,

$$\varrho \ = \ \sum_i p_i \, |e_i \otimes f_i\rangle\langle e_i \otimes f_i| \ = \ \sum_i p_i \, |e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|, \qquad p_i \geq 0, \quad \sum_i p_i = 1 \, ,$$

with $e_i \in \mathcal{H}_1$, $f_i \in \mathcal{H}_2$. Otherwise ϱ is said to be *entangled.*

Let us mention that, in contrast to the pure state situation, where one can test a vector for separability/entanglement by the Schmidt procedure, deciding separability of a mixed state is a very difficult task: in principle one would have to check all possible decompositions of ϱ into projections (note that the spectral resolution of ϱ is merely just one of them) to tell whether it is separable or entangled.

It is well known that if a linear map $\Lambda : \mathcal{A} \to \mathcal{A}$ sends the set $\mathcal{A}_h = \{X \in \mathcal{A}\,;\ X = X^*\}$ of all Hermitian elements of $\mathcal{A}$ into itself, then Λ can be represented in the form

$$\Lambda(X) \ = \ \sum_{i=1}^{\kappa} a_i K_i^* X K_i \, , \tag{5}$$

where $K_i \in \mathcal{A}$, and a_i, $i = 1, \ldots, \kappa$ are real numbers [1, 16]. In general, all maps of the form (5) are hermiticity-preserving. However, the representation (5) is not unique: in general, for a given Λ, there exist many possible representations of such form. The minimal length of Λ is defined to be the smallest κ among all expansions (5) of Λ. If we assume that the operators K_i for $i = 1, \ldots, \kappa$ are linearly independent, then κ in (5) must be minimal.

Recall that a map $\Lambda : \mathcal{A} \to \mathcal{A}$ which preserves the set $\mathcal{A}_h$ of Hermitian elements is called positive if $\Lambda(X) \geq 0$ whenever $X \in \mathcal{A}$ is positive, i.e. $(X\eta, \eta) \geq 0$ for all $\eta \in \mathcal{H}$.

A map Λ is called k-positive if its k-amplification $\Lambda_{(k)} := \mathbb{I}_k \otimes \Lambda$ that is the map

$$\mathbb{I}_k \otimes \Lambda : \mathcal{M}_k(\mathbb{C}) \otimes \mathcal{A} \to \mathcal{M}_k(\mathbb{C}) \otimes \mathcal{A} \tag{6}$$

is positive. $\mathcal{M}_k(\mathbb{C})$ denotes here the set of of all $k \times k$ complex matrices. We can identify $\mathcal{M}_k(\mathbb{C}) \otimes \mathcal{A}$ with the set of all $k \times k$ matrices $\mathcal{M}_k(\mathcal{A})$ with entries in $\mathcal{A}$ and in such notation one can represent $\Lambda_{(k)} : \mathcal{M}_k(\mathcal{A}) \to \mathcal{M}_k(\mathcal{A})$ simply by

(7)
$$\Lambda_{(k)}\begin{pmatrix} & \vdots & \\ \dots & X_{ij} & \dots \\ & \vdots & \end{pmatrix} := \begin{pmatrix} & \vdots & \\ \dots & \Lambda(X_{ij}) & \dots \\ & \vdots & \end{pmatrix}.$$

The map Λ is called *completely positive* if it is k-positive for all $k = 1, 2, \dots$. This terminology goes back to Stinespring [17]. It is well known that for $\mathcal{A} = \mathcal{L}(\mathcal{H})$, where $\mathcal{H}$ denotes an N-dimensional Hilbert space, N-positive maps on $\mathcal{A}$ are already completely positive.

Let us observe that all hermiticity-preserving maps which are not only positive but also completely positive can be written in the form (5) with positive a_i, $i = 1, \dots, \kappa$, i.e. equivalently by

(8)
$$\Lambda(X) := \sum_{i=1}^{\kappa} \widetilde{K}_i^* X \widetilde{K}_i \,,$$

where $\widetilde{K}_i = \sqrt{a_i} K_i$ and $\kappa \leq N^2$. Relation (8) is the so-called Kraus representation of a completely positive map Λ. This representation is very useful in quantum information theory. In particular, completely positive maps are used to describe the so-called quantum operations and quantum channels. In general, any map which is positive but not completely positive can be represented as a difference of two completely positive maps:

(9)
$$\Lambda(X) = \sum_{i=1}^{\kappa_1} K_i^* X K_i - \sum_{j=1}^{\kappa_2} M_j^* X M_j \,,$$

where operators $K_1, \dots, K_{\kappa_1}$, $M_1, \dots, M_{\kappa_2}$ are linearly independent and

$$\kappa = \kappa_1 + \kappa_2$$

denotes the minimal length of Λ.

In Section 2 we have stressed physical importance of complete positivity in connection with dynamical semigroup actions. Namely, the complete positivity of a superoperator guarantees that it maps states of a quantum system again to legitimate physical states regardless of the way the system is immersed in its surroundings. We have formulated it in the language of appropriate amplifications.

It turns out that allegedly nonphysical maps which are positive, or k-positive, but not completely positive, are also relevant for quantum physics. It is due to the fact that they provide theoretical tools allowing one to distinguish separable and entangled states of a bipartite quantum system. This is characterized by the famous Peres-Horodecki theorem [4, 15]: if $\varrho \in \mathcal{P}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is a mixed state of a bipartite quantum system then ϱ is separable iff for every positive map $\Lambda : \mathcal{A}(\mathcal{H}_2) \to \mathcal{A}(\mathcal{H}_2)$ the matrix $(\mathbb{I} \otimes \Lambda)\varrho$ is semipositive (here $\mathbb{I}$ denotes the identity map on $\mathcal{A}(\mathcal{H}_1)$).

The image of ϱ under $\mathbb{I} \otimes \Lambda$ is automatically semipositive if Λ is completely positive, but this condition may fail for some positive Λ: then ϱ is necessarily entangled.

An immediate example is provided by the transposition map $\Lambda = T$. Then $(\mathbb{I} \otimes T)\varrho$ is simply the partial transpose of ϱ with respect to the second subsystem, ϱ^{T_2}. So if such a partial transpose has a negative eigenvalue, then ϱ is entangled. Observe that semipositivity of the partial transpose alone is only necessary for the separability of ϱ. It happens to be also the sufficient condition in low-dimensional systems, i.e. 2×2 or 2×3. This fact is a direct consequence of particularly simple structure of low-dimensional positive maps which, in this case, all turn out to be *decomposable*, i.e. they can be represented in the form

$$\Lambda \;=\; \Lambda_1 + \Lambda_2 \circ T$$

with Λ_i being completely positive, $i = 1, 2$. In higher dimensions not all positive maps are decomposable and hence the transposition no longer plays such a distinguished role. Characterization of the structure of positive indecomposable maps has been a notoriously hard problem of contemporary mathematics.

An alternative method for the detection of entangled states is based on the so-called entanglement witnesses. By definition, a Hermitian operator $W \in L(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is an entanglement witness if it fulfils the following conditions:

(i) $((x \otimes y, Wx \otimes y)) \geq 0$ for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$,

(ii) $((\eta, W\eta)) < 0$ for some $\eta \in \mathcal{H}_1 \otimes \mathcal{H}_2$.

That is, an entanglement witness is not a semipositive operator (i.e. has a negative eigenvalue) but it is positive when restricted to product states (in quantum-information terminology: on separable pure states) or, using other terminology, it is block-positive.

From an experimentalist's point of view, entanglement witness is a nonlocal (in the sense that it extends over both parts of our system) observable whose expectation value, when measured in a state ϱ, i.e. the quantity $\mathrm{Tr}(\varrho W)$, can serve as a direct indicator of the entanglement present in ϱ. Often one can make use of "true" physical observables, like e.g. energy, which are relatively easy to measure in experiments. Appropriate techniques related to the spectral properties of such observables allow one to convert them into entanglement witnesses, see e.g. [14]. It is so far the best available way of detecting entanglement in laboratory experiments.

The relation between entanglement witnesses and positive but not completely positive maps giving rise to separability criteria is provided by the $\mathcal{J}$ isomorphism discussed in the previous section. Namely, from its properties it follows that if $((\eta, W\eta)) \geq 0$ for all $\eta \in \mathcal{H}_1 \otimes \mathcal{H}_2$, then $\Lambda = \mathcal{J}^{-1}(W)$ is completely positive. It means that positive maps which are not completely positive, that is ones having the form (9), are mapped to observables which are entanglement witnesses.

Let us conclude this section with an example of the application of $\mathcal{J}$ isomorphism. Suppose for simplicity that $\mathcal{H}_1 = \mathcal{H}_2 =: \mathcal{H}$ and let $\{e_i\}$, $i = 1, \ldots, n$, be its fixed orthonormal basis. For $i, j = 1, \ldots, n$ let E_{ij} be the operator defined by $E_{ij}e_j = e_i$ and $E_{ij}e_k = 0$ if $k \neq j$. The Hermitian operator on $\mathcal{H} \otimes \mathcal{H}$

$$V = \sum_{ij=1}^{n} E_{ij} \otimes E_{ji}$$

is called the *swap operator* in quantum information literature. Then using (3) V can be identified with the $\mathcal{J}$ image of the transposition T on $B(\mathcal{H})$. Indeed V is an entanglement witness: on the one hand we have

$$((x \otimes y, Vx \otimes y)) = |(x, y)|^2 \geq 0,$$

while on the other hand

$$V\psi = -\psi$$

for any antisymmetric vector ψ, so that V has eigenvalue -1.

5. Conclusion

Quantum information theory is today one of the most promising parts of physics. Enormous technological progress of the recent two decades has opened new practical applications of subtle quantum phenomena, like entanglement. Entanglement proves particularly useful in quantum cryptography, communication and information processing and is the very agent giving quantum technologies an advantage over their classical counterparts: security, efficiency and speed.

Forty years ago Professor R. S. Ingarden and his collaborators in Toruń outlined the programme of research whose goal was to incorporate information theoretic ideas and techniques into quantum physics. Its importance and meaning back then was purely theoretical. The results of this research prove even more important today in the context of developing quantum technologies.

References

[1] I. Bengtsson, K. Życzkowski, *Geometry of Quantum States*, Cambridge Univ. Press, Cambridge 2008.

[2] M. D. Choi, Lin. Alg. Appl. **12** (1975), 285.

[3] V. Gorini, A. Kossakowski, E. C. G. Sudarshan, J. Math. Phys. **17** (1976), 149.

[4] R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, Rev. Mod. Phys. **81** (2009), 865.

[5] R. S. Ingarden, Rep. Math. Phys. **10** (1976), 43.

[6] R. S. Ingarden, A. Kossakowski, and M. Ohya, *Information Dynamics and Open Systems*, Kluwer, Dordrecht 1997.

[7] A. Jamiołkowski, Rep. Math. Phys. **3** (1972), 273.

[8] A. Jamiołkowski, Intern. J. Theor. Phys. **22** (1983), 369.

[9] A. Jamiołkowski, Rep. Math. Phys. **21** (1985), 101.

[10] A. Jamiołkowski, Open Sys. Inf. Dyn. **11** (2004), 63.

[11] A. Kossakowski, Rep. Math. Phys. **3** (1972), 247.

[12] K. Kraus, Ann. of Phys. **64** (1971), 311.

[13] G. Lindblad, Comm. Math. Phys. **48** (1976), 119.

[14] M. Michalski, in: *Quantum Bio-Informatics III*, L. Accardi, W. Freudenberg, M. Ohya, eds., Quantum Probability and White Noise Analysis XXVI, p. 217, World Scientific, Singapore 2009.

[15] A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer, Dordrecht 1993.

[16] J. de Pilis, Pacific J. Math. **23** (1967), 129.

[17] W. F. Stinespring, Proc. AMS **6** (1955), 211.

[18] R. F. Streater, Contemporary Mathematics **203** (1997), 117.

[19] R. F. Streater, J. Math. Phys. **41** (2000), 3556.

[20] E. C. G. Sudarshan, et al., Phys. Rev. **121** (1961), 920.

[21] B. Terhal, Phys. Lett. A **271** (2000), 319.

[22] K. Yosida, *Functional Analysis*, Springer, Berlin 1978.

Institute of Physics
Nicolaus Copernicus University
Grudziądzka 5, PL-87-100 Toruń
Poland
e-mail: romp@phys.uni.torun.pl

Presented by Jakub Rembieliński at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on September 29, 2011

KWANTOWA TEORIA INFORMACJI – SZKOŁA TORUŃSKA

S t r e s z c z e n i e

Roman S. Ingarden jest jednym z twórców kwantowej teorii informacji (QIT). Z jednej strony QIT może być uważana za gałąź współczesnej mechaniki kwantowej, z drugiej zaś stanowi ona także uogólnienie klasycznej, shannonowskiej teorii informacji w oparciu o niekomutatywną teorię prawdopodobieństwa. Niniejsza praca stawia sobie za cel przybliżenie roli, jaką odegrali Profesor Ingarden i zespół jego współpracowników w tworzeniu, a następnie w rozwijaniu podstaw kwantowej teorii informacji.

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 23–36

Dedicated to Professor Roman Stanisław Ingarden
on the occasion of his ninetieth birthday

Massimo Vaccaro

BASICS OF LINEAR PARA-QUATERNIONIC GEOMETRY I

HERMITIAN PARA-TYPE STRUCTURE ON A REAL VECTOR SPACE

Summary In the present Part I of the paper we describe para-type structures on a real vector space and on a pseudo-Euclidean vector space. In Part II we shall investigate relevant classes of subspaces of a para-quaternionic Hermitian vector space, in particular the decomposition of a generic subspace. This article deals with basic definitions and results in para-quaternionic geometry. The standard para-quaternionic structure $\widetilde{Q}$ on the tensor product $H \otimes E$ of a pair of real vector spaces of dimension 2 and $2n$ respectively is defined as the Lie algebra $\widetilde{Q} = \mathfrak{sl}(H)$ of the special linear group $\mathrm{SL}(H)$ of volume preserving automorphisms on H. Any para-quaternionic vector space $(V, \widetilde{Q})$ is isomorphic to $(H^2 \otimes E^n, \mathfrak{sl}(H))$. Furthermore if (H, ω^H) and (E, ω^E) are symplectic spaces, the 2-form $\omega^H \otimes \omega^E$ defines a $\widetilde{Q}$-Hermitian metric on $(H^2 \otimes E^n, \mathfrak{sl}(H))$ and any Hermitian para-quaternionic vector space $(V, \widetilde{Q}, g)$ is isomorphic to $(H^2 \otimes E^{2n}, \mathfrak{sl}(H), \omega^H \otimes \omega^E)$.

Introduction

In Section 1 we recall the (pairwise dual) definitions of (para)-complex, (para)-hypercomplex and (para)-quaternionic structures. A para-hypercomplex structure $\{I, J, K\}$ on a real vector space V is a left module structure over the Clifford algebra of para-quaternions which is the real algebra $\widetilde{\mathbb{H}}$ generated by unity 1 and imaginary units i, j, k satisfying

$$-i^2 = j^2 = k^2 = 1, \quad ij = -ji = k. \tag{1}$$

For a para-quaternionic structure $\widetilde{Q}$ the left module structure is defined up to coniugation in $\widetilde{\mathbb{H}}$.

The real algebra of para-quaternions is isomorphic to $\mathrm{Mat}(2,\mathbb{R})$ (3), then from Wedderburn theorem it follows (Proposition 1.1) that any vector space V with a para-hypercomplex or para-quaternionic structure is the direct sum of 2-dimensional irreducible components; this implies that $\dim V = 2n$. This represents one difference respect to a hypercomplex and quaternionic vector space whose dimension is necessarily a multiple of 4. Other differences together with some analogies are listed afterwards.

In Section 2 we consider a real vector space V endowed with a pseudo Euclidean scalar product. A para-complex (resp. para-hypercomplex, resp. para-quaternionic) structure preserving (in the sense of Lie algebra) such a metric is called a *para-Hermitian* (resp. *Hermitian para-hypercomplex*, resp. *Hermitian para-quaternionic*) *structure*. The eigenspaces of a para-complex structure are totally isotropic, then an Hermitian metric is always neutral.

There exists a one to one correspondence between a para-Hermitian structure (g,K) on V, a pseudo Euclidean vector space (V,g) with a decomposition into a pair of totally isotropic subspaces, or equivalently a symplectic vector space (V,ω) with a bi-Lagrangian decomposition (Proposition 2.3).

The dimension of a Hermitian para-hypercomplex and Hermitian para-quaternionic vector space is $4n$ (Proposition 2.4).

The prototype of a para-hypercomplex Hermitian vector spaces is the n-dimensional para-quaternionic numerical space $\widetilde{\mathbb{H}}^n$ which is a real vector space of dimension $4n$, a $\widetilde{\mathbb{H}}$-module with respect to left multiplication by para-quaternions and is endowed with the canonical Hermitian product

$$h \cdot h' = \sum_{\alpha=1}^{n} h_\alpha \overline{h'_\alpha}\,; \quad h = (h_1, \dots, h_n), \;\; h' = (h'_1, \dots, h'_n) \in \widetilde{\mathbb{H}}^n.$$

The real part of the Hermitian product defines a pseudo-Euclidean (canonical) scalar product of neutral signature on the real vector space $\widetilde{\mathbb{H}}^n \simeq \mathbb{R}^{2n,2n}$. Moreover, left multiplications by i, j, k, respectively, induce real endomorphisms of $\widetilde{\mathbb{H}}^n$ satisfying (1) and skew-symmetric with respect to the metric. As an $\widetilde{\mathbb{H}}$-module, on a para-hypercomplex Hermitian vector space $(V^{4n}, \{I,J,K\}, g)$ we define the ($\widetilde{\mathbb{H}}$-valued)-Hermitian product $(\cdot) = (\cdot)_{\{I,J,K\}}$ by:

$$\begin{array}{rrcl} (\cdot): & V \times V & \to & \widetilde{\mathbb{H}}, \\ & (X,Y) & \mapsto & X \cdot Y = g(X,Y) + ig(X,IY) - jg(X,JY) - kg(X,KY). \end{array}$$

On a para-quaternionic Hermitian vector space $(V^{4n}, \widetilde{Q}, g)$, by using admissible para-hypercomplex bases $\{I,J,K\}$ of $\widetilde{Q}$, the ($\widetilde{\mathbb{H}}$-valued)-Hermitian product is defined up to inner automorphisms of $\widetilde{\mathbb{H}}$.

A tensor product $H \otimes E$ of two real vector spaces of dimension 2 and $2n$ respectively carries a *standard para-quaternionic structure* $\widetilde{Q}$ which can be identified with the Lie algebra $\mathfrak{sl}(H)$ of the special linear group $\mathrm{SL}(H)$ of volume preserving linear operators on H. Any para-quaternionic vector space $(V, \widetilde{Q})$ is isomorphic to $(H^2 \otimes E^n, \mathfrak{sl}(H))$, (Corollary (1.2)). Moreover if ω^H and ω^E are symplectic forms on H and E respectively, the 2-form $\omega^H \otimes \omega^E$ is an Hermitian metric on $(H \otimes E, \mathfrak{sl}(H))$ and we call $(\mathfrak{sl}(H), \omega^H \otimes \omega^E)$ a *standard Hermitian para-quaternionic structure on* $H \otimes E$. In Proposition (2.5) we prove that any Hermitian para-quaternionic vector space $(V, \widetilde{Q}, g)$ is isomorphic to some $(H^2 \otimes E^{2n}, \mathfrak{sl}(H), \omega^H \otimes \omega^E)$.

In the second part [8], by referring to the tensorial presentation of a para-quaternionic Hermitian vector space

$$(V, \widetilde{Q}, g) \simeq (H^2 \otimes E^{2n}, \mathfrak{sl}(H), \omega^H \otimes \omega^E),$$

we will characterize some relevant classes of subspaces of V defined with respect to the structure group of $\widetilde{Q}$ and $(\widetilde{Q}, g)$, respectively, and give a "para-quaternionic decomposition" of any vector subspace of V.

1. Para-type structures on a vector space

We first recall the definitions of some well known structures in complex and quaternionic geometry [1, 2].

Definition 1.1. Let V be a real vector space.

A *complex structure on* V^{2n} is an endomorphism $J \in \mathrm{End}(V)$ such that $J^2 = -\mathrm{Id}$.

A *hypercomplex structure* H *on* V^{4n} is a triple $(J_\alpha) = (J_1, J_2, J_3)$ of anti-commuting complex structures on V satisfying $J_1 J_2 = J_3$; it defines on V the structure of left vector space over quaternions $\mathbb{H} = \mathrm{span}_{\mathbb{R}}\{1, i, j, k\}$ such that multiplications by i, j and k are given by J_1, J_2 and J_3.

A *quaternionic structure on* V^{4n} is the 3-dimensional subspace $Q \subset \mathrm{End}(V)$ spanned by a hypercomplex structure H i.e. $Q = \mathrm{span}_{\mathbb{R}}\{J_1, J_2, J_3\}$. We say that the hypercomplex structure H is *subordinate* to the quaternionic structure Q or equivalently that it is an *admissible basis* of Q.

Note $Q \subset \mathfrak{gl}(V)$ is a Lie subalgebra isomorphic to $\mathfrak{sp}(1) \cong \mathrm{Im}\,\mathbb{H} = \mathrm{span}(i, j, k)$, the Lie algebra of the Lie group $\mathrm{Sp}(1) = S^3 \subset \mathbb{H} = \mathrm{span}\{1, i, j, k\}$ of unit quaternions.

Definition 1.2. A Euclidean scalar product $g =< \cdot, \cdot >$ on (V, J) (resp. (V, H), (V, Q)) is called *J-Hermitian* (resp. *H-Hermitian*, *Q-Hermitian*) if J (resp. H, Q) is a skew-symmetric endomorphism (resp. consists of skew-symmetric endomorphisms) of $(V, < \cdot, \cdot >)$. A vector space V endowed with a complex structure J (resp. hypercomplex structure H, quaternionic structure Q) and an Hermitian scalar product g is called an *Hermitian vector space* (V, J, g) (resp. *hypercomplex Hermitian vector space* (V, H, g), *quaternionic Hermitian vector space* (V, Q, g)).

Let us give now the corresponding definitions for para-geometry.

Definition 1.3. Let V be a real vector space of dimension n and $K \in \mathrm{End}(V)$ such that $K^2 = \mathrm{Id}$. Let denote V_K^+ and V_K^- the $+1$ and -1 eigenspaces of K. Then K is called a *product structure on* V if $\dim V_K^+, \dim V_K^- > 0$. A *para-complex structure on* V is a product structure with $\dim V_K^+ = \dim V_K^-$.

A triple (J_1, J_2, J_3) of anticommuting endomorphisms of V satisfying the relations:

$$-J_1^2 = J_2^2 = J_3^2 = \mathrm{Id}, \quad J_1 J_2 = J_3 \tag{2}$$

is called a *para-hypercomplex structure* on V. Observe that (J_1 is a complex structure and) J_2 and J_3 are para-complex structures on V. In fact, since J_1 and J_2 anticommute, $J_1(V_{J_2}^+) \subseteq V_{J_2}^-$ and $J_1(V_{J_2}^-) \subseteq V_{J_2}^+$, which implies $\dim V_{J_2}^+ = \dim V_{J_2}^-$, and analogously for J_3. A Lie subalgebra $\widetilde{Q} \subset \mathfrak{gl}(V)$ is called a *para-quaternionic structure* on V if there exists a basis J_1, J_2, J_3 satisfying the relations (2). Such a para-hypercomplex structure is called an *admissible basis* of $\widetilde{Q}$.

A para-hypercomplex structure (J_1, J_2, J_3) defines on V the structure of a left module over the Clifford algebra of para-quaternions $\widetilde{\mathbb{H}}$ [9] which is the real algebra generated by unity 1 and generators i, j, k satisfying

$$-i^2 = j^2 = k^2 = 1, \quad ij = -ji = k.$$

We recall the isomorphisms

$$\widetilde{\mathbb{H}} := \mathrm{Cl}_{1,1}(\mathbb{R}) =< 1, e_1 = i, e_2 = j, e_3 = e_1 e_2 = -e_2 e_1 = k >$$

or equivalently

$$\widetilde{\mathbb{H}} := \mathrm{Cl}_{2,0}(\mathbb{R}) =< 1, e_1 = j, e_2 = k, e_3 = e_1 e_2 = -e_2 e_1 = -i >,$$

and also that $\widetilde{\mathbb{H}}$ is isomorphic, as a real algebra, to the algebra $\mathrm{Mat}_2(\mathbb{R})$ of real (2×2)-matrices, the isomorphism being given by

$$\Phi : \mathbf{q} = q_0 + q_1 i + q_2 j + q_3 k \mapsto \begin{pmatrix} q_0 - q_3 & q_2 - q_1 \\ q_2 + q_1 & q_0 + q_3 \end{pmatrix}, \tag{3}$$

where $\mathcal{N}(q) := q\bar{q} = q_0^2 + q_1^2 - q_2^2 - q_3^2 = \det(\Phi(\mathbf{q}))$.

A basic example of a para-hypercomplex structure is the *standard para-hypercomplex structure* $(\mathcal{I}, \mathcal{J}, \mathcal{K})$ *of* $\mathbb{R}^2$ represented, in the canonical basis, by

$$\mathcal{I} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \mathcal{J} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathcal{K} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4}$$

Observe that $< \mathcal{I}, \mathcal{J}, \mathcal{K} >_{\mathbb{R}} \simeq \mathfrak{sl}_2(\mathbb{R})$ is the matrix Lie algebra of (2×2)-matrices of zero trace of the unimodular Lie group $SL_2(R)$ of matrices preserving any volume form of $\mathbb{R}^2$.

Generalizing, we define the *standard para-hypercomplex structure* $\widetilde{H} = (I, J, K)$ *of* R^{2n} represented, in the canonical basis, by

$$I = \mathcal{I} \oplus \mathcal{I} \oplus \ldots \oplus \mathcal{I}; \quad J = \mathcal{J} \oplus \mathcal{J} \oplus \ldots \oplus \mathcal{J}; \quad K = \mathcal{K} \oplus \mathcal{K} \oplus \ldots \oplus \mathcal{K} \tag{5}$$

with $\mathcal{I}, \mathcal{J}, \mathcal{K}$ given in (4).

By the identification $\widetilde{\mathbb{H}} \cong \mathrm{Mat}_2(\mathbb{R})$ given in (3), and from Wedderburn theorem, stating that every representation of a unitary, associative, semisimple algebras is direct sum of standard representations, we can than affirm that

Proposition 1.1.

- *There exists a unique, up to isomorphism, irreducible $\widetilde{\mathbb{H}}$-module $H^2 \simeq \mathbb{R}^2$.*
- *Every $\widetilde{\mathbb{H}}$-module V^{2n} is reducible as a direct sum $V = H^2 \oplus \ldots \oplus H^2$.*

Note that to have a direct sum decomposition of the $\widetilde{\mathbb{H}}$-module $(V^{2n}, \widetilde{H} = \{I, J, K\})$, into invariant 2-dimensional subspaces $U_1, \ldots, U_n$, let consider a basis e_i^+ of V_J^+, eigenspace of the para-complex structure J associated to the eigenvalue 1 (then $Ke_1^+, \ldots, Ke_n^+$ is a basis of V_J^-). The 2-dimensional subspaces

$$U_i =< e_i^+, Ke_i^+ >, \qquad i = 1, \ldots, n \tag{6}$$

are clearly $\widetilde{H}$-invariant, irreducible and isomorphic as $\widetilde{\mathbb{H}}$-modules. Choosing the basis $< e_i^+ - Ke_i^+, e_i^+ + Ke_i^+ >$ in each U_i, $\widetilde{H}$ corresponds to the standard para-hypercomplex structure of R^{2n} given in (5).

Let H^2 and E^n be real vector spaces. For any fixed basis (h_1, h_2) of H, one has the identification $H \simeq \mathbb{R}^2$: we define a corresponding *standard para-hypercomplex structure* $\{I, J, K\}$ *on* $H^2 \otimes E^n$ by

$$I = I(h \otimes e) = \mathcal{I}h \otimes e, \quad J = J(h \otimes e) = \mathcal{J}h \otimes e, \quad K = K(h \otimes e) = \mathcal{K}h \otimes e \tag{7}$$

with $\mathcal{I}, \mathcal{J}, \mathcal{K}$ given in (4) and the *standard para-quaternionic structure* $\mathfrak{sl}_2(\mathbb{R}) \otimes \mathrm{Id}$ *on* $H^2 \otimes E^n$ generated by any standard para-hypercomplex structure.

Since $\mathfrak{sl}_2(\mathbb{R}) \simeq \mathfrak{sl}(H)$, the Lie algebra of the Lie group $SL(H)$ of unimodular transformations of H, we will use the equivalent notation

$$\mathfrak{sl}_2(\mathbb{R}) \otimes \mathrm{Id} \simeq \mathfrak{sl}_2(\mathbb{R}) \simeq \mathfrak{sl}(H).$$

For any basis $\{e_1, \ldots, e_n\}$ in E^n, any standard para-hypercomplex structure on $H^2 \otimes E^n$ associated to the basis $\{h_1, h_2\}$ of H is represented in the basis $\{h_1 \otimes e_i, h_2 \otimes e_i,\ i = 1, \ldots, n\}$ by (5); we can then state the following:

Proposition 1.2. *Any vector space V^{2n} with a para-hypercomplex structure $\{I, J, K\}$ is isomorphic to $H^2 \otimes E^n$ with a standard para-hypercomplex structure. Consequently any para-quaternionic vector space $(V^{2n}, \widetilde{Q})$ is isomorphic to $(H^2 \otimes E^n, \mathfrak{sl}(H))$.*

More explicitly, for any basis $e_1, \ldots, e_n$ of V_J^+, h_1, h_2 of H^2 and $f_1, \ldots, f_n$ of E^n an isomorphism is given by

$$(e_i - Ke_i) \leftrightarrow h_1 \otimes f_i, \quad (e_i + Ke_i) \leftrightarrow h_2 \otimes f_i,$$

where $(e_i + Ke_i) \in V_K^+$ and $(e_i - Ke_i) \in V_K^-$. For $i = 1, \ldots, n$, we have the following other correspondences

$$U_i =< e_i, Ke_i >=< (e_i - Ke_i), \qquad (e_i + Ke_i) >\leftrightarrow H \otimes f_i;$$

$$e_i \leftrightarrow \frac{1}{2}(h_1 + h_2) \otimes f_i, \qquad Ke_i \leftrightarrow -\frac{1}{2}(h_1 - h_2) \otimes f_i;$$

$$V_J^+ \leftrightarrow (h_1 + h_2) \otimes E, \qquad V_J^- \leftrightarrow (h_1 - h_2) \otimes E;$$

$$V_K^+ \leftrightarrow h_2 \otimes E, \qquad V_K^- \leftrightarrow h_1 \otimes E.$$

We underline some analogies and some differences between quaternionic and para-quaternionic spaces.

- *A quaternionic vector space V has dimension $4n$, a para-quaternionic has dimension $2n$ with $n \in \mathbb{N}$.*

  In fact any irreducible $\mathbb{H}$-submodule has dimension 4: it follows from Wedderburn theorem since the simple algebra $\mathbb{H}$ is isomorphic to some ring of 4×4 real matrices. On the other hand, as already stated, for any $X \neq 0$ such that $TX \neq \lambda X$, $\lambda \in \mathbb{R}$, $T \in \tilde{a}$, the $\widetilde{\mathbb{H}}$-invariant submodule $< X, J_1X, J_2X, J_3X >$ is reducible since

  $$\begin{aligned} < X, J_1X, J_2X, J_3X > &= < X + J_3X, J_2(X + J_3X) > \\ &\qquad \oplus < X - J_3X, J_2(X - J_3X) > . \end{aligned}$$

- If $\dim V = 4n$, in quaternionic and para-quaternionic case, there always exists a basis of V of the following type:

  $$\{X_1, \dots, X_n, J_1X_1, \dots, J_1X_n, J_2X_1, \dots, J_2X_n, J_3X_1, \dots, J_3X_n\}$$

  for any admissible basis J_1, J_2, J_3.

- Let $\mathcal{H}$ be a hypercomplex structure $\mathcal{H} = H$ or respectively a para-hypercomplex structure $\mathcal{H} = \widetilde{H}$ on a vector space V. Let $\mathcal{Q} =< \mathcal{H} >$ be the corresponding quaternionic structure Q (resp. para-quaternionic structure $\widetilde{Q}$). The 3-dimensional vector space

  $$\mathrm{End}(V) \supset \mathcal{Q} =< \mathcal{H} >= \mathbb{R}I + \mathbb{R}J + \mathbb{R}K$$

  has a natural Euclidean (resp. pseudo-Euclidean) norm defined by

  $$L^2 = -\|L\|^2 \, Id, \; L \in \mathcal{Q}.$$

  Namely, if $\mathcal{Q} = Q$ is a quaternionic structure and

  $$\mathcal{Q} \ni L = aI + bJ + cK, \quad a, b, c \in \mathbb{R},$$

  since $L^2 = (-a^2 - b^2 - c^2)$ Id, then

  $$\|L\|^2 = (a^2 + b^2 + c^2).$$

  In the same way, if $\mathcal{Q} = \widetilde{Q}$ is a para-quaternionic structure and $L \in \mathcal{Q}$, from $L^2 = (-a^2 + b^2 + c^2)$ Id, we get

  $$\|L\|^2 = a^2 - b^2 - c^2.$$

Clearly, complex (resp. para-complex) structures have norm equal 1 (resp. -1). Notice the existence in $\widetilde{Q}$ of null vectors, i.e. $L \in \widetilde{Q}$ such that $\|L\|^2 = 0$, corresponding to nilpotent endomorphisms.

- Two admissible bases of $\mathcal{Q}$ are related by an orthogonal matrix $A = (A^\alpha_\beta) \in \mathrm{SO}(3)$ in the case of $\mathcal{Q} = Q$, or by a pseudo-orthogonal matrix $A = (A^\alpha_\beta) \in \mathrm{SO}(1,2)$ in the case of $\mathcal{Q} = \widetilde{Q}$.

- A complex structure or para-complex structure L on V is called *compatible* with $\mathcal{H} = (I, J, K)$ if it belongs to $\mathcal{Q} =< \mathcal{H} >$. The set of complex and para-complex structures compatible with $\mathcal{H}$ is denoted by $S(Q)$. If $\mathcal{H} = H$ is a hypercomplex structure then

$$S(Q) := S^- = \{L = aI + bJ + cK, a, b, c \in \mathbb{R},\ \|L\|^2 \equiv a^2 + b^2 + c^2 = 1\}$$

is a 2-sphere of complex structure. If $\mathcal{H} = \widetilde{H}$ is a para-hypercomplex structure then

$$S(Q) = S^+(Q) \cup S^-(Q),$$

where

$$S^+(Q) = \{L = aI + bJ + cK, a, b, c \in \mathbb{R},\ \|L\|^2 = a^2 - b^2 - c^2 = -1\},$$

is a one-sheet hyperboloid consisting of para-complex structures, and

$$S^-(Q) = \{L = aI + bJ + cK, a, b, c \in \mathbb{R},\ \|L\|^2 = a^2 - b^2 - c^2 = 1\},$$

is a two-sheets hyperboloid consisting of complex structures [2]. Observe that the set of nilpotent endomorphisms

$$\{A = aI + bJ + cK, a, b, c \in \mathbb{R},\ \|L\|^2 = a^2 - b^2 - c^2 = 0\}$$

in $\widetilde{Q}$ is a cone.

2. Para-type structures on a pseudo-Euclidean vector space

Definition 2.1. Let (V, K) be a $2n$-dimensional para-complex vector space. A pseudo-Euclidean scalar product $g =< \cdot, \cdot >$ on (V, K) is called *$K-$Hermitian* if K is a skew-symmetric endomorphism of $(V, < \cdot, \cdot >)$.

A vector space V endowed with a para-complex structure K and a K-Hermitian scalar product g is called a *para-Hermitian vector space* (V, K, g). The reason why we do not consider n-dimensional vector spaces endowed with a product structure not para-complex is that, as it will be stated afterwards, the metric on such spaces is always degenerate.

A para-hypercomplex structure (J_1, J_2, J_3) on V is called *para-hypercomplex Hermitian structure* with respect to the pseudo-Euclidean scalar product g if its endomorphisms are skew-symmetric with respect to g.

A para-quaternionic structure $\widetilde{Q}$ on V is called *para-quaternionic Hermitian structure* with respect to g if some (and hence any) admissible basis is Hermitian with respect to g.

We give the following

Definition 2.2. Let (V, g) be a pseudo-Euclidean vector space. A subspace $W \subset V$ is called *degenerate* if the restriction $g|_W$ is degenerate i.e. if there exists a non zero $y \in W$ such that $g(x, y) = 0$, $\forall x \in W$, and is called *totally isotropic* if $g|_W \equiv 0$ i.e. if $g(x, y) = 0$, $\forall x, y \in W$.

We need also the following lemma of linear algebra:

Lemma 2.1. *A pseudo-Euclidean vector space (V^{2n}, g) has signature (n, n) if and only if V admits a decomposition $V = U_1 \oplus U_2$ into a direct sum of two totally isotropic subspaces. Moreover* $\dim U_1 = \dim U_2 = n$.

Proof. Let define

$$\begin{array}{cccc} \alpha : & U_1 & \rightarrow & U_2^*, \\ & X & \mapsto & \alpha_X = g(X, \cdot) \end{array}$$

such that $\alpha_X(Y) = g(X, Y), X \in U_1,\ Y \in U_2$. The map α is clearly linear. Moreover, since g is non degenerate, it is injective. This implies $\dim U_1 \leq \dim U_2^*$. Analogously, defining the map

$$\begin{array}{cccc} \alpha' : & U_2 & \rightarrow & U_1^*, \\ & Y & \mapsto & \alpha'_Y = g(Y, \cdot) \end{array}$$

we get that $\dim U_1 = \dim U_2 = n$ and $\alpha : U_1 \rightarrow U_2^*$ is an isomorphism. Let choose $(f_1, \ldots, f_n)$ a basis of U_2 and denote by $(f_1^*, \ldots, f_n^*)$ the dual basis of U_2^*, i.e. $f_i^*(f_j) = \delta_{ij}$. Then $(e_i = \alpha^{-1}(f_i^*),\ i = 1, \ldots, n)$ is a basis of U_1. With respect to the basis $(e_1, \ldots, e_n, f_1, \ldots, f_n)$ of V the Gram matrix of g is

$$g = \begin{pmatrix} 0 & \mathrm{Id} \\ \mathrm{Id} & 0 \end{pmatrix}.$$

Then, with respect to the basis

$$(u_i = \frac{1}{2}(e_i + f_i), u'_i = \frac{1}{2}(e_i - f_i), i = 1, \ldots, n),$$

we have

$$g = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & -\mathrm{Id} \end{pmatrix}.$$

Viceversa, considering the basis $(e_1, \ldots, e_n, f_1, \ldots, f_n,)$ of V, with respect to which g is diagonal with $g(e_i, e_i) = 1$ and $g(f_i, f_i) = -1$, the subspaces

$$U_1 =< e_1 + f_1, \ldots, e_n + f_n > \quad \text{and} \quad U_2 =< e_1 - f_1, \ldots, e_n - f_n >$$

are totally isotropic of same dimension.

In the following, by *neutral signature* we will refer both to pseudo-Euclidean metrics (indicating a (n, n) signature) as well as to degenerate metrics (meaning a (n, s, n) signatures with $s = \dim \ker g$).

More generally, it is straightforward to prove the

Lemma 2.2. *Let (V^n, g) be a vector space with an eventually degenerate scalar product g. The signature of g is neutral if and only if V admits a decomposition into a direct sum of a pair of totally isotropic subspaces i.e. $V = (V_1)^h \oplus (V_2)^k$. More precisely the signature is $(r, n-2r, r)$ where $r = rk\, g(V_1 \times V_2)$ or equivalently the signature is*

$$(\frac{n-s}{2}, s, \frac{n-s}{2}),$$

where

$$s = \dim \ker\, g(V_1 \times V_2) + \dim \ker\, g(V_2 \times V_1).$$

Proposition 2.3. *Let V be a vector space. There exists a $1-1$ correspondence between the following objects:*
1) *para-Hermitian structure (g, K) on V,*
2) *pseudo-Euclidean metric g (of neutral signature) together with a decomposition $V = V^+ \oplus V^-$ such that V^+ and V^- are totally isotropic subspaces i.e.*

$$g(V^\pm, V^\pm) = 0,$$

3) *nondegenerate skew-symmetric bilinear form ω (symplectic form), together with a decomposition $V = V^+ \oplus V^-$ such that V^+ and V^- are Lagrangian subspaces, that is $\omega(V^\pm, V^\pm) = 0$.*

Proof. We shall distinguish three steps.

- 1) $\to$ 2): the metric is the same given metric g: $V^+ = V_K^+$ and $V^- = V_K^-$ are the eigenspaces of K; from the g-skew symmetry of K, they are totally isotropic. Moreover, by Lemma (2.1), g has signature (n, n).

- 2) $\to$ 3): Define $K = \mathrm{Id}$ on V^+, $K = -\mathrm{Id}$ on V^-. Then K is g-skew symmetric, and hence $\omega = g \circ K$ is a symplectic 2-form. Moreover V^+ and V^- are ω-Lagrangian subspaces.

- 3) $\to$ 1): Define $K = \mathrm{Id}$ on V^+, $K = -\mathrm{Id}$ on V^-. Then $g = \omega \circ K^{-1} = \omega \circ K$ is symmetric. Bilinearity and nondegeneracy of g follow from bilinearity and nondegeneracy of ω. Moreover the para-complex structure K is g-skew symmetric.

The existence of a nondegenerate, indefinite Hermitian metric on a para-hypercomplex (resp. para-quaternionic) Hermitian vector space leads to the following

Proposition 2.4. *The dimension of a vector space V, endowed with a para-hypercomplex (resp. para-quaternionic) Hermitian structure $(\widetilde{H}, g)$ (resp. $(\widetilde{Q}, g)$), is a multiple of 4. Moreover g has neutral signature (n, n), n even.*

Proof. Let $(\widetilde{H} = \{I, J, K\}, g)$ be a para-hypercomplex Hermitian structure (resp. $(\widetilde{Q} =< I, J, K >, g)$ be a para-quaternionic Hermitian structure) of V. Let moreover $(e_1, \ldots, e_n)$ be a basis of V_J^+, the eigenspace corresponding to the $+1$ eigenvalue of the para-complex structure J. Observe that $(Ke_1, \ldots, Ke_n)$ is a basis of V_J^-. The subspace V_J^+ (resp. V_J^-) is totally isotropic since, by the skew-symmetry of J,

$$0 = g(JX, Y) + g(X, JY) = 2g(X, Y) \qquad \forall X, Y \in V_J^+$$

(resp. $\forall X, Y \in V_J^-$) and $V = V_J^+ \oplus V_J^-$ which, by Lemma (2.1), implies that g has neutral signature. With respect to the basis $\{e_1, \ldots, e_n, Ke_1, \ldots, Ke_n\}$ the metric g can be written as

$$g = \begin{pmatrix} 0 & A \\ A^t & 0 \end{pmatrix},$$

where A is a skew-symmetric $(n \times n)$-matrix by the Hermitian hypothesis. Then g nondegenerate on V implies n even.

Remark 2.7. In the proof above we have proved that the subspaces V_J^+ and V_J^- are totally isotropic. When dealing with a para-quaternionic Hermitian structure the eigenspaces associated to any admissible para-complex structure are always maximal totally isotropic.

As a consequence of Proposition (2.4), the decomposition

$$V^{4n} = U_1 \oplus U_2 \oplus \ldots \oplus U_{2n}$$

of the $\widetilde{Q}$-module (V^{4n}, g, Q) into direct sum of $\widetilde{Q}$-invariant 2-dimensional subspaces $U_i =< e_i, Ke_i >$ defined in (6) is not orthogonal since each U_i is totally isotropic whereas g is nondegenerate on V.

The prototype of para-hypercomplex Hermitian vector spaces is the n-dimensional para-quaternionic numerical space $\widetilde{\mathbb{H}}^n$ which is a real vector space of dimension $4n$, a $\widetilde{\mathbb{H}}$-module with respect to left multiplication by para-quaternions i.e.

$$q(h) = q(h_1, \ldots, h_n) := (qh_1, \ldots, qh_n) \quad \forall h \in \widetilde{\mathbb{H}}^n, \ q \in \widetilde{\mathbb{H}}$$

and is endowed with the canonical Hermitian product

$$h \cdot h' = \sum_{\alpha=1}^{n} h_\alpha \overline{h'_\alpha}; \qquad h = (h_1, \ldots, h_n), \ \ h' = (h'_1, \ldots, h'_n) \in \widetilde{\mathbb{H}}^n. \tag{8}$$

The real part of the Hermitian product,

$$\mathrm{Re}(h \cdot h') = \mathrm{Re}\,(h_1 \overline{h'_1}) + \ldots + \mathrm{Re}\,(h_n \overline{h'_n}), \ h_i, h'_i \in \widetilde{\mathbb{H}}, \tag{9}$$

defines a pseudo-Euclidean (canonical) scalar product of neutral signature on the real vector space $\widetilde{\mathbb{H}}^n \simeq \mathbb{R}^{2n,2n}$.

Left multiplications by i, j, k, respectively, induce real endomorphisms, that we denote by the same name and symbol, on the real vector space $\widetilde{\mathbb{H}}^n$ satisfying (2). With respect to the scalar product (9), i is an isometry, while j and k are anti-isometries of $\widetilde{\mathbb{H}}^n$. All three endomorphisms i, j, k are skew-symmetric with respect to the metric. Observe that i, j, k are not automorphisms of $\widetilde{\mathbb{H}}^n$ regarded a vector space over $\widetilde{\mathbb{H}}$. In fact in general $i(qX) \neq q(iX)$, $X \in \widetilde{\mathbb{H}}^n$, $q \in \widetilde{\mathbb{H}}$ unless q is real.

Any para-hypercomplex Hermitian vector space

$$\mathbb{V} = (V^{4n}, \{I, J, K\}, g)$$

is isomorphic to $(\widetilde{\mathbb{H}}^n, \{i, j, k\}, \mathrm{Re}(\,\cdot\,))$. As an $\widetilde{\mathbb{H}}$-module, on a para-hypercomplex Hermitian vector space $\mathbb{V}$ we define the ($\widetilde{\mathbb{H}}$-valued)-Hermitian product $(\cdot) = (\cdot)_{\{I,J,K\}}$ by

$$\begin{aligned} (10)\quad (\cdot) : V \times V &\to \widetilde{\mathbb{H}}, \\ (X, Y) &\mapsto X \cdot Y = g(X, Y) + ig(X, IY) - jg(X, JY) - kg(X, KY). \end{aligned}$$

Claim 2.8. For any isomorphism $\phi : \mathbb{V} \to (\widetilde{\mathbb{H}}^n, \{i, j, k\}, \mathrm{Re}(\,\cdot\,))$ of Hermitian para-hypercomplex vector spaces, we have

$$X \cdot Y = \phi(X) \cdot \phi(Y), \quad \forall X, Y \in V.$$

Proof. Let $\phi : \mathbb{V} \to (\widetilde{\mathbb{H}}^n, \{i, j, k\}, \mathrm{Re}(\,\cdot\,))$ be an isomorphism of Hermitian para-hypercomplex vector spaces and denote $h = \phi(X), h' = \phi(Y)$, $X, Y \in V$. We prove that $g(X, IY)$, $-g(X, JY)$, $-g(X, KY)$, are the coefficients of i, j, k in $(h \cdot h')$, respectively. We get

$$g(X, IY) = \mathrm{Re}(h \cdot ih') = \mathrm{Re}[\sum_{\alpha=1}^{n} (h_\alpha \overline{ih'_\alpha})] = \mathrm{Re}[-\sum_{\alpha=1}^{n} (h_\alpha \overline{h'_\alpha}) i] = -\mathrm{Re}[(h \cdot h')i].$$

Considering that $j^2 = k^2 = 1$, the conclusion follows.

When taking into account a para-quaternionic Hermitian vector space $(V, \widetilde{Q}, g)$, we observe that the ($\widetilde{\mathbb{H}}$-valued)-Hermitian product defined in (10) depends on the chosen admissible basis $\{I, J, K\} \in \widetilde{Q}$. Two Hermitian products $(\;\cdot\;)_{\{I,J,K\}}$, $(\;\cdot\;)_{\{I',J',K'\}}$, referred to different admissible bases, are related by an inner automorphism of $\widetilde{\mathbb{H}}$. This implies that

$$\mathcal{N}((X \cdot Y)_{\{I,J,K\}}) = \mathcal{N}((X \cdot Y)_{\{I',J',K'\}}), \quad \forall X, Y \in V,$$

or equivalently, since the real part of the norm $\mathcal{N}((X \cdot Y))$ is independent of the basis $\{I, J, K\}$,

$$\mathcal{N}(\mathrm{Im}(X \cdot Y)_{\{I,J,K\}}) = \mathcal{N}(\mathrm{Im}(X \cdot Y)_{\{I',J',K'\}}), \quad \forall X, Y \in V.$$

Let consider now the standard para-hypercomplex vector space $(H^2 \otimes E^{2n}, \{I, J, K\})$, (resp. para-quaternionic vector space $(H^2 \otimes E^{2n}, \mathfrak{sl}(H))$). Let ω^E be a symplectic form on E and $\omega^H = h_1^* \wedge h_2^*$ a (standard) volume form on H. Observe that $\mathfrak{sl}(H) \simeq \mathfrak{sp}_{\omega^H}(H)$, the Lie algebra of transformations preserving ω^H.

The 2-form $\omega^H \otimes \omega^E$ is bilinear, symmetric and nondegenerate, defining a metric g on $H^2 \otimes E^{2n}$. In fact bilinearity and nondegeneracy follow from bilinearity and nondegeneracy of both ω^H and ω^E; furthermore, calculating on decomposable vectors, we get

$$g(h' \otimes e, \tilde{h} \otimes e') = \omega^H(h', \tilde{h})\omega^E(e, e') = \omega^H(\tilde{h}, h')\omega^E(e', e) = g(\tilde{h} \otimes e', h' \otimes e),$$

and hence g is symmetric. Observe that $g = \omega^H \otimes \omega^E$ is a para-hypercomplex Hermitian (resp. para-quaternionic Hermitian) metric on $H \otimes E$. In fact, for any $A \in \mathfrak{sl}(H)$, and calculating again on decomposable vectors, we obtain

$$\begin{aligned} g(Ah' \otimes e, \tilde{h} \otimes e') = \omega^H(Ah', \tilde{h})\omega^E(e, e') &= -\omega^H(h', A\tilde{h})\omega^E(e, e') \\ &= -g(h' \otimes e, A\tilde{h} \otimes e'). \end{aligned}$$

Definition 2.9. The $4n$-dimensional space $(H^2 \otimes E^{2n}, \{I, J, K\}, \omega^H \otimes \omega^E)$ (resp. $(H^2 \otimes E^{2n}, \mathfrak{sl}(H), \omega^H \otimes \omega^E)$ is *a standard para-hypercomplex Hermitian space* (resp. *the standard para-quaternionic Hermitian space*).

Proposition 2.5. *Let V^{4n} be a vector space with a para-quaternionic Hermitian structure $(\widetilde{Q}, g)$. Then the para-quaternionic Hermitian space $(V, \widetilde{Q}, g)$ is isomorphic to a standard para-quaternionic Hermitian space.*

Proof. By Corollary (1.2) we identify $(V^{4n}, \widetilde{Q})$ with $(H^2 \otimes E^{2n}, \mathfrak{sl}(H))$.

Then the given para-quaternionic Hermitian metric g on $H^2 \otimes E^{2n}$ can be written as $g = \omega^H \otimes \omega^E$ where $\omega^H = h_1^* \wedge h_2^*$ is the standard volume form on H and ω^E is defined by

$$\omega^E(e, e') := \frac{g(h \otimes e, h' \otimes e')}{\omega^H(h, h')},$$

for one (and hence any) pair of linearly independent vectors h, h'. It is straightforward to prove that the right member is well defined by observing that, by hermitianity, decomposable vectors are always isotropic (recall Remark 2.7) and $g(h_1 \otimes e, h_2 \otimes e') + g(h_2 \otimes e, h_1 \otimes e') = 0$. Moreover ω^E is clearly symplectic.

We conclude this first part with the following proposition which generalizes to the degenerate case a well known result valid for Hermitian vector spaces.

Proposition 2.6. *Let g be an indefinite and (possibly) degenerate scalar product in a $2n$-dimensional (resp. n-dimensional) vector space V endowed with a g-skew-symmetric complex (resp. product) structure I. Then the signature of g is of type $(2p, 2s, 2q)$ (resp. (m, t, m)) and there exist vectors $X_1, \ldots, X_n$ of V such that*

$$\{X_1, \ldots, X_n, IX_1, \ldots, IX_n\}$$

is a orthogonal basis of V and

$$\|X_i\| = \|IX_i\| = 1 \quad \textit{for} \quad i = 1, \dots, p,$$

$$(\textit{resp.}\ \|X_i\|) = -\|IX_i\| = 1 \quad \textit{for} \quad i = 1, \dots, m),$$

$$\|X_i\| = \|IX_i\| = -1 \quad \textit{for} \quad i = p+1, \dots, p+q,$$

$$(\textit{resp.}\ \|X_i\| = -\|IX_i\| = -1 \quad \textit{for} \quad i = m+1, \dots, 2m),$$

and

$$\|X_i\| = \|IX_i\| = 0 \quad \textit{for} \quad i = p+q+1, \dots, 2n$$

$$(\textit{resp.}\ i = 2m+1, \dots, 2n).$$

Acknowledgment

Work done under the programs of GNSAGA-INDAM of C.N.R. and PRIN07 "Riemannian metrics and differentiable structures" of MIUR (Italy) and in the framework of the direct cooperation agreement between the Universities of Roma "La Sapienza" and of Łódź.

References

[1] D. V. Alekseevsky, S. Marchiafava, *Quaternionic structures on a manifold and subordinated structures*, Ann. Mat. Pura Appl. **171** (1996), 205–273.

[2] D. V. Alekseevsky, V. Cortés, *The twistor spaces of a para-quaternionic Kaehler manifold*, Osaka J. Math. **45**, no. 1 (2008), 215–251.

[3] M. Bruni, *Su alcune proprietá di geometria euclidea ed hermitiana in uno spazio vettoriale quaternionale*, Ann. Mat. Pura Appl. **72** (1965), 59–77.

[4] V. Cortés, *A new construction of homogeneous quaternionic manifolds and related geometric structures*, Mem. Amer. Math. Soc. **147** (2000), no. 700.

[5] S. Salamon, *Quaternionic Kaehler manifolds*, Invent. Math. **67** (1982), 143–171.

[6] S. Salamon, *Differential geometry of quaternionic manifolds*, Ann. Scient. Ec. Norm. Sup. (4) 19 (1986), 31–55.

[7] M. Vaccaro, *Subspaces of a para-quaternionic Hermitian vector space*, Internat. J. Geom. Methods in Modern Phys. **8**, no. 7 (2011), in print; arXiv:1011.2947v1 [math.DG].

[8] M. Vaccaro, *Basics of linear para-quaternionic geometry II: Decomposition of a generic subspace of a para-quaternionic Hermitian vector space*, Bull. Soc. Sci. Lettres Łódź Sér. Rech. Déform. **61**, no. 2 (2011), to appear.

[9] S. Vukmirovic, *Para-quaternionic reduction* (2003) math. DG/0304424.

Dipartimento dell'Ingegneria di Informazione
e Matematica Applicata
Università di Salerno
I-84084 Fisciano (SA)
Italy
e-mail: massimo_vaccaro@libero.it

Presented by Włodzimierz Waliszewski at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on December 22, 2010

PODSTAWY LINIOWEJ GEOMETRII PARA-KWATERNIONOWEJ I

STRUKTURA PARA-TYPÓW HERMITOWSKICH NA RZECZYWISTEJ PRZESTRZENI WEKTOROWEJ

S t r e s z c z e n i e

W pracy przedstawiamy podstawowe definicje i wyniki z geometrii para-kwaternionowej. Standardowa struktura para-kwaternionowa $\widetilde{Q}$ na iloczynie tensorowym $H \otimes E$ odpowiednio pary rzeczywistych przestrzeni wektorowych wymiaru 2 i $2n$ jest określona jako algebra Liego $\widetilde{Q} = \mathfrak{sl}(H)$ specjalnej grupy liniowej $\mathrm{SL}(H)$ automorfizmów przestrzeni H zachowujących objętość. Każda para-kwaternionowa przestrzeń wektorowa $(V, \widetilde{Q})$ jest izomorficzna z przestrzenią $(H^2 \otimes E^n, \mathfrak{sl}(H))$. Co więcej, jeśli (H, ω^H) i (E, ω^E) są przestrzeniami symplektycznymi, to 2-forma $\omega^H \otimes \omega^E$ określa metrykę $\widetilde{Q}$-hermitowską na $(H^2 \otimes E^n, \mathfrak{sl}(H))$ i dowolna hermitowska para-kwaternionowa przestrzeń wektorowa $(V, \widetilde{Q}, g)$ jest izomorficzna z przestrzenią $(H^2 \otimes E^{2n}, \mathfrak{sl}(H), \omega^H \otimes \omega^E)$.

W części I pracy opisujemy struktury para-typów na przestrzeni wektorowej i na pseudoeuklidesowej przestrzeni wektorowej. W części II będziemy badali stosowne klasy podprzestrzeni para-kwaternionowych hermitowskich przestrzeni wektorowych, a w szczególności rozkład przestrzeni generującej.

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 37–46

Dedicated to Professor Roman Stanisław Ingarden
on the occasion of his ninetieth birthday

Jerzy Rutkowski, Leszek Wojtczak, and Claude Surry

SINGULAR PERTURBATION PROBLEM.
HOW TO APPROXIMATE DIRAC FUNCTION

Summary

In the present paper we consider the behaviour of elliptic equations of the type $\varepsilon u - u'' = f$ or $u - \varepsilon u'' = f$ with the Dirichlet or Newman boundary conditions on an interval $I =]0,1[$. The second case is a limit problem for $\varepsilon = 0$ of a strictly lower order for problems with $\varepsilon > 0$. In a second part, we define a Dirac function at the point $x = 1/2$ on $I =]0,1[$ and we approximate this kind of Dirac jump force by a still continuous function using the second type of the equation $u - \varepsilon u'' = 0$ whose classical solution is a stiff hyperbolic sinus with a force applied at the point $x = 1/2$ [1–3].

Introduction

We study the singular perturbation problem for elliptic equations of the type

(A) $$\varepsilon u - u'' = f$$

or

(B) $$u - \varepsilon u'' = f$$

with $\varepsilon \to 0$ in the case of the Dirichlet or Newman boundary conditions, where the function u'' denotes the second derivative of u with respect to the space variable x. We consider the Dirac function in the interval $I =]0,1[$ and we introduce the Sobolev space denoted by $H_0^1(I)$ and equipped for the norm

(1) $$|u|_1 = \left[\int_I u'^2 dx\right]^{1/2}$$

$$H_0^1 \left\{ v | v \in L^2(I);\ v' \in L^2(I) | v(0) = v(1) = 0 \right\}$$

$$|v|_1^2 = \int_I {v'}^2 dx,$$

while we note

$$|u| = \left[\int_I u^2 dx \right]^{1/2}. \tag{2}$$

A.I. Dirichlet problem

We want to show that it exists a unique $w_\varepsilon \in H_0^1$ for $\varepsilon > 0$, so that

$$\varepsilon \int_I w_\varepsilon v dx + \int_I w_\varepsilon' v' dz = \int_I f v dx. \tag{3}$$

In this purpose we introduce a continuous bilinear form $a_\varepsilon(u, v)$ which is coercive on H_0^1 and it satisfies the definition

$$a_\varepsilon(u, v) = \varepsilon \int_I u v dx + \int_I u' v' dx. \tag{4}$$

We assume that C is the Poincaré constant. Then

$$|a(u, v)| \le \left(C^2 \varepsilon + 1 \right) |u|_1 |v|_1, \tag{5}$$

$$a_\varepsilon(u, v) \ge |u_1|\, |v_1|. \tag{6}$$

Next, we take into account the function $v \to \int_I f v dx$ which is also a continuous bilinear form on H_0^1. Then, the Lax-Milgram theorem reads

$$\forall v \in H_0^1 : a_\varepsilon \left(w_\varepsilon, v \right) = \int f v dx, \tag{7}$$

and it allows us to state that $w_\varepsilon \to w_0$ on H_0^1 with

$$\forall v \in H_0^1 : \int_I w_0' v' dx = \int_I f v dx, \tag{8}$$

where w_ε is bounded in H_0^1. Taking $v = w_\varepsilon$ in (4) we get

$$|w_\varepsilon|_1^2 < a_\varepsilon(w_\varepsilon, w_\varepsilon) = \int_I f w_\varepsilon dx \le |f|\, |w_\varepsilon| \le |f|\, |w_\varepsilon|_1 \times C \tag{9}$$

with

$$|w_\varepsilon|_1 \le C |f| \tag{10}$$

for $\forall v \in H_0^1$, hence

$$\forall v \in H_0^1 : \varepsilon \int_I w_\varepsilon v dx + \int_I w_\varepsilon' v' dx = \int f v dx, \tag{11}$$

$$f_I w_0' v' dx = \int_I f v dx.$$

Taking into account the test function $v = w_\varepsilon - w_0$ we get

(12) $$\varepsilon \int_I w_\varepsilon (w_\varepsilon - w_0) dx + \int_I (w_\varepsilon' - w_0')^2 dx = 0$$

or, with (10), we can write

(13) $$\int_I (w_\varepsilon' - w_0')^2 dx \leq |w_\varepsilon| \, |w_\varepsilon - w_0| \times \varepsilon,$$

where

$$|w_\varepsilon' - w_0'|_1^2 \leq C^2 \varepsilon |f| \, |w_\varepsilon - w_0|_1$$

and

$$|w_\varepsilon - w_0|_1 \leq C^2 \varepsilon |f| \to 0 \quad \text{when} \quad \varepsilon \to 0.$$

A.II. Newman problem

Define $H^1 = H^1(I)$ equiped with the norm

(14) $$v \longrightarrow \|v\| = |v| + |v|_1.$$

We want to show how to find $w_\varepsilon \in H^1$ so that

(15) $$\varepsilon \int_I w_\varepsilon v dx + \int_I w_\varepsilon' v' dx = \int_I f v dx.$$

The existence and the unicity of the problem is given by Lax-Milgram theorem and we see that the constant of coercivity tends to 0 with ε. Take $v = 1$ in (15) we obtain:

(16) $$\varepsilon \int_I w_\varepsilon dx = \int_I f dx,$$

and if $\int_I f dx \neq 0$ we have $| \int_I w_\varepsilon dx| \to +\infty$. The sequence w_ε is not bounded in this case in $L^p, p \in]1, \infty[$ and we cannot get the convergence (weak or strong) in this kind of space.

Consider now the sequence

(17) $$\tilde{w}_\varepsilon = w_\varepsilon - \frac{\int_I f dx}{\varepsilon}.$$

We have

(18) $$\int_I \tilde{w}_\varepsilon dx = 0$$

for which we can use the Poincare-Wirtinger inequality:

(19) $$\tilde{w}_\varepsilon \leq |w_\varepsilon|_1.$$

We have

$$\forall v \in H^1 : \varepsilon \int_I \tilde{w}_\varepsilon v dx + \int_I \tilde{w}'_\varepsilon v' dx = \int_I f v dx - \int_I f dx \times \int_I v dx \tag{20}$$

where $\tilde{w}_\varepsilon$ is bounded in H^1 (take $v = \tilde{w}_\varepsilon$ in (20)). We consider a strong convergence in H^1 for which the solution is of the form

$$\forall v \varepsilon H^1 : \int_I w'_0 v' dx = \int_I f v dx - \int_I f dx \times \int_I v dx, \tag{21}$$

for $\varepsilon \to 0$, and then $\omega_\varepsilon \to \omega_o$. When f has its average equal to zero

$$\int_I f dx = 0 \tag{22}$$

we get the same results as in the case of H_0^1.

B.I. Dirichlet problem

Define a new bilinear form

$$a_\varepsilon(u, v) = \int_I u v dx + \varepsilon \int_I u' v' dx \tag{23}$$

in the space H_0^1 for C being the Poincaré constant, we obtain:

$$|a_\varepsilon(u, v)| \leq |u|\,|v| + \varepsilon |u|_1 |v|_1 \leq (C^2 + \varepsilon)|u|_1 |v|_1. \tag{24}$$

The coercivity gives $\forall u \in H_0^1$ where

$$a_\varepsilon(u, u) \geq \varepsilon |u_1|^2. \tag{25}$$

The Lax-Milgram theorem gives $\exists u_\varepsilon \forall v \in H_0^1$ that

$$a_\varepsilon(u, v) = \int_I f u dx. \tag{26}$$

Taking $v = u_\varepsilon$ in (26) we get

$$\int_I u_\varepsilon^2 dx = \int_I f u_\varepsilon dx - \varepsilon \int_I |u'_\varepsilon|^2 dx \leq \int_I f u_\varepsilon dx \tag{27}$$

and

$$|u_\varepsilon| \leq |f| \quad \text{for every} \quad \varepsilon > 0.$$

Take a test function $\varphi = C_c^\infty(I)$ (space of infinitely, derivable functions having a compact support in I), we obtain

$$\varphi' = \frac{d\varphi}{dx} \in H_0^1$$

and the integration by parts leads to the result:

$$\int_I u_\varepsilon \varphi' dx = -\int_I u_\varepsilon \varphi'' dx \tag{28}$$

and

$$\left| \int_I u'_\varepsilon \varphi' dx \right| \leq |u_\varepsilon| \, |\varphi''| \leq |f| \, |\varphi''| \tag{29}$$

for every $\varphi \in C_c^\infty(i)$, hence we have

$$\int_I u_\varepsilon \varphi dx + \varepsilon \int_I u'_\varepsilon \varphi' dx = \int_I f \varphi dx \tag{30}$$

by the use of (29)

$$\lim_{\varepsilon \to 0} \int_I u_\varepsilon \varphi dx = \int_I f \varphi dx. \tag{31}$$

The sequence u_ε is bounded in L^2 and C_c^∞ is dense in L^2 (L^2 is the dual space of L^2). We get the weak convergence of u_ε to f in L^2. As u_ε converges weakly to f in L^2, we get

$$|f| \leq \lim\ \mathrm{inf} |u_\varepsilon| \tag{32}$$

and we have

$$|u_\varepsilon| \leq |f| \Rightarrow \limsup |u_\epsilon| \leq |f| \tag{33}$$

and

$$u_\varepsilon \to f \quad \text{as} \quad \varepsilon \to 0 \tag{34}$$

in $L^2(I)$.

B.II. Newman problem

The crossing from H_0^1 to H^1 does not modify the steps and the results. The coercivity constant associated to the bilinear form

$$a_\varepsilon(u, v) = \int_I uv dx + \varepsilon \int_I u'v' dx \tag{35}$$

tends to 0 as $\varepsilon \to 0$ as in the case of H_0^1. We have the same walk as in H_0^1 till the strong convergence of u_ε to f in L^2.

C. Unit force the Dirichlet problem solution

We consider the functional

$$\int_I u_\varepsilon v dx + \int_I \varepsilon u'_\varepsilon v' dx = v\left(\frac{1}{2}\right) \quad \text{for} \quad \forall v \in H_0^1 \tag{36}$$

for which we assume the value $\alpha = 1/2$.

The space H_0^1 is continuously embedded in $C(I)$ (space of continuous functions) equipped with the norm of uniform convergence. We assume

$$v \to v(\alpha) \tag{37}$$

is a linear continuous form on H_0^1.

In this case the Lax-Milgram theorem insures the existence and the unicity of $u_\varepsilon \in H_0^1$ so that

$$\forall v \in H_0^1 \tag{38}$$

$$\int_I u_\varepsilon v dx + \varepsilon \int_I u'_\varepsilon v' dx = v(\alpha).$$

Taking a test function $\varphi \in C_c^1(]0, \frac{1}{2}[)$ we have

$$\int_I u'_\varepsilon \varphi' dx = -\int_I u_\varepsilon \varphi'' dx \tag{39}$$

and

$$u'_\varepsilon \in H^1\left(\left]0, \frac{1}{2}\right[\right) \tag{40}$$

whose derivative is u''_ε, so that

$$u_\varepsilon - \varepsilon u''_\varepsilon = 0. \tag{41}$$

The walk is the same for x belonging to the intervals $]0, 1/2[$ and $]1/2, 1[$. We consider now a test function which do not cancel necessarily for $x = 1/2$. In terms of the variational formulation we write:

$$\forall v \in H_0^1 \int_I u_\varepsilon v dx + \varepsilon \int_{]0,\frac{1}{2}[} u'_\varepsilon v' dx + \varepsilon \int_{]\frac{1}{2},1[} u'_\varepsilon v' dx = v(\alpha) \tag{42}$$

with (40) and (41). The restriction of u''_ε is determined to each interval open in L^2. We can integrate by parts each integral having u'_ε, and we get

$$\begin{aligned} &\int_{]0,\frac{1}{2}[} (u - \varepsilon u''_\varepsilon)\, v dx + \int_{]\frac{1}{2},1[} (u - \varepsilon u''_\varepsilon)\, v dx \\ &+ \varepsilon v(\alpha) \left[u'_\varepsilon(\alpha^-) - u'_\varepsilon(\alpha^+)\right] = v(\alpha). \end{aligned} \tag{43}$$

The first two integrals are cancelled by (41). We get

$$u'_\varepsilon\left(\frac{1}{2}^-\right) - u'_\varepsilon\left(\frac{1}{2}^+\right) = \frac{1}{\varepsilon}. \tag{44}$$

Let us define

$$w(x) = u_\varepsilon(x) - u_\varepsilon(1-x) \tag{45}$$

where $x \in]0, \frac{1}{2}[$ while $w(x)$ belongs to $H^2(]0,1/2[)$. We obtain

$$w - \varepsilon w'' = 0 \tag{46}$$

by continuity of u_ε in I, we can write

$$w(0) = 0, \qquad w\left(\frac{1}{2}\right) = 0 \tag{47}$$

where w is the unique function solution of

$$u - \varepsilon u'' = 0 \quad \text{for} \quad x \in \left]0, \frac{1}{2}\right[\tag{48}$$

with cancelled values for $x = 0$, and $x = 1/2$ while $u_\varepsilon(x) = u_\varepsilon(1-x)$ on $]1/2, 1[$, $]0, 1/2[$, u_ε is the classical solution of

$$u''_\varepsilon = \frac{u}{\varepsilon}. \tag{49}$$

As $u_\varepsilon(0) = 0$, u_ε is bounded by $\beta \in R$, namely

$$\mu_\varepsilon(x) = \beta \sinh\left(\frac{x}{\sqrt{\varepsilon}}\right) \tag{50}$$

and

$$u'_\varepsilon\left(\frac{1}{2}^-\right) = \frac{\beta}{\sqrt{\varepsilon}} \cosh\left(\frac{1}{2\sqrt{\varepsilon}}\right). \tag{51}$$

By the symetry, introduced in (45), we obtain

$$u'_\varepsilon\left(\frac{1}{2}^+\right) = -\frac{\beta}{\sqrt{\varepsilon}} \cosh\left(\frac{1}{2\sqrt{\varepsilon}}\right), \tag{52}$$

and with (44), we get $\forall \in]0, 1/2[$

$$\beta = \frac{1}{2\sqrt{\varepsilon}} \sinh \frac{x}{\sqrt{\varepsilon}} \frac{1}{\cosh\left(\frac{1}{2\sqrt{\varepsilon}}\right)}.$$

Next, we obtain the expression on the interval $]1/2, 1[$ by the use

$$u_\varepsilon(x) = u_\varepsilon(1-x).$$

Every compact K of $]0, 1/2[$, $]1/2, 1[$ is at a positive distance from $x = 1/2$ and is included in a set of type:

$$(]0, \eta[\cup]1-\eta, 1[) \quad \text{with} \quad \eta \in \left]0, \frac{1}{2}\right[. \tag{53}$$

For every $x \in \hat{K}$, we have $0 \leq \eta \leq \frac{1}{2}$:

$$0 \leq u_\varepsilon(x) \leq \frac{1}{2\sqrt{\varepsilon}} \sinh\left(\frac{\eta}{\sqrt{\varepsilon}}\right) \frac{1}{\cosh\left(\frac{1}{2\sqrt{\varepsilon}}\right)} \leq \frac{1}{2\sqrt{\varepsilon}} e^{\frac{\eta - \frac{1}{2}}{\sqrt{\varepsilon}}} \tag{54}$$

with the left-hand side of (54) tending to zero as $\varepsilon \to 0$. We consider a uniform convergence on K. We have

$$\begin{aligned}\int_I u_\varepsilon dx = 2\int_0^{1/2} u_\varepsilon dx &= \frac{1}{\sqrt{\varepsilon}\cosh\left(\frac{1}{2\sqrt{\varepsilon}}\right)} \int_0^{1/2} \sinh\left(\frac{x}{\sqrt{\varepsilon}}\right) dx \\ &= \frac{1}{\cosh\left(\frac{1}{2\sqrt{\varepsilon}}\right)}\left[\cosh\left(\frac{1}{2\sqrt{\varepsilon}}\right) - 1\right]\end{aligned} \tag{55}$$

and

$$\int_I u_\varepsilon dx \to 1 \quad \text{if} \quad \varepsilon \to 0.$$

Addendum

We use the Poincaré-Wirtinger inequality for a bounded interval $\omega \in H$. We introduce then the avarage of v on ω, namely:

$$m(v) = \frac{\int_w v dx}{|\omega|}.$$

We have

$$v(y) - v(x) = \int_x^y v'(t)dt.$$

Hence the average with respect to x on ω

$$|v(y) - m(v)| \leq \frac{1}{\omega}\int_\omega dx \int_x^y v'(t)dt \leq \frac{|v'|}{w}\int_w dx = |v'| \tag{56}$$

where

$$|v - m(v)|L^\infty \leq |v'|$$

and

$$|v| \leq m(v) + |v'|.$$

Take $\varphi \in C^0(I)$ (space of continuous functions on I). By continuity of φ at $x = 1/2$, we can see $\forall\varphi > 0$ it exists $\eta \in]0, 1/2[$ so that

$$\left|\varphi(x) - \varphi\left(\frac{1}{2}\right)\right| < \rho \qquad \text{for} \qquad \forall x \in |1, 1-\eta|. \tag{57}$$

Using (54) and (55), we can see that, it exists ε_ρ which below u_ε is overestimated in absolute value by ρ on the compact $K =]0, \eta[\cup]1-\eta, 1[$ and so that

$$\int_{\eta}^{1-\eta} u_\varepsilon dx - 1 < \rho$$

we have $\forall \varepsilon < \varepsilon_\rho$

$$\left| \int_I u_\varepsilon \varphi dx - \varphi\left(\frac{1}{2}\right) \right| \leq \left| \int_{\eta}^{1-\eta} u_\varepsilon \left[\varphi - \varphi\left(\frac{1}{2}\right)\right] dx \right| \tag{58}$$

$$+ \left| \varphi\left(\frac{1}{2}\right) \left[\int_{\eta}^{1-\eta} (u_\varepsilon - 1) dx \right] \right| + \int_K u_\varepsilon \varphi dx$$

and

$$\lim_{\varepsilon \to 0} \int_I u_\varepsilon \varphi dx = \varphi\left(\frac{1}{2}\right). \tag{59}$$

Conclusions

The pertutbated term for tending to zero is approximated by the different limits. The perturbation affects particularly the term of the high derivative of the function u. The application of the Lax-Milgram theorem is governed by the higher derivative of the function in the elliptic equations. In the considered cases for the Dirichlet or Newman perturbated problems, the computing techniques can be applied in the same manner.

References

[1] H. Brézis, *Analyse Fonctionelle – Théorie et Applications*, ed. Masson, Paris 1983.

[2] C. Surry, L. Wojtczak, J. Rutkowski, and I. Zasada, *Application of the finite elements method for thin films description. A basic formulation*, Bull. Soc. Sci. Lettres Łódź **55** Sér. Rech. Déform. **46** (2005), 9–19.

[3] G. Bertholon, C. Dupuy, C. Surry, R. Redon, and H. Zahouani, *Calculations of the constraints in a thin film deposited on the substrate*, Acta Phys. Polon. **83** (1993), 581.

Department of Solid State Physics
University of Łódź
Pomorska 149/153, PL-90-236 Łódź
Poland
e-mail: jerzyrut@uni.lodz.pl

Laboratoire Félix Trombe
Institut de Sciences et de Génie
de Matériaux et Procédés
CNRS, B.P. 5 Odeillo
F-66 125 Font Romeu Cédex
France
e-mail: claude.surry@orange.fr

Presented by Leszek Wojtczak at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on September 29, 2011

PROBLEM OSOBLIWOŚCI W RACHUNKU ZABURZEŃ. JAK PRZYBLIŻYĆ FUNKCJĘ DIRACA

S t r e s z c z e n i e

W niniejszej pracy rozważamy zachowanie się równań eliptycznych typu $\varepsilon u - u'' = f$ oraz $u - \varepsilon u'' = f$ z warunkami brzegowymi Dirichleta lub Newmana w przedziale $I =]0, 1[$. Drugim przykładem jest problem graniczny dla $\varepsilon = 0$ niższego uporządkowania dla $\varepsilon > 0$. Definiujemy funkcję Diraca w punkcie $x = 1/2$ i próbujemy przybliżać ten rodzaj siły jako skok funkcji Diraca o charakterze funkcji ciągłej używając równania drugiego typu $u - \varepsilon u'' = f$, którego rozwiązanie klasyczne jest dane przez sinus hiperboliczny z siłą punktową przyłożoną w punkcie $x = 1/2$ [1–3].

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 47–67

Dedicated to Professor Roman Stanisław Ingarden
on the occasion of his ninetieth birthday

Maciej Skwarczyński

DE BRANGES THEOREM AND GENERALIZED HYPERGEOMETRIC FUNCTIONS I

BIEBERBACH CONJECTURE AND MILIN FUNCTIONAL

Summary

We are going to report the most remarkable mathematical discovery of the previous century. The conjecture stated in 1916 by Ludwig Bieberbach was affirmed in 1984 by Louis de Branges, after 68 years of intensive research by top mathematicians.

0. Foreword

0.1. Editor's foreword

The editor highly recommends reading of the present memoir consisting of three parts. The exposition is entirely different from that of [Skw 09a, b]. First, it takes into account that the de Branges discovery was a part of a larger attempt toward the Riemann Hypothesis. Second, it is remarkably original by distinguishing considerations on the Bieberbach conjecture in relation with the Milin functional, de Branges functional vs. the hypergeometric equation, and basic properties of the Branges functions.

0.2. Author's foreword

Recently I got interested in early comments made in connection with the result of Luigi de Branges (1984). In the first place I should mention here articles of FitzGerald and Pommerenke (1985), Korevaar (1986), Kazarinoff (1988), and Grinshpan

(1999). The early version of my memoir was presented at the XV-th Conference on Analytic Functions and Related Topics (Chełm, July 5–9, 2009). The attention was focused on the work of N. Kazarinoff who unveiled the importance of generalized hypergeometric functions for de Branges reasoning. Kazarinoff points to the Clausen identity, Gegenbauer formula and Rainville integral as main ingredients in the final part of de Branges proof. The Bieberbach conjecture stated in 1916 was affirmed in the famous paper of de Branges (1984, [Brn 85]).

Besides the above articles I was helped by some books in which de Branges result has been discussed. See the trilogy by Henrici 1986 (third volume dedicated to S. Bergman), Conway 1995 and Gong 1999. The present small monograph attemts to report the general reception of [Brn 85] as of 2009.

In order to round off the general picture I offered two articles [Skw 09a,b] dealing with functions ${}_2F_1, {}_3F_2$. In the first Clausen identity is derived using direct computation and rudimentary properties of Fuchsian singularity. In the second Watson lemma is used to verify initial conditions in the solution to de Branges differential system.

Quite recently A. K. Rathie and R. B. Paris have published [R,P] a new beautiful proof of Watson summation theorem (which implies Watson lemma). The present text reflects all these developments. The author hopes that the importance of the subject and some originality in the arrangement of topics will prevail over the remaining insufficiency.

0.3. Initial remark

Acronyms below usually consist of first three consonants of author's name, followed by the year of publication. List of references is constructed alphabetically according to letters in the acronym. When no ambiquity results acronyms on this list appear without the year of publication.

In Section 2 we shall discuss Löwner parametric method which in 1923 led him to the proof of inequality $|a_3| \leq 3$ and plays a fundamental role in the affirmation of Bieberbach conjecture. Carathéodory convergence is used in deriving analytic conclusions with this method.

1. Bieberbach conjecture. Distorsion lemma. Carathéodory theorem

1.1. Univalent functions and Bieberbach conjecture

According to the well known Riemann mapping theorem every simply connected domain $D \subset \mathbb{C}$, with exception of $\emptyset, \mathbb{C}$, can be mapped conformally onto the unit disc Δ. Note that the inverse mapping $f : \Delta \to D$ is conformal. In general any function f which maps conformally Δ onto $f(\Delta) \subset \mathbb{C}$ is called *univalent* or *schlicht*. The Bieberbach class S consists of univalent functions $f(z)$, $z \in \Delta$ which satisfy

$f(0) = 0, f'(0) = 1$. Bieberbach proved (1916) [Bbr] (his portrait is presented in [B,D,D,M]) that for every $f \in S$ the power development

$$f(z) = z + a_2 z^2 + a_3 z^3 + a_4 z^4 + \dots, \qquad z \in \Delta \tag{1}$$

satisfies $|a_2| \leq 2$; moreover $|a_2 = 2$ if and only if f (up to rotation) is the Koebe function

$$\mathbf{K}(z) := \frac{z}{(1-z)^2} = z + 2z^2 + 3z^3 + \dots. \tag{2}$$

In a footnote (*vielleicht uberhaupt...*) Bieberbach expressed an expectation that $|a_n| \leq n$ for every $n = 2, 3, \dots$. In 1984 this *Bieberbach conjecture* was finally affirmed, Monographs [Drn], [Gdz] were published too early to mention this achievement. But in subsequent years many authors took up the subject. Their work is of definite interest. Since famous proofs usually evolve with time, critical remarks are welcomed by textbook authors. The present text was inspired (above all) by [F,P 85], [Krv 86], [Kzr 88], [Grn 99].

1.2. Distorsion of absolute value in S

Distortion lemma follows from Bieberbach inequality $|a_2| \leq 2$. We need it to prove the Carathéodory (his portrait is presented in Wikipedia: http://pl.wikipedia.org /wiki/Constantin_Catheodory) convergence theorem. It plays an eminent role in de Branges (his portrait is presented in [B,D,D,M]) proof of $|a_n| \leq n$, $n = 3, 4, \dots$.

(1) Distorsion Lemma (cf. [Ahl], pp. 84–85). *All functions $f \in S$ satisfy inequalities*

$$\frac{1-r}{(1+r)^3} \leq |f'(z)| \leq \frac{1+r}{(1-r)^3}, \qquad 0 < r = |z| < 1, \tag{3}$$

$$\frac{r}{(1+r)^2} \leq |f(z)| \leq \frac{r}{(1-r)^2}, \qquad 0 < r = |z| < 1. \tag{4}$$

Proof. The reasoning consists of three steps

1. A composition of conformal automorphism $\gamma : \Delta \to \Delta$ with $f \in S$ has development

$$(f \circ \gamma)(z) = (f \circ \gamma)(0) + (f \circ \gamma)'(0)z + +(1/2)(f \circ \gamma)''(0)z^2 + \dots. \tag{5}$$

Composition (5) is conformal. It becomes an element of S but after suitable normalization. This, together with Bieberbach inequality, yields

$$\left| \frac{1}{2} \, \frac{(f \circ \gamma)''(0)}{(f \circ \gamma)'(0)} \right| \leq 2. \tag{6}$$

Take arbitrary point $\xi \in \Delta$ and choose automorphism $\gamma : \Delta \to \Delta$ which maps 0 onto ξ

$$\gamma(z) := \frac{z + \xi}{1 + \bar{\xi} z}, \qquad z \in \Delta. \tag{7}$$

By immediate computation

$$\gamma'(z) = \frac{1-|\xi^2|}{(1+\bar{\xi}z)^2}, \qquad \gamma'(0) = 1-|\xi|^2, \tag{8}$$

$$\gamma''(z) = (1-|\xi|^2)\frac{-2\bar{\xi}}{(1+\bar{\xi}z)^3}, \qquad \gamma''(0) = (-2\bar{\xi})(1-|\xi|^2). \tag{9}$$

As a consequence

$$\begin{aligned} (f\circ\gamma)'(z) &= f'(\gamma(z)\gamma'(z)),\\ (f\circ\gamma)''(z) &= f''(\gamma(z))(\gamma'(z))^2 + f'(\gamma(z))\gamma''(z), \end{aligned} \tag{10}$$

$$\begin{aligned} (f\circ\gamma)'(0) &= f'(\xi)(1-|\xi|)^2,\\ (f\circ\gamma)''(0) &= f''(\xi)(1-|\xi|^2)^2 + f'(\xi)(-2\bar{\xi})(1-|\xi|^2). \end{aligned} \tag{11}$$

Hence inequality (6), divided by $1-|\xi|^2$ takes the form

$$\left|\left(\frac{f''(\xi)}{f'(\xi)} - \frac{2\bar{\xi}}{1-|\xi|^2}\right)\right| \le \frac{4}{1-|\xi^2|}. \tag{12}$$

2. Absolute value of Cauchy integral does not exceed the integral of absolute value with respect to arc length. In view of (12) its integration over $[0,z]$, $z\in\Delta$ yields

$$\left|\int\limits_{[0,z]}\left(\frac{f''(\xi)}{f'(\xi)} - \frac{2\bar{\xi}}{1-|\xi|^2}\right)d\xi\right| \le \int\limits_{[0,z]}\frac{4}{1-|\xi|^2}|d\xi|. \tag{13}$$

By immediate calculation

$$\int\limits_{[0,z]}\frac{f''(\xi)}{f'(\xi)}d\xi = \ln f'(\xi)\Big|_{\xi=0}^{\xi=z} = \ln f'(z). \tag{14}$$

Moreover, with parametrization $\xi(s) = sz$, $s\in[0,1]$

(15)

$$\int\limits_{[0,z]}\left(\frac{2\bar{\xi}}{1-|\xi|^2}\right)d\xi = \int\limits_0^1 \frac{2s|z|^2}{1-s^2|z|^2}dm(s) = -\ln(1-s^2|z|^2)\big|_{s=0}^{s=1} = -\ln(1-|z|^2).$$

The same parametrization is used to compute the right-hand side in (13). Namely

$$\begin{aligned} \int\limits_{[0,z]}\frac{4}{1-|\xi|^2}|d\xi| &= 2|z|\int\limits_0^1\left(\frac{1}{1+s|z|}+\frac{1}{1-s|z|}\right)dm(s)\\ &= 2|z|\left(\frac{1}{|z|}\ln(1+s|z|) - \frac{1}{|z|}\ln(1-s|z|)\right)\big|_{s=0}^{s=1} = 2\ln\frac{1+|z|}{1-|z|}. \end{aligned} \tag{16}$$

In view of (14), (15), (16) inequality (13) is rewritten as

$$\big|\ln f'(z) + \ln(1-|z|^2)\big| \le 2\ln\frac{1+|z|}{1-|z|}. \tag{17}$$

On the left-hand side the expression under absolute value can be replaced by its real part. This results in double inequality

$$\ln\frac{1}{1-|z|^2} - 2\ln\frac{1+|z|}{1-|z|} \le \ln|f'(z)| \le \ln\frac{1}{1-|z|^2} + 2\ln\frac{1+|z|}{1-|z|} \tag{18}$$

and exponentiation yields the desired inequality (3).

3. In the last step we prove the double inequality (4). Estimate from above follows easily from (3). Indeed, direct computation with parametrization $\xi(s) = sz$, $s \in [0,1]$ yields

$$\begin{aligned} |f(z)| &= \left|\int_{[0,z]} f'(\xi)d\xi\right| \le \int_{[0,z]} \frac{1+|\xi|}{(1-|\xi|)^3}|d\xi| \\ &= |z|\int_0^1 \frac{1+s|z|}{(1-z|z|)^3}dm(s) \le |z|\int_0^1 \frac{1+|z|}{(1+|z|)^3}dm(s) = \frac{|z|}{(1+|z|)^2}. \end{aligned} \tag{19}$$

It the following we are concerned with the estimate from below. For $z \in \Delta$, $r := |z|$ consider the circle

$$C_r := \{\xi \in \Delta; |\xi| = r\} \tag{20}$$

and denote $m(r) = |w|$ where $w \in f(C_r) = \{f(\xi); |\xi| = r\}$ is any point with the smallest distance to $f(0) = 0$. The segment in Δ joining 0 with w can be thought of as an image of a smooth curve $L \subset \Delta$ under the conformal mapping f, see Fig. 1 below.

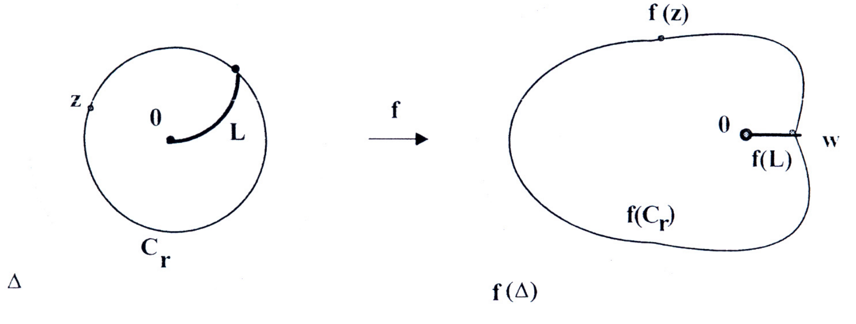

Fig. 1: The line $L \subset \Delta$ corresponds to the segment $f(L) = [0, w]$ of length $m(r)$.

Conformal mapping f distorts length at $\xi \in \Delta$ multiplying it by $|f'(\xi)|$. The length of $f(L)$ is expressed by an integral with respect to arc length of L. Therefore

$$(21)\qquad |f(z)| \geq m(r) = \text{ length of } f(L) = \int_L |f'(\xi)|\,|d\xi| \geq \int_L \frac{1-|\xi|}{(1+|\xi|)^3}|d\xi|.$$

In the end the inequality (3) was used. We want to show that

$$(22)\qquad \int_L \frac{1-|\xi|}{(1+|\xi|)^3}|d\xi| \geq \int_0^r \frac{1-\rho}{(1+\rho)^3}d\rho = \left(\frac{-1}{(1+\rho)^2}+\frac{1}{1+\rho}\right)\Bigg|_0^r = \frac{r}{(1+r)^2}.$$

In (22) the first inequality requires an explanation. Consider the integral over $[0,r]$ and let σ be the *lower* riemannian sum related to the division $0=\rho_0<\rho_1<\ldots<\rho_s=r$.

By definition

$$(23)\qquad \sigma := \sum_{i=0}^{s-1} m_i(\rho_{i+1}-\rho_i) \ \text{ where } \ m_i := \inf\left\{\frac{1-\rho}{(1+\rho)^3}; \rho \in [\rho_i,\rho_{i+1}]\right\}.$$

Let L_i be *arbitrarily chosen* segment on L beginning at ξ_i and ending at ξ_{i+1} where $|\xi_i|=\rho_i$, $|\xi_{i+1}|=\rho_{i+1}$. Replacing it by a subsegment (if necessary) we may assume, that all inner points of L_i lie in the open annulus $\sigma_i<|z|<\sigma_{i+1}$. Hence different segments L_i have disjoint interiors. In general the segments L_i do not sum up to L, so we have an inequality

$$(24)\qquad \int_L \frac{1-|\xi|}{(1+|\xi|)^3}|d\xi| \geq \sum_i \frac{1-|\xi|}{(1+|\xi|)^3}|d\xi|.$$

Obviously $|\xi_{i+1}-\xi_i| \geq \rho_{i+1}-\rho_i$. Moreover, by definition of m_i,

$$(25)\qquad \begin{aligned}\sum_i \int_{L_i} \frac{1-|\xi|}{(1+|\xi|)^3}|d\xi| &\geq \sum \int_{L_i} m_i|d\xi| = \sum_i m_i \ \text{ (length of } L_i) \\ &\geq \sum_i m_i|\xi_{i+1}-\xi_i| \geq \sum_i m_i(\rho_{i+1}-\rho_i) = \sigma.\end{aligned}$$

In the inequality which follows from (24), (25) we pass to the limit when tolerance of relevant division goes to 0. Then σ converges to the second integral in (22). With (21) and (22) the desired estimate from below in (4) is established. □

(2) Corollary (Bieberbach). *For every $f \in S$ the image $f(\Delta)$ contains the disc with center* 0 *and radius* 1/4.

Proof. When $r \to 1$ the lower limit of $m(r)$ is not smaller than 1/4 in view of (4). □

(3) *Remark.* The number 1/4 in this corollary cannot be improved. It is called *Koebe constant*, see p. 65.

1.3. The kernel of a sequence of domains

Carathéodory's result on sequences of univalent mappings $f_m : \Delta \to f_m(\Delta)$ occupies central place in the theory of univalent functions. A detailed proof of this theorem is presented in the next section. We need two important notions.

(4) *Definition.* Let $G_m \subset \mathbb{C}$, $m = 1, 2, \dots$ be a sequence of domains which are simply connected and contain a fixed point w_0. We consider two possible cases

1. There is a neighbourhood of w_0 contained in every G_m. Denote by G the set of all points in $\mathbb{C}$ which, together with some neighbourhood, are included in G_m for all sufficiently large m. Obviously G is open and nonvoid (note that $w_0 \in G$). The point w_0 belongs to (unique) component G_{w_0} of G. We call G_{w_0} *the kernel of* G_m with respect to w_0.

2. In the opposite case, when no neighbourhood of w_0 is contained in *all* G_m, we declare *the kernel of* G_m w.r.t. w_0 equal to $\{w_0\}$.

(5) *Remark.* See [Glz 52], p. 62, [Mrk 68], vol. 2, pp. 37–38, and [Gng 99], p. 36. The definition in [Crt 32], p. 91, is more general.

(6) *Definition.* Let $G_m \subset \mathbf{C}$, $m = 1, 2, \dots$ be a sequence of simply connected domains with kernel G_0. If every subsequence G_{m_k}, $k = 1, 2, \dots$, has G_0 as its kernel, we say that G_m is *kernel convergent* (briefly: k-convergent) to G_0. Symbolically,

$$G_n \xrightarrow{\mathrm{k}} G_0. \tag{26}$$

(7) *Examples* (from [Mrk 68] p. 38).

(a) Fig. 2 shows fixed disjoint rectangles Q', Q'' joined by a horizontal rectangle Q_m of height $1/m$. Choose fixed points $w' \in Q'$, $w'' \in Q''$. Let $G_m := Q' \cup Q_m \cup Q''$. Then $G_{w'} = Q'$, $G_{w''} = Q''$. It is easy to see that G_m is k-convergent both to $G_{w'}$ and $G_{w''}$.

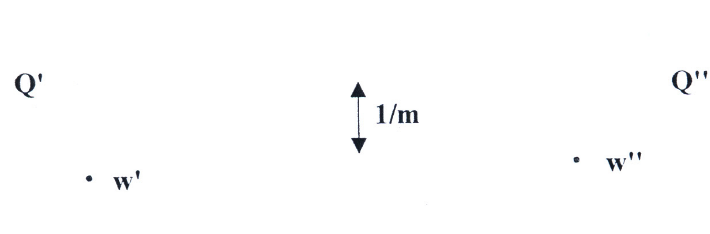

Fig. 2: Sequence G_m is k-convergent to $G_{w'} = Q'$, $G_{w''} = Q''$.

(b) Fig. 3 shows fixed rectangles Q', Q'' overlapping along the rectangle $Q = Q' \cap Q''$. Choose fixed point $w_0 \in Q$. Define $G_m = Q'$ for $m = 2k - 1$ and for $m = 2k$. Obviously $G_{w_0} = Q$. Since the kernel of G_{2k-1} with respect to w_0 is Q' and the kernel of G_{2k} with respect to w_0 is Q'', the sequence G_m is not k-convergent to G_{w_0}.

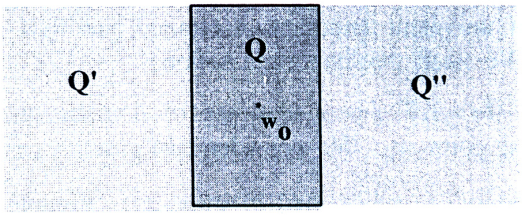

Fig. 3: Sequence $Q', Q'', Q', \ldots$ is not k-convergent to $Q = Q' \cap Q''$ w.r.t. w_0.

1.4. Carathéodory convergence theorem

The famous Carathéodory monograph [Crt 32] deals with mappings which are only locally conformal. For our purpose this less general result is sufficient. Since it plays a fundamental role in de Branges reasoning we supply the relevant proof (see also [Glz 52], pp. 62–67, and [Gng 99], pp. 36–37).

(8) Carathéodory convergence theorem. *Consider a sequence $f_m : \Delta \to D_m$ of conformal mappings, normalized by $f_m(0) = 0$, $f'_m(0) > 0$. Note that $0 \in D_m = f_m(\Delta)$ for all m. The sequence f_m converges locally uniformly to f if and only if D_0, the kernel w.r.t. 0 of D_m, is different from $\mathbb{C}$ and $D_m \xrightarrow{\mathrm{k}} D_0$. These (equivalent) conditions imply that $f(\Delta) = D_0$. Under the additional assumption $D_0 \neq \{0\}$ the mapping f is conformal, its image D_0 is simply connected and the sequence of inverse mappings $f_m^{-1} : D_m \to \Delta$ converges locally uniformly to $f^{-1} : D_0 \to \Delta$.*

Proof. We prove the main equivalence. (Remaining claims should become obvious in view of supplied arguments). The proof consists of two parts, concerned with necessity and sufficiency, respectively.

Part I. Assume that f_m converges locally uniformly to f. By Weierstrass' theorem f is holomorphic. Hence $f_m : \Delta \to \mathbb{C}$ is locally bounded. There are two possibilities.

1) The case $f = \text{const}$. Then $f \equiv 0$. We shall prove that $f(\Delta) = \{0\}$ is the kernel of D_m, $m = 1, 2, \ldots$, w.r.t. 0. Assume, to the contrary, that $D_0 \neq \{0\}$. By Definition (3) there is a disc $|w| < \rho$ contained in each of $D_m = f_m(\Delta)$. Schwarz lemma applied to f_m^{-1} yields $(f_m^{-1})'(0) \leq 1/\rho$, and hence $f'_m(0) \geq \rho$. Since, by the Weierstrass theorem, $f'_m \to (0)' = 0$, we have a contradiction. Hence $D_0 = \{0\}$ as claimed. The above reasoning applies to any subsequence of f_m. It follows that D_m is k-convergent to $\{0\}$.

2) The case $f \neq \text{const}$. Then, by the Hurwitz theorem, the limit $f : \Delta \to f(\Delta)$ is conformal. Our reasoning in this case consists of four steps.

1. We shall show first that $f(\Delta) \subset D_0$. Since $f(\Delta)$ is connected it suffices to prove that every point $b = f(a)$, where $a \in \Delta$, has a neighbourhood contained in every D_m with m large enough. To this aim we shall apply the Rouché

theorem. Choose discs U, V with radii $r, 2r$, common centre a and closures contained in Δ. For $z \in U$ the function $f(\xi) - f(z)$ has in V exactly one zero (attained at $\xi = z$) and is bounded away from 0 on bV. Since on bV the sequence $f_m(\xi) - f(z)$ converges uniformly to $f(\xi) - f(z)$, we have the inequality

$$|(f_m(\xi) - f(z)) - (f(\xi) - f(z))| < |f(\xi) - f(z)|, \quad \xi \in \mathrm{b}V \tag{27}$$

for m large enough. Note the identity

$$f_m(\xi) - f(z) = [f(\xi) - f(z)] + [(f_m(\xi) - f(z)) - (f(\xi) - f(z))]. \tag{28}$$

In view of (28) and (27), for m large enough, both $f_m(\xi) - f(z)$ and $f(\xi) - f(z)$ have the same number of zeros in V (Rouché's theorem). Hence the value $f(z)$ is attained by $f_m(\xi)$, $\xi \in V$ iff it is attained by $f(\xi)$, $\xi \in V$. Since $z \in U$ the latter means that $f(z)$ is attained by $f(\xi)$, $\xi \in U$. It follows that $f(U) = f_m(V) \subset D_m$. Note that $f(U)$ is a neighbourhood of b (independent of m) contained in D_m for all m large enough. Since b was chosen arbitrarily, we have $f(\Delta) \subset D_0$, as claimed.

2. We show the reverse inclusion $D_0 \subset f(\Delta)$. Consider arbitrary $w_0 \in D_0$. It belongs to a domain W such that $W \subset D_m$ for sufficiently large m. We may assume that W contains 0. For sufficiently large m the domain W is mapped conformally by $\varphi_m := f_m^{-1}$ and $\varphi_m(W) \subset \Delta$. Hence there exists a subsequence φ_{m_k} which converges locally uniformly in W to a holomorphic function φ. Note that

$$\varphi'(0) = \lim_{k\to\infty} \varphi'_{m_k}(0) = \lim_{k\to\infty} \frac{1}{f'_{m_k}(0)} = \frac{1}{f'(0)} \neq 0. \tag{29}$$

This shows that φ is nonconstant, hence (by Hurwitz's theorem) it maps conformally W onto $\varphi(W) \subset \Delta$. It follows that

$$\varphi_{m_k}(w_0) \to \varphi(w_0) \in \Delta \tag{30}$$

and, by composing (30) with f,

$$f(\varphi(w_0)) = \lim_{k\to\infty} f(\varphi_{m_k}(w_0)) = \lim_{k\to\infty} f_{m_k}(\varphi_{m_k}(w_0)) = \lim_{k\to\infty} w_0 = w_0. \tag{31}$$

Since $w_0 \in D_0$ was chosen arbitrarily this yields $D_0 \subset f(\Delta)$ as claimed. The converse inclusion has been proven in Step 1, and hence $G_0 = f(\Delta)$. One cannot map conformally Δ onto $\mathbb{C}$, so $D_0 \neq \mathbb{C}$. In addition (31) implies that $\varphi = f^{-1}$.

3. The reasoning in Step 2 can be applied to any subsequence f_{m_k} of f_m. Since such a subsequence converges locally uniformly to $f \neq \mathrm{const}$ we conclude that D_{m_k} has $f(\Delta) = D_0$ as its kernel. Hence D_m is k-convergent to D_0 as claimed.

4. The reasoning in Step 2 can be applied to any subsequence φ_{m_k} which is convergent in W. Such a subsequence φ_{m_k} converges in W to the limit $\varphi = f^{-1}$ (independent of m_k). It follows that φ_m itself converges in W to $\varphi = f^{-1}$. By the Vitali theorem the sequence $\varphi_m : D_m \to \Delta$ converges locally uniformly in D_0 to $\varphi = f^{-1}$ as claimed.

Part II. Assume that $G_m := f_m(\Delta)$ is k-convergent to $G_0 \neq \mathbb{C}$. We shall show that f_m converges locally uniformly in Δ. Distortion lemma applied to f_m yields

$$|f_m(z)| \leq |f'_m(0)| \frac{|z|}{(1-|z|)^2}, \qquad z \in \Delta. \tag{32}$$

There are two possibilities.

1) The case of $D_0 = \{0\}$. Then we have $f'_m(0) \to 0$. Indeed, assume to the contrary that there is the subsequence $f'_{m_k}(0)$ bounded away from 0. By Corollary (2) there is a neighbourhood of 0 contained in all D_{m_k}. Hence D_{m_k} has kernel different from $\{0\}$ contradicting the assumption that D_m is k-convergent to $\{0\}$.

Since $|f'_m(0)| \to 0$, inequality (32) implies that the sequence f_m converges to the constant 0 locally uniformly in Δ, as claimed.

2) The case $D_0 \neq \{0\}, \mathbb{C}$. Then the sequence $f'_m(0)$ is bounded. Indeed, assume the opposite. Then some subsequence $f'_{m_k}(0)$ diverges to ∞. Hence, by Koebe 1/4 theorem the kernel of D_{m_k} equals $\mathbb{C}$ contrary to the assumption that $D_0 \neq \{0\}$. Since $f'_m(0)$ is bounded the inequality (32) implies that the sequence f_m is locally bounded in Δ. By part I, the limit f of any convergent subsequence f_{m_k} maps Δ onto D_0, the kernel of D_{m_k}. The case $D_0 = \{0\}$ was excluded, $f \neq$ const is *the normalized* conformal mapping of Δ onto D_0. This mapping is independent of m_k and hence the sequence f_m converges locally uniformly to f according to the claim.

□

2. Löwner chain. Löwner equation. Milin functional

2.1. Löwner chain

This section splits into four steps.

1. For $f \in S$ and every $r \in (0,1)$ the function $r^{-1}f(rz)$, $z \in \Delta$, is holomorphic and invertible in some neighbourhood of the closed unit disc. Note that $r \to 1$ implies that $r^{-1}f(rz) \to f(z)$ locally uniformly in Δ. Therefore we may assume (without loss of generality) that the original function f maps biholomorphically a neighbourhood of clΔ. As a consequence, f maps the unit disc Δ onto a domain D bounded by an analytic Jordan curve C, see Fig. 4.

2. Since $0 \in D$ there is the halfline $(-\infty, w_C) \subset (-\infty, 0) \cap (\mathbf{C} \setminus D)$ such that $w_C \in C = f(b\Delta)$; see Fig. 5.

3. Fig. 6 below shows the Jordan arc L_m which consists of interval $(-\infty, w_C]$ followed by a part of C from w_C to a point w_m on C. For a slit domain $D_m := \mathbb{C} \setminus \mathrm{cl} L_m$ let $f_m : \Delta \to D_m$ be the Riemann mapping, normalized by $f_m(0) = 0$, $f'_m(0) > 0$.

Following a remarkable paper of de Branges [Brn 86] we take the liberty to apply the term *Riemann mapping* not only to a mapping onto Δ, but also. The intended meaning is visible from the context.

Assume that the arcs $L_1, L_2, \ldots$ are increasing and $w_m \to w_C$ when $m \to +\infty$. The exterior of $\cup L_m$ is disconnected. Its component D, determined by $0 = f(0)$, equals $f(\Delta)$. By Carathéodory's convergence theorem the sequence $f_m : \Delta \to D_m$ converges locally uniformly in Δ to $f : \Delta \to D$. Hence $f_m'(0) \to f'(0) = 1$ and $f_m/f_m'(0) \to f$. It is therefore enough to prove the Bieberbach conjecture for suitable normalized mappings of Δ onto the slited plane D_m.

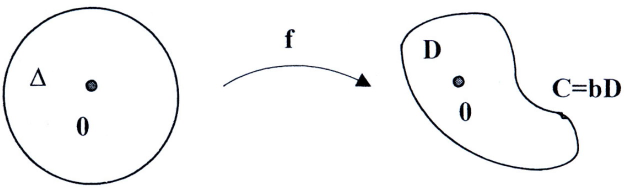

Fig. 4: Jordan curve $C = f(b\Delta)$.

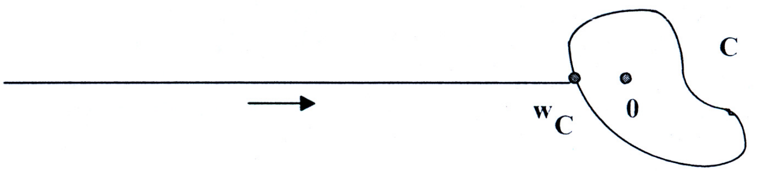

Fig. 5: The halfline $(-\infty, w_C)$.

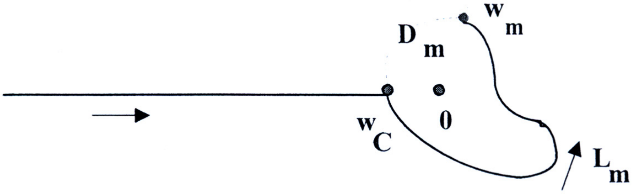

Fig. 6: Slited plane $D_m = f_m(\Delta)$.

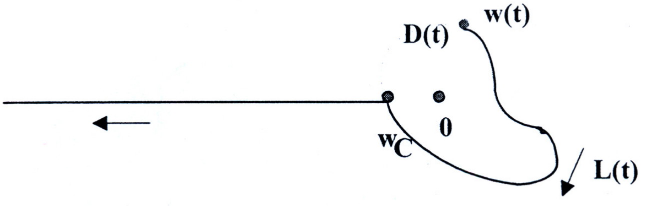

Fig. 7: After reparametrization. Increasing family $D(t)$ $t \in (0, +\infty)$.

4. Finally we modify previous considerations by replacing countable the family L_m with the continuous family $L(t)$ $t \in (-\infty, 0)$, where the arc $L(t)$ joins $-\infty$ with a point $w(t)$ on the slit. Consider the Jordan arc L parametrized by $w(t)$, $t \in (0, +\infty)$, and denote by $D(t)$ the complement of $\operatorname{cl} L(t)$. We note that $D(t)$ *decreases* when t increases and k-converges to D when $t \to 0$. After suitable reparametrization (the new parameter again denoted by t) which changes the orientation of L, we obtain *an increasing* family $D(t)$, $t \in (0, +\infty)$. Note that $D(t)$ approaches $\mathbb{C}$ when $t \to +\infty$; see Fig. 7.

Consider now the Riemann mapping $f_t : \Delta \to D(t)$ with the Taylor development

$$f_t(z) = a(t)z + \dots \quad a(t) > 0. \tag{33}$$

When $0 < s < t < +\infty$, the composition $f_t^{-1} \circ f_s$ maps Δ into Δ and Schwarz's lemma yields

$$|(f_t^{-1} \circ f_s)'(0)| < 1. \tag{34}$$

In view of $f_s(0) = 0$ it follows that $f_s'(0) < f_t'(0)$. Equivalently,

$$a(s) < a(t). \tag{35}$$

Thus the first coefficient $a(t)$ in (33) increases from 1 to $+\infty$ with $t \in (0, +\infty)$. Its logarithm varies from 0 to $+\infty$ and can be taken as a new parameter. There is no loss of generality to assume that

$$f_t(z) = e^t z + a_2(t)z^2 + a_3(t)z^3 + \dots \tag{36}$$

with $D(t) := f_t(\Delta)$, $t \in (0, +\infty)$, increasing. Note that in the limit, $D(0) = D$, $D(+\infty) = \mathbb{C}$. The function of two variables

$$f(z,t) := f_t(z), \qquad z \in \Delta, \quad t \in (0, +\infty), \tag{37}$$

is caled the *Löwner chain* (the portrait of Ch. Löwner is presented in Wikipedia: http://en.wikipedia.org/wiki/Charles_Loewner). For $t = 0$ it reduces to $f \in S$ while for sufficiently large t it reduces to the Riemann mapping onto the plane slited along a negative ray.

2.2. Löwner equation

A Löwner chain $f(z,t)$ satisfies the *Löwner equation*

$$\frac{\partial f}{\partial f} = \cdot p(z,t) \left(z \frac{\partial f}{\partial z} \right), \tag{38}$$

where $p(z,t)$ is holomorphic in z and

$$\operatorname{Re} p(z,t) > 0, \qquad p(0,t) = 1. \tag{39}$$

The original proof in [Lwn 23] is difficult. The reasoning below follows from [Drn 83], [Krv 86], [Ahl 73]. For $s < t$ consider the composition $f_t^{-1} \circ f_s : \Delta \to \Delta$; see Fig. 8. Recall that $D(t) \supset D(s)$. Denote by Γ_{st} the part of the arc L between $w(t)$ and $w(s)$. Note that f_t^{-1} maps $D(s)$ onto Δ without $f_t^{-1}(D(t) \setminus D(s))$. From known results on boundary correspondence follows that f_t^{-1} maps $D(s)$ onto the disc Δ

slited along an arc $f_t^{-1}(\Gamma_{st})$. The begining $\gamma(t)$ of this slit lies on $b\Delta$ as the image of $w(t)$ under f_t^{-1}. The end lies inside Δ as the image of $w(s) \in D(t)$ under f_t^{-1}. The map $\gamma : (0, +\infty) \to b\Delta$ (independent of s) is continuous, as proven in [Drn 83], p. 85. In view of (36) $f_t^{-1} \circ f_s$ has the Taylor development

$$\varphi(z) := f_t^{-1}(f_s(z)) = e^{s-t}z + \dots \quad z \in \Delta. \tag{40}$$

Let us eliminate the only zero in (40) and take the branch of logarithm. This yields

$$\Phi(z) := \ln \frac{(f_t^{-1} \circ f_s)(z)}{z} = \ln \frac{e^{s-t}z + \dots}{z} = \ln(e^{s-t} + \dots), \qquad z \in \Delta, \tag{41}$$

where $\Phi(0) = s - t$.

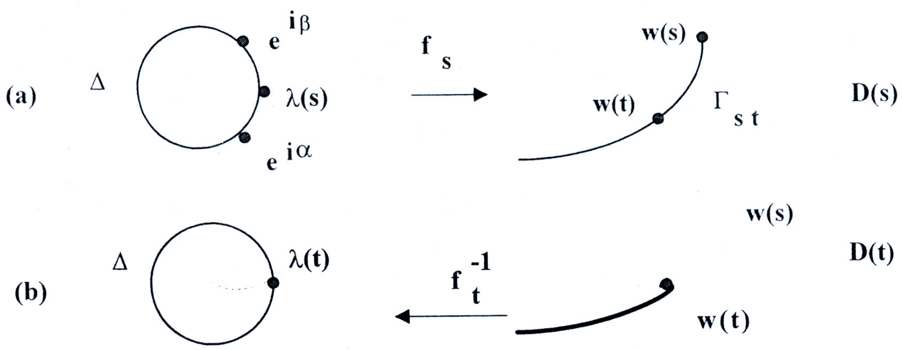

Fig. 8: Composition $f_t^{-1} \circ f_s : \Delta \to \Delta$. Note that $D(t) \supset D(s)$.

As shown in Fig. 8 the arc on bΔ (between $e^{i\alpha}$ and $e^{i\beta}$) goes under f_s onto Γ_{st}, and is mapped by $f_t^{-1} \circ f_s$ into the interior of Δ. Other points of bΔ are mapped under $f_t^{-1} \circ f_s$ into bΔ. This has important consequences. The harmonic function

$$\operatorname{Re}\Phi(z) = \ln \left| \frac{(f_t^{-1} \circ f_s)(z)}{z} \right| = \ln |(f_t^{-1} \circ f_s)(z)|, \quad z \in \Delta, \tag{42}$$

considered on bΔ is negative on the arc between e^{ia} and e^{ib} and is zero at other points of bΔ. Let us recall the Poisson kernel for Δ:

$$P(z, \lambda) = \operatorname{Re} \frac{\lambda + z}{\lambda - z}, \quad z \in \Delta, \quad \lambda \in \mathrm{b}\Delta. \tag{43}$$

A function holomorphic in Δ is determined by its real part up to imaginary constant. The relevant integral formula is known as the Schwarz-Poisson representation. For $\Phi(z)$ this representation yields

$$\begin{aligned}\Phi(z) &= \frac{1}{2\pi}\int\limits_0^{2\pi}\left[\mathrm{Re}\Phi(e^{i\theta})\right]\frac{e^{i\theta}+z}{e^{i\theta}-z}dm(\theta) \\ &= \frac{1}{2\pi}\int\limits_\alpha^\beta \ln\left|(f_t^{-1}\circ f_s)(e^{i\theta})\right|\frac{e^{i\theta}+z}{e^{i\theta}-z}dm(\theta).\end{aligned}\tag{44}$$

There is no additional constant since at $z=0$ all expressions in (44) are real in view of definition (41). Since $\Phi(0)=s-t$, (44) yields

$$s-t=\frac{1}{2\pi}\int\limits_\alpha^\beta \ln\left|(f_t^{-1}\circ f_s)(e^{i\theta})\right|dm(\theta).\tag{45}$$

More generally, for fixed $z=f_s^{-1}(w)$ we have by (41)

$$\Phi(z)=\ln\frac{(f_t^{-1}\circ f_s)(z)}{z}=\ln\frac{f_t^{-1}(w)}{f_s^{-1}(w)}\tag{46}$$

and (44) yields

$$\ln\frac{f_t^{-1}(w)}{f_s^{-1}(w)}=\frac{1}{2\pi}\int\limits_\alpha^\beta \ln\left|(f_t^{-1}\circ f_s)(e^{i\theta})\right|\frac{e^{i\theta}+f_s^{-1}(w)}{e^{i\theta}-f_s^{-1}(w)}dm(\theta).\tag{47}$$

We divide both sides of (47) by $s-t$ and consider the limit when $s\to t$. The quotient on the left-hand side converges to the derivative $-(\partial/\partial t)\ln f_t^{-1}(w)$. With $s\to t$ the slit in Δ reduces to the single point $\lambda(t)\in \mathrm{b}\Delta$ and the second factor under integral sign in (47) converges to a constant

$$\frac{\lambda(t)+f_t^{-1}(w)}{\lambda(t)-f_t^{-1}(w)}.\tag{48}$$

In view of (45) the integral of the first factor under the integral sign in (47) equals $s-t$. These observations lead to the equality

$$\frac{\partial}{\partial t}\ln f_t^{-1}(w)=-\frac{\lambda(t)+f_t^{-1}(w)}{\lambda(t)-f_t^{-1}(w)}.\tag{49}$$

In the final part we follow [Ahl 73] on p. 96. Formula (49) can be rewritten as p.d.e. satisfied by $f_t(z)$. Note that for fixed z

$$f_t^{-1}(w)=z.\tag{50}$$

We differentiate both sides of latter equality with respect to t. The right-hand side yields 0. Since, by the inverse function theorem, the left-hand side of (50) has derivative with respect to w; the desired derivative with respect to t can be computed with the chain rule for functions of two variables. Since $\partial w/\partial t=\partial f_t(z)/\partial t$, the result is

$$\frac{\partial f_t^{-1}(w)}{\partial t}+\frac{\partial f_t^{-1}(w)}{\partial w}\;\frac{\partial f_t(z)}{\partial t}=0.\tag{51}$$

We now use (51) to calculate the left-hand side of (49). It follows that

$$\frac{\partial f_t^{-1}(w)}{\partial t} = -z\frac{\lambda(t)+z}{\lambda(t)-z}. \tag{52}$$

Substituting (52) into (51) and dividing by $\partial f_t^{-1}(w)/\partial w$ or (what is the same) by $1 : \partial f_t(z)/\partial z$ yields

$$\frac{\partial f_t(z)}{\partial t} = z\frac{\lambda(t)+z}{\lambda(t)-z}\frac{\partial f_t(z)}{\partial z} = zp(z,t)\frac{\partial f_t(z)}{\partial z}. \tag{53}$$

This is the famous Loewner equation (38).

It is easy to verify that for $z \in \Delta$, $t \in [0, +\infty)$, the factor $p(z,t)$ has positive real part (belongs to the right halfplane). Since $|\lambda(t)| = 1$, we have

$$2\mathrm{Re}\; p(z,t) = \frac{\lambda+z}{\lambda-z} + \frac{\bar\lambda+\bar z}{\bar\lambda-\bar z} = 2\frac{|\lambda|^2-|z|^2}{|\lambda-z|^2} > 0. \tag{54}$$

□

Some geometric interpretation follows. Löwner's equation (38) describes dynamics of the family $f_t : \Delta \to D_t$. At $z \in \mathrm{b}\Delta_r$ vector z is orthogonal to $\mathrm{b}\Delta_r$. Under variable t the point $f_t(z)$ describes "trajectory" of the point z. The partial derivative $\partial f_t(z)/\partial t \in \mathbb{C}$ describes the vector tangent to this trajectory. The differential of conformal mapping $f_t(z)$ amounts to multiplication by $\partial f_t(z)/\partial z$. The latter operator preserves angles, and hence (at $f_t(z)$) the vector $(\partial f_t(z)/\partial z)\cdot z$ is orthogonal to $\mathrm{b}f(\Delta_r)$. Since $p(z,t)$ belongs to the right halfplane, its argument has absolute value smaller than $\pi/2$. Therefore the Löwner equation (38) indicates that at $f_t(z)$ the trajectory goes out of $f_t(\Delta_r)$. Loosely speaking, Löwner chain $f_t(z) = f(z,t)$ describes "an expanding flow". Since $|\lambda(t)| = 1$ we may write

$$p(z,t) = \frac{\partial(t)+z}{\partial(t)-z} = \frac{1+\kappa(t)z}{1-\kappa(t)z}, \quad \text{where} \quad \kappa(t) := \overline{\lambda(t)}. \tag{55}$$

2.3. Robertson inequalities

Another conjecture, apparently more complicated, brought an important progress to the Bieberbach problem. In 1936 M. S. Robertson (his portrait is presented in [B,D,D,M]) expressed expectation that for every function $f(z) = z + a_2z^2 + a_3z^3 + \ldots$ in S the *odd* function

(56)
$$\sqrt{f(z^2)} = \sqrt{z^2}\,\sqrt{1+a_2z^2+a_3z^4+\ldots} = z + b_3z^3 + b_5z^5 + \ldots \in S, \quad b_1 = 1,$$

satisfies

$$\sum_{k=1}^{n} \left|b_{2k-1}\right|^2 \le n, \qquad n = 2, 3, \ldots . \tag{57}$$

This is the *Robertson conjecture* stated in terms of *Robertson inequalities* (57).

Since a_n in the development of $f(z^2)$ appears at z^{2n}, comparison with coefficients in $(z + b_3z^3 + b_5z^5 + \dots)^2$ yields

$$a_n = b_1b_{2n-1} + b_3b_{2n-3} + \dots + b_{2n-1}b_1. \tag{58}$$

Hence, by applying the Schwarz inequality to (58), conditions (57) give

$$|a_n| \le \sqrt{\sum_{k=1}^{n} |b_{2k-1}|^2} \cdot \sqrt{\sum_{k=1}^{n} |b_{2k-1}|^2} \le \left(\sqrt{n}\right)^2 = n, \quad n = 2, 3, \dots . \tag{59}$$

Therefore the Robertson conjecture implies Bieberbach conjecture, see See [Rbr 36].

2.4. Exponentiating a power series

After 1950 Lebedev and Milin (the portrait of I. M. Milin is presented in [B,D,D,M], but confused there with that of Ch. Löwner) began systematic investigation of exponentiation. This resulted in a number of general L-M inequalities. In the next section we shall use the inequality

$$\sum_{k=0}^{n} |\beta_k|^2 \le (n+1) \exp\left\{ \frac{1}{n+1} \sum_{k+1}^{n} (n+1-k) \left(k|\alpha_k|^2 - \frac{1}{k} \right) \right\} \tag{60}$$

which connects coefficients in $\varphi(z) = \alpha_1 z + \alpha_2 z^2 + \dots$ and in $\exp\varphi(z) = \beta_0 + \beta_1 z + \beta_2 z^2 + \dots$. Both sides of (60) contain coefficients at z^k where $k \le n$, but summation starts with $k = 0$ for $\exp\varphi$ and with $k = 1$ for φ.

In the present section, following Chapter 5 of [Drn 83], we recall a remarkable proof of (60) due to D. Aharonov. Denote

$$A_n := \sum_{k=1}^{n} k^2 |\alpha_k|^2, \qquad B_n := \sum_{k=0}^{n} |\beta_k|^2. \tag{61}$$

Differentiation of $e^{\varphi(z)}$ yields $e^{\varphi(z)}\varphi'(z)$. Therefore

$$(\beta_0 + \beta_1 z + \beta_2 z^2 + \dots)' = (\beta_0 + \beta_1 z + \beta_2 z^2 + \dots)\,(\alpha_1 z + \alpha_2 z^2 + \dots)'. \tag{62}$$

Calculating the Cauchy product on the right-hand side of (62) and comparing coefficients at z^{n-1} yields

$$n\beta_n = \sum_{k=1}^{n} k\alpha_k \beta_{n-k}, \qquad \beta_0 = 1. \tag{63}$$

Hence, by Schwarz's inequality,

$$n^2 |\beta_n|^2 \le \left(\sum_{k=1}^{n} k^2 |\alpha_k|^2 \right) \left(\sum_{k=0}^{n-1} |\beta_k|^2 \right) = A_n B_{n-1}. \tag{64}$$

We use (64) to estimate $|\beta_n|^2$. It follows that

$$\begin{aligned} B_n &= B_{n-1} + |\beta_n|^2 \le B_{n-1} + \frac{1}{n^2} A_n B_{n-1} = \left(1 + \frac{1}{n^2} A_n\right) B_{n-1} \\ &= \frac{n+1}{n} \left(1 + \frac{A_n - n}{n(n+1)} \right) B_{n-1} \le \frac{n+1}{n} B_{n-1} \exp\left(\frac{A_n - n}{n(n+1)} \right). \end{aligned} \tag{65}$$

In the latter line the elementary estimation $1+x \leq e^x$, valid for all $x \in \mathbf{R}$, was used.

From (65) and analogous estimates for $B_{n-1}, B_{n-2}, \ldots$ we infer the inequality

$$B_n \leq (n+1) \exp\left(\sum_{k=1}^{n} \frac{A_k - k}{k(k+1)} \right). \tag{66}$$

Since

$$\sum_{k=1}^{n} \frac{1}{k+1} + \sum_{k=2}^{n+1} + -1 + \sum_{k=1}^{n+1} \frac{1}{k}, \tag{67}$$

we may rewrite (66) as

$$B_n \leq (n+1) \exp\left\{ \sum_{k=1}^{n} \frac{A_k}{k(k+1) - \sum\limits_{k=1}^{n} \frac{1}{k+1}} \right\}. \tag{68}$$

We now come to the central part of the reasoning. In view of

$$s_n := \sum_{k=1}^{n} \frac{1}{k(k+1)} = 1 - \frac{1}{n+1}, \qquad n = 1, 2, \ldots \tag{69}$$

and $s_0 = 0$, summation by parts yields

$$\begin{aligned} \sum_{k=1}^{n} A_k \frac{1}{k(k+1)} &= \sum_{k=1}^{n} A_k (s_k - s_{k-1}) \\ &= (A_n s_n - A_n s_{n-1}) + (A_{n-1} s_{n-1} - A_{n-1} s_{n-2}) + \ldots + A_2(s_2 - s_1) + A_1 s_1 \\ &= A_n s_n - (A_n - A_{n-1}) s_{n-1} - \ldots - (A_2 - A_1) s_1 \\ &= A_n s_n - \sum_{k=1}^{n} (A_k - A_{k-1}) s_{k-1} \\ &= \sum_{k=1}^{n} k^2 |\alpha_k|^2 \left(1 - \frac{1}{n+1}\right) - \sum_{k=1}^{n} k^2 |\alpha_k|^2 \left(1 - \frac{1}{k}\right) \\ &= \frac{1}{k} \sum_{k=1}^{n} k^2 |\alpha_k|^2 - \frac{1}{n+1} \sum_{k=1}^{n} k^2 |\alpha_k|^2. \end{aligned} \tag{70}$$

After substituting (70) into (68) we see that the claim (60) follows from (68) provided that

$$\begin{aligned} &\sum_{k=1}^{n} k |\alpha_k|^2 - \frac{1}{n+1} \sum_{k=1}^{n} k^2 |\alpha_k|^2 - \sum_{k=1}^{n} \frac{1}{k+1} \\ &= \frac{1}{n+1} \sum_{k=1}^{n} (n+1-k) \left(k |\alpha_k|^2 - \frac{1}{k} \right). \end{aligned} \tag{71}$$

In order to establish (71) note that the central term on the left cancels easily, so we need only to verify that

$$\sum_{k=1}^{n} k|\alpha_k|^2 - \sum_{k=1}^{n} \frac{1}{k+1} = \frac{1}{n+1}\sum_{k=1}^{n}(n+1)\left(k|\alpha_k|^2 - \frac{1}{k}\right) + \frac{1}{n+1}\sum_{k=1}^{n} 1. \tag{72}$$

Now the first term on the left cancels and we are left with an obvious identity

$$-\sum_{k=2}^{n+1}\frac{1}{k} = \left(-\sum_{k=1}^{n}\frac{1}{k}\right) + \left(1 - \frac{1}{n+1}\right). \tag{73}$$

Hence (L-M) inequality (60) has been proved. □

2.5. Logarithmic coefficients and Milin conjecture

I. M. Milin (1919–1992) was a mathematician from Leningrad (now again St. Petersburg). In 1971 he formulated a new conjecture, which implied the Robertson conjecture and (as a consequence) the Bieberbach conjecture). The concern was with logarithmic coefficients of $f \in S$. See [Mln 71], remarks before Theorem 3.2.

Recall that for $f \in S$ the quotient $f(z)/z$ $z \in \Delta$ is holomorphic and does not admit value 0. Consider the branch $\ln(f(z)/z)$ which vanishes at $z = 0$ and its Taylor development

$$\ln(f(z)/z) = c_1 z + c_2 z^2 + \dots . \tag{74}$$

Numbers c_k, $k = 1, 2, \dots$ are called *logarithmic coefficients* of f. I. M. Milin expressed expectation that every $f(z) = z + a_2 z^2 + \ldots \in S$ satisfies

$$I_n[f] \le 0, \qquad n = 2, 3, \dots \tag{75}$$

where

$$I_n[f] := \sum_{k=1}^{n-1}(n-k)\left(k\left|c_k\right|^2 - \frac{4}{k}\right). \tag{76}$$

This is *Milin conjecture* stated in terms of *Milin inequalities* (75). Note that *Milin functional* $I_n[f]$ in (76) consists of $n-1$ terms. We are going to present

(9) Theorem. *Milin conjecture implies Bieberbach conjecture.*

Proof. For $f \in S$ consider the odd function

$$\sqrt{f(z^2)} = b_1 z + b_3 z^3 + b_5 z^5 + \dots , \qquad b_1 = 1. \tag{77}$$

It follows that

$$\sqrt{\frac{f(z^2)}{z^2}} = \frac{1}{z}\sqrt{f(z^2)} = b_1 + b_3 z^2 + b_5 z^4 + \dots . \tag{78}$$

Recall that the main branch of square root can be expressed by the main branch of logarithm. Therefore

$$\sqrt{\frac{f(z^2)}{z^2}} = \exp\left[\frac{1}{2}\ln\frac{f(z^2)}{z^2}\right] = \exp\left(\frac{c_1}{2}z^2 + \frac{c_2}{2}z^4 + \dots\right). \tag{79}$$

With the abbreviation $w := z^2$ we infer from (78), (73) that

$$b_1 + b_3 w + b_5 w^2 + \ldots = \exp\left(\frac{c_1}{2}w + \frac{c_2}{2}w^2 + \dots\right). \tag{80}$$

In the next step we apply the second (L-M) inequality (cf. Sect. 2.4):

$$\sum_{k=0}^{n}|\beta_k|^2 \le (n+1)\exp\left\{\frac{1}{n+1}\sum_{k=1}^{n}(n+1-k)\left(k|\alpha_k|^2 - \frac{1}{k}\right)\right\} \tag{81}$$

to powers w^m, $m \le n-1$ in (80). It follows that

$$\begin{aligned}|b_1|^2 + |b_3|^2 + \ldots + |b_{2n-1}|^2 &\le n\exp\left\{\frac{1}{n}\sum_{k=1}^{n-1}(n-k)\left(k\left|\frac{c_k}{2}\right|^2 - \frac{1}{k}\right)\right\}\\ &= n\exp\left(\frac{I_n[f]}{4n}\right).\end{aligned} \tag{82}$$

The assumed Milin's conjecture implies $I_n[f] \le 0$ for $n = 2, 3, \ldots$, and hence the right-hand side is not greater than n. In view of (82) Robertson inequalities (57) hold for every $f \in S$. It suffices now to recall that Robertson conjecture implies Bieberbach conjecture.

2.6. Milin's functional vanishes on the Koebe function

We have seen that Milin conjecture implies Bieberbach conjecture. Indeed, in order to prove inequalities $|a_n| \le n$ for a given $f \in S$ it suffices to verify the inequalities $I_n[f] \le 0$. Already Bieberbach knew that the Koebe function

$$\mathbf{K}(z) := \frac{z}{(1-z)^2} = z + 2z^2 + 3z^3 + \ldots \qquad z \in \Delta \tag{83}$$

is (up to rotation) the only function in S with $|a_2| = 2$. The Koebe function was intensively investigated in the context of Bieberbach conjecture. The following considerations will help to motivate the de Branges construction, although, from the formal point of view, the latter is rather independent. We have seen in Section 1 that

$$\mathbf{K}(z) = \frac{1}{4}\left\{\left(\frac{1+z}{1-z}\right)^2 - 1\right\}, \qquad z \in \Delta, \tag{84}$$

maps Δ onto the slited plane $\mathbb{C} \setminus (-\infty, -1/4)$, see Fig. 9.

We shall now calculate logarithmic coefficients of $\mathbf{K}$. The development

$$\ln\frac{\mathbf{K}(z)}{z} = -2\ln(1-z) = 2\left(z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \dots\right) \tag{85}$$

yields $c_k = \frac{2}{k}$, $k = 1, 2, \ldots$. Hence $k(2/k)^2 - 4/k = 0$. As a consequence we have

$$I_n[\mathbf{K}] := \sum_{m=1}^{n-1} (n-m) \left(m|c_m^2 - \frac{4}{m} \right) = 0. \tag{86}$$

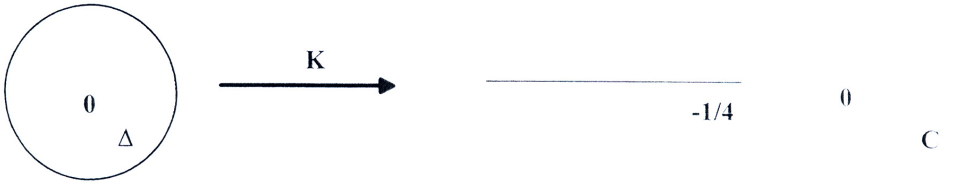

Fig. 9: The image of $\mathbf{K}(z), z \in \Delta$.

(10) *Remark* (important). For (regular) $f \in S$ consider the Löwner chain $f(z,t)$, $t \in (0, +\infty)$, and normalized Riemann mappings

$$g_t(z) := \frac{f(z,t)}{e^t} \in S, \qquad t \in (0, +\infty), \tag{87}$$

(see the beginning of this chapter). For sufficiently large t the function $g_t(z)$, $z \in \Delta$, maps the unit disc onto the complement of a subray of the negative halfaxis. The original mapping $f_t(z) = f(z,t)$, up to a multiplicative constant, equals $\mathbf{K}$ and normalization brings it again to $g_t(z) = \mathbf{K}(z)$. We see that $I_n(t) := I_n[g_t] = I_n[\mathbf{K}] = 0$ for $n = 1, 2, \ldots$. Moreover $I_n(0) := I_n[g_0] = I_n[f]$. From an additional *hypothetical* assumption, that $I_n(t)$, $t \in (0, +\infty)$, is *nondecreasing* it follows that $I_n(0) \leq 0$. But the latter is Milin's inequality $I_n[f] \leq 0$. Hence, to prove Milin inequalities for (regular) $f \in S$ it suffices to show that

$$I_n'(t) \geq 0, \qquad t \in (0, +\infty). \tag{88}$$

(11) *Remark* (important). A similar idea motivates de Branges proof of Milin inequalities. But de Branges reasoning is "more flexible". He is not considering $I_n(t) = I_n[g_t]$ but instead introduces and investigates another functional $\Omega_n(t) := \Omega_n[g_t]$. For details see the next chapter.

Acknowledgment

My thanks go to Dr. Eugeniusz Szpakowski, heart surgeon. Without his decisive help this memoir would never been finihed. I am also grateful to Profs. Z. Jakubowski and J. Ławrynowicz for their part in detecting and correcting some errors. The author hopes that the importance of the subject and some originality in the arrangement of topics will prevail over the remaining insufficiency.

References

[A,A,R]–[Wts] See this issue, pp. 98–101.

Department of Mathematical Analysis
Cardinal Stefan Wyszyński University
Dewajtis 5, PL-01-815 Warszawa, Poland
e-mail: skwarczynski@uksw.edu.pl

Presented by Zbigniew Jakubowski at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on July 16, 2010

TWIERDZENIE DE BRANGES'A A UOGÓLNIONE FUNKCJE HIPERGEOMETRYCZNE I

HIPOTEZA BIEBERBACHA I FUNKCJONAŁ MILINA

S t r e s z c z e n i e

Przedstawiamy najbardziej zaskakujące matematyczne odkrycie ubiegłego stulecia – dowód twierdzenia de Branges'a nawiązujący do własności uogólnionych funkcji hipergeometrycznych. Twierdzenie wypowiedziane jako hipoteza w roku 1916 przez Ludwiga Bieberbacha zostało udowodnione w roku 1984 przez Louisa de Branges'a po 68 latach intensywnych badań najwybitniejszych matematyków.

W paragrafie 2 analizujemy metodą parametryczną Löwnera, która w roku 1923 doprowadziła go do dowodu nierówności $|a_3| \leq 3$ i pełni podstawową rolę w potwierdzeniu słuszności hipotezy Bieberbacha. Do wyprowadzenia analitycznych wniosków z metody użyta jest zbieżność w sensie Carathéodory'ego.

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 69–88

Dedicated to Professor Roman Stanisław Ingarden
on the occasion of his ninetieth birthday

Maciej Skwarczyński

DE BRANGES THEOREM AND GENERALIZED HYPERGEOMETRIC FUNCTIONS II

DE BRANGES FUNCTIONAL AND HYPERGEOMETRIC EQUATION

Summary

Original article [Brn 84] contains the key explanation: *... the problem is to propagate information by means of a differential equation. For this purpose information has to be coded in a convenient form and then carried over from one end of an interval to the another.*

In other words: Löwner differential equation affects propagation of logarithmic coefficients. Soon enough the general insight of de Branges gained wider acceptance. Carl FitzGerald and Christian Pommerenke in [F,P 85] offered their own variant. Still another report was presented in [Krv 86] (see especially pp. 511–513. Korevaar's article was awarsed Chauvenet prize for mathematical exposition). The present chapter attempts to indicate the general plan of de Branges proof.

The time has come to discuss the role of Gauss hypergeometric function ${}_2F_1$ and its generalizations. Simplest of such generalizations, the Clausen function ${}_3F_2$ plays an eminent role in the final part of de Branges proof. We present two proofs of Clausen identity: one very short and one much longer. With such preparation we will derive in the next chapter the inequality $\tau' \leq 0$, thereby clearing the condition (2).

0. Initial remark

Formulae numbers (1) etc. and statement numbers (1) etc. referring to Part I of the memoir are quoted as (I.1) etc. and (I.1) etc., respectively. Acronyms below [Brn 84] etc. usually consist of first three consonants of author's name, followed by the year of publication. List of references is constructed alphabetically according to letters of the acronym. When no ambiguity results acronyms of this list appear without the year of publication.

1. De Branges functional. De Branges differential system

1.1. Derivative of logarithmic coefficient in a Löwner chain

Consider $f \in S$, its Löwner chain $f(z,t)$, $z \in \Delta$, $t \in (0,+\infty)$, and the logarithmic coefficients $c_k(t)$ of normalized function $g_t(z) := f(z,t)/e' \in S$. By definition

$$\ln \frac{f(z,t)}{e^t z} = \sum_k c_k(t) z^k. \tag{1}$$

Recall that in the Löwner equation

$$\frac{\partial f(z,t)}{\partial t} = zp(z,t) \frac{\partial f(z,t)}{\partial z} \tag{2}$$

the factor $p(z,t)$ has positive real part. From (I.59) follows Taylor development

$$p(z,t) = \frac{1+\kappa(t)z}{1-\kappa(t)z} = [1+\kappa(t)z] \sum_{k=0}^{\infty} [\kappa(t)z]^k = 1 + 2\sum_{k=1}^{\infty} \kappa(t)^k z^k. \tag{3}$$

Let us differentiate, with respect to t, both sides in (1). The logarithmic derivative of $e^t z$ is 1 hence

$$\frac{1}{f(z,t)} \frac{\partial}{\partial t} f(z,t) = 1 + \sum_{k=1}^{\infty} c_k'(t) z^k. \tag{4}$$

After replacing $(\partial/\partial t)\ f(z,t)$ by the right-hand side of (2) we find that

$$p(z,t) z \frac{1}{f(z,t)} \frac{\partial}{\partial t} f(z,t) = 1 + \sum_{k=1}^{\infty} c_k'(t) z^k. \tag{5}$$

Now differentiate again both sides in (1), this time with respect to z. The logarithmic derivative of $e^t z$ is now $1/z$, hence

$$\frac{1}{f(z,t)} \frac{\partial}{\partial z} f(z,t) = \frac{1}{z} + \sum_{k=1}^{\infty} k c_k(t) z^{k-1}, \tag{6}$$

$$\frac{z}{f(z,t)} \frac{\partial}{\partial z} f(z,t) = 1 + \sum_{k=1}^{\infty} k c_k(t) z^k. \tag{7}$$

Substituting (3) and (7) for (5) we get power series equality

$$\left(1 + 2\sum_{k=1}^{\infty} \kappa(t)^k z^k\right)\left(1 + \sum_{k=1}^{\infty} k c_k(t) z^k\right) = 1 + \sum_{k=1}^{\infty} c_k'(t) z^k. \tag{8}$$

Finally, comparing coefficients on both sides of (8) yields

$$c_k'(t) = k c_k(t) + 2\kappa(t)^k + 2\sum_{j=1}^{k-1} j c_j(t) \kappa(t)^{k-j}. \tag{9}$$

With abbreviations

$$\sigma_0(t) := 0, \quad \sigma_k(t) := \sum_{j=1}^{k} j c_j(t)\kappa(t)^{-j}, \quad k = 1, 2, \ldots, \tag{10}$$

we rewrite (9) as

(11)

$$c_k'(t) = 2\kappa(t)^k + \sum_{j=1}^{k-1} j c_j(t)\kappa(t)^{k-j} + \sum_{j=1}^{k} j c_j(t)\kappa(t)^{k-j} = \kappa(t)^k \left[2 + \sigma_{k-1}(t) + \sigma_k(t)\right].$$

1.2. De Branges differential system

For n fixed de Branges defined $\Omega_n(t)$ by replacing factors $n-k$, $k = 1, \ldots, n-1$, in the Milin functional (I.76) with conveniently chosen weights $\tau_k(t)$, see def. **(1)** below. Symbolically

$$\Omega_n(t) := \sum_{k=1}^{n-1} \left\{ k|c_k(t)|^2 - \frac{4}{k} \right\} \tau_k(t), \qquad t \in [0, +\infty). \tag{12}$$

(The dependence of τ_k on $n = 2, 3, \ldots$ is suppressed in order not to overload the notation).

(1) *Definition.* De Branges weights $\tau_k(t)$, $k = 1, \ldots, n-1$ are defined as follows:

$$\tau_k(t) := k \sum_{\nu=0}^{n-k-1} (-1)^\nu \frac{(2k+\nu+1)_\nu \cdot (2k+2\nu+2)_{n-k-1-\nu}}{(k+\nu)\cdot \nu! \cdot (n-k-\nu-1)!} e^{-\nu t - kt}, \quad k = 1, \ldots, n-1, \tag{13}$$

with the usual meaning of the *Pochhammer symbol*:

$$(\gamma)_0 := 1, \quad (\gamma)_\nu := \gamma(\gamma+1)(\gamma+2)\ldots(\gamma+\nu-1), \quad \nu \in \mathbb{N}. \tag{14}$$

(2) *Remark* (de Branges differential system). W shall see later that τ_k, $k = 1, \ldots, n-1$, are characterized as the unique solution in $(0, +\infty)$ to the system of ordinary differential equations (with $\tau_n :\equiv 0$):

$$\tau_k - \tau_{k+1} = -\left(\frac{\tau_k'}{k} + \frac{\tau_{k+1}'}{k+1} \right), \qquad k = 1, \ldots, n-1, \tag{15}$$

subject to initial conditions

$$\tau_k(0) = n-k, \qquad k = 1, \ldots, n-1. \tag{16}$$

Note that conditions (16) imply $\Omega_n(0) = I_n(0)$. This system can be solved successively. Unknown τ_{n-1} is determined first, τ_{n-2} next, and so forth ending with τ_1. At each step one meets a linear equation (of the first order) with one unknown and constant coefficients.

1.3. Unexpected change in notation

The area related to the de Branges theorem is so vast that we have to change notation in the midst of a reasoning. *The change is small and easy to control.* But anyway, a clear explanation is in order.

Inequalities which appear in Bieberbach problem are like rooms in the Hilbert hotel. They can be numbered either by $n = 2, 3, 4, \dots$ or by $n = 1, 2, 3, \dots$. The first way was appropriate at the early stage of investigations, when attention was centered on individual results (Bieberbach for a_2, Löwner for a_3, Charzyński and Schiffer for a_4). Situation is different in the case of de Branges proof, where all coefficient inequalities are treated at once and generalized hypergeometric functions enter the picture. It is more convenient to assume that n runs through the numbers $1, 2, \dots$. As a consequence in each individual formula index n pertains to $|a_{n+1}| \leq n+1$, not to $|a_n| \leq n$.

For example, replacing n by $n+1$ in the formula (13) we now write

$$\tau_k(t) := k \sum_{\nu=0}^{n-k} (-1)^\nu \frac{(2k+\nu+1)_\nu \cdot (2k+2\nu+2)_{n-k-\nu}}{(k+\nu) \cdot \nu! \cdot (n-k-\nu)!} e^{-\nu t - kt}, \quad k = 1, 2, \dots n. \tag{17}$$

To avoid collision we shall reserve the name de Branges weights for (13) and call (17) *de Branges functions*. The latter satisfy the system of equations (with $\tau_{n+1} :\equiv 0$)

$$\tau_k - \tau_{k+1} = -\left(\frac{\tau_k'}{k} + \frac{\tau_{k+1}'}{k+1} \right), \qquad k = 1, \dots, n, \tag{18}$$

subject to initial conditions

$$\tau_k(0) = n - k + 1, \qquad k = 1, \dots, n. \tag{19}$$

Moreover, the relevant expression for de Branges functional becomes

$$\Omega_n(t) := \sum_{k=1}^{n} \left\{ k|c_k(t)|^2 - \frac{4}{k} \right\} \tau_k(t). \tag{20}$$

Note that from (17) by direct differentiation follows

$$-\tau_k'(t) = k \sum_{\nu=0}^{n-k} (-1)^\nu \frac{(2k+\nu+1)_\nu \cdot (2k+2\nu+2)_{n-k-\nu}}{\nu! \cdot (n-k-\nu)!} e^{-\nu t - kt}, \quad n = 1, 2, \dots. \tag{21}$$

1.4. What lies ahead

It is the proper moment to present a precise plan for remaining reasoning. The general idea is to establish for $\Omega_n(t)$ the properties desired of $I_n(t)$ and then to use $\Omega_n(t)$, $n = 1, 2, \dots$, instead of $I_n(t)$, $n = 2, 3, \dots$. To be specific, we want to establish inequality $\Omega_n'(t) \geq 0$, $t \in (0, +\infty)$ and the limits $\Omega_n(+\infty) = 0$, $\Omega_n(0) = I_n(0)$. To achieve this we need to investigate the functions τ_k and their derivatives τ_k', defined by (17) and (21), respectively. We shall prove *key conditions*

(**1**) τ_k satisfy equations (18) (elementary),
(**2**) τ_k satisfy inequality $\tau_k' \le 0$ (nontrivial),
(**3**) τ_k satisfy condition $\tau_k(+\infty) = 0$ (elementary),
(**4**) τ_k satisfy initial conditions (19) (nontrivial).

Meanwhile the desired properties of $\Omega_n(t)$ will appear as corollaries. Note that (**1**), (**4**) justify characterization τ_k in terms of de Branges differential system. The present chapter establishes contitions (**1**), (**3**) (see Lemmas **(3)** and **(5)** below) and reduces everything else to (**2**), (**4**). Proofs of (2), (4) rely on g.h.f. (generalized hypergeometric functions) and will be given later. The Clausen identity (see chapter 5) will be used in chapter 6 to prove condition (2). In final chapter 7 we shall prove condition (4) using Watson's summation lemma.

1.5. De Branges functional vanishes at infinity

(3) Lemma to establish condition (**3**). *For fixed $n \in \mathbb{N}$ functions τ_k, $k = 1, \dots, n$, satisfy $\tau_k(+\infty) = 0$.*

Proof. The (finite) sum in (17) is a linear combination of exponentials

$$e^{-(\nu+k)tt}, \qquad \nu = 0, 1, \dots, n-k, \tag{22}$$

and hence converges to 0 when $t \to +\infty$. □

(4) Corollary. *De Branges functional $\Omega_n(t)$ satisfies $\Omega_n(+\infty) = 0$.*

Proof. Folows from (20) since $\tau_k \to 0$ and $c_k \to 2/k$. □

1.6. Derivative of the de Branges functional

(5) Lemma to establish (**1**). *For fixed $n \in \mathbb{N}$ de Branges functions* (17) *satisfy equations* (18)*. of de Branges differential system.*

Proof. With $\tau_{n+1} \equiv 0$ equations (18) can be written as

$$\tau_k + \frac{\tau_k'}{k} = \tau_{k+1} + \frac{\tau_{k+1}'}{k+1}, \quad k = 1, \dots, n. \tag{23}$$

We shall directly verify (23). From

$$\tau_k(t) := k \sum_{\nu=0}^{n-k} (-1)^\nu \frac{(2k+\nu+1)_\nu \cdot (2k+2\nu+2)_{n-k-\nu}}{(k+\nu) \cdot \nu! \cdot (n-k-\nu)!} e^{-\nu t - kt} \tag{24}$$

and

$$-\tau_k'(t) = k \sum_{\nu=0}^{n-k} (-1)^\nu \frac{(2k+\nu+1)_\nu \cdot (2k+2\nu+2)_{n-k-\nu}}{\nu! \cdot (n-k-\nu)!} e^{-\nu t - kt} \tag{25}$$

we have

$$(26)\qquad \left(\frac{\tau_k}{k}\right)' = \left(\sum_{\nu=0}^{n-k}(-1)^\nu \frac{(2k+\nu+1)_\nu \cdot (2k+2\nu+2)_{n=k-\nu}}{(k+\nu)\cdot \nu! \cdot (n-k-\nu)!} e^{-\nu t-kt}\right)' = -\tau_k$$

and, after replacing k with $k+1$, we can see that both sides of (23) vanish.

We now come to the very important result obtained by C. FitzGerald and Ch. Pommerenke in [F,P 85]. The paper was instrumental for general recognition of de Branges theorem. It offers explicit formulae which relate derivatives of logarithmic coefficients to the derivative of de Branges functional. Recall that formula (11) for derivatives of logarithmic coefficients involves functions σ_k related to the Löwner equation. In fact this is the only place where the de Branges construction makes essential use of the Löwner equation.

(6) FitzGerald-Pommerenke Lemma. *The derivative of de Branges functional*

$$(27)\qquad \Omega_n(t) := \sum_{k=1}^{n}\left\{k|c_k(t)|^2 - \frac{4}{k}\right\}\tau_k(t), \qquad n=1,2,\dots,$$

satisfies

$$(28)\qquad \Omega_n'(t) = -\sum_{k=1}^{n}|\sigma_{k-1}(t)+\sigma_k(t)+2|^2\frac{\tau_k'(t)}{k}, \qquad n=1,2,\dots.$$

Proof. We follow [F,P], p. 686. For brevity we shall suppress t, n. First of all, from Definition (10) of σ_k it follows that

$$(29)\qquad (\sigma_k-\sigma_{k-1})\kappa^k\left(\sum_{j=1}^{k} jc_j\kappa^{-1} - \sum_{j=1}^{k-1} jc_j(t)\kappa^{-j}\right)\kappa^k = (kc_k\kappa^{-k})\kappa^k = kc_k.$$

We now differentiate (27) substituting (29) for kc_k and (11) for c_k'. This yields

$$\begin{aligned}(30)\qquad \Omega' &= \sum_{k=1}^{n}\left(kc_k\bar c_k - \frac{4}{k}\right)'\tau_k + \sum_{k=1}^{n}\left(kc_k\bar c_k - \frac{4}{k}\right)\tau_k' \\ &= \sum_{k=1}^{n}\left(kc_k'\bar c_k + kc_k\bar c_k'\right)\tau_k + \sum_{k=1}^{n}\left(kc_k k\bar c_k - 4\right)\frac{\tau_k'}{k} \\ &= \sum_{k=1}^{n}\left[2\mathrm{Re}\,(kc_k\bar c_k')\right]\tau_k + \sum_{k=1}^{n}(kc_k k\bar c_k - 4)\frac{\tau_k'}{k} \\ &= 2\sum_{k=1}^{n}\left[\mathrm{Re}\,((\sigma_k+\sigma_{k-1}+2)(\bar\sigma_k-\bar\sigma_{k-1}))\right]\tau_k \\ &\quad + \sum_{k=1}^{n}\left((\sigma_k-\sigma_{k-1})(\bar\sigma_k-\bar\sigma_{k-1}) - 4\right)\frac{\tau_k'}{k}.\end{aligned}$$

In the latter line $\kappa\bar\kappa = |\kappa|^2 = 1$ was used.

In order to obtain (28) the first sum on the right side of (30) is transformed by parts. Since

$$
\begin{aligned}
(31)\qquad & \mathrm{Re}\left[(\sigma_k+\sigma_{k-1}+2)(\bar\sigma_k-\bar\sigma_{k-1})\right]\\
&=\mathrm{Re}\left[|\sigma_k|^2-\sigma_k\bar\sigma_{k-1}+\sigma_{k-1}\bar\sigma_k-|\sigma_{k-1}|^2+2\bar\sigma_k-2\bar\sigma_{k-1}\right]\\
&=\left(|\sigma_k|^2+2\mathrm{Re}\,\sigma_k\right)-\left(|\sigma_{k-1}|^2+2\mathrm{Re}\,\sigma_{k-1}\right),
\end{aligned}
$$

we have (using $\sigma_0=0$, $\tau_{n+1}=0$):

$$
\begin{aligned}
(32)\qquad & 2\sum_{k=1}^{n}\left[\mathrm{Re}(\sigma_k+\sigma_{k-1}+2)(\bar\sigma_k-\bar\sigma_{k-1})\right]\tau_k\\
&=2\sum_{k=0}^{n-1}\left(|\sigma_k|^2+2\mathrm{Re}\,\sigma_k\right)-2\sum_{k=0}^{n-1}\left(|\sigma_{k-1}|^2+2\mathrm{Re}\,\sigma_{k-1}\right)\tau_k\\
&=2\sum_{k=1}^{n}\left(|\sigma_k|^2+2\mathrm{Re}\,\sigma_k\right)\tau_k-2\sum_{k=1}^{n}\left(|\sigma_k|^2+2\mathrm{Re}\,\sigma_k\right)\tau_{k+1}\\
&=2\sum_{k=1}^{n}\left(|\sigma_k|^2+2\Re\,\sigma_k\right)(\tau_k-\tau_{k+1})\\
&=2\sum_{k=1}^{n}\left(|\sigma_k|^2+2\mathrm{Re}\,\sigma_k\right)\frac{\tau_k'}{k}-2\sum_{k=1}^{n}\left(|\sigma_k|^2+2\mathrm{Re}\,\sigma_k\right)\frac{\tau_{k+1}'}{k+1}\\
&=-2\sum_{k=1}^{n}\left(|\sigma_k|^2+2\mathrm{Re}\,\sigma_k+|\sigma_{k-1}|^2+2\mathrm{Re}\,\sigma_{k-1}\right)\frac{\tau_k'}{k}.
\end{aligned}
$$

The next to last line was obtained with Lemma **(5)**.

Returning to (30) we find

$$
\begin{aligned}
(33)\qquad \Omega' &= -2\sum_{k=1}^{n}\left(|\sigma_k|^2+2\mathrm{Re}\sigma_k+|\sigma_{k-1}|^2+2\mathrm{Re}\,\sigma_{k-1}\right)\frac{\tau_k'}{k}\\
&\quad+\sum_{k=1}^{n}\left[(\sigma_k-\sigma_{k-1})(\bar\sigma_k-\bar\sigma_{k-1})-4\right]\frac{\tau_k'}{k}\\
&=\sum_{k=1}^{n}\left[-2\sigma_k\bar\sigma_k-2(\sigma_k+\bar\sigma_k)-2\sigma_{k-1}\bar\sigma_{k-1}-2(\sigma_{k-1}+\bar\sigma_{k-1})-4\right]\frac{\tau_k'}{k}.
\end{aligned}
$$

On the other hand, we rewrite the right-hand side of (28) as follows:

$$
\begin{aligned}
(34)\qquad & -\sum_{k=1}^{n}|\sigma_{k-1}+\sigma_k+2|^2\frac{\tau_k'}{k}\\
&=\sum_{k=1}^{n}\left(-\sigma_{k-1}\bar\sigma_{k-1}-\sigma_{k-1}\bar\sigma_k-2\sigma_{k-1}-\sigma_k\bar\sigma_{k-1}-\sigma_{k-1}\bar\sigma_k-2\sigma_k-2\sigma_{k-1}-4\right)\frac{\tau_k'}{k}.
\end{aligned}
$$

Right-hand sides of (34), (23) are equal, and hence left-hand sides are equal as well, thus proving (28).

(7) *Remark.* In the above proof Lemma **(5)** has intervined, but explicit form (17) was not used in other way. (For deeper properties (**2**) and (**4**) of τ_k we shall use (17) directly.) The following corollary shows that the condition (**2**) implies the inequality $\Omega_n' \geq 0$ which (by Corollary **(4)**) implies

$$\Omega_n(0) \leq 0. \tag{35}$$

(8) Corollary. *Assume for every $n \in \mathbb{N}$ that $\tau_k' \leq 0$. Then $\Omega_n' \geq 0$.*

Proof. Immediate by inspecting the identity (28) in FitzGerald-Pommerenke lemma. □

1.7. De Branges functional and Milin inequalities

(9) Lemma. *Assume for every $n \in \mathbb{N}$ that $\tau_k(0) = n - k + 1$. Then*

$$\Omega_n(0) = \sum_{k=1}^{n} \left\{ k|c_k(0)|^2 - \frac{4}{k} \right\} \tau_k(0) = I_n(0). \tag{36}$$

It follows that $\Omega_n \leq 0$ implies Milin inequalities.

Proof. Immediate by the definition (4.20) of de Brange functional $\Omega_n(t)$.

(10) *Remark.* From Lemma **(9)** follows that the conditions (**2**), (**4**) imply the Milin conjecture, hence the Robertson conjecture, and hence the Bieberbach conjecture. Establishing (**2**), (**4**) (see the following chapters) concludes de Branges' proof of the Bieberbach conjecture.

1.8. Korevaar's examples

We close this chapter by quoting very attractive examples from [Krv 86], p. 508.

(11) *Example* ($n = 1$). Note that $\tau_2 = 0$. De Branges system (18), (19) consists of one equation

$$\tau_1 = -\tau_1', \qquad \tau_1(0) = 2 - 1 = 1. \tag{37}$$

Its solution is $\tau_1(t) = e^{-t}$. From (17) with $n = 1$, $k = 1$, follows the same result, namely

$$\tau_1(t) = 1 \cdot 1 \cdot \frac{(3)_0 \cdot (4)_0}{1 \cdot 1 \cdot 1} e^{-t \cdot 1} = e^{-t}. \tag{38}$$

This function satisfies key conditions (**2**), (**4**) on p. 68 in Part II. This yields the Milin inequality and, as a consequence, $|a_2| \leq 2$.

(12) *Example* ($n = 2$). Note that $\tau_3 = 0$. De Branges system (18) consists of two equations

$$\tau_1 - \tau_2 = -\left(\frac{\tau_1'}{1} + \frac{\tau_2'}{2}\right), \qquad (39)$$
$$\tau_2 = -\frac{\tau_2'}{2}$$

with initial conditions $\tau_1(0) = 3 - 1 = 2$ and $\tau_2(0) = 3 - 2 = 1$. The second equation yields $\tau_2(t) = e^{-2t}$. After substituting this into the first equation one finds

$$\tau_1 - e^{-2t} = -\tau_1' \frac{1}{2}(e^{-2t})' = -\tau_1' + e^{-2t}. \qquad (40)$$

This yields $\tau_1(t) = 4e^{-t} - 2e^{-2t}$. From (17) with $n = 2$, $k = 1$ follows the same result, namely

$$\tau_1 = 1 \cdot \frac{(3)_0 \cdot (4)_1}{1 \cdot 0! \cdot 1!} e^{-t \cdot 1} - 1 \cdot \frac{(4)_1 \cdot (6)_0}{2 \cdot 1! \cdot 0!} e^{-t \cdot 2} = 4e^{-t} - 2e^{-2t}. \qquad (41)$$

Key conditions (**2**), (**4**) are obviously satisfied. Hence $|a_3| \le 3$ (first proved by K. Löwner in 1923; see [Lwn 23]).

Remark. Despite of its attractiveness, example **(11)** cannot be treated as a mathematical *proof* of Bieberbach inequality $|a_2| \le 2$. It relies on de Branges theory which was derived with Carathéodory's convergence theorem. Yet, Carathéodory's convergence theorem rests itself on the inequality $|a_2| \le 2$. This remark does not extend to examples **(12), (13)** which *prove* $|a_3| \le 3$ with overwhelming ease. Chapeaux bas!

2. A proof of Clausen identity using differential equations

2.1. Generalized hypergeometric functions

Inequality $\tau_k' \le 0$ was originally confirmed by D. Askey, who deduced it from the work [A,G 76] on Jacobi polynamials, written jointly by Askey and Gasper. Following N. Kazarinoff we modify this part of reasoning by working directly in terms of *generalized hypergeometric functions* (abbreviation: g.h.f.):

$$w(z) := {}_AF_B(a;b;z) = {}_AF_B\left(\begin{matrix} a_1, a_2, \dots, a_A \\ b_1, b_2, \dots, b_B \end{matrix} \,\middle|\, z\right) := \sum_{\nu=0}^{+\infty} \frac{(a_1)_\nu \cdots (a_A)_\nu}{(b_1)_\nu \cdots (b_B)_\nu} \frac{z^\nu}{\nu!}. \qquad (42)$$

Notation in (1) comes from E. W. Barnes. For typographic reason the expression in (1) is often written as ${}_AF_B(a_1, \dots, a_A; b_1, \dots, b_B; z)$. Numbers $a_1, \dots, a_A$ are called *upper parameters*; numbers $b_1, \dots, b_B$ are called *lower parameters*. (The order of upper parameters is inessential and so is the order of lower parameters.) It is assumed that none of the lower parameters is zero or negative integer. When an upper parameter is zero or a negative integer, the series in (1) terminates (has only finitely many nonzero terms) and hypergeometric function reduces to a hypergeometric polynomial.

It is known that (1) satisfies ordinary an differential equation of order $\max(A, B)$. This equation is linear, homogeneous and singular. For $A = B + 1$ the series in (1) converges for $|z| < 1$. For a systematic account see [Slt].

Functions ${}_AF_B$ generalize Gauss' hypergeometric function ${}_2F_2(a_1, a_2; b_1; z)$. In order to simplify indices, Gauss function is often written as

$$ {}_2F_1(a; b; c; z) = \sum_{\nu=0}^{+\infty} \frac{(a)_\nu (b)_\nu}{(c)_\nu} \frac{z^\nu}{\nu!}. \tag{43}$$

It satisfies a differential equation of order 2; namely ([Slt], formula (1.2.1) on p. 5):

$$ L[w] := z(1-z)w'' + [c - (a+b+1)z]w' - abw = 0. \tag{44}$$

Most special functions of mathematical physics are particular cases of (2). W. Sawyer wrote in 1955: *There must be many universities today where* 95%, *if not* 100% *of the functions studied by physics, engeneering, and even mathematics students, are covered by symbol* ${}_2F_1$.

The above does not mean that other functions ${}_pF_q$ should be ignored. The most simple generalization of Gauss function

$$ w(z) = {}_3F_2(a', b', c'; d', e'; z) = \sum_{\nu=0}^{+\infty} \frac{(a')_\nu (b')_\nu (c')_\nu}{(d')_\nu (e')_\nu} \frac{z^\nu}{\nu!} \tag{45}$$

was investigated by T. Clausen in [Cls] (1828) (his portrait is presented in Wikipedia: http://en.wikipedia.org/wiki/Thomas_Clausen_(mathematician)). Clausen function satisfies a differential equation of the third order:

$$ \begin{aligned} M[w] := {} & z^2(1-z)w''' + [(1+d'+e)z - (3+a'+b'+c')z^2]w'' \\ & + [d'e' - (1+a'+b'+c'+a'b'+a'c'+b'c')z]w' - a'b'c'w = 0. \end{aligned} \tag{46}$$

This formula appears in [Gng 99], p. 113, and is slightly misprinted in [Kzr 88]. Letter c is prone to errors isnce it often appears as the lower parameter in ${}_2F_1$ and as the upper aparameter in ${}_3F_2$. We are interested in the case

$$ a' = 2\alpha, \quad b' = 2\beta, \quad c' = \alpha + \beta, \quad d' = 2(\alpha+\beta), \quad e' = \alpha + \beta + \frac{1}{2} $$

(see below). Then (46) takes the form

$$ \begin{aligned} M[w] := {} & z^2(1-z)w''' + \left[\left(3\alpha + 3\beta + \frac{3}{2}\right) z - 3(1+\alpha+\beta)z^2\right] w'' \\ & + \left[2(\alpha+\beta)\left(\alpha+\beta+\frac{1}{2}\right) - \left(1 + 3(\alpha+\beta) + 4\alpha\beta + 2(\alpha+\beta)^2 2a'b'\right) z\right] w' \\ & - 4\alpha\beta(\alpha+\beta)w = 0. \end{aligned} \tag{47}$$

2.2. Generalized hypergeometric equation

Generalized function $w(z) = {}_AF_B(a;b;z)$ (Barnes notation) satisfies the *generalized hypergeometric equation* (abbreviation: g.h.e.):

$$\theta(\theta + b_1 - 1)\dots(\theta + b_B - 1)w = z(\theta + a_1 \dots (\theta + a_A)w \tag{48}$$

where $\theta := z(d/dz)$ is the Aronhold differential operator; see [Rnv 60], p. 75. The Aronhold operator is a cornerstone of the classical theory of algebraic invariants. We recall a proof offered in [Rnv 60]. Note that operators of the form $\theta + c$ commute with each other. From $\theta z^\nu = \nu z^\nu$ it follows that

$$(\theta + b_j - 1)z^\nu = (\nu + b_j - 1)z^\nu \quad \text{and} \quad (\theta + a_i)z^\nu = (\nu + a)z^\nu.$$

Applying the operator on the left-hand side of (48) to every term in the development of w results in

$$\begin{aligned}
\left\{\theta \prod_{j=1}^{q}(\theta + b_j - 1)\right\} w &= \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\prod_{i=1}^{p}(a_i)_\nu}{\prod_{j=1}^{q}(b_j)_\nu} \left\{\theta \prod_{j=1}^{q}(\theta + b_j - 1)\right\} z^\nu \\
&= \sum_{\nu=1}^{\infty} \frac{1}{\nu!} \frac{\prod_{i=1}^{p}(a_i)_\nu}{\prod_{j=1}^{q}(b_j)_\nu} \left(\nu \prod_{j=1}^{q}(\nu + b_j - 1)\right) z^\nu \\
&= \sum_{\nu=1}^{\infty} \frac{1}{(\nu-1)!} \prod_{i=1}^{p}(a_i)_\nu \left(\prod_{j=1}^{q} \frac{(\nu + b_j - 1)}{\prod_{j=1}^{q}(b_j)_\nu}\right) z^\nu \\
&= \sum_{\nu=1}^{\infty} \frac{\prod_{i=1}^{p}(a_i)_\nu}{\prod_{j=1}^{q}(b_j)_\nu} \frac{z^\nu}{(\nu-1)!}.
\end{aligned} \tag{49}$$

The right-hand side of (48) is calculated analogously:

$$\begin{aligned}
z\left\{\prod_{i=1}^{p}(\theta + a_i)\right\} w &= z\sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\prod_{i=1}^{p}(a_i)_\nu}{\prod_{j=1}^{q}(b_j)_\nu} \left\{\prod_{i=1}^{p}(\theta + a_i)\right\} z^\nu \\
&= z\sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\prod_{i=1}^{p}(a_i)_\nu}{\prod_{j=1}^{q}(b_j)_\nu} \left(\prod_{i=1}^{p}(\nu + a_i)\right) z^\nu \\
&= z\sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\prod_{i=1}^{p}(a_i)_\nu(\nu + a_i)}{\prod_{j=1}^{q}(b_j)_\nu} z^\nu \\
&= \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\prod_{i=1}^{p}(a_i)_{\nu+1}}{\prod_{j=1}^{q}(b_j)_\nu} z^{\nu+1} \\
&= \sum_{\nu=1}^{\infty} \frac{\prod_{i=1}^{p}(a_i)_\nu}{\prod_{j=1}^{q}(b_j)_{\nu-1}} \frac{z^\nu}{(\nu-1)!}.
\end{aligned} \tag{50}$$

Expressions obtained in (49) and (50) are equal as claimed. Simple examples of (48) are discussed in next two sections.

2.3. Gauss differential equation

Consider Gauss hypergeometric function ${}_2F_1(a,b;c;z)$. We rewrite (48) as $L[w]=0$ with

$$L := \frac{1}{z}\theta(\theta+c-1)-(\theta+a)(\theta+b). \tag{51}$$

In view of

$$\theta = z\frac{d}{dz}, \qquad \theta^2 = z^2\frac{d^2}{dz^2}+z\frac{d}{dz}, \tag{52}$$

one finds

$$\frac{1}{2}\theta(\theta+c-1) = z\frac{d^2}{dz^2}+\frac{d}{dz}+(c-1)\frac{d}{dz} = z\frac{d^2}{dz^2}+c\frac{d}{dz}, \tag{53}$$

$$(\theta+a)(\theta+b) = \theta^2+(a+b)\theta+ab = z^2\frac{d^2}{dz^2}+z(a+b+1)\frac{d}{dz}+ab. \tag{54}$$

By substracting (54) from (53) we obtain an explicit expression for the operator L:

$$L[w] = z(1-z)w''+[c-(a+b+1)z]w'-ab \tag{55}$$

which agrees with (44).

2.4. Clausen differential equation

In this section we restrict our attention to the Clausen function ${}_3F_2(a,b,c;d,e;z)$. We write (48) as $M[w]=0$, where

$$M = \frac{1}{z}\theta(\theta+d-1)(\theta+e-1)-(\theta+a)(\theta+b)(\theta+c). \tag{56}$$

Using Viéte formulae as well as the identities

$$\theta = z\frac{d}{dz}, \qquad \theta^2 = z\left(z\frac{d^2}{dz^2}+\frac{d}{dz}\right), \qquad \theta^3 = z\left(z^2\frac{d^3}{dz^3}+3z\frac{d^2}{dz^2}+\frac{d}{dz}\right), \tag{57}$$

one rewrites the first product in (56):

$$\begin{aligned}
&\frac{1}{z}\theta^3+\frac{1}{z}(d+e-2)\theta^2+\frac{1}{z}(de-d-e+1)\theta \\
&= \left(z^2\frac{d^3}{dz^3}+3z\frac{d^2}{dz^2}+\frac{d}{dz}\right)+(d+e-2)\left(z\frac{d^2}{dz^2}+\frac{d}{dz}\right)+(de-d-e+1)\frac{d}{dz} \\
&= z^2\frac{d^3}{dz^3}+(d+e+1)z\frac{d^2}{dz^2}+de\frac{d}{dz}.
\end{aligned} \tag{58}$$

Analogously the second product in (56) yields

$$
\begin{aligned}
(\theta+a)(\theta+b)(\theta+c) &= \theta^3+(a+b+c)\theta^2+(ab+ac+bc)\theta+abc \\
&= z\left(z^2\frac{d^3}{dz^3}+3z\frac{d^2}{dz^2}+\frac{d}{dz}\right)+(a+b+c)z\left(z\frac{d^2}{dz^2}+\frac{d}{dz}\right) \\
&\quad +(ab+ac+bc)z\frac{d}{dz}+abc \\
&= z^3\frac{d^3}{dz^3}+(3+a+b+c)z^2\frac{d^2}{dz^2} \\
&\quad +[1+a+b+c+ab+ac+bc]z\frac{d}{dz}+abc.
\end{aligned} \tag{59}
$$

We get (56) by considering the difference between (58) and (59). Indeed

$$
\begin{aligned}
M &= z^2(1-z)\frac{d^3}{dz^3}+[z(d+e+1)-z^2(3+a+b+c)]\frac{d^2}{dz^2} \\
&\quad +[de-z(1+a+b+c+ab+ac+bc)]\frac{d}{dz}-abc.
\end{aligned} \tag{60}
$$

2.5. Derivative τ_k' of de Branges' function is represented as g.h.f.

As explained in the previous chapter, we are discussing formulae pertaining to

$$|a_{n+1}|\le n+1, \qquad n=1,2,\dots . \tag{61}$$

De Branges functions related to a_{n+1} are defined by

$$\tau_k(t)=k\sum_{\nu=0}^{n-k}(-1)^\nu\frac{(2k+\nu+1)_\nu(2k+2\nu+2)_{n-k-\nu}}{\nu!(k+\nu)(n-k-\nu)!}e^{-t(k+\nu)}, \tag{62}$$

$$k=1,2,\dots,n,$$

and initial conditions in de Branges' differential system are written as

$$\tau_k(0)=n+1-k, \qquad k=1,2,\dots n. \tag{63}$$

From (62) by direct differentiation follows

$$\tau_k'(t)=-ke^{-kt}\sum_{\nu=0}^{n-k}(-1)^\nu\frac{(2k+\nu+1)_\nu(2k+2\nu+2)_{n-k-\nu}}{\nu!(n-k-\nu)!}e^{-\nu t}. \tag{64}$$

The relevance of g.h.f. to de Branges' theorem becomes evident after representing the derivative (62) as the ${}_3F_2$ series. We follow an elegant reasoning by Henrici ([Hnr 68a], p. 605, vol. dedicated to S. Bergman). For reader's convenience some secondary details are offered.

Note the following properties of Pochhammer's symbol:

$$(2a)_{2j}=2^{2j}(a)_j(a+1/2)_j, \tag{65}$$

$$\sum_{\nu=0}^{n} \frac{(a)_\nu}{\nu!} = \frac{(a+1)_n}{n!}, \tag{66}$$

$$(a)_p(a+p)_r = (a)_{p+r}. \tag{67}$$

Formula (66) is easily proved by induction with respect to n. By

$$\frac{1}{(n-k-\nu)!} = (-1)^\nu \frac{(-n+k)_\nu}{(n-k)!} \tag{68}$$

and two cases of (67), we get

$$\begin{aligned} &(2k+1)_\nu(2k+\nu+1)_\nu = (2k+1)_{2\nu}, \\ &(2k+2)_{2\nu}(2k+2\nu+2)_{n-k-\nu} = (2k+2)_{n-k+\nu}, \end{aligned} \tag{69}$$

and may rewrite (64) as

$$\tau_k'(t) = \frac{-k}{(n-k)!} e^{-kt} \sum_{\nu=0}^{n-k} \frac{(-n+k)_\nu (2k+1)_{2\nu}(2k+2)_{n-k-\nu}}{(2k+1)_\nu(2k+2)_{2\nu}} \frac{e^{-\nu t}}{\nu!}. \tag{70}$$

Now, by (65) and (67),

$$(2k+1)_{2\nu} = 2^{2\nu}\left(k+\frac{1}{2}\right)_\nu (k+1)_\nu, \quad (2k+2)_{2\nu} = 2^{2\nu}(k+1)_\nu\left(k+\frac{3}{2}\nu\right), \tag{71}$$

$$(2k+2)_{n-k+\nu} = (2k+2)_{n-k}(n+k+2)_\nu, \tag{72}$$

and from (70), after cancelling $2^{2\nu}(k+1)_\nu$, follows

$$\begin{aligned} -\tau_k'(t) &= \frac{k}{(n-k)!} e^{-kt} \sum_{\nu=0}^{n-k} \frac{(-n+k)_\nu \left(k+\frac{1}{2}\right)_\nu (2k+2)_{n-k}(n+k+2)_\nu}{(2k+1)_\nu \left(k+\frac{3}{2}\right)_\nu} \frac{e^{-\nu t}}{\nu!} \\ &= k\frac{(2k+2)_{n-k}}{(n-k)!} e^{-kt} \sum_{\nu=0}^{n-k} \frac{(-n+k)_\nu \left(k+\frac{1}{2}\right)_\nu (n+k+2)_\nu}{(2k+1)_\nu \left(k+\frac{3}{2}\right)_\nu} \frac{e^{-\nu t}}{\nu!} \\ &= k\frac{(2k+2)_{n-k}}{(n-k)!} e^{-kt} {}_3F_2\left(\begin{matrix} -n+k, n+k+2, k+\frac{1}{2} \\ 2k+1, k+\frac{3}{2} \end{matrix}\middle| x\right), \qquad x = e^{-t}. \end{aligned} \tag{73}$$

Therefore, with the abbreviation $m := n-k \in 0, 1, \ldots, n$:

$$-\tau_k'(t) = k\frac{(2k+2)_m}{m!} e^{-kt} {}_3F_2\left(\begin{matrix} -m, m+2k+2, k+\frac{1}{2} \\ 2k+1, k+\frac{3}{2} \end{matrix}\middle| x\right). \tag{74}$$

In particular, for $t=0$ we have $x=1$, and hence

$$-\tau_k'(t) = k\frac{(2k+2)_m}{m!} {}_3F_2\left(\begin{matrix} -m, m+2k+2, k+\frac{1}{2} \\ k+\frac{3}{2}, 2k+1 \end{matrix}\middle| 1\right). \tag{75}$$

2.6. Clausen identity as a corollary from Cayley-Orr theorem

Functions ${}_2F_1$ and ${}_3F_2$ (with special choice of parameters) are related by a remarkable identity discovered by T. Clausen in 1828:

$$\left\{ {}_2F_1\left(\begin{array}{c} \alpha, \beta \\ \alpha+\beta+\frac{1}{2} \end{array} \middle| z \right) \right\}^2 = {}_3F_2\left(\begin{array}{c} 2\alpha, 2\beta, \alpha+\beta \\ 2\alpha+2\beta, \alpha+\beta+\frac{1}{2} \end{array} \middle| z \right). \tag{76}$$

This identity was used in Askey and Gasper [A,G 76]. Let us recall a *short proof* of (76). In 1858 (thirty years after Clausen's paper) A. Cayley (his portrait is presented in Wikipedia: http://en.wikipedia.org/wiki/Artur_Cayley) stated the following theorem: *If*

$$(1-z)^{\alpha+\beta-\gamma} {}_2F_1(2\alpha, 2\beta; 2\gamma; z) = \sum_{\nu=0}^{\infty} A_\nu z^\nu \tag{77}$$

then

$${}_2F_1\left(\alpha, \beta; \gamma+\frac{1}{2}; z\right) \cdot {}_2F_1\left(\gamma-\alpha, \gamma-b; \gamma+\frac{1}{2}; z\right) = \sum_{\nu=0}^{\infty} \frac{(\gamma)_\nu}{\left(\gamma+\frac{1}{2}\right)_\nu} A_\nu z^\nu. \tag{78}$$

One may say jokingly that: Cayley's theorem shows how to multiply Taylor coefficients A_ν in (77) by the ratio $(\gamma)_\nu/(\gamma+\frac{1}{2})_\nu$.

In case $\gamma := \alpha+\beta$ the left-hand side of (78) becomes much simple (cf. [Bly], p. 86. The present author thanks Ms. K. Posacka for explicit calculations). Since, by definition of the Gauss hypergeometric function:

$${}_2F_1\left(\begin{array}{c} 2\alpha, 2\beta \\ 2\alpha+2\beta \end{array} \middle| z \right) := \sum_{\nu=0}^{\infty} \frac{(2\alpha)_\nu (2\beta)_\nu}{(2\alpha+2\beta)_\nu} \frac{z^\nu}{\nu!} \tag{79}$$

we have (77) with $A_\nu := [(2\alpha)_\nu(2\beta)_\nu]/[92\alpha+2\beta)_\nu \nu!]$. Hence, by the Cayley theorem, we have (78). In view of $\gamma+\alpha+\beta$ the left-hand sides of (76) and (78) are equal. The right-hand sides of (76) and (78) are also equal:

$$\begin{aligned} {}_3F_2\left(\begin{array}{c} 2\alpha, 2\beta, \alpha+\beta; z \\ 2\alpha+2\beta, \alpha+\beta+\frac{1}{2} \end{array} \right) &= \sum_{\nu=0}^{\infty} \frac{(\alpha+\beta)_\nu}{(\alpha+\beta+\frac{1}{2}_\nu} \frac{(2\alpha)_\nu (2\beta)_\nu}{(2\alpha+2\beta)_\nu} \frac{z^\nu}{\nu!} \\ &= \sum_{\nu=0}^{\infty} \frac{(\alpha+\beta)_\nu}{(\alpha+\beta+\frac{1}{2}_\nu} A_\nu. \end{aligned} \tag{80}$$

Hence (76) follows from (78), i.e. one gets the Clausen identity. □

The bad news is that Cayley's results was originally stated without proof. For an algebraic proof of the Cayley's result the reader is referred to [Bly], the basic monograph on generalized hypergeometric functions.

2.7. A proof of Clausen identity using g.h.e

The idea of another proof comes from differential equations. Clausen was interested in situations when a function satisfying an equation of second order determines a solution to an equation of third order. For this reason he was manipulating differential expressions. His ideas are reflected in [Kzr 88], [Gng 99]. In the beginning we shall follow these authors. But there is a difference. We plan to check the identity (76) while ignoring its background. More systematic approach to the Clausen identity can be found in [Hnr 86b]. This will make our task much easier.

Consider the differential operator M defined by (47). After reordering

$$\begin{aligned} M[w] &:= z^2(1-z)w''' + \left[\left(3\alpha+3\beta+\frac{3}{2}\right)z - 3(1+\alpha+\beta)z^2\right]w'' \\ &\quad + \left[2(\alpha+\beta)\left(\alpha+\beta+\frac{1}{2}\right) - \left(1+3(\alpha+\beta)+4\alpha\beta+2(\alpha+\beta)^2\right)z\right]w' \\ &\quad -4\alpha\beta(\alpha+\beta)w. \end{aligned} \tag{81}$$

Formula (81) becomes even simpler with abbreviations

$$s_1 := \alpha+\beta, \quad s_2 = \alpha\beta. \tag{82}$$

Namely

$$\begin{aligned} M[w] &= z^2(1-z)w''' + \left[\left(3s_1+\frac{3}{2}\right)z - (3s_1+3)z^2\right]w'' \\ &\quad + \left[s_1(2s_1+1) - \left(1+3s_1+4s_2+2s_1^2\right)z\right]w' - 4s_1s_2w. \end{aligned} \tag{83}$$

In order to derive $M[w^2]$ note that

$$(w^2)' = 2ww', \quad (w^2)'' = 2(w')^2 + 2ww'', \quad (w^2)''' = 6w'w'' + 2ww''', \tag{84}$$

and, as a consequence

$$\begin{aligned} M[w^2] &= z^2(1-z)[6w'w'' + 2ww'''] \\ &\quad +[(6s_1+3)z - (6s_1+6)z^2]\,[(w')^2 + ww''] \\ &\quad +[s_1(2s_1+1) - (1+3s_1+4s_2+2s_1^2)z]ww' - 4s_1s_2w^2. \end{aligned} \tag{85}$$

Assume now that L stands for differential operator (44) with $a := \alpha$, $b := \beta$ and $c := \gamma = \alpha+\beta+\frac{1}{2}$. With abbreviations (82):

$$L[w] := z(1-z)w'' + \left[s_1 + \frac{1}{2} - (s_1+1)z\right]w' - s_2w \tag{86}$$

and, as a consequence,

$$zL[w] = z^2(1-z)w'' + \left[\left(s_1+\frac{1}{2}\right)z - (s_1+1)z^2\right]w' - s_2zw, \tag{87}$$

$$\begin{aligned} (zL[w])' &= z^2(1-z)w''' + \left[\left(s_1+\frac{1}{2}\right)z - (s_1+1)z^2\right]w'' - s_2zw' \\ &\quad +(2z-3z^2)w'' + \left[s_1+\frac{1}{2} - (2s_1+2)z\right]w' - s_2w. \end{aligned} \tag{88}$$

Consider now auxiliary operators

(89) $$N_1[w] := [(4s_1 - 2)\,w + 6zw'] \cdot L[w]$$
$$= [(4s_1 - 2)\,w + 6zw'] \cdot \left\{ z(1-z)w'' + \left[\left(s_1 + \frac{1}{2} \right) - (s_1+1)z \right] w' - s_2 w \right\},$$

(90) $$N_2[w] := 2w(zL[w])'$$
$$= 2z^2(1-z)ww'' + [2s_1 + 1 - (2s_1+2)z]ww' - 2s_2 w^2.$$

Note that $L[w] = 0$ implies both $N_1[w] = 0$ and $N_2[w] = 0$. We shall soon prove in this paper that

(91) $$M[w^2] = N_1[w] + N_2[w].$$

Hence $L[w] = 0$ implies $M[w^2] = 0$. It follows that both sides of Clausen formula (76) solve the same singular differential equation (47). We shall see later that these solutions are equal.

Now we give the promised proof of (91). The idea is to treat both sides of (91) as polynomials in w, w', w'', w''' and verify that corresponding coefficients are equal. Indeed, from (87), (89) and (90) follows that on both sides of (91) we have:

1). product $f'f''$ appears with
$6z^2(1-z) = (6z)z(1-z)$;

2). product ff''' appears with
$2z^2(1-z) = 0 + 2z^2(1-z)$;

3). square $(f')^2$ appears with
$(6s_1+3)z - (6s_1+6)z^2 = 6z\left[\left(s_1 + \frac{1}{2}\right) - (s_1+1)z\right]$;

4). product ff'' appears with
$(6s_1+3)z - (6s_1+6)z^2 = (4s_1-2)(z-z^2) + (2s_1+5)z - (8+2s_1)z^2$;

5). product ff' appears with
$4s_1 + 2s_1 - (2 + 6s_1 + 4s_1^2 + 8s_2)z = (2s_1-1)\left[(2s_1 + 1 - (2+2s_1)z\right] - 6s_2 z$
$+2s_1 + 1 - (4s_1 + 2s_2 + 4)z$;

6). square f^2 appears with
$-4s_1 s_2 = -s_2(4s_1 - 2) - 2s_2$.

Finally we shall show that both solutions u, ν to the Clausen equation (47):

(92)
$$u := \left[{}_2F_1\left(\alpha, \beta, \alpha+\beta+\frac{1}{2}; z\right)\right]^2 \qquad \nu := {}_3F_2\left(2\alpha, 2\beta, \alpha+\beta; 2\alpha+2\beta, \alpha+\beta+\frac{1}{2}; z\right)$$

are identical.

Assume *provisionally* that in the Clausen equation (47) no integer appears as lower parameter or as a difference between two lower parameters. Then according the theory of singular differential equations, all functions solving (46) and holomorphic in a neighbourhood of $z = 0$ form a one dimensional subspace. (For more details see the next section.) Our *provisional* assumption is removed by taking into account that ${}_2F_1$, ${}_3F_2$ are holomorphic in parameters (see [Rnv 60], Theorem 19 on p. 56). It follows that u, ν in (92) belong to the same one dimensional subspace. Since $u(0) = 1 = \nu(0)$ this implies $u \equiv \nu$ and the Clausen identity (76) is proved.

2.8. Fuchsian singularity and indicial equation

Consider a linear homogeneous differential equation of order n with meromorphic coefficients; see [Inc 26], p. 363. It can be written the form

$$w^{(n)} + p_1(z)w^{(n-1)} + \ldots + p_n(z)w = 0. \tag{93}$$

Assume further that

$$p_1(z) = \frac{P_1(z)}{z},\ p_2(z) = \frac{P_2(z)}{z^2}, \ldots, p_n(z) = \frac{P_n(z)}{z^n}, \tag{94}$$

where functions $P_1, P_2, \ldots, P_n$ are holomorphic in a neighbourhood of $z = 0$. In such situation one says that (93) has *Fuchsian singularity* at the origin. (Sometimes *regular singularity* is used. We propose to avoid this oxymoron.) Consider the algebraic equation in which unknown ρ appears via Pochhammer symbols

$$(\rho)_n + (\rho)_{n-1}P_1(0) + \ldots + (\rho)_1 P_{n-1}(0) + P_n(0). \tag{95}$$

It is called the indicial equation for (93). Its complex roots $\rho_1, \rho_2, \ldots, \rho_n$ are essential to describe the solution space for sigular equation (93). Assume that for $k, j \in \{1, \ldots, n\}$:

$$(k \neq j) \Rightarrow (\rho_k - \rho_j \notin \mathbb{Z}). \tag{96}$$

Then (93) has a fundamental system of solutions

$$z^{r_i} f_i(z), \qquad i = 1, \ldots, n, \tag{97}$$

where each f_i is holomorphic in a neighbourhood of $z = 0$ and $f_i(0) \neq 0$; see Forsyth [Frs 02] vol. IV, p. 95.

Consider now the generalized hypergeometric equation (48), satisfied by ${}_AF_B$ where $A = B + 1$. As usual, denote lower parameters by $b_1, b_2, \ldots b_B$. According to [Slt 66] formula 2.1.2.6, p. 43, therelevant indicial equation is

$$\rho(\rho + b_1 - 1)(\rho + b_2 - 1)\ldots(\rho + b_B - 1) = 0 \tag{98}$$

and the zeros of (98) are

$$0, 1 - b_1, 1 - b_2, \ldots, 1 - b_B. \tag{99}$$

Assumption (96) is satisfied iff no integer appears among the lower parameters or among differences of lower parameters. In such case (48) has a fundamental solution system of the form (97). One-dimensional solution subspace associated with the exponent $\rho = 0$ contains all solutions which are holomorphic in a neighbourhood of $z = 0$.

2.9. An exceptional situation

A simple second order equation

$$w'' - \frac{2}{z}w' + \frac{2}{z^2} = 0 \tag{100}$$

has Fuchsian singularity at $z = 0$. The relevant indicial equation

$$\rho(\rho - 1) - 2\rho + 2 = 0 \tag{101}$$

has zeros $\rho = 1$ and $\rho = 2$. Their difference is an integer; hence the reasoning described in the previous section does not apply. It is easy to check that solutions

$$w_1(z) := z, \qquad w_2(z) = z^2 \tag{102}$$

form a fundamental system. In this example solutions which are holomorphic in a neighbourhood of $z = 0$ form a multidimensional space.

(13) *Remark.* In the next chapter Clausen identity and Gegenbauer formula will be used to deduce the inequality $\tau_k' \leq 0$. By of (75) it suffices to show that

$$F(x) := {}_3F_2\left(\begin{matrix} -m+k, m+2k+2, k+\frac{1}{2} \\ 2k+1, k+\frac{3}{2} \end{matrix} \middle| x\right) \geq 0, \quad x \in (0,1) \tag{103}$$

for $k = 1, \ldots, n$ and $m = n - k$.

References

[A,A,R]–[Wts] See this issue, pp. 98–101.

Department of Mathematical Analysis
Cardinal Stefan Wyszyński University
Dewajtis 5, PL-01-815 Warszawa, Poland
e-mail: skwarczynski@uksw.edu.pl

Presented by Zbigniew Jakubowski at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on July 16, 2010

TWIERDZENIE DE BRANGES'A A UOGÓLNIONE FUNKCJE HIPERGEOMETRYCZNE II

FUNKCJONAŁ DE BRANGES'A I RÓWNANIE HIPERGEOMETRYCZNE

S t r e s z c z e n i e

Oryginalny artykuł [BRN 84] zawiera kluczowe wyjaśnienie: *zagadnienie polega na przekazywaniu informacji za pomocą równania różniczkowego. W tym celu informacja powinna być zakodowana w dogodnej postaci, a następnie przekazana z jednego końca przedziału do drugiego.*

Równanie różniczkowe Löwnera dotyczy propagacji współczynników logarytnicznych. Dostatecznie szybko spojrzenie de Branges'a uzyskało szerszą akceptację. Carl FitzGerald i Christian Pommerenke [F,P 85] dołożyli swoje własne ujęcie. Jeszcze inne przedstawienie znajdujemy w artykule [Krv 85] (zob. szczególnie s. 511–513. Artykuł ten uzyskał nagrodę Chauveneta za opracowanie matematyczne). W obecnym rozdziale staramy się zaprezentować ogólny plan dowodu de Branges'a.

Przychodzi czas na przedyskutowanie roli funkcji hipergeometrycznej Gaussa ${}_2F_1$ i jej uogólnień. Najprostsze z tych uogólnień to funkcja Clausena ${}_3F_2$, która spełnia zasadniczą rolę w końcowej części dowodu de Branges'a. Przedstawiamy dwa dowody tożsamości Clausena: jeden bardzo krótki, a drugi znacznie dłuższy, lecz chyba bardziej bezpośredni. Po takim przygotowaniu wyprowadzamy w następnym rozdziale nierówność $r \leq 0$, skąd już łatwo wynika warunek (2).

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 89–103

Dedicated to Professor Roman Stanisław Ingarden
on the occasion of his ninetieth birthday

Maciej Skwarczyński

DE BRANGES THEOREM AND GENERALIZED HYPERGEOMETRIC FUNCTIONS III

BASIC PROPERTIES OF DE BRANGES FUNCTIONS

Summary

A look at (II.37) shows that the Clausen identity is applicable to ${}_3F_2$ when one of the upper parameters is the arithmetic mean of remaining upper parameters and moreover, when multiplied by 2 or added to 1/2, yields lower parameters. If Clausen were directly identity applicable to the function

(1) $$F(x) := {}_3F_2\left(\begin{matrix} -m, m+2k+2.k+\frac{1}{2} \\ 2k+1, k+\frac{3}{2}\end{matrix}\,\middle|\, x\right),$$

the inequality $F(x) \geq 0$ would be obvious. Unfortunately this is not the case. Nevertheless one can prove $F(x) \geq 0$ by representing F as a finite sum of terms, to which the Clausen identity applies. Following Kazarinoff [Kzr 88] one reduces the problem to hypergeometric functions ${}_2F_1$ using the Rainville operator and Gegenbauer formula.

In the final chapter we conclude de Branges proof by verifying the condition (II.**4**). Required initial conditions $\tau_k(0) = n - k + 1$ will be derived from a classical result of G. N. Watson.

0. Initial remark

Formulae numbers (1) etc. and statement numbers **(1)** etc. referring to part I, II od the paper are quoted as (I.1), (II.1) etc. and **(I.1), (II.1)** etc., respectively. Conditions (**1**)–(**4**) on p. 68 in Part II are quoted as (II.**1**)–(II.**4**). Acronyms below [Rnv 60] etc. usually consist of first three consonants of author's name, followed by the year of publication. List of references is constructed alphabetically according to latters of the acronym. When no ambiguity results acronyms of this list appear without the year of publication.

1. De Branges functions have negative derivative

1.1. Rainville inegral representation

We begin with

(1) Rainville Theorem. *If* $p \leq q+1$, *if* $\operatorname{Re} b_1 > \operatorname{Re} a_1 > 0$, *if none of* $b_1, b_2, \ldots, b_q$ *is zero or a negative integer, and if* $|z| < 1$ *then*

$$(2)\qquad {}_pF_q\left(\begin{matrix} a_1, a_2, \ldots, a_p \\ b_1, b_2, \ldots, b_q \end{matrix} \middle| z\right)$$

$$= \frac{\Gamma(b_1)}{\Gamma(a_1)\Gamma(b_1-a_1)} \int_0^1 t^{a_1-1}(1-t)^{b_1-a_1-1}{}_{p-1}F_{q-1}\left(\begin{matrix} a_2, \ldots, a_p \\ b_2, \ldots, b_q \end{matrix} \middle| zt\right) dt.$$

For a proof see [Rnv 60], p. 85.

Denote by $R = R(a_1; b_1)$ the *Rainville integral operator* appearing in (2). It is a generalization of the Pochhammer integral representation; cf. [Slt 66], formula 1.6.6. For connection with fractional integration and the beta integral operator see Bertram Ross [Rss 75]. In acting on hypergeometric functions ${}_{p-1}F_{q-1}$ it has an effect of adjoining two parameters: upper a_1 and lower b_1. Note that it is easy to represent $F = {}_3F_2$ in (1) as an image of a suitable $G + {}_2F_1$ under the Rainville operator. Namely,

(3)

$$R_{k+(1/2);2k+1}\left[{}_2F_1\left(\begin{matrix} -m, m+2k+2 \\ k+\frac{3}{2} \end{matrix} \middle| x\right)\right] = {}_3F_2\left(\begin{matrix} -m, m+2k+2, k+\frac{1}{2} \\ k+\frac{3}{2}, 2k+1 \end{matrix} \middle| x\right).$$

1.2. Gegenbauer formula

This is second (very important) ingredient in the proof of $F(x) \geq 0$. Gegenbauer polynomials $C_m^\lambda(x)$, $m = 0, 1, \ldots$ (of order $\lambda > -1/2$) are defined by the relation

$$(4)\qquad (1-2xu+u^2)^{-\lambda} = \sum_{m=0}^{\infty} C_m^\lambda(x) u^m.$$

Direct calculation yields

$$(1-2xu+u^2)^{-\lambda} = \sum_{m=0}^{\infty} \frac{(\lambda)_m}{m!}(2xu-u^2)^m = \sum_{m=0}^{\infty}\sum_{j=0}^{m} \frac{(\lambda)_m}{j!(m-2j)!}(2x)^{m-2j}u^{m+j}$$

$$(5)\qquad = \sum_{m=0}^{\infty} u^m \sum_{j=0}^{[m/2]} (-1)^j \frac{(\lambda)_{m-j}}{m!(m-j)!}(2x)^{m-j}u^{m+j}.$$

In the latter transformation we have used a summation trick ([Rnv 60], p. 58):

$$\sum_{m=0}^{\infty}\sum_{j=0}^{m} C(m,j) = \sum_{m=0}^{\infty}\sum_{j=0}^{[m/2]} C(j,m-j). \tag{6}$$

By definition (4) the development (5) yields an explicit formula

$$C_m^{\lambda}(x) = \sum_{j=0}^{[m/2]} \frac{(-1)^j(\lambda)_{m-j}}{j!(m-2j)!}(2x)^{m-2j} \tag{7}$$

which in turn leads to the hypergeometric representation

$$C_m^{\lambda}(x) = \frac{\Gamma(m+2\lambda)}{\Gamma(m+1)\Gamma(2\lambda)}\,{}_2F_1\left(\begin{matrix} -m, m+2\lambda \\ \lambda+\frac{1}{2}\end{matrix}\,\middle|\,\frac{1-x}{2}\right); \tag{8}$$

see [B,E 75] or [Rnv 60], p. 279, formula (15). Note that the polynomial C_j^{λ} has the degree j. This implies that polynomials $C_0^{\lambda}, C_1^{\lambda}, \dots, C_m^{\lambda}$ form a basis in the space of polynomials with degrees not exceeding m.

(5.2) *Remark.* Determination of coordinates with respect to the abovementioned basis may become easier with information that the system C_j^{λ}, $j = 0, 1, \dots, m$, is orthogonal over the interval $[-1, 1]$ with respect to the elementary weight

$$g(x) := (1-x^2)^{\lambda-\frac{1}{2}}. \tag{9}$$

An important example is

$$\frac{(2x)^m}{m!} = \sum_{j=0}^{[m/2]} \frac{(\lambda+m-2j)}{j!(\lambda)_{m+1-j}} C_{m-2j}^{\lambda}(x); \tag{10}$$

see [Rnv 60], p. 283. It leads to the following fundamental identity:

(5.3) *Gegenbauer formula.* For $\nu \in (\lambda, +\infty)$:

$$C_m^{\nu}(x) = \sum_{j=0}^{[m/2]} c_j C_{m-2j}^{\lambda}(x), \tag{11}$$

where

$$c_j = \frac{(m-2j+\lambda)\Gamma(\lambda)(\nu-\lambda)_j\Gamma(m+\nu-j)}{j!\Gamma(\nu)\Gamma(m+\lambda-j+1)}. \tag{12}$$

With representation (8) one rewrites the Gegenbauer formula (11) as hypergeometric identity. Namely, for every $\lambda \in (0, \nu)$:

$${}_2F_1\left(\begin{matrix} -m, m+2\nu \\ \nu+\frac{1}{2}\end{matrix}\,\middle|\,\frac{1-x}{2}\right) = \sum_{j=0}^{[m/2]} \rho_j \cdot {}_2F_1\left(\begin{matrix} 2j-m, m-2j+2\lambda \\ \lambda+\frac{1}{2}\end{matrix}\,\middle|\,\frac{1-x}{2}\right) \tag{13}$$

with positive coefficients

$$\rho_j := \frac{m!(2\lambda)_{m-2}}{(m-2j)!(2\nu)_m} c_j > 0, \qquad j = 1, \dots, [m/2]. \tag{14}$$

Proof. Here we expand the reasoring given in [Kzr 88]; for a proof of Gegenbauer's formula see also [A,G 86]. In order to determine coefficients $c_j = c_j(m,\nu,\lambda)$ in (11) one substitutes (10) into (7), where in the latter ν is written instead of λ. This yields

$$c_j = \frac{\Gamma(\lambda)}{\Gamma(\nu)}\,\frac{(\lambda+m-2j)}{j!}\sum_{p=0}^{j}(-1)^p\frac{\binom{j}{p}\Gamma(\nu+m-p)}{\Gamma(\lambda+m-p+j+1)}. \tag{15}$$

The sum in (15) is rewritten as follows:

$$\begin{aligned}
&\sum_{p=0}^{j}(-1)^p\frac{\binom{j}{p}\Gamma(\nu+m-p)}{\Gamma(\lambda+m-p+j+1)} \\
&= \frac{\Gamma(\nu+m)}{\Gamma(\lambda+m-j+1)}\sum_{p=0}^{\infty}(-1)_p\frac{(j-m-\lambda)_p}{p!(1-m-\nu)_p} \\
&= \frac{\Gamma(\nu+m)}{\Gamma(\lambda+m-j+1)}\,{}_2F_1\left(\left.\begin{matrix}-j, j-m-\lambda\\ 1-m-\nu\end{matrix}\right| 1\right) \\
&= \frac{\Gamma(\nu+m)}{\Gamma(\lambda+m-j+1)}\,\frac{(1+\lambda-j-\nu)_j}{(1-\nu-m)_j},
\end{aligned} \tag{16}$$

where (in the last transformation) Chu-Vandermonde theorem was used; see [Slt 66], p. 28 or [Ths 92], p. 39.

Finally (15) and (16) yield the desired representation (11) with

$$C_j(m,\nu,\lambda) = \frac{\Gamma(\lambda)}{\Gamma(\nu)}\,\frac{(m-2j+\lambda)(\nu-\lambda)_j}{j!}\,\frac{\Gamma(m+\nu-j)}{\Gamma(m+\lambda-j+1)}. \tag{17}$$

Equivalently; see [Hua 63]:

$$c_j(m,\nu,\lambda) = \frac{(m-2j+\lambda)}{j!}\,\frac{\Gamma(\lambda)}{\Gamma(\nu)}\,\frac{\Gamma(j+\nu-\lambda)}{\Gamma(\nu-\lambda)}\,\frac{\Gamma(m+\nu-j)}{\Gamma(m+\lambda-j+1)}. \tag{18}$$

1.3. De Branges functions have negative derivatives

In view of (II.75) the desired inquality $\tau_k' \le 0$ is reduced to

$$F(x) := {}_3F_2\left(\left.\begin{matrix}-m+k, m+2k+2, k+\frac{1}{2}\\ 2k+1, k+\frac{3}{2}\end{matrix}\right| x\right) \ge 0, \quad x \in (0,1). \tag{19}$$

The latter follows by applying a suitable Rainville operator to both sides of (13) provided independent (and intelligent) choices of ν and λ are made.

Indeed, let us choose ν in such a way that parameters of ${}_2F_1$ on the left-hand side of (13) agree with initial parameters of $F = {}_3F_2$ in (19). This can be done with $\nu := k+1$. Consider now the Rainville operator which takes ${}_2F_1$ in (13) onto $F = {}_3F_2$ in (19). It is $R(k+\frac{1}{2}; 2k+1)$. Applying this Rainville operator to *both sides* of (13) (with $(1-x)/2$ replaced by x) yields

$$F(x) = \sum_{j=0}^{[m/2]} \rho_j \cdot {}_3F_2\left(\begin{matrix} 2j-m, m-2j+2\lambda, k+\frac{1}{2} \\ \lambda+\frac{1}{2}, 2k+1 \end{matrix} \middle| x \right), \quad x \in (0,1). \tag{20}$$

Obviously F does not depend on λ. But an intelligent choice of λ makes the Clausen identity applicable to all ${}_3F_2$ on the right-hand side of (20). Indeed, it is sufficient to take $\lambda := k + \frac{1}{2}$. Since the right-hand side of (20) is now positive we can see that $F \geq 0$. □

With the above we have cleared the condition (II.**2**) on. p. 68. The remaining condition (II.**4**) will be delt with in the next (last) chapter.

2. De Branges functions satisfy initial conditions

2.1. Watson summation lemma

This result concerns terminating ${}_3F_2$ series. G. N. Watson published it under the title *A note on generalized hypergeometric series* in the Proc. London Math. Soc. (2) **23** (1925) p. xiii (his portrait is presented in Biogr. Mems Fell. Royal Soc. (London) **1966**, no. 12, 520–530 in the obituary article by E. T. Whittaker; see doi:10.1098/rsbm.1066.0026). He originally stated it as follows. For n even

$$\begin{aligned} &{}_3F_2\left(\begin{matrix} -n, \lambda, 2\lambda+2\mu+n-1 \\ 2\lambda, \lambda+\mu \end{matrix} \middle| 1 \right) \\ &= \frac{n!\Gamma\left(\lambda+\frac{1}{2}n\right)\Gamma\left(\mu+\frac{1}{2}n\right)\Gamma(2\lambda)\Gamma(\lambda+\mu)}{\left(\frac{1}{2}n\right)!\Gamma\left(\lambda+\mu+\frac{1}{2}n\right)\Gamma(2\lambda+n)\Gamma(\lambda)\Gamma(\mu)}, \end{aligned} \tag{21}$$

while for n odd the left-hand side of (21) equals 0. For details see [Wts 25].

In order to conform (21) to our notation (m instead of n, c instead of λ, and b instead of $c + \mu - 2^{-1}$) we restate it as follows.

(4) Watson summation lemma. *Assume $m \in \{0, 1, 2, \ldots\}$. For even m the Clausen function satisfies*

$${}_3F_2\left(\begin{matrix} -m, 2b+m, c \\ b+\frac{1}{2}, 2c \end{matrix} \middle| 1 \right) = \frac{\left(\frac{1}{2}\right)_{m/2}\left(b-c+\frac{1}{2}\right)_{m/2}}{\left(b+\frac{1}{2}\right)_{m/2}\left(c+\frac{1}{2}\right)_{m/2}}, \tag{22}$$

while for m ***odd*** *the left-hand side of* (22) *is* ***zero.***

(5) *Remark.* This formulation agrees with [Hnr 86a], p. 609; see also [Bly 35], p. 16, and [Whp 29], p. 118. We shall deduce Lemma **(4)** from a more general Watson theorem; see the next section.

Now, with $m := n-k$, $b = k+1$, $c = k + \frac{1}{2}$, we apply Lemma **(4)** to ${}_3F_2$ in the formula (75). This yields $\tau_k'(0) = 0$ for $n-k$ odd. For $n-k$ even, say $n-k = 2s$, we have

$$
{}_3F_2\left(\begin{matrix} -m, m+2k+2, k+\frac{1}{2} \\ k+\frac{3}{2}, 2k+1 \end{matrix}\right) = \frac{\left(\frac{1}{2}\right)_s (1)_s}{\left(k+\frac{3}{2}\right)_s (k+1)_s} \tag{23}
$$

and by

$$
\frac{(2k+2)_{2s}}{(2s)!} = \frac{2^{2s}(k+1)_s \left(k+\frac{3}{2}\right)_s}{2^{2s}\left(\frac{1}{2}\right)_s (1)_s} \tag{24}
$$

the final result is

$$
-\frac{\tau_k'(0)}{k} = \begin{cases} 0 & \text{if} \quad n-k \quad \text{is odd,} \\ 1 & \text{if} \quad n-k \quad \text{is even.} \end{cases} \tag{25}
$$

We know that de Branges' functions satisfy de Branges equations. Hence, at $t = 0$,

$$
\tau_k(0) - \tau_{k+1}(0) = -\left(\frac{\tau_k'(0)}{k} + \frac{\tau_{k+1}'(0)}{k+1}\right), \qquad k = 1, \ldots, n. \tag{26}
$$

In view of (24) the right-hand side is always 1. Summing up the latter $n+1-k$ equations (25) yields the desired condition (II.**4**)

$$
\tau_k(0) = \sum_{\nu=k}^{n} \tau_\nu(0) - \tau_{\nu+1}(0) = \sum_{\nu=k}^{n} 1 = n+1-k. \tag{27}
$$

This ends our detailed description of de Branges proof. □

(6) *Remark.* With (22) the condition (II.**4**) is cleared. De Branges' theorem has been verified again, twenty five years after the original discovery.

(7) *Remark.* It would be interesting to derive the Watson formula (22) from a more general summation theorem. Yet, the role of convergence conditions in the non-terminating case needs further explanations. So far for proof of Watson's formula we have to refer to a very important and elegant idea in Watson's original paper [Wts 25].

2.2. Watson summation theorem

There are many proofs of Watson's lemma. Perhaps one should make the distinction between "geometric" and "algebraic" approach. Original proof in [Wts 25] is geometric in the sense that it appeals to orthogonality. On the other hand the proof in [Bly 35], p. 16, is based on formal operations on series. We verify Watson's lemma immediately, replacing b by $2b-a$ and then a by $-m$ in the following, more general result

(8) Watson summation theorem. *We have*

$$
{}_3F_2\left(\begin{matrix} a, b, c \\ \frac{1}{2}(a+b+1), 2c \end{matrix} \middle| 1\right) = \Gamma\left(\begin{matrix} \frac{1}{2}, 2c, \frac{1}{2}a+\frac{1}{2}b+\frac{1}{2}, \frac{1}{2}-\frac{1}{2}a-\frac{1}{2}b+c \\ \frac{1}{2}+\frac{1}{2}a, \frac{1}{2}+\frac{1}{2}b, \frac{1}{2}-\frac{1}{2}a+c, -\frac{1}{2}b+c \end{matrix}\right). \tag{28}
$$

Quite recently a new proof of (28) has been found by Arjun K. Rathie and R. B. Paris [R,P]. We shall describe it in the next section. It uses only two summation theorems for ${}_2F_1$ and an auxiliary lemma. The rest of the present section is devoted to preliminaries.

(9) First summation theorem of Gauss. *Hypergeometric function ${}_2F_1$ satisfies*

$$(29)\qquad {}_2F_1\left(\begin{array}{c} a, b \\ c \end{array}\middle| 1\right) = \Gamma\left(\begin{array}{c} c, c-a-b \\ c-a, c-b \end{array}\right), \qquad \mathrm{Re}(c-a-b) > 0.$$

For the proof see Lucy Joan Slater [Slt 66], p. 29.

(10) Second summation theorem of Gauss. *Hypergeometric function ${}_2F_1$ satisfies*

$$(30)\qquad {}_2F_1\left(\begin{array}{c} a, b; \frac{1}{2} \\ \frac{a}{2}+\frac{b}{2}+\frac{1}{2} \end{array}\right) = \Gamma\left(\begin{array}{c} \frac{1}{2}, \frac{a}{2}+\frac{b}{2}+\frac{1}{2} \\ \frac{a}{2}+\frac{1}{2}, \frac{b}{2}+\frac{1}{2} \end{array}\right).$$

For the proof see Bateman, Erdélyi, vol. I, formula 50. Note that in [Slt 66] the formula 1.7.1.9 on p. 32 is incompatible with the result quoted in [B,E 77].

(11) Auxiliary lemma (Rathie and Paris). *Assume $2c \neq -1, -2, \ldots$. For every $k \in \{0, 1, 2, \ldots\}$ we have*

$$(31)\qquad \frac{(c)_k}{(2c)_k} = \sum_{m=0}^{[k/2]} \frac{2^{-k-2m} k!}{\left(c+\frac{1}{2}\right)_m m!(k-2m)!},$$

where $[k/2]$ denotes the integer part of $k/2$.

Proof. The ratio of Pochhammer symbols on the left-hand side of (31) is rewritten as Γ expression. The latter is calculated by reversing the first summation theorem of Gauss. The result is

$$(32)\qquad \frac{(c)_k}{(2c)_k} = 2^{P-k}{}_2F_1\left(\begin{array}{c} -\frac{1}{2}k, \frac{1}{2}-\frac{1}{2}k; 1 \\ c+\frac{1}{2} \end{array}\right) = 2^{-k} \sum_{m=0}^{[k/2]} \frac{\left(-\frac{1}{2}k\right)_m \left(\frac{1}{2}-\frac{1}{2}k\right)_m}{\left(c+\frac{1}{2}\right)_m m!}.$$

The needed formula (31) follows in view of an elementary identity

$$(33)\qquad \left(-\frac{1}{2}k\right)_m \left(\frac{1}{2}-\frac{1}{2}k\right)_m = \frac{2^{-2m} k!}{(k-2m)!}.$$

We can now present what follows.

2.3. A new proof of Watson's theorem by Rathie and Paris

Denote by S the left-hand side of (28) written as Clausen series. By (32):

$$S = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{\left(\frac{a}{2}+\frac{b}{2}+\frac{1}{2}\right)_k} \frac{1}{k!} \frac{(c)_k}{(2c)_k} \tag{34}$$

$$= \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{\left(\frac{a}{2}+\frac{b}{2}+\frac{1}{2}\right)_k k!} \sum_{m=0}^{[k/2]} \frac{2^{-k-2m} k!}{\left(c+\frac{1}{2}\right)_m m!(k-2m)!}$$

$$= \sum_{k=0}^{\infty} \sum_{m=0}^{[k/2]} \frac{(a)_k (b)_k}{\left(\frac{a}{2}+\frac{b}{2}+\frac{1}{2}\right)_k} \frac{2^{-k-2m}}{\left(c+\frac{1}{2}\right)_m m!(k-2m)!}.$$

We now change the order of summation according to the formula (8) on p. 57 in [Rnv]:

$$\sum_{k=0}^{\infty} \sum_{m=0}^{[k/2]} A(m,k) = \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} A(m,k+2m). \tag{35}$$

The result is

$$S = \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{(a)_{k+2m} (b)_{k+2m} 2^{-k-4m}}{\left(\frac{a}{2}+\frac{b}{2}+\frac{1}{2}\right)_{k+2m} \left(c+\frac{1}{2}\right)_m m! k!} \tag{36}$$

$$= \sum_{m=0}^{\infty} \frac{(a)_{2m} (b)_{2m} 2^{-4m}}{\left(\frac{a}{2}+\frac{b}{2}+\frac{1}{2}\right)_{2m} \left(c+\frac{1}{2}\right)_m m!} \sum_{k=0}^{\infty} \frac{(a+2m)_k (b+2m)_k 2^{-k}}{\left(\frac{a}{2}+\frac{b}{2}+\frac{1}{2}+2m\right)_k k!}.$$

In the last line the sum over k is rewritten as the ${}_2F_1$ series. Its value is found with the second Gauss summation theorem. Namely:

$${}_2F_1\left(\begin{matrix} a+2m, b+2m; \frac{1}{2} \\ \frac{a}{2}+\frac{b}{2}+\frac{1}{2}+2m \end{matrix}\right) = \Gamma\left(\begin{matrix} \frac{1}{2}, \frac{a}{2}+\frac{b}{2}+\frac{1}{2} \\ \frac{a}{2}+\frac{1}{2}, \frac{b}{2}+\frac{1}{2} \end{matrix}\right) \frac{\left(\frac{a}{2}+\frac{b}{2}+\frac{1}{2}\right)_{2m}}{\left(\frac{a}{2}+\frac{1}{2}\right)_m \left(\frac{b}{2}+\frac{1}{2}\right)_m}. \tag{37}$$

When (37) is substituted into (36) we get after a simple computation with Pochhammer symbols

$$F = \Gamma\left(\begin{matrix} \frac{1}{2}, \frac{a}{2}+\frac{b}{2}+\frac{1}{2} \\ \frac{a}{2}+\frac{1}{2}, \frac{b}{2}+\frac{1}{2} \end{matrix}\right) \sum_{m=0}^{\infty} \frac{\left(\frac{1}{2}a\right)_m \left(\frac{1}{2}b\right)_m}{\left(c+\frac{1}{2}\right)_m}, \tag{38}$$

where, as in (1), F stands for ${}_3F_2$. Evaluating the sum with the first Gauss summation theorem we get on the right-hand side of (38) four Γ-s in the numerator and four Γ-s in the denominator. This is the desired formula (28). □

The last section is devoted to the past.

2.4. Understanding the past

Great mathematical discovery gives an opportunity for deeper understanding of previous developments. We have mentioned the work of A. Cayley. First published proof of the product formula (78) was given by W. Orr (1899). Another proof, based on systematic use of g.h.f., was offered in 1927 by F. W. J. Whipple, an astronomer famous for discovering elliptical character of meteorite trajectories. The next impulse came from G. Hardy (1877–1947) who got interested in results on g.h.f. rediscovered by S. Ramanujan, a talented mathematician from Madras. Hardy inspired W. N. Bailey, a lecturer from Manchester, to write a monograph [Bly 35] devoted to known results on g.h.f. This book quotes the Watson summation lemma.

G. N. Watson, a pupil of F. S. Macaulay at St. Paul School, was matriculated to Trinity College, Cambridge. His teachers, besides Hardy, were Whittaker (who soon moved to Edinburgh) and Barnes (bishop of Birmingham). He attended also some lectures given by Berry, Hobson, Forsyth and conducted correspondence with Lord Rayleigh. Watson graduated as Senior Wrangler in 1907, meaning that he was ranked first among those who were awarded First Class degrees. He become Trinity fellow in 1910. His interest included solvable cases of quintic equation. From 1918 to 1951 Watson was Mason professor of Pure Mathematics at Birmingham. In [Wts 25] (quoted by Bailey) Watson wrote:

> *In a recently published paper* Proc. Camb. Phil. Soc. **21** (1923), 492–503, *entitled "A chapter from Ramanujan's Note-Book", Prof. Hardy has given a catalogue of all the known cases in which a series of the type ${}_3F_2$ with last element unity is expressible in terms of Gamma functions. Incidentally he quoted formula which I had discovered* (...) *Prof. Hardy pointed out that it was a special case of a formula discovered by Ramanujan, and I did not attach any particular importance to it. But since the formula plays a moderately important part in Mr. Whipple's paper, it seems worth while to supplement his paper by giving my own proof of it. This proof was suggested to me by the proof which I constructed of a formula concerning the square of a Bessel function discovered by Prof. Jolliffe.* (See my Theory of Bessel functions, par. 16.3).

(12) *Remark.* Macaulay is better known as the author of the *The Algebraic Theory of Modular Systems*, Cambridge 1916; see [ME 77], vpl. 3, p. 69.

(13) *Remark.* In 1995 an unpublished Watson's lecture on solvable quintics (1918) has been found in the library of Birmingham University by Bruce L. Berndt (Urbana, Illinois).

Now let us look at Oxford. J. L. Burchnall has completed with honours his undergraduate education at Christ Church just before the outbreak of the First World War. After brave service he taught at Army School at Oxford, then took an academic position at Durham. In the years 1939–1951 Burchnall was Professor of Mathematics at Durham. A large part of his mathematical work was done jointly with T. W. Chaundy. Their research on differential operators brought connections with algebraic geometry, and the knowledge of differential equations was a guiding light in their persevered research on special functions.

In 1926 the study of complex differential equations, much in the tradition of English school (Forsyth and others) was undertaken in places far away from London.

In a newly founded Egyptian University in Cairo the chair of Mathematics was offered to E. L. Ince, who published there well known monograph *Ordinary differential equations.* Recall that Ince was a student of Chrystal and Whittaker.

The influence of English school was also felt beyond the Atlantic. In 1922 Mathematics Department of Michigan University was joined in by Ruel Vance Churchill who worked there until 1966. The selection of materials for his book *Fourier Series and the Boundary Problems* was influenced (among others) by results of Carslaw, Watson and Hobson. He was an advisor to Earl D. Rainville who in 1939 defended Ph. D. dissertation *Linear differential invariance under operators related to the Laplace transformation.*

Earl D. Rainville (1907–1966) received his B. A. at the University of Colorado (1930). In 1941 he began working at Michigan, where he wrote several very readable textbooks including now classic *Special Functions.* He advised eight Ph. D. students; the first of them (1946) was Sister Mary Celine Fasenmyer (see Amer. Math. Monthly **56** (1949), 14). Her work *Some generalized hypergeometric polynomials* contains an algorithm for deducing recurrence relations between hypergeometric expressions. Similar algorithms were used in recent computer experiments related to de Brange's theorem.

Afer 1950 a great centre for classical analysis was created at Stanford, California. Among prominent members of the Mathematics Department were Bergman, Löwner, Pólya, Royden, Schiffer, and Szegö, The name of M. Schiffer (1911–1997) [F,O,O] is inseparable from the Bieberbach conjecture (his paper with Z. Charzyński [C,S 60] presents an application of Grunsky inequalities to the proof of $|a_4| \leq 4$). He entered the University in Bonn where he studied mathematics under Bieberbach, Schmidt and Issai Schur. Schiffer's first paper *Finiteness theorems of invariant theory*, published in 1934 in Mathematische Zeitschrift, was written under supervision of I. Schur (Schur's lectures in Berlin [Sch 68] were prepared for print by H. Grunsky). Later, motivated by the Bieberbach conjecture, he developed a method now known as Schiffer's variation. It got him Ph. D. at the Hebrew University of Jerusalem in 1938. In 1952 Schiffer became professor of mathematics at Stanford University.

(13) *Remark.* De Branges' discovery was a part of a larger attempt (de Branges' response [Brn 94] to the Steele prize). We are very fortunate. Some of us may follow the indicated path. To quote "Apology": *David Hilbert is said to have assigned the Riemann Hypothesis as a thesis problem to his student Erhard Schmidt.* Contemporary mathematicians are invited to enhance their efforts toward the Riemann Hypothesis. De Branges' "Apology" deserves very careful reading.

References

[A,A,R] G. E. Andrews, R. Askey, and R. Roy, *Special functions*, Encyclopedia of Math. Appl. **71**, Cambridge Univ. Press, Cambridge 1999.

[A,G 76] R. Askey and G. Gasper, *Positive Jacobi polynomial sums II*, Amer. J. Math. **98** (1976), 709–737.

[A,G 86] —, —, *Inequalities for polynomials*, in: [B,D,D,M], pp. 7–32.

[Ahl] L. V. Ahlfors, *Conformal Invariants*, McGraw-Hill, New York 1973.

[B,D,D,M] *The Bieberbach conjecture, Proceedings of the symposium on the occasion of the proof*, A. Baernstein, D. Drasin, P. Duren, A. Marden eds., AMS, Providence, RI, 1986.

[B,E] H. Bateman and A. Erdelyi, *Higher Transcendental functions vol. 3*, Amer. Math. Soc., Providence, RI, 1977.

[Bbr] L. Bieberbach, *Über die Koeffizienten derjenige Potenzreihen, welche eine schlichte Abbildung des Einheitkreises vermitteln*, Sitzungsber. Preuss. Akad. Wiss. **38** (1916), 940–955.

[Bly] W. N. Bailey, *Generalized Hypergeometric Series*, Cambridge Univ. Press, Cambridge 1935.

[Bri 52] S. D. Bernardi, *A survey of the development of the theory of schlicht functions*, Duke Math. J. **19** (1952), 263–287.

[Bri 83] —, *Bibliography of schlicht functions*, Mariner, Tampa, FL 1983.

[Brn 85] L. de Branges, *A proof of the Bieberbach conjecture*, Acta Math. **154** (1985), 137–152.

[Brn 86] —, *Unitary linear systems whose transfer funnctions are Riemann mapping functions*, in: Operator Theory and Systems, Birkhauser, Basel 1986, pp. 105–124.

[Brn 94] —, *Responce to 1994*, in: Notices AMS, **41** (1994), 907–910.

[C,S] Z. Charzyński and M. M. Schiffer, *A new proof of the Bieberbach conjecture for the fourth coefficient*, Arch. Rat. Mech. Anal. **5** (1960), 187–193.

[Cls] T. Clausen, *Über die Falle wenn die Reihe ... ein quadrat von der Form ... hat*, Journal für Math. **3** (1828), 80–95.

[Cnw] J. B. Conway, *Functions of One Complex Variable I, II*, Springer-Verlag, Heidelberg-New York 1995.

[Crt] C. Carathéodory, *Conformal Representations*, Cambridge Tracts in Mathematics and Mathematical Physics **28**, Cambridge University Press, London 1932.

[Drn] P. L. Duren, *Univalent functions*, Springer-Verlag, Heidelberg-New York 1983.

[F,O,O] R. Finn, B. Osgood, and R. Osserman, *Menahem Max Schiffer 1911–1997* (Letter to the Editor), Notices AMS **49** (2002), 886.

[F,P] C. H. FitzGerald and Ch. Pommerenke, *The de Branges theorem on univalent functions*, Trans. Amer. Math. Soc. **290** (1985), 683–690.

[Frs] A. R. Forsyth, *Theory of Differential Equations*, six volumes, Cambridge Univ. Press, Cambridge 1900–1902, especially vol. 4, pp. 78–97.

[Ftz] C. H. FitzGerald, *The Bieberbach Conjecture: Retrospective*, Notices AMS **32** (1985), 2–6.

[G,L] A. W. Goodman and R. J. Libera, *A review of M. S. Robertson work*, Complex Variables and Elliptic Equations **3** (1984), 11–25.

[G lz] G. M. Goluzin, *Geometric Theory of Functions of a Complex Variable*, Moskva 1952 (Russian); Engl. transl., Amer. Math. Soc., Providence, RI, 1969.

[Gng] S. Gong, *The Bieberbach Conjecture*, Studies in Advanced Math., AMS/IP 1999.

[Grn 72] A. Z. Grinshpan, *Logarithmic coefficients of functions in the class S*, Sibirsk Mat. Zh. **13** (1972), 1145–1157; Siberian Math. J. **13** (1972), 793–801 (Russian).

[Grn 99] —, *The Bieberbach conjecture and Milin's functionals. Dedicated to Isaak Moiseevich Milin (1919–1992)*, The AMS Monthly **106** (1999), 203–214.

[Gsp] G. Gasper, *A short proof of an inequality used by de Branges in his proof of the Bieberbach, Robertson and Milin conjectures*, Complex Variables **7** (1986), 45–50.

[Gsy] H. Grunsky, *Koeffizientenbedingungen für schlichte abbilde meromorphe Funktionen*, Math. Z. **45** (1939), 29–61.

[Hll] E. Hille, *Analytic Function Theory I, II*, Ginn and Co., New York 1962.

[Hnr 86a] P. Henrici, *Applied and Computational Complex Analysis*, three volumes, Vol. 3 (dedicated to S. Bergman). Wiley, New York 1986.

[Hnr 86b] —, *Product theorems for formal hypergeometric series*, Preprint 09318 -86 Mathematical Sciences Research Institute, Berkeley 1986.

[Hua] L. K. Hua, *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains*, Translations of Math. Monographs **6**, AMS, Providence, 1963.

[Inc] E. L. Ince, *Ordinary Differential Equations*, Dover, New York 1926.

[J,Z,Z] Z. J. Jakubowski, A. Zielińska, and K. Zyskowska, *Sharp estimation of even coeffitients of bounded symmetric univalent functions*, Ann. Polon. Math. **40** (1983), 193–206.

[K,Ł] J. Krzyż and J. Ławrynowicz, *Elementy analizy zespolonej*, WNT, Warszawa 1981.

[K,S] W. Koepf and D. Schmersau, *On de Branges theorem*, Complex Variables **15** (1996), 213–230.

[Krv] J. Korevaar, *Ludwig Bieberbach's conjecture and its proof by Louis de Branges*, Amer. Math. Monthly **93** (1986), 505–514.

[Kzr] N. Kazarinoff, *Special functions and the Bieberbach conjecture*, Amer. Math. Monthly **95** (1988), 689–696.

[Lwn] K. Löwner, *Untersuchungen über schlichte konforme Abbildungen des Einheitskreises I*, Math. Ann. **89** (1923), 103–121.

[ME] *Matematicheskaja Encyklopedja I–V*, ed. I. M. Vinogradov, Moskva 1977.

[Mcr] T. M. MacRobert, *Functions of a Complex Variable*, MacMillan, London 1917.

[Mln 67] I. M. Milin, *On the cofficients of univalent functions*, Dokl. Akad. Nauk SSSR **176** (1967), 1015–1017; Soviet. Math. Dokl. **8** (1967), 1225–1258 (Russian).

[Mln 71] —, *Univalent functions and orthonormal systems*, Nauka, Moskva 1971 (Russian).

[Mrk] A. Markushevich, *Tieorja analiticheskich funkcij I, II*, Nauka, Moskva 1968.

[Pmm 75] Ch. Pommerenke, *Univalent Functions*, Vanderhoeck and Ruprecht, Göttingen 1975.

[Pmm 85] —, *The Bieberbach conjecture*, Math. Intelligencer **7** (1985), 23–25.

[R,P] A. K. Rathie and R. B. Paris, *A new proof of Watson's theorem for the series* ${}_3F_2$, Applied Mathematical Sciences **3** (2009), 161–164.

[Rbr] M. S. Robertson, *A remark on the odd schlicht functions*, Bull. Amer. Math. Soc. (1936), 366–370.

[Rnv] E. D. Rainville, *Special Functions*, Macmillan, New York 1960.

[Rss] B. Ross, *A brief history and exposition of the fundamental theory of fractional calculus*, in: Fractional Calculus and Its Applications (Springer Lecture Notes 457), Springer-Verlag, Heidelberg-New York 1975, pp. 1–36.

[Sch] I. Schur, *Vorlesungen ueber Invariantentheorie*, Springer-Verlag, Berlin-Göttingen-Heidelberg 1968.

[Skw 09a] M. Skwarczyński, *On the Bieberbach conjecture and Clausen identity. The role of Clausen identity*, Bull. Soc. Sci. Lettres Łódź **59** Ser. Rech. Deform. **58** (2009), 19–32.

[Skw 09b] —, *On the Bieberbach conjecture and Clausen identity. The role of Watson lemma*, Bull. Soc. Sci. Lettres Łódź **59** Ser. Rech. Deform. **59**, no. 1 (2009), 9–23.

[Slt] L. J. Slater, *Generalized Hypergeometric Functions*, Cambridge Univ. Press, Cambridge 1966.

[Smr] W. Smirnow, *Matematyka wyższa*, vol. 3, cz. 2, PWN, Warszawa 1965.

[Spn] D. C. Spencer, *Some problems in conformal mapping*, Bull. Amer. Math. Soc. **53** (1947), 417–439.

[Stn] P. K. Suetin, *Classical Orthogonal Polynomials*, Nauka, Moskva 1979 (Russian).

[Ths] W. H. Thomas II, *Introduction to real orthogonal polynomials*, Thesis, Naval Postgraduate School, Monterey, CA 1992.

[Whp 25] F. J. W. Whipple, *A group of generalized hypergeometric series...*, Proc. London Math. Soc. (2) **23** (1925), 104–114.

[Whp 29] —, *On a formula implied in Orr's theorems concerning the products of hypergeometric series*, Journal London Math. Soc. **4** (1929), 48–50

[Wts] G. N. Watson, *A note on generalized hypergeometric series*, Proc. London Math. Soc. (2) **23** (1925), xiii–xv.

Department of Mathematical Analysis
Cardinal Stefan Wyszyński University
Dewajtis 5, PL-01-815 Warszawa, Poland
e-mail: skwarczynski@uksw.edu.pl

Presented by Zbigniew Jakubowski at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on July 16, 2010

TWIERDZENIE DE BRANGES'A A UOGÓLNIONE FUNKCJE HIPERGEOMETRYCZNE III

PODSTAWOWE WŁASNOŚCI FUNKCJI DE BRANGES'A

S t r e s z c z e n i e

Spojrzenie na relację (II.37) wskazuje, że tożsamość Clausena stosuje się do ${}_3F_2$, gdy jeden z górnych parametrów jest średnią arytmetyczną pozostałych górnych parametrów i – co więcej – po pomnożeniu przez 2 lub dodaniu do 1/2 daje parametry dolne. Gdyby tożsamość Clausena dała się bezpośrendio zastosować do funkcji (1), nierówność $F(x) \geq 0$ byłaby oczywista. Niestety, sytuacja ta nie zachodzi. Pomimo to, nierówność tę można udowodnić poprzez przedstawienie F jako skończonej sumy wyrazów, do której tożsamość Clausena da się zastosować. Za Kazarinowem [Kzr 88] sprowadzamy zagdnienie do funkcji hipergeometrycznych ${}_2F_1$ używając operatora Rainville'a i wzoru Gegenbauera.

W ostatnim rozdziale kończymy dowód de Branges'a przez sprawdzenie warunku (II.**4**). Potrzebne warunki początkowe $\tau_k(0) = n - k + 1$ są wyprowadzone z klasycznego wyniku G. N. Watsona.

Subject index

Aharonov's proof of L-M inequality 62
Aronhold operator 79
Barnes notation ${}_AF_B$ 77
Bieberbach conjecture 49
Carathéodory convergence theorem 54
Cayley-Orr theorem 83
Clausen function ${}_3F_2$ 78
Clausen differential equation 80
Clausen identity 83
Clausen identity (proof using g.h.e.) 83
conditions (**1**)–(**4**) 73
conformal automorphism 49
de Branges differential system 72
de Branges functions τ_k 72
de Branges functional 72
derivative in a Löwner chain 70
distorsion lemma 49
expanding flow 61
FitzGerald-Pommerenke lemma 74
Fuchsian singularity at $z = 0$ 86
g.h. (generalized hypergeometric) 77
g.h.e. (g.h. equation) 79
g.h.f. (g.h. function) 77
g.h.f. form of τ_k' 44
Gauss function ${}_2F_1$ 78
Gauss differential equation 80
Gauss second summation theorem 95
Gauss first summation theorem 94
Gegenbauer formula 91
indicial equation 86
kernel convergence 53
Koebe 1/4-theorem 52
Koebe function 49
Korevaar's examples 76–77
L-M inequality 62
logarithmic coefficients 64
lower parameters 77
Löwner chain 58
Löwner equation 58
Milin functional 64
Milin conjecture 64
Milin inequalities 64
monotonicity property 66
Pochhammer symbol 71
Poisson kernel 59
Rainville representation 90
Rainville integral operator 90
Riemann mapping 56–57
Robertson inequalities 61
Robertson conjecture 61
schlicht (= univalent) function 48
Schwarz-Poisson formula 59
terminating series 77
understanding 96–97
unexpected change of notation 72
univalent (= schlicht) function 48
upper parameters 77
Watson summation lemma 93
Watson summation theorem 94
Watson theorem according to Rathie and Paris 95

Name index

Aharonov (Dov) . . . 62
Ahlfors . . . 49
Aronhold . . . 79
Askey . . . 77, 83
Bailey . . . 61, 96
Barnes . . . 77, 79, 96
Bateman . . . 94
Bergman . . . 48, 81, 98
Berndt . . . 97
Berry . . . 96
Bessel . . . 97
Bieberbach . . . 48, 49, 52, 61, 64, 65, 72, 98
Burchnall . . . 97
Carathéodory . . . 49, 52, 54, 77
Carslaw . . . 97
Cauchy . . . 50, 62
Cayley . . . 83, 96
Charzyński . . . 72, 98
Chaundy . . . 97
Chauvenet . . . 69
Chrystal . . . 97
Chu . . . 92
Churchill . . . 97
Clausen . . . 48, 69, 73, 78, 80, 83–87, 89, 93
de Branges 47, 48, 49, 56, 65, 66, 69–73, 76, . . . 77, 81, 89, 92–94, 97, 98
Erdélyi . . . 94
Conway . . . 48
Fasenmyer Mary Celine . . . 97
FitzGerald . . . 47, 69, 74, 76
Forsyth . . . 86, 96, 97
Fuchs . . . 48, 86
Gasper . . . 77, 83
Gauss . . . 69, 78, 80, 83
Gegenbauer . . . 48, 87, 89–91, 94–96
Gong . . . 48
Grinshpan . . . 47
Grunsky . . . 98
Hardy . . . 48, 96
Henrici . . . 48, 81
Hilbert . . . 72, 98
Hobson . . . 96, 97
Hurwitz . . . 55
Ince . . . 97
Jacobi . . . 77
Jakubowski . . . 66
Jollife . . . 97
Kazarinoff . . . 47, 48, 77,
Koebe . . . 49, 52, 65
Korevaar . . . 47, 69, 76
Ławrynowicz . . . 66
Lebedev . . . 62
Löwner . . . 48, 56, 58, 61, 66, 69, 70, 72, 74, . . . 77, 98
Macaulay . . . 96, 97
Milin . . . 62, 64
Orr . . . 71, 76, 83, 96–66
Paris . . . 48, 94, 95
Pochhammer . . . 71, 86, 90, 95, 96
Poisson . . . 59
Pólya . . . 98
Pommerenke . . . 47, 69, 74, 76
Posacka . . . 83
Rainville . . . 48, 89, 90, 93, 97
Ramanujan . . . 48, 96, 97
Rathie . . . 48, 94, 95
Rayleigh . . . 96
Riemann . . . 48, 49, 56, 57, 66, 98
Robertson . . . 61, 64, 65, 76
Ross . . . 90
Rouché . . . 55
Royden . . . 98
Sawyer . . . 78
Schiffer . . . 72, 98
Schmidt . . . 98
Schur (Issai) . . . 98
Schwarz . . . 59, 62
Slater . . . 94
Szegö . . . 98
Szpakowski . . . 66
Taylor . . . 70
Vandermonde . . . 92
Viète . . . 80
Watson . . . 48, 73, 89, 93–97
Whipple . . . 96, 97
Whittaker . . . 96, 97

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 105–118

Dedicated to Professor Roman Stanisław Ingarden
on the occasion of his ninetieth birthday

Janusz Garecki

TELEPARALLEL EQUIVALENT OF GENERAL RELATIVITY: A CRITICAL REVIEW

Summary

After reminder of some facts concerning general relativity (**GR**) we pass to *teleparallel gravity*. We are confining to the special model of the teleparallel gravity, which is popular recently, called *the teleparallel equivalent of general relativity* (**TEGR**). We are finishing with conclusion and some general remarks.

1. Introduction and standard formulation of GR

As it is known **GR** is a modern geometrical theory of gravity which simultaneously gives a mathematical model of the physical spacetime.

The mathematical model of the physical spacetime in **GR** is given by a pseudo-Riemannian differential manifold (Haussdorff, paracompact, connected, inextensible, orientable) (M_4, g_L). Here g_L means a Lorentzian metric which satisfies *Einstein equations*

$$G_\mu{}^\nu = \frac{8\pi G}{c^4} T_\mu{}^\nu \tag{1}$$

$(\alpha,\ \beta,\ \gamma, ..., \mu,\ \nu, ..., = 0, 1, 2, 3)$. We will identify geometrical objects with the sets of their components. Greek indices mean *coordinate components* of the geometrical objects.

So, g_L, is a *dynamical object*.

Here G_μ^ν is the so-called *Einstein tensor*, $T_\mu{}^\nu$ is the *matter energy-momentum tensor* (the source of the gravitational field), c is the velocity of light in vacuum, and G means Newtonian gravitational constant.

The mathematical model of the physical spacetime in **GR** is originated from *Einstein Equivalence Principle* (**EEP**) [1]. The main ingredient of this Principle is *universality of the free falls* of the test bodies in a given gravitational field.

GR reduces the gravitational interactions to some geometric aspects of the spacetime. Namely, we have:

1. g_L = gravitational potentials,
2. $\left\{{}^{\alpha}_{\beta\ \gamma}\right\}$ = gravitational strengths, and
3. $R^{\alpha}{}_{\beta\gamma\delta}(\{\})$ = strengths of the gravitational tidal forces.

The symmetry group of the **GR** is the infinite group **Diff M$_4$**.

The Levi-Civita connection $\left\{{}^{\alpha}_{\beta\ \gamma}\right\}$ is symmetric, metric and torsion-free.

Usually one uses in **GR** a maximal atlas of the local charts (local maps, coordinate patches) and *implicite* coordinate frames (natural frames, holonomic frames) and coframes $\big(\{\partial_\mu\},\ \{dx^\alpha\}\big)$ and coordinate components of the geometrical objects.

Every coordinate transformation

$$x^{\alpha'} = x^{\alpha'}(x^\beta), \quad \det\Big[\frac{\partial x^{\alpha'}}{\partial x^\beta}\Big] \neq 0 \tag{2}$$

changes coordinate frames and coframes, and coordinate components of the geometrical objects in standard way.

In the introductory relativity textbooks [2] one usually says about coordinate transformations and about transformations of the coordinate components of the geometrical objects. In fact, it is sufficient. Also some conservative specialist on tensor analysis follow this way [3]. But one can use in **GR** (and in tensor calculus also) arbitrary frames, especially *non-holonomic* (or anholonomic) frames and coframes $(\{h_a{}^\mu\ (x)\},\ \{h^b{}_\alpha(x)\}) :\ h_a{}^\mu(x)h^b{}_\mu(x) = \delta^b_a,\ (a, b, c, d, ..., = 0, 1, 2, 3)$. Latin indices (= anholonomic indices) numerate vectors and covectors.

The anholonomic frames and coframes *are not connected with local coordinates*, e.g., they are neutral under coordinate transformations. Instead of we have

$$\partial_\alpha = h^b{}_\alpha(x)\partial_b, \quad dx^\alpha = h_a{}^\alpha(x)dx^a, \tag{3}$$

or, equivalently,

$$\vec{e}_a := \partial_a = h_a{}^\beta(x)\partial_\beta, \quad \vartheta^b := dx^b = h^b{}_\mu(x)dx^\mu. \tag{4}$$

Here $(x) := \{x^\alpha\}$ are *spacetime coordinates*, and $\{x^a\}$ mean *tangent space coordinates*. In **GR** every tangent space is endowed with Minkowski structure.

For coordinate frames and coframes one has

$$\vec{e}_a = \delta^\beta_a\partial_\beta, \quad \vartheta^b = \delta^b_\mu dx^\mu. \tag{5}$$

Some remarks are in order:

1. $\{\vec{e}_a(x)\} \equiv \{\partial_a(x)\}$ is a coordinate frame in tangent space $T_x(M_4, g_L)$, and $\{\vartheta^b\} \equiv \{dx^b\}$ is a coordinate coframe in the dual space space $T^*_x(M_4, g_L)$.

Differential forms $\vartheta^b = dx^b = h^b{}_\mu(x)dx^\mu$ *are not integrable* for anholonomic frames $\{h^b{}_\mu(x)\} : d\vartheta^b \neq 0$.

2. Henceforth we will consequently use an old tensorial terminology of J. A. Schouten, and S. Gołąb, i.e., we will call $\{h_a{}^\beta(x)\}$ "frame" instead of $\{\vec{e}_a(x)\}$, and $\{h^b{}_\mu(x)\}$ "coframe" instead of $\{\vartheta^b\}$. It will useful in passing to teleparallel gravity because majority of the authors working in this field uses this terminology.

3. We permanently use standard Einstein summation convention.

As we see, anholonomic frames and coframes in our terminology *connect* the partial derivatives ∂_α and ∂_b, and differentials dx^α with dx^a. They also connect anholonomic components of the geometrical objects (denoted by Latin indices) with their coordinate components (denoted by Greek indices). Namely, one has (coordinates $\{x^\mu\}$ are fixed) for a tensor field of the type (r,s)

(6) $$T^{a_1\ldots a_r}{}_{b_1\ldots b_s}(x) = h^{a_1}{}_{\mu_1}(x)\ldots h^{a_r}{}_{\mu_r}(x)h_{b_1}{}^{\nu_1}(x)\ldots h_{b_s}{}^{\nu_s}(x)T^{\mu_1\ldots\mu_r}{}_{\nu_1\ldots\nu_s}(x),$$

and, conversely

(7) $$T^{\mu_1\ldots\mu_r}{}_{\nu_1\ldots\nu_s}(x) = h_{a_1}{}^{\mu_1}(x)\ldots h_{a_r}{}^{\mu_r}(x)h^{b_1}{}_{\nu_1}(x)\ldots h^{b_s}{}_{\nu_s}(x)T^{a_1\ldots a_r}{}_{b_1\ldots b_s}(x).$$

For a linear and metric connection ω one obtains. From here we confine to anholonomic tetrads and cotetrads (see below).

(8) $$\omega^a{}_{bc}(x) = h_c{}^\nu(x)\omega^a{}_{b\nu}(x),$$

where

(9) $$\omega^a{}_{b\nu}(x) = h^a{}_\lambda(x)\Gamma^\lambda{}_{\mu\nu}(x)h_b{}^\mu(x) + h^a{}_\rho(x)\partial_\nu h_b{}^\rho(x)$$

is so-called *spin connection*. Conversely, we have

(10) $$\Gamma^\rho{}_{\mu\nu}(x) = h_a{}^\rho(x)h^b{}_\mu(x)\omega^a{}_{b\nu}(x) + h_a{}^\rho(x)\partial_\nu h^a{}_\mu(x).$$

In **GR** one usually uses the anholonomic frames $\{h_a{}^\mu(x)\}$ and dual coframes $\{h^b{}_\mu(x)\}$ which form the so-called *orthonormal tetrad* and *cotetrad* fields. These fields are defined as follows

(11) $$h^a{}_\mu(x)h^b{}_\nu(x)\eta_{ab} = g_{\mu\nu}(x),$$

or, equivalently

(12) $$h_a{}^\mu(x)h_b{}^\nu(x)g_{\mu\nu}(x) = \eta_{ab}.$$

Here $\eta_{ab} = \mathrm{diag}(1,-1,-1,-1)$ is the Minkowski metric of the tangent spaces $T_x(M_4, g_L)$ and $g_{\mu\nu}(x)$ means the spacetime metric g_L.

The transformations of the spacetimes coordinates act only on spacetime indices (Greek indices) in standard way, whereas on the tangent space indices (Latin indices) act only *local or global Lorentz transformations*, e.g.,

(13) $$h'^a{}_\mu = \Lambda^a{}_b(x)h^b{}_\mu(x),$$

where

$$\Lambda^a{}_b(x)\eta_{ac}\Lambda^c{}_d(x) = \eta_{bd}. \tag{14}$$

For a global Lorentz transformation one has $\Lambda^a{}_b = const.$

Tetrads are not uniquely determined by the given spacetime metric $g_{\mu\nu}(x)$ but only up to local Lorentz transformations, i.e., up to six arbitrary functions. It is because a metric has only ten independent components and a tetrad field has sixteen independent components. So, for a given metric $g_{\mu\nu}(x)$ there exists ∞^6 different classes of tetrad fields $\{h_a{}^\mu(x)\}$ which satisfy (11–12). One class of the tetrad $[\{h_a{}^\mu(x)\}]$ means these tetrads which are connected by a global Lorentz transformation.

Contrary, given tetrad field $\{h_a{}^\mu(x)\}$ *determines unique metric*

$$g_{\mu\nu}(x) = h^a{}_\mu(x)h^b{}_\nu(x)\eta_{ab}, \tag{15}$$

where

$$h^a{}_\mu(x)h_b{}^\mu(x) = \delta^a_b. \tag{16}$$

In **GR** fundamental role plays the spacetime metric $g_{\mu\nu}(x)$ (it is an observable), whereas the orthonormal tetrads (they are not observables) play only an auxiliary role: they simplify calculations and they enable us to introduce spinors into spacetime structure.

The physical foundations and standard formulation of the **GR** have very good observational evidence. Observational consequences of the Einstein equations were confirmed up to $0,003\%$ in Solar System (weak gravitational field), and up to $0,05\%$ in binary pulsars (strong gravitational field). We mean here **EEP**, Einstein equations and mathematical model (M_4, g_L) of the physical spacetime. Universality of the free falls was confirmed up to 10^{-14} and some other consequences of the **EEP** were confirmed up to 10^{-23} (see, e.g., [1].).

So, up to now, we *do not need to modify or generalize* **GR** (Ockham razor).

We would like to emphasize that *we have no free parameter* in **GR**.

Fascinating is that despite this the theory has passed all the stringent tests with favour. In the proposed *generalized gravity theories* one has many free parameters, e.g., one has 28 free parameters in metric-affine gravity. These parameters can be adjusted in order to have agreement with experience.

2. Teleparallel gravity

This is a gravity with *an absolute parallelism*, i.e., with *curve independent parallelism* of distant vectors and tensors.

In this old approach (since 1928; renewed recently) the mathematical model of the physical spacetime is based on *Weitzenböck geometry* (= teleparallel geometry or geometry with absolute parallelism).

The geometry of such a kind *is uniquely determined* by the given tetrad field $\{h_a{}^\mu\}(x)$. Namely, one has (coordinates $\{x^\alpha\}$ are fixed):

1. Metric $g_{\mu\nu}(x) := h^a{}_\mu(x) h^b{}_\nu(x)\eta_{ab}$.

2. Teleparellel connection (Weitzenböck's connection) $\Gamma^\rho{}_{\mu\nu} := h_a{}^\rho(x)\partial_\nu h^a{}_\mu(x)$.

Here $h_a{}^\mu(x) h^b{}_\mu(x) = \delta^b_a$.

The teleparallel Weitzenböck connection has non-vanishing torsion

$$T^\rho{}_{\mu\nu} := \Gamma^\rho{}_{\nu\mu} - \Gamma^\rho{}_{\mu\nu}$$

iff the tetrads $\{h_a{}^\mu(x)\}$ are anholonomic, and it has identically vanishing curvature $R^\rho{}_{\theta\mu\nu}(\Gamma)$, where

$$R^\rho{}_{\theta\mu\nu}(\Gamma) := \partial_\mu \Gamma^\rho{}_{\theta\nu} - \partial_\nu \Gamma^\rho{}_{\theta\mu} + \Gamma^\rho{}_{\sigma\mu}\Gamma^\sigma{}_{\theta\nu} - \Gamma^\rho{}_{\sigma\nu}\Gamma^\sigma{}_{\theta\mu}. \tag{17}$$

Important remarks are in order:

1. Weitzenböck connection *is metric*, i.e.,

   $$\nabla_\rho g_{\mu\nu} := \partial_\rho g_{\mu\nu} - \Gamma^\alpha{}_{\mu\rho} g_{\alpha\nu} - \Gamma^\alpha{}_{\nu\rho} g_{\mu\alpha} \equiv 0. \tag{18}$$

   But the other possible covariant derivative

   $$\tilde{\nabla} g_{\mu\nu}(x) := \partial_\rho g_{\mu\nu} - \Gamma^\alpha{}_{\rho\mu} g_{\alpha\nu} - \Gamma^\alpha{}_{\rho\nu} g_{\mu\alpha}, \tag{19}$$

   is different from zero because Weitzenböck connection is not symmetric.

2. Torsion of the Weitzenböck connection *is entirely determined* by the Schouten-Van Danzig *anholonomy object* $\Omega^a{}_{bc}(x)$, where

   $$\Omega^a{}_{bc}(x) := h_b{}^\beta(x) h_c{}^\gamma(x)\big[\partial_\gamma h^a{}_\beta(x) - \partial_\beta h^a{}_\gamma(x)\big]. \tag{20}$$

   The anholonomity object measures anholonomy of the used tetrad field: for a holonomic tetrads $\{h_a{}^\mu(x)\}$ one has $\Omega^a{}_{bc}(x) \equiv 0$. Namely, we have

   $$T^\rho{}_{\mu\nu}(x) = h_a{}^\rho(x) h^b{}_\mu(x) h^c{}_\nu(x)\Omega^a{}_{bc}(x). \tag{21}$$

3. One has the following relation between the components of the Weitzenböck connection $\Gamma^\rho{}_{\mu\nu}(x)$ and between the components $\{{}^\rho{}_{\mu\nu}\}(x)$ of the Levi-Civita connection for the metric $g_{\mu\nu}(x)$

   $$\Gamma^\rho{}_{\mu\nu}(x) = \{{}^\rho{}_{\mu\nu}\}(x) + K^\rho{}_{\mu\nu}(x), \tag{22}$$

   where

   $$K^\rho{}_{\mu\nu}(x) := \frac{1}{2}\big(T_\mu{}^\rho{}_\nu + T_\nu{}^\rho{}_\mu - T^\rho{}_{\mu\nu}\big) \tag{23}$$

   is the *contortion tensor.*

4. For Weitzenböck connection $\Gamma^\rho{}_{\mu\nu}(x)$

   $$\omega^a{}_{b\nu}(x) \equiv 0 \Rightarrow \omega^a{}_{bc} \equiv 0, \tag{24}$$

   i.e., this *connection identically vanishes* in the tetrads $\{h_a{}^\mu(x)\}$ which have determined it.

Greek, i.e., holonomic indices are raised and lowered with the spacetime metric $g_{\mu\nu}$ and the Latin, i.e., anholonomic indices, are raised and lowered with the Minkowski metric η_{ab}.

The class of the tetrads $\left[\{h_a{}^{\mu}(x)\}\right]$ connected by global Lorentz transformations with $\Lambda^a{}_b = const$ determines the same Weitzenböck connection and geometry. On the other hand, the any two tetrad fields $\{h'^a{}_{\mu}(x)\}$, $\{h^a{}_{\mu}(x)\}$ which are connected by a local Lorentz transformation

$$h'^a{}_{\mu}(x) = \Lambda^a{}_b(x) h^b{}_{\mu}(x) \tag{25}$$

determine two different Weitzenböck connections, $\bar{\Gamma}^{\rho}{}_{\mu\nu}(x)$ and $\Gamma^{\rho}{}_{\mu\nu}(x)$ and *two different Weitzenböck geometries.*

So, the set of the all tetrads $\left(\{h_a{}^{\mu}(x)\}\right)$ splits onto *disjoint classes* (∞^6 classes) which determine *different Weitzenböck connections and geometries.* ∞^6 classes because the local Lorentz transformations depend on six arbitrary functions.

In consequence, the symmetry group of a teleparallel gravity consists of the group **Diff $\mathbf{M_4}$** and the global Lorentz group.

In the following we will confine to the very special case of the teleparallel gravity, namely we will confine to the so-called *teleparallel equivalent of general relativity* (**TEGR**).

The **TEGR** is a recent approach to teleparallel gravity which is mainly developed by mathematicians and physicists from Brasil (see, e.g., [4]).

One can look on **TEGR** as a new trial to rescue torsion in theory of gravity because, up to now, *no experiment confirmed the Riemann-Cartán torsion.* The Riemann-Cartan torsion is the torsion in the Riemann-Cartán geometry. This generalized metric geometry endowed with curvature and torsion was proposed by many authors since 1970 [5] as a geometric model of the physical spacetime. In our opinion lack of experimental evidence, many ambiguities to whose torsion leads, topological triviality of torsion and *Ockham razor* rather disqualify this model [6]..

The details of the standard approach to **TEGR** read.

One starts with the given metric $g_{\mu\nu}(x)$. This metric determines (up to local Lorentz transformations) the anholonomic tetrad $\{h_a{}^{\mu}(x)\}$ and dual cotetrad $\{h^a{}_{\mu}(x)\}$ fields, which satisfy

$$h^a{}_{\mu}(x) h^b{}_{\nu}(x) \eta_{ab} = g_{\mu\nu}(x), \tag{26}$$

$$h^a{}_{\mu}(x) h_b{}^{\mu}(x) = \delta^a_b. \tag{27}$$

Then, these fields determine the Weitzenböck connection

$$\Gamma^{\rho}{}_{\mu\nu}(x) = h_a{}^{\rho}(x) \partial_{\nu} h^a{}_{\mu}(x), \tag{28}$$

which satisfies

$$\left\{{}^{\rho}{}_{\mu\nu}\right\}(x) = \Gamma^{\rho}{}_{\mu\nu}(x) - K^{\rho}{}_{\mu\nu}(x). \tag{29}$$

Here $\left\{{}^{\rho}{}_{\mu\nu}\right\}(x)$ is the Levi-Civita connection for the metric $g_{\mu\nu}(x)$.

For the Weitzenböck connection $\Gamma^{\rho}{}_{\mu\nu}(x)$ one has

(30) $$R^{\rho}{}_{\theta\mu\nu}(\Gamma) \equiv R^{\rho}{}_{\theta\mu\nu}(\{\}) + Q^{\rho}{}_{\theta\mu\nu} \equiv 0.$$

Here

(31) $$R^{\rho}{}_{\theta\mu\nu}(\Gamma) := \partial_\mu \Gamma^{\rho}{}_{\theta\nu} - \partial_\nu \Gamma^{\rho}{}_{\theta\mu} + \Gamma^{\rho}{}_{\sigma\mu}\Gamma^{\sigma}{}_{\theta\nu} - \Gamma^{\rho}{}_{\sigma\nu}\Gamma^{\sigma}{}_{\theta\mu},$$

(32) $$R^{\rho}{}_{\theta\mu\nu}(\{\}) := \partial_\mu \{^{\rho}{}_{\theta\nu}\} - \partial_\nu \{^{\rho}{}_{\theta\mu}\} + \{^{\rho}{}_{\sigma\mu}\}\{^{\sigma}{}_{\theta\nu}\} - \{^{\rho}{}_{\sigma\nu}\}\{^{\sigma}{}_{\theta\mu}\},$$

and

(33) $$Q^{\rho}{}_{\theta\mu\nu} := D_\mu K^{\rho}{}_{\theta\nu} - D_\nu K^{\rho}{}_{\theta\ \mu} + K^{\rho}{}_{\sigma\mu} K^{\sigma}{}_{\theta\nu} - K^{\rho}{}_{\sigma\nu} K^{\sigma}{}_{\theta\mu}.$$

D_μ is the Levi-Civita covariant derivative expressed in terms of the Weitzenböck connection, i.e.,

(34) $$D_\rho v^\mu := \partial_\rho v^\mu + \big(\Gamma^{\mu}{}_{\lambda\rho} - K^{\mu}{}_{\lambda\rho}\big) v^\lambda.$$

$R^{\rho}{}_{\theta\mu\nu}(\Gamma)$ is the *main* curvature tensor of the Weitzenböck geometry. Main curvature tensor because one can consider other curvatures in Weitzenböck geometry, e.g., Riemannian curvature [7].

The Authors which work on **TEGR**, by use the fundamental formulas (26, 29, 30) of the Weitzenböck geometry, *rephrase*, step by step, all the formalism of the purely metric **GR** in terms of the Weitzenböck connection $\Gamma^{\rho}{}_{\mu\nu}(x)$ and its torsion $T^{\rho}{}_{\mu\nu}(x)$ (mainly in terms of torsion).

For example:

1. The Einstein Lagrangian for **GR**

   (35) $$L_E = (-)\alpha \sqrt{|g|} R(\{\}) + \partial_\mu w^\mu,$$

   where $g := \det\big[g_{\mu\nu}\big]$, and

   (36) $$w^\mu := \alpha \sqrt{|g|}\big(g^{\alpha\beta}\{^{\mu}{}_{\alpha\beta}\} + g^{\alpha\mu}\{^{\gamma}{}_{\alpha\gamma}\}\big)$$

   is rephrased to the form

   (37) $$\alpha h S^{\rho\mu\nu} T_{\rho\mu\nu} =: L_{TEGR},$$

   where $h = \det\big[h^{a}{}_{\mu}\big] = \sqrt{|g|}$. One obtains in fact ∞^6 different L_{TEGR} because L_{TEGR}, like L_E is invariant only under global Lorentz group. Despite that the field equations (39, 40) are locally Lorentz invariant. We could get localy Lorentz invariant L_{TEGR} if we rephrased $L = (-)\alpha\sqrt{|g|}R(\{\})$ and

   (38) $$S^{\rho\mu\nu} = (-)S^{\rho\nu\mu} := \frac{1}{2}\big[K^{\mu\nu\rho} - g^{\rho\nu} T^{\alpha\mu}{}_{\alpha} + g^{\rho\mu} T^{\alpha\nu}{}_{\alpha}\big].$$

2. The vacuum Einstein equations

   (39) $$\big[R_{\lambda}{}^{\rho}(\{\}) - \frac{1}{2}\delta_{\lambda}{}^{\rho} R(\{\})\big]\sqrt{|g|} = 0$$

   are rephrased to the form

   (40) $$\partial_\sigma\big(h S_{\lambda}{}^{\sigma\rho}\big) - 4\alpha^{(-1)}\big(h t_{\lambda}{}^{\rho}\big) = 0,$$

where

$$t_\lambda^{\ \rho} = h^a{}_\lambda J_a{}^\rho + 4\alpha \Gamma^\mu{}_{\lambda\nu} S_\mu{}^{\nu\rho}, \tag{41}$$

and

$$J_a{}^\rho = (-)4\alpha h_a{}^\lambda S_\mu{}^{\nu\rho} T^\mu{}_{\nu\lambda} + 4\alpha h_a{}^\rho S^{\alpha\beta\gamma} T_{\alpha\beta\gamma}, \tag{42}$$

and so on,

$\alpha := c^4/16\pi G$.

Then, these authors call the obtained formal reformulation of **GR** in terms of the Weitzenböck geometry *the teleparallel equivalent of the general relativity* (**TEGR**) and conclude: "Gravitational interaction can be described *alternatively* in terms of curvature, as it is usually done in **GR**, or in terms of torsion, in which case we have the so-called teleparallel gravity. *Whether gravitation requires a curved or torsional spacetime, therefore, turns out to be a matter of convention*". They also assert that **TEGR** "is better than the original **GR**" because, e.g., "in **TEGR** one can separate gravity from inertia (on the connection level) and this separation reads"

$$\left\{{}^{\alpha}_{\ \beta\gamma}\right\} = \Gamma^\alpha{}_{\beta\gamma} - K^\alpha{}_{\beta\gamma}. \tag{43}$$

Following the authors which work on **TEGR**, the left hand side term of the above "separation formula", ($\left\{{}^{\alpha}_{\ \beta\gamma}\right\}$), represents gravity and inertia and the right hand side terms describe inertia, ($\Gamma^\alpha{}_{\beta\gamma}$), and gravitation, ($K^\alpha{}_{\beta\gamma}$), respectively.

Of course, such separation *contradicts* **EEP** and *is impossible* in standard formulation of the **GR**.

We *cannot agree with such statements*. In our opinion, the "teleparallel equivalent of **GR**" (What kind of equivalence?) is only *formal and geometrically trivial*, non-unique (see below) rephrase of **GR** in terms of the Weitzenböck geometry. Such rephrase is, of course, *always possible* not only with **GR** but also with any other purely metric theory of gravity.

In our opinion, *we have no profound physical motivation* for expression of the gravitational interaction in terms of the teleparallel torsion because the Weitzenböck torsion *is entirely expressed in terms of the Van Danzig and Schouten anholonomity object* $\Omega^a{}_{bc}(x)$. So, the torsion of the teleparallel Weitzenböck connection *describes only anholonomy* of the used tetrad field and, therefore, *it is not connected neither with the real geometry of the physical spacetime nor with real gravity*, e.g., one can introduce Weitzenböck torsion already in flat Minkowski spacetime.

Weitzenböck torsion could only describe the inertial forces in the framework of the special relativity. In special relativity anholonomic tetrads really represent non-inertial frames..

Contrary, the Levi-Civita part of the Weitzenböck connection, *as independent of tetrads*, can have *and surely has* the physical and geometrical meaning. The Levi Civita connection depends only on metric. It is independent of the tetrads which determine the same spacetime metric.

Further ctitical remarks on **TEGR**.

1. **TEGR** is nothing new. In fact, it is exactly the old tetrad formulation of **GR** given in the very distant past by C. Möller [8] but expressed in terms of anholonomy of the tetrads instead of in terms of tetrads exclusively (As it was in Möller papers). For example, despite that the **TEGR** field equations are expressed in terms of torsion of the Weitzenböck geometry, they form the system of the 10 partial differential equations of the 2^{nd} order on 16 tetrads components, like the 10 field equations of the Möller's tetrad formulation of **GR**. Solving the **TEGR** equations in vacuum (or in matter) we are looking for the tetrad components $\{h_a{}^{\mu}(x)\}$ for *apriori given general form of the metric* $g_{\mu\nu}(x)$; not for the components of torsion. Weitzenböck connection and its torsion are calculated later [9].

   Therefore, the notation of the Lagrangian and the field equations of **TEGR** in terms of Weitzenböck torsion *is only a camouflage*: **TEGR** is simply the Möller's tetrad formulation of **GR**, and, like Möller's formulation of **GR**, determines uniquely the metric only.

   We would like to emphasize that one can find all the results of the **TEGR** including the **TEGR** energy-momentum tensor for pure gravity in the old Möller's papers. This 'tensor" is one of the most important results obtained in the framework of **TEGR**.

2. **TEGR** *is not unique*. This follows from the fact: given metric, $g_{\mu\nu}(x)$ has 10 intrinsic components and determines only 10 components of the tetrads field $\{h_a{}^{\mu}(x)\}$ which has 16 intrinsic components. It is a consequence of the known fact that a given metric *determines tetrad field up to local Lorentz transformations*, which form the local, six-parameters, ortochronous Lorentz group $L_+^{\uparrow}$ defined as follows

$$\begin{aligned} L_+^{\uparrow} &= \big\{\Lambda^a{}_b(x) : \Lambda^a{}_b(x)\eta_{ac}\Lambda^c{}_d(x) = \eta_{bd}, \\ \det\big[\Lambda^a{}_b(x)\big] &= 1, \;\; \Lambda^0{}_0 \geq 1\big\}. \end{aligned} \tag{44}$$

   The ten field equations of **GR** (or **TEGR**) determine the metric and also determine only ten components of the tetrad field. The remaining six components are lefting arbitrary functions of the spacetime coordinates $\{x^{\alpha}\}$ and can be arbitrarily established. It is a consequence of the local Lorentz invariance of the **TEGR** and **GR** field equations.

   So, for the given metric, $g_{\mu\nu}(x)$, (**GR**) there exist ∞^6 different classes of tetrad fields (**TEGR**) and, in consequence, ∞^6, different Weitzenböck connections $\Gamma^{\rho}{}_{\mu\nu}(x)$ (and geometries). Each of these connections satisfies the equations

$$\left\{{}^{\rho}{}_{\mu\nu}\right\}(x) = \Gamma^{\rho}{}_{\mu\nu}(x) - K^{\rho}{}_{\mu\nu}(x). \tag{45}$$

   In the above equations the left hand side *is independent of tetrads*; it depends only on metric $g_{\mu\nu}(x)$, whereas the both terms on the right hand side *depend on the class of the tetrads*. (One) class of tetrads := the set of tetrads $\big[\{h_a{}^{\mu}(x)\}\big]$

which are connected by global Lorentz transformations. Class of tetrads determines the same Weitzenböck connection and geometry. Different classes of tetrads are connected by local Lorentz transformations and determine different Weitzenböck connections and geometries.

As a result we obtain ∞^6 different Lagrangians (37) for **TEGR** and ∞^6 different **TEGR**. This fact was already known C. Möller in context of his tetrad formulation of **GR**. Namely, Möller, in fact, also has obtained ∞^6 different tetrad formulations of **GR** because, the 10 field equations of his tetrad formulation of **GR**, identical with Einstein equations (1), determine the tetrad field up to local Lorentz transformations, i.e., up to six arbitrary functions. These field equations determine the metric only. The same situation we have of course in the framework of the **TEGR** because the 10 field equations (40), like Möller's equations, are locally Lorentz invariant. In order to have field equations which would determine tetrad field completely (apart from constant Lorentz rotations) Möller has developed *tetrad theory of gravity* in which one has sixteen field equations onto sixteen tetrad components.

3. The authors which work on **TEGR** assert that the formula (43) (or (45)) gives *separation* of inertia $\left(\Gamma^{\rho}{}_{\mu\nu}(x)\right)$ from gravity $\left(K^{\rho}{}_{\mu\nu}(x)\right)$.

   Such speculative separation allows them, among other things, to introduce an energy-momentum tensor for gravity. It is in fact a family of ∞^6 different tensors the same as the family of the tensors which has been obtained many years ago by C. Möller *without any separation* in his tetrad formulation of **GR**. But this separation *is illusoric* because there exist ∞^6 different separations of the form (43) (or (45)) for given $\left\{{}^{\alpha}_{\beta\gamma}\right\}$, i.e., we *have no separation inertia from gravity* in **TEGR** (in agreement with **EEP**).

   In consequence, *we have no unique gravitational energy-momentum tensor* in **TEGR**.

4. The experts on **TEGR** transform trivially the geodesic equations of **GR**

$$\frac{d^2x^\alpha}{ds^2} + \left\{{}^{\alpha}_{\beta\gamma}\right\}\frac{dx^\beta}{ds}\frac{dx^\gamma}{ds} = 0 \tag{46}$$

   onto the *forces equations*

$$\frac{d^2x^\alpha}{ds^2} + \Gamma^{\alpha}{}_{\beta\gamma}\frac{dx^\beta}{ds}\frac{dx^\gamma}{ds} = K^{\alpha}{}_{\beta\gamma}\frac{dx^\beta}{ds}\frac{dx^\gamma}{ds} \tag{47}$$

   by putting in (46):

$$\left\{{}^{\alpha}_{\beta\gamma}\right\} = \Gamma^{\alpha}{}_{\beta\gamma} - K^{\alpha}{}_{\beta\gamma}. \tag{48}$$

   The *forces equations* (47) remind the **GR** equations of motion for a charged test particle when the both fields, electromagnetic and gravitational, simultaneously act on the particle

$$\frac{d^2x^\alpha}{ds^2} + \left\{{}^{\alpha}_{\beta\gamma}\right\} \frac{dx^\beta}{ds}\frac{dx^\gamma}{ds} = \frac{Q}{m} F^\alpha{}_\beta \frac{dx^\beta}{ds}. \tag{49}$$

Here Q, m denote electric charge and mass of the particle respectively and $F^\alpha{}_\beta$ mean electromagnetic field acting on the particle. The right hand side of (49) is the electromagnetic force per unit mass which acts on the particle.

The specialists on **TEGR** try to attach some physical meaning to the *force equations* (47), namely following them, the right hand side of (47) describes *gravitational force* acting on the particle, whereas the term $\Gamma^\alpha{}_{\beta\gamma}(dx^\beta/ds)\,(dx^\gamma/ds)$ describes *inertial force*.

But there exist ∞^6 different reformulatios of the geodesic equations (46) to the form (47) with different $\Gamma^\alpha{}_{\beta\gamma}$ and $K^\alpha{}_{\beta\gamma}$. Which one of them is correct, i.e., *which one of them gives correct inertial force and correct gravitational force*?

Talking about *equivalence* of **TEGR** with **GR** *is misleading* because there exist ∞^6 different **TEGR** in consequence of the local Lorentz invariance of the field equations (40). But we must emphasize that every **TEGR** determines unique and the same metric structure of the spacetime as **GR** does. So, from the metric point of view, the different **TEGR** are equivalent.

Here we have the same kind of *"equivalence"* as the *"equivalence"* between a given metric $g_{\mu\nu}(x)$ (10 functions) and a tetrad field (16 functions), which satisfies

$$h^a{}_\mu(x) h^b{}_\nu(x)\eta_{ab} = g_{\mu\nu}(x)$$

i.e., *we have no equivalence*. Remark also that metric and tetrads are different geometric objects.

Incorrect is also statement of the specialists on **TEGR** that Weitzenböck geometry is flat, like Minkowski geometry. In fact, e.g., *Riemannian curvature of such geometry is non-zero*. Also the curvature tensor $\tilde{R}^\alpha{}_{\beta\gamma\delta}(\Gamma)$ where

$$\tilde{R}^\alpha{}_{\beta\gamma\delta}(\Gamma) := \partial_\gamma \Gamma^\alpha{}_{\delta\beta} - \partial_\delta \Gamma^\alpha{}_{\gamma\beta} + \Gamma^\alpha{}_{\gamma\sigma}\Gamma^\sigma{}_{\delta\beta} - \Gamma^\alpha{}_{\delta\sigma}\Gamma^\sigma{}_{\gamma\beta} \tag{50}$$

is different from zero.

The tensor $\tilde{R}^\alpha{}_{\beta\gamma\delta}(\Gamma)$ differs from the former *main curvature tensor* $R^\alpha{}_{\beta\gamma\delta}(\Gamma)$ (see the formula (31)) by transposition lower indices in $\Gamma^\alpha{}_{\beta\gamma}(x)$. For Riemannian geometry, owing to symmetry of the Levi-Civita connection, these both tensors are identically equal.

Resuming, in our opinion, **TEGR** is nothing new. It is camouflaged, the very old tetrad formulation of **GR** given by C. Möller, and it, *by no means is better* than standard **GR**. Contrary, standard **GR** *is surely better* than any **TEGR** because **GR** is invariant under any change of tetrads, whereas **TEGR** is not. **TEGR**, like any teleparallel gravity, is invariant only under global Lorentz rotations of tetrads.

We will finish with some general remarks about teleparallel gravity.

It should be emphasized that there exist many other approaches to teleparellel gravity, different from **TEGR**, and which generalize **GR**. At the first time such approach to gravity was considered already by A. Einstein ("Fernparallelismus" in 1928 [10]) and then by C. Möller (1978), Pellegrini and Plebański [11], Hayashi and Shirafuji [12], and others. Recently the teleparallel approach to gravity is developed by F. B. Estabrook, Y. Itin, and L. Schücking [13].

In these other approaches to teleparallel gravity the gravitational Lagrangian is built from irreducible torsion componets or from tetrads immediately, and contains, in general, three free parameters to be determined by experiments. This Lagrangian is invariant under **Diff $\mathbf{M_4}$** and has also global Lorentz symmetry.

The fundamental geometric object are tetrads which determine spacetime metric and Weitzenböck connection, and, therefore, all the local Weitzenböck geometry of the physical spacetime.

In vacuum, we have in these approaches sixteen 2^{nd} order field equations on sixteen tetrad components. The field equations should determine the tetrads field $h_a{}^{\mu}(x)$ up to constant Lorentz rotations, i.e., up to global Lorentz group, and owing that, should determine a unique Weitzenböck geometry. But tetrads *are not observables*: they are very alike to the electromagnetic potentials. Moreover, there are problems with physical interpretation of the six additional tetrads components (10 components can describe gravitational field, but what about remaining 6 components?) and these theories suffer from badly posed Cauchy problem [14].

Acknowledgments

The Author would like to thank Professor Julian Ławrynowicz for possibility to deliver this lecture during the *Hypercomplex Seminar 2010* dedicated to Professor Roman S. Ingarden on the occasion of His 90th birthday, and the Mathematical Institute of the University of Szczecin for financial support (grant 503-4000-230351). Author also thanks Professor Friedrich W. Hehl for the most useful critical remarks.

References

[1] C. M. Will, *Theory and Experiment in Gravitational Physics*, Cambridge University Press, Cambridge 1993; *The Confrontation between General Relativity and Experiment*, arXiv:gr-qc/0510072.

[2] B. F. Schutz, *A First Course in General Relativity*, Cambridge University Press, Cambridge 1985 (Polish edition: B. F. Schutz, *Wstęp do ogólnej teorii względności*, Wydawnictwo Naukowe PWN, Warszawa 2002); L. D. Landau, E. M. Lifshitz, *The Classical Theory of Fields*, 4th edition, Oxford 2002 (Polish edition: L. D. Landau, E. M. Lifszyc, *Teoria pola*, Wydawnictwo Naukowe PWN, Warszawa 2009); J. Foster, J. D. Nightingale, *A Short Course in General Relativity*, Longman, London and New York 1979 (Polish edition: J. Foster, J. D. Nightingale, *Ogólna teoria względności*, PWN, Warszawa 1985); J. Plebański and A. Krasiński, *An Introduction to General Relativity and Cosmology*, Cambridge University Press, Cambridge 2006; S. Carrol, *Spacetime and Geometry. An Introduction to General Relativity*, Addison Wesley, 2004; W. Rindler, *Relativity, Special, General, and Cosmological*, Oxford University Press, Oxford 2004; A. Trautman, W. Kopczyński, *Czasoprzestrzeń i grawitacja*, PWN, Warszawa 1984 (There exists English edition).

[3] J. A. Schouten, *Ricci–Calculus*, Springer-Verlag, Berlin 1954; S. Gołąb, *Tensor Calculus*, PWN, Warsaw 1966 (in Polish. There exists English edition).

[4] V. C. de Andrade, L. C. T. Guillen and J. G. Pereira, Int. J. Mod. Phys. **D13** (2004), 2193 (arXiv:gr-qc/0501017); R. Aldrovandi, T. G. Lucas and J. G. Pereira, *Does a tensorial energy-momentum density for gravitation exist?*, arXiv:0812.0034 [gr-qc]; R. Aldrovandi, T. G. Lucas, and J. G. Pereira, *Inertia and gravitation in teleparallel gravity*, arXiv:0812.0034 v.2 [gr-qc]; V. C. de Andrade, H. I. Arcos, and J,. G. Pereira, *Torsion as Alternative to Curvature in the Description of Gravitation*, arXiv: gr-qc/0412034; H. I. Arcos, T. G. Lucas, and J. G. Pereira, *A Consistent Gravitationally–Coupled Spin–2Field Theory*, arXiv:1001.3407 [gr-qc]; R. Aldrovandi, J. G. Pereira, and K. H. Vu, *Doing without the Equivalence principle*, arXiv:gr-qc/0410042; R. A. Mosua and J. G. Pereira, Gen. Rel. Gravit. **36** (2004), 2525 (arXiv:gr-qc/0312093); R. Aldrovandi and J. G. Pereira, *Gravitation: in search of the missing torsion*, arXiv:0801.4148 [gr-qc]; H. I. Arcos and J. G. Pereira, *Torsion and the gravitational interaction*, arXiv:gr-qc/0408096; J. G. Pereira, T. Vargas, and C. M. Zhang, *Axial–Vector Torsion and the Teleparallel Kerr Spacetime*, arXiv:gr-qc/0102070; V. C. de Andrade, L. C. T. Guillen, and J. G. Pereira, *Teleparallel gravity: an overview*, arXiv:gr-qc/0011087; T. G. Lucas, Y. N. Obukhov, and J. G. Pereira, *Regularizing role of teleparallelism*, arXiv:0909:2418 [gr-qc]; H. I. Arcos, V. C. de Andrade, and J. G. Pereira, *Torsion and Gravitation: a New View*, arXiv:gr-qc/0403074; J. W. Maluf, F. F. Faria, and K. H. Castello-Branco, Class. Quantum Grav. **20** (2003), 4683; V. C. de Andrade, L. C. T. Guillen, and J. G. Pereira, *Teleparellel Spin Connection*, arXiv:gr-qc/0104103; Y. N. Obukhov and J. G. Pereira, Phys. Rev. **D67** (2003), 044008; A. A. Sousa, R. B. Pereira, and A. C. Silva, *Energy and angular momentum densities in a Gödel–type universe in the teleparallel geometry*, arXiv:0803.1481 [gr-qc]; J. F. da Rocha-Neto and K. H. Castello-Branco, *Gravitational Energy of Kerr and Kerr Anti-de Sitter Space-Times in the Teleparallel Geometry*, arXiv:gr-qc/0205028; J. W. Maluf, J. F. da Rocha-Neto, T. M. L. Toribio, and K. H. Castello-Branco, *Energy and angular momentum of the gravitational field in the teleparallel geometry*, arXiv:gr-qc/0204035; M. Sharif and A. Jawal, *Energy Content of Some Well-Known Solutions in Teleparallel Gravity*, arXiv:1005.5203 [gr-qc]; R. Aldrovandi, J. G. Pereira, and K. H. Vu, *Selected Topics in Teleparallel Gravity*, arXiv:gr-qc/0312008; R. Aldrovandi, P. B. Barros, and J. G. Pereira, Gen. Rel. Gravit. **35** (2003), 991; V. C. de Andrade, L. C. T. Guillen, and J. G. Pereira, *Teleparallel Gravity and the Gravitational Energy-Momentum Density*, arXiv:gr-qc/0011079; R. Aldrovandi and J. G. Pereira, *An introduction to teleparallel gravity*, Instituto de Fisica Teorica, UNESP, São Paulo, Brasil 2007.

[5] A. Trautman, *On the Structure of the Einstein-Cartán Equations*, Istituto Nazionale di Alta Matematica, Symposia Mathematica **12** (1973), 139; F. W. Hehl, Gen. Rel. Gravit. **4** (1973), 333; F. W. Hehl, ibidem **5** (1974), 491; F. W. Hehl et al., Rev. Mod. Phys. **48** (1976), 393; F. W. Hehl et al., *Gravitation and the Poincaré Gauge Field Theory with Quadratic Lagrangian*, in: General Relativity and Gravitation, Vol. I., Ed. A. Held, Plenum Publishing Corporation 1980; A. Trautman, *Fiber Bundles, Gauge Fields and Gravitation*, in: General Relativity and Gravitation, Vol. I. Ed. A. Held, Plenum Publishing Corporation 1980; A. Trautman, *Einstein-Cartan Theory*, arXiv:gr-qc/0606062.

[6] J. Garecki, *On Torsion in a Theory of Gravity*, in: Relativity, Gravitation, Cosmology, Eds. V. Dvoeglazov and A. Espinoza Garrido, 2004 Nova Science Publishers, Inc.

[7] M. I. Wanas, *Absolute Parallelism Geometry: Developments, Applications and Problems*, arXiv:gr-qc/0209050.

[8] C. Möller, *Conservation laws in the tetrad theory of gravitation*, in: Relativistic Theories of Gravitation, Ed. L. Infeld, Pergamon Press, Oxford-London-Edinburgh-New York-Paris-Frankfurt. Copyright 1964 by PWN, Warszawa 1964; C. Möller, Mat. Fys. Medd. Dan. Vid. Selsk. **35** (1966), 14pp. (NORDITA Publications no.190); C. Möller, *On the Crisis in the Theory of Gravitation and a Possible Solution*, Mat. Fys. Medd. Dan. Vid. Selsk. **39** (1978), 31pp, Köbenhavn 1978.

[9] G. Zet, *Schwarzschild Solution on a Space-Time with Torsion*, arXiv:gr-qc/0308078.

[10] T. Sauer, *Field equations in teleparallel spacetime: Einstein's Fernparallelismus approach towards unified field theory*, arXiv:physics/0405142.

[11] J. Plebański, *Tetrads and Conservation Laws*, in: Relativistic Theories of Gravitation, Ed. L. Infeld, Pergamon Press, Oxford-London-Edinburgh-New York-Paris-Frankfurt. Copyright 1964 by PWN, Warszawa 1964.

[12] K. Hayashi and T. Shirafuji, Phys. Rev. **D 19** (1979), 3524; F. Müller-Hoissen, *On the tetrad theory of gravity*, MPI–PAE/Pth61/84, 17pp.

[13] Y. Itin, *Coframe geometry and gravity*, arXiv:0711.4209 [gr-qc]; Y. Itin, *Does the coframe geometry can serve as a unification scheme?*, arXiv:gr-qc/0409071; E. Schücking, *Einstein's Apple and Relativity's Gravitational Field*, arXiv:0903.3768 [physics.hist-ph]; F. B. Estabrook, *Conservation Laws for Vacuum Tetrad Gravity*, arXiv:gr-qc/0508081.

[14] W. Kopczyński, J. Phys. **A 15** (1982), 493.

Institute of Mathematics
University of Szczecin
Wielkopolska 15, PL-70-451 Szczecin
Poland
e-mail: garecki@wmf.univ.szczecin.pl

Presented by Jakub Rembieliński at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on October 27, 2011

TELEPARALELNY EKWIWALENT OGÓLNEJ TEORII WZGLĘDNOŚCI: UWAGI KRYTYCZNE

S t r e s z c z e n i e

Po przedstawieniu podstawowych faktów z ogólnej teorii względności oraz z teleparalelnej grawitacji, ograniczam się do analizy specjalnego modelu teleparalelnej grawitacji nazwanego przez jego twórców *teleparalelnym ekwiwalentem ogólnej teorii względności* (w skrócie **TEGR**). Model ten był (i jest) ostatnio intensywnie badany głównie przez matematyków i fizyków z Brazylii.

W pracy pokazuję, że **TEGR** jest zakamuflowanym, starym, tetradowym sformułowaniem ogólnej teorii względności, dokonanym w latach 60-tych i 70-tych XX-go wieku przez C. Möllera i podkreślam, że **TEGR** jest niejednoznacznym i trywialnym przeformułowaniem ogólnej teorii względności, które nie może dać nic lepszego niż standardowe sformułowanie tej teorii (moim zdaniem, przeformułowanie to jest gorsze).

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 119–128

Dedicated to Professor Roman S. Ingarden
on the occasion of his ninetieth birthday

Andrzej Krzysztof Kwaśniewski

GRADED POSETS INVERSE ZETA MATRIX FORMULA IIA

THE FORMULA OF INVERSE ζ-MATRIX FOR GRADED POSETS WITH THE FINITE SET OF MINIMAL ELEMENTS VIA NATURAL JOIN OF MATRICES AND DIGRAPHS TECHNIQUE – A. RELABELING AND EXERCISES

Summary

We arrive at the explicit formula for the inverse of zeta matrix for any graded posets with the finite set of minimal elements following the first reference which is referred to as SNACK that is Sylvester Night Article on Cobweb posets and KoDAG graded digraphs. We start with a training in relabeling: examples and exercises.

1. Training in relabeling – Exercise

As we were and are to compare formulas from papers using different labeling – write and/or learn to see formulas from the above and below Observations, definitions etc. as for $x, y, k, s \in N \cup \{\mathbf{0}\}$ on one hand and as for $x, y, k, s \in \boldsymbol{N}$ on the other hand. Because of the comparisons reason we shall tolerate and use both being indicated explicitly.

Let us start with picture Examples 9, 10, 11 of inverse zeta matrices subsequently corresponding to picture Examples 1, 2, 5. For that to do it is enough for now to use the recurrent definition of the Möbius function

$$\mu(x,y) = \begin{cases} 1, & x = y \\ -\sum_{x\leq z<y} \mu(x,z), & x < y \end{cases}.$$

Before doing that note that we deal with F-*graded* posets and contact Remark 1 for notation and typical relations relevant below.

Recall 2. What form of the August Ferdinand Möbius matrix we do expect by now. Recall: (see Observation 3) – in the case of Möbius $\mu = \zeta^{-1}$ matrix as it is obligatory $\mathbf{c}_{r,r+1} = \mathbf{-1}$.

Recall (Remark 1) Markov property and observe by inspection that – in the case of Möbius $\mu = \zeta^{-1}$ matrix *for cobweb posets* it is obligatory to put

$$\begin{aligned} &M\left(r_F \times (r+2)_F\right) \\ &= -[I\left(r_F \times (r+1)_F\right) I\left((r+1)_F \times (r+2)_F\right) - I\left(r_F \times (r+2)_F\right)] \end{aligned}$$

i.e.

$$M\left(r_F \times (r+2)_F\right) = -[\mathbf{(r{+}1)}_F - \mathbf{1}] I\left(r_F \times (r+2)_F\right),$$

thereby:

$$c_{r,r+2} = -[\mathbf{(r{+}1)}_F - \mathbf{1}] c_{r,r+1}, \quad c_{r,r+1} = -1.$$

– What about then with arbitrary F-graded posets $(P, \leq)$?

In what follows we consider (consult the Remark 1) motivating examples and then representative Examples 9, 10, 11, 12 of Möbius matrix. After that the looked for Theorem 2 is stated for arbitrary F-graded posets $(P, \leq)$.

Motivating examples

Example 1. Let $i = 1, ..., r_F$, $k = 1, ..., (r+1)_F$, $j = 1, ..., (r+2)_F$ as now we consider (Remark 1) $x_{r,i} \prec \cdot x_{r+1,k}$ where $\{x_{r,i}\} = \Phi_r$ and $\{x_{r+1,k}\} = \Phi_{r+1}$ are independent sets. Then

$$\mu(x_{r,i}, x_{r+2,j}) = -\sum_{x_{r,i} \leq z < x_{r+2,j}} \mu(x_{r,i}, z) = -\left(1 + \sum_{k=1}^{(r+1)_F} \mu(x_{r,i}, x_{r+1,k})\right),$$

i.e.

$$\mu(x_{r,i}, x_{r+2,j}) = +[(r+1)_F - 1] = c_{r,r+2} = -[(r+1)_F - 1] c_{r,r+1}.$$

Example 2. From Example 1 we infer that as $\mu(x_{r,i}, x_{r+2,j}) = \mu(x_r, x_{r+2})$ then it is now enough to consider what follows (x_r, x_{r+3} any fixed):

$$\begin{aligned} \mu(x_r, x_{r+3}) &= -\sum_{x_r \leq z < x_{r+3}} \mu(x_r, z) = -\left(1 + \sum_{x_{r+1} \leq z < x_{r+3}} \mu(x_r, z)\right) \\ &= -\left(1 + (r+1)_F \mu(x_r, x_{r+1}) + \sum_{x_{r+2} \leq z < x_{r+3}} \mu(x_r, z)\right) \\ &= -\left(1 - (r+1)_F + (r+2)_F \mu(x_r, x_{r+2})\right), \end{aligned}$$

i.e.

$$\mu(x_r, x_{r+3}) = -[(r+2)_F - 1] c_{r,r+2} = -[(r+2)_F - 1][(r+1)_F - 1].$$

Via straightforward induction we conclude that now for arbitrary $r, s \in N \cup \{0\}$ and for any cobweb poset the following is true.

Theorem 2 for cobweb posets. ($N \cup \{0\}$.)

$$c_{r,s} = [s=r] - [s=r+1] + [s>r+](-1)^{s-r}\left((s-r-1)_F - 1)\right)\ldots(3_F - 1))\ (+1) =$$

$$= [s=r] - [s=r+1] + [s>r+1](-1)^{s-r} \prod_{i=r+1}^{s-1} (i_F - 1).$$

Let us see now how it works and how this theorem may be extended to general case of arbitrary F-denominated poset. At first the representative Examples 9, 10, 11, 12 of Möbius matrix follow which might be derived right from the recurrent definition of Möbius function without even referring to the above theorem.

$$\begin{bmatrix}
\mathbf{1} & -1 & -1 & +1 & +1 & +1 & -2 & -2 & -2 & -2 & +6 & +6 & +6 & +6 & +6 & -24\cdots \\
0 & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & +2 & -6 & -6 & -6 & -6 & -6 & +24\cdots \\
0 & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2 & +2 & -6 & -6 & -6 & -6 & -6 & +24\cdots \\
0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & -1 & +3 & +3 & +3 & +3 & +3 & -12\cdots \\
0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & -1 & +3 & +3 & +3 & +3 & +3 & -12\cdots \\
0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & -1 & +3 & +3 & +3 & +3 & +3 & -12\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & -1 & -1 & +4\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & -1 & -1 & +4\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & -1 & -1 & +4\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & +4\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -1\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -1\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1\cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1}\cdots \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & .\cdots
\end{bmatrix}$$

Example 3: ζ_N^{-1}. The Möbius function matrix $\mu = \zeta^{-1}$ for the natural numbers i.e. N – cobweb poset:

$$\mu_N = \begin{bmatrix}
\mathbf{I}_{1\times 1} & -\mathbf{I}(1\times 2) & +I(1\times 3) & -2I(1\times 4) & +6I(1\times 5) \\
\mathbf{O}_{2\times 1} & \mathbf{I}_{2\times 2} & -\mathbf{I}(2\times 3) & -2I(2\times 4) & -6I(2\times 5) \\
O_{3\times 1} & \mathbf{O}_{3\times 2} & \mathbf{I}_{3\times 3} & -\mathbf{I}(3\times 4) & +3I(3\times 5) \\
O_{4\times 1} & O_{4\times 2} & \mathbf{O}_{4\times 3} & \mathbf{I}_{4\times 4} & -\mathbf{I}(4\times 5) \\
O_{5\times 1} & O_{5\times 2} & O_{5\times 3} & \mathbf{O}_{5\times 4} & \mathbf{I}_{5\times 5} \\
\ldots & \text{etc.} & \ldots & \text{and so on} & \ldots
\end{bmatrix}$$

Note. μ has of course natural join inherited structure, of course.

Example 4: $\mu_N = \zeta_N^{-1}$. The *block presentation* of the Möbius function matrix $\mu = \zeta^{-1}$ for the natural numbers i.e. N-cobweb poset.

The secret (?) code for this KoDAG is given by its KoDAG self-evident code-triangle of the *coding matrix* $C(\mu_F)$(a starting part of it shown below):

$$C(\mu_N) = \begin{bmatrix} +1 & -1 & +1 & -2 & +6 & -24 \\ -0 & +1 & -1 & +2 & -6 & +24 \\ +0 & -0 & +1 & -1 & +3 & -12 \\ -0 & +0 & -0 & +1 & -1 & +4 \\ +0 & -0 & +0 & -0 & +1 & -1 \\ . & . & . & . & . & \end{bmatrix}$$

$$\begin{bmatrix}
\mathbf{1} & -1 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & \cdots \\
0 & \mathbf{1} & -1 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & \cdots \\
0 & 0 & \mathbf{1} & -1 & -1 & +1 & +1 & +1 & -2 & -2 & -2 & -2 & -2 & +8 & +8 & +8 & \cdots \\
0 & 0 & 0 & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & +2 & +2 & -8 & -8 & -8 & \cdots \\
0 & 0 & 0 & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2 & +2 & +2 & -8 & -8 & -8 & \cdots \\
0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & -1 & -1 & +4 & +4 & +4 & \cdots \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & -1 & -1 & +4 & +4 & +4 & \cdots \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & +4 & +4 & +4 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & \cdots \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & .\cdots
\end{bmatrix}$$

Example 5: ζ_F^{-1}. The Möbius function matrix $\mu = \zeta^{-1}$ for F=Fibonacci sequence:

$$\mu_F = \begin{bmatrix} I_{1\times 1} & -I(1\times 1) & 0I(1\times 1) & 0I(1\times 2) & 0I(1\times 3) \\ O_{1\times 1} & I_{1\times 1} & -I(1\times 1) & 0I(1\times 2) & 0I(1\times 3) \\ O_{1\times 1} & O_{1\times 1} & I_{1\times 1} & -I(1\times 2) & +I(1\times 3) \\ O_{2\times 1} & O_{2\times 1} & O_{2\times 1} & I_{2\times 2} & -I(2\times 3) \\ O_{3\times 1} & O_{3\times 1} & O_{3\times 1} & 0_{3\times 2} & I_{3\times 3} \\ \ldots & \text{etc.} & \ldots & \text{and so on} & \ldots \end{bmatrix}$$

Example 6: ζ_F^{-1}. The *block presentation* of the Möbius function matrix $\mu = \zeta^{-1}$ for F=Fibonacci sequence.

Recall then and note here up and below the block structure:

$$\sigma = \begin{bmatrix} I_{1_F\times 1_F} & B(1_F\times 2_F) & B(1_F\times 3_F) & B(1_F\times 4_F) & B(1_F\times 5_F) & B(1_F\times 6_F) \\ 0_{2_F\times 1_F} & I_{2_F\times 2_F} & B(2_F\times 3_F) & B(2_F\times 4_F) & B(2_F\times 5_F) & B(2_F\times 6_F) \\ 0_{3_F\times 1_F} & 0_{3_F\times 2_F} & I_{3_F\times 3_F} & B(3_F\times 4_F) & B(3_F\times 5_F) & B(3_F\times 6_F) \\ 0_{4_F\times 1_F} & 0_{4_F\times 2_F} & 0_{4_F\times 3_F} & I_{4_F\times 4_F} & B(4_F\times 5_F) & B(4_F\times 6_F) \\ ... & \text{etc.} & ... & \text{and so on} & ... & \end{bmatrix}$$

where $B(k_F\times(k+1)_F)$ denote corresponding constant $k_F\times(k+1)_F$ matrices in the case of ζ or ζ^{-1} matrices for example, with matrix elements from the ring $R=$ $2^{\{1\}}$, $Z_2=\{0,1\}$, Z etc.

$$\begin{bmatrix}
\mathbf{1} & -1 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & \cdots \\
0 & \mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4 & +8 & +8 & +8 & -16 & -16 & \cdots \\
0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4 & +8 & +8 & \cdots \\
0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4 & +8 & +8 & \cdots \\
0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4 & +8 & +8 & \cdots \\
0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & \cdots \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & \cdots \\
0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 & -1 & -1 & +2 & +2 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1 & -1 & -1 & +2 & +2 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & -1 & -1 & +2 & +2 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 & -1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & -1 & -1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & -1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \cdots \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & \cdots
\end{bmatrix}$$

Example 7: ζ_F^{-1}. The Möbius function matrix $\mu=\zeta^{-1}$ for ($1_F=2_F=1$ and $n_F=3$ for $n\geq 2$) the $F=\mathit{Fibonacci}$ relative special sequence $\mathbf{F}$ constituting the label sequence denominating cobweb poset associated to F-KoDAG Hasse digraph:

$$\mu_F = \begin{bmatrix} I_{1\times 1} & -I(1\times 1) & +0I(1\times 3) & -0I(1\times 3) & +0I(1\times 3) \\ O_{1\times 1} & +I_{1\times 1} & -I(1\times 3) & +2I(1\times 3) & -4I(1\times 3) \\ O_{3\times 1} & -O_{3\times 1} & +I_{3\times 3} & -I(3\times 3) & +2I(3\times 3) \\ O_{3\times 1} & +O_{3\times 1} & -O_{3\times 3} & +I_{3\times 3} & -I(3\times 3) \\ O_{3\times 1} & -O_{3\times 1} & +O_{3\times 3} & -0_{3\times 3} & +I_{3\times 3} \\ ... & \mathit{etc} & ... & \mathit{and\ so\ on} & ... \end{bmatrix}$$

Example 8: ζ_F^{-1}. The *block presentation* of the Möbius function matrix $\mu=\zeta^{-1}$ for ($1_F=2_F=1$ and $n_F=3$ for $n\geq 2$) the $F=\mathit{Fibonacci}$ relative special sequence $\mathbf{F}$ constituting the label sequence denominating cobweb poset associated to F-KoDAG Hasse digraph.

The secret (?) code for this KoDAG is given by its KoDAG self-evident code-triangle of the coding matrix $C(\mu_F)$ (a starting part of it shown below):

$$C(\mu_F) = \begin{bmatrix} 1 & -1 & +0 & -0 & +0 & -0 \\ 0 & +1 & -1 & +2 & -4 & +8 \\ 0 & -0 & +1 & -1 & +2 & -4 \\ 0 & +0 & -0 & +1 & -1 & +2 \\ 0 & -0 & +0 & -0 & +1 & -1 \\ . & . & . & . & . & \end{bmatrix}$$

$$\begin{bmatrix}
\mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4 & +8 & +8 & +8 & -16 & -16 & -16\cdots \\
0 & +\mathbf{1} & +\mathbf{0} & +\mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4 & +8 & +8 & +8\cdots \\
0 & -\mathbf{0} & +\mathbf{1} & +\mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4 & +8 & +8 & +8\cdots \\
0 & +\mathbf{0} & -\mathbf{0} & +\mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4 & +8 & +8 & +8\cdots \\
0 & -0 & +0 & -0 & +\mathbf{1} & +\mathbf{0} & +\mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4\cdots \\
0 & +0 & -0 & +0 & -\mathbf{0} & +\mathbf{1} & +\mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4\cdots \\
0 & -0 & +0 & -0 & +\mathbf{0} & -\mathbf{0} & +\mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2 & -4 & -4 & -4\cdots \\
0 & +0 & -0 & +0 & -0 & +0 & -0 & +\mathbf{1} & +\mathbf{0} & +\mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2\cdots \\
0 & -0 & +0 & -0 & +0 & -0 & +0 & -\mathbf{0} & +\mathbf{1} & +\mathbf{0} & -1 & -1 & -1 & +2 & +2 & +2\cdots \\
0 & +0 & -0 & +0 & -0 & +0 & -0 & +\mathbf{0} & -\mathbf{0} & +\mathbf{1} & -1 & -1 & -1 & +2 & +2 & +2\cdots \\
0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +\mathbf{1} & +\mathbf{0} & +\mathbf{0} & -1 & -1 & -1\cdots \\
0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -\mathbf{0} & +\mathbf{1} & \mathbf{0} & -1 & -1 & -1\cdots \\
0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +\mathbf{0} & -\mathbf{0} & +\mathbf{1} & -1 & -1 & -1\cdots \\
0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +\mathbf{1} & +\mathbf{0} & +\mathbf{0}\cdots \\
0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +\mathbf{1} & +\mathbf{0}\cdots \\
0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +0 & -0 & +\mathbf{1}\cdots \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & .\cdots
\end{bmatrix}$$

Example 9: ζ_F^{-1}. The Möbius function matrix $\mu = \zeta^{-1}$ for ($1_F = 1$ and $n_F = 3$ for $n \geq 2$) the N relative special sequence $\mathbf{F}$ constituting the label sequence denominating cobweb poset associated to F-KoDAG Hasse digraph:

$$\mu_F = \begin{bmatrix} I_{1\times 1} & -I(1\times 1) & +2I(1\times 3) & -4I(1\times 3) & +8I(1\times 3) \\ O_{1\times 1} & +I_{1\times 1} & -I(1\times 3) & +2I(1\times 3) & -4I(1\times 3) \\ O_{3\times 1} & -O_{3\times 1} & +I_{3\times 3} & -I(3\times 3) & +2I(3\times 3) \\ O_{3\times 1} & +O_{3\times 1} & -O_{3\times 3} & +I_{3\times 3} & -I(3\times 3) \\ O_{3\times 1} & -O_{3\times 1} & +O_{3\times 3} & -0_{3\times 3} & +I_{3\times 3} \\ \ldots & \text{etc.} & \ldots & \text{and so on} & \ldots \end{bmatrix}$$

Example 10: ζ_F^{-1}. The *block presentation* of the Möbius function matrix $\mu = \zeta^{-1}$ for ($1_F = 1$ and $n_F = 3$ for $n \geq 2$) the N relative special sequence $\mathbf{F}$ constituting the label sequence denominating cobweb poset associated to F-KoDAG Hasse digraph.

The secret (?) *code* for this KoDAG is given by its KoDAG self-evident code-triangle of the *coding matrix* $C(\mu_F)$ (a starting part of it shown below):

$$C(\mu_F) = \begin{bmatrix} 1 & -1 & +2 & -4 & +8 & -16 \\ 0 & +1 & -1 & +2 & -4 & +8 \\ 0 & -0 & +1 & -1 & +2 & -4 \\ 0 & +0 & -0 & +1 & -1 & +2 \\ 0 & -0 & +0 & -0 & +1 & -1 \\ . & . & . & . & . & \end{bmatrix}$$

From Observation 2 we infer what follows as obvious.

Oservation 3 (compare with Remark 1). *The block structure of ζ and consequently the block structure of μ for any graded poset with finite set of minimal elements (including cobwebs) is of the type:*

$$\zeta = \begin{bmatrix} I_1, B_1... & & & \\ & I_2, B_2... & & \\ & & I_3, B_3... & \\ & & ... & \\ & & & I_n, B_n... \end{bmatrix},$$

$$\mu = \begin{bmatrix} I_1, -B_1... & & & \\ & I_2, -B_2... & & \\ & & I_3, -B_3... & \\ & & ... & \\ & & & I_n, -B_n... \end{bmatrix},$$

$n \in N \cup \{\infty\}, \zeta, \mu \in I(\Pi; R)$, *where* $I_r = I_{r_F \times r_F}$ *and* $B_r = B(r_F \times (r+1)_F)$ *as introduced by Observation* 2.

Recall 3. Recall then and note here up and below the block structure ζ and consequently the block structure of μ *for any graded poset* P with finite set of minimal elements (including cobwebs) which is proprietary characteristic for any $\sigma \in I(P; R)$ where the ring $R = 2^{\{1\}}$, $Z_2 = \{0, 1\}$, Z etc.:

$$\sigma = \begin{bmatrix} I_{1_F \times 1_F} & M(1_F \times 2_F) & M(1_F \times 3_F) & M(1_F \times 4_F) & M(1_F \times 5_F) & M(1_F \times 6_F) \\ 0_{2_F \times 1_F} & I_{2_F \times 2_F} & M(2_F \times 3_F) & M(2_F \times 4_F) & M(2_F \times 5_F) & M(2_F \times 6_F) \\ 0_{3_F \times 1_F} & 0_{3_F \times 2_F} & I_{3_F \times 3_F} & M(3_F \times 4_F) & M(3_F \times 5_F) & M(3_F \times 6_F) \\ 0_{4_F \times 1_F} & 0_{4_F \times 2_F} & 0_{4_F \times 3_F} & I_{4_F \times 4_F} & M(4_F \times 5_F) & M(4_F \times 6_F) \\ ... & \text{etc.} & ... & \text{and so on} & ... & \end{bmatrix}$$

where in the case of $\oplus\!\!\rightarrow$-natural ζ or ζ^{-1} matrices, with matrix elements from the ring $R = 2^{\{1\}}$, $Z_2 = \{0, 1\}$, Z etc the rectangle *non-zero* block matrices $M(k_F \times (k+1)_F)$ denote corresponding connected *graded poset characteristic* $k_F \times (k+1)_F$ matrices.

Note then that

$$M(k_F \times (k+1)_F)_{r,s} = c_{i,j,k} B(k_F \times (k+1)_F)_{i,j}$$

for $i = 1, ..., k_F$ and $i = 1, ..., (k+1)_F$, where the rectangular "zero-one" $B(k_F \times (k+1)_F)$ matrices were introduced by the Observation 2. Consult Remark 1 – apart from the *Petitio Principi* motivating examples – for $i = 1, ..., k_F$ and $i = 1, ..., (k+1)_F$ as the layer $\langle \Phi_k \longrightarrow \Phi_{k+1} \rangle$ variables.

Note now the *important* fact. The relation

$$M(k_F \times (k+1)_F)_{i,j} = c_{i,j,k} B(k_F \times (k+1)_F)_{i,j},$$

where

$$i = 1, ..., k_F, \quad i = 1, ..., (k+1)_F$$

does not fix uniquely *the layer* $\langle \Phi_k \longrightarrow \Phi_{k+1} \rangle$ *coding matrix* $C_{k,k+1} = (c_{i,j,k})$, $i = 1, ..., k_F$, $i = 1, ..., (k+1)_F$ for F-denominated *arbitrary graded poset* – *except for* cobweb posets for which

$$B(k_F \times (k+1)_F) = I(k_F \times (k+1)_F).$$

In order to delimit this layer coding matrix *uniquely* we define *en bloc* the coding matrix $\mathbf{C}(\mu_F)$ for all layers.

Definition. F-graded poset $\langle \Phi, \mu_F \rangle$ *coding matrix* $\mathbf{C}(\mu_F)$.

Let $k, r, s \in N \cup \{\mathbf{0}\}$. Then we define $\mathbf{C}(\mu_F)$ via $\oplus\!\!\rightarrow$ originated blocks as follows:

$$\mathbf{C}(\mu_F) = (\mathbf{c}_{r,s})$$

where $\mathbf{c}_{r,s}$ are coding matrix elements for F-denominated cobweb poset, hence

$$\mu_F = \left([r = s] I_{r_F, r_F} + [s > r] \mathbf{c}_{r,s} B(r_F \times s_F) \right),$$

and where

$$c_{i,j,k} \equiv M(k_F \times (k+1)_F)_{i,j} = c_{i,j} B(k_F \times (k+1)_F)_{i,j};$$

thus the following identifications are self-evident:

$$\langle \Phi, \mu_F \rangle \equiv \langle \Phi, \zeta_F \rangle \equiv \langle \Phi, \leq \rangle \equiv \langle \Phi, \mathbf{C}(\mu_F) \rangle.$$

Result: $\mathbf{C}(\mu_F)$ as well as block sub-matrices $M(k_F \times (k+1)_F) = (c_{i,j,k})$ where $k \in N \cup \{\mathbf{0}\}$ are defined i.e. are given unambiguously.

Specifically, in *cobweb posets case*: for ζ function (matrix) we have

$$M(k_F \times (k+1)_F) = I(k_F \times (k+1)_F),$$

while for $\zeta^{-1} = \mu$ Möbius function (matrix) – from already considered examples' prompt we have already deduced these unambiguous $\mathbf{c}_{r,s}$ (see Theorem 2 for cobweb posets – above). Namely:

$$M(r_F \times (r+1)_F) = \mathbf{c}_{r,r+1} I(r_F \times (r+1)_F).$$

What about any F-denominated graded posets then? The answer *now* is of course secured now to be the same as for F-cobweb posets. The answer is automatically secured by the Definition 6. Just replace in the above Theorem 2 for cobweb posets $I(r_F \times (r+1)_F)$ by $B(r_F \times (r+1)_F)$ and-or see the Theorem 2 below for the corresponding recurrence equivalent to that from the *Petitio Principi* motivating examples resulting recurrence relation definition for $\mathbf{c}_{r,s}$.

In order to be complete also with the next section content another important example - the example of cover relation $\kappa_\Pi \in I(\Pi, R)$ matrix follows. Recall for that purpose now Observation 1 and the Remark 1 as to conclude what follows.

Observation 4. $(n \in N \cup \{\infty\})$ *The block structure of cover relation* $\kappa_\Pi \in I(\Pi, R)$ $(\chi(\prec\cdot_\Pi) \equiv \kappa_\Pi,)$ *is the following*

$$\kappa_\Pi = \oplus\!\!\rightarrow_{k=1}^{n} \kappa_k =$$

$$= \begin{bmatrix} 0_{1_F\times 1_F} & I(1_F\times 2_F) & 0_{1_F\times\infty} & & \\ 0_{2_F\times 1_F} & 0_{2_F\times 2_F} & I(2_F\times 3_F) & 0_{2_F\times\infty} & \\ 0_{3_F\times 1_F} & 0_{3_F\times 2_F} & 0_{3_F\times 3_F} & I(3_F\times 4_F) & 0_{3_F\times\infty} \\ & & \ldots & & \\ 0_{n_F\times 1_F} & \ldots & 0_{n_F\times n_F} & I(n_F\times(n+1)_F) & 0_{n_F\times\infty} \end{bmatrix},$$

where κ_k *is a cover relation of di-biclique*

$$\langle \Phi_k \rightarrow \Phi_{k+1}\rangle, \quad I_k \equiv I(k_F \times (k+1)_F), \quad k = 1, \ldots, n$$

and where – recall – $I(s \times k)$ *stays for* $(s \times k)$ *matrix of ones i.e.* $[I(s\times k)]_{ij} = 1$; $1 \le i \le s, 1 \le j \le k$. *while* $n \in N \cup \{\infty\}$ *and consequently the block structure of reflexive cover relation* $\eta_\Pi \in I(\Pi, R)$ $(\chi(\le\cdot_\Pi) = \prec\cdot_\Pi + \delta \equiv \eta_\Pi)$ *is given by*

$$= \begin{bmatrix} I_{1_F\times 1_F} & I(1_F\times 2_F) & 0_{1_F\times\infty} & & \\ 0_{2_F\times 1_F} & I_{2_F\times 2_F} & I(2_F\times 3_F) & 0_{2_F\times\infty} & \\ 0_{3_F\times 1_F} & 0_{3_F\times 2_F} & I_{3_F\times 3_F} & I(3_F\times 4_F) & 0_{3_F\times\infty} \\ & & \ldots & & \\ 0_{n_F\times 1_F} & \ldots & I_{n_F\times n_F} & I(n_F\times(n+1)_F) & 0_{n_F\times\infty} \end{bmatrix}.$$

Specifically, if restricting to cobweb posets: for ζ function (matrix) we have $B(k_F \times (k+1)_F) = I(k_F \times (k+1)_F)$, while for $\zeta^{-1} = \mu$ Möbius function (matrix) we would expect

$$B(r_F \times (r+1)_F) = c_{r,r+1} I(r_F \times (r+1)_F)$$

where $c_{k,k+1} = [C(\mu_F)]_{k,(k+1)}$.

What is then the explicit formula for $c_{k,k+1}$? It is of course equivalent to the question: what is then the explicit formula for $c_{r,s}$? Let us recapitulate our experience till now in order to infer the closing answer: Theorem 2 and its equivalent proof method.

Training in relabeling – Exercise. As we were and are to compare formulas from papers using different labeling – write and learn to see formulas from the above and below Observations as for $x, y, k, s \in N \cup \{\mathbf{0}\}$ on one hand and as for $x, y, k, s \in \boldsymbol{N}$ on the other hand. Because of the comparisons repeatedly reason we shall tolerate and use both being indicated explicitly if needed.

References

[1]–[46] See this issue, pp. 138–140.

Institute of Combinatorics and its Applications
High School of Mathematics and Applied Informatics
Kamienna 17, PL-15-021 Białystok
Poland
e-mail: kwandr@gmail.com

Presented by Julian Ławrynowicz at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on December 22, 2010

FORMUŁA NA MACIERZ MÖBIUSA DOWOLNEGO, CZĘŚCIOWO UPORZĄDKOWANEGO ZBIORU Z GRADACJĄ II
– A. PRZENUMEROWANIA I PRZYKŁADY

S t r e s z c z e n i e

W części II wskazuje się bezpośredni sposób otrzymywania macierzy Möbiusa dowolnego częściowo uporządkowanego zbioru z gradacją z wyprowadzonej jawnej formuły na postać tej macierzy dla szczególnych częściowo uporządkowanych zbiorów ze stopniowaniem zwanych ,,cobweb posets". W szczególności część IIA zawiera informacje trenujące przenumerowania: przykłady i ćwiczenia.

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 129–141

Dedicated to Professor Roman S. Ingarden
on the occasion of his ninetieth birthday

Andrzej Krzysztof Kwaśniewski

GRADED POSETS INVERSE ZETA MATRIX FORMULA IIB

THE FORMULA OF INVERSE ζ-MATRIX FOR GRADED POSETS WITH THE FINITE SET OF MINIMAL ELEMENTS VIA NATURAL JOIN OF MATRICES AND DIGRAPHS TECHNIQUE – B. WEIGHTED REFLEXIVE REACHABILITY RELATION

Summary

We arrive at the explicit formula for the inverse of zeta matrix for any graded posets with the finite set of minimal elements following the first reference which is referred to as SNACK that is Sylvester Night Article on Cobweb posets and KoDAG graded digraphs. In SNACK the way to arrive at formula of the zeta matrix for any graded posets with the finite set of minimal elements was delivered and explicit form was given. We present here effective way toward the formula for the inverse of zeta matrix which is being unearthed via adjacency and zeta matrix description of bipartite digraphs chains, the representatives of graded posets with sine qua non essential use of digraphs and matrices natural join introduced by the present author.

Namely, the bipartite digraphs elements of such chains amalgamate so as to form corresponding cover relation graded poset digraphs with corresponding adjacency matrices being amalgamated throughout natural join constituting adequate special database operation. As a consequence apart from zeta function also the Möbius function explicit expression for any graded posets with the finite set of minimal elements is being arrived at.

Purposely, on the way – special number theoretic code-triangles for KoDAGs are proposed and apart from the author combinatorial interpretation of F-nomial coefficients another related interpretation is inferred while referring to the number of all maximal chains in the corresponding poset interval. The formula for August Ferdinand Möbius matrix is also interpreted combinatorially.

2. Further training in relabeling

Recapitulation 2.1 (notation and the formula). The code $C(\mu_F)$ matrix no more secret.

Notation. Upside down notation development continuation.

Recall:

$$n^{\overline{k}} = n(n+1)(n+2)...(n+k-1),$$

denote:

$$n_F^{\overline{k}} \equiv n_F(n+1)_F(n+2)_F...(n+k-1)_F.$$

Denote (valid whenever defined for corresponding functions f of the natural number argument or of an argument from any chosen ring):

$$f(r_F)^{\overline{k}} = f(r_F)f([r+1]_F)...f([r+k-1]_F),\ n^{\overline{0}} \equiv 1,\ n \in N \cup\ \{\boldsymbol{0}\}, Z, R, etc.,$$

$$f(r_F)^{\underline{k}} = f(r_F)f([r-1]_F)...f([r-k+1]_F),\ n^{\underline{0}} \equiv 1,\ n \in N \cup\ \{\boldsymbol{0}\}, Z, R, etc..$$

Define Krot-on-shift-functions $K_s,\ s, r, i \in N \cup \{0\}$ or *Kroton* functions in brief -(Kroton = Croton = Codiaeum).

Definition ($\mathbf{N}\ \cup \{\mathbf{0}\}$ labels)

$$K_s(r_F) = [s > r][(r+1)_F - 1]^{\overline{s-r}}.$$

These of course constitute an upper triangle matrix with zeros on the diagonal for $s, r \in N \cup \{0\}$, ($\mathbf{r}$ = labels **r**ows).

Note two cases:

Let $s - r - 1 \neq 0$. Then

$$K_s(r_F) = [s > r] \prod_{i=r+1}^{s-1} (i_F - 1).$$

Let $s - r - 1 = 0$. Then

$$K_s(r_F) = [s > r].$$

Now – with this $\mathbf{N}\ \cup \{\mathbf{0}\}$ labeling as established in this note (Remark 2.1) – perform simple calculations. *Fibonacci* sequence $F = \langle \mathbf{1}, 1, 2, 3, 5, 8, 13, 21, 34, ... \rangle$ case Example.

$K_2(1_F) = 1$, $K_s(1_F) = 0$ for $s > 2$;
$K_3(2_F) = 1$, $K_s(2_F) = 0$ for $s > 3$;
$K_4(3_F) = 1$, $K_5(3_F) = 1$, $K_6(3_F) = 2$, $K_7(3_F) = 2 \cdot 4 = 8$, $K_8(3_F) = 8 \cdot 12 = 96$, $K_9(3_F) = 96 \cdot 20 = 1920$, and so on,
$K_5(4_F) = 1$, $K_6(4_F) = 1 \cdot 4$, $K_7(4_F) = 4 \cdot 7 = 14$, $K_8(4_F) = 14 \cdot 12 = 168$, $K_9(4_F) = 168 \cdot [F_8 - 1] =?$, $K_{10}(4_F) = 3360 \cdot [9_F - 1] =?$, and so on. Note that in the course of the above the following was used ($N \cup \{0\}$ – labeling).

Lemma 2.1. ($r, s \in N \cup \{0\}$. Obvious)

$$K_{s+1}(r_F) = K_s(r_F) \bullet [s_F - 1], \quad K_{r+1}(r_F) = 1,$$

N sequence case Example. This exercise has obvious outcomes in view of Lemma 2.1. For the just check of results see absolute values of *coding matrix* matrix elements from Example 9.

The next fact we mark as Lemma because of its importance.

Lemma 2.2. (Obvious – recapitulation). *Let $R = N, Z, ...,$ any commutative ring. For any graded F-denominated poset (hence connected) i.e. for any chain of subsequent natural joins of bipartite digraphs (di-bicliques for KoDAGs) and with the linear labeling of nodes fixed ($s, r \in N \cup \{0\}$ as in Remark 2.1. or $s, r \in N$):*

$$\mu = (\delta_{r,s} I_{r_F \times r_F} + [s > r] C(\mu_F)_{r,s} B(r_F \times s_F))$$

where $C(\mu_F)_{r,s} \in R$ are given by Definition 6. while $B(r_F \times s_F)$ are nonzero matrices introduced in the Observation 2.

Bearing in mind Definitions 6 and 7 and the the above Lemma 2.2 we see that the Theorem 2 for cobweb posets extends to be true for all F-denominated posets.

Theorem 2. (Kwaśniewski). *Let F be any natural numbers valued sequence. Then for arbitrary F-denominated graded poset (cobweb posets included)*

$$C(\mu_F)_{r,s} = c_{r,s} = [r = s] + K_s(r_F)(-1)^{s-r} = [r = s] + [s > r](-1)^{s-r}[(r+1)_F - 1]^{\overline{s-r}},$$

with matrix elements from N or the ring $R= 2^{\{1\}}$, $Z_2 = \{0, 1\}$, Z etc.
i.e. for cobweb posets

$$\mu = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} & c_{1,3}I(1_F \times 3_F) & c_{1,4}I(1_F \times 4_F) & c_{1,5}I(1_F \times 5_F) & c_{1,6}I(1_F \times 6_F) \\ \mu_{2,1} & \mu_{2,2} & c_{2,3}I(2_F \times 3_F) & c_{2,4}I(2_F \times 4_F) & c_{2,5}I(2_F \times 5_F) & c_{2,6}I(2_F \times 6_F) \\ \mu_{3,1} & \mu_{3,2} & I_{3_F \times 3_F} & c_{3,4}I(3_F \times 4_F) & c_{3,5}I(3_F \times 5_F) & c_{3,6}I(3_F \times 6_F) \\ \mu_{4,1} & \mu_{4,2} & 0_{4_F \times 3_F} & I_{4_F \times 4_F} & c_{4,5}I(4_F \times 5_F) & c_{4,6}I(4_F \times 6_F) \\ ... & etc. & ... & and\ so\ on & ... & ... \end{bmatrix}$$

with

$$\mu_{1,1} = I_{1_F \times 1_F},\ \mu_{2,1} = 0_{2_F \times 1_F},\ \mu_{3,1} = 0_{3_F \times 1_F},\ \mu_{4,1} = 0_{4_F \times 1_F},$$

$$\mu_{1,2} = c_{1,2} I_{1_F \times 2_F},\ \mu_{2,2} = I_{2_F \times 2_F},\ \mu_{3,2} = 0_{3_F \times 2_F},\ \mu_{4,2} = 0_{4_F \times 2_F},$$

where $I(k_F \times (k+1)_F)$ denotes (recall) $k_F \times (k+1)_F$ matrix of all entries equal to one. For any F-denominated poset replace $I(k_F \times (k+1)_F)$ by $B(k_F \times (k+1)_F)$ obtained from $I(k_F \times (k+1)_F)$ via replacing adequately (in accordance with Hasse digraph) corresponding ones by zeros.

Another Proof. One may prove the above also as follows.

From motivating examples we know that $\mu(x_{r,i}, x_{s,j}) = \mu(x_r, x_s)$. Observe then how the recurrent definition of Möbius function matrix μ gives birth to daughter descendant of μ i.e. the block structure of Möbius function coding matrix $C(\mu)$ implying for $C(\mu)$ a recurrence allowing simple solution simultaneously with combinatorial interpretation of Kroton matrix $K = (K_s(r_F)) \equiv (K_{r,s})$, where $K_s(r_F) = |C(\mu)_{r,s}|$.

For that to do call back the recurrent definition of the Möbius function where $x, y \in \Phi$ for $\Pi = (\Phi, \leq)$ and where – note: $\mu(x,y) = -1$ for $x \prec \cdot y$:

$$\mu(x,y) = \begin{cases} 1 & ; \quad x = y \\ -\sum_{x \leq z < y} \mu(x,z), & x < y \end{cases} .$$

The above recurrent definition of the Möbius function becomes – after **linear** order labeling has been applied – either $r, s, i \in N \cup \{0\}$ – as fixed-stated in this note, Remark 2.1 or $r, s \in N$ – whereby r, s are block-row and block-column indexes correspondingly – say it again – the above recurrent definition Möbius function in the case of F-denominated graded posets becomes $(c_{r,r+1} = -1)$

$$c_{r,s} = \begin{cases} 1 & ; \quad s = r \\ -\sum_{r \leq i < s} c_{r,i}, & r < s \end{cases} .$$

For that to see note that $\forall x, y, z \in \Phi$, $\exists\ r, s, i \in N$ such that $x_r \in \Phi_r$, $y_s \in \Phi_s$, $z_i \in \Phi_i$, hence for $x_r < y_s \ \equiv\ r < s$ where (important!) r, s, i stay now for *labels of independent sets* (levels) $\{\Phi_k\}$ i.e. label steps of La Scala i.e. label blocks. Thereby

$$c_{r,s} = \mu(x_r, y_s) = -\sum_{x_r \leq z < y_s} \mu(x_r, z) = -\sum_{x_r \leq z_i < y_s} \mu(x_r, z_i) = \sum_{r \leq i < s} c_{r,i}.$$

(Bear in mind Lemma 2.2. in order to get back to μ matrix unblocked appearance if needed.) From this recurrence the thesis follows.

How does this happens? 1) Let us put $r = 1$ just for the moment in order to make an inspection via example (r stays for *block-row* label and $k > 1$) and 2) use the Russian babushka in Babushka inspection i.e. apply the recurrent relation above subsequently till the end - till the smallest of size 1 babushka is encountered which is here $c_{r,r+1} = -1$. Use then trivial induction to state the validity of what follows below for all relevant values of variables $r, s \in N$.

$$c_{1,k} = -\sum_{1 \leq i < k} c_{1,i} = \left(-\sum_{1 \leq i < k_F}\right)\left(-\sum_{1 \leq i < (k-1)_F}\right) ... \left(-\sum_{1 \leq i < 3_F}\right) c_{1,2},$$

i.e.

$$c_{1,k} = (-1)^{k-1} \left(\sum_{1 \leq i < k_F}\right) ... \left(\sum_{1 \leq i < 4_F}\right)\left(\sum_{1 \leq i < 3_F}\right)(+1),$$

i.e.

$$c_{1,k} = -[1+1=k] + [k > 2](-1)^{k-1}\,(k_F - 1))\,...\,(3_F - 1))\ (+1) =$$

$$= -[1+1=k] + [k > 2](-1)^{k-1} \prod_{i=2+1}^{k} (i_F - 1).$$

Similarly we conclude that now for arbitrary $r, s \in N$

$$c_{r,s} = [s = r] - [s = r+1] + [s > r+1](-1)^{s-r}\,(s_F - 1))\,...\,(3_F - 1))\ (+1) =$$

$$= [s = r] - [s = r + 1] + [s > r + 1](-1)^{s-r} \prod_{i=r+2}^{s} (i_F - 1).$$

Equivalently we conclude that now for arbitrary $r, s \in N \cup \{0\}$

$$c_{r,s} = [s = r] - [s = r + 1] + [s > r+](-1)^{s-r} \left((s - r - 1)_F - 1)\right) \ldots (3_F - 1)) \; (+1) =$$

$$= [s = r] - [s = r + 1] + [s > r + 1](-1)^{s-r} \prod_{i=r+1}^{s-1} (i_F - 1).$$

To colligate and to imagine hint. Starting from the left upper corner of La Scala of ζ, μ,...,$\sigma \in I(\Pi, R)$ *down* $\Downarrow$ is biunivoquely starting from the "bottom" or "root" minimal elements level Φ_0 *up* $\Uparrow$ the Hasse digraph $(\Pi, \prec\cdot)$ uniquely representing the "much, much more cobwebbed tree"' – the digraph $(\Pi, \leq)$.

Descriptive – combinatorial interpretation: Once the formula has been observed-derived as above the following turns out perceptible. Namely note that

1. for $F = N$, $[s \neq r]$, the Kroton matrix element $|\mathbf{C}(\mu_N)_{r,s}|$, where

$$\mathbf{C}(\mu_N)_{r,s} = \mathbf{c}_{r,s} = [s > r](-1)^{s-r}[(r + 1)_N - 1]^{\overline{s-r}}$$

is equal to the number of heads' dispositions of maximal chains tailed at one vertex of the $r - th$ level and headed up at one vertex of the s-th level. This biunivoquely corresponds to the number of summands $= |\mathbf{C}(\mu_N)_{r,s}|$ entering the recurrence calculation of the $\mathbf{C}(\mu_N)$ matrix ("the Russian babushka in Babushka introspection" with interchangeable signs) being in one to one correspondence with climbing up Hasse digraph i.e. descending down the matrix μ La Scala along the way uniquely encoded by the subjected to their heads disposition maximal chains

$$c =< x_{\mathbf{r}}, x_{r+1}, ..., x_{s-1}, x_{\mathbf{s}} >, \; x_i \in \Phi_i, \; i = \mathbf{r}, r + 1, ..., s - 1, \mathbf{s}$$

with the tail $\mathbf{r}$ and the head $\mathbf{s}$ fixed as start and the end points of the descending down the La Scala blocks trip ($\equiv$ climbing up the levels of the graded Hasse digraph $\langle\Phi, \prec\cdot\rangle$).

2. For the same interpretation in the general F-case apply the Upside Down Notation Principle.

According to and from the above one extracts the obvious now property of Kroton functions i.e. matrix elements of Kroton matrix $K = (K_s(r_F)) \equiv (K_{r,s})$.

Lemma 2.3. $(r, s \in N \cup \{0\})$.

$$K_{s+1}(r_F) = K_s(r_F) \bullet [s_F - 1], \quad K_{r+1}(r_F) = 1$$

is equivalent to

$$K_{r,s} = - \sum_{r \leq i < s} (-1)^{s-i} K_{r,i} \quad K_{r+1}(r_F) = 1, \quad s > r.$$

Remark 5. Colligation. Scrape together and proceed to collocate the above combinatorial interpretation with hyper-boxes from [9].

Recall Definitions 4 and 5. Recall: $C_{\max}(\Pi_n)$ is the set of all maximal chains of Π_n. Recall: $C_{\max}^{k,n} = \{\text{maximal chains in } \langle \Phi_k \to \Phi_n \rangle\}$. Consult now Section 3. in [9] in order to view $C_{\max}(\Pi_n)$ or $C_{\max}^{k,n}$ as the hyper-box of points.

Namely [9] denoting with $V_{k,n}$ the discrete finite rectangular F-hyper-box or $(k,n) - F$-hyper-box or in everyday parlance just (k,n)-box

$$V_{k,n} = [k_F] \times [(k+1)_F] \times ... \times [n_F]$$

we identify (see Figure 7.) the following two just by agreement according to the F-natural identification:

$$C_{\max}^{k,n} \equiv V_{k,n}$$

i.e.

$$C_{\max}^{k,n} = \{\text{maximal chains in } \langle \Phi_k \to \Phi_n \rangle\} \equiv V_{k,n}.$$

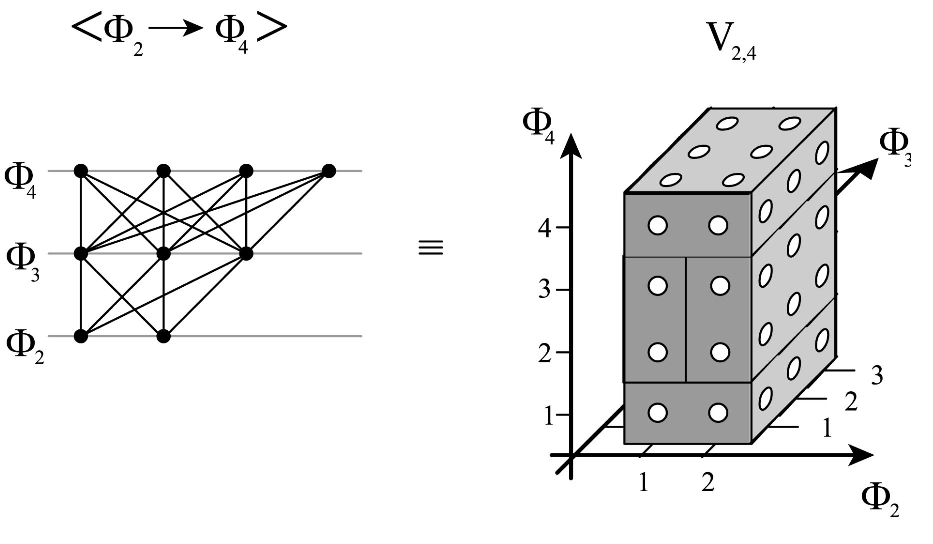

Fig. 7: A cobweb layer $\langle \Phi_2 \to \Phi_4 \rangle$ and equivalent hyper-box $V_{2,4}$.

Exercise. Deliver the descriptive combinatorial interpretation of Kroton matrix in the language of hyper-boxes from [9].

Recapitulation 2.2 (natural join). Recall that both $\leq$ partial order and $\prec\cdot$ cover relations are *natural join* of their bipartite correspondent chains, and this is exactly the reason and the very source of the Theorem 2 validity and shape. This is also the obvious clue statement for what follows. Note also that all on structure of any P

poset's information is coded by the ζ matrix – a characteristic function of $\leq \in \ P = \langle\Phi, \leq\rangle$. In short: ζ and equivalently $\mu = \zeta^{-1}$ are the Incidence algebra of P coding elements. In brief – recall – the following identifications are self-evident:

$$\langle\Phi, \mu_F\rangle \equiv \langle\Phi, \zeta_F\rangle \equiv \langle\Phi, \leq\rangle \equiv \langle\Phi, \mathbf{C}(\mu_F)\rangle .$$

3. F-nomial coefficients and $[Max]$ matrix of the N weighted reflexive reachability relation

Call back now the Remark 1. Then consider the incidence algebra of the *cobweb poset* Π as the algebra over (simultaneously) the ring R and the Boolean algebra $2^{\{1\}}$. Denote this incidence algebra by $I(\Pi, R, 2^{\{1\}}))$.

In the case $R = 2^{\{1\}}$ denote it by

$$I(\Pi, 2^{\{1\}}) \equiv I(\Pi, 2^{\{1\}}, 2^{\{1\}}).$$

Then for $\zeta \in I(\Pi, 2^{\{1\}})$ we have of course $\zeta^{-1} = \zeta$ ("reflexive reachability"), $\zeta_{\leq\cdot}^{-1} = \zeta_{\leq\cdot}$ (reflexive "cover") and so on. This is of course true for any poset relevant algebra i.e. for $I(P, 2^{\{1\}})$ – graded posets with finite set of minimal elements – included.

Consider now the algebra $I(\Pi, \mathbf{Z}, 2^{\{1\}}))$. We shall define now another characteristic matrix $[Max]$ as the matrix of the "N weighted" reflexive reachability relation. For that to do recall that in case of $I(\Pi, 2^{\{1\}})$

$$\leq = \prec\cdot^{*} = \text{reflexive reachability of } \prec\cdot$$

$$\prec\cdot^{*} \equiv (I - \prec\cdot)^{-1} = \prec\cdot^{0@} + \prec\cdot^{1@} + \prec\cdot^{2@} + \ldots + \prec\cdot^{k@} + \ldots \ \equiv \ \bigcup_{k\geq 0} \prec\cdot^{k},$$

where binary relations $\leq \subset \ \Phi\times\Phi$ and $\prec\cdot \ \subset \ \Phi\times\Phi$ etc. as subsets are identified with their matrices (see SNACK, [3, 2]), for example $\prec\cdot \equiv \kappa$. In the above the Boolean powers of κ were in action while here below this are to be powers over the $R = N, Z, 2^{\{1\}}$, etc.

The $[Max]$ matrix of the N weighted reflexive reachability relation is defined by the over the ring Z power series formula

$$[Max] = (I - \prec\cdot)^{-1} = \prec\cdot^{0} + \prec\cdot^{1} + \prec\cdot^{2} + \ldots + \prec\cdot^{k} + \ldots = \sum_{k\geq 0} \kappa^{k} = (I - \kappa)^{-1}.$$

Naturally

$$[Max]^{-1} = \delta - \kappa == \begin{bmatrix} I_1 & -B_1 & \text{zeros} & & & \\ & I_2 & -B_2 & \text{zeros} & & \\ & & I_3 & -B_3 & \text{zeros} & \\ & & \ldots & & & \\ & & & I_n & -B_n & \text{zeros} \end{bmatrix}$$

where (recall from Section I. 1.5)

$$[Max]_F = \mathbf{A}_F^0 + \mathbf{A}_F^1 + \mathbf{A}_F^2 + ... = (1 - \mathbf{A}_F)^{-1} =$$

$$= \left[\begin{array}{llllll} I_{1_F \times 1_F} & B(1_F \times 2_F) & B(1_F \times 3_F) & B(1_F \times 4_F) & B(1_F \times 5_F) & ... \\ 0_{2_F \times 1_F} & I_{2_F \times 2_F} & B(2_F \times 3_F) & B(2_F \times 4_F) & B(2_F \times 5_F) & ... \\ 0_{3_F \times 1_F} & 0_{3_F \times 2_F} & I_{3_F \times 3_F} & B(3_F \times 4_F) & B(3_F \times 5_F) & ... \\ 0_{4_F \times 1_F} & 0_{4_F \times 2_F} & 0_{4_F \times 3_F} & I_{4_F \times 4_F} & B(4_F \times 5_F) & ... \\ ... & \text{etc.} & ... & \text{and so on} & ... & \end{array}\right].$$

Comment 6. Combinatorial interpretation of $[Max]$.

$$[Max]_{s,t} = \text{the number of all maximal chains in the poset interval}$$
$$[x_{s,i}, x_{t,j}] = [x_s, x_t] \equiv [s, t],$$

where $x_{s,i}, x_s \in \Phi_s$ and $x_{t,j}, x_t \in \Phi_t$ for , say , $s \leq t$ with the reflexivity (loop) convention adopted i.e. $[Max]_{t,t} = 1$.

The above obvious statement being taken into the account, in view and in conformity with the environment of the Theorem 1 we arrive at the trivial and powerful Theorem 3.

Theorem 3. *Consider any F-cobweb poset with F being a natural numbers valued sequence. Let $x_k \equiv k \in \Phi_k$ and $x_t \equiv t \in \Phi_n$. Then*

$$\sum_{i \in \Phi_n} [Max]_{k,i} \equiv \sum_{i=1}^{n_F} [Max]_{k,i} = |C_{\max}\langle \Phi_{k+1} \to \Phi_n \rangle| = n_F^{\underline{m}},$$

where $m = n - k$.

Note that k, m, n are level labels (vertical) while $i = 1, ..., n_F$ stays for horizontal – along the fixed level – label. With that in mind fixed we observe what follows.

Corollary 3.1. *Consider any F-cobweb poset with F being a cobweb admissible sequence. Let $x_k \equiv k \in \Phi_k$ and $x_n \equiv n \in \Phi_n$. Let $n \geq k \equiv (n - m) \geq 2$. Then*

$$[Max]_{k,n} \, |\Phi_n| = n_F^{\underline{m}}$$

i.e.

$$[Max]_{k,n} = \binom{n-1}{k-2}_F (n - k + 1)_F!$$

Corollary 3.2. *Colligate with heads dispositions allied to the Theorem* 2.

Consider any F-cobweb poset with F being a *cobweb admissible* sequence. Let $x_k \equiv k \in \Phi_k$ and $x_m \equiv n \in \Phi_n$. Let $l + 1 = n \geq k \equiv (n - m) \geq 2$. Then

$$[Max]_{k,n} \, |\Phi_n| = n_F^{\underline{m}},$$

i.e.

$$\binom{n-1}{n-1-k}_F (n-1-k)_F! = \binom{n-1}{k}_F (n-1-k)_F! = [Max]_{k-2,n},$$

$$\binom{n-1}{n-1-k}_F = \frac{[Max]_{k-2,n}}{(n-1-k)_F!},$$

i.e. $(n-1=l)$

$$\binom{l}{k}_F = \binom{l}{l-k}_F = \frac{[Max]_{k-2,l+1}}{(l-k)_F!}.$$

Note that k, m, n, l are level labels (vertical) and this is convention to be kept till the end of this note.

The above obvious statement being taken into the account, in view and in conformity with the environment of Theorems 1 and 2 we are prompt to extract the trivial and powerful statement as the Theorem 4.

Theorem 4. *Consider any F-cobweb poset with F being a cobweb admissible sequence. Let $x_k \equiv k \in \Phi_k$ and $x_m \equiv n \in \Phi_n$. Let $(l+1) \geq k \geq 2$. Then*

$$\binom{l}{k}_F = \binom{l}{l-k}_F = \frac{[Max]_{k-2,l+1}}{(l-k)_F!},$$

i.e. $\binom{l}{k}_F = (l-k)_F!$'th fraction of the number of all maximal chains in the poset interval $[x_{k-2}, x_{l+1}]$, where $x_l \in \Phi_l$ and $x_k \in \Phi_k$ with the reflexivity (loop) convention adopted i.e. $[Max]_{n,n} = 1$.

Farewell Exercises

Problem-Exercise 3.1. Rewrite Markov property in F-nomials language.

Problem-Exercise 3.2. Find the inverse of $\binom{l}{k}_F$ using the Theorem 4 and the knowledge of $[Max]^{-1}$. Compare with [11].

Acknowledgments

Thanks are expressed here to the now student of Gdańsk University Maciej Dziemiańczuk for applying his skillful TeX-nology with respect most of my articles since three years as well as for his general assistance and cooperation on KoDAGs investigation. Maciej Dziemiańczuk was not allowed to write his diploma with me being supervisor – while Maciej studied in the local Bialystok University where my professorship till 2009.09.30 comes from.

The author expresses his gratitude also Dr Ewa Krot-Sieniawska for her several years' cooperation and vivid application of the alike material deserving, students' admiration for her being such a comprehensible and reliable teacher.

References

[1] A. K. Kwaśniewski, *Graded posets zeta matrix formula*, Bull. Soc. Sci. Lettres Łódź **60** Sér. Rech. Déform. **60**, no. 3 (2010), 99–115; arXiv:0901.0155v1 [v1] Thu, 1 Jan 2009 01:43:35 GMT (15 pages Sylvester Night paper).

[2] A. K. Kwaśniewski, *Some cobweb posets digraphs' elementary properties and questions*, arXiv:0812.4319v1, [v1] Tue, 23 Dec 2008 00:40:41 GMT,

[3] A. K. Kwaśniewski, *Cobweb posets and KoDAG digraphs are representing natural join of relations, their di-Bigraphs and the Corresponding Adjacency Matrices*, arXiv:math/0812.4066v1,[v1] Sun, 21 Dec 2008 23:04:48 GMT.

[4] A. K. Kwaśniewski, *Fibonomial cumulative connection constants*, arXiv:math/ 0406006v2 [v6] Fri, 20 Feb 2009 02:26:21 GMT, upgrade of Bulletin of the ICA **44** (2005), 81–92.

[5] A. K. Kwaśniewski, *On cobweb posets and their combinatorially admissible sequences*, Adv. Studies Contemp. Math. **18**, no. 1 (2009), 17–32; ArXiv:0512578v4 [v5] Mon, 19 Jan 2009 21:47:32 GMT;

[6] A. K. Kwaśniewski, *How the work of Gian Carlo Rota had influenced my group research and life*, arXiv:0901.2571 [v4] Tue, 10 Feb 2009 03:42:43 GMT.

[7] A. K. Kwaśniewski and M. Dziemiańczuk, *Cobweb posets – recent results*, ISRAMA 2007, December 1-17 2007 Kolkata, INDIA, Adv. Stud. Contemp. Math. **16**, no. 2 (2008), 197–218; arXiv:0801.3985v1 [v1] Fri, 25 Jan 2008 17:01:28 GMT.

[8] A. K. Kwaśniewski, *Cobweb posets as noncommutative prefabs*, Adv. Stud. Contemp. Math. **14**, no. 1 (2007), 37–47; arXiv:math/0503286v4 1: Tue, 15 Mar 2005 04:26:45 GMT.

[9] A. K. Kwaśniewski and M. Dziemiańczuk, *On cobweb posets' most relevant codings*, ArXiv:0804.1728v1, [v1] 10 Apr 2008 15:09:26 GMT, [v2] Fri, 27 Feb 2009 18:05:33 GMT.

[10] M. Dziemiańczuk, *On Cobweb posets tiling problem*, Adv. Stud. Contemp. Math. **16**, no. 2 (2008), 219–233.

[11] M. Dziemiańczuk, *On multi F-nomial coefficients and inversion formula for F-nomial coefficients*, ArXiv:0806.3626, 23 Jun 2008.

[12] M. Dziemiańczuk, *On cobweb admissible sequences – the production theorem*, Proceedings of The 2008 International Conference on Foundations of Computer Science (FCS'08), July 14-17, 2008, Las Vegas, USA, pp.163-165.

[13] M. Dziemiańczuk and W. Bajguz, *On GCD-morphic sequences*,IeJNART **3** (2009), 33–37; arXiv:0802.1303v1, [v1] Sun, 10 Feb 2008 05:03:40 GMT.

[14] M. Dziemiańczuk, *Cobweb posets website*, http://www.faces-of-nature.art.pl/cobwebposets.html.

[15] A. K. Kwaśniewski, *Combinatorial interpretation of Fibonomial coefficients*, Inst. Comp. Sci. UwB Preprint **52** November (2003).

[16] A. K. Kwaśniewski, *More on combinatorial interpretation of Fibonomial coefficients*, Inst. Comp. Sci. UwB Preprint **56** November (2003).

[17] A. K. Kwaśniewski, *Fibonomial cumulative connection constants*, Inst. Comp. Sci. UwB Preprint **58** December (2003); Bulletin of the ICA **44** (2005), 81–92; arXiv:math/0406006v1 [v1] Tue, 1 Jun **2004** 00:59:23 GMT [v6]Fri, 20 Feb 2009 02:26:21 GMT.

[18] A. K. Kwaśniewski, *Comments on combinatorial interpretation of Fibonomial coefficients – an email style letter*, Bulletin of the ICA **42** September (2004), 10–11,; arXiv:0802.1381.

[19] A. K. Kwaśniewski, *Combinatorial interpretation of the recurrence relation for Fibonomial coefficients*, Bull. Soc. Sci. Lettres Łódź **54** Ser. Rech. Deform. **44** (2004), 23–38; arXiv:math/0403017v2 [v1] Mon, 1 Mar **2004** 02:36:51 GMT

[20] A. K. Kwaśniewski, *More on combinatorial interpretation of Fibonomial coefficients*, Bull. Soc. Sci. Lettres Łódź **54** Ser.: Rech. Deform. **44** (2004), 23–38; arXiv:math/0402344v2 [v1] 23 Feb **2004** Mon, 20:27:13 GMT.

[21] D. E. Knuth, *Two Notes on Notation*, Amer. Math. Monthly **99**, no. 5 (1992), 403–422; arXiv:math/9205211v1, [v1] Fri, 1 May 1992 00:00:00 GMT.

[22] E. Krot, *A note on Möbius function and Möbius inversion formula of Fibonacci cobweb poset*, Bull. Soc. Sci. Lettres Łódź **54** Ser. Rech. Deform. **44** (2004), 39–44; arXiv:math/0404158v2, [v1] Wed, 7 Apr **2004** 10:23:38 GMT [v2] Wed, 28 Apr 2004 07:37:14 GMT.

[23] E. Krot, *The first ascent into the incidence algebra of the Fibonacci cobweb poset*, Advanced Studies in Contemporary Mathematics **11**, no. 2 (2005), 179–184; arXiv:math/0411007v1 [v1] Sun, 31 Oct *2004* 12:46:51 GMT.

[24] E. Krot-Sieniawska, *On incidence algebras description of cobweb posets*, arXiv:0802.3703v1 [v1] Tue, 26 Feb **2008** 13:12:43 GMT.

[25] E. Krot-Sieniawska, *Reduced incidence algebras description of cobweb posets and KoDAGs*, arXiv:0802.4293v1, [v1] Fri, 29 Feb **2008** 05:43:27 GMT.

[26] E. Krot-Sieniawska, *On characteristic polynomials of the family of cobweb posets*, Proc. Jangjeon Math. Soc. **11**, no. 2 (2008), 105–111; arXiv:0802.2696v1 [v1] Tue, 19 Feb **2008** 18:53:38 GMT.

[27] A. K. Kwaśniewski, *Towards ψ-extension of finite operator calculus of Rota*, Rep. Math. Phys. **48**, (3) (2001), 305-342; arXiv:math/0402078v1, [v1] Thu, 5 Feb 2004 13:02:30 GMT.

[28] A. K. Kwaśniewski, *On extended finite operator calculus of Rota and quantum groups*, Integral Transforms and Special Functions **2**, (4) (2001), 333.

[29] A. K. Kwaśniewski, *On simple characterizations of Sheffer ψ-polynomials and related propositions of the calculus of sequences*, Bull. Soc. Sci. Lett. Łódź Ser. Rech. Deform. **52**, Ser. Rech. Deform. **36** (2002), 45–65; arXiv:math/0312397v1 ,[v1] Sat, 20 Dec 2003 23:21:51 GMT.

[30] A. K. Kwaśniewski, *Main theorems of extended finite operator calculus*, Integral Transforms and Special Functions **14** (2003), 333.

[31] A. K. Kwaśniewski, *On basic Bernoulli-Ward polynomials*, Bull. Soc. Sci. Lettres Łódź **54** Ser. Rech. Deform. **45** (2004), 5–10; ArXiv: math/0405577v1 [v1] Sun, 30 May 2004 00:32:47 GMT.

[32] A. K. Kwaśniewski, *Graded posets inverse zeta matrix formula IIA. The formula of inverse zeta-matrix for graded posets with the finite set of minimal elements via natural join of matrices and digraphs technique – A. Relabeling and exercises*, ibid. **61**, no. 1 (2011), 119–128.

[33] A. K. Kwaśniewski, *The logarithmic Fib-binomial formula*, Adv. Stud. Contemp. Math. **9**, no. 1 (2004), 19–26; arXiv:math/0406258v1 [v1] Sun, 13 Jun 2004 17:24:54 GMT.

[34] A. K. Kwaśniewski, *Fibonacci-triad sequences*, Advan. Stud. Contemp. Math. **9**, (2) (2004), 109-118.

[35] A. K. Kwaśniewski, *Fibonacci q-Gauss sequences*, Adv. Stud. Contemp. Math. **8**, no. 2 (f2004), 121–124; ArXive: math.CO/0405591 31 May 2004.

[36] A. K. Kwaśniewski, *q-Poisson, q-Dobinski, Rota and coherent states-a fortieth anniversary memoir*, Proc. Jangjeon Math. Soc. **7**, (2) (2004), 95–98; arXiv:math/0402254v2 v2] Tue, 17 Feb 2004 21:49:39 GMT.

[37] A. K. Kwaśniewski, *Cauchy $\hat{q}_\psi$-identity and $\hat{q}_\psi$-Fermat matrix via $\hat{q}_\psi$-muting variables for Extended Finite Operator Calculus*, Proc. Jangjeon Math. Soc. **8**, no. 2 (2005), 191–196; Inst. Comp. Sci. UwB Preprint **60** December (2003); arXiv:math/0403107v1 [v1] Fri, 5 Mar 2004 09:57:32 GMT.

[38] A. K. Kwaśniewski, *Cauchy identity- and-Fermat matrix via muting variables for Extended Finite Operator Calculus*, Proc. Jangjeon Math. Soc. **8**, no. 2 (2005), 191–196; arXiv:math/0403107v1 [v1] Fri, 5 Mar 2004 09:57:32 GMT.

[39] A. K. Kwaśniewski, *ψ-Pascal and $\hat{q}_\psi$-Pascal matrices - an accessible factory of one source identities and resulting applications*, Advanced Stud. Contemp. Math. **10**, no. 2 (2005), 111–120; ArXiv:math.CO/0403123 v1 7 March 2004.

[40] A. K. Kwaśniewski: *Information on some recent applications of umbral extensions to discrete mathematics*, Review Bulletin of Calcutta Mathematical Society **13** (2005), 1–10; ArXiv:math.CO/0411145v1 [v2] Wed, 21 Sep 2005 14:12:33 GMT.

[41] A. K. Kwaśniewski, *Graded posets inverse zeta matrix formula I. Reference information and zeta-matrix formula*, Bull. Soc. Sci. Lettres Łódź **60** Sér. Rech. Déform. **60**, no. 3 (2010), 117–147.

[42] A. K. Kwaśniewski, *On umbral extensions of Stirling numbers and Dobinski-like formulas*, Adv. Stud. Contemp. Math. **12**, no. 1 (2006), 73–100; arXiv:math/0411002v5 [v5] Thu, 20 Oct 2005 02:12:47 GMT.

[43] A. K. Kwaśniewski, *First observations on prefab posets Whitney numbers*, Advances in Applied Clifford Algebras **18**, no. 1 (2008), 57–73; arXiv:0802.1696v1, [v1] Tue, 12 Feb 2008 19:47:18 GMT.

[44] Ewa Krot, *Further developments in finite fibonomial calculus*, Inst. Comp. Sci. UwB Preprint **64** February (2004); arXiv:math/0410550v2 [v1] Tue, 26 Oct 2004 10:37:43 GMT, [v2] Wed, 27 Oct 2004 08:01:59 GMT.

[45] J. Julve and M. Tonin, *Quantum gravity with higher derivative terms* , Il Nuovo Cimento B (1971-1996) **46**, no. 1 (1978), 137–152.

[46] A. K. Kwaśniewski, *Glimpses of the octonions and quaternions history and todays applications in quantum physics*, arXiv:0803.0119v1 [v1] Sun, 2 Mar 2008 12:39:16 GMT.

Institute of Combinatorics and its Applications
High School of Mathematics and Applied Informatics
Kamienna 17, PL-15-021 Białystok
Poland
e-mail: kwandr@gmail.com

Presented by Julian Ławrynowicz at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on December 22, 2010

FORMUŁA NA MACIERZ MÖBIUSA DOWOLNEGO, CZĘŚCIOWO UPORZĄDKOWANEGO, ZBIORU Z GRADACJĄ II

– B. RELACJA WAŻONEJ REFLEKSYJNEJ OSIĄGALNOŚCI

S t r e s z c z e n i e

W części IIB wskazuje się bezpośredni sposób otrzymywania macierzy Möbiusa dowolnego częściowo uporządkowanego zbioru z gradacją z wyprowadzonej jawnej formuły na postać tej macierzy dla szczególnych częściowo uporządkowanych zbiorów ze stopniowaniem zwanych „cobweb posets".

Cel ten jest osiągnięty dzięki owych „cobweb posets" jako i dowolnych (posets) częściowo uporządkowanych zbiorów ze stopniowaniem (gradacją) o skończonej liczbie elementów minimalnych utożsamieniu ze złączeniem naturalnym (natural join) łańcuchów grafów dwudzielnych.

Odzwierciedla to skutkująco struktura i macierzy sąsiedztwa i macierzy Möbiusa wszystkich acyklicznych grafów skierowanych zwanych diagramami Hasse tych „posetów" z gradacją. Jest to mianowicie postać sekwencyjnego złączenia naturalnego macierzy składowych łańcuchów grafów dwudzielnych.

W przypadku szczególnych częściowo uporządkowanych zbiorów ze stopniowaniem zwanych „cobweb posets" stanowiących w złączaniu naturalnym ciągi Kompletnych Grafów dwudzielnych – uporządkowanych (ordered) oraz skierowanych i acyklicznych (DAG's) autor na cześć Profesora Kazimierza Kuratowskiego nazwał owe grafy Hasse'go – KoDAGs.

PL ISSN 0459-6854

B U L L E T I N

DE LA SOCIÉTÉ DES SCIENCES ET DES LETTRES DE ŁÓDŹ

2011 Vol. LXI

Recherches sur les déformations no. 1

pp. 143–156

Dedicated to Professor Roman Stanisław Ingarden
on the occasion of his ninetieth birthday

Stanisław Bednarek and Tomasz Bednarek

THE SUBSTITUTE RESISTANCES, CAPACITANCES AND INDUCTANCES OF SOME FRACTAL NETWORKS

Summary

In the paper calculations of substitute resistances, capacitances and inductances of electric network having fractal shaped parameters have been made. Components of the network: resistors, capacitors or inductors are connected either in series or parallel, the values of whose consecutive elements constitute a geometric progression. The derived formulae for the border cases have been discussed. The obtained results are not only limit but may also be helpful in designing some electric network.

1. Introduction

The electrical resistance R is a physical quantity that is defined by the formula:

$$R = \frac{U}{I}, \tag{1}$$

where U is the voltage and I the current intensity [1]. In practice, the resistance is realized by the resistor that is represented graphically by the symbol below, see Fig. 1. The resistor's resistance is conventionally denoted by the letter R.

Resistors can be connected together in series, parallel or combinations of both, to produce more complex chains or networks. A good example of a simple circuit made up of a combination of series and parallel resistors is shown in Fig. 2. The simplest connection of resistors connected in series as shown in Fig. 3 will be examined. The chain should be replaced with one resistor that will be as effective as two resistors coupled together. The resistance of this single resistor is called a substitute resistance of the combination and denoted by R_S (the letter S stands for the English word to substitute).

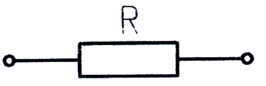

Fig. 1: The symbol of a resistor.

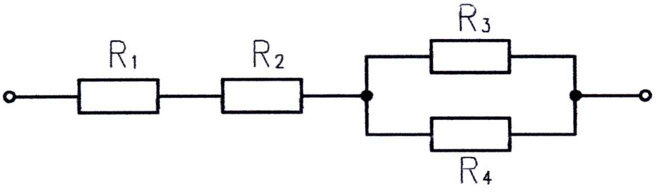

Fig. 2: The simple network of resistors.

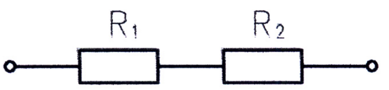

Fig. 3: Connection two resistors in series.

The value of the substitute resistance R_s, see Fig. 4, can be calculated employing the formula (1). Thus,

$$R_1 = \frac{U_1}{I}, \tag{2}$$

$$R_2 = \frac{U_2}{I}, \tag{3}$$

$$R_S = \frac{U}{I}. \tag{4}$$

The supply voltage U is equal to the sum of voltages U_1 and U_2, in accordance with the Kirchhoff's second law, and the current intensity is the same for both resistors (Kirchhoff's first law). It can be written as:

$$U = U_1 + U_2. \tag{5}$$

From the equations (2–4), $U_1,\ U_2, U$ are determined and replaced with (5). Thus, we obtain

$$R_S I = R_1 I + R_2 I. \tag{6}$$

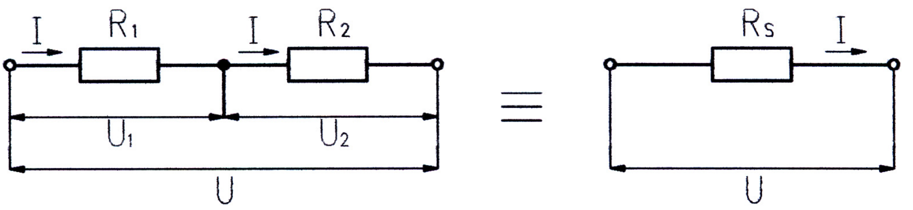

Fig. 4: Calculation of the substitute resistance two resistors connected in series.

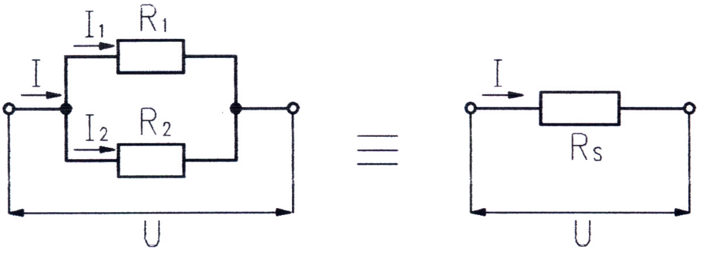

Fig. 5: Calculation of the substitute resistance two resistors connected in parallel.

By dividing both sides (6) by I we obtain

$$R_S = R_1 + R_2. \tag{7}$$

In a general case of n resistors connected in series, there is

$$R_S = \sum_{i=1}^{n} R_i. \tag{8}$$

Two resistors connected in parallel will be examined, see Fig. 5. Similarly as before, the equations can be written as follows:

$$R_1 = \frac{U}{I_1}, \tag{9}$$

$$R_2 = \frac{U}{I_2}, \tag{10}$$

$$R_S = \frac{U}{I}. \tag{11}$$

In this case, in accordance with the Kirchhoff's second law, the voltage across both resistors is the same and is equal to U, whereas current intensities I_1, I_2 flowing through each resistor satisfy the condition

$$I = I_1 + I_2. \tag{12}$$

From the equations (9–11), I_1, I_2, I are determined and replaced with (12), and then both sides of the obtained equation are divided by U. Thus, we obtain

$$\frac{1}{R_S} = \frac{1}{R_1} + \frac{1}{R_2}, \tag{13}$$

or

$$R_S = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}}. \tag{14}$$

In a general case of n resistors in parallel, there occurs

$$R_S = \frac{1}{\sum\limits_{i=1}^{n} \frac{1}{R_i}}. \tag{15}$$

The electrical capacity C is an electrical quantity that is defined by the equation

$$C = \frac{Q}{U}, \tag{16}$$

where Q is the electric charge and U is the voltage. The capacitance is realized by means of the capacitor that has been represented graphically in Fig. 6. The capacitor's capacitance is denoted conventionally by C. Similarly as before, our deliberations lead to the following formulae for the substitute capacitance C_S of the capacitors connected in series

$$C_S = \frac{1}{\sum\limits_{i=1}^{n}} \frac{1}{C_i}, \tag{17}$$

as well as for those connected in parallel

$$C_S = \sum_{i=1}^{n} C_i. \tag{18}$$

In the equation (16), U is the denominator, and therefore with the capacitors connected in series, the reciprocals of their individual capacitances should be added up, and with capacitors in parallel, their capacitances need to be summed up. In other words, quite the reverse as it is in the case of the resistors in parallel vs. in series. The inductance L is a physical quantity that is defined by the equation

$$L = \frac{\varphi}{I}, \tag{19}$$

where φ is the magnetic stream and I is current intensity. In practice, inductance is realized by means of a coil that is represented graphically in Fig. 7, and its inductance is conventionally denoted by L. Similarly as before, our deliberation leads us to a conclusion that if the adjacent coils do not affect each other magnetically, e.g. if they are separated by a considerable distance, then for the coils in series, the formula applies

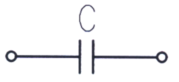

Fig. 6: The symbol of a capacitor.

Fig. 7: The symbol of a coil.

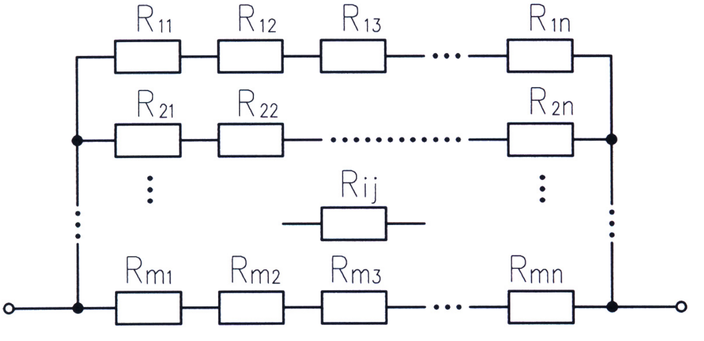

Fig. 8: Scheme of the complex network of resistors.

(20)
$$L_S = \sum_{i=1}^{n} L_i,$$

and for the parallel coils

(21)
$$L_S = \frac{1}{\sum\limits_{i=1}^{n} \frac{1}{L_i}}.$$

With a more complex combination of resistors made up of m chains with n resistors in each chain, as shown in Fig. 8, we can derive from the equations (15) and (18)

(22)
$$R_S = \frac{1}{\sum\limits_{i=1}^{m} \frac{1}{\sum_{j=1}^{n} R_{ij}}},$$

In the last case, the resistances R_{ij} may be treated as the elements of a rectangular matrix with m rows and n columns.

Let us now examine the properties of fractals. No comprehensive, broad or narrow definition of a fractal will be provided here, as it is still a matter of debate among mathematicians [2]. We will only consider the major defining feature of a fractal, which is self-similarity – that is the structure of a fractal is made up of parts that look like the original structure itself, or that its structure is similar at all scales. A good example of elementary fractals are the Cantor Set and the Sierpiński Gasket, also known as the Sierpiński Carpet [3] named after the Polish mathematician Waclaw Sierpiński [3].

The method of constructing the Cantor Set has been demonstrated in Fig. 9. The set is divided successively into three parts and the centre of the partition is removed. This process is iterated ad infinitum over all of the subsets that arise after successive iterations. A similar technique can be applied to construct the Sierpinski Carpet, see Fig. 10. In this case, an equilateral triangle is divided into four smaller, upside down triangles whose side lengths are equal to half the side length of the original or proceeding triangle. Then, the central, reversed sub-triangle is removed out the divided triangle. This process is iterated an infinite number of times over all of the triangles that have been obtained after successive divisions.

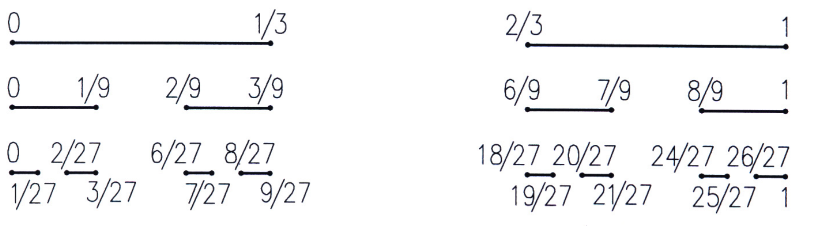

Fig. 9: Construction of the Cantor's holed set.

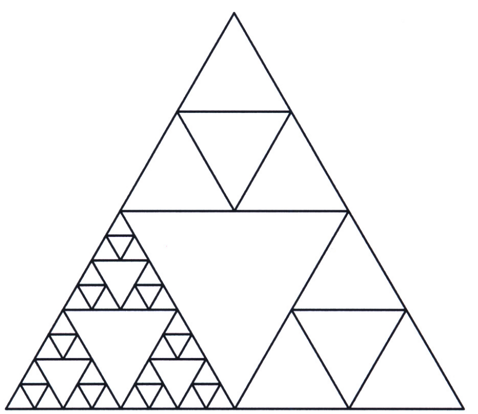

Fig. 10: Construction of the Sierpiński gasket.

2. The substitute resistance of the network of fractal resistors

A chain of n resistors in series will be reviewed and presented in Fig. 11. Suppose that the resistance i of this resistor satisfies the condition

$$R_i = R_1 q^{i-1} \qquad (1 \leq i \leq n),\ R_1 > 0,\ q \in R_+. \tag{23}$$

This is valid for $q > 0$, as negative resistances have no physical meaning with respect to ordinary resistors. If $n \to \infty$, then the chain is said to be self-similar. Each element of the chain that has appeared after the removal of one resistor on the left-hand side and marked with a dashed line has the same structure as the chain before the removal. The substitute resistance R_S of the chain, according to the obtained equation (8), is defined by the equation

$$R_S = \sum_{i=1}^{n} R_i. \tag{24}$$

The resistances follow a geometric progression whose sum of the n terms R_S, called a row is determined by the following formulae [4]:

$$R_S = R_1 \frac{1 - q^n}{1 - q} \quad \text{for} \quad q \neq 1, \tag{25}$$

$$R_S = nR_1 \quad \text{for} \quad q = 1. \tag{26}$$

Let's find out how R_S behaves at $n \to \infty$ for different values of q.

If $q > 0$, then from the equation (25) is obtained

$$\lim_{n\to\infty} R_S = R_1 \lim \frac{1 - q^n}{1 - q} = \infty, \quad \text{because} \quad \lim_{n\to\infty} q^n = \infty. \tag{27}$$

This implies that that such a chain has infinite resistance and in accordance with the formula (4) the electric current cannot flow through it.

If $q < 1$, then from the equation (25) is obtained

$$\lim_{n\to\infty} R_S = R_1 \lim_{n\to\infty} \frac{1 - q^n}{1 - q} = \frac{R_1}{1 - q}, \quad \text{because} \quad \lim_{n\to\infty} q^n = 0. \tag{28}$$

The implication is that the resistance of the chain is finite. The closer isq to 1, the higher is the resistance, and in accordance to the formula (4) the electric current is able to flow through it. The lower is the intensity I of the electric current, the closer is the value of q to 1.

Suppose $q = 1$, from the equation (26) is obtained

$$\lim_{n\to\infty} R_S = R_1 \lim_{n\to\infty} n = \infty \tag{29}$$

which implies that, in accordance with the equation (4), the current will not flow through such a chain, as its resistance is infinite.

Let's examine the network of n resistors in parallel, as shown in Fig. 12, if the i resistance of the resistor satisfies the condition (23), that is, $R_1 = R_1 q_i^{i-1}$ Also this circuit has a fractal structure. In accordance with (15), the substitute resistance of this circuit is determined by

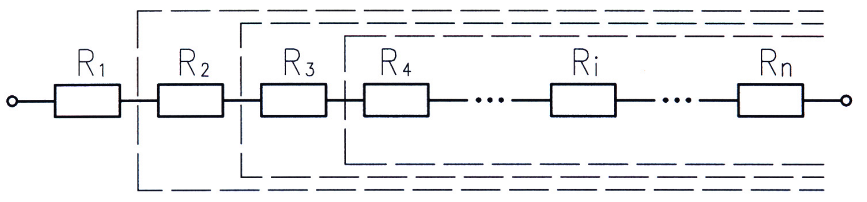

Fig. 11: The fractal network of resistors connected in series.

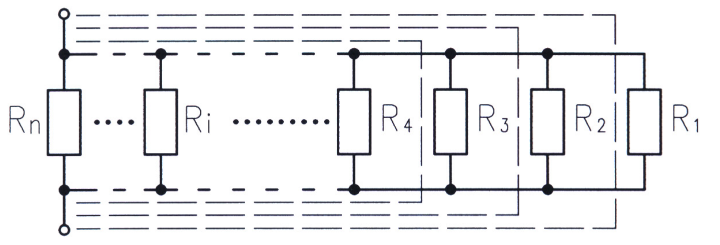

Fig. 12: The fractal network of resistors connected in parallel.

$$R_S = \frac{1}{\sum\limits_{i=1}^{n} \frac{1}{R_i}}. \tag{30}$$

To compute the sum in the denominator of the equation (30) it is worth noting the relationship

$$\frac{1}{R_1} = \frac{1}{R_1 q^{i-1}} = \left(\frac{1}{R_1}\right)\left(\frac{1}{q}\right)^{i-1}. \tag{31}$$

The equation (31) describes a geometrical progression whose i term is equal to $1/R_1$, the first term $1/R_1$, and the quotient is $1/q$. According to formula (25) and (26), the sum in the denominator of (30) equals:

$$\sum_{i=1}^{n} \frac{1}{R_1} = \frac{1}{R_1} \frac{1-\left(\frac{1}{q}\right)^n}{1-\frac{1}{q}} \quad \text{for} \quad q \neq 1, \tag{32}$$

$$\sum_{i=1}^{n} \frac{1}{R_1} = \frac{n}{R_1} \quad \text{for} \quad q = 1. \tag{33}$$

After replacing equations (32) and (33) with (30) we obtain:

$$R_S = R_1 \frac{1-\frac{1}{q}}{1-\left(\frac{1}{q}\right)^n} \quad \text{for} \quad q \neq 1, \tag{34}$$

$$R_S = \frac{R_1}{n} \quad \text{for} \quad q = 1. \tag{35}$$

Now, let us examine how R_S behaves, if $n \to \infty$ at different values of q. If $q = 1$. The formula (34) implies

$$\lim_{n\to\infty} R_S = R_1 \lim_{n\to\infty} \frac{1-\frac{1}{q}}{1-\left(\frac{1}{q}\right)^n} = R_1\left(1-\frac{1}{q}\right) \quad \text{because} \quad \lim_{n\to\infty}\left(\frac{1}{q}\right)^n = 0. \tag{36}$$

This carries an implication that, in this case, the resistance is finite and obviously positive, which enables the electric current to flow through such a circuit.

If $q < 1$, it is given from the formula (34)

$$\lim_{n\to\infty} R_S = R_1 \lim_{n\to\infty} \frac{1-\frac{1}{q}}{1-\left(\frac{1}{q}\right)^n} = 0 \quad \text{because} \quad \lim_{n\to\infty}\left(\frac{1}{q}\right)^n = \infty. \tag{37}$$

This implies that the resistance in such a circuit is zero, and according to the equation (4), this would result in an instantaneous jump to an infinitely strong current that could cause a short circuit. In this case, in accordance with the formula (23), an infinite number of decreasingly low resistances would be connected, so that eventually the resistance would reach the zero value.

Let's examine one more case, where $q = 1$. It follows from the equation (35) that

$$\lim_{n\to\infty} R_S = R \lim_{n\to\infty} \frac{1}{n} = 0. \tag{38}$$

Also in this case, there would be a short circut, as successive similar resistors would be connected, however, in increasing numbers, and eventually the network's resistance would decrease to zero.

3. The substitute capacitance of the network of fractal capacitors

Let's consider a chain of n capacitors connected in series, as in Fig. 13, in which i capacitance of the capacitor is expressed by

$$C_i = C_1 q^{i-1} (1 \le i \le n), \quad C_1 > 0, \quad q \in R_+. \tag{39}$$

By applying (17) and acting in a similar way as in the case of a chain of resistors in parallel, we obtain

$$C_S = C_1 \frac{1-\frac{1}{q}}{1-\left(\frac{1}{q}\right)^n} \quad \text{for} \quad q \ne 1, \tag{40}$$

and

$$C_S = \frac{C_1}{n} \quad \text{for} \quad q = 1. \tag{41}$$

After using (36), the border value C_S for $n \to \infty$ and $q > 1$ similarly as for the resistors in parallel is

$$C_S = C_1 \left(1 - \frac{1}{q}\right) \quad \text{for} \quad q > 1. \tag{42}$$

which means that in order to obtain the finite capaticance of such a chain it is necessary to connect capacitors with increasingly higher capacitance. It will be possible then to accumulate a finite charge on this chain. If $q < 1$, then the border value C_S, after applying the equation (37) is $C_S = 0$, and thus the chain has zero capacitance and one cannot accumulate an electric charge on it. Similarly, for $q = 1$ after applying (38), $C_S = 0$ is obtained.

Now a network of capacitors will be taken into consideration with a structure as presented in Fig. 14. By applying (18) and acting in a similar way as in the case of the resistors in series, we can derive formulas:

$$C_S = C_1 \frac{1 - q^n}{1 - q} \quad \text{for} \quad q \neq 1, \tag{43}$$

and

$$C_S = nC_1 \quad \text{for} \quad q = 1. \tag{44}$$

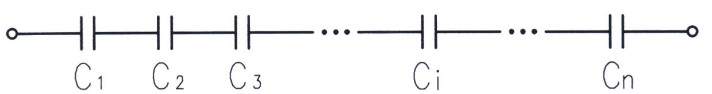

Fig. 13: The fractal network of capacitors connected in series.

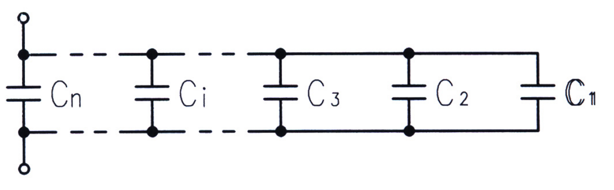

Fig. 14: The fractal network of capacitors connected in parallel.

The border values C_S for $n \to \infty$ can be obtained by analogy from formulae (27–29). There occurs then:

1. $C_S = \infty$ for $q > 1$, which means that the capacitors with increasingly high capacitance are connected, and the network has an infinite capacitance, and so an infinite charge would accumulate on it, which is physically impossible;

2. $C_S = C_1/(1-q)$, which means that capacitors of decreasing capacitances are connected and the network's capacitance as well and the charge accumulated on it are finite,

3. $C_S = \infty$ for $q = 1$, and so capacitors are connected, of constant capacitance, but their numbers, and the circuit capacitance and the charge are also infinite.

4. The substitute inductance of the network of fractal coils

It has been assumed that the inductance of the i coil in the chain of coils connected in series, see Fig. 15, is determined by the formula

$$L_i = L_1 q^{i-1} (1 \leq i \leq n), \quad L_1 > 0, \quad q \in R_+. \tag{45}$$

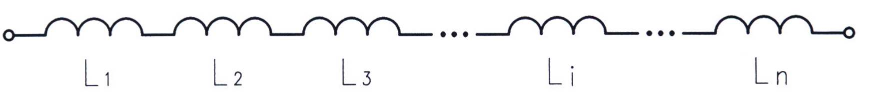

Fig. 15: The fractal network of coils connected in series.

In this case, we make use of an analogy with the series resistors and obtain:

$$L_S = L_1 \frac{1-q^n}{1-q} \quad \text{for} \quad q \neq 1, \tag{46}$$

and

$$L_S = nL_1 \quad \text{for} \quad q = 1. \tag{47}$$

The results from the analysis of the border cases for $n \to \infty$ are similar to those of resistors in series.

Suppose the i inductivity in the network of parallel coils, see Fig. 16, is satisfied by the formula (45). Drawing on analogy of resitors connected in parallel, the formulas are derived:

$$L_S = L_1 \frac{1-\frac{1}{q}}{1-\left(\frac{1}{q}\right)^n} \quad \text{for} \quad q \neq 1, \tag{48}$$

$$L_S = \frac{L_1}{n} \quad \text{for} \quad q = 1. \tag{49}$$

The analysis of the border cases for the coil network is similar to that for the parallel resistors.

For the clarity's sake and easier comparisons, computation results for the substitute resistances, capacitances ind inductivities have been gathered in Tab. 1.

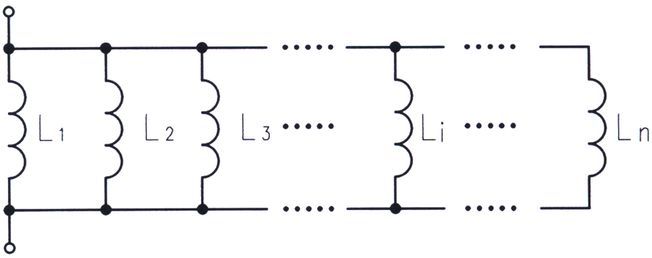

Fig. 16: The fractal network of coils connected in parallel.

Table 1. The list of the formulas on the substitute resistances R_S, capacitances C_S and inductances L_S of the fractal networks.

No.	Quantity	Type of connection	Value of q $q \in R_+$	Formula for the substitute values	Border values for $n \to \infty$	
1	Resistance R_s	In series	$q \neq 1$	$R_S = R_1 \frac{1-q^n}{1-q}$	$q<1$	$R_S = \frac{R_1}{1-q}$
					$q>1$	∞
			$q=1$	$R_S = nR_1$	∞	
		In parallel	$q \neq 1$	$R_S = R_1 \frac{1-\frac{1}{q}}{1-\left(\frac{1}{q}\right)^n}$	$q<1$	0
					$q>1$	$R_S = R_1\left(1-\frac{1}{q}\right)$
			$q=1$	$R_S = \frac{R_1}{n}$	0	
2	Capacitance C_s	In series	$q \neq 1$	$C_S = C_1 \frac{1-\frac{1}{q}}{1-\left(\frac{1}{q}\right)^n}$	$q<1$	0
					$q>1$	$C_S = C_1\left(1-\frac{1}{q}\right)$
			$q=1$	$C_S = \frac{C_1}{n}$	0	
		In parallel	$q \neq 1$	$C_S = C_1 \frac{1-q^n}{1-q}$	$q<1$	$C_S = \frac{C_1}{1-q}$
					$q>1$	∞
			$q=1$	$C_S = nC_1$	∞	
3	Inductance L_s	In series	$q \neq 1$	$L_S = L_1 \frac{1-q^n}{1-q}$	$q<1$	$L_S = \frac{L_1}{1-q}$
					$q>1$	∞
			$q=1$	$L_S = nL_1$	∞	
		In parallel	$q \neq 1$	$L_S = L_1 \frac{1-\frac{1}{q}}{1-\left(\frac{1}{q}\right)^n}$	$q<1$	0
					$q>1$	$L_S = L_1\left(1-\frac{1}{q}\right)$
			$q=1$	$L_S = \frac{L_1}{n}$	0	

5. Conclusions

1. By applying a geometric progression, we could calculate the substitute resistances, capacitances and inductances of some networks of fractal resistors, capacitors and coils, on the assumption that the resistance, capacitance and inducance ratios of consecutive elements are constant.

2. It has been demonstrated that the results obtained are correct physicswise in respect to the border case, where the numbers of n, m elements are becoming infinite $(n, m \to \infty)$.

3. Reasoning by analogy played a crucial role in obtaining fast results.

4. The method presented can also be applied to calculate substitute or resultant values of other physical quantities, e.g. equivalent focal values of the thin set, close-up lenses, or the resultant electric field strength of a series of point charges [5, 6]. Our knowledge of the values of these substitute or resultant quantities is of vital importance in the field of physics and technology.

References

[1] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, part 3, John Wiley and Sons, Inc., New York 2002.

[2] H. O. Peitgen, H. Jürgens, and D. Saute, *Chaos and Fractals, New Frontiers of Science*, part. 1, Springer, New York 1992.

[3] J. Budrewicz, *Fractals and chaos*, Science and Technology Edition, Warsaw 1996.

[4] I. Dziubiński, T. Świątkowski, *Mathematical handbook*, part. 1, Polish Scientific Edition, Warsaw 1985.

[5] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, part. 4, John Wiley and Sons, Inc., New York 2003.

[6] Ch. Kittel, *Introduction to Solid State Physics*, John Wiley and Sons, Inc., New York 1984.

Chair of Modelling the Teaching and Learning Processes
University of Łódź
Pomorska 149/153, PL-90-236 Łódź
Poland
e-mail: bedastan@uni.lodz.pl

Low Secondary School No 1
Sterlinga 24, PL-90-212 Łódź
Poland

Presented by Leszek Wojtczak at the Session of the Mathematical-Physical Commission of the Łódź Society of Sciences and Arts on December 22, 2010

ZASTĘPCZE OPORNOŚCI, POJEMNOŚCI I INDUKCYJNOŚCI PEWNYCH SIECI O STRUKTURZE FRAKTALNEJ

S t r e s z c z e n i e

W pracy obliczono zastępcze parametry, takie jak: oporności, pojemności elektryczne i indukcyjności, pewnych sieci o strukturze złożonych tych elementów. Oporniki, kondensatory i cewki indukcyjne zostały połączone szeregowo albo równolegle. Wartości wszystkich rozpatrywanych elementów tworzyły ciąg geometryczny. Przedyskutowano wyprowadzone wzory w przypadkach granicznych. Wzory te mają nie tylko znaczenie poznawcze, ale mogą być również użyteczne przy projektowaniu obwodów elektrycznych.

Rapporteurs – Referees

CONTENU DU VOLUME LXI, no. 2

1. **L. Kozma, J. Ławrynowicz, and L. Tamássy**, Obituary: Roman Stanisław Ingarden (1920–2011) .... ca. 7 pp.

2. **M. Vaccaro**, Basics of linear para-quaternionic geometry II. Decomposition of a generic subspace of a para-quaternionic Hermitian vector space .... ca. 18 pp.

3. **L. Wojtczak**, A remark on surface phenomena .... ca. 5 pp.

4. **D. Partyka**, The Generalized Fourier coefficients and extremal quasiconformal extension of a quasisymmetric automorphism of the unit circle .... ca. 14 pp.

5. **J. Garecki**, Is torsion needed in a theory of gravity? A reappraisal I. Motivation for introducing and lack of experimental evidence .... ca. 11 pp.

6. **S. Bednarek and J. Krysiak**, Application of the cylindrical leneses in educational physical experiments .... ca. 8 pp.

7. **A. Polka**, Space modelling with multidimensional vector products .... ca. 18 pp.

8. **M. Nowak-Kępczyk**, Surface segregation in binary alloy thin films in Valenta-Sukiennicki model vs. the experimental data . ca. 12 pp.

9. **A. Niemczynowicz**, The diagonal form of the Hamiltonian in a Zwanzig-type chain .... ca. 10 pp.

10. **K. Pomorski and P. Prokopow**, Numerical solutions of time-dependent Ginzburg-Landau equations for various superconducting structures .... ca. 14 pp.