EFFECTS OF FUZZY SETTING IN KOROVKIN THEORY VIA P_p-STATISTICAL CONVERGENCE

TUĞBA YURDAKADIM AND EMRE TAŞ

ABSTRACT. In this paper our aim is to approximate a function with the use of fuzzy positive linear operators when the fuzzy limit fails by defining the fuzzy analog of P_p-statistical convergence. It is effective to use this type of convergence since a sequence can still be P_p-statistical convergent while it is neither convergent nor statistically convergent. By considering fuzzy positive linear operators, we obtain Korovkin type approximation results for these operators in the sense of P_p-statistical convergence. The rate of approximation by fuzzy modulus of continuity is also presented.

Mathematics Subject Classification (2010): 03E72, 40C15, 40G15, 41A36
Key words: Fuzzy Korovkin theory, fuzzy rate, power series method, statistical convergence, fuzzy logic.

Article history:
Received: February 9, 2022
Received in revised form: May 29, 2022
Accepted: May 30, 2022

1. Introduction and Preliminaries

The class of smart boys, the class of clever students or the class of all apples which are red enough do not construct sets in the usual mathematical sense of these terms. However, the fact lies under such uncertainly defined sets play significant role in human thinking, pattern recognition, machine learning. This motivates Zadeh [30] to define fuzzy sets by assigning to each element a grade of membership ranging from 0 to 1. It is effective to use membership function to overcome the uncertainty. Later, many researchers have extended the well known concepts of classical set theory to fuzzy setting. There are also many studies on fuzzy topology since it is applicable to quantum particle physics [21], [22]. Recently the generalizations of fuzzy topology such as intuitionistic fuzzy topology, Pythagorean fuzzy topology have been studied in [10], [24], [29]. Furthermore fuzzy logic has also been used in different areas of mathematics, for example, while studying metric and topological spaces [28], matrix and linear systems [8], [25], approximation theory. Gal has presented some results dealing with approximation theory in fuzzy setting [18]. Korovkin type approximation results in fuzzy setting by using different types of convergences instead of ordinary convergence have been presented in [1], [2], [3], [9]. Statistical approximation of fuzzy trigonometric functions and fuzzy differentiable functions have been studied in [4], [5], [12], [13]. The corresponding statistical rates in the fuzzy approximation have been obtained in [14]. In ordinary convergence, all of the terms of the sequence except finite number have to belong to an arbitrarily small neighborhood of the limit. This is a critical weakness of ordinary convergence and by flexing this condition only for a majority of elements, statistical convergence has been defined. The aim of obtaining stronger results than the classical ones, different types of convergences have been defined and used in approximation theory.

In this study, by considering fuzzy positive linear operators we present some Korovkin type approximation results with the use of P_p-statistical convergence. We also obtain the rate of this approximation by fuzzy modulus of continuity. Furthermore we construct examples to show the strength of our results.

Now let us recall the basic definitions and notations.
If the limit

$$\delta(K) := \lim_{k\to\infty} \frac{1}{k+1} |\{n \leq k : n \in K\}|$$

exists then it is said to be the density of the subset $K \subseteq \mathbb{N}_0$. Here by $|.|$, we denote the number of the elements of enclosed set and $\mathbb{N}_0$ is the set of all nonnegative integers. If for every $\varepsilon > 0$, $\delta(K_\varepsilon) = 0$ where $K_\varepsilon = \{n \in \mathbb{N}_0 : |x_n - l| \geq \varepsilon\}$, then it is said that $x = (x_n)$ converges statistically to l [16], [17], [26].

Let (p_n) be a real sequence such that $p_0 > 0$, p_1, p_2, ...≥ 0, and $p(t) := \sum_{n=0}^{\infty} p_n t^n$ has radius of convergence R with $0 < R \leq \infty$. If the limit

$$\lim_{t\to R^-} \frac{1}{p(t)} \sum_{n=0}^{\infty} x_n p_n t^n = l$$

exists then it is said that $x = (x_n)$ is convergent to l in the sense of power series method [7], [20]. The next example shows that ordinary convergence is not as effective as power series method, i.e., power series method is more useful. Let $x = (1, -1, 1, -1, ...)$, $R = \infty$, $p(t) = e^t$ and for $n \geq 0$, $p_n = \frac{1}{n!}$. Then we immediately see that

$$\lim_{t\to\infty} \frac{1}{e^t} \sum_{n=0}^{\infty} \frac{x_n t^n}{n!} = \lim_{t\to\infty} \frac{1}{e^t} \sum_{n=0}^{\infty} \frac{(-1)^n t^n}{(n)!} = \lim_{t\to\infty} \frac{1}{e^t} e^{-t} = 0.$$

Hence while the sequence $x = (x_n)$ converges to 0 in the sense of power series method, it does not converge in the ordinary sense.
If $\lim x = l$ implies $P_p - \lim x = l$, then it is said that P_p is regular [7]. The regularity of power series method is equivalent to

$$\lim_{t\to R^-} \frac{p_n t^n}{p(t)} = 0$$

holds for each $n \in \mathbb{N}_0$ [7].

Let P_p be regular and $K \subset \mathbb{N}_0$. If the limit

$$\delta_{P_p}(K) := \lim_{t\to R^-} \frac{1}{p(t)} \sum_{n\in K} p_n t^n$$

exists then it is said to be the P_p-density of K.

The sequence $x = (x_n)$ of real numbers P_p-statistically converges to l if for every $\varepsilon > 0$, $\delta_{P_p}(K_\varepsilon) = 0$ that is for every $\varepsilon > 0$

$$\lim_{t\to R^-} \frac{1}{p(t)} \sum_{n\in K_\varepsilon} p_n t^n = 0.$$

An example of a sequence such that statistical convergent but not P_p-statistical convergent and an example of a sequence such that P_p-statistical convergent but not statistically convergent have been presented in [27].
If the followings are satisfied for a function $\nu : \mathbb{R} \longrightarrow [0, 1]$

- ν is normal, i.e., there exists $x_0 \in \mathbb{R}$ such that $\nu(x_0) = 1$,

- ν is convex, i.e., $\nu(\lambda x + (1-\lambda)y) \geq \min\{\nu(x), \nu(y)\},\ for\ all\ x, y \in \mathbb{R},\ \gamma \in [0, 1]$

- upper semi-continuous on $\mathbb{R}$ and
- the closure of the set $supp(\nu)$ is compact, where

$$supp(\nu) := \{x \in \mathbb{R} : \nu(x) > 0\}$$

then ν is said to be a fuzzy number and $\mathbb{R}_{\mathbb{F}}$ denotes the set of such elements.
Let

$$[\nu]^0 := \overline{\{x \in \mathbb{R} : \nu(x) > 0\}} \text{ and } [\nu]^r := \{x \in \mathbb{R} : \nu(x) \geq r\}, (0 < r \leq 1).$$

Recall from [19] that, for each $r \in [0,1]$, the set $[\nu]^r$ is an interval which is closed and bounded in $\mathbb{R}$. For any $q, s \in \mathbb{R}_{\mathbb{F}}$ and $\gamma \in \mathbb{R}$, the operations sum $q \oplus s$ and product $\gamma \odot q$ can be defined uniquely as follows:

$$[q \oplus s]^r = [q]^r + [s]^r \text{ and } [\gamma \odot q]^r = \gamma[q]^r, 0 \leq r \leq 1.$$

The interval $[q]^r$ can be denoted by $[q_-^{(r)}, q_+^{(r)}]$ where $q_-^{(r)} \leq q_+^{(r)}$ and $q_-^{(r)}, q_+^{(r)} \in \mathbb{R}$ for $r \in [0,1]$. Then define the following for $q, s \in \mathbb{R}_{\mathbb{F}}$

$$q \preceq s \leftrightarrow q_-^{(r)} \leq s_-^{(r)} \text{ and } q_+^{(r)} \leq s_+^{(r)}, \text{ for all } 0 \leq r \leq 1.$$

On the other hand consider the following metric

$$d : \mathbb{R}_{\mathbb{F}} \times \mathbb{R}_{\mathbb{F}} \longrightarrow \mathbb{R}_+$$

by

$$d(q,s) = \sup_{r \in [0,1]} \max\{|q_-^{(r)} - s_-^{(r)}|, |q_+^{(r)} - s_+^{(r)}|\}.$$

Note that $(\mathbb{R}_{\mathbb{F}}, d)$ is complete. Then for the fuzzy number valued functions f, g defined on $[a,b]$, the distance is introuced by

$$d^*(f,g) = \sup_{x \in [a,b]} \sup_{r \in [0,1]} \max\{|f_-^{(r)} - g_-^{(r)}|, |f_+^{(r)} - g_+^{(r)}|\}.$$

By using this metric, the statistical convergence has been introduced in fuzzy setting in [23] as follows: Let $(\nu_n)_{n \in \mathbb{N}_0}$ be a sequence of fuzzy numbers. If for every $\varepsilon > 0$,

$$\lim_k \frac{|n \leq k : d(\nu_n, \nu) \geq \varepsilon|}{k+1} = 0$$

holds then it is said that $(\nu_n)_{n \in \mathbb{N}_0}$ converges statistically to ν and we denote it by

$$st - \lim_n d(\nu_n, \nu) = 0.$$

Then in [3], A-statistical convergence has also been defined in fuzzy setting as follows: we say that $(\nu_n)_{n \in \mathbb{N}}$ converges A-statistically to $\nu \in \mathbb{R}_{\mathbb{F}}$ and we denote it by

$$st_A - \lim_n d(\nu_n, \nu) = 0,$$

if for every $\varepsilon > 0$

$$\lim_j \sum_{n : d(\nu_n, \nu) \geq \varepsilon} a_{jn} = 0$$

holds. If $A = C_1$, the Cesáro matrix of order one, then we get statistical convergence recalled above. Again in the case A is the identity matrix, then we get fuzzy convergence.
The main tool of the paper is P_p-statistical convergence and now we are ready to define it in fuzzy setting. If

$$\lim_{t \to R^-} \frac{1}{p(t)} \sum_{n \in K_\varepsilon} p_n t^n = 0$$

holds for every $\varepsilon > 0$ then it is denoted by $st_{P_p} - \lim d(\nu_n, \nu) = 0$ where $K_\varepsilon = \{n : d(\nu_n, \nu) \geq \varepsilon\}$.

2. Fuzzy Korovkin Theory in P_p-Statistical Sense

This section is devoted to our main results dealing with Korovkin type approximation and the P_p-statistical rate of approximation. We also provide examples to illustrate that it is still possible to approximate a function by fuzzy positive linear operators when the fuzzy limit fails. Therefore it is benefical to recall some of the well known concepts in fuzzy setting.

Let f be a function defined on $[a,b]$ with fuzzy number values. Then the fuzzy continuity of f at $x_0 \in [a,b]$ is defined as follows: if $x_n \to x_0$, then $d(f(x_n), f(x_0)) \to 0$ as $n \to \infty$. If f is continuous at every point $x \in [a,b]$, then it is said that f is fuzzy continuous on $[a,b]$. $C_{\mathbb{F}}[a,b]$ is the set of all fuzzy continuous functions on $[a,b]$. It is important to recall that $C_{\mathbb{F}}[a,b]$ is not a vector space but a cone. Now let $T : C_{\mathbb{F}}[a,b] \longrightarrow C_{\mathbb{F}}[a,b]$ be an operator. If for every $\alpha, \beta \in \mathbb{R}$, $f, g \in C_{\mathbb{F}}[a,b]$ and $x \in [a,b]$,

$$T(\alpha \odot f \oplus \beta \odot g; x) = \alpha \odot T(f;x) \oplus \beta \odot T(g;x)$$

holds then it is said that T is fuzzy linear. Also T is called fuzzy positive linear operator if it is fuzzy linear and $T(f;x) \preceq T(g;x)$ whenever $f, g \in C_{\mathbb{F}}[a,b]$, and all $x \in [a,b]$ with $f(x) \preceq g(x)$.

The following Korovkin type theorem in fuzzy setting has been given by Anastassiou [2].

Theorem 2.1. *Let T_n be fuzzy positive linear operators for every $n \in \mathbb{N}$ from $C_{\mathbb{F}}[a,b]$ into itself. Suppose that there exists a corresponding positive linear operators $\tilde{T_n}$ from $C[a,b]$ into itself with the property*

$$\{T_n(f;x)\}_{\pm}^{(r)} = \{\tilde{T_n}\}(f_{\pm}^{(r)}; x)$$

for all $x \in [a,b]$, $r \in [0,1], n \in \mathbb{N}$, $f \in C_{\mathbb{F}}[a,b]$. If

$$\lim_n ||\{\tilde{T_n}\}(x^i) - x^i|| = 0, i = 0,1,2,$$

then for all $f \in C_{\mathbb{F}}[a,b]$, we have

$$\lim_n d^*(T_n(f), f) = 0.$$

Anastassiou and Duman have given the A-statistical analog of this theorem in [3]. Now it is time to give our main result.

Theorem 2.2. *Let P_p be regular and T_n be fuzzy positive linear operators for every $n \in \mathbb{N}_0$ from $C_{\mathbb{F}}[a,b]$ into itself. Suppose that there exists a corresponding positive linear operators $\tilde{T_n}$ from $C[a,b]$ into itself with the property*

$$\{T_n(f;x)\}_{\pm}^{(r)} = \{\tilde{T_n}\}(f_{\pm}^{(r)}; x)$$

for all $x \in [a,b]$, $r \in [0,1], n \in \mathbb{N}_0$, $f \in C_{\mathbb{F}}[a,b]$. If

$$st_{P_p} - \lim_n \|\{\tilde{T_n}\}(x^i) - x^i\| = 0, i = 0,1,2,$$

then for all $f \in C_{\mathbb{F}}[a,b]$, we have

$$st_{P_p} - \lim_n d^*(T_n(f), f) = 0.$$

Proof. Let $f \in C_{\mathbb{F}}[a,b]$, $x \in [a,b]$ and $r \in [0,1]$. For every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f_{\pm}^{(r)}(y) - f_{\pm}^{(r)}(x)| < \varepsilon$ holds for every $y \in [a,b]$ satisfying $|y - x| < \delta$ since $f_{\pm}^{(r)} \in C[a,b]$. As in classical Korovkin theory, we have that

$$|f_{\pm}^{(r)}(y) - f_{\pm}^{(r)}(x)| \leq \varepsilon + 2H_{\pm}^{(r)} \frac{(y-x)^2}{\delta^2}$$

holds for all $y \in [a,b]$ where $2H_{\pm}^{(r)} := \|2f_{\pm}^{(r)}\|$. $\{\tilde{T_n}\}$ is positive and linear,

$$\begin{aligned}|\{\tilde{T_n}\}(f_{\pm}^{(r)}; x) - f_{\pm}^{(r)}(x)| &\leq \{\tilde{T_n}\}(|f_{\pm}^{(r)}(y) - f_{\pm}^{(r)}|; x) + H_{\pm}^{(r)}|\{\tilde{T_n}\}(1;x) - 1| \\ &\leq \varepsilon + (\varepsilon + H_{\pm}^{(r)}|\{\tilde{T_n}\}(1;x) - 1|) + \frac{2H_{\pm}^{(r)}}{\delta^2}|\{\tilde{T_n}\}((y-x)^2; x)|\end{aligned}$$

holds for each $n \in \mathbb{N}_0$ and it implies

$$|\{\tilde{T_n}\}(f_{\pm}^{(r)};x)-(f_{\pm}^{(r)};x)|\leq \varepsilon+(\varepsilon+H_{\pm}^{(r)}+2h^2\frac{H_{\pm}^{(r)}}{\delta^2})|\{\tilde{T_n}\}(1;x)-e_1|$$
$$+4h\frac{H_{\pm}^{(r)}}{\delta^2}|\{\tilde{T_n}\}(t;x)-x|+2\frac{H_{\pm}^{(r)}}{\delta^2})|\{\tilde{T_n}\}(t^2;x)-x^2|$$

where $h:=\max\{|a|,|b|\}$. Pick

$$H_{\pm}^{(r)}(\varepsilon):=\max\{\varepsilon+H_{\pm}^{(r)}+2h^2\frac{H_{\pm}^{(r)}}{\delta^2},4h\frac{H_{\pm}^{(r)}}{\delta^2},2\frac{H_{\pm}^{(r)}}{\delta^2}\}$$

and take supremum over $x\in[a,b]$, then we have that

$$\|\{\tilde{T_n}\}(f_{\pm}^{(r)})-f_{\pm}^{(r)}\|\leq \varepsilon+H_{\pm}^{(r)}(\varepsilon)\{\|\{\tilde{T_n}\}(1)-1\|+\|\{\tilde{T_n}\}(x)-x\|+\|\{\tilde{T_n}\}(x^2)-x^2\|\}.$$

Then, by the property in hypothesis, we obtain that

$$\begin{aligned}d^*(T_n(f),f)&=\sup_{x\in[a,b]}d(T_n(f;x)-f(x))\\&=\sup_{x\in[a,b]}\sup_{r\in[0,1]}\max\{|\{\tilde{T_n}\}(f_-^{(r)};x)-f_-^{(r)}(x)|,|\{\tilde{T_n}\}(f_+^{(r)};x)-f_+^{(r)}(x)\}\\&=\sup_{r\in[0,1]}\max\{\|\{\tilde{T_n}\}(f_-^{(r)})-(f_-^{(r)})\|,\|\{\tilde{T_n}\}(f_+^{(r)})-(f_+^{(r)})\|\}.\end{aligned}$$

Considering the above inequalities, we obtain that

$$d^*(T_n(f),f)\leq \varepsilon+H(\varepsilon)\{\|\{\tilde{T_n}\}(1)-1\|+\|\{\tilde{T_n}\}(x)-x\|+\|\{\tilde{T_n}\}(x^2)-x^2\|\}$$

where $H(\varepsilon):=\sup_{r\in[0,1]}\max\{H_-^{(r)}(\varepsilon),H_+^{(r)}(\varepsilon)\}$. Now for a given $\varepsilon^{'}$, choose $\varepsilon>0$ such that $0<\varepsilon<\varepsilon^{'}$ and also define

$$K:=\{n\in\mathbb{N}_0:d^*(T_n(f),f)\geq\varepsilon^{'}\},$$
$$K_0:=\{n\in\mathbb{N}_0:\|\{\tilde{T_n}\}(1)-1\|\geq\frac{\varepsilon^{'}-\varepsilon}{3H(\varepsilon)}\},$$
$$K_1:=\{n\in\mathbb{N}_0:\|\{\tilde{T_n}\}(x)-x\|\geq\frac{\varepsilon^{'}-\varepsilon}{3H(\varepsilon)}\},$$
$$K_2:=\{n\in\mathbb{N}_0:\|\{\tilde{T_n}\}(x^2)-x^2\|\geq\frac{\varepsilon^{'}-\varepsilon}{3H(\varepsilon)}\}.$$

Using the above inequalities, we have $K\subseteq K_0\cup K_1\cup K_2$ which implies that

$$\frac{1}{p(t)}\sum_{n\in K}p_nt^n\leq\frac{1}{p(t)}\{\sum_{n\in K_0}p_nt^n+\sum_{n\in K_1}p_nt^n+\sum_{n\in K_2}p_nt^n\}.$$

By taking limit as $0<t\to R^-$ on the both sides and using the hypothesis, we immediately obtain that

$$\lim_{0<t\to R^-}\frac{1}{p(t)}\sum_{n\in K}p_nt^n=0.$$

Hence the proof is completed. □

Example 2.3. *Let the sequences (p_n) and (a_n) defined as follows:*

$$p_n=\begin{cases}1 & , \quad n=2k\\ 0 & , \quad n=2k+1\end{cases},\qquad a_n=\begin{cases}0 & , \quad n=2k+1\\ 1 & , \quad n=2k\end{cases}.$$

One can immediately obtain that the method P_p is regular and

$$\delta_{P_p}(K_\varepsilon)=0$$

where $K_\varepsilon = \{n \in \mathbb{N}_0 : |a_n - 1| \geq \varepsilon\}$ *holds for every* $\varepsilon > 0$*. That is* $st_{P_p} - \lim a_n = 1$*. Notice that* (a_n) *is neither convergent in the ordinary sense nor statistically convergent. Now construct the fuzzy Bernstein-type operators as follows*

$$T_n^{\mathbb{F}}(f;x) = \begin{cases} a_n \odot \oplus_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \odot f(\frac{k}{n}) & , \quad n \in \mathbb{N} \\ f(x) & , \quad n = 0 \end{cases}$$

where $f \in C_{\mathbb{F}}[0,1],\ x \in [0,1]$*.*
In this case, one can also write

$$\{T_n^{\mathbb{F}}(f;x)\}_{\pm}^{r} = \{\tilde{T_n}\}(f_{\pm}^{(r)};x) = a_n \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f_{\pm}^{(r)}(\frac{k}{n})$$

where $f_{\pm}^{(r)} \in C[0,1]$*. Notice that*

$$\{\tilde{T_n}\}(1;x) = a_n,$$
$$\{\tilde{T_n}\}(t;x) = x a_n,$$
$$\{\tilde{T_n}\}(t^2;x) = [x^2 + \frac{x(1-x)}{n}]a_n.$$

Then

$$st_{P_p} - \lim_n \|\{\tilde{T_n}\}(x^i) - x^i\| = 0$$

holds for $i = 0, 1, 2$ *then*

$$st_{P_p} - \lim_n d^*(T_n^{\mathbb{F}}(f), f) = 0$$

holds for all $f \in C_{\mathbb{F}}[a,b]$ *follows from our main result. Notice that since the sequence* (a_n) *is not convergent,* $\{T_n^{\mathbb{F}}(f)\}_{n \in \mathbb{N}_0}$ *is not fuzzy convergent to* f*.*

Example 2.4. *Let* (p_n) *and* (a_n) *be defined as follows:*

$$p_n = \begin{cases} 1 & , \quad n = 2k \\ 0 & , \quad n = 2k+1 \end{cases}, \qquad a_n = \begin{cases} 0 & , \quad n = 2k \\ 1 & , \quad n = 2k+1 \end{cases}.$$

It is easy to see that P_p *is regular and*

$$K_\varepsilon = \{n \in \mathbb{N}_0 : |a_n - 0| \geq \varepsilon\} \subseteq \{n = 2k+1 : k \in \mathbb{N}_0\}$$

holds for every $\varepsilon > 0$*. Then we have*

$$\delta_{P_p}(K_\varepsilon) = \lim_{0<t \to R^-} \frac{1}{p(t)} \sum_{n \in K_\varepsilon} p_n t^n = 0$$

i.e., that (a_n) *is* P_p*-statistically convergent to 0. Construct the following fuzzy Bernstein-type operators:*

$$T_n^{\mathbb{F}}(f;x) = \begin{cases} (1+a_n) \odot \oplus_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \odot f(\frac{k}{n}) & , \quad n \in \mathbb{N} \\ f(x) & , \quad n = 0 \end{cases}$$

where $f \in C_{\mathbb{F}}[0,1],\ x \in [0,1]$*.*
Notice that since the sequence (a_n) *is not convergent,* $\{T_n^{\mathbb{F}}(f)\}_{n \in \mathbb{N}_0}$ *is not fuzzy convergent to* f *but still one can approximate* f *by* $T_n^{\mathbb{F}}(f)$ *with the use of* P_p*-statistical convergence.*

Now recall the modulus of continuity in fuzzy setting. Let f be a function defined on $[a,b]$ and fuzzy number valued. Then the fuzzy modulus of continuity of f is defined in [18] as follows:

$$w_1^{\mathbb{F}} := \sup_{x \in [a,b] : |x-y| \leq \delta} d(f(x), f(y))$$

for any $0 < \delta \leq b - a$. The rates of this approximation have been presented in [3] by this notion. Statistical rate of convergence has been defined and studied in [11], [15]. By modificating these concepts, P_p- statistical rate of convergence has been introduced in [6].

Definition 2.5. *Let (a_n) be a positive non-increasing sequence of real numbers and let P_p be regular. A sequence $x = (x_n)$ is P_p-statistically converges to the number l with rate $o\,(a_n)$ if for every $\varepsilon > 0$*

$$\lim_{0<t\to R^-}\left[\frac{1}{p\,(t)}\sum_{n:|x_n-l|\geq\varepsilon a_n} p_n t^n\right] = 0$$

and we denote it by $x_n - l = st_{P_p} - o\,(a_n)$, as $n \to \infty$.

Theorem 2.6. *Let P_p be regular and T_n be fuzzy positive linear operators $n \in \mathbb{N}_0$ from $C_{\mathbb{F}}[a,b]$ into itself. Suppose that there exists a corresponding positive linear operators $\tilde{T_n}$ of from $C[a,b]$ into itself with the property*

$$\{T_n(f;x)\}_{\pm}^{(r)} = \{\tilde{T}\}_n(f_{\pm}^{(r)};x)$$

for all $x \in [a,b]$, $r \in [0,1], n \in \mathbb{N}_0$, $f \in C_{\mathbb{F}}[a,b]$. If $(a_n),(b_n)$ are positive non-increasing sequences and also the operators $\{\tilde{T_n}\}$ satisfy the following conditions:

$$\|\{\tilde{T_n}\}(1) - 1\| = st_{P_p} - o(a_n)$$

$$w_1^{\mathbb{F}}(f,\nu_n) = st_{P_p} - o(b_n),$$

then for all $f \in C_{\mathbb{F}}[a,b]$, we have

$$d^*(T_n(f),f) = st_{P_p} - o(c_n).$$

Here $\nu_n := \sqrt{\|\{\tilde{T_n}\}(\phi)\|}$, $\phi(y) = (y-x)^2$ for each $x \in [a,b]$ and $c_n = \max\{a_n, b_n, a_n b_n\}$, for every $n \in \mathbb{N}_0$.

Proof. By Theorem 3 of [2], one can get, for each $n \in \mathbb{N}_0$ and $f \in C_{\mathbb{F}}[a,b]$, that

$$d^*(T_n(f),f) \leq H\|\{\tilde{T_n}\}(1) - 1\| + \|\{\tilde{T_n}\}(1) + 1\|w_1^{\mathbb{F}}(f,\nu_n)$$

where $H := d^*(f,\chi_0)$ and χ_0 denotes the neutral element for $\oplus$. Then we have that

$$d^*(T_n(f),f) \leq H\|\{\tilde{T_n}\}(1) - 1\| + \|\{\tilde{T_n}\}(1) + 1\|w_1^{\mathbb{F}}(f,\nu_n) + 2w_1^{\mathbb{F}}(f,\nu_n).$$

By using the similar idea in Lemma 4 in [11] and taking care of the right hand side of the following equality we obtain the desired result. Therefore the proof is completed. □

3. Conclusion

The fact lying under uncertainly defined sets play important role in human thinking, pattern recognition and machine learning and this motivates Zadeh to introduce fuzzy sets by attaching a grade of membership to each element. Since it is effective to overcome uncertainty, fuzzy theory has become an active area of research. Fuzzy logic and fuzzy settings of well-known concepts have been studied. In the present paper, by considering fuzzy positive linear operators we have obtained Korovkin type approximation results via P_p-statistical convergence. We have also studied the rate of this approximation with the use of fuzzy modulus of continuity. It is important to mention that our results are stronger than the results in the existing literature since P_p-statistical convergence is flexing the critical weakness of ordinary convergence. In order to show the strength of our results, we have provided some examples.

References

[1] M. Aiyub, K. Saini and K. Raj, *Korovkin type approximation theorem via lacunary equi-statistical convergence in fuzzy spaces*, J. Math. Computer Sci. **25** (2022), 312-321.

[2] G.A. Anastassiou, *On basic fuzzy Korovkin theory*, Stud. Univ. Babes-Bolyai Math. **50** (2005), 3-10.

[3] G.A. Anastassiou and O. Duman, *Statistical fuzzy approximation by fuzzy positive linear operators*, Comput. and Math. Appl. **55** (2008), 573-580.

[4] G.A. Anastassiou and O. Duman, *High order statistical fuzzy Korovkin theory*, Stoch. Anal. Appl. **27(3)** (2009), 543-554.

[5] G.A. Anastassiou and O. Duman, *Towards Intelligent Modelling: Statistical approximation theory*, Intelligent Systems Reference Library, Springer-Verlag, Berlin **14** (2011).

[6] Ö.G. Atlıhan, T. Yurdakadim and E. Taş, *A new approach to approximation by positive linear operators in weighted spaces*, Accepted for publication in Ukrainian J. Math.

[7] J. Boos, *Classical and Modern Methods in Summability*, Oxford University Press, 2000.

[8] J.J. Buckley, *Solving systems of linear fuzzy equations*, Fuzzy Sets and Syst. **43** (1991), 33-43.

[9] M. Burgin and O. Duman, *Approximations by linear operators in spaces of fuzzy continuous functions*, Positivity **15** (2011), 57-72.

[10] D. Coker, *An introduction to intuitionistic fuzzy topological spaces*, Fuzzy Sets and Syst. **88** (1997), 81-89.

[11] O. Duman, M.K. Khan and C. Orhan, *A-statistical convergence of approximating operators*, Math. Inequal. Appl. **6** (2003), 689-699.

[12] O. Duman and G.A. Anastassiou, *On statistical fuzzy trigonometric Korovkin theory*, J. Comput. Anal. Appl. **10(3)** (2008), 333-344.

[13] O. Duman, *Statistical fuzzy approximation to fuzzy differentiable functions by fuzzy linear operators*, Hacet. J. Math. Stat. **39(4)** (2010), 497-514.

[14] O. Duman, *Fuzzy approximation based on statistical rates*, Publ. Math. Debrecen **76(3-4)** (2010), 453-464.

[15] J.A. Fridy, H.I. Miller and C. Orhan, *Statistical rates of convergence*, Acta Sci. Math. **69** (2003), 147-157.

[16] H. Fast, *Sur la convergence statistique*, Colloq. Math. **2** (1951), 241-244.

[17] J.A. Fridy, *On statistical convergence*. Analysis **5** (1985), 301-313.

[18] S. Gal, *Approximation theory in fuzzy setting. Chapter 13 in Handbook of Analytic-Computational Methods in Applied Mathematics*, Chapman and Hall/CRC, Boca Raton (2000), 617-666.

[19] R.J. Goetschel and W. Voxman, *Elementary fuzzy calculus*, Fuzzy Sets and Syst. **18** (1986), 31-43.

[20] W. Kratz and U. Stadtmüller, *Tauberian theorems for J_p-summability*, J. Math. Anal. Appl. **139** (1989), 362-371.

[21] M.S. El Naschie, *On uncertainty of Cantorian geometry and two-slit experiment*, Chaos Solitons and Fractals **9** (1998), 517-529.

[22] M.S. El Naschie, *A review of E-infinity theory and mass spectrum of high energy particle physics*, Chaos Solitons and Fractals **19** (2004), 209-236.

[23] F. Nuray and E. Savaş, *Statistical convergence of sequences of fuzzy numbers*, Math. Slovaca. **45** (1995), 269-273.

[24] M. Olgun, M. Ünver and Ş. Yardımc ı, *Pythagorean fuzzy topological spaces*, Complex and Intelligent Systems **5** (2019), 177-183.

[25] K. Peeva, *Fuzzy linear systems*, Fuzzy Sets and Syst. **49** (1992), 339-355.

[26] T. Salat, *On statistically convergent sequences of real numbers*, Mat. Slovaca. **30** (1980), 139-150.

[27] M. Ünver and C. Orhan, *Statistical convergence with respect to power series methods and applications to approximation theory*, Numer. Funct. Anal. Optim. **40** (2019), 535-547.

[28] A. George and P. Veeramani, *On some results in fuzzy metric space*. Fuzzy Sets and Syst. **4** (1980), 395-399.

[29] R.R. Yager, *Pythagorean fuzzy subsets*, In:2013 joint IFSA world congres and NAFIPS annual meeting (IFSA/NAFIPS), IEEE, Edmenton, AB, Canada, 24-28 June 2013, 57-61.

[30] L.A. Zadeh, *Fuzzy sets* Inform. and Control **8** (1965), 338-353.

Bilecik Şeyh Edebali University, Department of Mathematics, Bilecik 11230, Turkey
Email address: tugbayurdakadim@hotmail.com

Kirşehir Ahi Evran University, Department of Mathematics, Kirşehir 40100, Turkey
Email address: emretas86@hotmail.com

THE STEINHAUS THEOREM FOR λ-REGULAR MATRICES IN ULTRAMETRIC FIELDS

P.N. NATARAJAN

Abstract. Throughout the present paper, K denotes a non-trivially valued, ultrametric (or non-archimedean) field, which is complete under the valuation of K. Sequences, infinite series and infinite matrices have their entries in K. The purpose of the present paper is to introduce λ-regular matrices and prove the Steinhaus theorem for λ-regular matrices in K.

Mathematics Subject Classification (2010): 40C05, 40D05, 40H05, 46S10.
Key words: Ultrametric (or non-archimedean) field, λ-conservative matrix, τ-multiplicative matrix, λ-regular matrix, the Steinhaus theorem.

Article history:
Received: March 9, 2022
Received in revised form: October 2, 2022
Accepted: October 3, 2022

1. Introduction and Preliminaries

In the present paper, K denotes a non-trivially valued, ultrametric (or non-archimedean) field, which is complete under the valuation of K. Entries of sequences, infinite series and infinite matrices are in K.

For a given sequence $x = \{x_k\}$ in K and an infinite matrix $A = (a_{nk})$, $a_{nk} \in K$, $n, k = 0, 1, 2, \ldots,$ define

$$(Ax)_n = \sum_{k=0}^{\infty} a_{nk}x_k,\ n = 0, 1, 2, \ldots,$$

where we suppose that the series on the right converge. $A(x) = \{(Ax)_n\}$ is called the A-transform of the sequence $x = \{x_k\}$.

If X, Y are sequence spaces, we write

$$A = (a_{nk}) \in (X, Y),$$

if $\{(Ax)_n\} \in Y$, whenever $x = \{x_k\} \in X$. In what follows, c, c_0 respectively denote the ultrametric Banach spaces of convergent and null sequences in K under the ultrametric norm

$$\|x\| = \sup_{k \geq 0} |x_k|,\ x = \{x_k\} \in c, c_0.$$

Following [1], the author of the present paper introduced the analogues in ultrametrix analysis of the concepts of λ-convergence, λ-boundedness etc. and made a study in [5, 6, 7]. We make a further study in the present paper. For an extensive study of the above concepts of λ-convergence, λ-boundedness etc. in the classical case, a standard reference in [1].

To make the paper self-contained, we recall the following definitions [5, 6, 7].

Definition 1.1. Let $\lambda = \{\lambda_n\}$ be a sequence in K such that

$$0 < |\lambda_n| \nearrow \infty,\ n \to \infty.$$

A sequence $\{x_n\}$ in K is said to be convergent with speed λ or λ-convergent if $\{x_n\} \in c$ with $\lim_{n\to\infty} x_n = s$ (say) and

$$\lim_{n\to\infty} \lambda_n(x_n - s) \text{ exists.}$$

Let c^λ denote the set of all λ-convergent sequences in K. From the definition, we have,

$$c^\lambda \subset c.$$

We note that the sequences $e_k = \{0, 0, \dots, 0, 1, 0, \dots\}$, 1 occurring in the kth place only, $k = 0, 1, 2, \dots$;

$$e = \{1, 1, 1, \dots\};$$

and

$$e^\lambda = \left\{\frac{1}{\lambda_0}, \frac{1}{\lambda_1}, \dots\right\},$$

all belong to c^λ.

Definition 1.2. A sequence $\{x_n\}$ in K is said to be bounded with speed λ or λ-bounded, if $x = \{x_n\} \in c$ with $\lim_{n\to\infty} x_n = s$ and the sequence

$$\{\lambda_n(x_n - s)\} \text{ is bounded.}$$

Let m^λ denote the set of all λ-bounded sequences in K. Note again that

$$c^\lambda \subset m^\lambda \subset c.$$

We need the following results, which can be easily proved.

Theorem 1.3 (see [4]). *$A = (a_{nk}) \in (c_0, c_0)$ if and only if*

$$\sup_{n,k} |a_{nk}| < \infty; \tag{1.1}$$

and

$$\lim_{n\to\infty} a_{nk} = 0, \ k = 0, 1, 2, \dots. \tag{1.2}$$

Theorem 1.4 (Kojima-Schur) (see [4]). *$A = (a_{nk}) \in (c, c)$, i.e., A is conservative or convergence preserving if and only if (1.1) holds,*

$$\lim_{n\to\infty} a_{nk} = a_k, \ k = 0, 1, 2, \dots; \tag{1.3}$$

and

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} = a. \tag{1.4}$$

In such a case,

$$\lim_{n\to\infty} (Ax)_n = \sum_{k=0}^{\infty} a_k(x_k - s) + sa, \tag{1.5}$$

where $x = \{x_k\} \in c$ with $\lim_{k\to\infty} x_k = s$.

Let $\mu = \{\mu_n\}$ be a sequence in K such that

$$0 < |\mu_n| \nearrow \infty, \ n \to \infty.$$

The following characterization of the matrix class (c^λ, c^μ) was proved in [5].

Theorem 1.5 (see [5]). *$A = (a_{nk}) \in (c^\lambda, c^\mu)$ if and only if*

$$A(e), A(e^\lambda), A(e_k) \in c^\mu, \ \ k = 0, 1, 2, \ldots; \tag{1.6}$$

$$\sup_{n,k} \left| \frac{a_{nk}}{\lambda_k} \right| < \infty; \tag{1.7}$$

and

$$\sup_{n,k} \left| \frac{\mu_n(a_{nk} - a_k)}{\lambda_k} \right| < \infty, \tag{1.8}$$

where $\lim_{n\to\infty} a_{nk} = a_k$, $k = 0, 1, 2, \ldots$.

Definition 1.6. If $A = (a_{nk}) \in (c^\lambda, c^\lambda)$, A is said to be λ-convergence preserving or λ-conservative.

Remark 1.7. Note that we get a characterization of λ-conservative matrices by putting $\mu = \lambda$, i.e., $\mu_n = \lambda_n$, $n = 0, 1, 2, \ldots$ in Theorem 1.5. Thus $A = (a_{nk})$ is λ-conservative if and only if

$$A(e), A(e^\lambda), A(e_k) \in c^\lambda, \ \ k = 0, 1, 2, \ldots; \tag{1.9}$$

(1.7) holds and

$$\sup_{n,k} \left| \frac{\lambda_n(a_{nk} - a_k)}{\lambda_k} \right| < \infty, \tag{1.10}$$

where $\lim_{n\to\infty} a_{nk} = a_k$, $k = 0, 1, 2, \ldots$.

We need the following characterization too (see [7]).

Theorem 1.8. *$A = (a_{nk}) \in (m^\lambda, c^\mu)$ if and only if*

$$A(e), A(e_k) \in c^\mu, \ \ k = 0, 1, 2, \ldots; \tag{1.11}$$

$$\lim_{k\to\infty} \frac{a_{nk}}{\lambda_k} = 0, \ \ n = 0, 1, 2, \ldots; \tag{1.12}$$

$$\lim_{n\to\infty} \sup_{k\geq 0} \left| \frac{a_{n+1,k} - a_{nk}}{\lambda_k} \right| = 0; \tag{1.13}$$

$$\lim_{k\to\infty} \frac{a_{n,k} - a_k}{\lambda_k} = 0, \ \ n = 0, 1, 2, \ldots; \tag{1.14}$$

and

$$\lim_{n\to\infty} \sup_{k\geq 0} \left| \frac{\mu_{n+1}(a_{n+1,k} - a_k) - \mu_n(a_{nk} - a_k)}{\lambda_k} \right| = 0. \tag{1.15}$$

2. Characterization of λ-regular matrices and Steinhaus theorem for λ-regular matrices

Definition 2.1. A conservative matrix $A = (a_{nk})$ is said to be τ-multiplicative if there exists $\tau \in K$ such that

$$\lim_{n\to\infty} (Ax)_n = \tau \lim_{k\to\infty} x_k,$$

$x = \{x_k\} \in c$.

Theorem 2.2. *The matrix $A = (a_{nk})$ is τ-multiplicative if and only if $A \in (c_0, c_0)$ and*

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} a_{nk} = \tau. \tag{2.1}$$

Proof. Let $A = (a_{nk})$ be τ-multiplicative, i.e.,

$$\lim_{n\to\infty}(Ax)_n = \tau \lim_{k\to\infty} x_k,$$

$x = \{x_k\} \in c$. If $x = \{x_k\} \in c_0$, then

$$\lim_{n\to\infty}(Ax)_n = \tau.0 = 0.$$

So $A \in (c_0, c_0)$. For the sequence $x = \{x_k\}$, $x_k = 1$, $k = 0, 1, 2, \ldots,$ $\lim_{k\to\infty} x_k = 1$. For this sequence,

$$\lim_{n\to\infty}(Ax)_n = \tau.1 = \tau,$$

$$\text{i.e., } \lim_{n\to\infty}\sum_{k=0}^{\infty} a_{nk} = \tau,$$

i.e., (2.1) holds.

Conversely, let $A \in (c_0, c_0)$ and (2.1) holds. Let $x = \{x_k\} \in c$ with $\lim_{k\to\infty} x_k = s$ (say). Consider the sequence $y = \{y_k\}$, where $y_k = x_k - s$, $k = 0, 1, 2, \ldots.$ Then $y = \{y_k\} \in c_0$. Since $A \in (c_0, c_0)$,

$$\begin{aligned}
&\lim_{n\to\infty}(Ay)_n = 0,\\
&\text{i.e., } \lim_{n\to\infty}\sum_{k=0}^{\infty} a_{nk}y_k = 0,\\
&\text{i.e., } \lim_{n\to\infty}\sum_{k=0}^{\infty} a_{nk}(x_k - s) = 0,\\
&\text{i.e., } \lim_{n\to\infty}\sum_{k=0}^{\infty} a_{nk}x_k - s\sum_{k=0}^{\infty} a_{nk} = 0,\\
&\text{i.e., } \lim_{n\to\infty}(Ax)_n - s\tau = 0, \text{ using (2.1)},\\
&\text{i.e., } \lim_{n\to\infty}(Ax)_n = \tau s,\\
&\text{i.e., } \lim_{n\to\infty}(Ax)_n = \tau \lim_{k\to\infty} x_k,
\end{aligned}$$

i.e., A is τ-multiplicative, completing the proof of the theorem. □

Remark 2.3. $A = (a_{nk})$ is regular if and only if it is 1-multiplicative.

Definition 2.4. A λ-conservative matrix $A = (a_{nk})$ is said to be λ-regular if for any $x = \{x_k\} \in c^{\lambda}$,

$$\lim_{n\to\infty}(Ax)_n = 0 \text{ if and only if } \lim_{n\to\infty} x_n = 0; \tag{2.2}$$

and

$$\lim_{n\to\infty} d_n(x) = 0 \text{ if and only if } \lim_{n\to\infty} b_n(x) = 0, \tag{2.3}$$

where,

$$b_n(x) = \lambda_n\left\{x_n - \lim_{n\to\infty} x_n\right\}$$

and

$$d_n(x) = \lambda_n\left\{(Ax)_n - \lim_{n\to\infty}(Ax)_n\right\}.$$

We now have the following characterization of λ-regular matrices in K. The proof of this characterization is very similar to that of its analogue in the classical case (for details, see [2]).

Theorem 2.5. *A matrix $A = (a_{nk})$ is λ-regular if and only if (1.2) holds;*

$$A(e) \in c^{\lambda} \backslash z^{\lambda} \tag{2.4}$$

and the infinite matrix $B = (b_{nk})$, defined by

$$b_{nk} = \frac{\lambda_n a_{nk}}{\lambda_k}, \ n, k = 0, 1, 2, \ldots$$

is τ-multiplicative, where τ is defined by

$$\tau = \lim_{n\to\infty} \lambda_n \sum_{k=0}^{\infty} \frac{a_{nk}}{\lambda_k} \neq 0,$$

$$z^{\lambda} = \left\{ x = \{x_k\} \in c_0 \middle/ \{b_n(x)\} \in c \right\}.$$

As a consequence of Theorem 1.8 and Theorem 2.5, we prove the Steinhaus theorem for λ-regular matrices in K.

Theorem 2.6. *A λ-regular matrix $A = (a_{nk})$ cannot belong to the class $(m^{\lambda}, c^{\lambda})$.*

Proof. Suppose $A = (a_{nk})$ is λ-regular and belongs to the class $(m^{\lambda}, c^{\lambda})$. In view of (1.14) with $\mu_n = \lambda_n$, $n = 0, 1, 2, \ldots,$

$$\lambda_{n+1} a_{n+1,k} - \lambda_n a_{nk} \to 0, n \to \infty, \text{ uniformly with respect to } k,$$
$$\text{i.e., } \lambda_n a_{n,k} \to l \text{ (say)}, n \to \infty, \text{ uniformly with respect to } k.$$

However, for $x = e_k$, $k = 0, 1, 2, \ldots,$

$$\lim_{n\to\infty} (e_k)_n = 0.$$

Since A is λ-regular,

$$\lim_{n\to\infty} (Ae_k)_n = 0.$$

Also for this sequence $x = e_k$,

$$\lim_{n\to\infty} b_n(x) = 0.$$

Since A is λ-regular,

$$\lim_{n\to\infty} d_n(x) = 0,$$
$$\text{i.e., } \lim_{n\to\infty} \lambda_n (Ae_k)_n = 0,$$
$$\text{i.e., } \lim_{n\to\infty} \lambda_n a_{nk} = 0, \ k = 0, 1, 2, \ldots.$$

Consequently,

$$\lambda_n a_{nk} \to 0, n \to \infty, \text{ uniformly with respect to } k. \tag{2.5}$$

Now,

$$\tau = \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{\lambda_n a_{nk}}{\lambda_k} = \sum_{k=0}^{\infty} \lim_{n\to\infty} \frac{\lambda_n a_{nk}}{\lambda_k} = 0, \text{ using (2.5)},$$

which is a contradiction, proving the theorem. □

References

[1] A. Aasma, H. Dutta and P.N. Natarajan, *An Introductory Course in Summability Theory*, Wiley, 2017.

[2] A. Aasma and P.N. Natarajan, *Matrix transforms between sequence spaces defined by speeds of convergence*, Filomat (accepted for publication).

[3] P.N. Natarajan, *The Steinhaus theorem for Toeplitz matrices in non-archimedean fields*, Coment. Math. Pace Mat. **20** (1978), 417–422.

[4] P.N. Natarajan, *Functional Analysis and Summability*, CRC Press, 2021.

[5] P.N. Natarajan, *On convergence and summability with speed in ultrametric fields*, An. Stiint. Univ. Al. I. Cuza Iasi. Mat. (N.S.) (accepted for publication).

[6] P.N. Natarajan, *On boundedness with speed λ in ultrametric fields* (submitted for publication).

[7] P.N. Natarajan, *Characterization of some new matrix classes over ultrametric fields* (submitted for publication).

Old No. 2/3, New No. 3/3, Second Main Road, R.A. Puram, Chennai 600 028, India
Email address: `pinnangudinatarajan@gmail.com`

PELL AND PELL-LUCAS, AND MODIFIED PELL NUMBERS WITH AN EXPONENTIAL GROWER FACTOR

XHEVAT Z. KRASNIQI

Abstract. In this paper we have generalized the classical Pell, Pell-Lucas, and Modified Pell numbers. The new numbers are named Pell, Pell-Lucas, and Modified Pell numbers with an exponential grower factor, respectively. Moreover, we have listed their first ten terms, then we have found the families of generating functions with some of their particular graphs, their Binet's formulae, their related basic identities and sums. Our results covers all basic results for classical Pell numbers, Pell-Lucas numbers, and Modified Pell numbers obtain previously by others.

Mathematics Subject Classification (2020): 11B37, 11B39.
Key words: Pell number, Pell-Lucas number, Modified Pell number, Binet's formulae, d'Ocagne's identitity.

Article history:
Received: September 24, 2022
Received in revised form: October 18, 2022
Accepted: October 20, 2022

1. Introduction

Since the nth Fibonacci number F_n is defined recursively by

$$F_n = F_{n-1} + F_{n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.1}$$

with initial terms $F_0 = F_1 = 1$, there are several other well-known of the Fibonacci type numbers defined by linear recurrence relation.

Let us recall some of them.

1° Lucas numbers:

$$L_n = L_{n-1} + L_{n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.2}$$

with initial terms $F_0 = 2$, $F_1 = 1$.

2° Pell numbers:

$$P_n = 2P_{n-1} + P_{n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.3}$$

with initial terms $P_0 = 0$, $P_1 = 1$.

3° Pell-Lucas numbers:

$$Q_n = 2Q_{n-1} + Q_{n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.4}$$

with initial terms $Q_0 = Q_1 = 2$.

4° Let $k \geq 1$ be an integer. The k-Pell numbers (see [1]):

$$P_{k,n} = 2P_{k,n-1} + kP_{k,n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.5}$$

with initial terms $P_{k,0} = 0$, $P_{k,1} = 1$.

5° Generalization of one-parameter of Pell numbers (see [5]):

$$P_{k,n} = kP_{k,n-1} + (k-1)P_{k,n-2}, \quad k \geq 2, \quad n \in \{2, 3, \dots\}, \tag{1.6}$$

with initial terms $P_{k,0} = 0$, $P_{k,1} = 1$.

6° Also, the generalized one-parameter Pell-Lucas numbers (see [5]) are defined by:

$$Q_{k,n} = kQ_{k,n-1} + (k-1)Q_{k,n-2}, \quad k \geq 2, \quad n \in \{2, 3, \dots\}, \tag{1.7}$$

with initial terms $Q_{k,0} = Q_{k,1} = 2$.

7° In the accessible literature we also encounter (not so rarely) the so-called modified Pell numbers (see [3]):

$$q_n = 2q_{n-1} + q_{n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.8}$$

with initial terms $q_0 = q_1 = 1$.

In this paper, we are going to introduce and study as well a new two-parameter generalization of the Pell numbers, the Pell-Lucas numbers, and the modified Pell numbers which is the main objective of this paper.

Let $r \geq 1$ be an integer and $a \geq 1$ a real number with same meaning throughout this paper. We define recursively, the two-parameter generalized of the Pell numbers $P_{a,r,n}$ (we named these *Pell numbers with an exponential grower factor*), as in the sequel:

$$P_{a,r,n} = 2a^r P_{a,r,n-1} + a^{r-1} P_{a,r,n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.9}$$

with initial terms $P_{a,r,0} = 0$, $P_{a,r,1} = 1$.

In the following, we recursively the two-parameter generalized of the Pell-Lucas numbers $Q_{a,r,n}$ (we named these *Pell-Lucas numbers with an exponential grower factor*) by:

$$Q_{a,r,n} = 2a^r Q_{a,r,n-1} + a^{r-1} Q_{a,r,n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.10}$$

with initial terms $Q_{a,r,0} = Q_{a,r,1} = 2$.

The two-parameter generalized of the Pell-Lucas numbers $q_{a,r,n}$ (we named these *Pell numbers with an exponential grower factor*) by:

$$q_{a,r,n} = 2a^r q_{a,r,n-1} + a^{r-1} q_{a,r,n-2}, \quad n \in \{2, 3, \dots\}, \tag{1.11}$$

with initial terms $q_{a,r,0} = q_{a,r,1} = 1$.

One can observe that putting $a = 1$ in our new numbers $P_{r,a,n}$, $Q_{r,a,n}$, and $q_{r,a,n}$ we obviously obtain the classical Pell numbers, the classical Pell-Lucas numbers, and the modified classical Pell numbers, respectively.

Next, in Table 1, we have shown the first terms of the sequence $\{P_{r,a,n}\}$ for $a = 2$, $r \in \{1, 2, 3, 4\}$, and $n \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

	First ten Pell numbers with exponential grow factor 2^r, $r \in \{1, 2, 3, 4\}$									
n	0	1	2	3	4	5	6	7	8	9
$P_{1,r,n}$	0	1	2	5	12	29	70	169	408	985
$P_{2,1,n}$	0	1	4	17	72	305	1292	5473	23184	98209
$P_{2,2,n}$	0	1	8	68	560	4616	38048	313616	2585024	21307424
$P_{2,3,n}$	0	1	16	260	4224	68624	1114880	18112576	294260736	4780622080
$P_{2,4,n}$	0	1	32	1032	33280	540736	17569792	702791680	23613800448	764637347840

TABLE 1

Remark 1.1. *Note that numbers in third row of the Table 1 are Pell's numbers. Also, it is interesting to observe (in fourth column of the present table) that the sequence* $P_{1,r,2}$, $P_{2,1,2}$, $P_{2,2,2}$, $P_{2,3,2}$, $P_{2,4,3}, \dots,$ *forms a geometric sequence with its quotient* 2.

Table 2 shows the first terms of the sequence $\{Q_{r,a,n}\}$ for $a = 2$, $r \in \{1, 2, 3, 4\}$, and $n \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

	First ten Pell-Lucas numbers with exponential grow factor 2^r, $r \in \{1,2,3,4\}$									
n	0	1	2	3	4	5	6	7	8	9
$Q_{1,r,n}$	2	2	6	14	34	82	198	478	1154	2786
$Q_{2,1,n}$	2	2	10	42	178	754	3194	13530	57314	242786
$Q_{2,2,n}$	2	2	20	164	1352	11144	91856	757136	6240800	51440672
$Q_{2,3,n}$	2	2	40	648	7648	124960	2029952	32979072	535784960	8704475648
$Q_{2,4,n}$	2	2	80	2576	83072	2678912	86389760	2785903616	89840033792	2897168310272

TABLE 2

Remark 1.2. *Note that numbers in third row of the Table 2 are Pell-Lucas' numbers. Also, it is interesting to observe (fourth column of the present table) that the sequence* $Q_{2,1,2}$, $Q_{2,2,2}$, $Q_{2,3,2}$, $Q_{2,4,2}, \ldots,$ *forms a geometric sequence with its quotient* 2.

The recurrences (1.9), (1.10) and (1.11) generate their characteristic equation

$$s^2 - 2a^r s - a^{r-1} = 0. \tag{1.12}$$

Equation (1.12) has two distinct roots

$$s_1 = a^r - \sqrt{a^{2r} + a^{r-1}}, \tag{1.13}$$

and

$$s_2 = a^r + \sqrt{a^{2r} + a^{r-1}}. \tag{1.14}$$

Here we easy can conclude that $s_1 < 0 < s_2$ and consequently $|s_1| < |s_2|$. Also, the following relationship between s_1 and s_2 hold true:

$$s_1 + s_2 = 2a^r, \tag{1.15}$$

$$s_1 - s_2 = -2\sqrt{a^{2r} + a^{r-1}}, \tag{1.16}$$

and

$$s_1 s_2 = -a^{r-1}. \tag{1.17}$$

These equalities are very usable and will be used later in this paper.

Next section is devoted to the generating functions of Pell numbers $P_{a,r,n}$, Pell-Lucas numbers $Q_{a,r,n}$, and modified Pell numbers $q_{a,r,n}$ with an exponential grower factor, respectively.

2. GENERATING FUNCTIONS OF SEQUENCES $\{P_{a,r,n}\}$, $\{Q_{a,r,n}\}$, AND $\{q_{a,r,n}\}$

Let us find first the generating functions of Pell numbers $P_{a,r,n}$ with an exponential grower factor. Assume that $P_{a,r,k}$ are coefficients of some power series with center at the origin and $f_{a,r}(x)$ are the sums of the these series, i.e.,

$$f_{a,r}(x) = \sum_{k=0}^{\infty} P_{a,r,k} x^k. \tag{2.1}$$

So, the analytic functions $f_{a,r}(x)$ are generating functions for the sequences $\{P_{a,r,k}\}$. Based on (1.9) and taking into account the initial conditions $P_{a,r,0} = 0$ and $P_{a,r,1} = 1$, we get

$$
\begin{aligned}
f_{a,r}(x) &= \sum_{k=0}^{\infty} P_{a,r,k}x^k \\
&= P_{a,r,0} + P_{a,r,1}x + \sum_{k=2}^{\infty}\left[2a^r P_{a,r,k-1} + a^{r-1}P_{a,r,k-2}\right]x^k \\
&= x + 2a^r x\sum_{k=2}^{\infty} P_{a,r,k-1}x^{k-1} + a^{r-1}x^2\sum_{k=2}^{\infty} P_{a,r,k-2}x^{k-2} \\
&= x + 2a^r x\sum_{k=0}^{\infty} P_{a,r,k}x^k + a^{r-1}x^2\sum_{k=0}^{\infty} P_{a,r,k}x^k.
\end{aligned}
$$

So, by (2.1) we find that

$$f_{a,r}(x) = x + 2a^r x f_{a,r}(x) + a^{r-1}x^2 f_{a,r}(x),$$

and thus, the generating functions for the sequences $\{P_{a,r,k}\}$ are

$$f_{a,r}(x) = \frac{x}{1-2a^r x - a^{r-1}x^2}, \quad (a \geq 1,\ r \geq 1). \tag{2.2}$$

The graphs of the functions $f_{a,r}(x)$ for various values of a and r are shown in Figure 1.

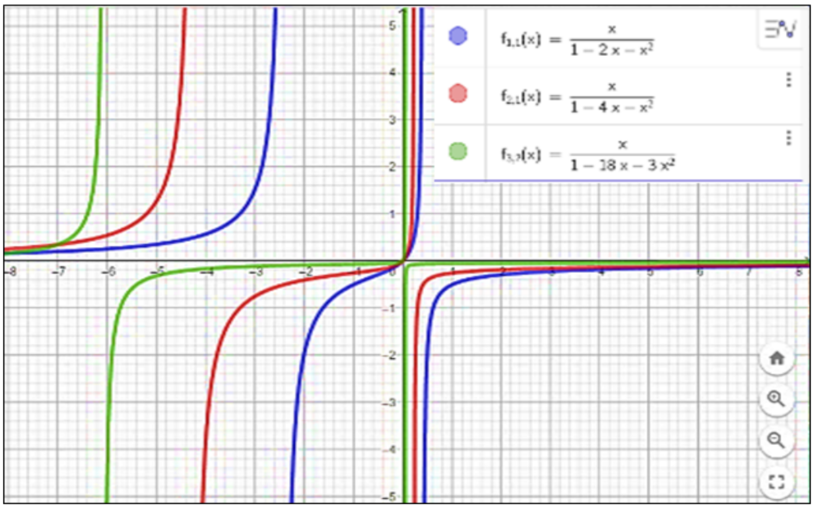

FIGURE 1. Graphs of generating functions: $f_{1,1}(x)$(blue), $f_{2,1}(x)$(red), and $f_{3,2}(x)$(green).

Further, we are going to find the generating functions for the sequences $\{Q_{a,r,k}\}$. Indeed, similarly we suppose that $Q_{a,r,k}$ are coefficients of some power series with center at the origin and $g_{a,r}(x)$ are the sums of the these series, i.e.,

$$g_{a,r}(x) = \sum_{k=0}^{\infty} Q_{a,r,k}x^k. \tag{2.3}$$

Using (1.10) and the initial conditions $Q_{a,r,0} = Q_{a,r,1} = 2$, we obtain

$$\begin{aligned} g_{a,r}(x) &= \sum_{k=0}^{\infty} Q_{a,r,k}x^k \\ &= Q_{a,r,0}x^0 + Q_{a,r,1}x + \sum_{k=2}^{\infty} \left[2a^r Q_{a,r,k-1} + a^{r-1}Q_{a,r,k-2}\right] x^k \\ &= 2 + 2x + 2a^r x \sum_{k=2}^{\infty} Q_{a,r,k-1}x^{k-1} + a^{r-1}x^2 \sum_{k=2}^{\infty} Q_{a,r,k-2}x^{k-2} \\ &= 2 + 2x + 2a^r x \sum_{k=0}^{\infty} Q_{a,r,k}x^k + a^{r-1}x^2 \sum_{k=0}^{\infty} Q_{a,r,k}x^k. \end{aligned}$$

So, by (2.3) we find that

$$g_{a,r}(x) = 2(1+x) + 2a^r x g_{a,r}(x) + a^{r-1}x^2 g_{a,r}(x),$$

and thus, the generating functions for the sequences $\{Q_{a,r,k}\}$ are

$$g_{a,r}(x) = \frac{2(1+x)}{1 - 2a^r x - a^{r-1}x^2}, \quad (a \geq 1,\ r \geq 1). \tag{2.4}$$

Finally, we find the generating functions for the sequences $\{q_{a,r,k}\}$. Indeed, let $q_{a,r,k}$ be the coefficients of some power series with center at the origin and $h_{a,r}(x)$ are the sums of the these series, i.e.,

$$h_{a,r}(x) = \sum_{k=0}^{\infty} q_{a,r,k}x^k. \tag{2.5}$$

Using (1.11) and the initial conditions $q_{a,r,0} = q_{a,r,1} = 2$, we obtain

$$\begin{aligned} h_{a,r}(x) &= \sum_{k=0}^{\infty} q_{a,r,k}x^k \\ &= q_{a,r,0}x^0 + q_{a,r,1}x + \sum_{k=2}^{\infty} \left[2a^r q_{a,r,k-1} + a^{r-1}q_{a,r,k-2}\right] x^k \\ &= 1 + x + 2a^r x \sum_{k=2}^{\infty} q_{a,r,k-1}x^{k-1} + a^{r-1}x^2 \sum_{k=2}^{\infty} q_{a,r,k-2}x^{k-2} \\ &= 1 + x + 2a^r x \sum_{k=0}^{\infty} q_{a,r,k}x^k + a^{r-1}x^2 \sum_{k=0}^{\infty} q_{a,r,k}x^k. \end{aligned}$$

So, by (2.5) we find that

$$h_{a,r}(x) = 1 + x + 2a^r x h_{a,r}(x) + a^{r-1}x^2 h_{a,r}(x),$$

and thus, the generating functions for the sequences $\{q_{a,r,k}\}$ are

$$h_{a,r}(x) = \frac{1+x}{1 - 2a^r x - a^{r-1}x^2}, \quad (a \geq 1,\ r \geq 1). \tag{2.6}$$

Of course, if we take $a = 1$ in (2.2), (2.4), and (2.6) we clearly obtain functions

$$f(x) = \frac{x}{1 - 2x - x^2},$$

$$g(x) = \frac{2(1+x)}{1 - 2x - x^2},$$

and

$$h(x) = \frac{1+x}{1-2x-x^2},$$

which, indeed, are the generating functions of the classical Pell numbers (1.3), classical Pell-Lucas numbers (1.4), and the modified classical Pell numbers (1.8), respectively.

3. Some explicit formulae

Here in this section, we are going to prove some very useful explicit formulae for the general terms of the sequences $\{P_{a,r,k}\}$, $\{Q_{a,r,k}\}$, and $\{q_{a,r,k}\}$. Then, we will apply them to derive several identities. To do this, we start with the following theorem.

Theorem 3.1 (Generalized Binet's formulae). *The nth terms of the sequences $\{P_{a,r,n}\}$, $\{Q_{a,r,n}\}$, and $\{q_{a,r,n}\}$ are of the form*

$$P_{a,r,n} = \frac{s_1^n - s_2^n}{s_1 - s_2}, \tag{3.1}$$

$$Q_{a,r,n} = \frac{2(1-s_2)}{s_1-s_2}s_1^n - \frac{2(1-s_1)}{s_1-s_2}s_2^n, \tag{3.2}$$

and

$$q_{a,r,n} = \frac{1-s_2}{s_1-s_2}s_1^n - \frac{1-s_1}{s_1-s_2}s_2^n, \tag{3.3}$$

where s_1 and s_2 are given in (1.13) and (1.14), respectively.

Proof. For the distinct roots s_1, s_2 of the equation (1.2), the numbers s_1^n, s_2^n are linearly invariant and form the basis for the space of all solutions of the equations (1.9), (1.10), and (1.11). Whence,

$$P_{a,r,n} = ps_1^n + qs_2^n, \quad p, q \in \mathbb{R}, \tag{3.4}$$

is the solution of the recurrence (1.9). If we put $n = 0$ and $n = 1$ in (3.4) we get $p + q = 0$ and $ps_1 + qs_2 = 1$. Subsequently, we find that $p = \frac{1}{s_1-s_2}$ and $q = -\frac{1}{s_1-s_2}$. Hence, putting these values to (3.4) we immediately obtain (3.1).

Similarly, the solution of the recurrence (1.10) is

$$Q_{a,r,n} = us_1^n + vs_2^n, \quad u, v \in \mathbb{R}. \tag{3.5}$$

If we put $Q_{a,r,0} = 2$ and $Q_{a,r,1} = 2$ in (3.5) we get $u + v = 2$ and $us_1 + vs_2 = 2$. Whence, we find that $u = \frac{2(1-s_2)}{s_1-s_2}$ and $v = -\frac{2(1-s_1)}{s_1-s_2}$. Thus, putting these values to (3.5) we immediately obtain (3.2).

At the end we prove (3.3). The solution of the recurrence (1.11) is

$$q_{a,r,n} = ts_1^n + ws_2^n, \quad t, w \in \mathbb{R}. \tag{3.6}$$

If we put $q_{a,r,0} = 1$ and $P_{a,r,1} = 1$ in (3.6) we get $t + w = 1$ and $ts_1 + ws_2 = 1$. Hence, we find that $t = \frac{1-s_2}{s_1-s_2}$ and $w = -\frac{1-s_1}{s_1-s_2}$. Thus, putting these values to (3.6) we immediately obtain (3.3).

The proof is completed. □

Using relations (1.13) and (1.14) we have:

Corollary 3.2. *The nth terms of the sequences $\{P_{a,r,n}\}$, $\{Q_{a,r,n}\}$, and $\{q_{a,r,n}\}$ are of the form*

$$P_{a,r,n} = \frac{\left(a^r + \sqrt{a^{2r}+a^{r-1}}\right)^n - \left(a^r - \sqrt{a^{2r}+a^{r-1}}\right)^n}{2\sqrt{a^{2r}+a^{r-1}}},$$

$$\begin{aligned} Q_{a,r,n} &= \frac{\left(1-a^r+\sqrt{a^{2r}+a^{r-1}}\right)\left(a^r+\sqrt{a^{2r}+a^{r-1}}\right)^n}{\phantom{\sqrt{a^{2r}}}} \cdots \\ &\quad \cdots \frac{-\left(1-a^r-\sqrt{a^{2r}+a^{r-1}}\right)\left(a^r-\sqrt{a^{2r}+a^{r-1}}\right)^n}{\sqrt{a^{2r}+a^{r-1}}}, \end{aligned}$$

and

$$q_{a,r,n} = \frac{\left(1-a^r+\sqrt{a^{2r}+a^{r-1}}\right)\left(a^r+\sqrt{a^{2r}+a^{r-1}}\right)^n}{} \cdots$$
$$\cdots \frac{-\left(1-a^r-\sqrt{a^{2r}+a^{r-1}}\right)\left(a^r-\sqrt{a^{2r}+a^{r-1}}\right)^n}{2\sqrt{a^{2r}+a^{r-1}}}.$$

Another consequence, for $a = 1$ and $r = 1$, is:

Corollary 3.3 (**Classical Binet's formulae**). *The nth terms of the sequences $\{P_n\}$, $\{Q_n\}$, and $\{q_n\}$ are of the form*

$$P_n = \frac{\sqrt{2}\left[\left(1+\sqrt{2}\right)^n-\left(1-\sqrt{2}\right)^n\right]}{4},$$
$$Q_n = \left(1+\sqrt{2}\right)^n+\left(1-\sqrt{2}\right)^n,$$

and

$$q_n = 2\left[\left(1+\sqrt{2}\right)^n+\left(1-\sqrt{2}\right)^n\right].$$

Next we are going to show the relationships between nth terms of the sequences:

(1) $\{P_{a,r,n}\}$ and $\{Q_{a,r,n}\}$,
(2) $\{P_{a,r,n}\}$ and $\{q_{a,r,n}\}$, and
(3) $\{Q_{a,r,n}\}$ and $\{q_{a,r,n}\}$.

Corollary 3.4. *For all $a \geq 1$, $r \geq 1$, and $n \geq 0$, we have*

(i) $Q_{a,r,n} = 2\left(P_{a,r,n} + a^{r-1}P_{a,r,n-1}\right)$,
(ii) $q_{a,r,n} = P_{a,r,n} + a^{r-1}P_{a,r,n-1}$, *and*
(iii) $Q_{a,r,n} = 2q_{a,r,n}$.

Proof. Based on Theorem 3.1 and (1.17) we have

$$\begin{aligned} Q_{a,r,n} &= \frac{2(1-s_2)}{s_1-s_2}s_1^n - \frac{2(1-s_1)}{s_1-s_2}s_2^n \\ &= \frac{2}{s_1-s_2}\left[(s_1^n-s_2^n) - s_1s_2(s_1^{n-1}-s_2^{n-1})\right] \\ &= 2\left(\frac{s_1^n-s_2^n}{s_1-s_2} - s_1s_2\frac{s_1^{n-1}-s_2^{n-1}}{s_1-s_2}\right) \\ &= 2\left(P_{a,r,n}+a^{r-1}P_{a,r,n-1}\right). \end{aligned}$$

Equality (ii) has been proved in very similar way, while (iii) is obvious.
The proof is completed. □

As a special case, for $a = 1$ and $r = 1$, we have:

Corollary 3.5 (**Basic classical relationships formulae**). *For all $n \geq 0$, we have*

(i) $Q_n = 2\left(P_n + P_{n-1}\right)$,
(ii) $q_n = P_n + P_{n-1}$, *and*
(iii) $Q_n = 2q_n$.

The following lemma is not important only for itself but also for other applications.

Lemma 3.6. *For all $a \geq 1$, $r \geq 1$, and $n \geq 0$, we have*

$$\lim_{n\to\infty}\frac{P_{a,r,n+1}}{P_{a,r,n}} = s_2.$$

Proof. Inequality $\left|\frac{s_1}{s_2}\right| < 1$ implies $\lim_{n\to\infty}\left(\frac{s_1}{s_2}\right)^n = 0$, and consequently

$$\lim_{n\to\infty}\frac{P_{a,r,n+1}}{P_{a,r,n}} = \lim_{n\to\infty}\frac{s_1^{n+1}-s_2^{n+1}}{s_1^n-s_2^n} = s_2.$$

The proof is completed. □

Putting $a=1$ and $r=1$ in this lemma we obtain:

Corollary 3.7. *For all $n \geq 0$, we have*

$$\lim_{n\to\infty}\frac{P_{n+1}}{P_n} = 1+\sqrt{2}.$$

We know that the number $1+\sqrt{2}$ is well-known as silver ratio.

Remark 3.8. *Note that we can obtain the "silver ratios" as many as we wish. For example, for the sequence $\{P_{2,1,n}\}$ the "silver ratio" is*

$$\lim_{n\to\infty}\frac{P_{2,1,n+1}}{P_{2,1,n}} = 2+\sqrt{5}.$$

Lemma 3.6 enable us to conclude that the radius of convergence of the series (2.1) is $\frac{1}{r_2}$. So, we can write

$$f_{a,r}(x) = \sum_{k=0}^{\infty} P_{a,r,k}x^k, \quad \forall x \in \left(-\frac{1}{r_2}, \frac{1}{r_2}\right).$$

Based on this fact and some prior knowledge we can prove easily next theorem which shows another form of the numbers $P_{a,r,n}$.

Theorem 3.9. *For all $a \geq 1$, $r \geq 1$, and $n \geq 0$, we have*

$$P_{a,r,n} = \frac{f_{a,r}^{(n)}(0)}{n!},$$

where $f_{a,r}^{(n)}$ denotes the n-th derivative of the function $f_{a,r}(x)$.

4. Some basic related identities

Generalized Binet's formulae (3.1), (3.2), and (3.3) given in the previous section are very useful to derive some identities for numbers $P_{a,r,n}$, $Q_{a,r,n}$, and $q_{a,r,n}$. Here we give those in several theorems.

Theorem 4.1 (**Generalized Catalan's identities**). *For all $a \geq 1$, $r \geq 1$, $s \geq 1$, and $n \geq 0$, we have*

$$P_{a,r,n-s}P_{a,r,n+s} - P_{a,r,n}^2 = -\left(-a^{r-1}\right)^{n-s}P_{a,r,s}^2, \tag{4.1}$$

$$Q_{a,r,n-s}Q_{a,r,n+s} - Q_{a,r,n}^2 = 4(2a^r + a^{r-1} - 1)\left(-a^{r-1}\right)^{n+2-s}Q_{a,r,s}^2, \tag{4.2}$$

and

$$q_{a,r,n-s}q_{a,r,n+s} - q_{a,r,n}^2 = (2a^r + a^{r-1} - 1)\left(-a^{r-1}\right)^{n+2-s}q_{a,r,s}^2, \tag{4.3}$$

Proof. Let us prove first (4.1). Namely, using (3.1) of Theorem 3.1, (1.16) and (1.17) we get

$$
\begin{aligned}
&P_{a,r,n-s}P_{a,r,n+s}-P_{a,r,n}^{2}\\
&\quad=\frac{(s_2^{n-s}-s_1^{n-s})(s_2^{n+s}-s_1^{n+s})-(s_2^{n}-s_1^{n})^2}{(s_2-s_1)^2}\\
&\quad=\frac{2(s_1s_2)^n-s_1^{n+s}s_2^{n-s}-s_2^{n+s}s_1^{n-s}}{\left(s_2-s_1\right)^2}\\
&\quad=\frac{2\left(-a^{r-1}\right)^n-\left(-a^{r-1}\right)^n\left(\frac{s_1}{s_2}\right)^s-\left(-a^{(r-1)}\right)^n\left(\frac{s_2}{s_1}\right)^s}{\left(s_2-s_1\right)^2}\\
&\quad=-\frac{\left(-a^{r-1}\right)^n}{(s_2-s_1)^2}\left[\frac{s_1^{2s}-2(s_1s_2)^s+s_2^{2s}}{(s_1s_2)^s}\right]\\
&\quad=-\frac{\left(-a^{r-1}\right)^n}{(s_2-s_1)^2}\left[\frac{(s_1^{s}-s_2^{s})^2}{\left(-a^{(r-1)}\right)^s}\right]\\
&\quad=-\frac{\left(-a^{r-1}\right)^n}{\left(-a^{r-1}\right)^s}\left(\frac{s_1^{s}-s_2^{s}}{s_2-s_1}\right)^2\\
&\quad=-\left(-a^{r-1}\right)^{n-s}P_{a,r,s}^{2}.
\end{aligned}
$$

For the proof of (4.2) first (for short notation) we denote $y=\frac{2(1-s_2)}{s_1-s_2}$ and $z=-\frac{2(1-s_1)}{s_1-s_2}$. Now using (3.2) of Theorem 3.1, (1.15) and (1.17) to obtain

$$
\begin{aligned}
&Q_{a,r,n-s}Q_{a,r,n+s}-Q_{a,r,s}^{2}\\
&\quad=\left(ys_1^{n-s}+zs_2^{n-s}\right)\left(ys_1^{n+s}+zs_2^{n+s}\right)-\left(ys_1^{n}+zs_2^{n}\right)^2\\
&\quad=yz\left(s_1^{n-s}s_2^{n+s}-2s_1^{n}s_2^{n}+s_1^{n+s}s_2^{n-s}\right)\\
&\quad=yz\left(s_1s_2\right)^n\left(\left(\frac{s_1}{s_2}\right)^s-2+\left(\frac{s_2}{s_1}\right)^s\right)\\
&\quad=-4\frac{(1-s_2)(1-s_1)}{(s_1-s_2)^2}\left(s_1s_2\right)^n\frac{\left(s_1^{s}-s_2^{s}\right)^2}{\left(s_1s_2\right)^s}\\
&\quad=-4(1+s_1s_2-(s_1+s_2))\left(s_1s_2\right)^{n-s}\left(\frac{s_1^{s}-s_2^{s}}{s_1-s_2}\right)^2\\
&\quad=-4(1-a^{r-1}-2a^r)\left(-a^{r-1}\right)^{n-s}Q_{a,r,s}^{2}\\
&\quad=4(2a^r+a^{r-1}-1)\left(-a^{r-1}\right)^{n+2-s}Q_{a,r,s}^{2}.
\end{aligned}
$$

Finally, taking into account that $q_{a,r,n}=\frac{1}{2}Q_{a,r,n}$, we immediately obtain (4.3).
The proof is completed. □

Putting the values $a=1$ and $r=1$ into relations (4.1) and (4.2) we obtain the following corollary.

Corollary 4.2 (Classical Catalan's identities). *For all $s\geq 1$ and $n\geq 0$, we have*

$$P_{n-s}P_{n+s}-P_{n}^{2}=\left(-1\right)^{n+1-s}P_{s}^{2},$$

and

$$Q_{n-s}Q_{n+s}-Q_{n}^{2}=8\left(-1\right)^{n+2-s}Q_{s}^{2},$$

and

$$q_{n-s}q_{n+s}-q_{n}^{2}=2\left(-1\right)^{n+2-s}q_{s}^{2}.$$

Moreover, inserting values $a=1$, $r=1$, and $s=1$ into relations (4.1) and (4.2), imply next corollary.

Corollary 4.3 (**Classical Simpson's identities** [2], [4]). *For all $n \geq 0$, we have*

$$P_{n-1}P_{n+1} - P_n^2 = (-1)^n ,$$
$$Q_{n-1}Q_{n+1} - Q_n^2 = 8\,(-1)^{n+1} ,$$

and

$$q_{n-1}q_{n+1} - q_n^2 = 2\,(-1)^{n+1} .$$

Some other identities are given in next theorem.

Theorem 4.4. *Let m, n be two positive integers such that $m \geq n \geq 0$. Then, for all $a \geq 1$ and $r \geq 1$, we have*

$$P_{a,r,m}P_{a,r,n+1} - P_{a,r,m+1}P_{a,r,n} = \big(-a^{r-1}\big)^n P_{a,r,m-n}, \tag{4.4}$$

$$Q_{a,r,m}Q_{a,r,n+1} - Q_{a,r,m+1}Q_{a,r,n} = 4(1 - a^{r-1} - 2a^r)(-a^{r-1})^n P_{a,r,m-n}, \tag{4.5}$$

and

$$q_{a,r,m}q_{a,r,n+1} - q_{a,r,m+1}q_{a,r,n} = (1 - a^{r-1} - 2a^r)(-a^{r-1})^n P_{a,r,m-n}. \tag{4.6}$$

Proof. For $m \geq n \geq 0$, we have

$$\begin{aligned}
&P_{a,r,m}P_{a,r,n+1} - P_{a,r,m+1}P_{a,r,n} \\
&\quad = \frac{\left(s_1^m - s_2^m\right)\left(s_1^{n+1} - s_2^{n+1}\right) - \left(s_1^{m+1} - s_2^{m+1}\right)\left(s_1^n - s_2^n\right)}{\left(s_1 - s_2\right)^2} \\
&\quad = \frac{s_1^n s_2^{m+1} - s_1^{n+1} s_2^m + s_1^{m+1} s_2^n - s_1^m s_2^{n+1}}{\left(s_1 - s_2\right)^2} \\
&\quad = \frac{s_1^n s_2^m (s_2 - s_1) - s_1^m s_2^n (s_2 - s_1)}{\left(s_1 - s_2\right)^2} \\
&\quad = (s_1 s_2)^n \frac{s_2^{m-n} - s_1^{m-n}}{s_2 - s_1} \\
&\quad = \big(-a^{r-1}\big)^n P_{a,r,m-n}.
\end{aligned}$$

As in the proof of Theorem 3.9, let $y = \frac{2(1-s_2)}{s_1 - s_2}$ and $z = -\frac{2(1-s_1)}{s_1-s_2}$. Then, for $m \geq n \geq 0$, we obtain

$$\begin{aligned}
&Q_{a,r,m}Q_{a,r,n+1} - Q_{a,r,m+1}Q_{a,r,n} \\
&\quad = \left(ys_1^m + zs_2^m\right)\left(ys_1^{n+1} + zs_2^{n+1}\right) - \left(ys_1^{m+1} + zs_2^{m+1}\right)\left(ys_1^n + zs_2^n\right) \\
&\quad = yz\big[s_1^m s_2^n (s_2 - s_1) - s_1^n s_2^m (s_2 - s_1)\big] \\
&\quad = -4\frac{(1 - s_2)(1 - s_1)}{(s_1 - s_2)^2}(-a^{r-1})^n (s_2 - s_1)\big(s_2^{m-n} - s_1^{m-n}\big) \\
&\quad = -4(1 + s_1 s_2 - (s_1 + s_2))(-a^{r-1})^n \frac{s_2^{m-n} - s_1^{m-n}}{s_1 - s_2} \\
&\quad = 4(1 - a^{r-1} - 2a^r)(-a^{r-1})^n P_{a,r,m-n}.
\end{aligned}$$

Relation (4.6) is a direct consequence of (4.5).
The proof is completed. □

For $a = r = 1$ in (4.4) and (4.5) we obtain next corollary.

Corollary 4.5 (**Classical d'Ocagne's identities** [2], [4]). *For all $m \geq n \geq 0$, we have*

$$P_m P_{n+1} - P_{m+1} P_n = (-1)^n P_{m-n},$$

and

$$Q_m Q_{n+1} - Q_{m+1} Q_n = 8(-1)^{n+1} P_{m-n},$$

$$q_m Q_{n+1} - q_{m+1} q_n = 2(-1)^{n+1} P_{m-n}.$$

Generalized Binet's formulae (3.1), (3.2), and (3.3), allow us to express the sum of the first terms of sequences $P_{a,r,n}$, $Q_{a,r,n}$, and $q_{a,r,n}$ in a simple form. Namely, we prove the following.

Theorem 4.6. *For all $a \geq 1$, $r \geq 1$, and $n \geq 0$, we have*

$$\sum_{j=0}^{n} P_{a,r,j} = \frac{a^{r-1} P_{a,r,n} + P_{a,r,n+1} - 1}{2a^r + a^{r-1} - 1}, \tag{4.7}$$

$$\sum_{j=0}^{n} Q_{a,r,j} = \frac{Q_{a,r,n+1} + a^{r-1} Q_{a,r,n} + 4(1-a^r)}{2a^r + a^{r-1} - 1}, \tag{4.8}$$

and

$$\sum_{j=0}^{n} q_{a,r,j} = \frac{q_{a,r,n+1} + a^{r-1} q_{a,r,n} + 4(1-a^r)}{2(2a^r + a^{r-1} - 1)}. \tag{4.9}$$

Proof. Using generalized Binet's formulae we have

$$\begin{aligned}
\sum_{j=0}^{n} P_{a,r,j} &= \frac{1}{s_1 - s_2} \sum_{j=0}^{n} (s_1^j - s_2^j) \\
&= \frac{1}{s_1 - s_2} \left(\frac{1 - s_1^{n+1}}{1 - s_1} - \frac{1 - s_2^{n+1}}{1 - s_2} \right) \\
&= \frac{s_1 - s_2 - (s_1^{n+1} - s_2^{n+1}) + s_1 s_2 (s_1^n - s_2^n)}{(1 - s_1)(1 - s_2)(s_1 - s_2)} \\
&= \frac{1}{1 - (s_1 + s_2) + s_1 s_2} \left(1 - \frac{s_1^{n+1} - s_2^{n+1}}{s_1 - s_2} + s_1 s_2 \frac{s_1^n - s_2^n}{s_1 - s_2} \right) \\
&= \frac{1}{2a^r + a^{r-1} - 1} \left(a^{r-1} P_{a,r,n} + P_{a,r,n+1} - 1 \right),
\end{aligned}$$

which proves (4.7).

To prove (4.8) we use once again notations $y = \frac{2(1-s_2)}{s_1 - s_2}$ and $z = -\frac{2(1-s_1)}{s_1 - s_2}$ and (3.2). Namely, after some transformations, we have

$$\begin{aligned}
\sum_{j=0}^{n} Q_{a,r,j} &= \sum_{j=0}^{n} \left(y s_1^j + z s_2^j \right) \\
&= y \frac{1 - s_1^{n+1}}{1 - s_1} + z \frac{1 - s_2^{n+1}}{1 - s_2} \\
&= \frac{y + z - (z r_1 + y s_2) - (y s_1^{n+1} + z s_2^{n+1}) + s_1 s_2 (y s_1^n + z s_2^n)}{(1 - s_1)(1 - s_2)}.
\end{aligned}$$

Since

$$\begin{aligned}
&y + z = 2, \\
&s_1 s_2 = -a^{r-1}, \\
&z s_1 + y s_2 = 2(2a^r - 1), \\
&(1 - s_1)(1 - s_2) = 1 - a^{r-1} - 2a^r,
\end{aligned}$$

then we get

$$\begin{aligned}\sum_{j=0}^{n} Q_{a,r,j} &= \frac{2-2(2a^r-1)-Q_{a,r,n+1}-a^{r-1}Q_{a,r,n}}{1-a^{r-1}-2a^r}\\ &= \frac{Q_{a,r,n+1}+a^{r-1}Q_{a,r,n}+4(1-a^r)}{2a^r+a^{r-1}-1}.\end{aligned}$$

Relation (4.9) follows from relation (4.8).

The proof is completed. □

If we take specific values $a = r = 1$ in Theorem 4.6, then we obtain next well-known equalities for Pell and Pell-Lucas numbers.

Corollary 4.7. *For all $n \geq 0$, we have*

$$\sum_{j=0}^{n} P_j = \frac{P_n + P_{n+1} - 1}{2},$$

$$\sum_{j=0}^{n} Q_j = \frac{Q_{n+1} + Q_n}{2},$$

and

$$\sum_{j=0}^{n} q_j = \frac{q_{n+1} + q_n}{4}.$$

Some other interesting relations are shown in next statement.

Theorem 4.8. *For all $a \geq 1$, $r \geq 1$, and $n \geq 0$, we have*

(i) $P_{a,r,3n} = 4(a^r + a^{r-1})P^3_{a,r,n} + 3(-a^{r-1})^n P_{a,r,n}$,

(ii) *The sequence $\{P_{a,r,2}\}$ is a geometric one with respect to r.*

(iii) *The following inequality $P_{a+1,r,2} \leq 2^r P_{a,r,2}$ hods true.*

Proof. (i) By Binet's formula have

$$\begin{aligned}P_{a,r,3n} &= \frac{s_1^{3n}-s_2^{3n}}{s_1-s_2}\\ &= \frac{s_1^{n}-s_2^{n}}{s_1-s_2}\left\{[s_1^{2n}-2(s_1s_2)^n+s_2^{2n}]+3(s_1s_2)^n\right\}\\ &= P_{a,r,n}\left\{(s_1-s_2)^2\left(\frac{s_1^{n}-s_2^{n}}{s_1-s_2}\right)^2+3(s_1s_2)^n\right\}\\ &= P_{a,r,n}\left\{4(a^r+a^{r-1})P^2_{a,r,n}+3(-a^{r-1})^n\right\}.\end{aligned}$$

(ii) By same argument, we also have

$$P_{a,r,2} = \frac{s_1^2-s_2^2}{s_1-s_2} = s_1+s_2 = 2a^r,$$

and consequently

$$\frac{P_{a,r+1,2}}{P_{a,r,2}} = 2a.$$

(iii) It is obvious that

$$\frac{P_{a+1,r,2}}{P_{a,r,2}} = \frac{(a+1)^r}{a^r} = \left(1+\frac{1}{a}\right)^r \leq 2^r, \quad \forall a, r \geq 1.$$

The proof is completed. □

Putting $a = 1$ into Theorem 4.8 we obtain:

Corollary 4.9. *For all $n \geq 0$, we have*

(i) $P_{3n} = 8P_n^3 + 3(-1)^n P_n$,
(ii) *The sequence $\{P_{1,r,2}\}$ is a geometric one with respect to r with its quotient 2.*
(iii) *The following equality $P_{2,r,2} = 2^r P_{1,r,2}$ hods true.*

References

[1] P. Catarino, *A note involving two-by-two matrices of the k-Pell and k-Pell-Lucas sequences*, Int. Math. Forum **8** (2013), 1561–1568.
[2] A. F. Horadam, *Pell identities*, Fibonacci Quart. **9(3)** (1971), 245–252.
[3] A. F. Horadam, *Applications of Modifed Pell Numbers to Representations*, Ulam Quart. **3(1)** (1994), 34–53.
[4] J. M. Mahon and A. F. Horadam, *Pell and Pell-Lucas polynomials*, Fibonacci Quart. **23(1)** (1985), 7–20.
[5] L. T. Spelina and I. Wloch, *On generalized Pell and Pell-Lucas Numbers*, Iran J. Sci. Technol. Trans. Sci. **43** (2019), 2871–2877 .

University of Prishtina "Hasan Prishtina", Faculty of Education, Department of Mathematics and Informatics, Avenue "Mother Theresa " no. 5, Prishtina 10000, Kosovo
Email address: `xhevat.krasniqi@uni-pr.edu`

HERMITE PRINCIPLE, LINDEMANN'S IDEA AND SIMPLE PROOFS FOR THE BASIC RESULTS IN THE IRRATIONALITY AND TRANSCENDENCE OF SOME NUMBERS

SEVER ANGEL POPESCU

A tribute to the 80^{th} birthday of Professor Gavriil Păltineanu

ABSTRACT. In this note we give some simple proofs for the irrationality of π^2 and e^q, where q is a nonzero rational number. We also use some basic ideas of Hermite, Hilbert, Hurwitz and Lindemann to give simple proofs for the transcendence of e and of e^a, where a is a nonzero algebraic number. In particular we obtain the transcendence of π. In the end we apply our previous results to develop a great idea of Wantzel for a complete solution of the three famous Greek geometrical problem: squaring a circle, duplicate a cube and trisection of an angle. We also supply all the elementary prerequisites of algebraic number theory such that the famous Lindemann's idea be easily understood even by an undergraduate student with some basic knowledge in Calculus and Algebra.

Mathematics Subject Classification (2010): 11J81, 11J99
Key words: Hermite principle, irrationality and transcendence of e and π, Lindemann theorem, Greek old geometrical conjectures.

Article history:
Received: September 30, 2022
Received in revised form: November 10, 2022
Accepted: November 15, 2022

1. Introduction

In this paper we try to present as elementary as possible (with complete and self-contained proofs) basic results related to the irrationality and transcendence of some fundamental real or complex numbers. The reader must have some knowledge of mathematics that is taught in the first year of the undergraduate studies. Namely, he (she) has to know the basic elementary results in differential and integral calculus, and in Linear Algebra. It would be better if the reader had some knowledge in complex function theory. However, we do not assume such a knowledge and we reduce everything to usual Riemann integrals for a real variable function with complex values. We also use a basic and elementary result from finite group theory.

Trying to make an elementary presentation, the author was inspired by the main ideas of some great mathematicians like C. Hermite [4], D. Hilbert [5], A. Hurwitz [6] and F. Lindemann [8]. Some other nice ideas came to us from the enlightened works of A. Baker [2] and M. Waldschmidt [10]. The last but not the least, we were encouraged by the elegant one page paper of I. Niven [9] and by the inspirational book *Proofs from THE BOOK* of Aigner and Ziegler [1].

In Section 2 we make a short introduction in the long history of the attempts to prove the irrationality and transcendence of e and π. Here we highlighted Hermite's contribution to what we call today Hermite-Padé approximation and the so called Hermite Principle.

In Section 3 we apply Hermite's Principle to supply a short and elementary proof for the irrationality of π^2 (as an immediate consequence it comes also the irrationality of π).

In Section 4 we apply the same Hermite's Principle to prove in a simple way the irrationality of e^q for $q \in \mathbb{Q} \setminus \{0\}$.

In Section 5 we prove that e is a transcendental number. This proof uses the original Hermite's idea, but extended with a very useful Hilbert's and Hurwitz's ideas relative to the nonzero property of an integer which is nonzero modulo p, where p is an arbitrary prime number. In this proof we find the seeds of the great ideas of Lindemann and Weierstrass for proving the independence over $\overline{\mathbb{Q}}$ (the field of all algebraic numbers-see Section 6) of a sequence of powers of e, where these powers are distinct algebraic numbers (see also Section 6).

In Section 6 we introduce in a self-contained manner some elementary facts from algebraic number theory. Here we assume that the reader has no previous knowledge of this topic. We use all of these for giving self-contained presentations in sections 7 and 8.

In Section 7 we supply a complete and simple proof for a slight generalization of the famous Lindemann's result [8] relative to the transcendence of π, namely we prove that e^α is a transcendental number, where α is a nonzero algebraic number. In particular we prove that π is a transcendental number. It is clear enough that the main result of Section 5 is a direct consequence of the main theorem 7.2 from Section 7. Theorem 7.2 is a particular case of another famous result of Weierstrass and Lindemann (see [2]) which has a non-elementary proof. We wanted to insist on the natural evolution of mathematical ideas and not to choose the "shortest path", because a clever reader who want to really understand these deep and beautiful ideas, will enjoy our "longer, but natural path".

In Section 8 we apply the results obtained in Sections 6 and 7 to develop an enlightened idea of P. L. Wantzel [11] to associate to a geometrical construction realized only with a compass and a straight-edge, a tower of algebraic number fields. Following his idea we give complete answers to the three famous Greek geometrical problem: squaring a circle, duplication of a cube and trisection of an angle.

2. Some historical notes

More than 2300 years ago, Aristotle (384-322 BC) said that the diameter and the circumference of a circle are not commensurable, i.e. π is not a rational number in our modern language. But a first proof came later in 1761, when J. H. Lambert proved the irrationality of $\tan q$ for q a nonzero rational number. So, if π was a rational number, then $\tan(\pi/4) = 1$ would be irrational, a contradiction. Lambert used the same idea as L. Euler, who proved in 1737 that e is an irrational number by using continuous fraction expansions. Trying to prove the transcendence of e, C. Hermite [4] discovered in 1873 a new type of approximation with rational functions for $f(x) = e^{bx}$, $b \in \mathbb{N}^*$. Now we call his discovery, Hermite Principle. Here is a short description of it. Let n_0, n_1, b be nonzero natural numbers and let z a nonzero real number. We denote

$$f(x) = \frac{x^{n_0}}{n_0!}(b-x)^{n_1},$$

a polynomial of degree $N = n_0 + n_1$ in $\mathbb{Q}[x]$, and let us consider the following Riemann integral with a parameter z:

$$I = I_{n_0,n_1,b}(z) = \int_0^b f(x)e^{-zx}dx.$$

We integrate it N-times by parts, and finally find:

$$I = \frac{1}{z^{N+1}}\left[F_z(0) - F_z(b)e^{-zb}\right],$$

where

$$F_z(x) = z^N f(x) + z^{N-1}f'(x) + ... + zf^{(N-1)}(x) + f^{(N)}(x).$$

Thus,

$$e^{zb} - \frac{F_z(b)}{F_z(0)} = \frac{z^{N+1}e^{bz}}{F_z(0)} I. \tag{2.1}$$

It is easy to see that $F_z(b)$ is a polynomial in z of degree n_0, with coefficients in $\frac{1}{n_0!}\mathbb{Z}[b]$, and $F_z(0)$ is a polynomial in z of degree n_1, with integer coefficients (in $\mathbb{Z}[b]$). Since the derivatives of $f(x)$ at $x = 0$ are zero up to the order $n_0 - 1$ inclusive, the polynomial $F_z(0)$ in the variable z has a fixed degree n_1 and fixed coefficients (they depend only on n_1 and b). So, it has a fixed minimal or maximal value on each interval $[A, B]$.

If we fix n_1 and make $n_0 \to \infty$, we see that the approximation $e^{zb} \approx \frac{F_z(b)}{F_z(0)}$ is an approximation of e^{zb} (as a function of z) with rational functions in z, i.e.

$$\frac{z^{N+1}e^{bz}}{F_z(0)} I \to 0,$$

when $n_0 \to \infty$ and $z \in [A, B]$, a closed and bounded interval which does not contain any root of $F_z(0)$ (see the above explanation on the polynomial $F_z(0)$). Indeed,

$$\left|\frac{z^{N+1}e^{bz}}{F_z(0)}\right| I \leq \frac{b^{n_0}M^{n_0}}{n_0!} \cdot \frac{(bM)^{n_1+1}}{\inf\limits_{z\in[A,B]} |F_z(0)|} \cdot C,$$

where $M = \sup\limits_{z\in[A,B]} |z|$ and $C = \sup\limits_{x\in[0,b], z\in[A,B]} e^{-z(x-b)}$. Since

$$\frac{(bM)^{n_1+1}}{\inf\limits_{z\in[A,B]} |F_z(0)|} \cdot C$$

is a positive constant K, and since $\frac{b^{n_0}M^{n_0}}{n_0!} \to 0$, when $n_0 \to \infty$, we see that $\frac{z^{N+1}e^{bz}}{F_z(0)} I$ goes uniformly (relatively to z) to zero, when $n_0 \to \infty$ and when $z \in [A, B]$. We also see that $I = I_{n_0,n_1,b}(z)$ itself uniformly goes to zero, when $n_0 \to \infty$ and when $z \in [A, B]$. Hence $e^{zb} \approx \frac{F_z(b)}{F_z(0)}$ is indeed a uniform approximation (relative to z) on each real interval $[A, B]$ which does not contain any root of $F_z(0)$. Such type of approximation with rational functions is called Hermite-Padé approximation. If one makes $z = 1$ in (2.1), one gets:

$$F_1(0) \cdot e^b - F_1(b) = e^b I_{n_0,n_1,b}(1). \tag{2.2}$$

If we assume that e^b is a rational number c/d, then (2.2) becomes:

$$cF_1(0) - dF_1(b) = de^b I_{n_0,n_1,b}(1). \tag{2.3}$$

If one could prove that the left side of this last equality is always a nonzero number outside of an interval $(-\varepsilon, \varepsilon)$ for an $\varepsilon > 0$, and since the right side goes to zero, when $n_0 \to \infty$, we would obtain a contradiction. This is in fact an example of the *Hermite Principle* applied to the problem of the irrationality of e^b. In [4] Hermite proved that the left side of (2.3) is always a nonzero quantity for $n_0 = n_1 = n \to \infty$, i.e. $|cF_1(0) - dF_1(b)| \geq \varepsilon$ for an $\varepsilon > 0$ and for any $n \geq N_0$ but, for the transcendence of e, he gives a very complicated method. He used a complicated asymptotic and obscure method which was extensively improved by Hilbert [5] and Huwitz [6] (twenty years later) by using a new algebraic idea. Let us explain Hurwitz' idea for the above example. For this, we make $n_1 = n_0 + 1$ in the definition of f, and see that $f^{(j)}(b)$ is a multiple of n_1 for every $j = 0, 1, ..., N = 2n_0 + 1$, while $f^{(j)}(0)$ is a multiple of n_1 only for $j \in \{0, 1, ..., N\} \setminus \{n_0\}$, and $f^{(n_0)}(0) = b^{n_1} \neq 0$. Now, if we take n_1 large enough and such that $(cb, n_1) = 1$ in (2.3), then $cF_1(0) - dF_1(b) \in \mathbb{Z} \setminus \{0\}$, whereas $de^b I_{n_0,n_1,b}(1) \to 0$, when $n_1 \to \infty$.

Using again Hermite Principle, in 1947, I. Niven [9] gave a one page elegant proof for the irrationality of π. For a reader with a more advanced knowledge in number theory, we recommend the deep, but not elementary study of the transcendental numbers in the book of A. Baker [2].

3. π^2 and π are irrational numbers

Lemma 3.1. *For any fixed $n \in \mathbb{N}^*$, we consider the following polynomial function:*

$$f_n(x) = \frac{x^{n-1}(1-x)^n}{(n-1)!}, \ x \in [0,1]. \tag{3.1}$$

Then its derivatives at $x = 0$ and $x = 1$ are of the following form:

$$\begin{cases} f_n^{(j)}(1) = M_j \cdot n, \ M_j \in \mathbb{Z}, \ j = 0,1,...,2n-1, \\ f_n^{(j)}(0) = L_j \cdot n, \ L_j \in \mathbb{Z}, \ j = 0,1,...,2n-1, \ j \neq n-1, \\ f_n^{(n-1)}(0) = 1. \end{cases} \tag{3.2}$$

Since f_n is a polynomial function of degree $2n-1$, $f_n^{(j)}(x) = 0$ for all $j > 2n-1$ and for all $x \in [0,1]$.

Proof. We have only to carefully use Leibniz formula for computing the derivative of a product of two functions. □

Now, let us use Hermite Principle (see Section 2) for the following integral ($n \in \mathbb{N}^*$, n an odd number):

$$I_n = \int_0^1 f_n(x)\sin(\pi x)\,dx,$$

where f is the polynomial function defined in (3.1). We integrate it by parts $(2n-1)$-times and find:

$$I_n = \sum_{j=0}^{n-1} (-1)^j \frac{1}{\pi^{2j+1}} f_n^{(2j)}(1) + \sum_{j=0}^{n-1} (-1)^j \frac{1}{\pi^{2j+1}} f_n^{(2j)}(0). \tag{3.3}$$

Assume that $\pi^2 = a/b$, $a, b \in \mathbb{N}^*$, $(a,b) = 1$. So, from (3.3) we get:

$$\pi a^{n-1} I_n = \sum_{j=0}^{n-1} (-1)^j a^{n-j-1} b^j f_n^{(2j)}(1) + \sum_{j=0}^{n-1} (-1)^j a^{n-j-1} b^j f_n^{(2j)}(0). \tag{3.4}$$

Now we apply Lemma 3.1 and find that all the terms in these two sums are multiple of n, except the term $j = \frac{n-1}{2}$ in the second sum ($j \in \mathbb{N}$, because n is an odd number).

Thus (3.4) becomes:

$$\pi a^{n-1} I_n = M \cdot n + (-1)^{\frac{n-1}{2}} a^{\frac{n-1}{2}} b^{\frac{n-1}{2}}, \ M \in \mathbb{Z}. \tag{3.5}$$

Since a and b are fixed numbers, we can take n such that n is prime with a and b. Thus, in (3.5), the right side is not zero modulo n, i.e. it cannot be zero for such numbers n.

Now we evaluate the left side of (3.5). It is clear that

$$\pi a^{n-1} I_n \leq \frac{a^{n-1}}{(n-1)!}\pi$$

for any $n = 1, 2, ...$ Thus $\pi a^{n-1} I_n \to 0$, when $n \to \infty$, taking values such that n is prime with a and b. Hence, we obtained a contradiction, namely, the nonzero integer numbers from the right cannot be closer and closer to zero. So π^2 is an irrational number.

In particular, if π were a rational number, then also π^2 would be a rational number, a contradiction. Hence π itself is an irrational number.

4. e^q, $q \in \mathbb{Q} \setminus \{0\}$, IS AN IRRATIONAL NUMBER

It is sufficient to prove that e^c, where $c \in \mathbb{N}^*$, is an irrational number. Indeed, if $q = m/n$, where $m, n \in \mathbb{N}^*$, $(m.\ n) = 1$, and if $e^q = a/b$, $a, b \in \mathbb{N}^*$, $(a, b) = 1$, then $e^m = (a/b)^n$ would be a rational number, etc.

Thus we assume that $q \in \mathbb{N}^*$ and $e^q = a/b$, $a, b \in \mathbb{N}^*$, $(a, b) = 1$. In order to use again the Hermite Principle we consider the following Riemann integral:

$$J_n = \int_0^1 f_n(x) e^{qx}\, dx,$$

where $f(x)$ is the function defined in (3.1). Now, we integrate by parts $(2n-1)$-times and get:

$$J_n = e^q \sum_{j=0}^{2n-1} (-1)^j \frac{1}{q^{j+1}} f_n^{(j)}(1) - \sum_{j=0}^{2n-1} (-1)^j \frac{1}{q^{j+1}} f_n^{(j)}(0).$$

We use again Lemma 3.1, formulas (3.2) and find for $e^q = a/b$:

$$bq^{2n} J_n = \sum_{j=0}^{2n-1} (-1)^j a q^{2n-j-1} f_n^{(j)}(1) - \sum_{j=0}^{2n-1} (-1)^j b q^{2n-j-1} f_n^{(j)}(0) =$$

$$= M \cdot n + (-1)^{n-1} b q^n.$$

Let us take now n such that n is prime with bq. So, $M \cdot n + (-1)^p bq^p$ is not divisible by n for such n. So, $M \cdot n + (-1)^{n-1} bq^n \neq 0$ for values of n. But,

$$bq^{2n} J_n \leq b \cdot q^2 \cdot \frac{(q^2)^{n-1}}{(n-1)!} \to 0,$$

when $n \to \infty$, so these nonzero integers, $M \cdot n + (-1)^{n-1} bq^n \to 0$, when n is large enough, such that n is prime with bq, a contradiction. Thus, e^q is an irrational number for any $q \in \mathbb{Q} \setminus \{0\}$.

Remark 4.1. *In this particular case, it is easy to see that $J_n > 0$ for any $n = 1, 2, \ldots$ So, it is not necessarily to take only those n which are prime with bq.*

5. e IS A TRANSCENDENTAL NUMBER

Assume that e is an algebraic number, i.e. it is a root of a polynomial with rational coefficients. This is equivalent to say that there exist $a_0, a_1, \ldots, a_k$ in $\mathbb{Z}$, with $a_0, a_k \neq 0$, such that

$$a_0 + a_1 e + \ldots + a_k e^k = 0. \tag{5.1}$$

For any $n \in \mathbb{N}^*$, let

$$g_n(x) = \frac{x^{n-1}[P(x)]^n}{(n-1)!} \tag{5.2}$$

be a polynomial of degree $d = n(k+1) - 1$, where $P(x) = (x-1) \cdot (x-2) \cdot \ldots \cdot (x-k)$. Like in Lemma 3.1, it is not difficult to see that for any $m \in \{1, 2, \ldots, k\}$ one has:

$$\begin{cases} g_n^{(j)}(m) = M_j \cdot n,\ M_j \in \mathbb{Z},\ j = 0, 1, \ldots, d, \\ g_n^{(j)}(0) = L_j \cdot n,\ L_j \in \mathbb{Z},\ j = 0, 1, \ldots, d,\ j \neq n-1, \\ g_n^{(n-1)}(0) = (-1)^{kn}[k!]^n \end{cases}. \tag{5.3}$$

For any $m \in \{1, 2, \ldots, k\}$ we consider the following Riemann integral:

$$I_n(m) = \int_0^m g_n(x) e^{-x}\, dx. \tag{5.4}$$

Since g_n is a polynomial of degree d, we integrate it by parts d-times and find:

$$I_n(m) = -e^{-m} \sum_{j=0}^{d} g^{(j)}(m) + \sum_{j=0}^{d} g^{(j)}(0). \tag{5.5}$$

Thus,

$$\sum_{m=1}^{k} a_m e^m I_n(m) = -\sum_{m=1}^{k} a_m \sum_{j=0}^{d} g^{(j)}(m) + \sum_{j=0}^{d} g^{(j)}(0) \sum_{m=1}^{k} a_m e^m. \tag{5.6}$$

From (5.1) we can write $-a_0$ instead of $\sum_{m=1}^{k} a_m e^m$. Thus, from (5.3), (5.6) becomes:

$$\sum_{m=1}^{k} a_m e^m I_n(m) = M \cdot n - a_0(-1)^{kn}[k!]^n.$$

Take now $n = p$, a prime number sufficiently large, say $p \geq p_0$, such that $a_0[k!]^n$ is not divisible by p for $p \geq p_0$. Hence, $M \cdot p - a_0(-1)^{kp}[k!]^p \in \mathbb{Z} \setminus \{0\}$ for any $p \geq p_0$. But,

$$\left| \sum_{m=1}^{k} a_m e^m I_p(m) \right| \leq \left[\sum_{m=1}^{k} |a_m| \, e^m \right] \cdot k^p \cdot \left[\sup_{x \in [0,k]} |P(x)|^p \right] \cdot \frac{1}{(p-1)!}.$$

Let us denote $U = \sum_{m=1}^{k} |a_m| \, e^m$ and $V = \sup\limits_{x \in [0,k]} |P(x)|$. So,

$$\left| \sum_{m=1}^{k} a_m e^m I_p(m) \right| \leq U \cdot V \cdot k \cdot \frac{(V \cdot k)^{p-1}}{(p-1)!} \to 0,$$

when $p \to \infty$. Thus, we get a contradiction. Hence e cannot be an algebraic number, i.e. it is a transcendental number.

6. Some elementary facts from algebraic number theory

There are some good books on algebraic number theory. Unfortunately, it is difficult for a reader with few basic knowledge in undergraduate mathematics to study them. This is why we give here only some prerequisites on algebraic numbers, which we extensively use in Sections 7 and 8.

Let $\mathbb{Q}$ be the rational number field, contained in any other subfield K of $\mathbb{C}$, the field of all complex numbers, which is algebraically closed, i.e. any polynomial $P \in \mathbb{C}[x]$ has all its roots in $\mathbb{C}$ ([7], Example 5, Ch. VI. See also the elegant and elementary proof of Fefferman [3]). A complex number $\alpha \in \mathbb{C}$ is said to be an *algebraic number* (over $\mathbb{Q}$) if there exist $a_0, a_1, ..., a_{n-1}$ in $\mathbb{Q}$ such that

$$a_0 + a_1\alpha + ... + a_{n-1}\alpha^{n-1} + \alpha^n = 0, \ n \in \mathbb{N}^*, \tag{6.1}$$

This means that α is a root of a monic polynomial $P(x)$ of degree $n \geq 1$, i.e. $P(\alpha) = 0$, where

$$P(x) = a_0 + a_1 x + ... + a_{n-1}x^{n-1} + x^n \in \mathbb{Q}[x]. \tag{6.2}$$

Let us denote $\overline{\mathbb{Q}}$ the subset of all algebraic numbers of $\mathbb{C}$. It is clear that $\mathbb{Q} \subset \overline{\mathbb{Q}}$. If in (6.1) $a_0, a_1, ..., a_{n-1} \in \mathbb{Z}$, we say that α is an *algebraic integer* (over $\mathbb{Q}$). Let $\mathbb{A}$ be the subset of all algebraic integers of $\mathbb{C}$. So, $\mathbb{Z} \subset \mathbb{A} \subset \overline{\mathbb{Q}}$. If α is an algebraic number, we denote $\mathbb{Q}[\alpha]$ the least subring of $\mathbb{C}$ which contains $\mathbb{Q}$ and α, i.e. the set of all polynomials in α,

$$f(\alpha) = b_0 + b_1\alpha + ... + b_k\alpha^k,$$

where $b_0, b_1, ..., b_k \in \mathbb{Q}$. Let α be an algebraic number and let $f_\alpha \in \mathbb{Q}[x]$ be a monic polynomial of the least degree such that $f_\alpha(\alpha) = 0$. Since f_α is unique, it is called the *minimal polynomial of* α (over $\mathbb{Q}$). The degree of f_α is called the degree of α and we denote it by $\deg \alpha$. It is easy to see that f_α is irreducible over $\mathbb{Q}$. If $\deg f_\alpha = \deg \alpha = n$, the other roots $\alpha_2, \alpha_3, ..., \alpha_n \in \mathbb{C}$ (in fact in $\overline{\mathbb{Q}}$) are called the *conjugates* *of* α. The set $\{\alpha_1 = \alpha, \alpha_2, ..., \alpha_n\}$ is called the *orbit of* α. Sometimes we say that $\alpha_1 = \alpha, \alpha_2, ..., \alpha_n$ are the conjugates of α.

Let P be the monic polynomial from (6.2) (not necessarily the minimal polynomial of α) and let $\alpha_1, \alpha_2, ..., \alpha_n$ be its roots in $\overline{\mathbb{Q}}$. Let $x_1, x_2, ..., x_n$ be n independent variables and let

$$\begin{array}{c} s_1(x_1, x_2, ..., x_n) = \sum_{1\le i\le n} x_i \\ s_2(x_1, x_2, ..., x_n) = \sum_{1\le i<j\le n} x_i x_j \\ \vdots \\ s_n(x_1, x_2, ..., x_n) = x_1 x_2 ... x_n \end{array}, \tag{6.3}$$

be the corresponding *fundamental symmetric polynomials* in variables $x_1, x_2, ..., x_n$. In general, a polynomial $f(x_1, x_2, ..., x_n) \in \mathbb{Q}[x_1, x_2, ..., x_n]$ is said to be a *symmetric polynomial* if

$$f(x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(n)}) = f(x_1, x_2, ..., x_n),$$

for any permutation $\sigma \in \sum_n$, the group of all permutations of the set $\{1, 2, ..., n\}$. It is clear that the polynomials $s_1, s_2, ..., s_n$ from (6.3) are symmetric polynomials. From Viète's formulas one has:

$$\begin{array}{c} s_1(\alpha_1, \alpha_2, ..., \alpha_n) = -a_{n-1} \\ s_2(\alpha_1, \alpha_2, ..., \alpha_n) = a_{n-2} \\ \vdots \\ s_n(\alpha_1, \alpha_2, ..., \alpha_n) = (-1)^n a_0 \end{array}, \tag{6.4}$$

where $\alpha_1, \alpha_2, ..., \alpha_n$ are the roots of the polynomial P from (6.2). It is clear enough that if $\alpha = \alpha_1$ is an algebraic integer, then $s_j(\alpha_1, \alpha_2, ..., \alpha_n) \in \mathbb{Z}$ for any $j = 1, 2, ..., n$.

Theorem 6.1. *With the above notation and definitions, let P be a monic polynomial with integer coefficients and let $\alpha_1, \alpha_2, ..., \alpha_n$ be all the roots of P (in $\mathbb{A}$). Let $f(x_1, x_2, ..., x_n)$ be a nonzero symmetric polynomial in the variables $x_1, x_2, ..., x_n$, with coefficients in $\mathbb{Z}$. Then $f(\alpha_1, \alpha_2, ..., \alpha_n)$ is an integer, i.e. it is in $\mathbb{Z}$.*

Proof. First of all we prove that $f(x_1, x_2, ..., x_n) = g(s_1, s_2, ..., s_n)$, where g has integer coefficients and $\deg g \le d = \deg f$, and $s_1, s_2, ..., s_n$ are the fundamental symmetric polynomials from (6.3) (see also [7], 4.6 for a proof in a more condensed form). If we do this, the statement of the theorem comes from (6.4).

Since for $n = 1$ the statement is obvious, we consider $n \ge 2$. We proceed by mathematical induction on n. Assume we proved the statement for $k = 1, 2, ..., n-1$ variables. Let us prove it for n variables [statement 1]. For this statement 1 we use mathematical induction on d, the degree of f (the greatest degree of its monomials). If the degree d of f is zero, we have nothing to prove. Assume that $d \ge 1$ and suppose that we have proved the statement 1 for any degree $0, 1, ..., d-1$ and let us prove statement 1 for f with degree d.

If $f(x_1, x_2, ..., x_{n-1}, 0) = 0$, then f is divisible by x_n and we continue the reasoning from (6.5) bellow, with $F(x_1, x_2, ..., x_n) = f(x_1, x_2, ..., x_{n-1}, 0)$. So, we may assume that $f(x_1, x_2, ..., x_{n-1}, 0)$ is not identical to zero.

Since $f(x_1, x_2, ..., x_{n-1}, 0)$ is a nonzero ($n \ge 2$) symmetric polynomial in $n-1$ variables, our induction hypothesis says that

$$f(x_1, x_2, ..., x_{n-1}, 0) = g_1(s'_1, s'_2, ..., s'_{n-1}),$$

where g_1 is a polynomial of degree $\le d = \deg f$ in $n-1$ variables with integer coefficients, and $s'_1, s'_2, ..., s'_{n-1}$ are the fundamental symmetric polynomials in these $n-1$ variables. In fact, $s'_j(x_1, x_2, ..., x_{n-1}) = s_j(x_1, x_2, ..., x_{n-1}, 0)$ for any $j = 1, 2, ..., n-1$.

Now, let us consider a new polynomial:

$$h(x_1, x_2, ..., x_n) = g_1(s_1, s_2, ..., s_{n-1}) \in \mathbb{Z}[x_1, x_2, ..., x_n].$$

It is clear that h is a symmetric polynomial of degree $\le d$ in n variables. Thus,

$$F(x_1, x_2, ..., x_n) = f(x_1, x_2, ..., x_n) - h(x_1, x_2, ..., x_n)$$

is a symmetric polynomial with integer coefficients and $F(x_1, x_2, ..., x_{n-1}, 0) = 0$. Thus, F is divisible by x_n. Since F is symmetric, it is also divisible by $s_n = x_1 x_2 ... x_n$. So,

$$F(x_1, x_2, ..., x_n) = s_n G(x_1, x_2, ..., x_n), \tag{6.5}$$

where G is a polynomial in n variables with integer coefficients and $\deg G \leq d - n < d - 1$. Now, the mathematical induction hypothesis on d says that

$$G(x_1, x_2, ..., x_n) = H(s_1, s_2, ..., s_n)$$

with $\deg H \leq \deg G \leq d - n$. Thus

$$f(x_1, x_2, ..., x_n) = s_n H(s_1, s_2, ..., s_n) + g_1(s_1, s_2, ..., s_{n-1}),$$

where the degree of the polynomial on the right side is $\leq d$. Thus, denoting

$$g(s_1, s_2, ..., s_n) = s_n H(s_1, s_2, ..., s_n) + g_1(s_1, s_2, ..., s_{n-1}),$$

we see that $f(x_1, x_2, ..., x_n) = g(s_1, s_2, ..., s_n)$, where $\deg g \leq d = \deg f$, i.e. the statement 1 for n variables is proved, and so, the induction on n is finished. Hence Theorem 6.1 is completely proved. $\square$

At the beginning of this section we gave some definitions relative to the basic field $\mathbb{Q}$ of rational numbers. For our further purposes we need to start with an arbitrary subfield K of $\mathbb{C}$, the complex number field. It is clear that $\mathbb{Q} \subset K \subset \mathbb{C}$. We say that $\alpha \in \mathbb{C}$ is an *algebraic number over* K if there exists a monic polynomial $f \in K[x]$ such that $f(\alpha) = 0$. Denote $f_{\alpha,K}$ such a monic polynomial of the least possible degree. By applying the Euclid division algorithm, we see that $f_{\alpha,K}$ is unique with these last properties: $f_{\alpha,K} \in K[x]$, $f_{\alpha,K}$ is monic, $f_{\alpha,K}(\alpha) = 0$ and $f_{\alpha,K}$ has the least degree with these previous three properties. This is why $f_{\alpha,K}$ is called the *minimal polynomial of* α *over* K. We also say that the degree $\deg_K \alpha$ of α over K is equal to $\deg f_{\alpha,K}$. It is easy to see that $f_{\alpha,K}$ is irreducible over K, i.e. in $K[x]$. Now, α is a simple root of $f_{\alpha,K}$, otherwise $f'_{\alpha,K}(\alpha) = 0$ and $\deg f'_{\alpha,K} < \deg f_{\alpha,K}$. Moreover, any other root β of $f_{\alpha,K}$ has also $f_{\alpha,K}$ as its minimal polynomial, i.e. $f_{\beta,K} = f_{\alpha,K}$. Indeed, since $f_{\alpha,K}(\beta) = 0$, using the Euclidean division algorithm in $K[x]$, one can see that $f_{\beta,K}$ is a factor of $f_{\alpha,K}$ in $K[x]$. Since $f_{\alpha,K}$ is irreducible and monic, we see that $f_{\beta,K} = f_{\alpha,K}$. Thus $f_{\alpha,K}$ is the minimal polynomial for all the conjugates of α, i.e. for all the roots of it.

Lemma 6.2. *Let α be an algebraic number over K, a subfield of $\mathbb{C}$, and let $f_{\alpha,K}$ be its minimal polynomial over K. Then $f_{\alpha,K}$ has only simple roots.*

Proof. Let β be a root of $f_{\alpha,K}$. Since $f_{\alpha,K} = f_{\beta,K}$, if β was a multiple root of $f_{\beta,K}$, then $f'_{\beta,K}(\beta) = 0$, which is not possible, because $\deg f'_{\beta,K} < \deg f_{\beta,K}$. $\square$

Lemma 6.3. *Let K be a subfield of $\mathbb{C}$ and let α be an algebraic number over K. Then $K[\alpha]$, the least subring of $\mathbb{C}$ which contains K and α, is a subfield of $\mathbb{C}$, which is a finite vector space over K and* $\dim_K K[\alpha] = K[\alpha] : K = \deg\, f_{\alpha,K}$. *Thus $K[\alpha] = K(\alpha)$, the least subfield of $\mathbb{C}$ which contains K and α. Conversely, if $\gamma \in \mathbb{C}$ such that $\mathbb{Q}[\gamma] : \mathbb{Q} < \infty$, then γ is an algebraic number over $\mathbb{Q}$.*

Proof. Let

$$f_{\alpha,K}(x) = x^n + a_1 x^{n-1} + ... + a_n \in K[x]$$

be the minimal polynomial of α. Thus, $a_n \neq 0$ and

$$\alpha^n = -a_1 \alpha^{n-1} - ... - a_n.$$

So,

$$\begin{aligned} \alpha^{n+1} &= -a_1 \alpha^n - ... - a_n \alpha = \\ &= a_1[a_1 \alpha^{n-1} + ... + a_n] - a_2 \alpha^{n-1} - ... - a_n \alpha = \\ &= b_1 \alpha^{n-1} + b_2 \alpha^{n-2} + ... + b_n \in K[\alpha]. \end{aligned}$$

Continuing in this way, we get:

$$\alpha^{n+2} = \alpha \left[b_1 \alpha^{n-1} + b_2 \alpha^{n-2} + ... + b_n \right] =$$

$$= b_1[-a_1\alpha^{n-1} - ... - a_n] + b_2\alpha^{n-1} + ... + b_n\alpha =$$
$$= c_1\alpha^{n-1} + c_2\alpha^{n-2} + ... + c_n \in K[\alpha].$$

Thus, for any $k \in \mathbb{N}^*$, one obtains

$$\alpha^{n+k} = \sum_{j=0}^{n-1} d_j\alpha^j \in K[\alpha].$$

Hence, $\{1, \alpha, ..., \alpha^{n-1}\}$ is a generating system for $K[\alpha]$. Now, if

$$\sum_{j=0}^{t} e_j\alpha^j = 0,\ e_j \in K,\ e_t \neq 0,\ j = 0, 1, ..., t,\ t \leq n-1,$$

then α is a root of the monic polynomial $g(x) = \sum_{j=0}^{n-1} e_j^* x^j \in K[x]$, where $e_j^* = e_j/e_t$, $j = 0, 1, ..., t$. Since $f_{\alpha,K}$ is the minimal polynomial of α and since $\deg g < \deg f_{\alpha,K}$, we obtain a contradiction. Hence $\{1, \alpha, ..., \alpha^{n-1}\}$ is also a linear independent set over K. This means that $\{1, \alpha, ..., \alpha^{n-1}\}$ is a basis of the vector space $K[\alpha]$ over K, so $K[\alpha] : K = n = \deg f_{\alpha,K}$.

Take now $\beta \in K[\alpha]$, $\beta \neq 0$. Since $\{1, \beta, \beta^2, ...\}$ cannot be linear independent over K ($\dim_K K[\alpha]) = n$), we see that there exists a liner combination

$$\lambda_0 + \lambda_1\beta + ... + \lambda_t\beta^t = 0,\ \lambda_j \in K, j = 0, 1, ..., t,$$

such that $\lambda_0, \lambda_t \neq 0$. We know that $\beta^{-1} \in \mathbb{C}$ and, multiplying this last equality by β^{-1}, we get:

$$\beta^{-1} = -\frac{1}{\lambda_0}\left(\lambda_1 + \lambda_2\beta + ... + \lambda_t\beta^{t-1}\right) \in K[\alpha].$$

So, $K[\alpha]$ is a subfield of $\mathbb{C}$, i.e. the least subfield of $\mathbb{C}$, generated by K and α.

For the last statement, we have to remark that the sequence $\{1, \gamma, ..., \gamma^n, ...\}$ cannot be linear independent over $\mathbb{Q}$. So, there exists a nontrivial linear combination

$$\sum_{j=0}^{N-1} l_j\gamma^j + \gamma^N = 0,\ l_j \in \mathbb{Q},$$

etc. □

Corollary 6.4. *Let α, β be algebraic numbers over $\mathbb{Q}$. Then β is an algebraic number over $K = \mathbb{Q}[\alpha]$ and so, $\mathbb{Q}[\alpha][\beta] = \mathbb{Q}[\alpha, \beta]$ is a subfield of $\mathbb{C}$, and*

$$\mathbb{Q}[\alpha, \beta] : \mathbb{Q} = (\mathbb{Q}[\alpha] : \mathbb{Q}) \cdot (\mathbb{Q}[\alpha, \beta] : \mathbb{Q}[\alpha]) = \tag{6.6}$$
$$= (\mathbb{Q}[\beta] : \mathbb{Q}) \cdot (\mathbb{Q}[\alpha, \beta] : \mathbb{Q}[\beta]) < \infty.$$

Hence, the set $\overline{\mathbb{Q}}$ of all the algebraic numbers (over $\mathbb{Q}$) is a subfield of $\mathbb{C}$.

Proof. Since $f_{\beta,\mathbb{Q}} \in \mathbb{Q}[x] \subset \mathbb{Q}[\alpha][x]$ and $f_{\beta,\mathbb{Q}}(\beta) = 0$, we see that $f_{\beta,\mathbb{Q}[\alpha]}$ is a divisor of $f_{\beta,\mathbb{Q}}$ in $\mathbb{Q}[\alpha][x]$. So, β is an algebraic number over $K = \mathbb{Q}[\alpha]$ and Lemma 6.3 says that $K[\beta]$ is a field, and $\mathbb{Q}[\alpha][\beta] : \mathbb{Q} < \infty$. Thus $\mathbb{Q}[\alpha, \beta] : \mathbb{Q} < \infty$ and so, the sets $\{1, \alpha \pm \beta, ..., (\alpha \pm \beta)^m, ...\}$ and $\{1, \alpha\beta, ..., (\alpha\beta)^m, ...\}$ are linear dependent over $\mathbb{Q}$. This means that there exist N and M natural nonzero numbers and $a_j \in \mathbb{Q}$, $j = 0, 1, ..., N-1$such that

$$\sum_{j=0}^{N-1} a_j(\alpha \pm \beta)^j + (\alpha \pm \beta)^N = 0,$$

and $b_j \in \mathbb{Q}$, $j = 0, 1, ..., M-1$ such that

$$\sum_{j=0}^{M-1} b_j(\alpha\beta)^j + (\alpha\beta)^M = 0.$$

It is clear that $\alpha \pm \beta$ is a root of the monic polynomial

$$P(x) = \sum_{j=0}^{N-1} a_j x^j + x^N \in \mathbb{Q}[x],$$

and $\alpha\beta$ is a root of the polynomial

$$Q(x) = \sum_{j=0}^{M-1} b_j x^j + x^M = 0 \in \mathbb{Q}[x].$$

Hence $\alpha \pm \beta$ and $\alpha\beta$ are also algebraic numbers (over $\mathbb{Q}$). $\overline{\mathbb{Q}}$ is a field, because an algebraic number $\gamma \neq 0$ has a reciprocal γ^{-1} even in $\mathbb{Q}[\gamma] \subset \overline{\mathbb{Q}}$ (Lemma 6.3). It remains to prove formula (6.6). Let $\{1, \alpha, ..., \alpha^{n-1}\}$ be a basis of the vector space $\mathbb{Q}[\alpha]$ over $\mathbb{Q}$. Here $n = \deg_{\mathbb{Q}} \alpha$ and let $\{1, \beta, ..., \beta^{m-1}\}$ be a basis of the vector space $\mathbb{Q}[\alpha, \beta]$ over $\mathbb{Q}[\alpha]$. Here $m = \deg_{\mathbb{Q}[\alpha]} \beta$. It is easy to prove that $\{\alpha^i \beta^j\}$, $i = 0, 1, ..., n-1$, $j = 0, 1, 2, ..., m-1$ is a basis of the vector space $\mathbb{Q}[\alpha, \beta]$ over $\mathbb{Q}$. So, one obtains that $\mathbb{Q}[\alpha, \beta] : \mathbb{Q} = nm$, i.e. formula (6.6) □

Remark 6.5. $\overline{\mathbb{Q}} : \mathbb{Q} = \infty$, *because we can find algebraic numbers over $\mathbb{Q}$ of any degree. For instance, let n be a positive integer and $f_n(x) = x^n + 2$. It is easy to see that this last polynomial is irreducible, so any root α_n of it has degree n over $\mathbb{Q}$, i.e. $f_{\alpha_n, \mathbb{Q}} = f_n$.*

Definition 6.6. *Let $K \subset L \subset \mathbb{C}$ be two subfields of $\mathbb{C}$ and let $\sigma : L \to \mathbb{C}$ be a field morphism of L into $\mathbb{C}$, i.e. $\sigma(\alpha \pm \beta) = \sigma(\alpha) \pm \sigma(\beta)$, $\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$ and $\sigma(1) = 1$ for any $\alpha, \beta \in L$. We say that σ is a K-embedding of L into $\mathbb{C}$ if $\sigma(\gamma) = \gamma$ for any $\gamma \in K$.*

Remark 6.7. *Any field morphism $\sigma : L \to \mathbb{C}$, where L is a subfield of $\mathbb{C}$, is a one-to-one mapping. Indeed, $\sigma(\alpha) = \sigma(\beta)$, or $\sigma(\alpha - \beta) = 0$ implies $\alpha = \beta$, otherwise $\delta = \alpha - \beta \neq 0$, so there exist δ^{-1} in L. Hence, $1 = \sigma(1) = \sigma(\delta\delta^{-1}) = \sigma(\delta)\sigma(\delta^{-1}) = 0$, a contradiction Thus, $\alpha = \beta$.*

Lemma 6.8. *Let α be an algebraic number over a subfield K of $\mathbb{C}$. Then there exist exactly $n = \deg_K \alpha = \deg f_{\alpha,K}$ K-embeddings $\sigma_1, \sigma_2, ..., \sigma_n$ of $K[\alpha]$ into $\mathbb{C}$, where $\sigma_j(\alpha) = \alpha_j$, $j = 1, 2, ..., n$, with $\alpha_1 = \alpha$, $\alpha_2, ..., \alpha_n$ the roots of $f_{\alpha,K}$, the minimal polynomial of α over K.*

Proof. Let

$$f_{\alpha,K}(x) = a_0 + a_1 x + ... + a_{n-1} x^{n-1} + x^n \in K[x]$$

be the minimal polynomial of α over K, and let σ be a K-embedding of K into $\mathbb{C}$. Then

$$0 = \sigma(a_0 + a_1\alpha + ... + a_{n-1}\alpha^{n-1} + \alpha^n) =$$

$$= a_0 + a_1\sigma(\alpha) + ... + a_{n-1}\sigma(\alpha)^{n-1} + \sigma(\alpha)^n.$$

Thus $\sigma(\alpha)$ is a root of $f_{\alpha,K}$. Since $f_{\alpha,K}$ has simple roots, $\sigma(\alpha) = \alpha_j$, one of the conjugate of α. Hence, there exist exactly n distinct K-embeddings of $K[\alpha]$ into $\mathbb{C}$. □

Lemma 6.9. *Let $K \subset L$ be two subfields in $\mathbb{C}$ and $\sigma_0 : L \to \mathbb{C}$ be a fixed K-embedding of L into $\mathbb{C}$. Let α be a fixed algebraic number over L with $n = \deg_L \alpha$. Then there exist exactly n K-embeddings μ of $L[\alpha]$ into $\mathbb{C}$ such that $\mu(\lambda) = \sigma_0(\lambda)$ for any $\lambda \in L$. We say that μ extends σ_0 to $L[\alpha]$.*

Proof. Let

$$b = a_0 + a_1\alpha + ... + a_{n-1}\alpha^{n-1}, \ a_0, a_1, ..., a_{n-1} \in L$$

be an arbitrary element in $L[\alpha]$ (Lemma 6.3) and let μ be a K-embedding of $L[\alpha]$ into $\mathbb{C}$ such that the restriction of it to L is σ_0. Thus,

$$\mu(b) = \sigma_0(a_0) + \sigma_0(a_1)\mu(\alpha) + ... + \sigma_0(a_{n-1})\mu(\alpha)^{n-1}.$$

From Lemma 6.8 we know that $\mu(\alpha) = \alpha_j$, where α_j is one of the roots of the minimal polynomial of α over L. Since $n = \deg f_{\alpha,L}$, we see that one has exactly n distinct extensions of σ_0 to $L[\alpha]$. □

Corollary 6.10. *Let α, β be two algebraic numbers over $\mathbb{Q}$. Then the number of the $\mathbb{Q}$-embeddings of $\mathbb{Q}[\alpha, \beta]$ into $\mathbb{C}$ is equal to $\mathbb{Q}[\alpha, \beta] : \mathbb{Q}$.*

Proof. For each $\mathbb{Q}$-embedding $\sigma_i : \mathbb{Q}[\alpha] \to \mathbb{C}$, $i = 1, 2, ..., n = \deg_{\mathbb{Q}} \alpha$ (Lemma 6.8) one has $\mathbb{Q}[\alpha, \beta] : \mathbb{Q}[\alpha] = m$, $\mathbb{Q}$-embeddings μ_{ij}, $j = 1, 2, ..., m$ which extends σ_i to $\mathbb{Q}[\alpha, \beta]$, i.e. such that $\mu_{ij}(\lambda) = \sigma_i(\lambda)$ for any $\lambda \in \mathbb{Q}[\alpha]$ (Lemma 6.9). It remains to prove that $\mathbb{Q}[\alpha, \beta] : \mathbb{Q} = (\mathbb{Q}[\alpha, \beta] : \mathbb{Q}[\alpha]) \cdot (\mathbb{Q}[\alpha] : \mathbb{Q})$. But this is exactly formula (6.6) which was proved in corollary 6.4. □

Remark 6.11. *One can easily generalize Corollary 6.10. Let $K \subset L$ be two subfields in $\mathbb{C}$ such that $L : K = m < \infty$. Let $\gamma_1, \gamma_2, ..., \gamma_n, ...$ numbers in L such that $\gamma_1 \notin K$, $\gamma_2 \notin K[\gamma_1], ..., \gamma_n \notin K[\gamma_1, \gamma_2, ..., \gamma_{n-1}], ...$ Since $K[\gamma_{j+1}] : K[\gamma_j] > 1$ and $L : K = m$, the previous construction can have only a finite number of steps, i.e. one gets:*

$$K \subset K[\gamma_1] \subset K[\gamma_1, \gamma_2] \subset ... \subset K[\gamma_1, \gamma_2, ..., \gamma_k] = L. \tag{6.7}$$

As in the proof of Corollary 6.4 we can prove that

$$m = L : K = (K[\gamma_1] : K) \cdot (K[\gamma_1, \gamma_2] : K[\gamma_1]) \cdot ... \cdot (L : K[\gamma_1, \gamma_2, ..., \gamma_{k-1}]).$$

We apply now Lemma 6.9 to each inclusion in (6.7) and find that the number of K-embeddings of L into $\mathbb{C}$ is exactly $m = L : K$.

Definition 6.12. *Let $\alpha \in \overline{\mathbb{Q}}$ be an algebraic number (over $\mathbb{Q}$) and let $f_{\alpha,\mathbb{Q}} \in \mathbb{Q}[x]$ be its minimal polynomial (over $\mathbb{Q}$). The set $\mathcal{O}(\alpha)$ of all conjugates of α (over $\mathbb{Q}$), i.e. the set of all the roots of $f_{\alpha,\mathbb{Q}}$ is said to be the orbit of α.*

It is clear that if $\alpha_1 = \alpha$, $\alpha_2, ..., \alpha_n$ ($n = \deg f_{\alpha,\mathbb{Q}}$) are all the roots of $f_{\alpha,\mathbb{Q}}$, then $\mathcal{O}(\alpha_j) = \mathcal{O}(\alpha)$ for any $j = 2, 3, ..., n$, because $f_{\alpha_j,\mathbb{Q}} = f_{\alpha,\mathbb{Q}}$ for $j = 2, 3, ..., n$. Moreover, if $\sigma : K \to \mathbb{C}$ is a $\mathbb{Q}$-embedding of a subfield K of $\mathbb{C}$, which contains $\mathbb{Q}[\alpha]$, then $\sigma(\mathcal{O}(\alpha)) = \mathcal{O}(\alpha)$.

Definition 6.13. *Let M be a finite set of algebraic numbers (over $\mathbb{Q}$). We say that M is closed to $\mathbb{Q}$-embeddings if for any $\mathbb{Q}$-embedding $\sigma : \overline{\mathbb{Q}} \to \mathbb{C}$ of the subfield $\overline{\mathbb{Q}}$ of $\mathbb{C}$, one has $\sigma(M) \subset M$ (in fact $\sigma(M) = M$, because σ is a one-to-one mapping-see Remark 6.7).*

Example 6.14. *An orbit $\mathcal{O}(\alpha)$ over $\mathbb{Q}$ is closed to $\mathbb{Q}$-embeddings.*

Proposition 6.15. *Any finite subset M of $\overline{\mathbb{Q}}$, which is closed to $\mathbb{Q}$-embeddings is a finite (disjoint) union of orbits over $\mathbb{Q}$.*

Proof. Let $m = |M|$ the number of elements of M. We proceed by induction on m. For $m = 1$, we see that $M = \{q\}$, $q \in \mathbb{Q}$, so $M = \mathcal{O}(q)$. Assume that our statement is true for $k = 1, 2, ..., m - 1$ and let us prove it for $k = m$. Let β_1 be in M. Since M is closed to $\mathbb{Q}$-embeddings, $\mathcal{O}(\beta_1) \subset M$. Since $|\mathcal{O}(\beta_1)| \geq 1$ and since $M \setminus \mathcal{O}(\beta_1)$ is also closed to $\mathbb{Q}$-embeddings, we apply the induction hypothesis and find that

$$M \setminus \mathcal{O}(\beta_1) = \bigcup_{j=2}^{N} \mathcal{O}(\beta_j),$$

where the last union is a disjoint one. So,

$$M = \bigcup_{j=1}^{N} \mathcal{O}(\beta_j).$$

□

Let $\alpha_1 = \alpha$, $\alpha_2, ..., \alpha_n$ ($n = \deg f_{\alpha,\mathbb{Q}}$) be the orbit of an algebraic number α over $\mathbb{Q}$. Since any $\mathbb{Q}$-embedding $\sigma : \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n] \to \mathbb{C}$ give rise to a permutation of $\alpha_1, \alpha_2, ..., \alpha_n$, we see that $K = \sigma(\mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n]) \subset \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n]$. Moreover, σ is a one-to-one mapping, so

$$K : \mathbb{Q} = \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n] : \mathbb{Q},$$

i.e. $K = \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n]$. Thus, any $\mathbb{Q}$-embedding σ of K is a one-to-one and onto field morphism on K, so the set of all such morphisms is a group under the composition law of functions. We call this group the

Galois group of K and denote it $Gal(K/\mathbb{Q})$, or simple $G(K)$ or G. If instead of $\mathbb{Q}$ we take an arbitrary subfield L of $K = \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n]$, we also get the Galois group $Gal(K/L)$ of all L-embeddings of K onto K. From Remark 6.11 we see that the number $|Gal(K/L)|$, the number of the elements of $Gal(K/L)$, the order of $Gal(K/L)$, is equal to $K : L$, i.e. with the degree of the extension $L \subset K$.

Since in Section 8 we use the Galois groups $G = Gal(K/L)$ of order a power of $p = 2$, we recall a basic result from finite group theory. We say that a subgroup H of G is a *normal subgroup* of G if $x^{-1}Hx = H$ for any $x \in G$. In this last case, the equivalence relation: $x \sim y$ if and only if $x^{-1}y \in H$, give rise to a group structure on the set of equivalence classes $G/\sim$. This last group of classes is denoted G/H. If G is a finite group of order $|G|$ (the number of its elements) then, by counting the number of classes (each class has the same number of elements), we obtain Lagrange's formula

$$|G| = |H| \cdot |G/H|.$$

Lemma 6.16. (*Corollary 6.6,* [7]) *Let G be a finite group of order $|G| = p^n$, where p is a prime number and $n > 0$. Then there exists a tower of normal subgroups $\{G_i\}$, $i = 0, 1, ..., n$,*

$$\{1\} = G_0 \subset G_1 \subset ... \subset G_n = G,$$

such that $|G_i/G_{i-1}| = p$ for all $i = 1, 2, ..., n$.

Let us assume now that $G = Gal(K/L)$, where $K = \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n]$, L is a subfield of K, $L \neq K$, $\{\alpha = \alpha_1, \alpha_2, ..., \alpha_n\} = \mathcal{O}(\alpha)$ and $|G| = 2^n$.

Corollary 6.17. *With the above notation and hypotheses, there exists a tower of subfields*

$$L = K_0 \subset K_1 \subset ... \subset K_n = K,$$

such that $K_i : K_{i-1} = 2$ for any $i = 1, 2, .., n$. Thus, there exist $\beta_1, \beta_2, ..., \beta_n \in K$ such that $K_i = K_{i-1}[\beta_i]$ for $i = 1, 2, ..., n$. In particular, $K = L[\beta_1, \beta_2, ..., \beta_n]$.

Proof. It is sufficient to take $K_i = \{x \in K : \sigma(x) = x, \sigma \in G_{n-i}\}$, where $\{G_i\}$, $i = 0, 1, ..., n$, is the tower of groups from Lemma 6.16 and $G = Gal(K/L)$. Since the number of K_{i-1}-embeddings of K_i into $\mathbb{C}$ is equal to $|G_i/G_{i-1}| = 2$, we see that $K_i : K_{i-1} = 2$ for any $i = 1, 2, .., n$. Thus, if we take any $\beta_i \in K_i \setminus K_{i-1}$, we see that $K_i = K_{i-1}[\beta_i]$ for $i = 1, 2, ..., n$. □

Definition 6.18. *Let K be a subfield of $\mathbb{C}$ such that $K : \mathbb{Q} = n$ (always $\mathbb{Q} \subset K$) and let $\sigma_1, \sigma_2, ..., \sigma_n$ the distinct $\mathbb{Q}$-embeddings of K into $\mathbb{C}$ (see Remark 6.11). For any $\alpha \in K$ (so α is an algebraic number over $\mathbb{Q}$, because $\mathbb{Q}[\alpha] : \mathbb{Q} \leq n$, Lemma 6.3) we define the following rational number (it is equal to a power of $(-1)^m f_{\alpha,\mathbb{Q}}(0)$), where $m = \deg f_{\alpha,\mathbb{Q}}$-see Lemma 6.20 bellow) :*

$$N_K(\alpha) = \sigma_1(\alpha) \cdot \sigma_2(\alpha) \cdot ... \cdot \sigma_n(\alpha),$$

and we call it the K-norm of α over $\mathbb{Q}$. If $K = \mathbb{Q}[\alpha]$, we simply denote $N_{\mathbb{Q}[\alpha]}(\alpha)$ by $N(\alpha)$, and call it the norm of α (over $\mathbb{Q}$).

Lemma 6.19. *Let $\alpha \in \overline{\mathbb{Q}}$ be an algebraic number (over $\mathbb{Q}$) and $f_{\alpha,\mathbb{Q}} \in \mathbb{Q}[x]$ be its minimal polynomial. Then $N(\alpha) = (-1)^n f_{\alpha,\mathbb{Q}}(0)$, where $n = \deg_{\mathbb{Q}} f_{\alpha,\mathbb{Q}}$.*

Proof. Let $\sigma_1, \sigma_2, ..., \sigma_n$ the distinct $\mathbb{Q}$-embeddings of $\mathbb{Q}[\alpha]$ into $\mathbb{C}$.

Since $\sigma_1(\alpha), \sigma_2(\alpha), ..., \sigma_n(\alpha)$ are all the conjugates of α, the last Viète formula says that

$$N(\alpha) = \sigma_1(\alpha) \cdot \sigma_2(\alpha) \cdot ... \cdot \sigma_n(\alpha) = (-1)^n f_{\alpha,\mathbb{Q}}(0).$$

□

Lemma 6.20. *Let K be a subfield of $\mathbb{C}$ such that $K : \mathbb{Q} = n$, and let α be in K such that $\alpha_1 = \alpha, \alpha_2, ..., \alpha_m$ are all its conjugates. Then*

$$N_K(\alpha) = (\alpha_1 \alpha_2 ... \alpha_m)^{K:\mathbb{Q}[\alpha]} = [(-1)^m f_{\alpha,\mathbb{Q}}(0)]^{K:\mathbb{Q}[\alpha]} \in \mathbb{Q}. \tag{6.8}$$

Proof. The last equality comes from Lemma 6.19. To prove the first equality, we have to see that any $\mathbb{Q}$-embedding $\sigma_i : \mathbb{Q}[\alpha] \to \mathbb{C}$ is completely defined by $\sigma_i(\alpha) = \alpha_i$, $i = 1, 2, ..., n$ (Lemma 6.8) and it can be extended to $l = K : \mathbb{Q}[\alpha]$ distinct $\mathbb{Q}$-embeddings μ_{ij}, $j = 1, 2, ..., l$ of K into $\mathbb{C}$ (Lemma 6.9 and Remark 6.11). □

Lemma 6.21. *With the above notation, definitions and hypotheses, the mapping $N_K : K \to \mathbb{Q}$ is a multiplicative mapping, i.e.*

$$N_K(\alpha\beta) = N_K(\alpha) \cdot N_K(\beta)$$

for any $\alpha, \beta \in K$.

Proof. We use only the multiplicative property of a field morphism. Indeed, let $\sigma_1, \sigma_2, ..., \sigma_n$ be all the distinct $\mathbb{Q}$-embeddings of K into $\mathbb{C}$. Thus,

$$N_K(\alpha\beta) = \sigma_1(\alpha\beta) \cdot \sigma_2(\alpha\beta) \cdot ... \cdot \sigma_n(\alpha\beta) =$$
$$= [\sigma_1(\alpha) \cdot ... \cdot \sigma_n(\alpha)] \cdot [\sigma_1(\beta) \cdot ... \cdot \sigma_n(\beta)] = N_K(\alpha) \cdot N_K(\beta).$$

□

Corollary 6.22. *Let q be in $\mathbb{Q}$ and let α be an algebraic number (over $\mathbb{Q}$) with $n = \deg_{\mathbb{Q}} \alpha$. Then*

$$N(q\alpha) = q^n N(\alpha). \tag{6.9}$$

Proof. Let $\sigma_1, \sigma_2, ..., \sigma_n$ be the $\mathbb{Q}$-embedding of $\mathbb{Q}[\alpha]$ into $\mathbb{C}$. So,

$$N(q\alpha) \overset{def}{=} N_{\mathbb{Q}[\alpha]}(q\alpha) = \sigma_1(q\alpha) \cdot \sigma_2(q\alpha) \cdot ... \cdot \sigma_n(q\alpha) =$$
$$= q\sigma_1(\alpha) \cdot q\sigma_1(\alpha) \cdot ... \cdot q\sigma_n(\alpha) = q^n N(\alpha).$$

□

Definition 6.23. *Let $\alpha \in \mathbb{C}$ be a complex number. If α is a root of a monic polynomial $P \in \mathbb{Z}[x]$, we say that α is an algebraic integer.*

Lemma 6.24. *Let $\alpha \in \mathbb{Q}$ be an algebraic integer. Then $\alpha \in \mathbb{Z}$.*

Proof. Assume that $\alpha = a/b$, $a, b \in \mathbb{Z}^*$, $b > 0$ and $(a, b) = 1$, i.e. there is no prime number which divide both a and b. Let

$$P(x) = c_0 + c_1 x + ... + c_{n-1}x^{n-1} + x^n \in \mathbb{Z}[x],$$

be such that $P(\alpha) = 0$. Thus

$$c_0 b^n + c_1 b^{n-1} a + ... + c_{n-1} b a^{n-1} + a^n = 0,$$

and so, if $b \neq 1$, take a prime number p which divide b and see from this last equality that p also divide a, a contradiction. Hence $b = 1$, i.e. $\alpha \in \mathbb{Z}$. □

Lemma 6.25. *For any algebraic number α (over $\mathbb{Q}$) there is a positive integer $d > 0$, such that $d\alpha$ is an algebraic integer. Moreover, $\deg_{\mathbb{Q}}(d\alpha) = \deg_{\mathbb{Q}} \alpha$ and $f_{d\alpha,\mathbb{Q}} \in \mathbb{Z}[x]$.*

Proof. Let

$$f_{\alpha,\mathbb{Q}}(x) = c_0 + c_1 x + ... + c_{n-1}x^{n-1} + x^n \in \mathbb{Q}[x]$$

be the monic minimal polynomial of α, and let $c_j = a_j/b_j$, $(a_j, b_j) = 1$, and $b_j > 0$ for any $j = 0, 1, ..., n-1$. Thus,

$$\frac{a_0}{b_0} + \frac{a_1}{b_1}\alpha + ... + \frac{a_{n-1}}{b_{n-1}}\alpha^{n-1} + \alpha^n = 0. \tag{6.10}$$

Let us multiply the equality (6.10) by d^n, where d is the LCM of $b_0, b_1, ..., b_{n-1}$, and find:

$$\frac{d^n a_0}{b_0} + \frac{d^{n-1}a_1}{b_1}(d\alpha) + ... + \frac{d a_{n-1}}{b_{n-1}}(d\alpha)^{n-1} + (d\alpha)^n = 0.$$

Since

$$e_0 = \frac{d^n a_0}{b_0}, e_1 = \frac{d^{n-1}a_1}{b_1}, ..., e_{n-1} = \frac{da_{n-1}}{b_{n-1}} \in \mathbb{Z},$$

we see that $d\alpha$ is a root of the polynomial

$$g(x) = e_0 + e_1 x + ... + e_{n-1}x^{n-1} + x^n \in \mathbb{Z},$$

so $d\alpha$ is an algebraic integer. Moreover, since $\mathbb{Q}[\alpha] = \mathbb{Q}[d\alpha]$ (both being fields and $d \in \mathbb{Z} \subset \mathbb{Q}$), $\deg_{\mathbb{Q}} \alpha = \mathbb{Q}[\alpha] : \mathbb{Q} = \mathbb{Q}[d\alpha] : \mathbb{Q} = \deg_{\mathbb{Q}}(d\alpha)$, so $f_{d\alpha,\mathbb{Q}} = g \in \mathbb{Z}[x]$. □

Lemma 6.26. *Let α, β be two algebraic integers (over $\mathbb{Q}$) in $\mathbb{C}$. Then $\alpha \pm \beta$ and $\alpha\beta$ are also algebraic integers (over $\mathbb{Q}$). In particular, $\mathbb{A}$ is a subring of $\overline{\mathbb{Q}}$.*

Proof. Let

$$h_\alpha(x) = a_0 + a_1 x + ... + a_{n-1}x^{n-1} + x^n \in \mathbb{Z}[x],$$
$$h_\beta(x) = b_0 + b_1 x + ... + b_{m-1}x^{m-1} + x^m \in \mathbb{Z}[x]$$

be two polynomials with integer coefficients, such that $h_\alpha(\alpha) = 0$ and $h_\beta(\beta) = 0$. So, any element s of $\mathbb{Z}[\alpha, \beta] = \mathbb{Z}[\alpha][\beta]$ can be written as

$$s = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} c_{ij}\alpha^i\beta^j,\ c_{ij} \in \mathbb{Z}.$$

Let us denote $\omega_1, \omega_2, ..., \omega_k$, $k = nm$, the set of generators $\{\alpha^i\beta^j\}$ of $\mathbb{Z}[\alpha, \beta]$, where $i = 0, 1, ..., n-1$, $j = 0, 1, ..., m-1$, and let us take an arbitrary element γ of $\mathbb{Z}[\alpha, \beta]$. Thus, for any $i = 1, 2, ..., k$ on can write:

$$\gamma\omega_i = \sum_{j=1}^{k} a_{ij}\omega_j,\ a_{ij} \in \mathbb{Z},\ i, j \in \{1, 2, ..., k\}. \tag{6.11}$$

Denote $A = (a_{ij})$ the $k \times k$ matrix with the integer entries a_{ij}, $i, j = 1, 2, ..., k$. So (6.11) can also be written:

$$B \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where $B = \gamma \cdot I - A$, with I the $k \times k$ identity matrix. Since $\omega_1, \omega_2, ..., \omega_k$ is a nontrivial solution of the above homogenous system, we see that $\det B = 0$. Thus γ is a root of the monic polynomial $P(x) = \det(xI - A)$ with coefficients in $\mathbb{Z}$. Hence γ is also an algebraic integer (over $\mathbb{Q}$). For $\gamma = \alpha \pm \beta$, $\alpha\beta$ we obtain the statement of the lemma. □

Remark 6.27. *Let α be an algebraic integer and $g \in \mathbb{Z}[x]$ be a monic polynomial of the least degree such that $g(\alpha) = 0$. Then $f_{\alpha,\mathbb{Q}} = g$, so $f_{\alpha,\mathbb{Q}} \in \mathbb{Z}[x]$. Indeed, since $g \in \mathbb{Z}[x] \subset \mathbb{Q}[x]$ and $g(\alpha) = 0$, we see that $f_{\alpha,\mathbb{Q}}$ is a divisor of g in $\mathbb{Q}[x]$. Since all the conjugates of α are roots of g, i.e, since $\sigma(g(\alpha)) = g(\sigma(\alpha))$ for any $\mathbb{Q}$-embedding of $\mathbb{Q}[\alpha]$, all these roots are algebraic integers. From Viète formulas and from the Lemmas 6.24, 6.26, we see that the coefficients of $f_{\alpha,\mathbb{Q}}$ are integers, i.e. $f_{\alpha,\mathbb{Q}} = g$ (because of the minimally of g). Therefore, α is an algebraic integer if and only if $f_{\alpha,\mathbb{Q}}$ has integer coefficients. However, practically, it is not a good idea to use this last characterization of an algebraic integer as a definition.*

Definition 6.28. *We say that an algebraic integer α is divisible by a nonzero integer n if there exists another algebraic integer β such that $\alpha = n\beta$ in $\mathbb{A}$.*

Lemma 6.29. *Let α be a nonzero algebraic integer. Then there exist only a finite number of prime numbers p which divide α.*

Proof. Assume that $\alpha = p\beta$, where p is a prime number and β is another algebraic integer. Let $K = \mathbb{Q}[\alpha,\beta]$, the least subfield of $\mathbb{C}$ which contains $\mathbb{Q}$, α and β. Thus,

$$N_K(\alpha) = p^{K:\mathbb{Q}} N_K(\beta),$$

(see Lemmas 6.20, 6.21). Since $N_K(\alpha)$ and $N_K(\alpha)$ are integers, we see that p is a divisor of the fixed number $N_K(\alpha)$, so such a p takes a finite number of values. □

7. A simple proof for the basic Lindemann's theorem

Remark 7.1. *Here are some easy remarks for a reader who has no background in complex function theory. Let $\alpha = a + ib$, $a, b \in \mathbb{R}$, $i = \sqrt{-1}$, be a complex number. Then, by e^α we understand the following complex number:*

$$e^\alpha = e^a \cos b + ie^a \sin b.$$

Using elementary trigonometry, it is easy to prove that

$$e^{\alpha+\beta} = e^\alpha \cdot e^\beta, \text{ and } e^{k\alpha} = (e^\alpha)^k,$$

where $\alpha, \beta \in \mathbb{C}$ and $k \in \mathbb{N}$. So, later in this section, we can freely use equalities of the form:

$$(e^{\alpha_1})^{k_1} \cdot (e^{\alpha_2})^{k_2} \cdot \ldots \cdot (e^{\alpha_n})^{k_n} = e^{\alpha_1 k_1 + \alpha_2 k_2 + \ldots + \alpha_n k_n}$$

for $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ and $k_1, \ldots, k_n \in \mathbb{N}$.

The next result is the famous Lindemann's Theorem [8]. In the proof of Theorem 7.2 we mix the main Lindemann idea with some ideas from the proof of a more general result, known as Weierstrass-Lindemann Theorem [2].

Theorem 7.2. *Let α be a nonzero algebraic number, i.e. a nonzero root of a polynomial with rational coefficients. Then e^α is a transcendental number. In particular, since $e^{\pi i} = -1$, π cannot be an algebraic number, i.e. π is a transcendental number.*

Proof. Lemma 6.25 says that there exists a positive integer $d > 0$ such that $d \cdot \alpha$ is an algebraic integer. If e^α was an algebraic number, then $(e^\alpha)^d$ would be also an algebraic number ($\overline{\mathbb{Q}}$ is a field-see Corollary 6.4). Thus, it is sufficient to prove the statement for α an algebraic integer. We also can assume that $\alpha \notin \mathbb{Z}$. Indeed, if α is a positive integer, we proved the statement in Section 5. If α is a negative integer, then $e^\alpha = (e^{-\alpha})^{-1}$ cannot be an algebraic number because $e^{-\alpha}$ is not an algebraic number (see Section 5). Since α is a nonzero algebraic integer, $\alpha \notin \mathbb{Z}$, its monic minimal polynomial $P_{\alpha,\mathbb{Q}}$ has integer coefficients (Remark 6.27). Thus,

$$P_{\alpha,\mathbb{Q}}(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0 \in \mathbb{Z}[x], \tag{7.1}$$

$a_0 \neq 0$, $P_{\alpha,\mathbb{Q}}(\alpha) = 0$ and $\deg_{\mathbb{Q}}(\alpha) \geq 2$ (because α is not in $\mathbb{Q}$).

Let $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$ be all the roots of $P_{\alpha,\mathbb{Q}}$, i.e. all the conjugates of α over $\mathbb{Q}$. Since $P_{\alpha,\mathbb{Q}}$ is the minimal polynomial of α, all these roots are distinct (Lemma 6.2) and they are also algebraic integers (Remark 6.27).

Now, let us assume that e^α is an algebraic number, i.e. its minimal monic polynomial $f_{e^\alpha,\mathbb{Q}}$ over $\mathbb{Q}$ has rational coefficients, say:

$$f_{e^\alpha,\mathbb{Q}}(x) = x^m + \frac{b_{m-1}}{b_m}x^{m-1} + \ldots + \frac{b_0}{b_m},$$

where $b_j \in \mathbb{Z}$ for $j = 0, 1, \ldots, m$, $m \geq 1$ and $b_m > 0, b_0 \neq 0$. So, if one denotes

$$Q(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_0 \in \mathbb{Z}[x],$$

we see that

$$Q_1 = Q(e^{\alpha_1}) = \sum_{i=0}^{m} b_i e^{i\alpha_1} = 0, \tag{7.2}$$

where $\alpha_1 = \alpha$.

In general, for any $j = 1, 2, ..., n$, we denote

$$Q_j = Q\left(e^{\alpha_j}\right) = \sum_{i=0}^{m} b_i e^{i\alpha_j}. \tag{7.3}$$

Thus,

$$R = Q_1 \cdot Q_2 \cdot ... \cdot Q_n = 0. \tag{7.4}$$

Let us denote

$$S = \{s = (k_1, k_2, ..., k_n) \in \mathbb{N}^n : 0 \le k_j \le m,\ j = 1, 2, ..., n\}.$$

Thus (7.4) can also be written:

$$R = \sum_{s=(k_1,k_2,...,k_n)\in S} b_{k_1} b_{k_2} ... b_{k_n} e^{\alpha(s)} = 0, \tag{7.5}$$

where $\alpha(s) = k_1\alpha_1 + k_2\alpha_2 + ... + k_n\alpha_n$.

Let us choose only the distinct such $\alpha(s)$ and let us denote them by $\beta_1, \beta_2, ..., \beta_t$. For such an algebraic integer $\alpha(s) = \beta_j$ we denote b_j^* the sum of all $b_{k_1} b_{k_2} ... b_{k_n}$ which appear as coefficients in front of a $e^{\alpha(s)}$, which is equal to e^{β_j}. If $b_j^* = 0$, we do not count the term $b_j^* e^{\beta_j}$ in the sum (7.5) and change the order of $\beta_1, \beta_2, ..., \beta_t$ such that all $b_j^* e^{\beta_j}$ are not zero and $\beta_1, \beta_2, ..., \beta_t$ are distinct (maybe with another value of t). So, the sum (7.5) can also be written:

$$R = \sum_{j=1}^{t} b_j^* e^{\beta_j} = 0, \tag{7.6}$$

where $b_j^* \in \mathbb{Z}^*$ (see the explanation bellow why we can take only those nonzero b_j^*), $j = 1, 2, ..., t$ and $\beta_1, \beta_2, ..., \beta_t$ are distinct algebraic integers.

This sum is not "empty", i.e. not all $b_j^* = 0$, $j = 1, 2, ..., t$. Indeed, let us consider on $\mathbb{C}$ the lexicographic order " $\prec$ " : $a + ib \prec c + id$, $i = \sqrt{-1}$, $a, b, c, d \in \mathbb{R}$, if $a < c$, or if $a = c$ and $b < d$. For any $l = 1, 2, ..., n$, let us choose the greatest element $k_{t_l}\alpha_l$ relative to " $\prec$ " in the set $T_l = \{k\alpha_l : k = 0, 1, ..., m, b_k \neq 0\}$ (see (7.2) and (7.3)). It is easy to see that $k_{t_l} = 1$, or m and it is unique in T_l with this property. Now, the algebraic integer

$$v = \sum_{l=1}^{n} k_{t_l}\alpha_l$$

is one of the $\{\beta_j\}_{j=1,2,...,t}$ in (7.6), say β_{j_0}. Since $b_0 \neq 0$ and $b_m \neq 0$, we see that $b_{j_0}^* = b_{k_{t_1}} ... b_{k_{t_n}} \neq 0$. Therefore, we can assume that all the integers b_j^* are not zero for any $j = 1, 2, ..., t$.

Relative to the sum in (7.6) we can make the following important remark. Let $\tau : \mathbb{Q}[\beta_1, \beta_2, ..., \beta_t] \to \mathbb{C}$ be a $\mathbb{Q}$-embedding. So,

$$\tau(\alpha(s)) = k_1\tau(\alpha_1) + k_2\tau(\alpha_2) + ... + k_n\tau(\alpha_n).$$

Since

$$\tau(\alpha(s)) = \sigma(k_1)\alpha_1 + \sigma(k_2)\alpha_2 + ... + \sigma(k_n)\alpha_n,$$

where $\{\sigma(k_1), \sigma(k_2), ..., \sigma(k_2)\}$ is a permutation σ of the set $\{k_1, k_2, ..., k_n\}$ and

$\{\tau(\alpha_1), \tau(\alpha_2), ..., \tau(\alpha_n)\}$ is a permutation of $\{\alpha_1, \alpha_2, ..., \alpha_n\}$), the sum of all products $b_{k_{j_1}} b_{k_{j_2}} ... b_{k_{j_n}}$ which appear in front of a $e^{\tau(\alpha(s))}$ is the same like the corresponding sum of $e^{\alpha(s)}$. Hence, if $\alpha(s) = \beta_j$ and $\tau(\alpha(s)) = \beta_l$, then $b_j^* = b_l^*$, where b_l^* is the coefficient which appears in front of e^{β_l}. Thus, $b_l^* \neq 0$, i.e. the sum in (7.6) is "invariable" to any $\mathbb{Q}$-embedding τ of $\mathbb{Q}[\beta_1, \beta_2, ..., \beta_t]$. Thus, all the elements of any orbit $\mathcal{O}(\beta_j)$, $j = 1, 2, ..., t$ are also in the set $\{\beta_1, \beta_2, ..., \beta_t\}$, i.e. this last set is a disjoint union of orbits $\mathcal{O}(\beta_{j_q})$, $q = 1, 2, ..., h$, for a subset $\{\beta_{j_1}, \beta_{j_2}, ..., \beta_{j_h}\}$ of $\{\beta_1, \beta_2, ..., \beta_t\}$.

Now, we come back to Hermite Principle and consider the polynomial

$$H(x)=\prod_{j=1}^{t}(x-\beta_j)\in\mathbb{A}[x],$$

where $\mathbb{A}$ is the ring of all algebraic integers in $\mathbb{C}$. Let us denote $W=\{\beta_1,\beta_2,...,\beta_t\}$, let us fix an arbitrary prime number p, and let us fix a $\gamma\in W$. With this notation, let us consider the following polynomial:

$$f_{p,\gamma}(x)=\frac{1}{(p-1)!}\frac{[H(x)]^p}{x-\gamma}\in\frac{1}{(p-1)!}\mathbb{A}[x]\subset\overline{\mathbb{Q}}[x],$$

where $\overline{\mathbb{Q}}$ is the field of all algebraic numbers. Here $\deg f_{p,\gamma}=tp-1=d_p$. This polynomial is a generalization of the polynomial g_p from (5.2) in Section 5.

The next integral is a complex integral of an analytic function, computed (for instance) on complex segments $[\gamma,\delta]$, $[0,\delta]$, $[0,\gamma]$, etc. We consider

$$I_{p,\gamma}(\delta)=\int_{\gamma}^{\delta}f_{p,\gamma}(x)e^{-x}\,dx, \tag{7.7}$$

where δ is another element of W. We integrate by parts $I_{p,\gamma}(\delta)$ d_p-times and find (see also (5.5)):

$$I_{p,\gamma}(\delta)=-e^{-\delta}F_{p,\gamma}(\delta)+e^{-\gamma}F_{p,\gamma}(\gamma), \tag{7.8}$$

where

$$F_{p,\gamma}(x)=f_{p,\gamma}(x)+f'_{p,\gamma}(x)+...+f_{p,\gamma}^{(d_p)}(x). \tag{7.9}$$

We prove now that $I_{p,\gamma}(\delta)\to 0$, when $p\to\infty$, for any fixed $\gamma,\delta\in W$. Let us take an open disc D of radius R such that $W\subset D\subset\mathbb{C}$ and take $M\geq 2R$ such that $|e^{-x}|\leq M$ for any $x\in D$. Since $|x-\beta_k|<2R\leq M$ for all $k=1,...,t$, it follows that

$$|I_{p,\gamma}(\delta)|\leq\frac{M^{tp}}{(p-1)!}\to 0,$$

as $p\to\infty$.

Now, let us come back to formula (7.9) and compute $f_{p,\gamma}^{(j)}(x)$ at $x=\delta\neq\gamma$ (if $\delta=\gamma$, $I_{p,\gamma}(\delta)=0$) for any $j=1,2,...,d_p$. Recall that

$$U_\gamma(x)=\prod_{\theta\in W,\ \theta\neq\gamma}(x-\theta)^p. \tag{7.10}$$

Thus, $f_{p,\gamma}^{(j)}(\delta)=0$ for any $0\leq j<p$ and

$$f_{p,\gamma}^{(j)}(\delta)=\frac{1}{(p-1)!}\sum_{k=0}^{j}\binom{j}{k}\left[(x-\gamma)^{p-1}\right]^{(j-k)}[U_\gamma(x)]^{(k)}\Big|_{x=\delta} \tag{7.11}$$

for $d_p\geq j\geq p$. Hence,

$$f_{p,\gamma}^{(j)}(\delta)=M_j\cdot p,$$

where M_j is an algebraic integer (inclusive zero) for any $j=0,1,...,d_p$.

We easily see that $f_{p,\gamma}^{(j)}(\gamma)=0$ for any $0\leq j<p-1$. For $p<j\leq d_p$, we see that

$$\begin{aligned}f_{p,\gamma}^{(j)}(\gamma)&=\frac{1}{(p-1)!}\sum_{k=0}^{j}\binom{j}{k}\left[(x-\gamma)^{p-1}\right]^{(j-k)}[U_\gamma(x)]^{(k)}\Big|_{x=\gamma}\\&=N_j\cdot p,\end{aligned} \tag{7.12}$$

where N_j is an algebraic integer. Now

$$f_{p,\gamma}^{(p-1)}(\gamma)=U_\gamma(\gamma)$$

which is a nonzero algebraic integer (see (7.10)). Finally, we see that

$$F_{p,\gamma}(\delta) = M \cdot p$$

for $\delta \neq \gamma$, and $F_{p,\gamma}(\gamma) = N \cdot p + U_\gamma(\gamma)$, where M, N are both algebraic integers. Now, let us come back to formulas (7.8) and (7.6), and compute

$$\sum_{i=1}^{t} b_i^* e^{\beta_i} I_{p,\gamma}(\beta_i) = -\sum_{i=1}^{t} b_i^* F_{p,\gamma}(\beta_i) + e^{-\gamma} F_{p,\gamma}(\gamma) \sum_{i=1}^{t} b_i^* e^{\beta_i}.$$

Since $R = \sum_{i=1}^{t} b_i^* e^{\beta_i} = 0$ (see (7.6)), we get:

$$\sum_{i=1}^{t} b_i^* e^{\beta_i} I_{p,\gamma}(\beta_i) = -\sum_{i=1}^{t} b_i^* F_{p,\gamma}(\beta_i) = L \cdot p - b_k^* U_{\beta_k}(\beta_k),$$

where $\beta_k = \gamma$ and L is an algebraic integer. Since $b_k^* U_{\beta_k}(\beta_k)$ is a nonzero algebraic integer, not divisible by p for p large enough (see Lemma 6.29), we see that for p large enough

$$S_k(p) \stackrel{def}{=} -\sum_{i=1}^{t} b_i^* e^{\beta_i} I_{p,\beta_k}(\beta_i)$$

is a nonzero algebraic integer. We see that

$$S(p) = \prod_{k=1}^{t} S_k(p) = (-1)^t \prod_{k=1}^{t} \left[\sum_{i=1}^{t} b_i^* F_{p,\beta_k}(\beta_i) \right] \tag{7.13}$$

is a symmetric polynomial in the "variables" $\beta_1, \beta_2, ..., \beta_t$ with integer coefficients. Now, the set $W = \{\beta_1, \beta_2, ..., \beta_t\} \subset \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n]$ is closed to any $\mathbb{Q}$-embedding $\sigma : \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_n] \to \mathbb{C}$. From Proposition 6.15 we see that $W = \cup_{j=1}^{h} \mathcal{O}(\beta_{i_j})$, where β_{i_j}, $j = 1, 2..., h$, are some elements of W, and $\mathcal{O}(\beta_{i_j})$ is the orbit of β_{i_j}. Let $V_{i_j} \in \mathbb{Z}[x]$ be the monic irreducible polynomial which has as the roots the element of $\mathcal{O}(\beta_{i_j})$. We see that the elements of W are the roots of the polynomial

$$V(x) = \prod_{j=1}^{h} V_{i_j}(x) \in \mathbb{Z}[x].$$

Now,

$$S(p) = (-1)^t \prod_{j=1}^{h} \prod_{\beta_{i_0} \in \mathcal{O}(\beta_{i_j})} \left[\sum_{i=1}^{t} b_i^* F_{p,\beta_{i_0}}(\beta_i) \right]$$

So, we can apply Theorem 6.1 for each polynomial $P = V_{i_j}$ and find that each $\prod_{\beta_{i_0} \in \mathcal{O}(\beta_{i_j})} \left[\sum_{i=1}^{t} b_i^* F_{p,\beta_{i_0}}(\beta_i) \right]$ is an integer number, i.e. $S(p) \in \mathbb{Z}$. Moreover, it is a nonzero integer for p large enough, because for p sufficiently large each i_0-factor, $\sum_{i=1}^{t} b_i^* F_{p,\beta_{i_0}}(\beta_i)$, in the product (7.13) is not zero ($b_{i_0}^* U_{\beta_{i_0}}(\beta_{i_0})$ is not divisible by p for p large enough). But $S(p) \to 0$, when $p \to \infty$, a contradiction. Indeed, for any $i_0 \in W$, $S_{i_0}(p) \to 0$, when $p \to \infty$, because each $I_{p,\beta_{i_0}}(\beta_i) \to 0$, when $p \to \infty$ for $i = 1, 2, ..., t$. Hence e^α cannot be an algebraic number, i.e it is a transcendental number.

In particular, if π was an algebraic number, then πi would be an algebraic number and so, $e^{\pi i} = -1$ is a transcendental number (from the first part of Theorem 7.2), a new contradiction. Hence π is a transcendental number. □

Theorem 7.2 has many consequences.

Corollary 7.3. *Let r be a nonzero real algebraic number (Exp. $\sqrt{2}$, $\sqrt[3]{5 + 4\sqrt[5]{3}}$, etc.). Then $\sin r$ ($r \neq 0$), $\cos r$, $\tan r$, $\ln r$ ($r > 0, r \neq 1$) are real transcendental numbers.*

Proof. Let us assume that $\sin r$ is an algebraic number. Since $\sin^2 r + \cos^2 r = 1$, we see that $\cos r = \pm\sqrt{1-\sin^2 r} \in \mathbb{Q}[\sin r][\cos r]$, where $\mathbb{Q}[\sin r][\cos r] : \mathbb{Q} \leq 2 \cdot [\mathbb{Q}[\sin r] : \mathbb{Q}] < \infty$, because $\sin r$ is an algebraic number and so, $\mathbb{Q}[\sin r] : \mathbb{Q} = \deg_{\mathbb{Q}} f_{\sin r,\mathbb{Q}}$, where $f_{\sin r,\mathbb{Q}} \in \mathbb{Q}[x]$ is the minimal polynomial of $\sin r$ over $\mathbb{Q}$. Thus, $e^{ri} = \cos r + i \sin r$ is an algebraic number, a contradiction (see Theorem 7.2 and the fact that $\overline{\mathbb{Q}}$ is a field). The same reasoning also works for $\cos r$. If $\tan r$ was an algebraic number, then

$$\frac{1}{\cos^2 r} = 1 + \tan^2 r$$

would be an algebraic number too. So,

$$\cos r = \pm\frac{1}{\sqrt{1+\tan^2 r}}$$

is in an extension of degree at most 2 of $\mathbb{Q}[\tan r]$, and $\mathbb{Q}[\tan r] : \mathbb{Q} < \infty$, because $\tan r$ was supposed to be an algebraic number over $\mathbb{Q}$. Hence, $\cos r$ would be an algebraic number, a contradiction.

Now, let us assume that $\ln r = \alpha$ is a nonzero algebraic number. Then $e^{\alpha} = r$ is a transcendental number, a contradiction (Theorem 7.2). $\square$

8. Wantzel's idea and the answer to the three famous Greek compass and straight-edge problems

Given a segment of a straight-line $[OA]$, $O \neq A$, the Old Greek Mathematical School was interested to construct, using only a compass and a straight-edge (in what follows we abbreviate this by a <<CS-construction>>), a new segment of a given length or a point with some geometrical properties. Here are three of the most known such problems which remained unsolved up to XIX-th century, when a french mathematician, P. L. Wantzel, made an enlightened connection between such geometrical problems and algebraic numbers [11]. In this way he succeeded to transform these geometrical problems into problems of algebraic numbers. In the following, using our results from Section 6 and Section 7, we give reasonable and rigorous answers to the next three geometrical Old Greek problems.

Problem 1. (Squaring a circle) Given a segment $[OA]$ of a straightline $d = OA$, with the length, say $r = 1$ unit, using only a CS-construction, find a square which has its area equal to the area of a circle centered at O and of radius $r = 1$.

Problem 2. (Duplication of a cube) Given a segment of length l, using only a CS-construction, find another segment of length L such that the volume of a cube of edge L is twice the volume of a cube of edge l.

Problem 3. (Trisection of an arbitrary angle) Given an arbitrary angle $\widehat{AOB}$, using only a CS-construction, find a new angle $\widehat{COD}$ such that the measure of $\widehat{AOB}$ is three times the measure of $\widehat{COD}$.

Now, we present in the following the great idea of Wantzel (1837) [11]. Let us start with a segment $[OA]$, $O \neq A$, and consider the unity measure the length of $[OA]$, i.e. $l[OA] = 1$. Let $\Delta = OA$ be the straight line generated by the distinct points O and A. We call it the *real number line,* because any other segment of length m (of a straight-line) can be "measured" on Δ; this means that always one can find a point M "on the right of O" such that the new segment $[OM]$ has its length equal to m. So, the "distance" from O to M is the length of our segment. We write $M(m)$ and call m the *coordinate* of the point M. It is clear that $O(0)$ and $A(1)$. Using only a compass one can easily construct all the points $N(n)$, where n is an integer ($n \in \mathbb{Z}$). Take now a half-line $O\delta \neq OA$, and k a nonzero arbitrary natural number. Using a compass one can construct on $O\delta$ the points $M_i(i)$, $i = 1, 2, ..., k$, such that the segment $[OM_i]$ has length equal to i. Now, using a CS-construction, we can draw parallel straight-lines λ_i to the straight-line AM_k (the straight-line which connect A and M_k) such that $M_i \in \lambda_i$ for any $i = 1, 2, ..., k-1$. Let $P(x_i) = \lambda_i \cap \Delta$ be the intersection point between the straight-line λ_i and the straight-line $OA = \Delta$. Thales theorem says that $x_i = i/k$, $i = 1, 2, ..., k$. Thus we succeeded "to construct $\mathbb{Q}_+$" on the real number line Δ. It is easy to make such a construction "on the left" of O, on Δ. Thus we just made a CS-construction for the rational number field $\mathbb{Q}$ on Δ.

Now we use a CS-construction to draw a perpendicular straight-line Π at O on Δ. We denote $\overrightarrow{i}$ the vector $\overrightarrow{OA}$ and $\overrightarrow{j} = \overrightarrow{OB}$, where B is "above" Δ such that the length of $[OB]$ is 1. Using the unity measure $[OB]$, we can represent like above all the rational numbers on the "imaginary number line" $\Pi = OB$. We call this couple (OA, OB), or $(O, \overrightarrow{i}, \overrightarrow{j})$, the complex plane.

Remark 8.1. *If $M_1(x) \in \Delta$, $x \in \mathbb{Q}$, and if $M_2(y) \in \Pi$, $y \in \mathbb{Q}$, then we can easily make a CS-construction to find the fourth vertex P of the rectangle M_2OM_1P. We write $P(x, y)$ and say that x and y are the (rational) coordinates of the point P. We associate to the point $P(x, y)$ the complex number $z = x + iy$. Thus, using only CS-constructions, we can represent in the complex plane $(O, \overrightarrow{i}, \overrightarrow{j})$ all the algebraic numbers of the algebraic field $\mathbb{Q}[i]$, where $i = \sqrt{-1}$.*

Definition 8.2. *A complex number $a + ib$, $a, b \in \mathbb{R}$ is said to be a CS-number if the point $M(a, b)$ in the complex plane $(O, \overrightarrow{i}, \overrightarrow{j})$ can be obtained by a CS-construction.*

It is easy to prove the following result.

Lemma 8.3. *A complex number $z = a + ib$, $a, b \in \mathbb{R}$ is a CS-number if and only if a, b are both CS-numbers.*

Lemma 8.4. *Let a, b be two nonzero real CS-numbers. Then $a \pm b$ and ab are also CS-real numbers.*

Proof. The first statement is obvious. Let us prove the second. Using the symmetry as a CS-construction we can assume that $a, b > 0$. In the complex plane $(O, \overrightarrow{i}, \overrightarrow{j})$ we consider the following CS-constructed points: $A(1, 0)$, $X(a, 0)$ and $Y(0, b)$. Let us use a CS-construction to draw a parallel straight-line μ to AY such that $X \in \mu$. Let $W(0, y)$ be the intersection between μ and $\Pi = OY = OB$. From Thales theorem we see that $1/a = b/y$, so $ab = y$, i.e. ab is a CS-number. □

Corollary 8.5. *Let $z_1 = a_1 + ib_1$ and $z_2 = a_2 + ib_2$ be two CS-numbers. Then $z_1 \pm z_2$ and z_1z_2 are also CS-numbers.*

Proof. Since z_1, z_2 are CS-numbers, a_1, b_1, a_2, b_2 are also CS-numbers (Lemma 8.3). Since $a_1 \pm a_2$, $b_1 \pm b_2$, $a_1a_2 - b_1b_2$ and $a_1b_2 + a_2b_1$ are CS-numbers (Lemma 8.4), we see that $z_1 \pm z_2$ and z_1z_2 are also CS-numbers (lemma 8.3). □

Corollary 8.6. (*see also Remark 8.1*) *$\mathbb{Q}$ and $\mathbb{Q}[i]$, $i = \sqrt{-1}$ contain only CS-numbers.*

Proof. We just proved this statement in Remark 8.1. □

Lemma 8.7. *Let $z = a + ib$, $a, b \in \mathbb{R}$ be a nonzero CS-number. Then*

$$\frac{1}{z} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}$$

is also a CS-number. Hence, the subset of all CS-numbers in $\mathbb{C}$ is a subfield of $\mathbb{C}$ (*see also Corollary 8.5*) *called the CS-field.*

Proof. Using symmetries relative to Δ and Π, one can assume that $a > 0, b \geq 0$. Lemma 8.4 says that a, b, and $a^2 + b^2$ are CS-numbers. From Lemma 8.3 and Corollary 8.5, it remains to prove that if $c \in \mathbb{R}$, $c \neq 0$ is a CS-number, then $1/c$ is also a CS-number. Indeed, let us consider the above complex plane $(O, \overrightarrow{i}, \overrightarrow{j})$ and the points: $C(c, 0)$ on Δ and $B(0, 1)$ on Π. Now, we draw a perpendicular straight-line δ on BC such that $B \in \delta$. Let D be the intersection point of δ and Δ. We apply the height theorem in the rightangle triangle CBD ($\widehat{CBD} = 90°$) and find $D(-1/c)$. Thus $1/c$ is a CS-number. □

Lemma 8.8. *Let d be a positive real CS-number. Then $\sqrt{d}$ is also a CS-number.*

Proof. In the complex plane $(O, \overrightarrow{i}, \overrightarrow{j})$ we take the points $A(1,0)$ and $D(d+1,0)$. Let M be the midpoint of $[OD]$. It is clear that M is a CS-constructible point. Now, we draw a circle $\mathcal{C}$ centered at M and of radius $r = (d+1)/2$. Let E be one of the two intersection points between this last circle $\mathcal{C}$ and the perpendicular straight line η on $\Delta = OA = OD$, with $A \in \eta$. Using again the height theorem in the rightangle triangle DEO ($\widehat{DEO} = 90°$), we see that the length of the segment $[AE]$ is exactly $\sqrt{d}$. Hence, $\sqrt{d}$ is a CS-number. □

Proposition 8.9. *Let $z = a + ib$, $a, b \in \mathbb{R}$ be a nonzero CS-number. Then $\pm\sqrt{z}$ is also a CS-number. In particular, if $K \subset \mathbb{C}$ is a subfield of $\mathbb{C}$ such that for any $z \in K$, z is a CS-number, and if w is a root of a quadratic polynomial $P \in K[x]$, then $K[w]$ contains only CS-numbers.*

Proof. Since always one can make a CS-construction to find the symmetric of a given point relative to the straight-lines $\Delta = OA$ and $\Pi = OB$, we take only the case $+\sqrt{z} \overset{def}{=} \sqrt{z}$, where $a \geq 0$ and $b > 0$. Let $|z| = \sqrt{a^2 + b^2}$ the absolute value of z. Since z is a CS-number, then $a, b, a^2 + b^2$ are also CS-numbers. So, Lemma 8.8 says that $|z|$ is a CS-number. Let θ be the unique solution of the system of equations in the variable θ ($\theta \in (0, \pi/2]$):

$$\begin{cases} \cos\theta = \frac{a}{|z|} \geq 0 \\ \sin\theta = \frac{b}{|z|} > 0 \end{cases}.$$

If $\theta = 90°$, we can easily make a CS-construction for it. So, we assume that $\theta < 90°$. Since a, b, $|z|$ are CS-numbers, $\sqrt{|z|}$ is also a CS-number. Consider the point $M(a, b)$, so $\theta = \widehat{AOM}$. Let us construct the bisectrix $O\delta$ of the angle $\widehat{AOM}$ and take the point $M_1 \in O\delta$ such that $OM_1 = \sqrt{|z|}$. It follows that M_1 corresponds to the complex number $\sqrt{z} = \sqrt{|z|}\left(\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right)$. □

Corollary 8.10. *The CS-field is closed to square roots, i.e. if $z \neq 0$ is a CS-number, then $\pm\sqrt{z}$ are also CS-numbers.*

Proposition 8.11. *Let α be a CS-number in $\mathbb{C}$, $\alpha \notin \mathbb{Q}$. Then α is an algebraic number, i.e. $\alpha \in \overline{\mathbb{Q}}$, $\deg_{\mathbb{Q}} \alpha = 2^n$ for a natural number n, and there exists a sequence $\alpha_0 = 1, \alpha_1, \alpha_2, ..., \alpha_{j_0-1}\ \alpha_{j_0} = \alpha$ of CS-numbers such that*

$$\mathbb{Q} = \mathbb{Q}[\alpha_0] \subset \mathbb{Q}[\alpha_1] \subset \mathbb{Q}[\alpha_1, \alpha_2] \subset ... \subset \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_{j_0-1}] \subset \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_{j_0-1}][\alpha],$$

where any simple extension, $\mathbb{Q}[\alpha_0, \alpha_1, ..., \alpha_{j-1}] \subset \mathbb{Q}[\alpha_0, \alpha_1, ..., \alpha_j]$, $j = 1, 2, ..., j_0$ in this tower, has degree 2.

Proof. Let $\alpha = a + ib$, $a, b \in \mathbb{R}$ be a nonzero CS-number. Let $M_k(a, b)$ be the point in the complex plane, which was successively obtained by k CS-constructions starting from O and $A(1, 0)$. So, we see that one can firstly construct a CS-number in $\mathbb{Q}$ or in $\mathbb{Q}[i]$ (Corollary 8.6). Let us denote this step by "step 1". The following step of CS-construction supplies a CS-number either in $\mathbb{Q}$ or in $\mathbb{Q}[i]$, or in a quadratic extension of $\mathbb{Q}$ or of $\mathbb{Q}[i]$. This is because at each step we either intersect two straight-lines with coefficients in $\mathbb{Q}$ or in $\mathbb{Q}[i]$, or we intersect a circle with a line (or two circles) with coefficients in $\mathbb{Q}$ or in $\mathbb{Q}[i]$. So, after the "step 2" we get a CS-number, say α_2, such that $\deg_{\mathbb{Q}} \alpha_2 = 1, 2$, or 4, i.e. a power of 2. Using mathematical induction on k, the number of steps, we easily find that $\alpha \in \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_s]$, where $0 \leq s \leq k$ ($s = 0$ means that we remained in $\mathbb{Q}$) and

$$\mathbb{Q} \subset \mathbb{Q}[\alpha_1] \subset \mathbb{Q}[\alpha_1, \alpha_2] \subset ... \subset \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_s],$$

such that $\mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_j] : \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_{j-1}] = 2$, $j = 1, 2, ..., s$. Thus $\deg_{\mathbb{Q}} \alpha = 2^s$ with $s \in \mathbb{N}$ and, if $\alpha \in \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_{j_0}] \setminus \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_{j_0-1}]$, then

$$\mathbb{Q} \subset \mathbb{Q}[\alpha_1] \subset \mathbb{Q}[\alpha_1, \alpha_2] \subset ... \subset \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_{j_0-1}] \subset \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_{j_0-1}][\alpha].$$

□

Theorem 8.12. *Any algebraic number α is a CS-number if and only if $\mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_s] : \mathbb{Q} = 2^n$, where $\alpha = \alpha_1, \alpha_2, ..., \alpha_s$ are the conjugates of α over $\mathbb{Q}$ and $n \in \mathbb{N}$.*

Proof. If $n = 0$, we have nothing to prove. Let n be a positive integer, assume that $\mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_s] : \mathbb{Q} = 2^n$, and let us consider the extensions of fields

$$\mathbb{Q} = L \subset K = \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_s].$$

Now we can apply Corollary 6.17 to the extension $\mathbb{Q} = L \subset K$ and find a tower of fields

$$\mathbb{Q} \subset \mathbb{Q}[\beta_1] \subset \mathbb{Q}[\beta_1, \beta_2] \subset ... \subset \mathbb{Q}[\beta_1, \beta_2, ..., \beta_s] = K,$$

such that

$$\mathbb{Q}[\beta_1, \beta_2, ..., \beta_i] : \mathbb{Q}[\beta_1, \beta_2, ..., \beta_{i-1}] = 2$$

for all $i = 1, 2, ..., s$. Here $\beta_0 = 1$. Since there exists $i_0 \in \{1, 2, ..., s\}$ such that

$$\alpha \in \mathbb{Q}[\beta_1, \beta_2, ..., \beta_{i_0}] \setminus \mathbb{Q}[\beta_1, \beta_2, ..., \beta_{i_0-1}],$$

we see that

$$\mathbb{Q}[\beta_1, \beta_2, ..., \beta_{i_0}] = \mathbb{Q}[\beta_1, \beta_2, ..., \beta_{i_0-1}][\alpha].$$

Since all $\beta_1, \beta_2, ..., \beta_{i_0-1}$ are CS-numbers (Proposition 8.9), we see that α itself is a CS-number, being an element in the quadratic extension

$$\mathbb{Q}[\beta_1, \beta_2, ..., \beta_{i_0-1}] \subset \mathbb{Q}[\beta_1, \beta_2, ..., \beta_{i_0-1}][\alpha].$$

Conversely, let us assume that α is a CS-number, $\alpha \in \mathbb{Q}$. So, there exists a tower of fields

$$\mathbb{Q} = L_0 \subset L_1 \subset ... \subset L_h = \mathbb{Q}[\alpha], \ h \in \{0, 1, ...\}, \tag{8.1}$$

such that $L_j : L_{j-1} = 2$ for any $j = 1, 2, ..., h$ (Proposition 8.11). Take $\gamma_j \in L_j \setminus L_{j-1}$ and find that $L_j = L_{j-1}[\gamma_j]$ for any $j = 1, 2, ..., h$. Let $\sigma_1, \sigma_2, ..., \sigma_s$, $s = 2^h = \deg_{\mathbb{Q}} \alpha$, be all the $\mathbb{Q}$-embeddings of $\mathbb{Q}[\alpha]$ into $\mathbb{C}$. For any $j = 1, 2, ..., s$, we take the range of the tower (8.1) through σ_j :

$$\mathbb{Q} = L_0 \subset \sigma_j(L_1) \subset ... \subset \sigma_j(L_h) = \mathbb{Q}[\sigma_j(\alpha)],$$

and see that all $\sigma_j(\alpha) = \alpha_j$, $j = 1, 2, ..., s$, the conjugates of α, are CS-numbers (Proposition 8.9). From Lemma 8.7 we conclude that any element of $\mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_s]$ is a CS-number. We use now the basic Primitive Element Theorem ([7], theorem 4.6) and find an element $\beta \in \mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_s]$ such that $\mathbb{Q}[\alpha_1, \alpha_2, ..., \alpha_s] = \mathbb{Q}[\beta]$. Since β is a CS-number, $\deg_{\mathbb{Q}} \beta = 2^n = \mathbb{Q}[\beta] : \mathbb{Q}$ (Proposition 8.11). □

Remark 8.13. *The condition $\deg_{\mathbb{Q}} \alpha = 2^n$, $n > 1$ is not sufficient for the CS-construction of α. Here is a counterexample and an application of Theorem 8.12. Let α be a root of the polynomial $P(x) = x^4 - x + 1 \in \mathbb{Q}[x]$. First of all let us prove that P is irreducible over $\mathbb{Q}$. Since the only possible roots of P in $\mathbb{Q}$ are ± 1, we see that P has no roots in $\mathbb{Q}$. So, the unique possibility for P to decompose itself over $\mathbb{Q}$ is:*

$$P(x) = (x^2 + ax + b)(x^2 + cx + d),$$

where $a, b, c, d \in \mathbb{Q}$ and both factors are irreducible over $\mathbb{Q}$. Identifying the coefficients we get the system:

$$\begin{cases} a + c = 0 \\ b + d + ac = 0 \\ ad + bc = -1 \\ bd = 1 \end{cases}.$$

It is easy to see that this system reduces to the equation

$$a^6 - 4a^2 - 1 = 0,$$

which has no solution in $\mathbb{Q}$. Hence P is irreducible over $\mathbb{Q}$. We prove now that the extension $\mathbb{Q} \subset \mathbb{Q}[\alpha]$ of degree 4 has no subextension of degree 2. Assume that there exists an element β of degree 2 in $\mathbb{Q}[\alpha]$. Since $\alpha^4 = \alpha - 1$, $\deg_{\mathbb{Q}} \alpha = 4$, and because we can assume that $\beta^2 = q \in \mathbb{Q}$, this β can be written as:

$$\beta = a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3, \ a_0, a_1, a_2, a_3 \in \mathbb{Q}.$$

and

$$\beta^2 = a_0^2 - a_2^2 + a_2^2\alpha + (a_1^2 - a_3^2)\alpha^2 + a_3^2\alpha^3 +$$
$$+2\left[-a_1a_3 + (a_0a_1 + a_1a_3 - a_2a_3)\alpha + (a_0a_2 + a_2a_3)\alpha^2 + (a_0a_3 + a_1a_2)\alpha^3\right].$$

Since $\{1, \alpha, \alpha^2, \alpha^3\}$ *are linear independent over* $\mathbb{Q}$ ($\deg f_{\alpha,\mathbb{Q}} = 4$, *not less*), a_0, a_1, a_2, a_3 *must be rational solutions of the system:*

$$\begin{cases} a_2^2 + 2a_0a_1 + 2a_1a_3 - 2a_2a_3 = 0 \\ a_1^2 - a_3^2 + 2a_0a_2 + 2a_2a_3 = 0 \\ a_3^2 + 2a_0a_3 + 2a_1a_2 = 0. \end{cases}.$$

If $a_3 = 0$, *we obtain* $a_1 = a_2 = a_3 = 0$, *so* $\beta = a_0 \in \mathbb{Q}$. *Let us assume that* $a_3 \neq 0$ *and divide the equations of the system by* a_3^2. *Now, denoting* $a_0/a_3 = b_0$, $a_1/a_3 = b_1$ *and* $a_2/a_3 = b_2$, *we finally obtain the following system:*

$$\begin{cases} b_2^2 + 2b_0b_1 + 2b_1 - 2b_2 = 0 \\ b_1^2 - 1 + 2b_0b_2 + 2b_2 = 0 \\ 1 + 2b_0 + 2b_1b_2 = 0 \end{cases}, \ b_0, b_1, b_2 \in \mathbb{Q}.$$

By the elimination of b_0, b_1 *we obtain the following equation in the variable* $b_2 \stackrel{def}{=} y$:

$$8y^7 - 16y^6 - 7y^4 + 24y^3 - 16y^2 - y + 1 = 0.$$

By checking $y = \pm 1$, $\pm 1/2$, $\pm 1/4$, $\pm 1/8$ *we see that these numbers cannot be solutions of the above equation. Thus,* $a_3 \neq 0$ *gives rise to a contradiction. Hence, in* $\mathbb{Q}[\alpha]$ *we have no subextension of degree* 2 *over* $\mathbb{Q}$, *i.e.* α *is not a CS-number* (*Proposition 8.11*). *From Theorem 8.12 we can immediately conclude that* $\mathbb{Q}[\alpha = \alpha_1, \alpha_2, \alpha_3, \alpha_4] : \mathbb{Q} = 12$, *or* 24. *where and* $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ *are all the conjugates of* α. *Indeed,* $P(x) = (x - \alpha)Q(x)$, $Q(x) \in \mathbb{Q}[\alpha][x]$, *in* $\mathbb{Q}[\alpha]$. *If* Q *had a factor of degree* 2 *in* $\mathbb{Q}[\alpha]$, *then* $\mathbb{Q}[\alpha_1, \alpha_2, \alpha_3, \alpha_4] : \mathbb{Q} = 4$, *or* 8, *i.e. a power of* 2. *From Theorem 8.12 we would obtain that* α *is a CS-number, a contradiction. Thus* $Q(x)$ *is irreducible over* $\mathbb{Q}[\alpha]$ *and its degree over* $\mathbb{Q}[\alpha]$ *is* 3. *Thus* $\mathbb{Q}[\alpha_1, \alpha_2, \alpha_3, \alpha_4] : \mathbb{Q} = 12$, *or* 24.

Remark 8.14. (*Answer to* ***Problem 1***) *Everything reduces to a CS-construction for* $\sqrt{\pi}$. *If one could do this, from Proposition 8.11 we find that* $\sqrt{\pi}$ *is an algebraic number, or that* π *itself is an algebraic number, a contradiction* (*Theorem 7.2*). *Hence, problem* ***Problem 1*** *has no solution at all.*

Remark 8.15. (*Answer to* ***Problem 2***) *The question reduces to a CS-construction for* $\sqrt[3]{2}$. *But* $\sqrt[3]{2}$ *is not a CS-number because* $\deg_{\mathbb{Q}} \sqrt[3]{2} = 3$, *which is not a power of* 2 (*see Proposition 8.11*). *Hence, problem* ***Problem 2*** *has no solution at all.*

Remark 8.16. (*Answer to* ***Problem 3***) *Let us assume that after a CS-construction we succeed to construct an angle of* $20^\circ = 60^\circ/3$. *It is clear that an angle of of* 60° *has a CS-construction. For instance, take a point* $Y \in \Pi$ (*in the complex plane* $(O, \overrightarrow{i}, \overrightarrow{j})$), *where* $\Pi = OB$, $B(0,1)$, *such that the length of* $[AY]$ *is* 2 *units. Then the angle* $\widehat{OAY}$ *has the measure equal to* 60°. *Since* $Y(0, \sqrt{3})$, *we see that* Y *is a CS-number. If we could find a CS-construction for an angle of* 20°, *then* $\alpha = \cos 20^\circ$ *is a CS-number. From the known formula:*

$$\cos 3x = 4\cos^3 x - 3\cos x,$$

we find (*for* $x = 20^\circ$) *that*

$$8\alpha^3 - 6\alpha - 1 = 0,$$

i.e. that α *is a root of the equation*

$$8x^3 - 6x - 1 = 0.$$

It is easy to see that this polynomial is irreducible over $\mathbb{Q}$ (*it has no rational root*). *So,* $\deg_{\mathbb{Q}} \alpha = 3$, *which is not a power of* 2. *Hence* α *is not a CS-number (Proposition 8.11), i.e.* ***Problem 3*** *has no solution.*

Acknowledgement. We are grateful to Prof. dr. Victor Alexandru (University of Bucharest) for a useful discussion on the statement of Theorem 8.12. We are also grateful to the referee for giving us a lot of suggestions in order to improve the first version of this paper.

References

[1] M. Aigner and G.H. Ziegler, *Proofs from THE BOOK*, Sixth Edition, Springer, 2018.
[2] A. Baker, *Transcendental Number Theory*, Cambridge University Press, New York, 1975.
[3] C. Fefferman, *An easyproof of the fundamental theorem of algebra*, The Amer. Math. Monthly **74(7)** (1967), 854-855.
[4] C. Hermite, *Sur la fonction exponentielle*, C. R. Acad. Sci. (Paris) **77** (1873), 18-24.
[5] D. Hilbert, *Über die Transcendenz der Zahlen e* und π, Math. Ann. **43** (1893), 216-219.
[6] A. Hurwitz, *Beweis der Transcendenz der Zahl e*, Math. Ann. **43** (1893), 220-222.
[7] S. Lang, *Algebra*, Springer, 2002.
[8] F. Lindemann, *Über die Zahl* π, Math. Ann. **20** (1882), 213-225.
[9] I. Niven, *A simple proof that* π *is irrational*, Bull. Amer. Math. Soc. **53** (1947), 509-509.
[10] M. Waldschmidt, *La méthode de Charles Hermite en théorie des nombres transcendents*, http://journals.openedition.org/bibnum/893?lang=en.
[11] P.L. Wantzel, *Recherches sur les moyens de reconnaître si un problème de géométrie peut se résoudre avec la règle et le compas*, J. Math. Pures Appl. **2** (1837), 366-372.

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania,
Email address: `angel.popescu@gmail.com`

GEOMETRIC INEQUALITIES FOR BI-SLANT SUBMANIFOLDS IN KENMOTSU SPACE FORMS

MOHAMMED MOHAMMED

Abstract. In the present paper we obtain Chen first inequality and Chen-Ricci inequality, respectively, for bi-slant submanifolds in Kenmotsu space form.

Mathematics Subject Classification (2010): 53C25; 53C40; 53C50.
Key words: Chen inequality, squared mean curvature, Ricci curvature, bi-slant submanifold, Kenmotsu space form.

Article history:
Received: October 21, 2022
Received in revised form: December 2, 2022
Accepted: December 5, 2022

1. Introduction

A well-known problem in submanifold theory is the immersibility of a Riemannian manifold in a Euclidean space. The embedding theorem of J.F. Nash [19] states that every Riemannian manifold can be isometrically embedded in Euclidean spaces with sufficiently high codimension. There were several reasons for which Nash's theorem was difficult to apply. One reason is that it generally requires a large codimension for a Riemannian manifold to support isometric embeddings in Euclidean spaces. Another reason is that at that time there were not known general optimal relationships between the classical intrinsic invariants and the principal extrinsic invariants for arbitrary submanifolds of Euclidean spaces, excepting the three fundamental equations of submanifolds. This leads to another fundamental problem in submanifold theory: Find simple relationship between the main extrinsic invariant (squared mean curvature) and intrinsic invariants of a submanifold. In order to provide some answers to this fundamental problem, B.Y. Chen in [7], [8] introduced new types of Riemannian invariants, known as Chen δ-invariants.

The Chen first invariant δ_M of a Riemannian manifold M is defined by

$$\delta_M(p) = \tau(p) - (\inf K)(p).$$

where τ is the scalar curvature of M and $K(\pi)$ denotes the sectional curvature of a plane section π in T_pM, $p \in M$.

For n-dimensional submanifolds M in a real space form $R(c)$ of constant sectional curvature c, the following basic inequality involving the intrinsic invariant δ_M and the squared mean curvature was established in [7]

$$\delta_M \leq \frac{n^2(n-2)}{2(n-1)}\|H\|^2 + \frac{1}{2}(n+1)(n-2)c, \tag{1.1}$$

where H is the mean curvature vector.

A slant submanifold [6] is a submanifold N of an almost Hermitian manifold (M, J) with constant Wirtinger angle (or Kaehler angle). The Wirtinger angle $\theta(X)$ of a tangent vector X to N at a point $p \in N$ is the angle between JX and the tangent space of N at p. Special cases are complex submanifolds

($\theta = 0$) and totally real submanifolds ($\theta = \frac{\pi}{2}$). Furthermore A. Lotta [17] introduced the class of slant submanifolds of almost contact metric manifolds.

A semi-slant submanifold of a Kahlerian manifold is a submanifold whose tangent bundle is the direct sum of a complex distribution and a slant distribution with the slant angle $\theta \neq 0$ (see [22]). Moreover, Cabrerizo et al. [4] introduced the class of a semi-slant submanifold of a Sasakian manifold. The authors defined and studied bi-slant and semi-slant submanifolds of an almost contact metric manifold, in particular a Sasakian one. They proved a characterization theorem for semi-slant submanifolds and obtain integrability conditions for the distributions which are involved in the definition of such submanifolds. Cioroboiu [12] established Chen inequalities for semi-slant submanifolds in Sasakian space forms by using subspaces orthogonal to the Reeb vector field.

There were many authors who studied Chen's inequalities for different submanifolds in different types of ambient spaces (see [1], [2], [5], [10], [13], [14], [16], [20], [23]).

2. Preliminaries

In this section, we recall some definitions and notations used throughout this paper.

Let $\overline{M}$ be a $(2m+1)$-dimensional almost contact metric manifold endowed with a Riemannian metric g, a tensor field ϕ of type $(1,1)$, a structure vector field ξ and a 1-form η which satisfy

$$\begin{aligned} &\phi^2 X = -X + \eta(X)\xi, \\ &\phi\xi = 0, \quad \eta(\xi) = 1, \quad \eta(\phi X) = 0, \quad \eta(X) = g(X,\xi), \\ &g(\phi X, \phi Y) = g(X,Y) - \eta(X)\eta(Y), \quad g(\phi X, Y) = -g(X, \phi Y). \end{aligned}$$

We denote by $\overline{\nabla}$ the Levi-Civita connection of g. If, in addition,

$$(\overline{\nabla}_X \phi)(Y) = g(\phi X, Y)\xi - \eta(Y)\phi X,$$

for any vector fields X, Y on $\overline{M}$, then $\overline{M}$ is said to be a Kenmotsu manifold. One also has

$$\overline{\nabla}_X \xi = X - \eta(X)\xi = -\phi^2 X.$$

A Kenmotsu manifold with constant ϕ-holomorphic sectional curvature c is called a Kenmotsu space form and is denoted by $\overline{M}(c)$. The curvature tensor $\overline{R}$ of a Kenmotsu space form is given by [15]

$$\begin{aligned} 4\overline{R}(X,Y)Z =&(c-3)[g(Y,Z)X - g(X,Z)X] + (c+1)[g(\phi Y, Z)\phi X \\ &- g(\phi X, Z)\phi Y - 2g(\phi X, Y)\phi Z + g(X,Z)\eta(Y)\xi \\ &- g(Y,Z)\eta(X)\xi + \eta(X)\eta(Z)Y - \eta(Y)\eta(Z)X], \end{aligned} \tag{2.1}$$

for any $X, Y \in T\overline{M}$.

Let M be an n-dimensional submanifold of a Kenmotsu space form $\overline{M}$ equipped with a Riemannian metric g. The Gauss and Weingarten formulae are given respectively by

$$\begin{aligned} \overline{\nabla}_X Y &= \nabla_X Y + h(X,Y), \\ \overline{\nabla}_X Y &= -A_N X + \nabla^{\perp}_X N, \end{aligned}$$

for all $X, Y \in TM$ and $N \in T^{\perp}M$, where ∇ is the Levi-Civita connection on M and $\nabla^{\perp}$ the normal connection, respectively. The second fundamental form h is related to the shape operator A by

$$g(h(X,Y), N) = g(A_N X, Y).$$

The equation of Gauss is given by

$$\overline{R}(X,Y,Z,W) = R(X,Y,Z,W) - g(h(X,Z), h(Y,W)) + g(h(X,W), h(Y,Z)), \tag{2.2}$$

for all X, Y, Z, W tangent to M .

Let $\{e_1,...,e_n\}$ be an orthonormal basis of the tangent space T_pM. The scalar curvature τ at p is defined by

$$\tau(p) = \sum_{1\leq i<j\leq n} K(e_i \wedge e_j), \tag{2.3}$$

where $K(e_i \wedge e_j)$ denotes the sectional curvature of the plane section spanned by e_i, e_j.

In particular, if we put $e_n = \xi$, then (2.3) implies

$$2\tau = \sum_{1\leq i\neq j\leq n-1} K(e_i \wedge e_j) + 2\sum_{i=1}^{n-1} K(e_i \wedge \xi). \tag{2.4}$$

Let L be a k-plane section of T_pM and X a unit vector in L; we choose an orthonormal basis $\{e_1,...,e_k\}$ of L such that $e_1 = X$. The Ricci curvature Ric_L of L at X is defined by

$$Ric_L(X) = K_{12} + K_{13} + \cdots + K_{1k}. \tag{2.5}$$

We simply call it the k-Ricci curvature.

The mean curvature vector $H(p)$ at $p \in M$ is given by

$$H(p) = \frac{1}{n}\sum_{i=1}^{n} h(e_i, e_i). \tag{2.6}$$

If we put $h^r_{ij} = g(h(e_i, e_j), e_r)$, $i, j = 1,...,n, r \in \{n+1,...,2m+1\}$, the squared norm of the second fundamental form h is

$$\|h\|^2 = \sum_{r=n+1}^{2m+1}\sum_{i,j=1}^{n} (h^r_{ij})^2.$$

For any $X \in TM$, we can write $\phi X = PX + FX$, where $PX$and FX are the tangential and normal components of ϕX, respectively. We denote

$$\|P\|^2 = \sum_{i,j=1}^{n} g^2(Pe_i, e_j).$$

Definition 2.1. A submanifold is called *totally geodesic* if the second fundamental form vanishes identically; it is *totally umbilical* if $h(X,Y) = g(X,Y)H$, for any tangent vectors X,Y on M.

Definition 2.2. [4] A differentiable distribution D on M is called a *slant distribution* if for each $p \in M$ and each non-zero vector $X \in D_p$ the angle $\theta_D(X)$ between ϕX and the vector subspace D_p is constant, i.e., independent on the choice of X; $\theta_D(X)$ is called the slant angle.

Definition 2.3. [4] A submanifold M tangent to the structure vector field ξ is said to be a *bi-slant submanifold* of $\overline{M}$ if there exist two orthogonal distributions D_1 and D_2 such that

(i) TM admits the orthogonal direct decomposition $TM = D_1 \oplus D_2 \oplus \{\xi\}$, where $\{\xi\}$ is the 1-dimensional distribution spanned by ξ,
(ii) D_1, D_2 are slant distributions with slant angles θ_1, θ_2.

Definition 2.4. [4] A submanifold M tangent to ξ is said to be a *semi-slant submanifold* of $\overline{M}$ if there exist two orthogonal distributions D_1 and D_2 on M such that

(i) TM admits the orthogonal direct decomposition $TM = D_1 \oplus D_2 \oplus \{\xi\}$,
(ii) The distribution D_1 is an invariant distribution, i.e., $\phi(D_1) = D_1$,
(iii) The distribution D_2 is a slant distribution with slant angle $\theta \neq 0$.

A bi-slant submanifold of an almost contact metric manifold $\overline{M}$ is called proper if the slant distributions D_1 and D_2 have the slant angles $\theta_1, \theta_2 \neq 0, \frac{\pi}{2}$.

Suppose M is a proper bi-slant submanifold with dimension $n = 2d_1 + 2d_2 + 1$ in $\overline{M}$. Let us consider an orthonormal basis of T_pM

$$e_1, e_2 = \sec\theta_1 Pe_1, \cdots, e_{2d_1-1}, e_{2d_1} = \sec\theta_1 Pe_{2d_1-1}, e_{2d_1+2} = \sec\theta_2 Pe_{2d_1+1},$$
$$\cdots, e_{2d_1+2d_2-1}, e_{2d_1+2d_2} = \sec\theta_2 e_{2d_1+2d_2-1},\ e_{2d_1+2d_2+1} = \xi.$$

Then

$$g^2(\phi e_i, e_{i+1}) = \begin{cases} \cos^2\theta_1, & for \quad i \in \{1, 3, ..., 2d_1 - 1\}, \\ \cos^2\theta_2, & for \quad i \in \{2d_1 + 1, ..., 2d_1 + 2d_2 - 1\}, \end{cases}$$

$$\sum_{i,j=1}^{n} g^2(\phi e_j, e_i) = 2(d_1\cos^2\theta_1 + d_2\cos^2\theta_2). \tag{2.7}$$

3. Chen First Inequality For Bi-Slant Submanifolds in Kenmotsu Space Forms

Pandey et al. [21] obtained B.Y. Chen inequalities for a bi-slant submanifold M of a Kenmotsu space form, when the structure vector field ξ is tangent to M. In this section, we prove Chen first inequality for proper bi-slant submanifolds in Kenmotsu space forms, by using orthogonal subspaces to the structure vector field ξ.

We need an algebric lemma from [7].

Lemma 3.1. [7] *Let $a_1, ...a_k, b$ be $k+1$ $(k \geq 2)$ real numbers such that*

$$\left(\sum_{i=1}^{k} a_i\right)^2 = (k-1)\left(\sum_{i=1}^{k} a_i^2 + b\right).$$

Then $2a_1a_2 \geq b$, with equality holding if and only if $a_1 + a_2 = a_3 = ... = a_k$.

Theorem 3.2. *Let $\psi : M \to \overline{M}(c)$ be an isometric immersion of an n-dimensional $(n \geq 3)$ proper bi-slant submanifold M in a $(2m+1)$-dimensional Kenmotsu space form $\overline{M}(c)$. Then*

(i) *For any plane section π invariant by P and tangent to D_1,*

$$\tau - K(\pi) \leq \frac{n^2(n-2)}{2(n-1)}\|H\|^2 + \frac{n(n-3)(c-3)}{8}$$
$$+ 3\frac{c+1}{4}[(d_1 - 1)\cos^2\theta_1 + d_2\cos^2\theta_2] - (n-1). \tag{3.1}$$

(ii) *For any plane section π invariant by P and tangent to D_2,*

$$\tau - K(\pi) \leq \frac{n^2(n-2)}{2(n-1)}\|H\|^2 + \frac{n(n-3)(c-3)}{8}$$
$$+ 3\frac{c+1}{4}[d_1\cos^2\theta_1 + (d_2 - 1)\cos^2\theta_2] - (n-1). \tag{3.2}$$

The equality case of the inequality (3.1) or (3.2) holds at a point $p \in M$ if and only if there exist an orthonormal basis $\{e_1, e_2, ..., e_n = \xi\}$ of T_pM and an orthonormal basis $\{e_{n+1}, ..., e_{2m}, e_{2m+1}\}$ of $T_p^\perp M$ such that the shape operators of M in $\overline{M}(c)$ at p have the following forms

$$A_{n+1} = \begin{pmatrix} a & 0 & 0 & \cdots & 0 \\ 0 & \mu - a & 0 & \cdots & 0 \\ 0 & 0 & \mu & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \mu \end{pmatrix},$$

$$A_r = \begin{pmatrix} h_{11}^r & h_{12}^r & 0 & \cdots & 0 \\ h_{12}^r & -h_{22}^r & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad r \in \{n+1, ..., 2m+1\}.$$

Proof. We assume that M is an $(n = 2d_1+2d_2+1)$-dimensional proper bi-slant submanifold of a Kenmotsu space form $\overline{M}(c)$ with constant ϕ-holomorphic sectional curvature c.

Let $p \in M$ and $\{e_1, e_2, ..., e_n = \xi\}$ an orthonormal basis of the tangent space T_pM and $\{e_{n+1}, e_{n+2}, ..., e_{2m+1}\}$ an orthonormal basis of $T_p^\perp M$.

From the equation (2.4), one has

$$2\tau = \sum_{i,j=1}^{n} R(e_i, e_j, e_j, e_i) = \sum_{i\neq j}^{n-1} R(e_i, e_j, e_j, e_i) + 2\sum_{i=1}^{n-1} R(e_i, \xi, \xi, e_i). \tag{3.3}$$

In the Gauss equation (2.2), we put $X = W = e_i$ and $Y = Z = e_j$, $\forall i, j \in 1, 2, ..., n$, and we take the summation over $1 \leq i, j \leq n$. We get

$$\sum_{i\neq j}^{n-1} R(e_i, e_j, e_j, e_i) = \sum_{i\neq j}^{n-1} \overline{R}(e_i, e_j, e_j, e_i) - \sum_{i\neq j}^{n-1} g(h(e_i, e_j), h(e_i, e_j)) + \sum_{i\neq j}^{n-1} g(h(e_i, e_i), h(e_j, e_j)). \tag{3.4}$$

By the formula (2.1) of the Riemannian curvature tensor of a Kenmotsu space form $\overline{M}(c)$, we have

$$\begin{aligned} \overline{R}(e_i, e_j, e_j, e_i) &= \frac{c-3}{4}\left[g(e_j, e_j)g(e_i, e_i) - g(e_i, e_j)g(e_j, e_i)\right] \\ &+ \frac{c+1}{4}\{\eta(e_i)\eta(e_j)g(e_j, e_i) - \eta(e_j)\eta(e_j)g(e_i, e_i) \\ &+ \eta(e_j)\eta(e_i)g(e_i, e_j) - \eta(e_i)\eta(e_i)g(e_j, e_j) \\ &- g(\phi e_i, e_j)g(\phi e_j, e_i) + g(\phi e_j, e_j)g(\phi e_i, e_i) \\ &+ 2g(e_i, \phi e_j)g(\phi e_j, e_i)\}. \end{aligned}$$

Then

$$\sum_{i\neq j}^{n-1} \overline{R}(e_i, e_j, e_j, e_i) = \frac{c-3}{4}(n-1)(n-2) + 3\frac{c+1}{4}\sum_{i\neq j}^{n-1} g^2(\phi e_i, e_j). \tag{3.5}$$

If we substitute the equation (3.5) in the equation (3.4), we get

$$\begin{aligned} \sum_{i\neq j}^{n-1} R(e_i, e_j, e_j, e_i) &= \frac{c-3}{4}(n-1)(n-2) + 3\frac{c+1}{4}\sum_{i\neq j}^{n-1} g^2(\phi e_i, e_j) \\ &- \sum_{i\neq j}^{n-1} g(h(e_i, e_j), h(e_i, e_j)) + \sum_{i\neq j}^{n-1} g(h(e_i, e_i), h(e_j, e_j)). \end{aligned} \tag{3.6}$$

The equation (3.3) becomes

$$\begin{aligned} 2\tau = \sum_{i,j=1}^{n} R(e_i, e_j, e_j, e_i) &= \frac{c-3}{4}(n-1)(n-2) + 3\frac{c+1}{4}\sum_{i\neq j}^{n-1} g^2(\phi e_i, e_j) \\ - \sum_{i\neq j}^{n-1} g(h(e_i, e_j), h(e_i, e_j)) &+ \sum_{i\neq j}^{n-1} g(h(e_i, e_i), h(e_j, e_j)) + 2\sum_{i=1}^{n-1} R(e_i, \xi, \xi, e_i). \end{aligned} \tag{3.7}$$

Calculate

$$\sum_{i=1}^{n-1} K(\xi \wedge e_j) = \sum_{j=1}^{n-1} R(e_i, \xi, \xi, e_i) = \sum_{i=1}^{n-1} \overline{R}(e_i, \xi, \xi, e_i)$$

$$+ \sum_{i=1}^{n-1} g(h(\xi, \xi), h(e_j, e_j)) - \sum_{i=1}^{n-1} g(h(\xi, e_j), h(\xi, e_j)). \tag{3.8}$$

From the properties of a Kenmotsu manifold one has $h(\xi, \xi) = h(\xi, X) = 0$. Then

$$\sum_{i=1}^{n-1} R(e_i, \xi, \xi, e_i) = -(n-1) \tag{3.9}$$

If we substitute the equation (3.9) in (3.7) we get

$$2\tau = \sum_{i,j=1}^{n} R(e_i, e_j, e_j, e_i) = \frac{c-3}{4}(n-1)(n-2)$$

$$+ 3\frac{c+1}{4}\|P\|^2 - \|h\|^2 + n^2\|H\|^2 - 2(n-1), \tag{3.10}$$

where

$$\|P\|^2 = \sum_{i,j=1}^{n} g^2(\phi e_i, e_j) = 2(d_1 \cos^2\theta_1 + d_2 \cos^2\theta_2).$$

Now denote by

$$\epsilon = 2\tau - \frac{n^2(n-2)}{n-1}\|H\|^2 - \frac{c-3}{4}(n-1)(n-2)$$

$$- 3\frac{c+1}{4}\|P\|^2 + 2(n-1). \tag{3.11}$$

The equation (3.10) is equivalent to

$$n^2\|H\|^2 = (n-1)(\epsilon + \|h\|^2). \tag{3.12}$$

Let $p \in M, \pi \subset T_pM$, dim $\pi = 2$, and π orthogonal to ξ and invariant by P.

Now, we consider the following two cases:

Case I. The plane section π is tangent to D_1. We may assume that π is spanned by the orthonormal basis $\{e_1, e_2\}$. We take e_{n+1} in the direction of mean curvature vector H. The relation (3.12) becomes

$$\left(\sum_{i=1}^{n} h_{ii}^{n+1}\right)^2 = (n-1)\left\{\sum_{i,j=1}^{n}\sum_{r=n+1}^{2m+1}\left(h_{ij}^r\right)^2 + \epsilon\right\}, \tag{3.13}$$

or equivalently,

$$(\sum_{i=1}^{n} h_{ii}^{n+1})^2 = (n-1)\left[\sum_{i=1}^{n}\left(h_{ii}^{n+1}\right)^2 + \sum_{i\neq j}\left(h_{ij}^{n+1}\right)^2 + \sum_{r=n+2}^{2m+1}\sum_{i,j}^{n}\left(h_{ij}^r\right)^2 + \epsilon\right]. \tag{3.14}$$

Using Lemma 3.1 we derive

$$2h_{11}^{n+1}h_{22}^{n+1} \geqslant \sum_{i\neq j}\left(h_{ij}^{n+1}\right)^2 + \sum_{r=n+2}^{2m+1}\sum_{i,j=1}^{n}\left(h_{ij}^r\right)^2 + \epsilon. \tag{3.15}$$

On the other hand, we have

$$K(\pi) = R(e_1, e_2, e_2, e_1) = g(h(e_1, e_1), h(e_2, e_2))$$
$$- g(h(e_1, e_2), h(e_1, e_2)) + \frac{c-3}{4} + 3\cos^2\theta_1 \frac{c+1}{4}.$$

Then

$$K(\pi) = \sum_{r=n+1}^{2m+1} \left[h_{11}^r h_{22}^r - (h_{12}^r)^2 \right] + \frac{c-3}{4} + 3\cos^2\theta_1 \left(\frac{c+1}{4} \right).$$

By using (3.15) we obtain

$$K(\pi) \geq \frac{c-3}{4} + 3\cos^2\theta_1 \left(\frac{c+1}{4} \right) + \frac{1}{2} \sum_{r=n+2}^{2m+1} \sum_{i,j=3}^{n-1} (h_{ij}^r)^2$$
$$+ \frac{1}{2} \sum_{r=n+2}^{2m+1} (h_{11}^r + h_{22}^r)^2 + \sum_{r=n+1}^{2m+1} \sum_{j>2} \left[(h_{1j}^r)^2 + (h_{2j}^r)^2 \right] + \frac{\epsilon}{2}.$$

From the last equation we get

$$K(\pi) \geqslant \frac{c-3}{4} + 3\cos^2\theta_1 \left(\frac{c+1}{4} \right) + \frac{\epsilon}{2}. \tag{3.16}$$

If we substitute (3.11) in above relation, we obtain

$$\tau - K(\pi) \leq \frac{n^2(n-2)}{2(n-1)} \|H\|^2 + \frac{n(n-3)(c-3)}{8}$$
$$+ 3\frac{c+1}{4} [2(d_1 - 1)\cos^2\theta_1 + 2d_2 \cos^2\theta_2] - (n-1), \tag{3.17}$$

which is the inequality (3.1)

Case II

If the plane section π is tangent to D_2 and invariant by P, similarly we get

$$\tau - K(\pi) \leq \frac{n^2(n-2)}{2(n-1)} \|H\|^2 + \frac{n(n-3)(c-3)}{8}$$
$$+ 3\frac{c+1}{4} [2d_1 \cos^2\theta_1 + 2(d_2 - 1)\cos^2\theta_2] - (n-1),$$

which is the inequality (3.2)

The case of equality of the inequality (3.1) at a point $p \in M$ holds if and only if equalities in inequalities (3.15), (3.16) and Lemma (3.1) hold, i.e.,

$$\begin{cases} h_{ij}^{n+1} = 0, \quad \forall i \neq j, \quad i, j > 2, \\ h_{ij}^r = 0, \quad \forall i \neq j \quad i, j > 2, r = n+1, ..., 2m+1, \\ h_{11}^r + h_{22}^r = 0, \qquad \forall r = n+2, ..., 2m+1, \\ h_{1j}^r = h_{2j}^r = 0, \qquad \forall j > 2,\ r = n+1, ..., 2m+1, \\ h_{11}^{n+1} + h_{22}^{n+1} = h_{33}^{n+1} = ... = h_{nn}^{n+1}. \end{cases} \tag{3.18}$$

Moreover we may choose e_1, e_2 such that $h_{12}^{n+1} = 0$ and we denote by $a = h_{11}^{n+1}$, $b = h_{22}^{n+1}$, $\mu = h_{33}^{n+1} = \cdots = h_{nn}^{n+1}$.

Then we obtain the desired forms for the shape operators A_r, $r \in \{n+1, \cdots, 2m+1\}$. Also, if $\mu = 0$, the submanifold is minimal. □

4. Ricci curvature and squared mean curvature

B.Y. Chen in [9] established a sharp relationship involving the Ricci curvature and the squared mean curvature on an n-dimensional submanifold M in a Riemannian space form $R(c)$ of constant sectional curvature c. For any unit tangent vector X at $p \in M$, its Ricci curvature satisfies

$$Ric(X) \leq (n-1)c + \frac{n^2}{4}\|H\|^2.$$

The above inequality is called Chen-Ricci inequality.

On the other hand, K. Arslan et al. [3] established a Chen-Ricci inequality for submanifolds in Kenmotsu space forms. Also I. Mihai [18] obtained a Chen-Ricci inequality for submanifolds in Sasakian space forms.

D.W. Yoon [24] established the following Chen-Ricci inequality for bi-slant submanifolds in a cosymplectic space form.

Theorem 4.1. [24] *Let M be an n-dimensional bi-slant submanifold satisfying $g(X, \phi Y) = 0$, for any $X \in D_1$ and any $Y \in D_2$, in a cosymplectic space form $\overline{M}(c)$. Then*

(1) *For each unit vector $X \in T_pM$ orthogonal to ξ and*
   (i) *X is tangent to D_1, we have*

$$Ric(X) \leq \frac{1}{4}\left\{(n-1)c + \frac{1}{2}(3\cos^2\theta_1 - 2)c + n^2 \parallel H \parallel^2\right\}. \tag{4.1}$$

   (ii) *X is tangent to D_2, we have*

$$Ric(X) \leq \frac{1}{4}\left\{(n-1)c + \frac{1}{2}(3\cos^2\theta_2 - 2)c + n^2 \parallel H \parallel^2\right\}. \tag{4.2}$$

(2) *If $H(p) = 0$, then a unit tangent vector X at p satisfies (4.1) or (4.2) if and only if $X \in N_p$, where $N_p = \{X \in T_pM | h(X,Y) = 0, \forall Y \in T_pM\}$.*
(3) *The equality case of (4.1) and (4.2) hold identically for all unit tangent vectors orthogonal to ξ at p if and only if p is a totally geodesic point.*

In this section we will study the relation between the Ricci curvature and the squared mean curvature for bi-slant submanifolds in a Kenmotsu space form.

Theorem 4.2. *Let M be an n-dimensional bi-slant submanifold in a $(2m+1)$-dimensional Kenmotsu space form $\overline{M}(c)$. Then:*

(1) *For each unit vector $X \in T_pM$ and*
   (i) *X tangent to D_1, we have*

$$Ric(X) \leq \frac{1}{4}n^2\|H\|^2 + (n-2)\frac{c-3}{4} + 3\frac{c+1}{8}\cos^2\theta_1 - 1. \tag{4.3}$$

   (ii) *X tangent to D_2, we have*

$$Ric(X) \leq \frac{1}{4}n^2\|H\|^2 + (n-2)\frac{c-3}{4} + 3\frac{c+1}{8}\cos^2\theta_2 - 1. \tag{4.4}$$

(2) *If $H(p) = 0$, then a unit tangent vector X at p satisfies (4.3) or (4.4) if and only if $X \in N_p$.*
(3) *The equality cases of (4.3) and (4.4) hold identically for all unit tangent vectors orthogonal to ξ at p if and only if p is a totally geodesic point.*

Proof.
(1) Let $X \in T_pM$ be a unit tangent vector X at p. We choose an orthonormal basis $\{e_1, \cdots, e_n = \xi, e_{n+1}, \cdots, e_{2m+1}\}$ in $T_p\overline{M}(c)$ such that $e_1, \cdots, e_n$ are tangent to M at p, with $e_1 = X$. We recall the equation (3.10)

$$2\tau = \sum_{i,j=1}^{n} R(e_i, e_j, e_j, e_i) = \frac{c-3}{4}(n-1)(n-2) + 3\frac{c+1}{4}\|P\|^2 - \|h\|^2 + n\|H\|^2 - 2(n-1). \tag{4.5}$$

Then we can write

$$n^2\|H\|^2 = 2\tau - \frac{c-3}{4}(n-1)(n-2) - 3\frac{c+1}{4}\|P\|^2 + \|h\|^2 + 2(n-1). \tag{4.6}$$

which leads to

$$n^2\|H\|^2 = 2\tau + \sum_{r=n+1}^{2m+1}\left[(h_{11}^r)^2 + (h_{22}^r + ... + h_{mm}^r)^2 + 2\sum_{i<j}(h_{ij}^r)^2\right] - 2\sum_{r=n+1}^{2m+1}\sum_{2\le i<j\le n} h_{ii}^r h_{jj}^r - \frac{c-3}{4}(n-1)(n-2) - 3\frac{c+1}{4}\|P\|^2 + 2(n-1).$$

It follows that

$$\begin{aligned} n^2\|H\|^2 &= 2\sum_{2\le i<j\le n} K_{ij} + 2Ric(X) + \frac{1}{2}\sum_{r=n+1}^{2m+1}(h_{11}^r + \cdots + h_{nn}^r)^2 \\ &+ \frac{1}{2}\sum_{r=n+1}^{2m+1}(h_{11}^r - h_{22}^r - \cdots - h_{nn}^r)^2 \\ &+ 2\sum_{r=n+1}^{2m+1}\sum_{1\le i<j\le n}(h_{ij}^r)^2 - 2\sum_{r=n+1}^{2m+1}\sum_{2\le i<j\le n} h_{ii}^r h_{jj}^r \\ &- \frac{c-3}{4}(n-1)(n-2) - 3\frac{c+1}{4}\|P\|^2 + 2(n-1). \end{aligned}$$

Then we get

$$\begin{aligned} n^2\|H\|^2 - \frac{1}{2}\sum_{r=m+2}^{2m+1}(h_{11}^r + \cdots + h_{nn}^r)^2 &= 2\sum_{2\le i<j} K_{ij} + 2Ric(X) \\ &+ \frac{1}{2}\sum_{r=n+1}^{2m+1}(h_{11}^r - h_{22}^r - \cdots - h_{nn}^r)^2 \\ &+ 2\sum_{r=n+1}^{2m+1}\sum_{i<j}(h_{ij}^r)^2 - 2\sum_{r=n+1}^{2m+1}\sum_{2\le i<j} h_{ii}^r h_{jj}^r \\ &- \frac{c-3}{4}(n-1)(n-2) - 3\frac{c+1}{4}\|P\|^2 + 2(n-1). \end{aligned} \tag{4.7}$$

We compute

$$\sum_{2\le i<j\le n} K_{ij} = \sum_{2\le i<j\le n-1} R(e_i, e_j, e_j, e_i) + 2\sum_{i=2}^{m} R(e_i, \xi, \xi, e_i). \tag{4.8}$$

From the Gauss equation

$$\sum_{2\le i<j}^{n-1} R(e_i,e_j,e_j,e_i) = \sum_{r=n+1}^{2m+1}\sum_{2\le i<j}\left[h^r_{ii}h^r_{jj} - (h_{ij})^2\right] + \sum_{2\le i<j}^{n-1}\overline{R}(e_i,e_j,e_j,e_i). \tag{4.9}$$

we get

$$\sum_{2\le i<j}^{n-1}\overline{R}(e_i,e_j,e_j,e_i) = \frac{c-3}{8}(n-2)(n-3) + 3\frac{c+1}{4}\sum_{2\le i<j\le n} g(\phi e_i,e_j). \tag{4.10}$$

By substituting the equation (4.10) in equation (4.9)

$$\begin{aligned}\sum_{2\le i<j\le n-1} R(e_i,e_j,e_j,e_i) &= \sum_{r=n+1}^{2m+1}\sum_{2\le i<j\le n-1}\left[h^r_{ii}h^r_{jj} - (h_{ij})^2\right]\\ &+ \frac{c-3}{8}(n-2)(n-3) + 3\frac{c+1}{4}\sum_{2\le i<j\le n} g(\phi e_i,e_j)\end{aligned} \tag{4.11}$$

and then by the equation (4.8) we have

$$\begin{aligned}\sum_{2\le i<j} K_{ij} &= \frac{c-3}{8}(n-2)(n-3) + 3\frac{c+1}{4}\sum_{2\le i<j\le n} g(\phi e_i,e_j)\\ &+ 2\sum_{r=n+1}^{2m+1}\sum_{2\le i<j}\left[h^r_{ii}h^r_{jj} - (h_{ij})^2\right] + 2\sum_{i=2}^{n-1} R(e_i,\xi,\xi,e_i).\end{aligned} \tag{4.12}$$

Similarly

$$\sum_{i=2}^{n-1} R(e_i,\xi,\xi,e_i) = -(n-2). \tag{4.13}$$

Then

$$\begin{aligned}\sum_{2\le i<j} K_{ij} &= \frac{c-3}{8}(n-2)(n-3) + 3\frac{c+1}{8}[\|P\|^2 - \|Pe_1\|^2]\\ &+ 2\sum_{r=n+1}^{2m+1}\sum_{2\le i<j}\left[h^r_{ii}h^r_{jj} - (h_{ij})^2\right] - (n-2).\end{aligned} \tag{4.14}$$

Therefore

$$Ric(X) \le \frac{1}{4}n^2\|H\|^2 + (n-2)\frac{c-3}{4} + 3\frac{c+1}{8}\|Pe_1\|^2 - 1. \tag{4.15}$$

(i) If X is tangent to D_1, we have $\|Pe_1\|^2 = \cos^2\theta_1$; then from (4.15) we obtain

$$Ric(X) \le \frac{1}{4}n^2\|H\|^2 + (n-2)\frac{c-3}{4} + 3\frac{c+1}{8}\cos^2\theta_1 - 1. \tag{4.16}$$

(ii) If X is tangent to D_2, we have $\|Pe_1\|^2 = \cos^2\theta_2$; then from (4.15) we obtain

$$Ric(X) \le \frac{1}{4}n^2\|H\|^2 + (n-2)\frac{c-3}{4} + 3\frac{c+1}{8}\cos^2\theta_2 - 1. \tag{4.17}$$

(2) Assume $H(p) = 0$. Equality holds in (4.3) or (4.4) if and only if

$$\begin{cases} h^r_{12} = h^r_{13} = \cdots = h^r_{1m} = 0,\\ h^r_{11} = h^r_{22} + \cdots + h^r_{mm}, \quad r \in \{m+2,\cdots,2m+1\}.\end{cases}$$

Then $h^r_{1j} = 0$, for all $j \in \{1,\cdots,m+1\}$, $r \in \{m+2,\cdots,2m+1\}$, that is $X \in N_p$.

(3) The equality cases of (4.3) and (4.4) hold for all unit tangent vectors at p if and only if

$$\begin{cases} h^r_{ij} = 0, \quad i \neq j, & r \in \{m+2, \cdots, 2m+1\}, \\ h^r_{11} + \cdots + h^r_{mm} - 2h^r_{ii} = 0, & i \in \{1, \cdots, m\}, \quad r \in \{m+2, \cdots, 2m+1\}. \end{cases}$$

It follows that p is a totally geodesic point.

□

Acknowledgements

The author thanks the Ministry of Higher Education and Scientific Research of Sudan and Al Neelain University.

References

[1] N. Aktan, M.Z. Sarikaya and E. Ozusaglam, *B.Y. Chen's inequality for semi-slant submanifolds in T-space forms*, Balkan J. Geom. Appl. **13(1)** (2008), 1-10.
[2] P. Alegre, A. Carriazo, Y.H. Kim and D.W. Yoon, *B.-Y. Chen's inequality for submanifolds of generalized space forms*, Indian J. Pure Appl. Math. **38(3)** (2007), 185-201.
[3] K. Arslan, R. Ezentas, I. Mihai, C. Murathan and C. Özgür, *Ricci curvature of submanifolds in Kenmotsu space forms*, Int. J. Math. Math. Sci. **29(12)** (2002), 719-726.
[4] J.L. Cabrerizo, A. Carriazo, L.M. Fernández, and M. Fernández, *Semi-slant submanifolds of a Sasakian manifold*, Geom. Dedicata. **78(2)** (1999), 183-199.
[5] A. Carriazo, L.M. Fernández and M.B. Hans-Uber, *B.Y. Chen's inequality for S-space forms:a pplications to slant immersions*, Indian J. Pure Appl. Math. **34(9)** (2003), 1287-1298.
[6] B.Y. Chen, *Slant immersions*, Bull. Austral. Math. Soc. **41(1)** (1990), 135-147.
[7] B.-Y. Chen, *Some pinching and classification theorems for minimal submanifolds*, Arch. Math. **60(1)** (1993), 568-578.
[8] B.-Y. Chen, *A Riemannian invariant and its applications to submanifold theory*, Results Math. **27** (1995), 17-26.
[9] B.-Y. Chen, *Relations between Ricci curvature and shape operator for submanifolds with arbitrary codimensions*, Glasg. Math. J. **41(1)** (1999), 33-44.
[10] D. Cioroboiu, *B.-Y. Chen inequalities for semislant submanifolds in Sasakian space forms*, Int. J. Math. Math. Sci. **27** (2003), 1731-1738.
[11] D. Cioroboiu, *B.-Y. Chen inequalities for bi-slant submanifolds in Sasakian space forms*, Demonstr Math. **36(1)** (2003), 179-187.
[12] D. Cioroboiu and A. Oiaga, *B.-Y. Chen's inequalities for slant submanifolds in Sasakian space forms*, Rend. Circ. Mat. Palermo **52(3)** (2003), 367-381.
[13] F. Defever, I. Mihai and L. Verstraelen, *B.-Y. Chen's inequality for C-totally real submanifolds of Sasakian space forms*, Boll. Un. Mat. Ital. **11(2)** (1997), 365-374.
[14] R.S. Gupta, *B.-Y. Chen's inequalities for bi-slant submanifolds in cosymplectic space forms*, Sarajevo. J. Math. **9(1)** (2013), 117-128.
[15] K. Kenmotsu, *A class of almost contact Riemannian manifolds*, Tohoku. Math. J. **24(1)** (1972), 93-103.
[16] J.S. Kim, Y.M. Song and M.M. Tripathi, *B.-Y. Chen's inequalities for submanifolds in generalized complex space forms*, Bull. Korean Math. Soc. **40(3)** (2003), 411-423.
[17] A. Lotta, *Slant submanifolds in contact geometry*, Bull. Math. Soc. Math. Roum. **39** (1996), 183-198.
[18] I. Mihai, *Ricci curvature of submanifolds in Sasakian space forms*, J. Aust. Math. Soc. **72(2)** (2002), 247-256.
[19] J.F. Nash, *The imbedding problem for Riemannian manifolds*, Annals of Mathematics **63(1)** (1956), 20-63.

[20] A. Oiaga and I. Mihai, *B.-Y. Chen's inequalities for slant submanifolds in complex space forms*, Demonstr. Math. **32(4)** (1999), 835-846.

[21] P.K. Pandey, R.S. Gupta and A. Sharfuddin, *B.-Y. Chen's inequalities for bi-slant submanifolds in Kenmotsu space forms*, Demonstr. Math. **4** (2010), 887-898.

[22] N. Papaghiuc, *Semi-slant submanifolds of a Kaehlerian manifold*, An. Stiint. Cuza. Univ. Iasi. **40** (1994), 62-78.

[23] G.E. Vâlcu, *B.-Y. Chen's inequalities for slant submanifolds in quaternionic space forms*, Turk. J. Math. **34(1)** (2010), 115-128.

[24] D.W. Yoon, *Inequality for Ricci curvature of slant submanifolds in cosymplectic space forms*, Turk. J. Math. **30** (2006), 43-56.

Department of Mathematics, Faculty of Mathematical Sciences and Statistics , Al Neelain University, Khartoum, Sudan

Email address: malsdeeg@gmail.com

FIXED POINT THEOREMS IN TRICOMPLEX VALUED BIPOLAR METRIC SPACES

G. SIVA AND L. RASIKANNAN

Abstract. The concept of tricomplex valued bipolar metric space is introduced in this article, and some properties are derived. For tricomplex valued bipolar metric spaces, some fixed point results of contravariant maps satisfying rational inequalities are also proved. Moreover, some examples are provided to illustrate our main results.

Mathematics Subject Classification (2010): 54E40, 54H25.
Key words: Tricomplex number, partial order, contravariant map, fixed point.

Article history:
Received: November 12, 2022
Received in revised form: December 7, 2022
Accepted: December 10, 2022

1. Introduction

Let $\mathbb{C}_0$ and $\mathbb{C}_1$ be the set of all real and complex numbers respectively. Bicomplex numbers are defined by C. Segre [10] as: $z = \mho_1 + \mho_2 i_1 + \mho_3 i_2 + \mho_4 i_1 i_2$, where $\mho_1, \mho_2, \mho_3, \mho_4 \in \mathbb{C}_0$, and the independent units i_1, i_2 are such that ${i_i}^2 = {i_2}^2 = -1$, and $i_1 i_2 = i_2 i_1$. We denote the set of bicomplex numbers $\mathbb{C}_2$ is defined as: $\mathbb{C}_2 = \{z : z = \mho_1 + \mho_2 i_1 + \mho_3 i_2 + \mho_4 i_1 i_2, \mho_1, \mho_2, \mho_3, \mho_4 \in \mathbb{C}_0\}$, i.e., $\mathbb{C}_2 = \{z : z = l_1 + i_2 l_2, l_1, l_2 \in \mathbb{C}_1\}$, where $l_1 = \mho_1 + \mho_2 i_1 \in \mathbb{C}_1$ and $l_2 = \mho_3 + \mho_4 i_1 \in \mathbb{C}_1$.

Tricomplex numbers are defined by G. B. Price [9] as: $\eta = \mho_1 + \mho_2 i_1 + \mho_3 i_2 + \mho_4 j_1 + \mho_5 i_3 + \mho_6 j_2 + \mho_7 j_3 + \mho_8 i_4$, where $\mho_1, \mho_2, \mho_3, \mho_4, \mho_5, \mho_6, \mho_7, \mho_8 \in \mathbb{C}_0$, and the independent units $i_1, i_2, i_3, i_4, j_1, j_2, j_3$ are such that ${i_i}^2 = {i_4}^2 = -1$, $i_4 = i_1 j_3 = i_1 i_2 i_3$, $j_2 = i_1 i_3 = i_3 i_1, {j_2}^2 = 1, j_1 = i_1 i_2 = i_2 i_1$ and ${j_1}^2 = 1$. We denote the set of tricomplex numbers $\mathbb{C}_3$ is defined as:
$\mathbb{C}_3 = \{\eta : \eta = \mho_1 + \mho_2 i_1 + \mho_3 i_2 + \mho_4 j_1 + \mho_5 i_3 + \mho_6 j_2 + \mho_7 j_3 + \mho_8 i_4, \mho_1, \mho_2, \mho_3, \mho_4, \mho_5, \mho_6, \mho_7, \mho_8 \in \mathbb{C}_0\}$, i.e., $\mathbb{C}_3 = \{\eta : \eta = z_1 + i_3 z_2, z_1, z_2 \in \mathbb{C}_2\}$, where $z_1 = \mho_1 + \mho_2 i_2 \in \mathbb{C}_2$ and $z_2 = \mho_3 + \mho_4 i_2 \in \mathbb{C}_2$.

If $\eta = z_1 + i_3 z_2$ and $\mu = w_1 + i_3 w_2$ be any two tricomplex numbers then the sum is $\eta \pm \mu = (z_1 + i_3 z_2) \pm (w_1 + i_3 w_2) = (z_1 \pm w_1) + i_3(z_2 \pm w_2)$ and the product is $\eta.\mu = (z_1 + i_3 z_2).(w_1 + i_3 w_2) = (z_1 w_1 - z_2 w_2) + i_3(z_1 w_2 + z_2 w_1)$.

Let $0, 1, e_1 = 1 + j_3/2, e_2 = 1 - j_3/2$ be four idempotent elements in $\mathbb{C}_3$ such that $e_1 + e_2 = 1$ and $e_1 e_2 = 0$. Every tricomplex number $\eta = z_1 + i_3 z_2$ can be uniquely be expressed as the combination of e_1 and e_2, i.e., $\eta = z_1 + i_3 z_2 = (z_1 - i_2 z_2)e_1 + (z_1 + i_2 z_2)e_2$. This representation of η is called the idempotent representation with respect to the idempotent components $\eta_1 = (z_1 - i_2 z_2)e_1$ and $\eta_2 = (z_1 + i_2 z_2)e_2$.

An element $\eta = z_1 + i_3 z_2 \in \mathbb{C}_3$ is called invertible if there exists an element μ in $\mathbb{C}_3$ such that $\eta\mu = 1$ where μ is called inverse of η. An element in $\mathbb{C}_3$ is called nonsingular element if it has an inverse in $\mathbb{C}_3$ and an element in $\mathbb{C}_3$ is called singular element if it does not have an inverse in $\mathbb{C}_3$.

An element $\mu = w_1 + i_2 w_2 \in \mathbb{C}_3$ is nonsingular iff $|{w_1}^2 + {w_2}^2| \neq 0$ and singular iff $|{w_1}^2 + {w_2}^2| = 0$. The inverse of μ is defined as $\mu^{-1} = \frac{w_1 - i_2 w_2}{{w_1}^2 + {w_2}^2}$.

The norm $||.||$ of $\mathbb{C}_3$ is a positive real valued function and $||.|| : \mathbb{C}_3 \to {\mathbb{C}_0}^+$ by

$$\begin{aligned} ||\eta|| &= ||z_1 + i_3 z_2|| \\ &= \{|z_1|^2 + |z_2|^2\}^{\frac{1}{2}} \\ &= \Big[\frac{|(z_1 - i_2 z_2)|^2 + |(z_1 + i_2 z_2)|^2}{2}\Big]^{\frac{1}{2}} \\ &= ({\mho_1}^2 + {\mho_2}^2 + {\mho_3}^2 + {\mho_4}^2 + {\mho_5}^2 + {\mho_6}^2 + {\mho_7}^2 + {\mho_8}^2)^{\frac{1}{2}}, \end{aligned}$$

where $\eta = \mho_1 + \mho_2 i_1 + \mho_3 i_2 + \mho_4 j_1 + \mho_5 i_3 + \mho_6 j_2 + \mho_7 j_3 + \mho_8 i_4 = z_1 + i_3 z_2 \in \mathbb{C}_3$.

Define a partial order $\precsim_{i_3}$ on $\mathbb{C}_3$ as follows. For $\eta = z_1 + i_3 z_2$ and $\mu = w_1 + i_3 w_2$ be any two tricomplex numbers. $\eta \precsim_{i_3} \mu$ if and only if $z_1 \precsim_{i_2} w_1$, and $z_2 \precsim_{i_2} w_2$. It follows that $\eta \precsim_{i_3} \mu$ if one of the following conditions is satisfied:

(i) $z_1 = w_1$, $z_2 = w_2$,
(ii) $z_1 \prec_{i_2} w_1$, $z_2 = w_2$,
(iii) $z_1 = w_1$, $z_2 \prec_{i_2} w_2$,
(iv) $z_1 \prec_{i_2} w_1$, $z_2 \prec_{i_2} w_2$.

In particular we will write $\eta \precnsim_{i_3} \mu$ if $\eta \precsim_{i_3} \mu$ and $\eta \neq \mu$ and one of (ii),(iii), and (iv) is satisfied, and we will write $\eta \prec_{i_3} \mu$ if only (iv) is satisfied. Note that

(I) $\eta \precsim_{i_3} \mu \Rightarrow ||\eta|| \leq ||\mu||$,
(II) $||\eta + \mu|| \leq ||\eta|| + ||\mu||$,
(III) $||a\eta|| = |a|||\eta||$, where a is a non negative real number,
(IV) $||\eta\mu|| \leq 2||\eta||||\mu||$, and the equality holds only when atleast one of η and μ is nonsingular,
(V) $||\eta^{-1}|| = ||\eta||^{-1}$ if η is a nonsingular,
(VI) $||\frac{\eta}{\mu}|| = \frac{||\eta||}{||\mu||}$, if μ is a nonsingular.

A. Azam et al introduced the concept of complex valued metric spaces in [1]. The notion of bicomplex valued metric spaces was introduced by J. Choi et al in [2]. In [5], G. Mani et al. introduced the idea of tricomplex valued metric spaces, developed some properties, and demonstrated common fixed point results for mappings satisfying a rational inequality.

Definition 1.1. [5] *Let $\Phi \neq \emptyset$ be a set. A tricomplex valued metric is a mapping $d : \Phi \times \Phi \to \mathbb{C}_3$ satisfying the following axioms:*

(i) $0 \precsim_{i_3} d(\varrho, \varpi)$, $\forall\ \varrho, \varpi \in \Phi$,
(ii) $d(\varrho, \varpi) = 0$ *if and only if* $\varrho = \varpi$ *in* Φ,
(iii) $d(\varrho, \varpi) = d(\varpi, \varrho)$, $\forall\ \varrho, \varpi \in \Phi$,
(iv) $d(\varrho, \varpi) \precsim_{i_3} d(\varrho, \theta) + d(\theta, \varpi)$, $\forall\ \varrho, \theta, \varpi \in \Phi$.

The pair (Φ, d) is called a tricomplex valued metric space.

A. Mutlu et al [7] introduced the notion of bipolar metric space to give a new definition of distance measurement between the members of two separate sets. Bipolar metric space is a metric space generalization. Recently, many articles have appeared on fixed point theory in bipolar metric spaces; see, for example, [3, 4, 6, 8, 12, 11] and the references therein.

Definition 1.2. [7] *Let $\Phi \neq \emptyset$ and $\Psi \neq \emptyset$ be two sets. A bipolar metric is a mapping $D : \Phi \times \Psi \to [0, \infty)$ satisfying the following axioms:*

(I) $D(\varrho, \varpi) = 0 \Rightarrow \varrho = \varpi$, *whenever* $(\varrho, \varpi) \in \Phi \times \Psi$,
(II) $\varrho = \varpi \Rightarrow D(\varrho, \varpi) = 0$, *whenever* $(\varrho, \varpi) \in \Phi \times \Psi$,
(III) $D(\varrho, \varpi) = D(\varpi, \varrho)$, $\forall\ \varrho, \varpi \in \Phi \cap \Psi$,
(IV) $D(\varrho_1, \varpi_2) \leq D(\varrho_1, \varpi_1) + D(\varrho_2, \varpi_1) + D(\varrho_2, \varpi_2)$, $\forall\ \varrho_1, \varrho_2 \in \Phi$, *and* $\varpi_1, \varpi_2 \in \Psi$.

The triple (Φ, Ψ, D) is called a bipolar metric space.

In this paper, we extend the domain of tricomplex valued metric to the Cartesian product of two non-empty sets, and we present a new definition of tricomplex valued bipolar metric space that generalizes the notion of tricomplex valued metric space. Also, we derive some properties of tricomplex valued bipolar metric spaces. Furthermore, in tricomplex valued bipolar metric space, we prove some fixed point results for contravariant maps satisfying various types of rational inequalities.

2. Tricomplex valued bipolar metric spaces

Definition 2.1. *Let $\Phi \neq \emptyset$ and $\Psi \neq \emptyset$ be two sets. A tricomplex valued bipolar metric is a mapping $d : \Phi \times \Psi \to \mathbb{C}_3$ satisfying the following conditions:*

(i) $0 \precsim_{i_3} d(\varrho, \varpi)$, *whenever* $(\varrho, \varpi) \in \Phi \times \Psi$,
(ii) $d(\varrho, \varpi) = 0 \Rightarrow \varrho = \varpi$, *whenever* $(\varrho, \varpi) \in \Phi \times \Psi$,
(iii) $\varrho = \varpi \Rightarrow d(\varrho, \varpi) = 0$, *whenever* $(\varrho, \varpi) \in \Phi \times \Psi$,
(iv) $d(\varrho, \varpi) = d(\varpi, \varrho)$, $\forall\ \varrho, \varpi \in \Phi \cap \Psi$,
(v) $d(\varrho_1, \varpi_2) \precsim_{i_3} d(\varrho_1, \varpi_1) + d(\varrho_2, \varpi_1) + d(\varrho_2, \varpi_2)$, $\forall\ \varrho_1, \varrho_2 \in \Phi$, *and* $\varpi_1, \varpi_2 \in \Psi$.

The triple (Φ, Ψ, d) is called a tricomplex valued bipolar metric space(or, TVBMS).

Remark 2.2. (i) *Let (Φ, Ψ, d) be a TVBMS. If $\Phi \cap \Psi = \emptyset$, then (Φ, Ψ, d) is called disjoint. The space (Φ, Ψ, d) is said to be a joint if $\Phi \cap \Psi \neq \emptyset$. The sets Ψ and Φ are called right pole and left pole of (Φ, Ψ, d), respectively.*
(ii) *Let (Φ, d) be a tricomplex valued metric space, then (Φ, Φ, d) is a TVBMS. Conversely, if (Φ, Ψ, d) is a TVBMS such that $\Phi = \Psi$, then (Φ, d) is a tricomplex valued metric space.*

Example 2.3. *Let $\Phi = (0, \infty)$ and $\Psi = (-\infty, 0]$. Let $d(\varrho, \varpi) = (i_2 i_3)|\varrho - \varpi|$, where $(\varrho, \varpi) \in \Phi \times \Psi$. Then (Φ, Ψ, d) is a disjoint TVBMS.*

Definition 2.4. *Let (Φ, Ψ, d) be a TVBMS. Where points of the sets Ψ, Φ, and $\Phi \cap \Psi$ are called right, left, and central points respectively. A sequence that contains only right(or left, or central) points is called a right (or left, or central) sequence in (Φ, Ψ, d).*

Definition 2.5. *Let (Φ, Ψ, d) be a TVBMS. A left sequence $(\varrho_n)_{n=1}^{\infty}$ converges to a right point ϖ(or $(\varrho_n)_{n=1}^{\infty} \to \varpi$) if and only if for every $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists an integer $n_0 \in \mathbb{N}$(Natural numbers) such that $d(\varrho_n, \varpi) \prec_{i_3} c$, $\forall\ n \geq n_0$. Also a right sequence $(\varpi_n)_{n=1}^{\infty}$ converges to a left point ϱ (or $(\varpi_n)_{n=1}^{\infty} \to \varrho$) if and only if for every $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists an integer $n_0 \in \mathbb{N}$ such that $d(\varrho, \varpi_n) \prec_{i_3} c$, $\forall\ n \geq n_0$. When it is given $(\theta_n)_{n=1}^{\infty} \to \vartheta$ for a TVBMS (Φ, Ψ, d) without precise data about the sequence, this means that either $(\theta_n)_{n=1}^{\infty}$ is a right sequence and ϑ is a left point, or $(\theta_n)_{n=1}^{\infty}$ is a left sequence and ϑ is a right point.*

Lemma 2.6. *Let (Φ, Ψ, d) be a TVBMS. Then a left sequence $(\varrho_n)_{n=1}^{\infty}$ converges to a right point ϖ if and only if $||d(\varrho_n, \varpi)|| \to 0$, and also a right sequence $(\varpi_n)_{n=1}^{\infty}$ converges to a left point ϱ if and only if $||d(\varrho, \varpi_n)|| \to 0$.*

Proof. Let $(\varrho_n)_{n=1}^{\infty}$ be a left sequence, and $(\varrho_n)_{n=1}^{\infty} \to \varpi \in \Psi$. For a given real number $\epsilon > 0$, let $c = \frac{\epsilon}{\sqrt{8}} + i_1 \frac{\epsilon}{\sqrt{8}} + i_2 \frac{\epsilon}{\sqrt{8}} + j_1 \frac{\epsilon}{\sqrt{8}} + i_3 \frac{\epsilon}{\sqrt{8}} + j_2 \frac{\epsilon}{\sqrt{8}} + j_3 \frac{\epsilon}{\sqrt{8}} + i_4 \frac{\epsilon}{\sqrt{8}}$. For every $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists an integer $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$, $d(\varrho_n, \varpi) \prec_{i_3} c$.

$$||d(\varrho_n, \varpi)|| < ||c|| = \epsilon,\ \forall\ n \geq n_0.$$

It follows that $||d(\varrho_n, \varpi)|| \to 0$ as $n \to \infty$. Conversely, suppose that $||d(\varrho_n, \varpi)|| \to 0$ as $n \to \infty$. Then given $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists a real number $\delta > 0$ such that for $z \in \mathbb{C}_3$

$$||z|| < \delta \Rightarrow z \prec_{i_3} c.$$

For this δ, there exists an integer $n_0 \in \mathbb{N}$ such that

$$||d(\varrho_n, \varpi)|| < \delta,\ \forall\ n \geq n_0.$$

This means that $d(\varrho_n, \varpi) \prec_{i_3} c, \forall n \geq n_0$. Hence $\varrho_n \to \varpi \in \Psi$.
Obviously, a right sequence $(\varpi_n)_{n=1}^{\infty}$ converges to a left point ϱ if and only if $||d(\varrho, \varpi_n)|| \to 0$ and this complete the proof. □

Lemma 2.7. *Let (Φ, Ψ, d) be a TVBMS. If a central point is a limit of a sequence, then it is the unique limit of the sequence.*

Proof. Let $(\varrho_n)_{n=1}^{\infty}$ be a left sequence, $(\varrho_n)_{n=1}^{\infty} \to \varrho \in \Phi \cap \Psi$, and $(\varrho_n)_{n=1}^{\infty} \to \varpi \in \Psi$. For a given real number $\epsilon > 0$, let $c = \frac{\epsilon}{\sqrt{8}} + i_1 \frac{\epsilon}{\sqrt{8}} + i_2 \frac{\epsilon}{\sqrt{8}} + j_1 \frac{\epsilon}{\sqrt{8}} + i_3 \frac{\epsilon}{\sqrt{8}} + j_2 \frac{\epsilon}{\sqrt{8}} + j_3 \frac{\epsilon}{\sqrt{8}} + i_4 \frac{\epsilon}{\sqrt{8}}$. For every $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists an integer $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$, we have $d(\varrho_n, \varrho) \prec_{i_3} \frac{c}{3}$, and $d(\varrho_n, \varpi) \prec_{i_3} \frac{c}{2}$, and then

$$d(\varrho, \varpi) \precsim_{i_3} d(\varrho, \varrho) + d(\varrho_n, \varrho) + d(\varrho_n, \varpi) \prec_{i_3} 0 + \frac{c}{2} + \frac{c}{2}.$$

$$||d(\varrho, \varpi)|| \leq ||d(\varrho, \varrho) + d(\varrho_n, \varrho) + d(\varrho_n, \varpi)|| < ||0 + \frac{c}{2} + \frac{c}{2}|| = ||c|| = \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have $d(\varrho, \varpi) = 0$ which implies $\varrho = \varpi$. □

Lemma 2.8. *Let (Φ, Ψ, d) be a TVBMS. If a left sequence $(\varrho_n)_{n=1}^{\infty}$ converges to ϖ and a right sequence $(\varpi_n)_{n=1}^{\infty}$ converges to ϱ, then $d(\varrho_n, \varpi_n) \to d(\varrho, \varpi)$ as $n \to \infty$.*

Proof. Let $(\varrho_n)_{n=1}^{\infty} \to \varpi \in \Psi$, and $(\varpi_n)_{n=1}^{\infty} \to \varrho \in \Phi$. For a given real number $\epsilon > 0$, let $c = \frac{\epsilon}{\sqrt{8}} + i_1 \frac{\epsilon}{\sqrt{8}} + i_2 \frac{\epsilon}{\sqrt{8}} + j_1 \frac{\epsilon}{\sqrt{8}} + i_3 \frac{\epsilon}{\sqrt{8}} + j_2 \frac{\epsilon}{\sqrt{8}} + j_3 \frac{\epsilon}{\sqrt{8}} + i_4 \frac{\epsilon}{\sqrt{8}}$. For every $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists an integer $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$, we have $d(\varrho_n, \varpi) \prec_{i_3} \frac{c}{2}$, and $d(\varrho, \varpi_n) \prec_{i_3} \frac{c}{2}$, then

$$d(\varrho, \varpi) \precsim_{i_3} d(\varrho, \varpi_n) + d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi)$$

implies

$$d(\varrho, \varpi) - d(\varrho_n, \varpi_n) \precsim_{i_3} d(\varrho, \varpi_n) + d(\varrho_n, \varpi) \prec \frac{c}{2} + \frac{c}{2},$$

$$||d(\varrho_n, \varpi_n) - d(\varrho, \varpi)|| \leq ||d(\varrho, \varpi_n) + d(\varrho_n, \varpi)|| < ||c|| = \epsilon, \forall n \geq n_0,$$

and hence $d(\varrho_n, \varpi_n) \to d(\varrho, \varpi)$ as $n \to \infty$. □

Definition 2.9. *Let (Φ_1, Ψ_1) and (Φ_2, Ψ_2) be two tricomplex valued bipolar metric spaces, and $f : \Phi_1 \cup \Psi_1 \to \Phi_2 \cup \Psi_2$.*

(i) *If $f(\Phi_1) \subseteq \Phi_2$ and $f(\Psi_1) \subseteq \Psi_2$, then f is called a covariant map from (Φ_1, Ψ_1) to (Φ_2, Ψ_2), and we write $f : (\Phi_1, \Psi_1) \rightrightarrows (\Phi_2, \Psi_2)$.*

(ii) *If $f(\Phi_1) \subseteq \Psi_2$ and $f(\Psi_1) \subseteq \Phi_2$, then f is called a contravariant map from (Φ_1, Ψ_1) to (Φ_2, Ψ_2), and we write $f : (\Phi_1, \Psi_1) \rightleftarrows (\Phi_2, \Psi_2)$.*

Remark 2.10. *Suppose d_1, and d_2 be two tricomplex valued bipolar metrics on (Φ_1, Ψ_1) and (Φ_2, Ψ_2) respectively. We can also use the symbols $f : (\Phi_1, \Psi_1, d_1) \rightrightarrows (\Phi_2, \Psi_2, d_2)$ and $f : (\Phi_1, \Psi_1, d_1) \rightleftarrows (\Phi_2, \Psi_2, d_2)$ in the place of $f : (\Phi_1, \Psi_1) \rightrightarrows (\Phi_2, \Psi_2)$ and $f : (\Phi_1, \Psi_1) \rightleftarrows (\Phi_2, \Psi_2)$.*

Definition 2.11. *Let (Φ, Ψ, d) be a TVBMS.*

(i) *A sequence (ϱ_n, ϖ_n) on the set $\Phi \times \Psi$ is called a bisequence on (Φ, Ψ, d).*

(ii) *If both $(\varrho_n)_{n=1}^{\infty}$ and $(\varpi_n)_{n=1}^{\infty}$ converges, then the bisequence (ϱ_n, ϖ_n) is called convergent. If both $(\varrho_n)_{n=1}^{\infty}$ and $(\varpi_n)_{n=1}^{\infty}$ converges to a same point $\varrho \in \Phi \cap \Psi$, then the bisequence is called biconvergent.*

(iii) *A bisequence (ϱ_n, ϖ_n) on (Φ, Ψ, d) is called a Cauchy bisequence, if for each $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there is an $n_0 \in \mathbb{N}$ such that $d(\varrho_n, \varpi_{n+m}) \prec_{i_3} c$, $\forall\ n \geq n_0$.*

Lemma 2.12. *Let (Φ, Ψ, d) be a TVBMS. Then (ϱ_n, ϖ_n) is a Cauchy bisequence if and only if $||d(\varrho_n, \varpi_{n+m})|| \to 0$ as $n \to \infty$.*

Proof. Let (ϱ_n, ϖ_n) is a Cauchy bisequence. For a given real number $\epsilon > 0$, let $c = \frac{\epsilon}{\sqrt{8}} + i_1\frac{\epsilon}{\sqrt{8}} + i_2\frac{\epsilon}{\sqrt{8}} + j_1\frac{\epsilon}{\sqrt{8}} + i_3\frac{\epsilon}{\sqrt{8}} + j_2\frac{\epsilon}{\sqrt{8}} + j_3\frac{\epsilon}{\sqrt{8}} + i_4\frac{\epsilon}{\sqrt{8}}$. For every $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists an integer $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$, $d(\varrho_n, \varpi_{n+m}) \prec_{i_3} c$.

$$||d(\varrho_n, \varpi_{n+m})|| < ||c|| = \epsilon, \ \forall\ n \geq n_0.$$

It follows that $||d(\varrho_n, \varpi_{n+m})|| \to 0$ as $n \to \infty$. Conversely, suppose that $||d(\varrho_n, \varpi_{n+m})|| \to 0$ as $n \to \infty$. Then given $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists a real number $\delta > 0$ such that for $z \in \mathbb{C}_3$

$$||z|| < \delta \Rightarrow z \prec_{i_3} c$$

For this δ, there exists an integer $n_0 \in \mathbb{N}$ such that

$$||d(\varrho_n, \varpi_{n+m})|| < \delta, \ \forall\ n \geq n_0.$$

This means that $d(\varrho_n, \varpi_{n+m}) \prec_{i_3} c, \forall n \geq n_0$. Hence (ϱ_n, ϖ_n) is a Cauchy bisequence. □

Proposition 2.13. *Let (Φ, Ψ, d) be a TVBMS. Then every biconvergent bisequence is a Cauchy bisequence.*

Proof. Let (ϱ_n, ϖ_n) be a bisequence, which is biconvergent to a point $\varrho \in \Phi \cap \Psi$. For a given real number $\epsilon > 0$, let $c = \frac{\epsilon}{\sqrt{8}} + i_1\frac{\epsilon}{\sqrt{8}} + i_2\frac{\epsilon}{\sqrt{8}} + j_1\frac{\epsilon}{\sqrt{8}} + i_3\frac{\epsilon}{\sqrt{8}} + j_2\frac{\epsilon}{\sqrt{8}} + j_3\frac{\epsilon}{\sqrt{8}} + i_4\frac{\epsilon}{\sqrt{8}}$. For every $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists an integer $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$, $d(\varrho_n, \varrho) \prec_{i_3} \frac{c}{2}$, and for every $n \geq n_0$, $d(\varrho, \varpi_{n+m}) \prec_{i_3} \frac{c}{2}$. Then we have

$$d(\varrho_n, \varpi_{n+m}) \precsim_{i_3} d(\varrho_n, \varrho) + d(\varrho, \varrho) + d(\varrho, \varpi_{n+m}) \prec_{i_3} \frac{c}{2} + 0 + \frac{c}{2}, \forall n \geq n_0.$$

$$||d(\varrho_n, \varpi_{n+m})|| \leq ||d(\varrho_n, \varrho) + d(\varrho, \varrho) + d(\varrho, \varpi_{n+m})|| < ||\frac{c}{2} + 0 + \frac{c}{2}|| = ||c|| = \epsilon, \forall n \geq n_0.$$

So (ϱ_n, ϖ_n) is a Cauchy bisequence. □

Proposition 2.14. *Let (Φ, Ψ, d) be a TVBMS. Then every convergent Cauchy bisequence is biconvergent.*

Proof. Let (ϱ_n, ϖ_n) be a Cauchy bisequence such that $(\varrho_n)_{n=1}^{\infty}$ convergent to ϖ in Ψ and $(\varpi_n)_{n=1}^{\infty}$ convergent to ϱ in Φ. For a given real number $\epsilon > 0$, let $c = \frac{\epsilon}{\sqrt{8}} + i_1\frac{\epsilon}{\sqrt{8}} + i_2\frac{\epsilon}{\sqrt{8}} + j_1\frac{\epsilon}{\sqrt{8}} + i_3\frac{\epsilon}{\sqrt{8}} + j_2\frac{\epsilon}{\sqrt{8}} + j_3\frac{\epsilon}{\sqrt{8}} + i_4\frac{\epsilon}{\sqrt{8}}$. For every $c \in \mathbb{C}_3$ with $0 \prec_{i_3} c$, there exists an integer $n_0 \in \mathbb{N}$ such that $d(\varrho_n, \varpi) \prec_{i_3} \frac{c}{3}$, $d(\varrho, \varpi_{n+m}) \prec_{i_3} \frac{c}{3}$, for all $n \geq n_0$, and $d(\varrho_n, \varpi_{n+m}) \prec_{i_3} \frac{c}{3}$, for all $n \geq n_0$. Then

$$d(\varrho, \varpi) \precsim_{i_3} d(\varrho, \varpi_{n+m}) + d(\varrho_n, \varpi_{n+m}) + d(\varrho_n, \varpi) \prec_{i_3} \frac{c}{3} + \frac{c}{3} + \frac{c}{3}, \forall n \geq n_0.$$

$$||d(\varrho, \varpi)|| \leq ||d(\varrho, \varpi_{n+m}) + d(\varrho_n, \varpi_{n+m}) + d(\varrho_n, \varpi)|| < ||\frac{c}{3} + \frac{c}{3} + \frac{c}{3}|| = ||c|| = \epsilon, \forall n \geq n_0.$$

Therefore $d(\varrho, \varpi) = 0$ and so that $\varrho = \varpi$. Then (ϱ_n, ϖ_n) is biconvergent. □

Definition 2.15. *A TVBMS (Φ, Ψ, d) is called complete, if every Cauchy bisequence is convergent, or equivalently, biconvergent.*

3. Main results

In this section, we shall prove some fixed point theorems for different types of contravariant mappings on TVBMS.

Theorem 3.1. *Let (Φ, Ψ, d) be a complete TVBMS with nonsingular $1 + d(\varrho, \varpi)$ and $||1 + d(\varrho, \varpi)|| \neq 0$, whenever $(\varrho, \varpi) \in \Phi \times \Psi$. If a contravariant map $f : (\Phi, \Psi, d) \rightleftarrows (\Phi, \Psi, d)$ satisfies $d(f(\varpi), f(\varrho)) \precsim_{i_3} \lambda d(\varrho, \varpi) + \frac{\mu d(\varrho, f(\varrho)) d(f(\varpi), \varpi)}{1 + d(\varrho, \varpi)}$, whenever $(\varrho, \varpi) \in \Phi \times \Psi$, for some $\lambda, \mu \in (0, 1)$ with $\lambda + 2\mu < 1$. Then the function $f : \Phi \cup \Psi \to \Phi \cup \Psi$ has a unique fixed point(or, UFP).*

Proof. Let $\varrho_0 \in \Phi$, $\varpi_0 = f(\varrho_0) \in \Psi$, and $\varrho_1 = f(\varpi_0)$. Suppose, $\varpi_n = f(\varrho_n)$ and $\varrho_{n+1} = f(\varpi_n)$, for all $n \in \mathbb{N}$. Then (ϱ_n, ϖ_n) is a bisequence on (Φ, Ψ, d). For all $n \in \mathbb{N}$, from

$$\begin{aligned}
d(\varrho_n, \varpi_n) &= d(f(\varpi_{n-1}), f(\varrho_n)) \\
&\precsim_{i_3} \lambda d(\varrho_n, \varpi_{n-1}) + \frac{\mu d(\varrho_n, f(\varrho_n))d(f(\varpi_{n-1}), \varpi_{n-1})}{1 + d(\varrho_n, \varpi_{n-1})} \\
&= \lambda d(\varrho_n, \varpi_{n-1}) + \frac{\mu d(\varrho_n, \varpi_n)d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_n, \varpi_{n-1})} \\
||d(\varrho_n, \varpi_n)|| &\leq ||\lambda d(\varrho_n, \varpi_{n-1}) + \frac{\mu d(\varrho_n, \varpi_n)d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_n, \varpi_{n-1})}|| \\
&\leq \lambda ||d(\varrho_n, \varpi_{n-1})|| + 2\mu ||d(\varrho_n, \varpi_n)||
\end{aligned}$$

we conclude that

$$||d(\varrho_n, \varpi_n)|| \leq \frac{\lambda}{1 - 2\mu} ||d(\varrho_n, \varpi_{n-1})||,$$

and

$$\begin{aligned}
d(\varrho_n, \varpi_{n-1}) &= d(f(\varpi_{n-1}), f(\varrho_{n-1})) \\
&\precsim_{i_3} \lambda d(\varrho_{n-1}, \varpi_{n-1}) + \frac{\mu d(\varrho_{n-1}, f(\varrho_{n-1}))d(f(\varpi_{n-1}), \varpi_{n-1})}{1 + d(\varrho_{n-1}, \varpi_{n-1})} \\
&= \lambda d(\varrho_{n-1}, \varpi_{n-1}) + \frac{\mu d(\varrho_{n-1}, \varpi_{n-1})d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_{n-1}, \varpi_{n-1})} \\
||d(\varrho_n, \varpi_{n-1})|| &\leq ||\lambda d(\varrho_{n-1}, \varpi_{n-1}) + \frac{\mu d(\varrho_{n-1}, \varpi_{n-1})d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_{n-1}, \varpi_{n-1})}|| \\
&\leq \lambda ||d(\varrho_{n-1}, \varpi_{n-1})|| + 2\mu ||d(\varrho_n, \varpi_{n-1})||,
\end{aligned}$$

so that

$$||d(\varrho_n, \varpi_{n-1})|| \leq \frac{\lambda}{1 - 2\mu} ||d(\varrho_{n-1}, \varpi_{n-1})||,$$

Therefore, by putting $\beta = \frac{\lambda}{1-2\mu}$, we have

$$||d(\varrho_n, \varpi_n)|| \leq \beta^{2n} ||d(\varrho_0, \varpi_0)||$$

and

$$||d(\varrho_n, \varpi_{n-1})|| \leq \beta^{2n-1} ||d(\varrho_0, \varpi_0)||.$$

For every $m, n \in \mathbb{N}$,

$$\begin{aligned}
d(\varrho_n, \varpi_m) &\precsim_{i_3} d(\varrho_n, \varpi_n) + d(\varrho_{n+1}, \varpi_n) + d(\varrho_{n+1}, \varpi_m) \\
&\precsim_{i_3} (\beta^{2n} + \beta^{2n+1}) d(\varrho_0, \varpi_0) + d(\varrho_{n+1}, \varpi_m) \\
&\precsim_{i_3} \ldots \\
&\precsim_{i_3} (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m-1}) d(\varrho_0, \varpi_0) + d(\varrho_m, \varpi_m) \\
&\precsim_{i_3} (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m}) d(\varrho_0, \varpi_0), \text{ if } m > n,
\end{aligned}$$

$$||d(\varrho_n, \varpi_m)|| \leq (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m}) ||d(\varrho_0, \varpi_0)||, \text{ if } m > n,$$

and similarly, if $m < n$, then

$$d(\varrho_n, \varpi_m) \precsim_{i_3} (\beta^{2m+1} + \beta^{2m+2} + \ldots + \beta^{2n+1}) d(\varrho_0, \varpi_0),$$

$$||d(\varrho_n, \varpi_m)|| \leq (\beta^{2m+1} + \beta^{2m+2} + \ldots + \beta^{2n+1}) ||d(\varrho_0, \varpi_0)||.$$

By $\beta \in (0,1)$, $||d(\varrho_n, \varpi_m)|| \to 0$, as $n, m \to \infty$, we conclude that (ϱ_n, ϖ_n) is a Cauchy bisequence. Since (Φ, Ψ, d) is complete, (ϱ_n, ϖ_n) converges, and biconverges to a point $\theta \in \Phi \cap \Psi$. Hence, $f(\varrho_n) = \varpi_n \to \theta \in \Phi \cap \Psi$ as $n \to \infty$ implies $d(f(\theta), f(\varrho_n)) \to d(f(\theta), \theta)$ as $n \to \infty$, by using Lemma 2.8. Also by taking the limit from

$$d(f(\theta), f(\varrho_n)) \precsim_{i_3} \lambda d(\varrho_n, \theta) + \frac{\mu d(\varrho_n, \varpi_n) d(f(\theta), \theta)}{1 + d(\varrho_n, \theta)}$$

we obtain

$$||d(f(\theta), f(\varrho_n))|| \leq \lambda ||d(\varrho_n, \theta)|| + \frac{\mu ||d(\varrho_n, \varpi_n) d(f(\theta), \theta)||}{||1 + d(\varrho_n, \theta)||},$$

as $n \to \infty$, we get $d(f(\theta), \theta) = 0$. Hence $f(\theta) = \theta$. Therefore θ is a fixed point of f.
If ϑ is another fixed point of f, then $f(\vartheta) = \vartheta$, $\vartheta \in \Phi \cap \Psi$, and hence,

$$d(\theta, \vartheta) = d(f(\theta), f(\vartheta)) \precsim_{i_3} \lambda d(\theta, \vartheta) + \frac{\mu d(\theta, f(\theta)) d(f(\vartheta), \vartheta)}{1 + d(\theta, \vartheta)} \precsim_{i_3} \lambda d(\theta, \vartheta).$$

Therefore $||d(\theta, \vartheta)|| = 0$ so that $\theta = \vartheta$. So f has a UFP. □

The above Theorem 3.1 generalizes a Corollary 5 of [1].

Example 3.2. *Let $\Phi = \{0, \frac{1}{2}, 2\}$ and $\Psi = \{0, \frac{1}{2}\}$. Let $d(\varrho, \varpi) = (1 + i_3)|\varrho - \varpi|$, where $(\varrho, \varpi) \in \Phi \times \Psi$. Then (Φ, Ψ, d) is a complete TVBMS. Define a contravariant map $f : (\Phi, \Psi, d) \rightleftarrows (\Phi, \Psi, d)$ by $f(0) = 0$, $f(\frac{1}{2}) = 0$, and $f(2) = \frac{1}{2}$. Then, f satisfies the inequality $d(f(\varpi), f(\varrho)) \precsim_{i_3} \lambda d(\varrho, \varpi) + \frac{\mu d(\varrho, f(\varrho)) d(f(\varpi), \varpi)}{1 + d(\varrho, \varpi)}$ for $\lambda = \frac{1}{3}$ and $\mu = \frac{1}{6}$. By Theorem 3.1, f has a UFP zero in $\Phi \cap \Psi$.*

Theorem 3.3. *Let (Φ, Ψ, d) be a complete TVBMS with nonsingular $1 + d(\varrho, \varpi)$ and $||1 + d(\varrho, \varpi)|| \neq 0$, whenever $(\varrho, \varpi) \in \Phi \times \Psi$. If a contravariant map $f : (\Phi, \Psi, d) \rightleftarrows (\Phi, \Psi, d)$ satisfies $d(f(\varpi), f(\varrho)) \precsim_{i_3} \lambda[d(\varrho, f(\varrho)) + d(f(\varpi), \varpi)] + \frac{\mu d(\varrho, f(\varrho)) d(f(\varpi), \varpi)}{1 + d(\varrho, \varpi)}$, whenever $(\varrho, \varpi) \in \Phi \times \Psi$, for some $\lambda, \mu \in (0,1)$ with $2\lambda + 2\mu < 1$. Then the function $f : \Phi \cup \Psi \to \Phi \cup \Psi$ has a UFP.*

Proof. Let $\varrho_0 \in \Phi$, $\varpi_0 = f(\varrho_0) \in \Psi$, and $\varrho_1 = f(\varpi_0)$. Suppose, $\varpi_n = f(\varrho_n)$ and $\varrho_{n+1} = f(\varpi_n)$, for all $n \in \mathbb{N}$. Then (ϱ_n, ϖ_n) is a bisequence on (Φ, Ψ, d). For all $n \in \mathbb{N}$, from

$$\begin{aligned}
d(\varrho_n, \varpi_n) &= d(f(\varpi_{n-1}), f(\varrho_n)) \\
&\precsim_{i_3} \lambda[d(\varrho_n, f(\varrho_n)) + d(f(\varpi_{n-1}), \varpi_{n-1})] + \frac{\mu d(\varrho_n, f(\varrho_n)) d(f(\varpi_{n-1}), \varpi_{n-1})}{1 + d(\varrho_n, \varpi_{n-1})} \\
&= \lambda[d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi_{n-1})] + \frac{\mu d(\varrho_n, \varpi_n) d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_n, \varpi_{n-1})} \\
||d(\varrho_n, \varpi_n)|| &\leq ||\lambda[d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi_{n-1})] + \frac{\mu d(\varrho_n, \varpi_n) d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_n, \varpi_{n-1})}|| \\
&\leq \lambda ||[d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi_{n-1})]|| + 2\mu ||d(\varrho_n, \varpi_n)||,
\end{aligned}$$

we conclude that

$$||d(\varrho_n, \varpi_n)|| \leq \frac{\lambda}{1 - \lambda - 2\mu} ||d(\varrho_n, \varpi_{n-1})||,$$

and

$$\begin{aligned}
d(\varrho_n, \varpi_{n-1}) &= d(f(\varpi_{n-1}), f(\varrho_{n-1})) \\
&\precsim_{i_3} \lambda[d(\varrho_{n-1}, f(\varrho_{n-1})) + d(f(\varpi_{n-1}), \varpi_{n-1})] \\
&\quad + \frac{\mu d(\varrho_{n-1}, f(\varrho_{n-1})) d(f(\varpi_{n-1}), \varpi_{n-1})}{1 + d(\varrho_{n-1}, \varpi_{n-1})} \\
&= \lambda[d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_n, \varpi_{n-1})] + \frac{\mu d(\varrho_{n-1}, \varpi_{n-1}) d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_{n-1}, \varpi_{n-1})} \\
||d(\varrho_n, \varpi_{n-1})|| &\leq ||\lambda[d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_n, \varpi_{n-1})] + \frac{\mu d(\varrho_{n-1}, \varpi_{n-1}) d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_{n-1}, \varpi_{n-1})}|| \\
&\leq \lambda||[d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_n, \varpi_{n-1})]|| + 2\mu||d(\varrho_n, \varpi_{n-1})||
\end{aligned}$$

so that

$$||d(\varrho_n, \varpi_{n-1})|| \leq \frac{\lambda}{1 - \lambda - 2\mu} ||d(\varrho_{n-1}, \varpi_{n-1})||,$$

Therefore, by putting $\beta = \frac{\lambda}{1-\lambda-2\mu}$, we have

$$||d(\varrho_n, \varpi_n)|| \leq \beta^{2n} ||d(\varrho_0, \varpi_0)||$$

and

$$||d(\varrho_n, \varpi_{n-1})|| \leq \beta^{2n-1} ||d(\varrho_0, \varpi_0)||.$$

For every $m, n \in \mathbb{N}$,

$$\begin{aligned}
d(\varrho_n, \varpi_m) &\precsim_{i_3} d(\varrho_n, \varpi_n) + d(\varrho_{n+1}, \varpi_n) + d(\varrho_{n+1}, \varpi_m) \\
&\precsim_{i_3} (\beta^{2n} + \beta^{2n+1}) d(\varrho_0, \varpi_0) + d(\varrho_{n+1}, \varpi_m) \\
&\precsim_{i_3} \ldots \\
&\precsim_{i_3} (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m-1}) d(\varrho_0, \varpi_0) + d(\varrho_m, \varpi_m) \\
&\precsim_{i_3} (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m}) d(\varrho_0, \varpi_0), \text{ if } m > n,
\end{aligned}$$

$$||d(\varrho_n, \varpi_m)|| \leq (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m}) ||d(\varrho_0, \varpi_0)||, \text{ if } m > n,$$

and similarly, if $m < n$, then

$$d(\varrho_n, \varpi_m) \precsim_{i_3} (\beta^{2m+1} + \beta^{2m+2} + \ldots + \beta^{2n+1}) d(\varrho_0, \varpi_0),$$

$$||d(\varrho_n, \varpi_m)|| \leq (\beta^{2m+1} + \beta^{2m+2} + \ldots + \beta^{2n+1}) ||d(\varrho_0, \varpi_0)||.$$

By $\beta \in (0, 1)$, $||d(\varrho_n, \varpi_m)|| \to 0$, as $n, m \to \infty$, we conclude that (ϱ_n, ϖ_n) is a Cauchy bisequence. Since (Φ, Ψ, d) is complete, (ϱ_n, ϖ_n) converges, and biconverges to a point $\theta \in \Phi \cap \Psi$. Hence, $f(\varrho_n) = \varpi_n \to \theta \in \Phi \cap \Psi$ as $n \to \infty$ implies $d(f(\theta), f(\varrho_n)) \to d(f(\theta), \theta)$ as $n \to \infty$, by using Lemma 2.8. Also by taking the limit from

$$d(f(\theta), f(\varrho_n)) \precsim_{i_3} \lambda[d(\varrho_n, \varpi_n) + d(f(\theta), \theta)] + \frac{\mu d(\varrho_n, \varpi_n) d(f(\theta), \theta)}{1 + d(\varrho_n, \theta)}$$

we obtain

$$||d(f(\theta), f(\varrho_n))|| \leq \lambda[||d(\varrho_n, \varpi_n) + d(f(\theta), \theta)||] + \frac{\mu||d(\varrho_n, \varpi_n) d(f(\theta), \theta)||}{||1 + d(\varrho_n, \theta)||},$$

as $n \to \infty$, we get $d(f(\theta), \theta) = 0$. Hence $f(\theta) = \theta$. Therefore θ is a fixed point of f.
If ϑ is another fixed point of f, then $f(\vartheta) = \vartheta$, $\vartheta \in \Phi \cap \Psi$, and hence,

$$d(\theta, \vartheta) = d(f(\theta), f(\vartheta)) \precsim_{i_3} \lambda[d(\theta, f(\theta)) + d(f(\vartheta), \vartheta)] + \frac{\mu d(\theta, f(\theta)) d(f(\vartheta), \vartheta)}{1 + d(\theta, \vartheta)}.$$

Therefore $||d(\theta,\vartheta)|| = 0$ so that $\theta = \vartheta$. So f has a UFP. $\square$

The above Theorem 3.3 generalizes a Theorem 3.2 of [11].

Example 3.4. *Let Φ be the collection of all singleton subsets of $\mathbb{R}$ and Ψ be the collection of all compact subsets of $\mathbb{R}$. Let $d(\varrho, E) = |\varrho - \inf(E)| + i_3|\varrho - \sup(E)|$, where $(\varrho, E) \in \Phi \times \Psi$. Then (Φ, Ψ, d) is a complete TVBMS. Define a contravariant map $f : (\Phi, \Psi, d) \rightleftarrows (\Phi, \Psi, d)$ by $f(E) = \frac{\inf(E)+\sup(E)+6}{8}$, for all $E \in \Phi\cup\Psi$. Then, f satisfies the inequality $d(f(E), f(\varrho)) \precsim_{i_3} \lambda[d(\varrho, f(\varrho))+d(f(E), E)]+\frac{\mu d(\varrho,f(\varrho))d(f(E),E)}{1+d(\varrho,E)}$ for $\lambda = \frac{1}{3}$ and $\mu = 0$. By Theorem 3.3, f has a UFP $\{1\} \in \Phi \cap \Psi$.*

Theorem 3.5. *Let (Φ, Ψ, d) be a complete TVBMS with nonsingular $1 + d(\varrho, f(\varrho)) + d(f(\varpi), \varpi)$ and $||1 + d(\varrho, f(\varrho)) + d(f(\varpi), \varpi)|| \neq 0$, whenever $(\varrho, \varpi) \in \Phi \times \Psi$. If a contravariant map $f : (\Phi, \Psi, d) \rightleftarrows (\Phi, \Psi, d)$ satisfies $d(f(\varpi), f(\varrho)) \precsim_{i_3} \lambda[d(\varrho, \varpi) + d(\varrho, f(\varrho)) + d(f(\varpi), \varpi)] + \frac{\mu d(\varrho,f(\varrho))d(f(\varpi),\varpi)}{1+d(\varrho,f(\varrho))+d(f(\varpi),\varpi)}$, whenever $(\varrho, \varpi) \in \Phi \times \Psi$, for some $\lambda, \mu \in (0, 1)$ with $3\lambda + 2\mu < 1$. Then the function $f : \Phi \cup \Psi \to \Phi \cup \Psi$ has a UFP.*

Proof. Let $\varrho_0 \in \Phi$, $\varpi_0 = f(\varrho_0) \in \Psi$, and $\varrho_1 = f(\varpi_0)$. Suppose, $\varpi_n = f(\varrho_n)$ and $\varrho_{n+1} = f(\varpi_n)$, for all $n \in \mathbb{N}$. Then (ϱ_n, ϖ_n) is a bisequence on (Φ, Ψ, d). For all $n \in \mathbb{N}$, from

$$\begin{aligned}
d(\varrho_n, \varpi_n) &= d(f(\varpi_{n-1}), f(\varrho_n)) \\
&\precsim_{i_3} \lambda[d(\varrho_n, \varpi_{n-1}) + d(\varrho_n, f(\varrho_n)) + d(f(\varpi_{n-1}), \varpi_{n-1})] \\
&\quad + \frac{\mu d(\varrho_n, f(\varrho_n))d(f(\varpi_{n-1}), \varpi_{n-1})}{1 + d(\varrho_n, f(\varrho_n)) + d(f(\varpi_{n-1}), \varpi_{n-1})} \\
&= \lambda[d(\varrho_n, \varpi_{n-1}) + d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi_{n-1})] + \frac{\mu d(\varrho_n, \varpi_n)d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi_{n-1})} \\
||d(\varrho_n, \varpi_n)|| &\leq ||\lambda[d(\varrho_n, \varpi_{n-1}) + d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi_{n-1})] + \frac{\mu d(\varrho_n, \varpi_n)d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi_{n-1})}|| \\
&\leq \lambda||[d(\varrho_n, \varpi_{n-1}) + d(\varrho_n, \varpi_n) + d(\varrho_n, \varpi_{n-1})]|| + 2\mu||d(\varrho_n, \varpi_n)||
\end{aligned}$$

we conclude that

$$||d(\varrho_n, \varpi_n)|| \leq \frac{2\lambda}{1 - \lambda - 2\mu}||d(\varrho_n, \varpi_{n-1})||,$$

and

$$\begin{aligned}
d(\varrho_n, \varpi_{n-1}) &= d(f(\varpi_{n-1}), f(\varrho_{n-1})) \\
&\precsim_{i_2} \lambda[d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_{n-1}, f(\varrho_{n-1})) + d(f(\varpi_{n-1}), \varpi_{n-1})] \\
&\quad + \frac{\mu d(\varrho_{n-1}, f(\varrho_{n-1}))d(f(\varpi_{n-1}), \varpi_{n-1})}{1 + d(\varrho_{n-1}, f(\varrho_{n-1})) + d(f(\varpi_{n-1}), \varpi_{n-1})} \\
&= \lambda[d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_n, \varpi_{n-1})] \\
&\quad + \frac{\mu d(\varrho_{n-1}, \varpi_{n-1})d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_{n-1}, f(\varrho_{n-1})) + d(f(\varpi_{n-1}), \varpi_{n-1})}
\end{aligned}$$

$$\begin{aligned}
||d(\varrho_n, \varpi_{n-1})|| &\leq ||\lambda[d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_n, \varpi_{n-1})] \\
&\quad + \frac{\mu d(\varrho_{n-1}, \varpi_{n-1})d(\varrho_n, \varpi_{n-1})}{1 + d(\varrho_{n-1}, f(\varrho_{n-1})) + d(f(\varpi_{n-1}), \varpi_{n-1})}|| \\
&\leq \lambda||[d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_{n-1}, \varpi_{n-1}) + d(\varrho_n, \varpi_{n-1})]|| + 2\mu||d(\varrho_n, \varpi_{n-1})||
\end{aligned}$$

so that

$$||d(\varrho_n, \varpi_{n-1})|| \leq \frac{2\lambda}{1 - \lambda - 2\mu}||d(\varrho_{n-1}, \varpi_{n-1})||,$$

Therefore, by putting $\beta = \frac{2\lambda}{1-\lambda-2\mu}$, we have

$$||d(\varrho_n, \varpi_n)|| \leq \beta^{2n}||d(\varrho_0, \varpi_0)||$$

and

$$||d(\varrho_n, \varpi_{n-1})|| \leq \beta^{2n-1}||d(\varrho_0, \varpi_0)||.$$

For every $m, n \in \mathbb{N}$,

$$\begin{aligned} d(\varrho_n, \varpi_m) &\precsim_{i_3} d(\varrho_n, \varpi_n) + d(\varrho_{n+1}, \varpi_n) + d(\varrho_{n+1}, \varpi_m) \\ &\precsim_{i_3} (\beta^{2n} + \beta^{2n+1})d(\varrho_0, \varpi_0) + d(\varrho_{n+1}, \varpi_m) \\ &\precsim_{i_3} \ldots \\ &\precsim_{i_3} (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m-1})d(\varrho_0, \varpi_0) + d(\varrho_m, \varpi_m) \\ &\precsim_{i_3} (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m})d(\varrho_0, \varpi_0), \text{ if } m > n, \end{aligned}$$

$$||d(\varrho_n, \varpi_m)|| \leq (\beta^{2n} + \beta^{2n+1} + \ldots + \beta^{2m})||d(\varrho_0, \varpi_0)||, \text{ if } m > n,$$

and similarly, if $m < n$, then

$$d(\varrho_n, \varpi_m) \precsim_{i_3} (\beta^{2m+1} + \beta^{2m+2} + \ldots + \beta^{2n+1})d(\varrho_0, \varpi_0),$$

$$||d(\varrho_n, \varpi_m)|| \leq (\beta^{2m+1} + \beta^{2m+2} + \ldots + \beta^{2n+1})||d(\varrho_0, \varpi_0)||.$$

By $\beta \in (0, 1)$, $||d(\varrho_n, \varpi_m)|| \to 0$, as $n, m \to \infty$, we conclude that (ϱ_n, ϖ_n) is a Cauchy bisequence. Since (Φ, Ψ, d) is complete, (ϱ_n, ϖ_n) converges, and biconverges to a point $\theta \in \Phi \cap \Psi$. Hence, $f(\varrho_n) = \varpi_n \to \theta \in \Phi \cap \Psi$ as $n \to \infty$ implies $d(f(\theta), f(\varrho_n)) \to d(f(\theta), \theta)$ as $n \to \infty$, by using Lemma 2.8. Also by taking the limit from

$$d(f(\theta), f(\varrho_n)) \precsim_{i_3} \lambda[d(\varrho_n, \theta) + d(\varrho_n, \varpi_n) + d(f(\theta), \theta)] + \frac{\mu d(\varrho_n, \varpi_n)d(f(\theta), \theta)}{1 + d(\varrho_n, \varpi_n) + d(f(\theta), \theta)}$$

we obtain

$$||d(f(\theta), f(\varrho_n))|| \leq \lambda[||d(\varrho_n, \theta) + d(\varrho_n, \varpi_n) + d(f(\theta), \theta)||] + \frac{\mu||d(\varrho_n, \varpi_n)d(f(\theta), \theta)||}{||1 + d(\varrho_n, \varpi_n) + d(f(\theta), \theta)||},$$

as $n \to \infty$, we get $d(f(\theta), \theta) = 0$. Hence $f(\theta) = \theta$. Therefore θ is a fixed point of f.
If ϑ is another fixed point of f, then $f(\vartheta) = \vartheta$, $\vartheta \in \Phi \cap \Psi$, and hence,

$$d(\theta, \vartheta) = d(f(\theta), f(\vartheta)) \precsim_{i_3} \lambda[d(\theta, \vartheta) + d(\theta, f(\theta)) + d(f(\vartheta), \vartheta)] + \frac{\mu d(\theta, f(\theta))d(f(\vartheta), \vartheta)}{1 + d(\theta, f(\theta)) + d(f(\vartheta), \vartheta)}.$$

Therefore $||d(\theta, \vartheta)|| = 0$ so that $\theta = \vartheta$. So f has a UFP. □

The above Theorem 3.5 generalizes a Theorem 3.3 of [11].

4. CONCLUSIONS

All tricomplex valued bipolar metric space fixed point theorems are generalisations of tricomplex valued metric space fixed point theorems, which are generalisations of bicomplex valued metric spaces and complex valued metric spaces. Because complex valued metric spaces are generalisations of metric spaces, studies of fixed point results in tricomplex valued bipolar metric spaces are important.

Acknowledgement. The authors first like to thank the reviewers and the editor for their valuable comments and corrections for improvement of this article. The first author's study was funded by the Council of Scientific and Industrial Research (CSIR), India (File No:09/688(0031)/2018-EMR-I).

References

[1] A. Azam, B. Fisher and M. Khan, *Common fixed point theorems in complex valued metric spaces*, Numerical Functional Analysis and Optimization **32** (2011), 243-253.

[2] J. Choi, S. K. Datta, T. Biswas and N. Islam, *Some fixed point theorems in connection with two weakly compatible mappings in bicomplex valued metric spaces*, Honam Mathematical J. **39** (2017), 115-126.

[3] G. N. V. Kishore, K. P. R. Rao, H. Isik, B. S. Rao and A. Sombabu, *Covariant mappings and coupled fixed point results in bipolar metric spaces*, Int. J. Nonlinear Anal. Appl. **12** (2021), 1-15.

[4] G. N. V. Kishore, R. P. Agarwal, B. S. Rao and R. V. N. S. Rao, *Caristi type cyclic contraction and common fixed point theorems in bipolar metric spaces with applications*, Fixed Point Theory Appl. **21** (2018), 1-13.

[5] G. Mani, A. J. Gnanaprakasam, A. U. Haq, I. A. Baloch and F. Jarad, *Common fixed point theorems on tricomplex valued metric space*, Mathematical Problems in Engineering (2022), 12p.

[6] C. G. Moorthy and G. Siva, *Bipolar multiplicative metric spaces and fixed point theorems of covariant and contravariant mappings*, Mathematical Analysis and Convex Optimization **2** (2021), 1-11.

[7] A. Mutlu and U. Gurdal, *Bipolar metric spaces and some fixed point theorems*, J. Nonlinear Sci. Appl. **9** (2016), 5362-5373.

[8] A. Mutlu, K. Ozkan and U. Gurdal, *Coupled fixed point theorems on bipolar metric spaces*, Eur. J. Pure Appl. Math. **10** (2017), 655-667.

[9] G. B. Price, *An Introduction to Multicomplex Spaces and Functions*, Marcel Dekker, New York, NY, USA, 1991.

[10] C. Segre, *Le rappresentazioni reali delle forme complesse a gli enti iperalgebrici*, Math. Ann. **40** (1892), 413-467.

[11] G. Siva, *Bicomplex valued bipolar metric spaces and fixed point theorems*, Mathematical Analysis and its Contemporary Applications **4** (2022), 29-43.

[12] B. Srinuvasa Rao and G. N. V. Kishore, *Common fixed point theorems in bipolar metric spaces with applications to integral equations*, Int. J. Eng. Techn. **7** (2018), 1022-1026.

Department of Mathematics, Alagappa University, Karaikudi-630 003, India.
Email address: `gsivamaths2012@gmail.com`

Department of Computer Science and Engineering, Alagappa Chettiar Government College of Engineering and Technology, Karaikudi-630 003, India.
Email address: `rasikannancse@accetedu.in`