ISSN: 0025-5742

THE MATHEMATICS STUDENT

Volume 89, Nos. 1-2, January - June (2020)
(Issued: May, 2020)

PUBLISHED BY
THE INDIAN MATHEMATICAL SOCIETY
www.indianmathsociety.org.in

THE MATHEMATICS STUDENT

Edited by M. M. SHIKARE

In keeping with the current periodical policy, THE MATHEMATICS STUDENT will seek to publish material of interest not just to mathematicians with specialized interest but to the postgraduate students and teachers of mathematics in India. With this in view, it will ordinarily publish material of the following type:

1. research papers,
2. the texts (written in a way accessible to students) of the Presidential Addresses, the Plenary talks and the Award Lectures delivered at the Annual Conferences.
3. general survey articles, popular articles, expository papers and Book-Reviews.
4. problems and solutions of the problems,
5. new, clever proofs of theorems that graduate / undergraduate students might see in their course work, and
6. articles that arouse curiosity and interest for learning mathematics among readers and motivate them for doing mathematics.

Articles of the above type are invited for publication in THE MATHEMATICS STUDENT. Manuscripts intended for publication should be submitted online in the LaTeX and .pdf file including figures and tables to the Editor M. M. Shikare on E-mail: msindianmathsociety@gmail.com along with a Declaration form downloadable from our website.

Manuscripts (including bibliographies, tables, *etc.*) should be typed double spaced on A4 size paper with 1 inch (2.5 cm.) margins on all sides with font size 11 pt. in LaTeX. Sections should appear in the following order: Title Page, Abstract, Text, Notes and References. Comments or replies to previously published articles should also follow this format. In LaTeX the following preamble be used as is required by the Press:

\ documentclass[11 pt,a4paper,twoside,reqno]{amsart}
\ usepackage {amsfonts, amssymb, amscd, amsmath, enumerate, verbatim, calc}
\ renewcommand{\ baselinestretch}{1.2}
\ textwidth=12.5 cm
\ textheight=20 cm
\ topmargin=0.5 cm
\ oddsidemargin=1 cm
\ evensidemargin=1 cm
\ pagestyle{plain}

The details are available on Society's website: www.indianmathsociety.org.in

Authors of articles / research papers printed in the the Mathematics Student as well as in the Journal shall be entitled to receive a *soft copy (PDF file)* of the paper published. There are no page charges. However, if author(s) (whose paper is accepted for publication in any of the IMS periodicals) *is (are) unable to send the LaTeX file of the accepted paper, then a charge Rs. 100 (US $ 10) per page will be levied towards LaTeX typesetting.*

All business correspondence should be addressed to S. K. Nimbhorkar, Treasurer, Indian Mathematical Society, C/O Dr. Mrs. Prachi Kulkarni, Ankur Hospital, Tilaknagar, Aurangabad 431 001 (MS), India. E-mail: treasurerindianmathsociety@gmail.com or sknimbhorkar@gmail.com

In case of any query, one may contact the Editor through the e-mail.

ISSN: 0025-5742

THE MATHEMATICS STUDENT

Volume 89, Nos. 1-2, January - June (2020)
(Issued: May, 2020)

PUBLISHED BY
THE INDIAN MATHEMATICAL SOCIETY
www.indianmathsociety.org.in

ISSN: 0025-5742

Published by Prof. Satya Deo for the Indian Mathematical Society, type set by M. M. Shikare, "Krushnakali", Servey No. 73/6/1, Gulmohar Colony, Jagtap Patil Estate, Pimple Gurav, Pune 411061 and printed by Dinesh Barve at Parashuram Process, Shed No. 1246/3, S. No. 129/5/2, Dalviwadi Road, Barangani Mala, Wadgaon Dhayari, Pune-411 041 (India).
Printed in India.

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020)

CONTENTS

****

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 1–27

CAN WE EXPLICITLY DETERMINE THE STRUCTURE OF RATIONAL GROUP ALGEBRAS ?*

GURMEET K. BAKSHI

Abstract. Wedderburn-Artin theorem on the structure of semisimple finite dimensional algebras is regarded by many as the first major result in the theory of noncommutative rings and has remained important from the early 20th century to the present. A fundamental problem in group rings is to determine the complete set of primitive central idempotents and precise description of the Wedderburn decomposition of a given semisimple group algebra. The problem has been a subject of intensive research for almost a century due to its relation to various other problems in group rings. The last few years saw tremendous development on the subject. This is an updated survey on the modern approach for handling this problem starting with the classical method.

1. Introduction

Let $\mathbb{Q}G$ be the group algebra of a finite group G over the field $\mathbb{Q}$ of rationals. A celebrated theorem due to Mashcke states that $\mathbb{Q}G$ is a semisimple ring, i.e., each left ideal of $\mathbb{Q}G$ is its direct summand. Consequently, by the theory of semisimple rings, $\mathbb{Q}G$ is a direct sum of its minimal two sided ideals, which, indeed, turn out to be simple Artinian rings. The Wedderburn-Artin theorem on the structure of simple Artinian rings says that each of these minimal two sided ideals is isomorphic to a matrix ring over a finite dimensional division algebra over $\mathbb{Q}$. So, in essense, there is a ring isomorphism

$$\mathbb{Q}G \cong \oplus_{1\leq i\leq k} M_{n_i}(D_i), \tag{1}$$

* This article is based on the text of the 30th Hansraj Gupta Memorial Award Lecture delivered at the 85th Annual conference of the IMS - An International Meet held at IIT Kharagpur, W. B. during November 22-25, 2019

Research supported by Science and Engineering Research Board (SERB), DST, Govt. of India under the scheme Mathematical Research Impact Centric Support (sanction order no MTR/2019/001342) is gratefully acknowledged

where $n_i \geq 1$ is an integer and D_i is a finite dimensional division algebra containing $\mathbb{Q}$ in its center for all $1 \leq i \leq k$. Furthermore, the integer k and the pairs (n_1, D_1), (n_2, D_2), $\cdots$, (n_k, D_k) are uniquely determined upto permutation by the group G with D_i's unique upto isomorphism. This decomposition is called the Wedderburn decomposition of $\mathbb{Q}G$ and the matrix rings $M_{n_i}(D_i)$, $1 \leq i \leq k$, are called simple components of $\mathbb{Q}G$. Further, it is well known in the theory of semisimple rings that each simple component, which is a minimal two sided ideal of $\mathbb{Q}G$, is generated by a central idempotent. For $1 \leq i \leq k$, let e_i be the central idempotent of $\mathbb{Q}G$ which generates the simple component $M_{n_i}(D_i)$. The simple components being minimal two sided ideals imply that each of these e_i's can not be further written as sum of two mutually orthogonal central idempotents of $\mathbb{Q}G$; the idempotents with this property are called primitive central idempotents. The intersection of any two distinct simple components being zero, it follows that distinct e_i and e_j are orthogonal, i.e., $e_ie_j = 0$, if $i \neq j$. Further the isomorphism in eqn 1 gives that the sum of e_i's, $1 \leq i \leq k$, is 1. Thus there is a set $\{e_i \mid 1 \leq i \leq k\}$ of central idempotents of $\mathbb{Q}G$ with the following properties:

(i): $e_1 + e_2 + \cdots + e_k = 1$;
(ii): for all $1 \leq i, j \leq k$, $i \neq j$, $e_ie_j = 0 = e_je_i$;
(iii): for all $1 \leq i \leq k$, e_i can't be written as sum of two central idempotents .

The set $\{e_i \mid 1 \leq i \leq k\}$ of central idempotents of $\mathbb{Q}G$ satisfying properties (i)-(iii) is unique and is called the complete and irredundant set of primitive central idempotents of $\mathbb{Q}G$.

In practice, it is quite a hard problem to explicitly determine the complete algebraic structure, i.e., the complete and irredundant set of primitive central idempotents and the explicit Wedderburn decomposition of a given semisimple group algebra $\mathbb{Q}G$. Furthermore, an understanding of this problem is a tool to deal with several problems concerning group algebras. For instance, the study of the automorphism group of $\mathbb{Q}G$, the unit group of the integral group ring $\mathbb{Z}G$, the isomorphism problem in group algebra are all linked with deep understanding of the complete algebraic structure of $\mathbb{Q}G$ (see e.g. (see [1, 2, 14, 15, 16, 18, 21, 22, 24, 26, 27, 28, 31, 32, 33, 34, 35, 38, 42, 43]). Also, thorough understanding of the primitive central idempotents and the explicit Wedderburn decomposition of $\mathbb{Q}G$ leads to the character theory of G over $\mathbb{Q}$.

The classical method to find primitive central idempotents of $\mathbb{Q}G$ begins with the computation of the primitive central idempotent $e(\chi) = \frac{\chi(1)}{o(G)}\sum_{g\in G}\chi(g)g^{-1}$ of the complex group algebra $\mathbb{C}G$, followed by summing up all the primitive central idempotents of the form $e(\sigma\circ\chi)$ with $\sigma\in \mathrm{Gal}(\mathbb{Q}(\chi)/\mathbb{Q})$, where $\chi\in \mathrm{Irr}(G)$ (the complex irreducible characters of G), $\mathbb{Q}(\chi)$ is the field obtained by adjoining all $\chi(g)$ for $g\in G$ to $\mathbb{Q}$ and $\mathrm{Gal}(\mathbb{Q}(\chi)/\mathbb{Q})$ is the Galois group of $\mathbb{Q}(\chi)$ over $\mathbb{Q}$ (see [44] for details). The primitive central idempotent so obtained, i.e., $\sum_{\sigma\in \mathrm{Gal}(\mathbb{Q}(\chi)/\mathbb{Q})} e(\sigma\circ\chi)$ of $\mathbb{Q}G$, is commonly denoted by $e_{\mathbb{Q}}(\chi)$. This approach has computational difficulty because the information on the complete character table of G may not be known. Even if it is known, it is difficult to know $\mathrm{Gal}(\mathbb{Q}(\chi)/\mathbb{Q})$ for a given χ. Furthermore, having computed $e_{\mathbb{Q}}(\chi)$, it does not give information about the simple component $\mathbb{Q}Ge_{\mathbb{Q}}(\chi)$. It is important to mention that the knowledge of any primitive central idempotent e need not immediately yield the structure of simple component $\mathbb{Q}Ge$ or vice versa. It very well depends upon the approach used for the computations.

A more recent approach is to write the algebraic structure of $\mathbb{Q}G$ completely from the subgroup structure of G. Let us first ask a simple question. Can a subgroup of G enable us to write some idempotent of $\mathbb{Q}G$? The answer is yes, because if H is a subgroup of a finite group G, then $\hat{H} = \frac{1}{o(H)}\sum_{h\in H} h$ is an idempotent of $\mathbb{Q}G$, called the idempotent determined by H. Here is the next question in sequel. Can such idempotents help us to understand arbitrary primitive central idempotents of $\mathbb{Q}G$? In this direction, it is known that if G is an abelian group, then any primitive central idempotent of of $\mathbb{Q}G$ is an integral linear combination of $\hat{H}$, where H runs through subgroups of G ([41], Proposition VI.1.16). However, the explicit description of the coefficient needs to given. For an abelian group, this type of description was given by Jespers, Leal and Milies ([30], also see Chapter VII of [20]). They gave explicit expressions of the primitive central idempotents completely in terms of G, from which the Wedderburn decomposition, which was earlier given by Perlis and Walker[40], follows as a consequence. In [29], Jespers, Leal and Paques gave a very precise fully internal description of the primitive central idempotents when G is a nilpotent group. The description is only determined by a lattice of subgroups without making use of the characters of G. As a consequence, they deduced that any primitive central idempotent of the rational group algebra of a nilpotent group is a $\mathbb{Z}$ linear combination of $\hat{H}$, where H runs through subgroups of G. This result gave a forceful blow for moving further. Soon Olivieri, del Río and Simon [39] realized that it is the monomality of the characters of nilpotent groups which is playing a crucial role in [29]. So, conceptualizing this idea, they tried to understand the expression of $e_{\mathbb{Q}}(\chi)$ completely using the subgroup structure of G, for the case when χ is a monomial character. They found that the calculations of $e_{\mathbb{Q}}(\chi)$ for a monomial character χ depends upon a specific pairs of subgroups of G, which they termed as Shoda pairs of G. Since that time, the method using Shoda pairs has been in the middle of active research that stimulated several mathematicians and tremedous progress has been made.

During the last decade, the author in various collaborative works has taken this study further in a series of papers which appeared in [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]. The progress is a far reaching generalization of the earlier work and also contributes to the understanding of the $\mathbb{Q}$-characters of monomial groups. The primitive central idempotents of the rational group algebra of some special classes of p-groups has been done in [4]. The beautiful work of Olivieri, del Río and Simon [39] reveals that the Shoda pairs of G yield primitive central idempotents of $\mathbb{Q}G$. However, to know the precise Wedderburn decomposition, one needs to know, when do two different Shoda pairs yield the same primitive central idempotent? To this end, the author with Kulkarni and Passi in [3] gave precise algorithm for metabelian groups, thus yielding the explicit Wedderburn decomposition in this case. In [13], jointly with Maheshwary, the complete algebraic structure of the rational group algebra of a normally monomial group was provided [13], which extended the work of [3]. This information has resulted in a significant impact in understanding normally monomial groups as well. Consequently, a search for normally monomial groups among the groups in `GAP` library of small groups, done in [5], indicates that this is a substantial class of monomial groups. Abstracting the theory developed in [3, 13], recently the author with Kaur developed an efficient technique of providing the algebraic structure of the rational group algebra of a large class of monomial groups [10]. The main ingredient in the theory has been

Isaac's notion of character triples and Clifford's correspondence theorem. The technique involves the construction of certain rooted directed trees whose particular leaves yield the desired structure of the group algebra. In [11], jointly with Kaur, the notion of generalized strongly monomial groups has been introduced, which is a natural but non-trivial generalization of the concept of strongly monomial groups introduced by Olivieri, del Río and Simon in [39]. In that paper, an answer to the problem under consideration has been given for a generalized strongly monomial group. It is proved in [11] that the class of generalized strongly monomial groups subsumes several well known classes of monomial groups invesigated by leading group theorists. Infact, at present, it is not known whether there exists a monomial group which does not lie in the class of generalized strongly monomial groups and looks challenging to find any such example.

The work on Shoda pairs, strong Shoda pairs done by Olivieri, del Río and Simon in [39] and later on extremely strong Shoda pairs by Bakshi, Maheshwary in [13] has been implemented in GAP package Wedderga [6] that features functions to compute strong Shoda pairs, extremely strong Shoda pairs, primitive central idempotents realized by strong (resp. extremely strong) Shoda pairs, information on Wedderburn Decomposition etc. Wedderga also includes functions to check whether a given group is strongly monomial/normally monomial or not. Some of the algorithms used in the implementation have appeared in [5] and [37].

Beyond monomial groups, there are only a few known results that describe the algebraic structure of their rational group algebras. This has been done by Giambruno-Jespers [19], Jespers-Olteanu-del Río [36], and Janssens [23]. In [3] and [4], an alternative approach has been used to write expressions of $e_{\mathbb{Q}}(\chi)$ for an irreducible character χ of G in terms of the Euler function φ and Mobius function μ as an explicit group ring element. Initially in [4], a constaint on χ was imposed and later in [3], the result was obtained for an arbitrary $\chi \in \mathrm{Irr}(G)$.

The purpose of this article is to provide an exhaustive survey of the above development.

Going little beyond the purview of the title, it is pertinent to mention that Shoda pair theory has also been used to describe the algebraic structure of semisimple finite group algebras, which has applications in coding theory, and the interested reader is referred to [7], [8], [9], [12] and [17] for the development. Furthermore, the applications of Shoda pair theory in understanding the unit group of integral group rings can be seen in [11, 14, 15, 16, 33, 34].

2. Abelian groups

We begin with the case when $G := C_n$, the cyclic group of order n. Let $\zeta_n = e^{2\pi i/n}$ be a complex primitive nth root of unity. For a divisor d of n, let

$$\phi_d(x) = \prod_{\substack{1\leq i\leq n \\ \gcd(i,n)=d}} \zeta_n^i.$$

It is know that the cyclotomic polynomial $\phi_d(x)$ has integral coefficients and it is irreducible over $\mathbb{Q}$. Furthermore, $x^n - 1 = \prod_{d|n} \phi_d(x)$ is the factorization of $x^n - 1$ into irreducible factors, and using Chinese remainder theorem one obtains the following isomorphism:

$$\mathbb{Q}C_n \cong \mathbb{Q}[x]/\langle x^n - 1\rangle \cong \oplus_{d|n}\mathbb{Q}[x]/\langle \phi_d(x)\rangle \cong \oplus_{d|n}\mathbb{Q}(\zeta_d).$$

This is precisely the Wedderburn decomposition of QC_n. To illustrate, consider $G := C_{10}$, the cyclic group of order 10. In this case, $x^{10} - 1 = \phi_1(x)\phi_2(x)\phi_5(x)\phi_{10}(x)$ and hence $\mathbb{Q}C_{10} \cong \mathbb{Q} \oplus \mathbb{Q}(\zeta_2) \oplus \mathbb{Q}(\zeta_5) \oplus \mathbb{Q}(\zeta_{10})$. As $\zeta_2 = -1$, $\mathbb{Q}(\zeta_2) = \mathbb{Q}$ and so $\mathbb{Q}C_{10} \cong \mathbb{Q} \oplus \mathbb{Q} \oplus \mathbb{Q}(\zeta_5) \oplus \mathbb{Q}(\zeta_{10})$. the desired Wedderburn decomposition of $\mathbb{Q}C_{10}$.

Generalizing this approach to abelian groups, Perlis and Walker [40] proved the following:

Theorem 1. ([25], Theorem 3.3.6) *If G is an abelian group of order n, then*

$$\mathbb{Q}G \cong \oplus_{d|n}\ a_d\mathbb{Q}(\zeta_d)$$

where a_d is the number of cyclic subgroups of G of order d.

Let us apply Perlis and Walker's theorem to the group $G := C_2 \times C_4$. The number of subgroups of order 1, 2 and 4 are 1, 3, 2 respectively. Hence in the Wedderburn decomposition of $\mathbb{Q}G$, there is one copy of $\mathbb{Q}$, three copies of $\mathbb{Q}(\zeta_2)$ and two copies of $\mathbb{Q}(\zeta_4)$. As $\zeta_2 = -1$ and $\zeta_4 = i = \sqrt{-1}$, we obtain $\mathbb{Q}G \cong \mathbb{Q} \oplus \mathbb{Q} \oplus \mathbb{Q} \oplus \mathbb{Q} \oplus \mathbb{Q}(i) \oplus \mathbb{Q}(i)$.

Having understood the Wedderburn decomposition of $\mathbb{Q}G$ when G is abelian, let us move to the progress made for understanding the primitive central idempotents in this case.

If G is a cyclic group of prime power order, Jespers, Leal amd Milies [30] proved the following:

Theorem 2. ([20], Chapter VII, Lemma 1.2) *Let $G =< x >$ be a cyclic group of order p^m, $m \geq 1$, p a prime. Let*

$$G = G_0 \supseteq G_1 \supseteq ... \supseteq G_m = 1$$

be the descending chain of all the subgroups of G, thus $G_i =< x^{p^i} >$. Then the primitive central idempotents of $\mathbb{Q}G$ are

$$e_0 = \hat{G} \text{ and } e_i = \hat{G}_i - \hat{G}_{i-1},\ 1 \leq i \leq m.$$

Furthermore, $\mathbb{Q}Ge_i \cong \mathbb{Q}(\zeta_{p^i})$, where ζ_{p^i} denotes a primitive p^ith root of unity.

The following is the generalization of the above theorem to abelian groups:

Theorem 3. ([20], Chapter VII, Theorem 1.4)

- **(i):** *Let p be a prime and G a finite abelian p-group. If H is a subgroup of G such that G/H is cyclic and H^*/H be the unique minimal subgroup of G/H, then*

$$e_H = \widehat{H} - \widehat{H}^*$$

  is a primitive central idempotent of $\mathbb{Q}G$ with $\mathbb{Q}Ge_H \cong \mathbb{Q}(\zeta_d)$, where ζ_d is primitive dth root of unity and $d = o(G/H)$. Furthermore e_H as H runs through all the subgroups of G with cyclic quotient group is a complete and irredundant list of primitive central idempotents of $\mathbb{Q}G$.
- **(ii):** *Let G be a finite abelian group. Write $G = G_1 \times G_2 \times \cdots \times G_l$ as the direct product of finite p_i-groups G_i for distinct primes $p_1, p_2, \cdots p_l$. Then the primitive central idempotents of $\mathbb{Q}G$ are of the form $e_1e_2, \cdots e_l$, where each e_i is a primitive central idempotent of $\mathbb{Q}G_i$. Furthermore,*
  - **(a):** *$(\mathbb{Q}G)e_1e_2, \cdots e_l \cong \mathbb{Q}(\zeta_d)$ for some d with $d|o(G)$;*
  - **(b):** *The number of primitive central idempotents e with $(\mathbb{Q}G)e \cong \mathbb{Q}(\zeta_d)$ equals the number of subgroups H of G with G/H cyclic of order d.*

Clearly Theorem 1, which is due to Perlis and Walker, follows as a consequence of the above theorem. A simpler description with a much shorter proof of the above theorem was later given by Jespers, Leal and Paques in [29]. In order to state it, we need to first introduce a notation. For $K \trianglelefteq H \leq G$, with H/K cyclic, define $\varepsilon(H, K) = \hat{H}$ if $H = K$, otherwise

$$\varepsilon(H, K) = \prod(\hat{K} - \hat{M}),$$

where M runs over all normal subgroups of H minimal with respect to the property of containing K properly. Given $\alpha = \sum_{g \in G} \alpha_g g$ in $\mathbb{Q}G$, $\mathrm{supp}(e)$ is the set of all those $g \in G$ such that $\alpha_g \neq 0$.

Theorem 4. ([29], Corollary 2.1) *Let G be an finite abelian group. The primitive central idempotents of $\mathbb{Q}G$ are precisely of the form $\varepsilon(G, N)$, with N a subgroup of G so that G/N is cyclic. In particular, if e is a primitive central idempotent of $\mathbb{Q}G$, then* $\mathrm{supp}(e)$ *is subgroup G and e is a $\mathbb{Z}$ linear combination of idempotents of the form $\hat{H}$, where H is a subgroup of G.*

After this stage, it is natural to go further to understand the primitive central idempotent $e_{\mathbb{Q}}(\chi)$ and the corresponding simple component $\mathbb{Q}Ge_{\mathbb{Q}}(\chi)$ when χ is a linear character of an arbitrary finite group G. Olivieri. del Río and Simon [39] proved the following:

Theorem 5. ([39], Lemma 1.2) *If χ is a linear character of a finite group G and $K = \ker \chi$, then $e_{\mathbb{Q}}(\chi) = \varepsilon(G, K)$ and $\mathbb{Q}Ge_{\mathbb{Q}}(\chi) \cong \mathbb{Q}(\zeta_d)$, where $d = o(G/K)$.*

3. Moving from abelian groups to nilpotent groups

The main objective of Jespers, Leal and Paques, while proving Theorem 4, was to provide a fully internal desciption of the primitive central idempotents for nilpotent G as stated below:

Theorem 6. ([29], Theorem 2.1) *Let G be a finite nilpotent group. The primitive idempotents of $\mathbb{Q}G$ are precisely all the elements of the form*

$$\sum(\varepsilon(G_m, H_m))^g$$

where the sum is over all G-conjugates of $\varepsilon(G_m, H_m)$, $(\varepsilon(G_m, H_m))^g = g^{-1}(\varepsilon(G_m, H_m))g$, H_m and G_m are subgroups of G satisfying the following properties:

(i): $H_0 \leq H_1 \leq \cdots \leq H_m \leq G_m \leq G_{m-1} \leq \cdots \leq G_0 = G$;

(ii): *for* $0 \leq i \leq m$, $H_i \trianglelefteq G_i$, $\mathcal{Z}(G_i/H_i)$, *the center of* G_i/H_i, *is cyclic;*
(iii): *for* $0 \leq i < m$, G_i/H_i *is not abelian and* G_m/H_m *is abelian;*
(iv): *for* $0 \leq i < m$, $G_{i+1}/H_i = \mathrm{Cen}_{G_i/H_i}(\mathcal{Z}_2(G_i/H_i))$, *where* $\mathcal{Z}_2(G_i/H_i)$ *is the 2nd center of* G_i/H_i*;*
(v): *for* $1 \leq i \leq m$, $\bigcap_{x \in G_{i-1}/H_{i-1}} H_i^x = H_{i-1}$, *where* $H^g = g^{-1}Hg$.

In Corollary 2.2 of [29], it is also shown as a consequence of the above theorem that any primitive central idempotent of the rational group algebra of a nilpotent group is a $\mathbb{Z}$ linear combination of $\hat{H}$, where H is a subgroup of G.

4. From nilpotent groups to arbitrary monomial groups: Shoda pair theory

A character χ of a group G is called monomial if it is induced from a linear character of a subgroup of G. It may be recalled that by a character of G, we always mean a complex character. The groups where all irreducible characters are monomial are called monomial groups. Some well known examples of monomial groups are nilpotent groups, metabelian groups, supersolvable groups, more generally abelian-by-supersolvable groups. The class of monomial groups is quite rich because it is well known that every monomial group is solvable and every solvable group can be embedded in a monomial group.

Olivieri. del Río and Simon [38] looked deep into the work in [29] and realized that it the monomality of nilpotent groups which is playing a key role there. Hence they tried to push the results to monomial groups by understanding $e_{\mathbb{Q}}(\chi)$ when χ is a monomial character of G.

Let χ be a monomial irreducible character of G. Suppose $\chi = \lambda^G$, where λ is a linear character of a subgroup H of G. If $K = \ker \chi$, then the pair (H, K) has the following properties:
(i) $K \trianglelefteq H$, H/K is cyclic and
(ii) If $g \in G$ and $[H, g] \cap H \subseteq K$ then $g \in H$.
The condition (i) is obvious and (ii) follows from Shoda's result (see [25], Corollary 3.2.3). Conversely if (H, K) is a pair of subgroups of G that satisfy (i) and (ii) and λ is any linear character on H with kernel K, then by Shoda's theorem λ^G is irreducible. Olivieri, del Río and Simon ([39],

Definition 1.4) called a pair (H, K) of G that satisfy conditions (i) and (ii) stated above as a Shoda pair of G. As desribed above, the advantage of Shoda pairs of G is that they give us monomial irreducible characters of G and conversely, every irreducible monomial character of G arises from a Shoda pair of G.

The following result describes the primitive central idempotents of $\mathbb{Q}G$ corresponding to a monomial irreducible characters of G and is a crucial result proved in [39]:

Theorem 7. ([39], Theorem 2.1) *Suppose χ is a monomial irreducible character of an arbitrary finite group G and suppose that χ is induced from a linear character λ on subgroup H with kernel K. Then*

$$e_{\mathbb{Q}}(\chi) = \alpha e(G, H, K),$$

where $\alpha = \frac{[\mathrm{Cen}_G(\varepsilon(H,K)):H]}{[\mathbb{Q}(\lambda):\mathbb{Q}(\chi)]}$ and $e(G, H, K)$ is the sum of distinct G-conjugates of $\epsilon(H, K)$.

The above theorem, in other words, says that the primitive central idempotents of $\mathbb{Q}G$ are realized by Shoda pairs of G (or Shoda pairs of G yield primitive central idempotents of $\mathbb{Q}G$). Hence, an alternative way to state it is as follows:

Theorem 8. ([39], Corollary 2.2) *If (H, K) is a Shoda pair of G, then a certain rational multiple of $e(G, H, K)$ is a primitive central idempotents of $\mathbb{Q}G$. Furthermore, the rational multiple is uniquely determined by (H, K).*

For monomial groups, all irreducible characters arise from Shoda pairs of G and therefore the following corollory follows:

Corollary 1. ([39], Corollary 2.3) *A finite group G is a monomial if and only if every primitive central idempotent of $\mathbb{Q}G$ is of the type $\alpha e(G, H, K)$ for $\alpha \in \mathbb{Q}$ and a Shoda pair (H, K) of G.*

The next question, which is of concern, is the structure of the simple component $\mathbb{Q}G\alpha e(G, H, K)$. This is a difficult question in general. However by imposing a constraint on Shoda pairs, Olivieri, del Río and Simon in [39] described this simple component very explicitly. Shoda pairs with the so called added constarint are termed as strong Shoda pairs and are defined as follows:

Definition 1. Strong Shoda Pair ([39], Definition 3.1, Proposition 3.3) *A Shoda pair* (H, K) *of* G *is called a strong Shoda pair if*
(i) $H \trianglelefteq N_G(K)$
(ii) $\epsilon(H, K)\epsilon(H, K)^g = 0$ *for all* $g \in G\backslash N_G(K)$, *where* $\epsilon(H, K)^g = g^{-1}\epsilon(H, K)g$.

Definition 2. Strongly monomial group ([25], p. 244) *A group* G *where every primitive central idempotents of* $\mathbb{Q}G$ *is realized by a strong Shoda pair of* G *is called strongly monomial group.*

The following result provides examples of strongly monomial groups.

Theorem 9. ([39], Theorem 4.4) *All abelian-by-supersolvable groups are strongly monomial. In particular, all metabelian groups are strongly monomial.*

For a ring R, denote by $R*^{\sigma}_{\tau}G$, the crossed product of the group G over the ring R with action σ and twisting τ (see [25], Chapter 2 for details on crossed product). The advantage of strong Shoda pairs lies in the following theorem:

Theorem 10. ([39], Proposition 3.4) *Let* (H, K) *be an strong Shoda pair of* G *and let* $k = [H : K]$. *Denote* $N_G(K)$ *by* N *and let* $n = [G : N]$. *Let* x *be a generator of* H/K *and* $\phi : N/H \mapsto N/K$ *a left inverse of the projection* $N/K \mapsto N/H$. *Then*

(i) $e(G, H, K)$ *is a primitive central idempotent of* $\mathbb{Q}G$, *i.e.,* $\alpha = 1$;

(ii)
$$\mathbb{Q}Ge(G, H, K) \cong M_n(\mathbb{Q}(\zeta_k) *^{\sigma}_{\tau} N/H),$$
where the action σ *and the twisting* τ *are given by* $\zeta_k^{\sigma(a)} = \zeta_k^i$, *if* $x^{\phi(a)} = x^i$; $\tau(a, b) = \zeta_k^j$, *if* $\phi(ab)^{-1}\phi(a)\phi(b) = x^j$, *for* $a, b \in N/H$ *and integers* i *and* j.

In the next theorem, we state a generalization of Theorem 10 for a larger class of Shoda pairs. The author in a joint work with Kaur gave, in [11], a natural but non trivial generalization of the concept of strong Shoda pairs, and called them generalized strong Shoda pairs. A little prepartion needs to be done in order to define them. First of all, the notation of $e(G, H, K)$ has been generalized as follows. Suppose $H = H_0 \leq H_1 \leq \cdots \leq H_n = G$ is an arbitrary chain of subgroups of G, set $\varepsilon^{(0)}(H, K) = \varepsilon(H, K)$ and $\varepsilon^{(i)}(H, K) =$ the sum of all the distinct H_i-conjugates of $\varepsilon^{(i-1)}(H, K)$ for

$1 \leq i \leq n$. Finally, denote the last step $\varepsilon^{(n)}(H,K)$ by $\mathfrak{e}(G,H,K)$. The reader should be warned about the distinction between $e(G,H,K)$ and $\mathfrak{e}(G,H,K)$. At this stage, the definition of $\mathfrak{e}(G,H,K)$ very well depends upon the subgroups H_i for $0 < i < n$. However, we will later mention that this is not the case for generalized strong Shoda pairs and special kind of chain $H = H_0 \leq H_1 \leq \cdots \leq H_n = G$ of subgroups of G (so called strong inductive chain from H to G).

Definition 3. Generalized strong Shoda pair ([11], p.422) *A Shoda pair (H,K) of G is called a generalized strong Shoda pair of G if there is a chain $H = H_0 \leq H_1 \leq \cdots \leq H_n = G$ of subgroups of G (called strong inductive chain from H to G) such that the following conditions hold for all $0 \leq i < n$:*

- **(i):** $H_i \trianglelefteq \mathrm{Cen}_{H_{i+1}}(\varepsilon^{(i)}(H,K))$;
- **(ii):** *the distinct H_{i+1}-conjugates of $\varepsilon^{(i)}(H,K)$ are mutually orthogonal.*

As the name suggests, a strong Shoda pair must be a generalized strong Shoda pair of G. This follows because for a strong Shoda pair (H,K) of G, $N_G(K)$ equals $\mathrm{Cen}_G(\varepsilon(H,K))$, by Lemma 3.2 of [39] and so the strong inductive chain '$H \leq G'$ works. Also one needs to see that the definition of a generalized strong Shoda pair given above is equivalent to the one given in [11]. For, if λ is a linear character of H with kernel K, then by Lemma 3 of [11], $\varepsilon^{(i)}(H,K)$ is a rational multiple of $e_{\mathbb{Q}}(\lambda^{H_i})$, and hence the centralizers of $\varepsilon^{(i)}(H,K)$ and $e_{\mathbb{Q}}(\lambda^{H_i})$ in H_{i+1} coincide. So the equivalence of the above definition of a generalized strong Shoda pair of G with that given in [11] follows. The following result follows by the repeated application of Lemma 3 of [11]:

Theorem 11. *If (H,K) is a generalized strong Shoda pair of G and $H = H_0 \leq H_1 \leq \cdots \leq H_n = G$ is a strong inductive chain from H to G, then $e_{\mathbb{Q}}(\lambda^{H_i})$, where λ is a linear character of H with kernel K, becomes equal to $\varepsilon^{(i)}(H,K)$ for all $0 \leq i \leq n$.*

Hence the following result follows:

Theorem 12. *If (H,K) is a generalized strong Shoda pair of G and λ is any linear character on H with kernel K, then*

$$e_{\mathbb{Q}}(\lambda^G) = \mathfrak{e}(G,H,K).$$

Hence, for any generalized strong Shoda pair (H, K) of G, $\mathfrak{e}(G, H, K)$ is a primitive central idempotent of $\mathbb{Q}G$

Remark 1. *The fact that $\mathfrak{e}(G, H, K)$ is equal to $e_{\mathbb{Q}}(\lambda^G)$ reveals that the construction of $\mathfrak{e}(G, H, K)$ is independent of the strong inductive chain $H = H_0 \leq H_1 \leq \cdots \leq H_n = G$. From now onwards, $\mathfrak{e}(G, H, K)$ appears only when (H, K) is a generalized strong Shoda pair of G and its computation is done using some strong inductive chain from H to G.*

Theorem 12 says that each generalized strong Shoda pair of G realizes in a character free way a primitive central idempotent of $\mathbb{Q}G$, namely $\mathfrak{e}(G, H, K)$. Furthermore, in [11], a description of simple component $\mathbb{Q}G\mathfrak{e}(G, H, K)$ for a generalized strong Shoda pair (H, K) of G is given. To be able to state that result, we recall the following notion of strong inductive source:

Definition 4. Strong Inductive source ([11], p. 422) *Let S be a subgroup of G and let $\psi \in \mathrm{Irr}(S)$. We say that ψ is a strong inductive source of G if the following conditions hold:*

(i): *ψ^G is irreducible;*
(ii): *$S \trianglelefteq \mathrm{Cen}_G(e_{\mathbb{Q}}(\psi))$;*
(iii): *the distinct G-conjugates of $e_{\mathbb{Q}}(\psi)$ are mutually orthogonal.*

Suppose (H, K) is a generalized strong Shoda pair of G, λ a linear character on H with kernel K, and $H = H_0 \leq H_1 \leq \cdots \leq H_n = G$ a strong inductive chain from H to G. Then, it is clear that λ^{H_i} is a strong inductive source of H_{i+1} for all $0 \leq i < n$. The description of the simple component $\mathbb{Q}Ge_{\mathbb{Q}}(\lambda^G)$ is obtained by recursively applying the next theorem that descibes the simple component $e_{\mathbb{Q}}(\psi^G)$, when ψ is a strong inductive source of G. For notational convenience, let us denote $\mathrm{Cen}_G(e_{\mathbb{Q}}(\psi))$ by C. For each $x \in C/S$, let $\overline{x} \in C$ be a fixed inverse image of x under the natural map $C \to C/S$. Since $\overline{x}$ centralizes $e_{\mathbb{Q}}(\psi)$, we have $\overline{x}\mathbb{Q}Se_{\mathbb{Q}}(\psi)\overline{x}^{-1} = \mathbb{Q}Se_{\mathbb{Q}}(\psi)$. Consequently, there is a map $\sigma_S : C/S \to \mathrm{Aut}(\mathbb{Q}Se_{\mathbb{Q}}(\psi))$, which maps x to the conjugation automorphism $(\sigma_S)_x$ on $\mathbb{Q}Se_{\mathbb{Q}}(\psi)$ induced by $\overline{x}$. Also $\overline{x}r = (\overline{x}r\overline{x}^{-1})\overline{x} = (\sigma_S)_x(r)\overline{x}$ for $x \in C/S$ and $r \in \mathbb{Q}Se_{\mathbb{Q}}(\psi)$, and an action occurs. Furthermore for $x, y \in C/S$, $S\overline{x}S\overline{y} = S\overline{xy}$ and so $\overline{x}.\overline{y} = s\overline{xy}$, where $s \in S$. Thus there is a map $\tau_S : C/S \times C/S \to \mathcal{U}(\mathbb{Q}Se_{\mathbb{Q}}(\psi))$ such that $\tau_S(x, y) = se_{\mathbb{Q}}(\psi) \in \mathcal{U}(\mathbb{Q}Se_{\mathbb{Q}}(\psi))$. Finally one can check that $\mathbb{Q}Ce_{\mathbb{Q}}(\psi) = \mathbb{Q}Se_{\mathbb{Q}}(\psi) *_{\tau_S}^{\sigma_S} C/S$.

Proposition 1. ([11], Proposition 2) *Let* $\psi \in \mathrm{Irr}(S)$ *be a strong inductive source of* G *and let* $T = \{t_i \mid 1 \leq i \leq m\}$ *be a right transversal of* C *in* G, *where* $C = \mathrm{Cen}_G(e_{\mathbb{Q}}(\psi))$. *Then the following hold:*

- **(i):** $\mathbb{Q}Ge_{\mathbb{Q}}(\psi^G)$ *is isomorphic to* $M_m(\mathbb{Q}Ce_{\mathbb{Q}}(\psi))$ *and the map which defines this isomorphism is given by* $\alpha \mapsto (\alpha_{ij})_{m\times m}$, *where* $\alpha_{ij} = e_{\mathbb{Q}}(\psi)t_j\alpha t_i^{-1}e_{\mathbb{Q}}(\psi)$;
- **(ii):** *the map* $\alpha \mapsto \sum_{t\in T}\alpha^t$ *defines an isomorphism from* $\mathcal{Z}(\mathbb{Q}Ce_{\mathbb{Q}}(\psi))$, *the center of* $\mathbb{Q}Ce_{\mathbb{Q}}(\psi)$, *to* $\mathcal{Z}(\mathbb{Q}Ge_{\mathbb{Q}}(\psi^G))$;
- **(iii):** C/S *acts on* $\mathcal{Z}(\mathbb{Q}Se_{\mathbb{Q}}(\psi))$ *by conjugation and* $\mathcal{Z}(\mathbb{Q}Se_{\mathbb{Q}}(\psi))^{C/S} = \{r \in \mathcal{Z}(\mathbb{Q}Se_{\mathbb{Q}}(\psi)) \mid (\sigma_S)_x(r) = r\ \forall x \in C/S\}$*is a subring of* $\mathcal{Z}(\mathbb{Q}Ce_{\mathbb{Q}}(\psi))$.

The repeated application of the above result yields the following:

Theorem 13. ([11], Theorem 3) *Let* (H,K) *be a generalized strong Shoda pair of* G *and* λ *be a linear character on* H *with kernel* K. *Let* $H = H_0 \leq H_1 \leq \cdots \leq H_n = G$ *be a strong inductive chain from* H *to* G, *then* $\mathbb{Q}Ge_{\mathbb{Q}}(\lambda^G)$,*i.e.,* $\mathbb{Q}G\mathfrak{e}(G,H,K)$ *is isomorphic to*
$M_{k_n}(M_{k_{n-1}}\cdots(M_{k_1}(\mathbb{Q}He_{\mathbb{Q}}(\lambda) \ *^{\sigma_{H_0}}_{\tau_{H_0}} \ C_0/H_0) \ *^{\sigma_{H_1}}_{\tau_{H_1}} \ C_1/H_1) \ *^{\sigma_{H_2}}_{\tau_{H_2}} \cdots *^{\sigma_{H_{n-1}}}_{\tau_{H_{n-1}}} C_{n-1}/H_{n-1})$,
where C_i, σ_{H_i}, τ_{H_i}, k_i *are as defined above.*

Definition 5. Generalized strongly monomial group ([11], p. 423) *A group* G *is called generalized strongly monomial if every primitive central idempotent of* $\mathbb{Q}G$ *is realized by a generalized strong Shoda pair of* G.

Theorems 11 and 13 tell us completely the primitive central idempotents and the corresponding simple components of $\mathbb{Q}G$, when G is a generalized strongly monomial group.

The next question that concern us is how big is the class of generalized strongly monomial groups. Since a strong Shoda pair of G is a generalized strong Shoda pair of G, the following result follows:

Theorem 14. ([11], p.423) *All strongly monomial groups are generalized strongly monomial.*

Recall that a group G is called subnormally monomial if every complex irreducible character of G is induced from a linear character of a subnormal subgroup of G. It is easy to see that a Shoda pair (H,K) of G with H

subnormal in G is a generalized strong Shoda pair of G. Consequently, the following holds:

Theorem 15. ([11], p.423) *All subnormally monomial groups are generalized strongly monomial.*

We now move to provide some more examples of subnormally monomial groups. Let $\mathcal{C}$ be the class of all finite groups G such that every subquotient, i.e., quotient group of subgroup of G, satisfy the following property that either it is abelian or it contains a non central abelian normal subgroup. It is proved in Proposition 1 of [10] that all subnormally monomial groups lie in $\mathcal{C}$. In [11], the following is proved:

Theorem 16. ([11], Theorem 1) *The groups in class $\mathcal{C}$ are generalized strongly monomial.*

The next thorem provides numerous examples of groups in $\mathcal{C}$ showing that the generalized strongly monomial groups are in abundance.

Theorem 17. ([11], Theorem 1) *The following groups are in class $\mathcal{C}$:*

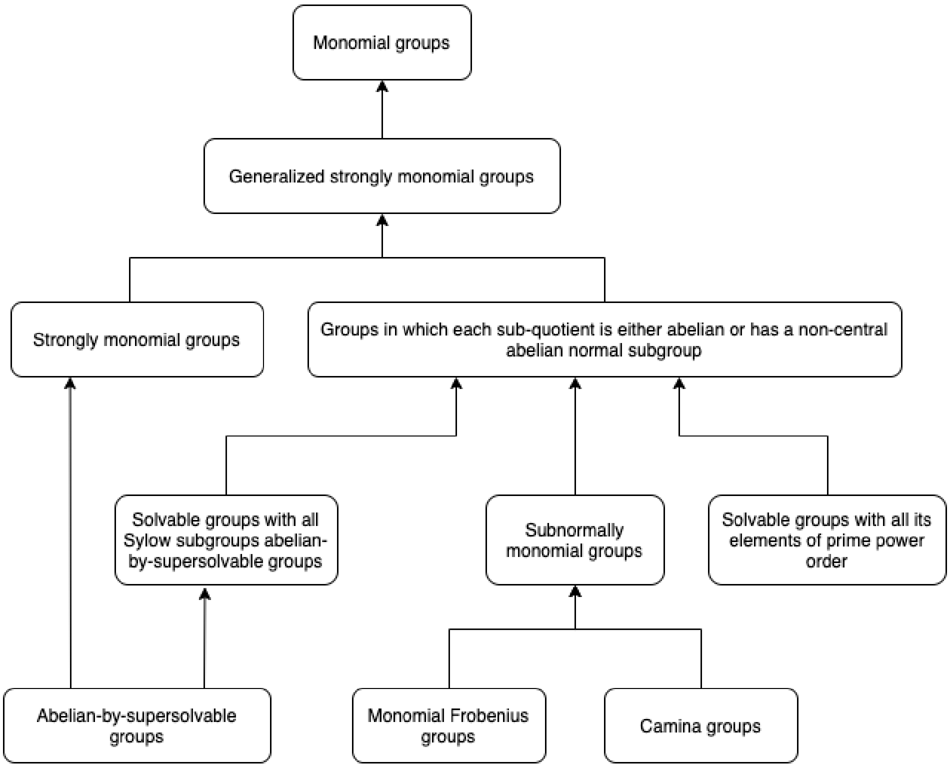
(i) *monomial Frobenius groups;*
(ii) *Camina groups;*
(iii) *groups such that all its non linear irreducible characters vanish only on the elements of order p, where p is a fixed prime;*
(iv) *solvable groups with the property that all its non linear irreducible characters of the same degree are Galois conjugate (this property holds in particular when all its non linear irreducible characters have distinct degree);*
(v) *monomial groups whose non linear irreducible characters are induced from abelian subgroups;*
(vi) *monomial groups such that all its non linear irreducible characters have same degree;*
(vii) *monomial groups of odd order such that there are exactly two non linear irreducible characters of each degree;*
(viii) *solvable groups with all its elements of prime power order;*
(ix) *the class $\mathcal{A}$ of solvable groups with all its Sylow subgroups abelian;*
(x) *$\mathcal{A}$-by-supersolvable groups, in particular, abelian-by-supersolvable groups;*
(xi) *solvable groups G satisfying the following condition: for all primes p dividing the order of G and for all subgroups A of G, $O^p(A)$,*

the unique smallest normal subgroup of A *such that* $A/O^p(A)$ *is a* p*-group, has no central* p*-factor.*

Remark 2. **(i):** *There do exist groups in* $\mathcal{C}$ *which are not strongly monomial groups. For example, in (*[10]*, section 2) it is shown that* SmallGroup(1000,86)*, which is a well known example of a group that is not strongly monomial, belongs to* $\mathcal{C}$*.*

(ii): *So far, there is no known example of a monomial group which is not generalized strongly monomial.*

The following picture sums up the vastness of generalized strongly monomial groups.



5. Wedderburn Decomposition

In order to find the precise Wedderburn decomposition of $\mathbb{Q}G$, the knowledge given so far is not sufficient. Let us elaborate on this point. Suppose S is the set of all Shoda pairs of G and E is the complete and irredundant set of all primitive central idempotents of $\mathbb{Q}G$. We have learnt in the previous section that each Shoda pair in S realizes a primitive central idempotent of $\mathbb{Q}G$. More explicitly we have the map $\theta : S \to E$ which sends (H, K) to $\alpha e(G, H, K)$, where $\alpha = \frac{[\mathrm{Cen}_G(\varepsilon(H,K)):H]}{[\mathbb{Q}(\lambda):\mathbb{Q}(\chi)]}$, and λ is any linear character on H with kernel K. Note that for any $(H, K) \in S$ and $x \in G$, both (H, K) and (H^x, K^x) have the same image under θ and hence θ is not one-one. In order to determine E, we need a subset $\mathcal{S}$ of S such that the restriction of θ to $\mathcal{S}$ is a bijection from $\mathcal{S}$ to E. Having known such a set $\mathcal{S}$ and the structure of the simple component of $\mathbb{Q}G$ associated with each Shoda pair, the Wedderburn decomposition of $\mathbb{Q}G$ is immediate, if G is monomial.

Since different Shoda pairs can give rise to same primitive central idempotent, an equivalence of Shoda pairs is defined as follows:

Definition 6. Equivalence of Shoda pairs *Two Shoda pairs of G are said to be equivalent, if they realize the same primitive central idempotent of $\mathbb{Q}G$.*

Definition 7. Complete and Irredundant set of Shoda pairs *A set of representatives of distinct equivalence classes of Shoda pairs of G is called a complete irredundant set of Shoda pairs of G.*

A necessary and sufficient condition for two Shoda pairs of G to realize the same primitive central idempotent of $\mathbb{Q}G$ is given by the following:

Theorem 18. ([39], Proposition 1.4) *Let (H_1, K_1) and (H_2, K_2) be two Shoda pairs of a finite group G and let $\alpha_1, \alpha_2 \in \mathbb{Q}$ such that $e_i = \alpha_i e(G, H_i, K_i)$ is a primitive central idempotent of $\mathbb{Q}G$ for $i = 1, 2$. Then $e_1 = e_2$ if, and only if, there exists $g \in G$ such that $H_1^g \cap K_2 = K_1^g \cap H_2$.*

Thus (H_1, K_1) and $(H_2, K_2) \in S$ are equivalent if and only if there exists $g \in G$ such that $H_1^g \cap K_2 = K_1^g \cap H_2$. A procedure to write a complete and irredundant set $\mathcal{S}$ of Shoda pairs, i.e., a set of representatives of distinct equivalence classses of Shoda pairs for some classes of groups is known:

- metabelian groups [3].
- normally monomial groups [13].
- groups in the class $\mathcal{C}$, i.e., any finite group G such that every subquotient of G satisfy the following property that either it is abelian or it contains a non central abelian normal subgroup [10].

5.1. **Metabelian groups.** In [3], a procedure to write a complete and irredundant set of Shoda pairs for a finite metabelian group G is given, and as a consequence one can write the explicit Wedderburn decomposition of $\mathbb{Q}G$ such a group.

Let G be a metabelian group and let $\mathcal{A}$ be a fixed maximal abelian subgroup of G containing the commutator subgroup G' of G. Consider the set $\mathcal{T}$ of all the subgroups D of G with $D \leq \mathcal{A}$ and $\mathcal{A}/D$ cyclic under the equivalence relation defined by conjugacy in G. Let $\mathcal{T}_G$ denote a set of representatives of the distinct equivalence classes. Set

$$S_G := \{(D, \mathcal{A}) \mid D \in \mathcal{T}_G,\ D \text{ core-free}\}.$$

If $N \trianglelefteq G$ with $\mathcal{A}_N/N$ a maximal abelian subgroup of G/N containing its commutator $(G/N)'$, (if $N = (1)$, $\mathcal{A}_N/N = \mathcal{A}$), similarly define

$$S_{G/N} := \{(D/N, \mathcal{A}_N/N) \mid D/N \in \mathcal{T}_{G/N},\ D/N \text{ core-free}\}.$$

Set

$$\mathcal{S} := \{(N, D/N, \mathcal{A}_N/N) \mid N \trianglelefteq G,\ S_{G/N} \neq \emptyset,\ (D/N, \mathcal{A}_N/N) \in S_{G/N}\}.$$

Let $\chi \in \mathrm{Irr}(G)$ and $N = \ker(\chi)$. Let $\overline{\chi}$ be the corresponding character of G/N. Since $\overline{\chi}$ is a faithful character of G/N, by Theorem 2 of [3], there exists a unique pair $(D/N, \mathcal{A}_N/N) \in S_{G/N}$ such that $\overline{\chi}$ is the character afforded by $\overline{\rho}^G$ for some linear representation $\overline{\rho}$ of $\mathcal{A}_N/N$ with kernel D/N. In this situation, the triple $(N, D/N, \mathcal{A}_N/N)$ is said to be associated with χ.

The following theorem is proved in [3], which gives a complete and irredundant set of Shoda pairs and primitive central idempotents, and the Wedderburn decomposition of $\mathbb{Q}G$.

Theorem 19. ([3], Theorem 3) *Let G be a finite metabelian group with $\mathcal{S}$ defined as above. Then the following staements hold:*

(i) If $\chi \in \mathrm{Irr}(G)$ with $(N, D/N, \mathcal{A}_N/N)$ its associated triple, then (A_N, D) is a strong Shoda pair of G that determine χ and

$$e_{\mathbb{Q}}(\chi) = e(G, \mathcal{A}_N, D).$$

(ii) $\{(A_N, D) \mid (N, D/N, \mathcal{A}_N/N) \in \mathcal{S}\}$ is a complete and

irredundant set of Shoda pairs of G.
(iii)

$$\mathbb{Q}G \cong \sum_{(N,\, D/N\, \mathcal{A}_N/N) \in \mathcal{S}} \mathbb{Q}Ge(G, \mathcal{A}_N, D).$$

Furthermore, for any $(N, D/N, \mathcal{A}_N/N) \in \mathcal{S}$,

$$\mathbb{Q}Ge(G, \mathcal{A}_N, D) = M_n(\mathbb{Q}(\xi_k) \star^{\sigma}_{\tau} (N_G(D)/\mathcal{A}_N)), \tag{2}$$

where $k = [\mathcal{A}_N : D]$, $n = [G : N_G(D)]$, x is a generator of $\mathcal{A}_N/D$, $\varphi : N_G(D)/\mathcal{A}_N \to N_G(D)/D$ denotes a left inverse of the canonical projection $N_G(D)/D \to N_G(D)/\mathcal{A}_N$, $\mathbb{Q}(\xi_k) \star^{\sigma}_{\tau} (N_G(D)/\mathcal{A}_N)$ denotes the crossed product of the group $N_G(D)/\mathcal{A}_N$ over the coefficient ring $\mathbb{Q}(\xi_k)$ with action σ and twisting τ given by $\xi_k^{\sigma(a)} = \xi_k^i$, if $x^{\phi(a)} = x^i$; $\tau(a, b) = \xi_k^j$, if $\varphi(ab)^{-1}\varphi(a)\varphi(b)D = x^j$, for $a, b \in N_G(D)/\mathcal{A}_N$ and integers i, j.

In section 4 of [3], the above theorem is also illustrated with several examples.

5.2. **Normally monomial groups.** The work of [3] has been generalized to normally monomial groups in [13]. Recall that a group G is called *normally monomial*, if every complex irreducible character of G is induced from a linear character of a normal subgroup of G. It may be mentioned that all metabelian groups are normally monomial and there do exist normally monomial groups which are not metabelian. In fact there do exist normally monomial groups which are not abelian-by-supersolvable. It may be pointed out that there also exist abelian-by-supersolvable which are not normally monomial. One can see examples in section 4 of [3] for more details.

A Shoda pair (H, K) of G is said to be an extremely strong Shoda pair if $H \trianglelefteq G$. It is easy to see that an extremely strong Shoda pair of G is a strong Shoda pair of G, because the first condition of a strong Shoda pair, i.e., $H \trianglelefteq N_G(K)$ holds trivially and the second condition, i.e., $\epsilon(H, K)$ and $\epsilon(H, K)^g$ are orthogonal for all $g \in G \setminus N_G(K)$ holds because, $N_G(K)$ being equal to $\mathrm{Cen}_G(\epsilon(H, K))$, both of them are distinct primitive central idempotents of $\mathbb{Q}H$. Two extremely strong Shoda pairs are said to be equivalent, if they are equivalent as Shoda pairs of G. A set of representatives of distinct equivalence classes of extremely strong Shoda pairs of G is called a complete irredundant set of extremely strong Shoda pairs of G.

For a finite group G, let $\mathcal{N}$ be the set of all the distinct normal subgroups of G. For $N \in \mathcal{N}$, let A_N be a normal subgroup of G containing N such that A_N/N is an abelian normal subgroup of maximal order in G/N. Note that the choice of A_N is not unique. However, one needs to fix one such A_N for a given N. For a fixed A_N, set

$\mathcal{D}_N$: the set of all subgroups D of A_N containing N such that $\mathrm{core}_G(D) = N, A_N/D$ is cyclic and is a maximal abelian subgroup of $N_G(D)/D$.

$\mathcal{T}_N$: a set of representatives of $\mathcal{D}_N$ under the equivalence relation defined by conjugacy of subgroups in G.

$\mathcal{S}_N$: $\{(A_N, D) \mid D \in \mathcal{T}_N\}$.

Theorem 20. ([13], Theorem 1) *Let G be a finite group. Then,*

(i): $\cup_{N\in\mathcal{N}}\mathcal{S}_N$ *is a complete irredundant set of extremely strong Shoda pairs of G.*

(ii): $\{e(G, A_N, D) \mid (A_N, D) \in \mathcal{S}_N,\ N \in \mathcal{N}\}$ *is the complete set of primitive central idempotents of $\mathbb{Q}G$, if, and only if, G is normally monomial.*

As a consequence, Corollary 1 of [13] provides explicit Wedderburn decomposition of $\mathbb{Q}G$ when G is a normally monomial group. Precise computation for some normally monomial groups (including examples which are not abelian-by-supersolvable) have been done in the section 4 of [13].

5.3. **Groups in $\mathcal{C}$.** The work presented here is an abstraction of that in the previous section and has been done by the author with Kaur in [10]. In Theorem 17 we mentioned that the class $\mathcal{C}$ consisting of all finite groups G such that every quotient group of a subgroup of G satisfy the property that either it is abelian or it contains a non central abelian normal subgroup is a large class of monomial groups that contain strongly monomial and normally monomial groups besides various other important families of monomial groups.

Here is a broad idea of constructing Shoda pairs of an arbitrary group $G \in \mathcal{C}$ given in [10]. Corresponding to each normal subgroup N of G, a rooted directed tree $\mathcal{G}_N$ has been constructed whose particular leaves correspond to Shoda pairs of G ([10],Theorem 2). The condition for the collection of Shoda pairs corresponding to these leaves of $\mathcal{G}_N$ as N runs over all the normal subgroups of G to be complete and irredundant has been investigated in Theorem 3 of [10]. The vertices of the tree $\mathcal{G}_N$ are some special character triples. Recall that a character triple (H, A, ϑ) of G means $H \leq G$, $A \trianglelefteq H$, and $\vartheta \in \mathrm{Irr}(A)$ invariant in H. It is said to be N-linear character triple of G, if, in addition, ϑ is linear and $\ker \vartheta^G = N$. The tree $\mathcal{G}_N$ is constructed as a directed subgraph of a bigger graph $\mathcal{G}$, described below.

Construction of graph $\mathcal{G}$

The vertex set $\mathcal{V}$ of $\mathcal{G}$ consists of N-linear character triples of G and for given (H, A, ϑ) and (H', A', ϑ') in $\mathcal{V}$, there is an edge from (H, A, ϑ) to (H', A', ϑ'), if $(H', A', \vartheta') \in Cl(H, A, \vartheta)$, where the construction of $Cl(H, A, \vartheta)$ is slightly technical and is described as follows:
For N-linear character triple (H, A, ϑ) of G, denote by

- $\mathcal{A}\mathrm{ut}(\mathbb{C}|\vartheta)$: the group of automorphisms of the field $\mathbb{C}$ of complex numbers which keep $\mathbb{Q}(\vartheta)$ fixed element wise.
- $\mathrm{Irr}(H|\vartheta)$: the set of all irreducible characters ψ of H which lie above ϑ, i.e., the restriction ψ_A of ψ to A has ϑ as a constituent.
- $\widetilde{\mathrm{Irr}}(H|\vartheta)$: subset of $\mathrm{Irr}(H|\vartheta)$ consisting of those ψ which satisfy $\ker \psi^G = N$.
- $\widetilde{\mathrm{Lin}}(H|\vartheta)$: subset of $\widetilde{\mathrm{Irr}}(H|\vartheta)$ consisting of linear characters.
- $\mathcal{A}_{(H,A,\vartheta)}$: a fixed normal subgroup of H of maximal order containing $\ker \vartheta$ such that $\mathcal{A}_{(H,A,\vartheta)}/\ker \vartheta$ is abelian. Note that there may be several choices of such $\mathcal{A}_{(H,A,\vartheta)}$, however, we fix one such choice for a given triple (H, A, ϑ).

For notational convenience, denote $\mathcal{A}_{(H,A,\vartheta)}$ by $\mathcal{A}$ and consider the action of $\mathcal{A}\mathrm{ut}(\mathbb{C}|\vartheta)$ on $\widetilde{\mathrm{Lin}}(\mathcal{A}|\vartheta)$ by setting

$$\sigma.\varphi = \sigma \circ \varphi,\ \sigma \in \mathcal{A}\mathrm{ut}(\mathbb{C}|\vartheta),\ \varphi \in \widetilde{\mathrm{Lin}}(\mathcal{A}|\vartheta).$$

Also H acts on $\widetilde{\mathrm{Lin}}(\mathcal{A}|\vartheta)$ by

$$h.\varphi = \varphi^h,\ h \in H,\ \varphi \in \widetilde{\mathrm{Lin}}(\mathcal{A}|\vartheta).$$

Notice that the two actions on $\widetilde{\mathrm{Lin}}(\mathcal{A}|\vartheta)$ are compatible in the sense that

$$\sigma.(h.\varphi) = h.(\sigma.\varphi),\ h \in H,\ \sigma \in \mathcal{A}\mathrm{ut}(\mathbb{C}|\vartheta),\ \varphi \in \widetilde{\mathrm{Lin}}(\mathcal{A}|\vartheta).$$

This consequently gives an action of $\mathcal{A}\mathrm{ut}(\mathbb{C}|\vartheta)\times H$ on $\widetilde{\mathrm{Lin}}(\mathcal{A}|\vartheta)$. Under this double action, denote by

- $\mathfrak{Lin}(\mathcal{A}|\vartheta)$: a set of representatives of distinct orbits of $\widetilde{\mathrm{Lin}}(\mathcal{A}|\vartheta)$.

Finally, if $H = A$, define $Cl(H, A, \vartheta)$ to be an empty set, otherwise set

$$Cl(H, A, \vartheta) := \{(I_H(\varphi), \mathcal{A}, \varphi) \mid \varphi \in \mathfrak{Lin}(\mathcal{A}|\vartheta)\},$$

where $I_H(\varphi) = \{g \in G \mid \varphi^g = \varphi\}$ is the inertia group of φ in H. Note that all the character triples in $Cl(H, A, \vartheta)$ are N-linear character triples of G.

Construction of graph $\mathcal{G}_N$ from $\mathcal{G}$

Clearly $(G, N, 1_N) \in \mathcal{V}$, where 1_N is the character of N which takes constant value 1. Let $\mathcal{V}_N$ be the set of those vertices $v \in \mathcal{V}$ for which there is a directed path from $(G, N, 1_N)$ to v. If $\mathcal{E}$ denotes the set of edges of $\mathcal{G}$, denote by $\mathcal{E}_N$ the set of ordered pairs $(u, v) \in \mathcal{E}$ so that $u, v \in \mathcal{V}_N$. Then $\mathcal{G}_N = (\mathcal{V}_N, \mathcal{E}_N)$ is a directed subgraph of $\mathcal{G}$.

The following theorem tells us that certain leaves of $\mathcal{G}_N$ yield Shoda pairs of G.

Theorem 21. ([10], Theorem 2) *Let $G \in \mathcal{C}$ and $\mathcal{N}$ the set of all normal subgroups of G.*

- **(i):** *For $N \in \mathcal{N}$, the following hold:*
  - **(a):** *$\mathcal{G}_N$ is a rooted directed tree with $(G, N, 1_N)$ as its root;*
  - **(b):** *the leaves of $\mathcal{G}_N$ of the type (H, H, ϑ) correspond to Shoda pairs of G. More precisely, if (H, H, ϑ) is a leaf of $\mathcal{G}_N$, then $(H, \ker\vartheta)$ is a Shoda pair of G.*
- **(ii):** *If (H', K') is any Shoda pair of G, then there is a leaf (H, H, ϑ) of $\mathcal{G}_N$, where $N = \mathrm{core}_G(K')$, such that (H', K') and $(H, \ker\vartheta)$ realize the same primitive central idempotent of $\mathbb{Q}G$.*

The next question whether the collection of these Shoda pairs obtained from the leaves of $\mathcal{G}_N$ as N runs through all the normal subgroups of G give us a complete and irredundant set of Shoda pairs is discussed in Theorem 3 of [10]. Denote by $\mathcal{L}_N$ those leaves of $\mathcal{G}_N$ which yield Shoda pairs of G. Let $\mathcal{S}_N$ be the set of Shoda pairs obtained from the leaves in $\mathcal{L}_N$. If $(H, H, \vartheta) \in \mathcal{L}_N$ and $v_1\ (v_1, v_2)\ v_2 \cdots (v_{n-1}, v_n)\ v_n$ is the directed path from $v_1 = (G, N, 1_N)$ to $v_n = (H, H, \vartheta)$, then we call n to be the height of (H, H, ϑ) and term v_i as the i^{th} node of (H, H, ϑ), $1 \le i \le n$.

Definition 8. Good leaf ([10], p.458) *Let* $(H, H, \vartheta) \in \mathcal{L}_N$ *be of height* n *with* (H_i, A_i, ϑ_i) *as its* i^{th} *node,* $1 \leq i \leq n$. *Then* (H, H, ϑ) *is said to be good if the following holds for all* $1 < i \leq n$*: given* $x \in N_{H_{i-1}}(\ker \vartheta_i)$*, there exist* $\sigma \in \mathcal{A}\mathrm{ut}(\mathbb{C})$ *such that* $(\vartheta^{H_i})^x = \sigma \circ \vartheta^{H_i}$.

Theorem 22. ([10], Theorem 3) *If* $G \in \mathcal{C}$ *and* $\mathcal{N}$ *is the set of all normal subgroups of* G*, then* $\bigcup_{N \in \mathcal{N}} \mathcal{S}_N$ *is a complete irredundant set of Shoda pairs of* G *if, and only if, the leaves in* $\mathcal{L}_N$ *are good for all* $N \in \mathcal{N}$.

For precise computation of the directed trees and corresponding Shoda pairs, one can look at section 7 of [10]. The computation has also been done for small(1000, 86) in GAP library which is not strongly monomial but, being in $\mathcal{C}$, is generalized strongly monomial.

6. Arbitrary finite groups

Beyond monomial groups, there are only a few known results that describe the algebraic structure of their rational group algebras. This has been done for alternating group by Giambruno and Jespers in [19].

For an arbitrary irreducible character χ of G, in [36], the authors have proved using Brauer induction theorem on induced characters that the expressions $e(G, H, K)$, where (H, K) run over strong Shoda pairs of G, are the building blocks to describe $e_{\mathbb{Q}}(\chi)$. More precisely the following is proved:

Theorem 23. ([36], Proposition 3.2) *Let* G *be a finite group of order* n *and* χ *an irreducible chorarcter of* G. *Then the primitive idempotent* $e_{\mathbb{Q}}(\chi)$ *of* $\mathbb{Q}G$ *associated with* χ *is of the form*

$$e_{\mathbb{Q}}(\chi) = \frac{\chi(1)}{[\mathbb{Q}(\zeta_n) : \mathbb{Q}(\chi)]} \sum_i \frac{a_i}{[G : C_i]} [\mathbb{Q}(\zeta_n) : \mathbb{Q}(\psi_i)] e(G, H_i, K_i)$$

where $a_i \in \mathbb{Z}$, (H_i, K_i) *are strong Shoda pairs of subgroups of* G, $C_i = \mathrm{Cen}_G(\epsilon(H_i, K_i))$ *and* ψ_i *are linear characters of* H_i *with kernel* K_i.

Using Artin induction theorem, a refinement in the expression of $e_{\mathbb{Q}}(\chi)$ given in the above theorem was later done by Janssens in [23]. Let C_1, C_2, $\cdots$,C_r be representatives of the conjugacy classes of cyclic subgroups of G and let c_i be a generator of C_i. For $\chi \in \mathrm{Irr}\, G$, set

$$b_{C_i} = \frac{[G : \mathrm{Cen}_G(c_i)]}{[G : C_i]} \sum \mu([C_i^* : C_i])\Big(\sum_{\sigma \in \mathrm{Gal}(\mathbb{Q}(\chi)/\mathbb{Q})} \sigma(\chi)\Big)(z^*),$$

where the first sum runs through all cyclic subgroups C_i^* of G which contain C_i, z^* is a generator of C_i^* and μ is Möbius mu function. With this notation, Janssens proved the following:

Theorem 24. ([23], Theorem 0.2) *Let G be a finite group and χ an irreducible chararcter of G. Then the primitive idempotent $e_{\mathbb{Q}}(\chi)$ of $\mathbb{Q}G$ associated with χ is of the form*

$$e_{\mathbb{Q}}(\chi) = \sum_{i=1}^{r} \frac{b_{C_i}\chi(1)}{[G : \mathrm{Cen}_G(\varepsilon(C_i, C_i))]} e(G, C_i, C_i) = \sum_{i=1}^{r} \frac{b_{C_i}\chi(1)}{[G : C_i]} (\sum_{k=1}^{m_i} \varepsilon(C_i, C_i)^{g_{i_k}}),$$

where $T_i = \{g_{i_1}, \cdots, g_{i_{m_i}}\}$ is a right transversal of C_i in G.

In [3] and [4], an alternative approach has been used to write expressions of $e_{\mathbb{Q}}(\chi)$ for an irreducible character χ of G in terms of the Euler function φ and Mobius function μ. Initially in [4], a constaint on χ was imposed and later in [3], the result was obtained for an arbitrary $\chi \in \mathrm{Irr}(G)$. The following theorem gives an expression for $e_{\mathbb{Q}}(\chi)$ as an explicit element of $\mathbb{Q}G$:

Theorem 25. ([3], Theorem 1) *Let G be a finite group of order n and $\chi \in \mathrm{Irr}(G)$. The primitive central idempotent $e_{\mathbb{Q}}(\chi)$ of $\mathbb{Q}G$ associated with χ is given by*

$$e_{\mathbb{Q}}(\chi) = \frac{\chi(1)}{\sum_{g\in G, \chi(g)\neq 0} \sum_{d|n} \left(\frac{\nu_d^{\chi}(g)\mu(d)}{\varphi(d)}\right)^2} \sum_{g\in G,\, \chi(g)\neq 0} \sum_{d|n} \frac{\nu_d^{\chi}(g)\mu(d)}{\varphi(d)} g.$$

Here $\nu_d^{\chi}(g)$ is the number of eigenvalues of ρ(g) of order d *and ρ is a representation of G affording the character χ.*

References

[1] P. J. Allen and C. Hobby, *A characterization of units in* $\mathbf{Z}[A_4]$, J. Algebra **66** (1980), no. 2, 534–543.

[2] P. J. Allen and C. Hobby, *A characterization of units in* $\mathbf{Z}S_4$, Comm. Algebra **16** (1988), no. 7, 1479–1505.

[3] Gurmeet K. Bakshi, R. S. Kulkarni, and I. B. S. Passi, *The rational group algebra of a finite group*, J. Algebra Appl. **12** (2013), no. 3.

[4] Gurmeet K. Bakshi, and I. B. S. Passi, *Primitive central idempotents in rational group algebras*, Comm. Algebra **40** (2012), no. 4, 1413–1426.

[5] Gurmeet K. Bakshi, and S. Maheshwary, *Extremely strong Shoda pairs with* `GAP`, J. Symbolic Comput. **76** (2016), no. 5, 97–106.

[6] Gurmeet K. Bakshi, Osnel Broche Cristo, Allen Herman, Alexander Konovalov, Sugandha Maheswary, Aurora Olivieri, Gabriela Olteanu, Ángel del Río, and Inneke van Geldar, *Wedderga — wedderburn decomposition of group algebras*, Version 4.10; (2018), (http://www.gap-system.org).

[7] Gurmeet K. Bakshi, Shalini Gupta, and Inder Bir S. Passi, *Semisimple metacyclic group algebras*, Proc. Indian Acad. Sci. Math. Sci. **121** (2011), no. 4, 379–396.

[8] Gurmeet K. Bakshi, Shalini Gupta, and Inder Bir S. Passi, *The structure of finite semisimple metacyclic group algebras*, J. Ramanujan Math. Soc. **28** (2013), no. 2, 141–158.

[9] Gurmeet K. Bakshi, Shalini Gupta, and Inder Bir S. Passi, *The algebraic structure of finite metabelian group algebras*, Comm. Algebra **43** (2015), no. 6, 2240–2257.

[10] Gurmeet K. Bakshi and Gurleen Kaur, *Character triples and Shoda pairs*, J. Algebra **491** (2017), 447–473.

[11] Gurmeet K. Bakshi and Gurleen Kaur, *A generalization of strongly monomial groups*, J. Algebra **520** (2019), 419–439.

[12] Gurmeet K. Bakshi and Gurleen Kaur, *Semisimple finite group algebra of a generalized strongly monomial group*, Finite Fields Appl. **60** (2019), 101571.

[13] Gurmeet K. Bakshi and Sugandha Maheshwary, *The rational group algebra of a normally monomial group*, J. Pure Appl. Algebra **218** (2014), no. 9, 1583–1593.

[14] Gurmeet K. Bakshi and Sugandha Maheshwary, *On the index of a free abelian subgroup in the group of central units of an integral group ring*, J. Algebra **434** (2015), 72–89.

[15] Gurmeet K. Bakshi, Sugandha Maheshwary, and Inder Bir S. Passi, *Integral group rings with all central units trivial*, J. Pure Appl. Algebra **221** (2017), no. 8, 1955–1965.

[16] Gurmeet K. Bakshi, Sugandha Maheshwary, and Inder Bir S. Passi, *Group rings and the RS-property*, Comm. Algebra **47** (2019), no. 3, 969–977.

[17] Osnel Broche and Ángel del Río, *Wedderburn decomposition of finite group algebras*, Finite Fields Appl. **13** (2007), no. 1, 71–79.

[18] S. P. Coelho, E. Jespers, and C. Polcino Milies, *Automorphisms of group algebras of some metacyclic groups*, Comm. Algebra **24** (1996), no. 13, 4135–4145.

[19] A. Giambruno and E. Jespers, *Central idempotents and units in rational group algebras of alternating groups*, Internat. J. Algebra Comput. **8** (1998), no. 4, 467–477.

[20] E. G. Goodaire, E. Jespers, and C. Polcino Milies, *Alternative loop rings*, North-Holland Mathematics Studies, vol. 184, North-Holland Publishing Co., Amsterdam, 1996.

[21] A. Herman, *On the automorphism groups of rational group algebras of metacyclic groups*, Comm. Algebra **25** (1997), no. 7, 2085–2097.

[22] Graham Higman, *The units of group-rings*, Proc. London Math. Soc. (2) **46** (1940), 231–248.

[23] Geoffrey Janssens, *Primitive central idempotents of rational group algebras*, J. Algebra Appl. **12** (2013), no. 1, 1250130, 5.

[24] E. Jespers and Á. del Río, *A structure theorem for the unit group of the integral group ring of some finite groups*, J. Reine Angew. Math. **521** (2000), 99–117.

[25] E. Jespers and Á. del Río, *Group ring Groups - Volume 1 : Orders and generic constructions of units*, De Gruyter Textbook, De Gruyter, 2015.

[26] E. Jespers, Á. del Río, and M. Ruiz, *Groups generated by two bicyclic units in integral group rings*, J. Group Theory **5** (2002), no. 4, 493–511.

[27] E. Jespers, Á. del Río, and I. Van Gelder, *Writing units of integral group rings of finite abelian groups as a product of Bass units*, Math. Comp. **83** (2014), no. 285, 461–473.

[28] E. Jespers and G. Leal, *Generators of large subgroups of the unit group of integral group rings*, Manuscripta Math. **78** (1993), no. 3, 303–315.

[29] E. Jespers, G. Leal, and A. Paques, *Central idempotents in the rational group algebra of a finite nilpotent group*, J. Algebra Appl. **2** (2003), no. 1, 57–62.

[30] E. Jespers, G. Leal, and C. Polcino Milies, *Idempotents in rational abelian group algebras*, Preprint.

[31] E. Jespers, G. Leal, and C. Polcino Milies, *Units of integral group rings of some metacyclic groups*, Canad. Math. Bull. **37** (1994), no. 2, 228–237.

[32] E. Jespers, G. Olteanu, and Á. del Río, *Rational group algebras of finite groups: from idempotents to units of integral group rings*, Algebr. Represent. Theory **15** (2012), no. 2, 359–377.

[33] E. Jespers, G. Olteanu, Á. del Río, and I. Van Gelder, *Group rings of finite strongly monomial groups: central units and primitive idempotents*, J. Algebra **387** (2013), 99–116.

[34] E. Jespers, G. Olteanu, Á. del Río, and I. Van Gelder, *Central units of integral group rings*, Proc. Amer. Math Soc. **142** (2014), 2193–2209.

[35] E. Jespers and M. M. Parmenter, *Construction of central units in integral group rings of finite groups*, Proc. Amer. Math. Soc. **140** (2012), no. 1, 99–107.

[36] Eric Jespers, Gabriela Olteanu, and Ángel del Río, *Rational group algebras of finite groups: from idempotents to units of integral group rings*, Algebr. Represent. Theory **15** (2012), no. 2, 359–377.

[37] A. Olivieri and Á. del Río, *An algorithm to compute the primitive central idempotents and the Wedderburn decomposition of a rational group algebra*, J. Symbolic Comput. **35** (2003), no. 6, 673–687.

[38] A. Olivieri, Á. del Río, and J. J. Simón, *The group of automorphisms of the rational group algebra of a finite metacyclic group*, Comm. Algebra **34** (2006), no. 10, 3543–3567.

[39] Aurora Olivieri, Ángel del Río, and Juan Jacobo Simón, *On monomial characters and central idempotents of rational group algebras*, Comm. Algebra **32** (2004), no. 4, 1531–1550.

[40] Sam Perlis and Gordon L. Walker, *Abelian group algebras of finite order*, Trans. Amer. Math. Soc. **68** (1950), 420–426.

[41] S. K. Sehgal, *Topics in group rings*, Monographs and Textbooks in Pure and Applied Math., vol. 50, Marcel Dekker, Inc., New York, 1978.

[42] S. K. Sehgal, *Units of integral group rings—a survey*, Algebraic structures and number theory (Hong Kong, 1988), World Sci. Publ., Teaneck, NJ, 1990, pp. 255–268.

[43] S. K. Sehgal, *Units in integral group rings*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 69, Longman Scientific & Technical, Harlow, 1993, With an appendix by Al Weiss.

[44] T. Yamada, *The Schur subgroup of the Brauer group*, Lecture Notes in Mathematics, Vol. 397, Springer-Verlag, Berlin, 1974.

Gurmeet K. Bakshi
Centre for Advanced Study in Mathematics
Panjab University, Chandigarh 160014, India
E-mail: gkbakshi@pu.ac.in

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 29–41

IDEAL CONVERGENCE, ROLE OF NICE IDEALS AND SUMMABILITY MATRICES*

PRATULANANDA DAS

1. Basic facts of ideal convergence, role of nice ideals

For the last several decades the summability theory and, in particular, the study of convergence of sequences has been one of the most important and active area of research works in Pure Mathematics and has extensively found application in Topology, Functional Analysis, Fourier Analysis, Measure Theory, Applied Mathematics, Mathematical Modeling, Computer Science etc.

The usual notion of convergence does not always capture in fine details the properties of vast class of sequences that are not convergent. Also many times in different investigations in Mathematics we come across sequences that are not convergent but almost all of its terms (in some sense) have the properties of a convergent sequence. So it always seems better to include more sequences under purview, while discussing convergence. One way of including more sequences under purview is to consider those sequences that are convergent when restricted to some 'big' set of natural numbers which is a big set in certain prevalent sense. If by a 'big' set of natural numbers one understands a co-finite subset of the set of all natural numbers then the usual notion of convergence arises. If by a 'big' set of natural numbers one understands a subset of the set of all natural numbers having asymptotic (or natural) density equal to zero, then the notion of statistical convergence arises.

Let $\mathbb{N}$ denote the set of natural numbers and for $K \subset \mathbb{N}$, $K(m,n)$ denote the cardinality of the set $K \cap [m,n]$. The upper and lower natural density of the subset K are defined by

$$\overline{d}(K) = \limsup_{n\to\infty} \frac{K(1,n)}{n} \text{ and } \underline{d}(K) = \liminf_{n\to\infty} \frac{K(1,n)}{n}.$$

If $\overline{d}(K) = \underline{d}(K)$ then we say that the natural density of K exists and it is denoted simply by $d(K)$. Clearly $d(K) = \lim_{n\to\infty} \frac{K(1,n)}{n}$. Intuitively, it is thought that there

* This article is based on the text of the 30th Srinivasa Ramanujan Memorial Award lecture delivered at the 85th Annual conference of the IMS - An International Meet held at IIT Kharagpur, W. B. during November 22-25, 2019

are more positive integers than perfect squares, or say, cubes, since every perfect square is already positive, and there are so many natural numbers which are not squares. However, the set of positive integers is not in fact larger in size than the set of perfect squares: both sets are infinite and countable and can therefore be put in one-to-one correspondence. Nevertheless if one goes through the natural numbers, the squares become increasingly scarce. The notion of natural density makes this intuition precise. The natural density (as well as some other types of densities) is in fact studied in probabilistic number theory. The notion of natural density was first introduced to define a more general notion of convergence by Fast [14] and independently by Steinhaus [28] in 1951. After the works of Šalăt [23] and particularly of Fridy and Connor [16, 17, 8, 9, 10, 11] it became one of the major thirst areas of Summability theory and since then a lot of work has been done on statistical convergence and its further generalizations.

Most importantly the idea of statistical convergence have been extended to two types of convergences, namely, $\mathcal{I}$ and $\mathcal{I}^*$ convergence by Kostyrko et al in 2000 [18] with the help of ideals (which have long been topics of research in Set Theory). This approach is much more general as most of the known convergence methods become special cases.

We start by recalling the basic notions of ideals and filters. A family $\mathcal{I} \subset 2^Y$ of subsets of a non-empty set Y is said to be an ideal in Y if

$$(i)\ A,\ B \in \mathcal{I} \text{ implies } A \cup B \in \mathcal{I},$$

$$(ii)\ A \in \mathcal{I},\ B \subset A \text{ imply } B \in \mathcal{I}.$$

Further an admissible ideal $\mathcal{I}$ of Y satisfies $\{x\} \in \mathcal{I}$ for each $x \in Y$. Such ideals are also called free ideals. If $\mathcal{I}$ is a proper non-trivial ideal in Y (i.e. $Y \notin \mathcal{I}$, $\mathcal{I} \neq \{\phi\}$), then the family of sets $\mathcal{F}(\mathcal{I}) = \{M \subset Y : \text{there exists } A \in \mathcal{I} : M = Y \setminus A\}$ is a filter in Y. It is called the filter associated with the ideal $\mathcal{I}$.

Definition 1.1. [18] *A sequence (x_n) is said to be $\mathcal{I}$-convergent to $\xi \in \mathbb{R}$ ($\xi = \mathcal{I}-\lim_{n\to\infty} x_n$) if and only if for each $\varepsilon > 0$ the set $A(\varepsilon) = \{n \in \mathbb{N} : |x_n - \xi| \geq \xi\} \in \mathcal{I}$. The element ξ is called the $\mathcal{I}$-limit of the sequence (x_n).*

Some of the examples of ideals and corresponding convergence notions are described below (see [18]).

Example 1.1. (a) If $\mathcal{I}$ is the class $\mathcal{I}_f$ of all finite subsets of $\mathbb{N}$ then $\mathcal{I}_f$ is a non-trivial admissible ideal and $\mathcal{I}_f$-convergence coincides with the usual convergence of real sequences.

(b) If $\mathcal{I}_d$ is the class of all $A \subset \mathbb{N}$ with $d(A) = 0$, then $\mathcal{I}_d$ is a non-trivial admissible ideal and $\mathcal{I}_d$-convergence coincides with the statistical convergence.

(c) The uniform density of a set $A \subset \mathbb{N}$ is defined as follows: For integers $t \geq 0$ and $s \geq 1$ let $A(t+1, t+s) = \mathrm{card}\{n \in A : t+1 \leq n \leq t+s\}$. Put

$$\beta_s = \liminf_{t\to\infty} A(t+1,t+s),\ \beta^s = \limsup_{t\to\infty} A(t+1,t+s).$$

It can be shown that the following limits exist :

$$\underline{u}(A) = \lim_{s\to\infty} \frac{\beta_s}{s},\ \overline{u}(A) = \lim_{s\to\infty} \frac{\beta^s}{s}.$$

If $\underline{u}(A) = \overline{u}(A)$, then $u(A) = \underline{u}(A)$ is called the uniform density of the set A. Put $\mathcal{I}_u = \{A \subset \mathbb{N} : u(A) = 0\}$. Then $\mathcal{I}_u$ is a non-trivial ideal and $\mathcal{I}_u$-convergence is said to be the uniform statistical convergence.

(d) A wide class of $\mathcal{I}$-convergences can be obtained as follows. For an infinite matrix of reals $A = (a_{i,k})$, a sequence of reals $\widetilde{x} = (x_k)$ and $n \in \mathbb{N}$ we write $A_n(\widetilde{x}) = \sum_{k=1}^{\infty} a_{n,k}x_k$.

Let $A = (a_{i,k})$ be an infinite matrix of reals. We say that a sequence $\widetilde{x} = (x_k)$ is A-summable if

(1) the series $A_n(\widetilde{x})$ is convergent for all but finitely many $n \in \mathbb{N}$,
(2) the sequence $(A_n(\widetilde{x}))$ is convergent.

The real $\lim_{n\to\infty} A_n(\widetilde{x})$ is called the A-limit of the sequence $\widetilde{x}$ and is denoted by $\lim^A \widetilde{x}$. The famous Silverman-Toeplitz theorem [Silverman, 1913 and independently Toeplitz, 1913] says that a matrix A is regular if and only if $\lim_{n\to\infty} A_n(\widetilde{x}) = \lim_{n\to\infty} x_n$ for every ordinary convergent sequence $\widetilde{x}$ which is equivalent to the following conditions.

(i) $\sup_j \big(\sum_{k=1}^{\infty} |a_{j,k}|\big) < +\infty$,
(ii) $\lim_j a_{j,k} = 0$ for each $k \in \mathbb{N}$,
(iii) $\lim_j \Big(\sum_{k=1}^{\infty} a_{j,k}\Big) = 1$.

Of particular interest have been the class of non-negative regular summability matrices. For $E \subset \mathbb{N}$ and a non-negative regular matrix $A = (a_{i,k})$, we put

$$d_A^{(n)}(E) = \sum_{k=1}^{\infty} a_{n,k}\ \chi_E(k)$$

for $n \in \mathbb{N}$. If $\lim_{n\to\infty} d_A^{(n)}(E) = d_A(E)$ exists, then $d_A(E)$ is called A-density of E [15]. From the regularity of A it follows that $\lim_{n\to\infty} \sum_{k=1}^{\infty} a_{n,k} = 1$ and from this we see that $d_A(E) \in [0,1]$ (if it exists). Consequently $\mathcal{I}_A = \{E \subset \mathbb{N} : d_A(E) = 0\}$ is a non-trivial ideal. Note that $\mathcal{I}_d$-convergence can be obtained from $\mathcal{I}_A$-convergence by choosing $a_{n,k} = \frac{1}{n}$ for $k \leq n$ and $a_{n,k} = 0$ for $k > n$. On the other hand if

$a_{n,k} = \frac{1}{s_n}$ for $k \le n$ and $a_{n,k} = 0$ for $k > n$ where $s_n = \displaystyle\sum_{j=1}^{n} \frac{1}{j}$ for $n \in \mathbb{N}$, then we obtain the notion of $\mathcal{I}_\delta$-convergence (logarithmic convergence). Finally choosing $a_{n,k} = \frac{\phi(k)}{n}$ for $k \le n, k|n$ and $a_{n,k} = 0$ for $k \le n, k$ does not divide n and $a_{n,k} = 0$ for $k > n$ we get ϕ-convergence of Schoenberg (see [24]), where ϕ is the Euler function.

Clearly if $\mathcal{I}$ is an admissible ideal then for a sequence of real numbers, the usual convergence implies $\mathcal{I}$-convergence with the same limit.

The following properties are the most familiar axioms of convergence.

(a) Every constant sequence $(\xi, \xi, ...)$ converges to ξ.

(b) The limit of any convergent sequence is uniquely determined.

(c) If a sequence (x_n) has the limit ξ then each of the subsequence has the same limit.

(d) If each subsequence of the sequence (x_n) has a subsequence which converges to ξ then the sequence (x_n) converges to ξ.

Theorem 1.1. *(i) The notion of $\mathcal{I}$-convergence satisfy (a),(b) and (d).*

(ii) If $\mathcal{I}$ contains an infinite set then $\mathcal{I}$-convergence does not satisfy (c).

Proof. (i) The proof of (a) is trivial. To prove (b) observe that for any $A_1, A_2 \in \mathcal{I}$ we have $(\mathbb{N} \setminus A_1) \cap (\mathbb{N} \setminus A_2) \neq \phi$. If there are two limits $\xi, \eta \in \mathbb{R}, \xi \neq \eta$ choose ε such that

$$0 < \varepsilon < \frac{1}{2}|\xi - \eta|$$

and put $A_1 = \{n \in \mathbb{N} : |x_n - \xi| \ge \varepsilon\}, A_2 = \{n \in \mathbb{N} : |x_n - \eta| \ge \varepsilon\}$.

Suppose now that (d) does not hold. Then there exists $\varepsilon_0 > 0$ such that

$$A(\varepsilon_0) = \{n \in \mathbb{N} : |x_n - \xi| \ge \varepsilon_0\} \notin \mathcal{I}.$$

But then $A(\varepsilon_0)$ is an infinite set since $\mathcal{I}$ is admissible. Let $A(\varepsilon_0) = \{n_1 < n_2 < ...\}$. Put $y_k = x_{n_k}$ for $k \in \mathbb{N}$. Then (y_k) is a subsequence of x without a subsequence $\mathcal{I}$-convergent to ξ.

(ii) Suppose that $A \in \mathcal{I}$ is an infinite set, $A = \{n_1 < n_2 < ...\}$. Let $B = \mathbb{N} \setminus A = \{m_1 < m_2 < ...\}$. The set B is also infinite set since $\mathcal{I}$ is a non-trivial ideal. Define $x = (x_n)$ by choosing $\xi_1, \xi_2 \in \mathbb{R}, \xi_1 \neq \xi_2$ and put $x_{n_k} = \xi_1, x_{m_k} = \xi_2$ for $k \in \mathbb{N}$. Obviously $\mathcal{I} - \lim x_k = \xi_2$, but the subsequence $y_k = x_{n_k},\ k \in \mathbb{N}$, $\mathcal{I}$-converges to ξ_1.

Obviously if $\mathcal{I}$ is an admissible ideal which does not contain any infinite set then $\mathcal{I}$-convergence coincides with the usual convergence and obviously fulfills (c).

Recall the following result from the theory of statistical convergence. A sequence (x_n) of real numbers is statistically convergent to ξ if and only if there exist a set $M = \{m_1 < m_2 < ...\} \subset \mathbb{N}$ such that $d(M) = 1$ and $\lim_{k\to\infty} x_{m_k} = \xi$ (See [23]).

This result influenced the introduction of the following concept of convergence, namely, $\mathcal{I}^*$-convergence.

Definition 1.2. [18] *A sequence $x = (x_n)$ is said to be $\mathcal{I}^*$-convergent to $\xi \in \mathbb{R}$ if and only if there exists a set $M \in \mathcal{F}(\mathcal{I})$, $M = \{m_1 < m_2 < ...\}$ such that* $\lim_{k\to\infty} |x_{m_k} - \xi| = 0$.

The similar results like that of the followings were originally proved in [18] in the more general settings of a metric space.

Theorem 1.2. *Let $\mathcal{I}$ be an admissible ideal. If $\mathcal{I}^* - \lim x_n = \xi$ then $\mathcal{I} - \lim x_n = \xi$.*

Proof. By assumption there exists a set $H \in \mathcal{I}$ such that for $M = \mathbb{N} \setminus H = \{m_1 < m_2 < ...\}$ we have

$$\lim_{k\to\infty} x_{m_k} = \xi.$$

Let $\varepsilon > 0$. Then there exists $k_0 \in \mathbb{N}$ such that $|x_{m_k} - \xi| < \varepsilon$ for each $k > k_0$. Then obviously

$$A(\varepsilon) = \{n \in \mathbb{N} : |x_n - \xi| \geq \varepsilon\} \subset H \cup \{m_1 < m_2 < ... < m_{k_0}\}. \quad (1.1)$$

The set on the right-hand side of (1.1) belong to $\mathcal{I}$ (since $\mathcal{I}$ is admissible). So $A(\varepsilon) \in \mathcal{I}$.

The converse implication between $\mathcal{I}$- and $\mathcal{I}^*$-convergence depends essentially on the structure of the metric space (X, ρ). and in [18] it was shown that if X has no accumulation point then $\mathcal{I}$- and $\mathcal{I}^*$- convergence coincide for each admissible ideal $\mathcal{I}$. Otherwise we can have a result like following.

Theorem 1.3. *There exist an admissible ideal $\mathcal{I}$ and a sequence (y_n) of real numbers such that $\mathcal{I} - \lim y_n = \xi$ but $\mathcal{I}^* - \lim y_n$ does not exist.*

Proof. Let $\xi \in \mathbb{R}$. Then we can always find a sequence (x_n) of distinct real numbers which is usually convergent to ξ i.e. $\lim x_n = \xi$ and the sequence $(|x_n - \xi|)$ is decreasing to 0. Put $\varepsilon_n = |x_n - \xi|$ for $n \in \mathbb{N}$. For $\mathcal{I}$ we take the ideal $\mathcal{I}_1$ as follows. Let $\mathbb{N} = \bigcup_{j=1}^{\infty} \triangle_j$ be a decomposition of $\mathbb{N}$ into infinite number of pairwise disjoint infinite subsets. $\mathcal{I}_1$ denotes the class of all subsets of $\mathbb{N}$ which intersect only finite number of sets $\triangle_j$'s. Then $\mathcal{I}_1$ is a nontrivial admissible ideal.

Define the sequence (y_n) by $y_n = x_j$ if $n \in \triangle_j$. Let $\eta > 0$. Choose $\nu \in \mathbb{N}$ such that $\varepsilon_\nu < \eta$. Then $A(\eta) = \{n : |y_n - \xi| \geq \eta\} \subset \triangle_1 \cup \triangle_2 \cup ... \cup \triangle_\nu$. Hence $A(\eta) \in \mathcal{I}_1$ and $\mathcal{I}_1 - \lim y_n = \xi$. But it can be shown that $\mathcal{I}_1^* - \lim y_n$ does not exist.

Next a necessary and sufficient condition (for ideal $\mathcal{I}$) is presented under which $\mathcal{I}$- and $\mathcal{I}^*$-convergence are equivalent. This condition (AP) is similar to the condition (APO) used in [11]. The nomenclature "AP" has actually been used from long ago in Set Theory in the following sense: A collection of subsets $\mathcal{D}$ of $\mathbb{N}$ having property "Q" has additive property if given any sequence of sets (A_n) in $\mathcal{D}$, there is a set $A \in \mathcal{D}$ such that $A \subset^* A_n$ for all n i. e. A is almost contained in every A_n meaning $A \setminus A_n$ is finite for every n.

Definition 1.3. *An admissible ideal $\mathcal{I}$ is said to satisfy the condition (AP) (or is called a P-ideal or sometimes AP-ideal) if for every countable family of mutually disjoint sets $(A_1, A_2, ...)$ from $\mathcal{I}$ there exists a countable family of sets $(B_1, B_2, ...)$ such that $A_j \triangle B_j$ is finite for each $j \in \mathbb{N}$ and $\bigcup_{k=1}^{\infty} B_k \in \mathcal{I}$.*

It is clear that $B_j \in \mathcal{I}$ for each $j \in \mathbb{N}$.

Theorem 1.4. *Let $\mathcal{I}$ be an admissible ideal.*
(i) If the ideal $\mathcal{I}$ has the property (AP) then for arbitrary sequence (x_n) of real numbers, $\mathcal{I} - \lim x_n = \xi$ implies $\mathcal{I}^* - \lim x_n = \xi$.
(ii) If for every arbitrary sequence (x_n) of real numbers, $\mathcal{I} - \lim x_n = \xi$ implies $\mathcal{I}^* - \lim x_n = \xi$, then $\mathcal{I}$ has the property (AP).

Proof. (i) Suppose that $\mathcal{I}$ satisfies the condition (AP). Let $\mathcal{I} - \lim x_n = \xi$. Then $A(\varepsilon) = \{n \in \mathbb{N} : |x_n - \xi| \geq \varepsilon\} \in \mathcal{I}$ for every $\varepsilon > 0$. Put $A_1 = \{n \in \mathbb{N} : |x_n - \xi| \geq 1\}$ and $A_n = \{n \in \mathbb{N} : \frac{1}{n} \leq |x_n - \xi| \leq \frac{1}{n-1}\}$ for $n \geq 2, n \in \mathbb{N}$. Obviously $A_i \cap A_j = \phi$ for $i \neq j$. By condition (AP) there exists a countable family of sets $(B_1, B_2, ...)$ such that $Aj \triangle B_j$ is finite for each $j \in \mathbb{N}$ and $B = \bigcup_{k=1}^{\infty} B_k \in \mathcal{I}$. It is sufficient to prove that for $M = \mathbb{N} \setminus B$ we have

$$\lim_{n \to \infty, n \in M} x_n = \xi. \quad (1.2)$$

Let $\eta > 0$. Choose $k \in \mathbb{N}$ such that $\frac{1}{k+1} < \eta$. Then $\{n \in \mathbb{N} : |x_n - \xi| \geq \eta\} \subset \bigcup_{j=1}^{k+1} A_j$. Since $A_j \triangle B_j, j = 1, 2, ..., (k+1)$ are finite sets, there exists $n_0 \in \mathbb{N}$ such that

$$(\bigcup_{j=1}^{k+1} B_j) \cap \{n \in \mathbb{N} : n > n_0\} = (\bigcup_{j=1}^{k+1} A_j) \cap \{n \in \mathbb{N} : n > n_0\}. \quad (1.3)$$

If $n > n_0$ and $n \notin B$ then $n \notin \bigcup_{j=1}^{k+1} B_j$ and, by (1.3), $n \notin \bigcup_{j=1}^{k+1} A_j$. But then $|x_n - \xi| < \frac{1}{n+1} < \eta$ and so (1.2) holds.
(ii) Choose any $\xi \in \mathbb{R}$. We can find a sequence (x_n) of distinct real numbers such that $\lim x_n = \xi$ and the sequence $(|x_n - \xi|)$ is decreasing to 0. For $n \in \mathbb{N}$ let

$\varepsilon_n = |x_n - \xi|$. Let (A_n) be a disjoint family of non-empty sets from $\mathcal{I}$. Define a sequence (y_n) by $y_n = x_j$, if $n \in A_j$. Let $\eta > 0$. Choose $m \in \mathbb{N}$ such that $\varepsilon_m < \eta$. Then $A(\eta) = \{n \in \mathbb{N} : |y_n - \xi| \geq \eta\} \subset A_1 \cup A_2 \cup ... \cup A_m$. Hence $A(\eta) \in \mathcal{I}$ and $\mathcal{I} - \lim y_n = \xi$. By virtue of our assumption we have also $\mathcal{I}^* - \lim y_n = \xi$. Hence there exists a set $B \in \mathcal{I}$ such that if $M = \mathbb{N} \setminus B = (m_1 < m_2 < ...)$ then

$$\lim_{k\to\infty} y_{m_k} = \xi. \ (1.4)$$

Put $B_j = A_j \cap B$ for $j \in \mathbb{N}$. Then $B_j \in \mathcal{I}$ for each j. Further $\bigcup_{j=1}^{\infty} B_j = B \cap (\bigcup_{j=1}^{\infty} A_j) \subset B$. Hence $\bigcup_{j=1}^{\infty} B_j \in \mathcal{I}$. Fix $j \in \mathbb{N}$. From (1.4) it follows that A_j has only a finite number of elements common with the set M. Thus there exists $k_0 \in \mathbb{N}$ such that $A_j \subset (A_j \cap B) \cup \{m_1, m_2, ..., m_{k_0}\}$. Hence $A_j \triangle B_j = A_j - B_j \subset \{m_1, m_2, ..., m_{k_0}\}$ which implies that $A_j \triangle B_j$ is a finite set. From the arbitrariness of $j \in \mathbb{N}$ it follows that $\mathcal{I}$ has the property (AP).

We have already seen how condition (AP) which is a combinatorial property of the ideals can come handy in the theory of ideal convergence. There are many more instances where P ideals come into picture, which interested readers can see from several papers on ideal convergence. However for certain investigations, one need to look further, typically into the topological aspects of ideals. Following is a classic case of such application.

Definition 1.4. *Let $x = (x_n)$ be a sequence of real numbers.*
(i) An element $\xi \in \mathbb{R}$ is said to be an $\mathcal{I}$-limit point of x provided that there is a set $M = (m_1 < m_2 < ...) \subset \mathbb{N}$ such that $M \notin \mathcal{I}$ and $\lim_{k\to\infty} x_{m_k} = \xi$.
(ii) An element $\xi \in \mathbb{R}$ is said to be an $\mathcal{I}$-cluster point of x if and only if for each $\varepsilon > 0$ we have $\{n \in \mathbb{N} : |x_n - \xi| < \varepsilon\} \notin \mathcal{I}$.

Denote by $\mathcal{I}(L_x)$ and $\mathcal{I}(C_x)$ the set of all $\mathcal{I}$-limit and $\mathcal{I}$-cluster points of x, respectively. The similar results like that of the following results may be found in [12, 20].

Theorem 1.5. *Let $\mathcal{I}$ be an admissible ideal. Then for each sequence $x = (x_n)$ of real numbers we have $\mathcal{I}(L_x) \subset \mathcal{I}(C_x)$.*

Proof. Let $\xi \in \mathcal{I}(L_x)$. Then there exists a set $M = \{m_1 < m_2 < ...\} \notin \mathcal{I}$ such that

$$\lim_{k\to\infty} |x_{m_k} - \xi| = 0.$$

Take $\delta > 0$. According to above, there exists $k_0 \in \mathbb{N}$ such that for $k > k_0$ we have $|x_{m_k} - \xi| < \delta$. Hence $\{n \in \mathbb{N} : |x_n - \xi| < \delta\} \supset M \setminus \{m_1, m_2, ..., m_{k_0}\}$ and so

$\{n \in \mathbb{N} : |x_n - \xi| < \delta\} \notin \mathcal{I}$, which means that $\xi \in \mathcal{I}(C_x)$.

It was also observed in [12] in the topological settings X that if $x = (x_n)$ and $y = (y_n)$ are two sequences in X such that $\{n \in \mathbb{N} : x_n \neq y_n\} \notin \mathcal{I}$ then $\mathcal{I}(C_x) = \mathcal{I}(C_y)$, $\mathcal{I}(L_x) = \mathcal{I}(L_y)$.

Theorem 1.6. [20] *Let $\mathcal{I}$ be an admissible ideal.*
(i) The set $\mathcal{I}(C_x)$ is closed for each sequence $x = (x_n)$ of real numbers.
(ii) Suppose that there exists a disjoint sequence of sets (M_n) such that $M_n \subset \mathbb{N}$ and $M_n \notin \mathcal{I}$ for $n \in \mathbb{N}$. Then for each closed set $F \subset \mathbb{R}$ there exists a sequence $x = (x_n)$ of real numbers such that $F = \mathcal{I}(C_x)$.

Next a characterization of the set of $\mathcal{I}$-limit points is presented for certain special class of ideals. The result was established in the most general structure of a topological space in [12].

Recall that after identifying the power set $\mathcal{P}(\mathbb{N})$ of $\mathbb{N}$ with the Cantor space $C = \{0,1\}^{\mathbb{N}}$ in a standard manner we may consider an ideal as a subset of C. An ideal is called an analytic ideal if it corresponds to an analytic subset of C. Solecki [26, 27] proved that every analytic P-ideal is determined by some lower semicontinuous submeasure on $\mathbb{N}$.

Let S be a set. We say that a map $\varphi : \mathcal{P}(S) \to [0, \infty]$ is a submeasure on S if it satisfies the following conditions:

- $\varphi(\phi) = 0$ and $\varphi(\{s\}) < \infty$ for every $s \in S$,
- φ is monotone: if $A \subset B \subset S$, then $\varphi(A) \leq \varphi(B)$,
- φ is subadditive: if $A, B \subset S$, then $\varphi(A \cup B) \leq \varphi(A) + \varphi(B)$.

A submeasure φ on $\mathbb{N}$ is lower semicontinuous if for every $A \subset \mathbb{N}$ we have $\varphi(A) = \lim\limits_{n\to\infty} \varphi(A \cap [1, n])$.

Note that a submeasure on $\mathbb{N}$ is lower semicontinuous if and only if it is lower semicontinuous as a function from $P(\mathbb{N})$ to $[0, \infty]$. Solecki in [27] proved that every analytic P-ideal $\mathcal{I}$ can be presented as

$$\mathcal{I} = \{A \subset \mathbb{N} \ : \ \lim_{n\to\infty} \varphi(A \backslash [1, n]) = 0\}$$

for some lower semicontinuous submeasure φ on $\mathbb{N}$.

Theorem 1.7. [12] *For any sequence (x_n) of real numbers, the set $\mathcal{I}(L_x)$ is a F_σ-set provided $\mathcal{I}$ is an analytic P-ideal.*

For the next theorem we shall need the following observation which is true in $\mathbb{R}$. For any closed set $F \subset \mathbb{R}$ there is a sequence $x = (x_n)$ of real numbers such that $F = L(x)$ ($z \in L(x)$ if for each neighborhood W of z, $\{n \in \mathbb{N} : x_n \in W\}$ is infinite).

Theorem 1.8. [12] *For each F_σ-set A in $\mathbb{R}$ there exists a sequence $x = (x_n)$ of real numbers such that $A = \mathcal{I}(L_x)$ provided $\mathcal{I}$ is an analytic P-ideal.*

There are several other instances where analytic P-ideals (or more precisely, the property of being "analytical") are found to be extremely helpful. If we assume that $\mathcal{I}$ is an analytic P-ideal then ideal limits of continuous functions behave like ordinary limits. In [18] it was shown that $\mathcal{I}_d$-limits of sequences of continuous functions are of the first Baire class. Finally it was generalized to all analytic P-ideals and all Baire classes in [19]. In [21] it was shown that if $\mathcal{I}$ is an analytic P-ideal then for any finite measure space (X, M, μ), real valued measurable functions $f_n, f (n \in \mathbb{N})$ defined almost everywhere on X such that (f_n) is pointwise $\mathcal{I}$-convergent to f almost everywhere on X and every $\varepsilon > 0$ there is an $A \in M$ such that $\mu(X\ A) < \varepsilon$ and f_n equi ideally converges to f on A i. e. an ideal version of the famous Egoroff's Theorem of Measure Theory holds for analytic P-ideals.

2. How many distinct nice ideals are there, and role of non-negative matrices

We have already seen that the class of all analytic P-ideals are the most important class which helps to obtain several deep and interesting results in the theory of ideal convergence. This class itself have been topics of research in the field of Set Theory which we would not dwell upon much here. It had all started long back in 1950's. Points of $\mathbb{N}^* = \beta\mathbb{N} \setminus \mathbb{N}$ (i.e. the remainder in the Stone-Cech compactification of the space of natural numbers with discrete topology) are identified with free ultrafilters on the set $\mathbb{N}$. Recall that ultrafilters are those filters which are not properly contained in any other filter. P-points are precisely those ultrafilters whose dual ideals are P-ideals. In 1956, Rudin showed that the space $\mathbb{N}^*$ has P-points if the continuum hypothesis is assumed. In view of all these, we can ask a very natural question as to how many distinct such ideals are there.

One can actually give an answer, but in order to provide the result, one needs to consider the following. The notion of natural density can be generalized as follows. Let $g : \mathbb{N} \to [0, \infty)$ be a function with $\lim\limits_{n\to\infty} g(n) = \infty$ and $n/g(n) \nrightarrow 0$. The upper simple density (or density of weight g) [1] was defined by the formula

$$\overline{d}_g(A) = \limsup_{n\to\infty} \frac{|A \cap [1, n]|}{g(n)}$$

for $A \subset \mathbb{N}$.

Theorem 2.1. [1] *If $g\colon \mathbb{N} \to [0, \infty)$ is such that $g(n) \to \infty$ and $n/g(n) \nrightarrow 0$, then the ideal $\mathcal{I}_g$ is equal to $Exh(\varphi)$ where*

$$\varphi(E) = \sup_{n\in\mathbb{N}} \frac{card(E \cap n)}{g(n)} \quad \textit{for } E \subset \mathbb{N},$$

and φ is a lower semicontinuos submeasure on $\mathbb{N}$. Consequently, $\mathcal{I}_g$ is an $F_{\sigma\delta}$ P-ideal on $\mathbb{N}$.

Now the following result provides the answer to our question and we can conclude that there are uncountably many distinct analytic P-ideals !!

Theorem 2.2. [1] *There exists a family $G_0 \subset G$ of cardinality c such that $\mathcal{I}_f$ is incomparable with $\mathcal{I}_d$ for every $f \in G$, and $\mathcal{I}_f$ and $\mathcal{I}_g$ are incomparable for any distinct $f, g \in G_0$.*

We have already seen that non-negative regular matrices also generate ideals. Since the class of these matrices is also huge (at least uncountable) and many of the known nice ideals like $\mathcal{I}_d$ are nothing but the ideals generated by suitable matrices, so Connor conjectured that the class of all analytic P-ideals coincide with the class of all ideals generated by non-negative regular matrices. In the rest of this article we state results from the very recent paper [13] which shows that the conjecture is false.

In order to understand how things work, we first break down the definition of these matrices and introduce notions of certain matrices which are not necessarily regular.

Definition 2.1. *We say that a matrix $A = (a_{i,k})$ is*

(1) nonnegative *if $a_{i,k} \geq 0$ for every $i, k \in \mathbb{N}$,*
(2) admissible *if $\lim_{i\to\infty} a_{i,k} = 0$ for every $k \in \mathbb{N}$.*

Definition 2.2. *A matrix $A = (a_{i,k})$ is semiregular if*

(1) A is admissible,
(2) $\lim_{i\to\infty} \sum_{k\in\omega} a_{i,k} = \infty$.

A semiregular matrix $A = (a_{i,k})$ is of

- type 1 *if $\sum_{k\in\mathbb{N}} |a_{i,k}| < \infty$ for all but finitely many i;*
- type 2 *if $\sum_{k\in\mathbb{N}} |a_{i,k}| = \infty$ for infinitely many i.*

It can be verified that the functions d_{ST1} and d_{ST2} defined as before (recall how a density function was defined for a regular matrix) for semiregular matrices of type 1 and type 2 respectively both satisfy basic properties of density functions and one can generate the corresponding ideals

$$\mathcal{I}(ST1) = \{E \subset \mathbb{N} : d_{ST1}(E) = 0\}$$

and

$$\mathcal{I}(ST2) = \{E \subset \mathbb{N} : d_{ST2}(E) = 0\}.$$

The following result obtained in [2] showed that indeed all non-negative regular matrices generate analytic P-ideals.

Theorem 2.3. [2] *Every ideal belonging to* $\mathcal{I}(REG)$ *is a* $F_{\sigma\delta}P$*-ideal where* $\mathcal{I}(REG) = \{\mathcal{I}_A : A \text{ is a regular matrix}\}$ *and* $\mathcal{I}_A = \{E \subset \mathbb{N} : d_A(E) = 0\}$ *(as defined in Section 1).*

Finally for semi-regular matrices we have the following results.

Theorem 2.4. [13] *Every member of* $\mathcal{I}(ST1)$ *is also a* $F_{\sigma\delta}P$*-ideal whereas members of* $\mathcal{I}(ST2)$ *are* $F_{\sigma\delta}$ *but not necessarily* P*-ideal.*

Now coming back to the folklore Connor's conjecture, it is not actually true. So in order to disprove Connor's conjecture we need to construct suitable example of ideals and for that the following techniques can come very handy.

Definition 2.3. *For ideals* $\mathcal{I}, \mathcal{J}$ *and* $A \notin \mathcal{I}$ *we define the following new ideals*

(1) $\mathcal{I} \oplus \mathcal{J} = \{A \subset \mathbb{N} \times \{0,1\} : \{n : (n,0) \in A\} \in \mathcal{I} \wedge \{n : (n,1) \in A\} \in \mathcal{J}\}$,

(2) $\mathcal{I} \otimes \mathcal{J} = \{A \subset \mathbb{N} \times \mathbb{N} : \{n : \{k : (n,k) \in A\} \notin \mathcal{J}\} \in \mathcal{I}\}$,

(3) $\mathcal{I} \otimes \emptyset = \{A \subset \mathbb{N} \times \mathbb{N} : \{n : \{k : (n,k) \in A\} \neq \emptyset\} \in \mathcal{I}\}$.

An ideal $\mathcal{I}$

(1) is dense if for every infinite $A \subset \mathbb{N}$ there is an infinite $B \in \mathcal{I}$ such that $B \subset A$,

(2) is summable if for every $f : \mathbb{N} \to [0,\infty)$ such that $\sum_{n=1}^{\infty} f(n) = \infty$ we define a *summable ideal generated by a function* f by $\mathcal{I}_f = \{B \subset \mathbb{N} : \sum_{n\in B} f(n) < \infty\}$. In particular, if $f(n) = 1/n$ we obtain the ideal $\mathcal{I}_{1/n} = \{B \subset \mathbb{N} : \sum_{n\in B} \frac{1}{n} < \infty\}$. It is known that summable ideals are F_σ P-ideals.

We are now in a position to state our final observations.

• The most common ideal $\mathcal{I}_d$ (w.r.t natural density) is in $\mathcal{I}(REG)$ which can not be generated by any semiregular matrix.

• $\mathcal{I}_{1/n}$ is a dense $F_{\sigma\delta}P$-ideal belonging to $\mathcal{I}(ST2)$ which can not be generated by any regular matrix or semiregular matrix of type 1.

• $\mathcal{I}_{1/n} \oplus Fin$ is a non-dense $F_{\sigma\delta}P$-ideal belonging to $\mathcal{I}(ST2)$ which can not be generated by any regular matrix or semiregular matrix of type 1.

• $Fin \otimes \emptyset$ is a $F_{\sigma\delta}$ non P-ideal generated by a semiregular matrix of type 2.

• $\mathcal{I}_u \oplus Fin$ is an ideal which can not be generated by any matrix, regular or semiregular.

References

1. M. Balcerzak, P. Das, M. Filipczak, J. Swaczyna, Generalized kinds of density and the associated ideals, Acta Math. Hungar, **147** (1) (2015), 97 - 115.

2. A. Bartoszewich, P. Das, S. Glab, On matrix summability of spliced sequences and A-density of points, Linear Algebra Appl., **487** (2015), 22 - 42.
3. S. Bhunia, P. Das and S. K. Pal, Restricting statistical convergence, Acta Math. Hungar., **134** (1-2) (2012), 153 - 161.
4. J. Boos, Classical and Modern Methods in Summability, Oxford Math. Monogr., Oxford Univ. Press, Oxford, 2000.
5. R. C. Buck, The measure theoretic approach to density, Amer. J. Math., **68** (1946), 560 - 580.
6. H. Cartan, Théorie des filtres, C. R. Acad. Sci. Paris, **205** (1937), 595 - 598.
7. R. G. Cooke, Infinite Matrices and Sequence Spaces, Dover Publ., New York, 1965.
8. J. Connor, The statistical and strong p-Cesaro convergence of sequences. Analysis. **8** (1988), 47-63.
9. J. Connor, On strong matrix summability with respect to a modulus and statistical convergence, Canad. Math. Bull., **32** (1989), 194 - 198.
10. J. Connor, Two valued measures and summability, Analysis, **10** (1990), 373 - 385.
11. J. Connor, R-type summability methods, Cauchy criteria, P-sets and statistical convergence. Proc. Amer. Math. Soc. **115** (1992), 319 - 327.
12. P. Das, Some further results on ideal convergence in topological spaces. Topology Appl. **159** (10-11) (2012), 2621 - 2626.
13. P. Das, R. Filipow, J. Tryba, A note on nonregular matrices and ideals associated with them, Colloq. Math., **159** (1) (2020), 29 - 45.
14. H. Fast, Sur la convergence statistique, Colloq. Math., **2** (1951), 241 - 244.
15. A.R. Freedman, J.J. Sember, Densities and summability, Pacific J. Math., **9** (1981), 293 - 305.
16. J.A. Fridy, On statistical convergence, Analysis, **5** (1985), 301 - 313.
17. J.A. Fridy, Statistical limit points, Proc. Amer Math. Soc., **118** (1993), 1187 - 1192.
18. P. Kostyrko, T. Šalăt, W. Wilczynski, $\mathcal{I}$-convergence, Real Anal. Exchange, **26** (2) (2000-2001), 669 - 686.
19. M. Laczkovich, I. Reclaw, Ideal limits of sequences of continuous functions, Fund. Math., **203** (1) (2009), 39 - 46.
20. B.K. Lahiri, P. Das, I and I^*-convergence in topological spaces, Math. Bohemica, **130** (2) (2005), 153 - 160.
21. N. Mrozek, Ideal version of Egorov's Theorem for analytic P-ideal, J. Math. Anal. Appl, **349** (2) (2009), 452 - 458.
22. F. Nuray, W.H. Ruckle, Generalized statistical convergence and convergence free spaces, J. Math. Anal. Appl., **245** (2000), 513 - 527.
23. T. Šalăt, On statistically convergent sequences of real numbers, Math. Slovaca, **30** (1980), 139 - 150.
24. I.J. Schoenberg, The integrability of certain functions and related summability methods. Amer. Math. Monthly. **66** (1959), 361 - 375.
25. L. L. Silverman, On the definition of the sum of a divergent series, Univ. of Missouri Stud. Math. Ser. 1, Univ. of Missouri, Columbia, MO, 1913.
26. S. Solecki, Analytic ideals, Bull. Symb. Logic, **92** (3) (1994), 339 - 348.
27. S. Solecki, Analytic ideals and their applications, Ann. Pure Appl. Logic, **99** (1-3) (1999), 51 - 72.
28. H. Steinhaus, Sur la convergence ordinaire et la convergence asymptotique, Colloq. Math. **2** (1951), 73 - 74.

29. O. Toeplitz, ber allgemeine lineare Mittelbildungen, Prace Matematyczno-Fizyczne, **22** (1913), 113 - 119.

Pratulananda Das
Department of Mathematics
Jadhavpur University, Kolkata-700032, W. B.
E-mail :pratulananda@yahoo.co.in

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 43–57

AN EXPOSITION OF ARTIN'S PRIMITIVE ROOT CONJECTURE

GAURI GUPTA

(Received : 24 - 08 - 2019 ; Revised : 14 - 01 - 2020)

Abstract. The main aim of this paper is to give an exposition of Artin's primitive root conjecture. In the first half of the paper, we discuss different density theorems such as Dirichlet, Frobenius and Chebotarev density theorems required to understand the Artin's Conjecture. In the later half, we also give a heuristic approach to this conjecture.

1. Introduction

Arithmetic density theorems are statements about the size of sets of prime numbers which are defined by arithmetic conditions. A typical example is the following. Let S be the set of prime numbers whose last digit in decimal representation is 1 and n be a prime number chosen randomly, then what will be the probability that $n \in S$. We have, $P(n \in S) = \frac{1}{4}$. In general, S can represent the set of prime numbers whose last digit in decimal representation is either 1, 3, 5 or 7.

In this paper we discuss different density theorems and the relation between them. Chebotarev density theorem may be regarded as a generalization of the Dirichlet's theorem on primes in arithmetic progressions and Frobenius theorem on the splitting of polynomials in finite Galois extensions. The Frobenius Density Theorem is a specialization of the Chebotarev density theorem in which the conjugacy class C is required to be a division of the galois group G. We say that two elements of G belong to the same division if the cyclic subgroups that they generate are conjugate in G. The partition of G into divisions, in general, is coarse than its partition into conjugacy classes and thus, Frobenius theorem is correspondingly weaker than Chebotarev density theorem.

2010 Mathematics Subject Classification: 11A07, 11R45
Key words and phrases: Dirichlet density, Riemann hypothesis, Frobenius density, Chebotarev density theorem, Root Conjecture

Now let $a \in \mathbb{Z}/\{0\}$ and $p \nmid a$ be a prime number. The unit group $\mathbb{F}_p^* = (\mathbb{Z}/m\mathbb{Z})^*$ is a cyclic group of order $p-1$. If the residue of a is a generator of $\mathbb{F}_p^*$, it is said to be a primitive root of a modulo p. Let us look at the decimal expansion of numbers of the form $\frac{1}{p}$ with p prime. For example, $\frac{1}{11} = 0.090909\ldots$ and for $p \nmid 10$, these decimal expansions are purely periodic. Denote by k, the period of decimal expansion of $\frac{1}{p}$. Then,

$$10^k \equiv 1 \bmod p.$$

This implies that k is the smallest integer for which the above congruence is satisfied. So, k is the multiplicative order of 10 modulo p and is denoted by $ord_p(10)$. By Fermat's little theorem we have

$$10^{p-1} \equiv 1 \bmod p$$

and $1 \leq k \leq p-1$. Thus, $k \mid p-1$. If $k = p-1$, we say that 10 is a primitive root mod p. In general if $p \nmid a$ and the smallest integer k such that

$$a^k \equiv 1 \bmod p$$

is $p-1$, then a is a primitive root mod p. The decimal expansion for $\frac{1}{7} = 0.142857142857\ldots$ shows that 10 is a primitive root modulo 7. We expect that there are infinitely many primes p having 10 as a primitive root modulo p.

In order to speak quantitatively about the distribution of certain types of primes, or integers in general, we need some sort of measure theory on the set of primes or the set of integers. We define two types of densities namely, natural density and Dirichlet density.

Let $S \subseteq T$ be two subsets of positive integers, with T being infinite. We define the natural density of S in T as

$$\delta(S) = \lim_{N\to\infty} \frac{\#\{n \in S : n \leq N\}}{\#\{n \in T : n \leq N\}},$$

provided this limit exists. Here # A denotes the cardinality of set A.

Properies. Let P be the set of prime numbers. The natural density δ, of $S \subset P$ has the following properties.

- $0 \leq \delta(S) \leq 1,\ \delta(P) = 1$
- If $S_1 \subset S_2$, then $\delta(S_1) \leq \delta(S_2)$
- If $S \subset P$ has a natural density δ, then $P \backslash S$ has density $1-\delta$.

- If S_1, S_2 has density δ_1, δ_2, and if the density of $S_1 \cup S_2$ exits and is equal to δ, then $max\{\delta_1, \delta_2\} \leq \delta \leq min\{\delta_1 + \delta_2, 1\}$.
- If $S_1 \triangle S_2$ is finite, then $\delta_1 = \delta_2$.

Example. Let S be the set of primes of the form n^2+1. Then the natural density $\delta(S) = 0$.

Solution. Since S is a subset of the set of numbers of the form n^2+1, this implies

$$\delta(S) \leq \lim_{x\to\infty} \frac{\sqrt{x-1}}{x/\log x} = 0.$$

Here we are using PNT (see Theorem **??**), that is, the number of primes $p \leq x$ asymptotically tend to $\frac{x}{\log x}$ Thus, $\delta(S) = 0$. Nevertheless, the set S is expected to be infinite.

1.1. **Dirichlet Density.** Let $S \subseteq T$ be two subsets of positive integers, with $\sum\limits_{n\epsilon T} n^{-s}$ divergent. Then the limit

$$d(S) = \lim_{\substack{s\to 1\\ s>1}} \frac{\sum\limits_{n\epsilon S} n^{-s}}{\sum\limits_{n\epsilon T} n^{-s}}$$

provided it exists, is called the Dirichlet density of S. (see [3, Theorem 8, Ch. VIII, §4]).

Lemma. Let $S \subseteq T$ be two subsets of $\mathbb{N}$ such that S has a natural density δ in T. Then S also has Dirichlet density δ in T.

However, the converse is false, that is, there can be sets such that they admit a Dirichlet density but not a natural density. For example, let S be the set of positive integers which have the first digit 1 when written in base 10. Then the set S has a Dirichlet density but the natural density of set S does not exist (lower natural density $\neq$ upper natural density). In this sense, Dirichlet density is a strictly more general notion. The main difference between the two concepts is that natural density rests on an ordering of the set of primes, whereas Dirichlet density does not.

Theorem. There are infinitely many primes p of the form $p \equiv 1$ mod 4.

Proof. Let S_1 be the set of primes p, such that $p \equiv 1$ mod 4 and let S_2 be the set of primes p, such that $p \equiv 3 \, mod \, 4$. Then, the two residue classes modulo 4 have the same Dirichlet density as given by

$$\lim_{\substack{s\to 1\\ s>1}} \frac{\sum\limits_{p\epsilon S_1} p^{-s}}{\sum\limits_{p\epsilon P} p^{-s}} = \lim_{\substack{s\to 1\\ s>1}} \frac{\sum\limits_{p\epsilon S_2} p^{-s}}{\sum\limits_{p\epsilon P} p^{-s}}.$$

Also since their sum has to be 1, this implies that for each set $d(S_1) = d(S_2) = \frac{1}{2} > 0$. That is, S_1 and S_2 contain infinitely many primes.

1.2. **Riemann zeta function.** The Riemann zeta function, $\zeta(s)$, is a function of a complex variable s, defined by the series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} ,$$

which converges absolutely and uniformly in the domain $\Re(s) \geq 1 + \delta$, for every $\delta > 0$.

It therefore represents an analytic function in the half plane $Re(s) > 1$. The zeta function can also be written as an Euler product

$$\zeta(s) = \prod_p \Big(1 - \frac{1}{p^s}\Big)^{-1},$$

where the product is over the set of all the primes. The product formula shows that $\zeta(s) \neq 0$ for $Re(s) > 1$. The zeta function admits a unique meromorphic continuation to the whole complex plane, with a simple pole of residue 1 at $s = 1$, and no other poles. Zeros of the Riemann zeta function $\zeta(s)$ come in two different types. The so-called trivial zeros occur at all negative even integers $s = -2, -4, -6, \ldots$, and nontrivial zeros lie in the critical strip $\{s = \sigma + it \mid 0 < \sigma < 1\}$ (see [6, Ch. VII, §1]).

1.3. **Riemann hypothesis.** The Riemann hypothesis is the conjecture that states the Riemann zeta function has its non trivial complex zeros on the line $Re(s) = \frac{1}{2}$.

The Riemann Hypothesis is equivalent to the assertion that all the roots of the Riemann xi function $\xi(s)$ are real.

$$\xi(s) = \frac{1}{2} s(s-1)\pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s)$$

for $s \in \mathbb{C}$, where $\Gamma(s)$ is the Gamma function. The functional equation for ξ is,

$$\xi(s) = \xi(1-s).$$

The Riemann Hypothesis if true has important implications for the distribution of prime numbers (see section **??** and [6, Ch. VII, §1, P. 432]).

1.4. **Generalised Riemann Hypothesis.** When the Riemann hypothesis is formulated for Dirichlet L-functions, it is known as the Generalized Riemann hypothesis (GRH). A Dirichlet character is a completely multiplicative function χ on the units of $\mathbb{Z}/k\mathbb{Z}$ such that that there exists a positive integer k with $\chi(n+k) = \chi(n)$ for all n and $\chi(n) = 0$ whenever $(n,k) \neq 1$. The corresponding Dirichlet L-function is defined as

$$L(\chi, s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

for every $s \in \mathbb{C}$ such that $\Re(s) > 1$. The function has an analytic continuation and can be extended to a meromorphic function on the whole complex plane. The generalized Riemann hypothesis asserts that, for every Dirichlet character χ and every complex number s with $L(\chi, s) = 0$, if s is not a negative real number, then the real part of s is $\frac{1}{2}$. If $\chi(n) = 1$ for all n, it gives the ordinary Riemann hypothesis.

1.5. **Prime Number Theorem. Theorem**. Let $\pi(x)$ be the number of primes $\leq x$, for any real number x. Then

$$\pi(x) \sim \frac{x}{\log x}\ , \text{ as } \ x \to \infty$$

where we write $f(x) \sim g(x)$ as x $\to \infty$, if $\frac{f(x)}{g(x)} \to 1$ as x $\to \infty$.

The Prime Number Theorem gives an asymptotic approximation of the number of prime numbers less than a given real number. However, Gauss noted that $\pi(x)$ is better approximated by the logarithmic integral $Li(x)$ defined by

$$Li(x) = \int_2^x \frac{dt}{\log t}\ ,$$

than by $\frac{x}{\log x}$. Gauss, however, did not manage to prove the prime number theorem. This was achieved much later (1896) by Hadamard and De la Vallée Poussin, inspired by an epoch making paper by Riemann (1859). De la Vallée Poussinn proved that there exists a constant $c > 0$ such that

$$\pi(x) = Li(x) + O\left(xe^{-c\sqrt{\log x}}\right) \qquad as\ x \to \infty.$$

An equivalent statement of Riemann hypothesis is that that for each $\epsilon > 0$ there exists a positive constant C such that

$$|\pi(x) - Li(x)| \leq Cx^{\frac{1}{2}+\epsilon}$$

for all $x \geq 2$ (see [2]). Note that the error term in the result of De la Vallée Poussin is much larger than $x^{\frac{1}{2}+\epsilon}$. Thus, the Riemann Hypothesis if true implies that $\pi(x)$ is very well approximated by the logarithmic integral, which makes many calculations involving primes much easier.

2. Arithmetic Density Theorems

The results below were originally proved for the Dirichlet density theorem, which is easier to manipulate. However, they are also valid for the natural density, but in this case the proofs require additional techniques, largely due to Hecke. The theorem of Frobenius is a special case of the Chebotarev density theorem. For many applications of the Chebotarev density theorem it suffices to have the Frobenius density theorem, which is both older (1880) and easier to prove than the Chebotarev density theorem (1922).

2.1. **Frobenius Density Theorem. Theorem.** Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial of degree n. Assume that the discriminant $\Delta(f)$ is non-zero, so that $f(x)$ has n distinct zeros $\alpha_1, \alpha_2, \ldots, \alpha_n$ in a suitable extension field of the field $\mathbb{Q}$. Let $K = \mathbb{Q}(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be the extension field generated by these zeroes. Let G be the Galois group of $f(x)$, that is, the group of field automorphisms of K. Each $\sigma \in G$ permutes the zeroes $\alpha_1, \alpha_2, \ldots, \alpha_n$ of f, and is completely determined by the way in which it permutes these zeroes. Hence, G may be considered as a subgroup of the group S_n of permutations of n symbols. Each element $\sigma \in G$, can be written as a product of disjoint cycles, including the cycles of length 1. The cycle lengths, which forms a partition of n, give a cycle pattern of σ. If p is a prime number not dividing $\Delta(f)$, then we can write f mod p as a product of distinct irreducible factors over $\mathbb{F}_p$. The degrees of these irreducible factors form the decomposition type of f modulo p; which is also a partition of n. The Frobenius theorem states that the density of the number of prime numbers p, with a given decomposition type, is proportional to the number of $\sigma \in G$ with the same cycle pattern.

Theorem Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial of degree n. Denote by G the Galois group of f and let N be the total number of elements in G. The natural density of the set of primes p for which f modulo p has a given decomposition type $n_1, n_2, \ldots, n_i$, exists, and is equal to $\frac{1}{N}$

times the number of $\sigma \in G$ with decomposition in disjoint cycle of the form $c_{n_1}, c_{n_2}, \ldots c_{n_i}$, where c_{n_k} is an n_k-cycle. (see [7, P. 36-38])

Corollary The set of primes p, for which f modulo p splits completely into linear factors has natural density $1/\#G$.

Proof. Only the identity permutation has the cycle pattern in which all n_i are equal to 1.

Example. Let $f(x) \in \mathbb{Z}[x]$, $f = x^4 + 3x^2 + 7x + 4$. We are interested in the decomposition pattern of f modulo prime numbers p. Modulo $p = 2$, the polynomial f decomposes into two irreducible factors, one linear and one of degree 3, as follows

$$f \equiv x(x^3 + x + 1) \bmod 2.$$

We say the decomposition type of f modulo 2 is $(1,3)$. Similarly, modulo $p = 11$ the polynomial f decomposes as

$$f \equiv (x^2 + 5x + 1)(x^2 - 5x - 4) \bmod 11$$

where the two factors are irreducible over $\mathbb{F}_{11}$. Thus, the decomposition type of f modulo 11 is $(2,2)$. This implies that f is irreducible over the ring of integers. Further experiments suggests that, that the primes with type $(1,3)$ have density $\frac{2}{3}$; that the primes with type $(2,2)$ have density $\frac{1}{4}$; that no prime at all exists with the type 4 or with type $(1,1,2)$; and, to make the densities add up to 1, that the primes with type $(1,1,1,1)$ have density $\frac{1}{12}$. This indicates that the Galois group for the polynomial has order 12 that is, it is the alternating group A_4. The alternating group A_4 contains, in addition to the identity element, eight elements of type $(1,3)$ and three elements of type $(2,2)$. This explains the fractions $8/12 = 2/3$ and $3/12 = 1/4$.

2.2. **Ramification and Frobenius elements.** For a finite field extension $\mathbb{K}/\mathbb{Q}$, let $O_{\mathbb{K}}$ be the ring of integers of the algebraic number field $\mathbb{K}$. For a prime $p \in \mathbb{Z}$, the ideal pO_K of O_K can be written uniquely as

$$pO_K = \mathfrak{p}_1^{e_1} \ldots \mathfrak{p}_k^{e_k}$$

where $\mathfrak{p}_i$ are distinct prime ideals in O_K (see [6, Theorem 3.3, Ch. I, §3]). The prime $p \in \mathbb{Z}$ is said to ramify in $\mathbb{K}$ if $e_i > 1$ for some i, and otherwise it is said to be unramified.

Let $\mathbb{K}$ be a finite Galois extension of $\mathbb{Q}$ with $G = Gal(\mathbb{K}/\mathbb{Q}$. Let p be a prime and $\mathfrak{p}$ be a prime ideal in the ring of integers O_K lying over p. For each $\mathfrak{p}|p$, the subgroup $D_\mathfrak{p}$ of $Gal(\mathbb{K}/\mathbb{Q})$ fixing $\mathfrak{p}$ is called the decomposition group of $\mathfrak{p}$. The extension $\mathbb{F}_\mathfrak{p}/\mathbb{F}_p$ is necessarily Galois and its Galois group is cyclic generated by the Frobenius automorphism $Frob_p : x \to x^p$ of $\mathbb{F}_\mathfrak{p}$. There is a natural surjective map $D_\mathfrak{p} \to Gal(\mathbb{F}_\mathfrak{p}/\mathbb{F}_p)$; and it's injectivity is equivalent to the assertion that $\mathfrak{p}$ is unramified in $\mathbb{K}/\mathbb{Q}$. There are finitely many primes p such that they ramify in $\mathbb{K}$.

$$G \supseteq D_\mathfrak{p} \hookrightarrow Gal(\mathbb{F}_\mathfrak{p}/\mathbb{F}_p) \equiv < Frob_p > .$$

$$\sigma_\mathfrak{p} \hookrightarrow (Frob_p : x \to x^p).$$

Suppose p is unramified in $\mathbb{K}/\mathbb{Q}$. Then the automorphism $\sigma_\mathfrak{p} \in D_\mathfrak{p}$ whose image in $Gal(\mathbb{F}_\mathfrak{p}/\mathbb{F}_p)$ is $Frob_p$ is a unique element of the Galois Group G called Frobenius element for $\mathfrak{p}|p$. For all prime ideals $\mathfrak{p}'|p$, the Frobenius elements $\sigma_\mathfrak{p}$ and $\sigma_{\mathfrak{p}'}$ are conjugate in G. The conjugacy class of the Frobenius element $\sigma_\mathfrak{p} \in G$ is the Frobenius class of p, denoted by $Frob_p$. It is common to refer to $Frob_p$ as a Frobenius element $\sigma_p \in G$ representing its conjugacy class (so $\sigma_\mathfrak{p} = \sigma_p$ for some $\mathfrak{p}|p$). Thus $Frob_p = \{\sigma_\mathfrak{p} : \mathfrak{p}|p\}$ is only determined up to conjugacy. If G is abelian then its conjugacy classes all consist of a single element, in which we case $Frob_p$ is a singleton set and there is a unique choice for σ_p.

2.3. **Chebotarev density theorem. Theorem.** Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial with degree n. Assume that the discriminant $\Delta(f)$ of f does not vanish so that f has distinct roots in suitable extension field of $\mathbb{Q}$. Let C be a conjugacy class of the Galois group G of f. Let S be the set of primes p, not dividing $\Delta(f)$ for which σ_p belongs to C,

$$S = \{p \mid \sigma_p \in C \subset G\}.$$

Then, S has a natural density and, $\delta(S) = \#C/\#G$. (see [6, Ch. VII, §13])

2.4. **Relation between Dirichlet, Frobenius and Chebotarev Density Theorems.** Let us consider the polynomial $f = x^m - 1$. It's splitting field $\mathbb{K}$ is the cyclotomic extension obtained by adjoining an m^{th} root of unity ζ_m to $\mathbb{Q}$, that is $\mathbb{K} = \mathbb{Q}(\zeta_m)$. The Galois group $G = Gal(\mathbb{K}/\mathbb{Q})$ has order $\phi(m)$ and is isomorphic to the group $(\mathbb{Z}/m\mathbb{Z})^*$ via the isomorphism,

$$(\mathbb{Z}/m\mathbb{Z})^* \longrightarrow Gal(\mathbb{K}/\mathbb{Q})$$

$$j \longrightarrow [\sigma_j : \zeta \mapsto \zeta^j].$$

The automorphisms σ_j that sends ζ to ζ^j with j and m co-prime, form a group that corresponds to $j \bmod m \in (\mathbb{Z}/m\mathbb{Z})^*$. Note that the Galois group G is abelian, the Frobenius substitution σ_p is a well-defined element of G, not just up to conjugacy. If p is a prime number not dividing m, then the Frobenius substitution σ_p is the element of G that under the isomorphism $G \cong (\mathbb{Z}/m\mathbb{Z})^*$ corresponds to $(p \bmod m)$.

Rephrasing the Dirichlet density theorem, let $f = x^m - 1$, then the set of primes p, such that the Frobenius substitution σ_p is equal to a given element of G, has a density and is equal to $1/\#G$. In other words, the Frobenius elements are equidistributed over the galois group if p varies over all primes not dividing m.

We can also formulate the Dirichlet density theorem using the Chebotarev density theorem which is as follows. Since $C_\sigma = \{\sigma\}$ for all $\sigma \in Gal(\mathbb{K}/\mathbb{Q})$, $\#C_\sigma = 1$. Then the set of primes p, not dividing m, such that $p \equiv a \bmod m$, has a density and is equal to $1/\phi(m)$. (Since $\#G = \phi(m)$).

Let $f = x^{12} - 1$. We find that the decomposition type depends only on the residue class of p modulo 12,

$p \equiv 1 \bmod 12 \ : \quad 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1$

$p \equiv 5 \bmod 12 \ : \quad 1, 1, 1, 1, 2, 2, 2, 2$

$p \equiv 7 \bmod 12 \ : \quad 1, 1, 1, 1, 1, 1, 2, 2, 2$

$p \equiv 11 \bmod 12 \ : \quad 1, 1, 2, 2, 2, 2, 2.$

Since the four decomposition type are pairwise distinct, the Frobenius theorem implies the special case $m = 12$ of Dirichlet's theorem. However, this does not work in all cases. For example, consider $m = 8$

$p \equiv 1 \bmod 8 \ : \quad 1, 1, 1, 1, 1, 1, 1, 1$

$p \equiv 5 \bmod 8 \ : \quad 1, 1, 1, 1, 2, 2$

$p \equiv 7 \bmod 8 \ : \quad 1, 1, 2, 2, 2$

$p \equiv 11 \bmod 8 \ : \quad 1, 1, 2, 2, 2.$ In this case, the Frobenius density theorem does not distinguish between the residue class 7 mod 8 and 11 mod 8, since these classes belong to the same division.

To differentiate between these residue classes, we need a stronger version of theorem. The Chebotarev density theorem provides this. Here the Galois group G is divided into conjugacy classes instead of the divisions based on the cycle pattern of the Frobenius element as in the Frobenius density theorem. We say, that two elements in G belong to the same class e if the cyclic subgroups that they generate are conjugate in G. Thus, the

Chebotarev density theorem directly implies the Frobenius density theorem but not vice-versa.

3. The Artin Primitive Root Conjecture

Let m be a positive integer. A primitive root modulo m is an integer a, such that every number co-prime to m can be expressed as a power of a modulo m. In other words, let $(a, m) = 1$, then a is called primitive root modulo m if and only if order$(a) = \phi(m)$. If there exists a primitive root modulo m, then the group of units, $(\mathbb{Z}/m\mathbb{Z})^*$ is generated by the primitive root modulo m, $(\mathbb{Z}/m\mathbb{Z})^* \equiv \langle a \bmod m \rangle$.

3.1. **Formulation of Artin's Primitive Root Conjecture.** Let $a \in \mathbb{Q}$ and $a \neq \{-1, 0, 1\}$. Denote by $S(a)$ the set of prime numbers p such that a is a primitive root modulo p, and set

$$S(a)(x) = \#\{p \in S(a) : p \leq x\}.$$

Let h be the largest integer such that $a = a_0^h$ where $a_0 \in \mathbb{Q}$. Artin's primitive root conjecture states that

- $S(a)$ is infinite if a is not a square of a rational number.
- as $x \to \infty$

$$S(a)(x) = \prod_{\ell \nmid h} \Big(1 - \frac{1}{\ell(\ell-1)}\Big) \prod_{\ell | h} \Big(1 - \frac{1}{\ell - 1}\Big) \frac{x}{logx} + o\Big(\frac{x}{logx}\Big)$$

$$S(a)(x) = A(h)\frac{x}{logx} + o\Big(\frac{x}{logx}\Big),$$

where $A(h)$ is a positive rational multiple of A. The constant

$$A(1) = A = \prod_{\ell} \Big(1 - \frac{1}{\ell(\ell-1)}\Big) = 0.373955813\ldots$$

is known as the Artin's Constant.

If h is even, $S(a)$ is finite and $A(h) = 0$. Thus the condition a is not a perfect square is necessary for $S(a)$ to be infinite. When h is odd, (see [1]). We can also find more information on Artin's Conjecture in [5].

3.2. **Heuristic Approach.** We make the following observation,

Lemma 3.2 We have, $p \in S(a)$ if and only if $a^{\frac{p-1}{\ell}} \not\equiv 1 \bmod p$ for every prime ℓ dividing $p-1$.

Proof $\Longrightarrow$ By definition.

$\Longleftarrow$ Let $p \notin S(a)$. Then $a^{\frac{p-1}{k}} \equiv 1 \bmod p$ for some divisor k of $p-1$. That is $a^{\frac{p-1}{\ell}} \equiv 1 \bmod p$ for some prime divisor ℓ of k. This is a contradiction.

Lemma. Let $\ell \mid p-1$. If the congruence $a^{\frac{p-1}{\ell}} \equiv 1 \bmod p$ holds, then the equation $x^\ell \equiv a \bmod p$ has a solution modulo p.

Proof Write $a = b^m$ where b is a primitive root mod p. Then, $b^{\frac{m(p-1)}{\ell}} \equiv 1 \bmod p$. A power of a primitive root can only be congruent to 1 mod p if the exponent is a multiple of $p-1$, hence we have $\ell \mid m$. Let $m = n\ell$, then $(b^n)^\ell \equiv a \bmod p$ and thus $x^\ell \equiv a \bmod p$ has a solution mod p.

Let us look at primes p such that,

$$p \equiv 1 \bmod \ell,\ a^{\frac{p-1}{\ell}} \equiv 1 \bmod p,$$

where ℓ is a fixed prime. Consider $f = x^\ell - 1$, then for primes p, such that $p \equiv 1 \bmod \ell$, f splits completely modulo l. That is, $x^\ell \equiv 1 \bmod p$ has ℓ distinct solutions. Let $\alpha_1, \alpha_2, \ldots \alpha_\ell$ be the distinct roots of f modulo p. Then using the lemma above we have, $\alpha_1 b^\ell, \alpha_2 b^\ell, \ldots \alpha_\ell b^\ell$ are the distinct roots of $x^\ell \equiv a \bmod p$.

We conclude that the prime p that satisfies $p \equiv 1 \bmod \ell$, $a^{\frac{p-1}{\ell}} \equiv 1 \bmod p$ splits completely in the number field, $\mathbb{K}_\ell = \mathbb{Q}(\zeta_\ell, a^{\frac{1}{\ell}})$.

Corollary Let p be a prime number. For all primes ℓ, such that $\ell \mid p-1$, a is a primitive root modulo p if and only if the prime p does not split completely in $\mathbb{K}_\ell = \mathbb{Q}(\zeta_\ell, a^{\frac{1}{\ell}})$.

Let us denote the degree of $\mathbb{K}_\ell$ by $n(\ell)$. Then we expect the density of the set of primes such that a is a primitive root modulo p is $\prod_\ell \left(1 - \frac{1}{n(\ell)}\right)$. But, if $\ell \mid h$, then $a^{\frac{p-1}{\ell}} = a_0^{\frac{h(p-1)}{\ell}} \equiv 1 \bmod p$, holds trivially and the density equals $1 - \frac{1}{\ell-1}$ and $\mathbb{K}_\ell = \mathbb{Q}(\zeta_\ell)$. Thus, assuming the various conditions on ℓ to be independent, we expect the natural density as,

$$\prod_{\ell \nmid h} \left(1 - \frac{1}{\ell(\ell-1)}\right) \prod_{\ell \mid h} \left(1 - \frac{1}{\ell - 1}\right) = A(h).$$

This is a heuristic, and some mathematical rigorous analysis is provided later.

Proposition. Let $a \in \mathbb{Q}^*$ be a square or $a = -1$. Then $S(a)$, the set of primes p such that a is a primitive root modulo p, is finite. That is the density of $S(a)$ is equal to 0.

Proof. Let a be a square, and let p be an odd prime co-prime to a. Since a is a square in $\mathbb{Q}^*$, then it's multiplicative order modulo p is at most $(p-1)/2$. Therefore, a square cannot be a primitive root modulo p, unless $p = 2$

The integer -1 is an element of order 2 in $\mathbb{Q}^*$. If $p \neq 2$, is an odd prime then the equality $\langle -1 \rangle = \{-1, 1\}$ holds, and it implies that -1 is a primitive rood modulo $p = \{2, 3\}$.

Thus, for such values of a, the set $S(a)$, is finite and hence the corresponding density is 0.

Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial and p be a prime number. The Frobenius density theorem tells us that the decomposition pattern of f modulo p is same as the cycle pattern of the Frobenius element σ_p. So, f has a root modulo p means f modulo p has a linear factor and thus the Frobenius element has a cycle of length 1, that means it keeps one element fixed. So instead of counting the number of primes p, such that f has a root modulo p, we can count the number of elements of the galois group that leave an element fixed.

Property Let $\ell \in \mathbb{N}$ be a prime and $a, b \in \mathbb{Z}/\{0\}$, $a > 0$, p be a prime such that $p \nmid a$. Assuming $p \neq 2$ and $\frac{b}{a}$ is not a ℓ^{th} power, the set

$$S = \{p \mid ax^\ell - b \text{ has a root } \mod p\}$$

has a natural density, equal to $\frac{\ell-1}{\ell}$.

Proof Let $f = x^\ell - \frac{b}{a}$ and let $\rho = \sqrt[l]{\frac{b}{a}}$ be any ℓ^{th} root of $\frac{b}{a}$. The roots of the polynomial f are $\rho\zeta^k$, $k = 0, 1, \ldots \ell - 1$, where ζ is a primitive ℓ^{th} root of unity. Let $\mathbb{K}$ be the splitting field of f. Then $\mathbb{K} = \mathbb{Q}(\rho, \zeta)$ and the degree of extension $[\mathbb{K} : \mathbb{Q}] = \ell(\ell - 1)$. Let $G = Gal(\mathbb{K}/\mathbb{Q})$ be the galois group of f with order $\ell(\ell-1)$ and each element $\sigma \in G$ is an automorphism that sends an element to its conjugate (a root of the minimal polynomial of the element). The minimal polynomial of ρ is $x^\ell - \frac{b}{a}$ and the minimal polynomial of ζ is the cyclotomic polynomial,

$$\frac{x^\ell - 1}{x - 1} = x^{\ell-1} + x^{\ell-2} + \ldots + x + 1.$$

Fix $\sigma \in G$. Hence σ maps

$$\rho \to \rho\zeta^e$$

$$\zeta \to \zeta^u$$

where $e \in \mathbb{F}_\ell$ and $u \in \mathbb{F}_\ell^*$. Since the order of G is $\ell(\ell-1)$. These are exactly the elements of the Galois group G. In general we have,

$$\sigma(\rho\zeta^j) = \rho\zeta^{uj+e}.$$

Given u and e, we are interested in deciding whether there exits an integer j such that

$$\sigma(\rho\zeta^j) = \rho\zeta^{uj+e} = \rho\zeta^j.$$

This is equivalent to deciding whether the system of congruences

$$uj + e \equiv j \bmod \ell$$

$$(u-1)j \equiv -e \bmod \ell$$

has a solution. The identity map ($u = 1, e = 0$) is one such map and for $u \neq 1$,

$$j \equiv \frac{-e}{u-1} \bmod \ell$$

there exists such j, if and only if $u - 1$ is invertible and e be any element in $\mathbb{F}_\ell$. There are $\ell - 2$ invertible elements for $u - 1$, so the total number of maps that fix an element is $\ell(\ell-2)+1$. We conclude that the density of set of the primes p such that f has a root modulo p is $\frac{\ell(\ell-2)+1}{\#G} = \frac{(\ell-1)^2}{\ell(\ell-1)} = \frac{\ell-1}{\ell}$.

For ℓ prime, let $S(a, \ell)$ be defined by

$$S(a, \ell) = \{p \mid \ a^{\frac{p-1}{\ell}} \not\equiv 1 \bmod p\}.$$

Then,

$$\delta(S(a,\ell)) = \begin{cases} \left(1 - \frac{1}{\ell(\ell-1)}\right), & \text{if } \ell \nmid h \\ \left(1 - \frac{1}{\ell-1}\right), & \text{otherwise} \end{cases} \tag{3.1}$$

The density $\delta(S(a))$ as proposed by Lemma **??** if it exists, would therefore be equal to,

$$\delta(S(a)) = \ \delta(\bigcap_\ell S(a,\ell))$$

where the intersection is taken over all primes ℓ. We say that two sets $S(a, \ell_1)$ and $S(a, \ell_2)$ are independent if the set of primes p in $S(a, \ell_1)$ do not affect the set of primes in $S(a, \ell_2)$. Assuming that the sets $S(a, \ell)$ are independent of each other, Artin conjectured that the density of $S(a)$ satisfies,

$$\delta(S(a)) = \prod_\ell \delta(S(a,\ell)) = A,$$

We observe that a prime splits completely in $\mathbb{K}_{\ell_1}$, $\mathbb{K}_{\ell_2}$, …, $\mathbb{K}_{\ell_n}$ if and only if it splits completely in $\mathbb{K}_{\ell_1\ell_2\dots\ell_n}$. If k is odd, then $\mathbb{K}_\ell$ for $\ell \mid k$ are completely linearly disjoint and we have,

$$n_k = \prod_{\ell \mid k} n_\ell$$

where n_ℓ is the degree of $\mathbb{K}_\ell$ over $\mathbb{Q}$. For even subscripts, $n_{2k} = n_k$ or $2n_k$ depending on whether $\sqrt{a}$ is contained in $\mathbb{Q}(\zeta_k)$. For example, consider the case $a = 5$, then $\sqrt{a} \in \mathbb{Q}(\zeta_5)$.

$$[\mathbb{K}_{10} : \mathbb{Q}] = 20 \neq 2 \cdot 20 \neq [\mathbb{K}_2 : \mathbb{Q}][\mathbb{K}_5 : \mathbb{Q}].$$

If $a = a_1 a_2^2$ with a_1 be squarefree, then $\sqrt{a}$ will be contained in $\mathbb{Q}(\zeta_k)$ with k squarefree, if $a_1 \mid k$ and $a_1 \equiv 1 \bmod 4$. Artin suggested the correction factor,

$$1 - \mu(a_1) \prod_{\ell \mid a_1} \frac{1}{n_\ell - 1}$$

where

$$\mu(k) = \begin{cases} (-1)^r, & \text{if } k \text{ is a square free number having } r \text{ prime factors} \\ 0, & \text{otherwise} \end{cases} \tag{3.2}$$

and the corrected formula is,

$$\delta(S(a)) = A'(h) = \begin{cases} A(h), \ \text{if } a_1 \not\equiv 1 \bmod 4 \\ A(h)\left(1 - \mu(a_1) \prod_{\substack{\ell \mid a_1 \\ \ell \mid h}} \left(\frac{1}{\ell-2}\right) \prod_{\substack{\ell \mid a_1 \\ \ell \nmid h}} \left(\frac{1}{\ell^2-\ell-1}\right)\right), \\ \text{otherwise.} \end{cases} \tag{3.3}$$

Using the General Riemann Hypothesis, Hooley [1] proved the corrected formula of Artin for the density of primes for which a is a primitive root modulo p, namely

$$\#\{p \mid p \in S(a) \text{ and } p \leq x\} = A'(h) \cdot \frac{x}{\log x} + \mathcal{O}\left(\frac{x \log\log x}{\log^2 x}\right),$$

where $\delta(S(a))$ is given by (3.3) (see [4]).

Acknowledgement

This article is a result of a visit of the author for an internship under Prof. Peter Jossen of ETH Zürich in May-July 2019. The author would like to thank Prof. Peter Jossen for his guidance and support.

References

[1] Hooley, Christopher *On Artin's Conjecture,* Journal für die reine und angewandte Mathematik, Vol 225 (1967), 209-220.

[2] Lagarias, Jeffrey C *An Elementary Problem Equivalent to the Riemann Hypothesis,* Amer. Math. Monthly 109, no. 6 (2002), 1-2.

[3] Lang, Serge *Algebraic Number Theory*, Second Edition, Springer-Verlag (1994).

[4] Moree, Pieter *Artin's Primitive Root Conjecture - a survey*, Integers, Vol 12, no. 6 (2012), 1305-1416.

[5] Murty, M. Ram *On Artin's Conjecture*, Journal of Number Theory, Vol 16 (1983), 147-168.

[6] Neukirch, Jürgen *Algebraic Number Theory*, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, 322 (1999).

[7] Pesiri, Alfonso *The Chebotarev Density Theorem and its Applications,* Graduation thesis in Mathematics, Universitá Delgi Sudi Roma Tre (2007), 1-46.

[8] Stevenhagen, Peter; Lenstra, Hendrik W. *Chebotarev and his Density Theorem*, The Mathematical Intelligencer, Vol 18, no.2 (1996), 26-37.

Gauri Gupta
Undergraduate Student
Department of Mathematics
Indian Institute of Technology Delhi
Haus Khas, New Delhi -110016
E-mail: gaurigupta.iitd@gmail.com

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 59–62

TWO IMPORTANT THEOREMS THAT ARE REALLY ONE

ROBERT F. BROWN
(Received : 31 - 03 - 2020 ; Revised : 15 - 04 - 2020)

Abstract. An elementary proof of the equivalence of the Intermediate Value Theorem and a fixed point theorem on the real line illustrates the equivalence of the Brouwer Fixed Point Theorem and the Poincaré-Miranda Theorem.

Section 2.8 of the widely used textbook *Single Variable Calculus* by John Rogawski and Colin Adams [4] is devoted to a theorem that the book states this way:

Theorem 1. *(Intermediate Value Theorem) If the function $f(x)$ is continuous on $[a, b]$ with $f(a) \neq f(b)$ and y is between $f(a)$ and $f(b)$, then there exists $c \in (a, b)$ such that $f(c) = y$.*

The point y is "between" $f(a)$ and $f(b)$ means either that $f(a) < y < f(b)$ or that $f(a) > y > f(b)$, depending on how $f(a)$ and $f(b)$ are related to each other.

In order to avoid "between", with its two cases, the theorem is sometimes stated in this neater way:

Theorem 2. *(Intermediate Value Theorem) If the function $g(x)$ is continuous on $[a, b]$ and $g(a)g(b) < 0$, then $g(c) = 0$ for some $c \in (a, b)$.*

What do I mean by claiming that Theorem 2 is also the Intermediate Value Theorem? Theorems 1 and 2 are the same theorem in the sense that they are *equivalent*, which means that each can be proved directly from the other. Certainly Theorem 2 is a special case of Theorem 1 since one of $g(a)$ and $g(b)$ is positive and the other is negative. But if you assume

2010 Mathematics Subject Classification: 54H25
Key words and phrases: Intermediate Value Theorem, Brouwer Fixed point Theorem, Poincaré-Miranda Theorem

that Theorem 2 is true, then so is Theorem 1 because if $f(x)$ satisfies the hypotheses of Theorem 1, then $g(x) = f(x) - y$ satisfies the hypotheses of Theorem 2 and $0 = g(c) = f(c) - y$ means that $f(c) = y$.

In that textbook, Problem 32 of Section 2.8 is to prove

Theorem 3. *Assume that $f(x)$ is continuous and that $a \leq f(x) \leq b$ for $a \leq x \leq b$. Then $f(c) = c$ for some $c \in [a, b]$.*[1]

Proof. If $f(a) = a$ or $f(b) = b$ there is nothing to prove so we can assume $f(a) > a$ and $f(b) < b$, then $g(x) = f(x) - x$ satisfies the hypotheses of Theorem 2 and $g(c) = 0$ implies that $f(c) = c$. □

The point c such that $f(c) = c$ is called a *fixed point* of the function f. Thus Theorem 3 is an example of a *fixed point theorem.*

Now we'll go a bit beyond what is in [4]. The following fixed point theorem has somewhat more general hypotheses because, rather than functions that take $[a, b]$ to itself, it concerns a class of function from $[a, b]$ to the real line.

Theorem 4. *If a function $f(x)$ is continuous on $[a, b]$ with $f(a) \geq a$ and $f(b) \leq b$, then there exists $c \in [a, b]$ such that $f(c) = c$.*

Proof. Again there is nothing to prove unless $f(a) > a$ and $f(b) < b$. Let $g(x) = x - f(x)$, then $g(a)g(b) < 0$ so by Theorem 2 there exists $c \in [a, b]$ such that g(c) = c - f(c) = 0 and thus $f(c) = c$. □

So Theorem 4 is also a consequence of the Intermediate Value Theorem. But the authors of [4] could have gone a step further, because we will show by a similar sort of argument that this fixed point theorem in turn *implies* the Intermediate Value Theorem and thus those theorems are equivalent.

Theorem 5. *Given that for a function $g(x)$ continuous on $[a, b]$ with $g(a) \geq a$ and $g(b) \leq b$ there exists $c \in [a, b]$ such that $g(c) = c$, then $g(a)g(b) < 0$ implies $g(c) = 0$ for some $c \in (a, b)$.*

Proof. If $g(a) > 0$ and $g(b) < 0$, let $f(x) = g(x) + x$ which satisfies the hypotheses of Theorem 4 so $f(c) = g(c)+c = c$ and thus $g(c) = 0$. The other possibility is that $g(a) < 0$ and $g(b) > 0$ in which case let $f(x) = x - g(x)$

[1] In [4] the problem is stated for $a = 0$ and $b = 1$, but this more general notation is consistent with the way that we are stating the other results and the solution is really the same either way.

and then $f(c) = c - g(c) = c$ means $g(c) = 0$. Therefore, the fixed point theorem implies the Intermediate Value Theorem and consequently these theorems are equivalent. □

Theorem 3 is a special case of our first "important theorem", as promised in the title: the Brouwer Fixed Point Theorem. Instead of the function $f(x)$ that takes the real line to itself, the theorem discovered by L. E. J. Brouwer early in the 20th century concerns a function that takes Euclidean n-space $\mathbb{R}^n$ to itself.

Theorem 6. *(Brouwer Fixed Point Theorem) Let*

$$B^n = \{x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \colon a_i \leq x_i \leq b_i, i = 1, 2, \ldots, n\}.$$

Suppose $f(x)$ is continuous and $f(x) \in B^n$ for all $x \in B^n$, then there exists $c \in B^n$ such that $f(c) = c$.

Here is some evidence that the Brouwer Fixed Point Theorem is important: it has been mentioned in more than 500 research papers. Since that list includes more than 50 papers published since 2015, we can see that this theorem is related to research topics of current interest and that is why it is important.

Theorem 3, the solution to Problem 32, is the $n = 1$ case of the Brouwer Fixed Point Theorem. The Intermediate Value Theorem is also the $n = 1$ case of a more general theorem, one that was stated by Henri Poincaré in 1883. It also concerns a function $g(x)$ that takes $\mathbb{R}^n$ to itself. We can think of such a function as a vector-valued function $g(x) = (g_1(x), g_2(x), \ldots, g_n(x))$.

Theorem 7. *Suppose $g(x) = (g_1(x), g_2(x), \ldots, g_n(x))$ is continuous on B^n such that*

$$g_i(x_1, \ldots x_{i-1}, a_i, x_{i+1}, \ldots, x_n) \leq 0$$

and

$$g_i(x_1, \ldots x_{i-1}, b_i, x_{i+1}, \ldots, x_n) \geq 0$$

where $a_i < b_i$ for $i = 1, 2, \ldots, n$, then $g(c) = g(c_1, c_2, \ldots, c_n) = (0, 0, \ldots, 0) = \mathbf{0}$ for some $c \in \mathbb{R}^n$.

In 1940, Carlo Miranda [3] proved that Poincaré's theorem is equivalent to the Brouwer fixed point theorem. As a result, Theorem 6 became known as the Poincaré-Miranda Theorem[2], and that is our other "important theorem". The paper [3] has appeared in the list of references of over 120 research papers, more than half of them published since 2015, so it is very much a part of contemporary mathematical research.

An elegant demonstration of the equivalence of the Brouwer Fixed Point Theorem and the Poincaré-Miranda Theorem has recently been published [2].

Thus, by proving that the fixed point theorem of Problem 32 of [4] is equivalent to the Intermediate Value Theorem, we have not only an interesting exercise that is related to that familiar theorem, but we have actually proved the $n = 1$ case of Miranda's surprising discovery that the Brouwer Fixed Point Theorem and an n-dimensional version of the Intermediate Value Theorem are equivalent. Those two important theorems state the same mathematical fact in different forms and so, as our title states, they really are one.

Acknowledgements

I thank Phoenix Kim for suggesting improvements in the exposition of this note and I thank the referee for a very prompt and helpful report.

References

[1] W. Kulpa, The Poincaré-Miranda Theorem, *Amer. Math. Monthly* **104**(1997), 545 - 550.

[2] J. Mawhin, Simple proofs of the Hadamard and Poincaré-Miranda theorems using the Brouwer fixed point theorem, *Amer. Math. Monthly* **126**(2019), 260 - 263.

[3] C. Miranda, Un'osservazione su un teorema di Brouwer, *Boll. Un. Mat. Ital.* **3**, 5 - 7 (1940).

[4] J. Rogawski and C. Adams, *Single Variable Calculus, Third Edition*, Freeman, 2015.

Robert F. Brown
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
e-mail: rfb@math.ucla.edu

[2] To find out more about the Poincaré-Miranda Theorem, in particular how it is proved, see [1]

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 63–85

A SYMMETRIC RANDOM WALK ON THE VERTICES OF A HEXAHEDRON

JYOTIRMOY SARKAR
(Received : 17 - 06 - 2019 ; Revised : 25 - 01 - 2020)

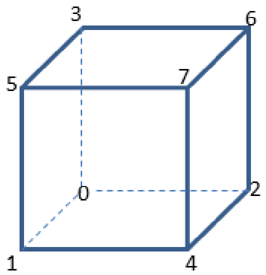
Abstract. A symmetric random walk on the vertices of a hexahedron (cube) starts at a vertex called the origin; and at each step it moves, to an adjacent vertex with probability 1/3 each. We find the means and the standard deviations of the number of steps needed to (1) return to origin, (2) visit all vertices, and (3) return to origin after visiting all vertices. We also find (i) the number of vertices visited before return to origin, (ii) the last vertex visited, and (iii) the number of vertices visited while returning to origin after visiting all vertices.

1. Introduction

Let $\{0, 1, 2, 3, 4, 5, 6, 7\}$ denote the vertices of a hexahedron (commonly known as a cube), with Vertex 0 adjacent to Vertices $1, 2, 3$; and every pair of diametrically opposite vertices adding up to 7. See Fig. 1. The 12 edges of the cube are given by the 0-1 adjacency matrix $C = (c_{ij})$, where $c_{ij} = 1$ if and only if there is an edge joining Vertex i and Vertex j, in which case we say that Vertices i and j are adjacent. On the contrary, four pairs of main diagonals, and 12 pairs of face diagonals are not directly connected.

We consider a symmetric random walk (RW) on the vertices of a hexahedron. Let $X(t)$ denote the location of a RW at epoch $t = 0, 1, 2, \ldots$. Without loss of generality, the RW starts at Vertex 0, which is called the origin; that is, $X(0) \equiv 0$. Thereafter, at each successive epoch, from the

2010 Mathematics Subject Classification: 60G50; 60G40
Key words and phrases: Absorbing state, cover time, dual graph, one-step transition, recursive relation, transient state.



$$C = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

FIGURE 1. Labeling the vertices and denoting the edges of a hexahedron

current vertex, the RW moves to one of the three adjacent vertices (including the one the RW came from) with equal probability 1/3 each. This is why the RW is called symmetric. We assume the successive moves are independent.

We study the distribution of the time until the following events happen:

(1) Return to origin (after leaving it),
(2) Visit to all vertices (at least once), and
(3) Return to origin after visiting all vertices.

The epochs when these events occur are respectively called the return time T_R, the cover time $\bar{T}$, and the additional time to return after visiting all vertices ${}_LT_R$, where $L = X(\bar{T})$ denotes the vertex that was visited the last among all vertices. Clearly, $X(T_R) = 0$ and $X({}_LT_R) = 0$, but $X(\bar{T}) = L \neq 0$. In particular, $T_R \neq \bar{T}$.

We also study the probability distribution of N_R, L, and ${}_LN_R$ defined below:

(i) The number of vertices (other than the origin) visited until return to origin is denoted by N_R. That is, N_R is the number of distinct elements among $\{X(1), X(2), \ldots, X(T_R - 1)\}$;
(ii) The last vertex visited among all vertices, or the location of the RW when all vertices have been visited at least once, is denoted by $L = X(\bar{T})$; and
(iii) The number of vertices visited while the RW returns to origin after visiting all vertices is denoted by ${}_LN_R$. That is, ${}_LN_R$ is the number of distinct values among $\{X(\bar{T}), X(\bar{T}+1), \ldots, X({}_LT_R - 1)\}$.

The study of RWs on the vertices of a connected graph has a rich literature. See, for example, Gödel and Jagers [1], who studied the expected recurrence time, first passage time and symmetrized first passage time (or commuting time). Letac and Takács [2] studied the limiting probability that a RW visits a particular vertex of a dodecahedron; and they related the result to the spectrum of a certain finite-dimensional Banach algebra. van Slijpe established the equality of the mean passage times to travel from one vertex to another and vice versa, for distance-regular graphs in [7], and extended the result to vertex-transitive graphs in [8]. Wildeberger [9] expressed the hypergroup structures (or character table) associated with any distance transitive graph, and found the exact probability that the RW returns to the origin after n steps using commutative harmonic analysis.

Sarkar [5] answered all six questions mentioned above in the context of an asymmetric random walk on a polygon. Maiti and Sarkar [3] answered the same questions on a finite linear path and on a finite circular network using mathematical induction, and for a symmetric RW on the vertices of a tetrahedron or an octahedron in [6]. For a RW on the vertices of a hexahedron, these questions are partially answered in [6] based on simulation results, but without rigorous justification. Here we prove those results.

Let us describe the strategy for proving the results: We will draw an appropriate stochastic transition diagram in which the nodes will represent a certain configuration of triplets—the origin, the current vertex and the set of vertices already visited. In all diagrams, to avoid clutter, the origin is hidden from view; an empty circle denotes a vertex already visited, a filled circle denotes the current location of the RW, and a vertex with no circle at all has not been visited yet. Henceforth, in the context of a particular transition diagram, we reserve the word 'vertex' to mean a point where the edges of a hexahedron meet, and the word 'node' to refer to a particular configuration of triplets. In each transition diagram, we will describe the transitions between nodes following an appropriate strategy of renumbering the vertices. Whenever entry into a node signifies mission accomplished (for whichever purpose the diagram is drawn), that node is called an *absorbing state*. For each random variable, we will express the entire probability distribution when possible; otherwise, we will report only the mean and the standard deviation (SD).

Let us introduce some terminologies and notation: The *distance* between two vertices i and j is the number of edges on the *shortest path* connecting i and j. Let ${}_iT_j$ denote the time (or the number of steps) taken by the RW to go from Vertex i to Vertex j. With this notation, without loss of generality (renumbering the vertices, other than the origin, if necessary), we have $T_R = {}_0T_0 = 1 + {}_1T_0$, and ${}_LT_R = {}_LT_0$. Additionally, let us denote the mean, the mean square and the variance of ${}_iT_j$, respectively by ${}_ie_j = E[{}_iT_j]$, ${}_is_j = E[{}_iT_j^2]$, and ${}_iv_j = V({}_iT_j) = {}_is_j - {}_ie_j{}^2$; where E denotes expectation and V denotes variance. The positive square root of the variance is called the SD.

From [6], readers can glean more insight into the above-mentioned strategy to study a symmetric RW. In particular, because the graphical representation of a tetrahedron is a complete graph K_4 on four vertices, the answers to the six questions for a tetrahedron are rather straight-forward. But the answers for an octahedron are relatively more intricate because three pairs of diametrically opposite vertices are not directly connected by an edge. As we will see in this paper, the answers for a hexahedron are even more intricate. We answer two questions each in Sections 2–4 in the order 1, (i); (ii), 2; and 3, (iii). Section 5 concludes the paper with a summary and several directions of future research.

2. Return to Origin

Having left the origin at time 1, without loss of generality, the RW goes to Vertex 1. Thereafter, sooner or later the RW will return to origin. When does that happen, and how many non-origin vertices are visited during the trip? To answer these questions, we obtain the distributions of T_R and N_R.

2.1. **Time to Return to Origin.**

To study T_R, in view of symmetry, it suffices to keep track of only the distance of the current vertex from the origin, and forgo the exact label of the current vertex and the pattern of already visited vertices. A simple transition diagram, depicted in Fig. 2, suffices to study T_R, where each node represents a different distance between the current vertex and the origin.

The RW on the vertices of a hexahedron starts at the origin and surely goes to Vertex 1 (after renumbering, if needed). This gives rise to Node 1. Hence, $T_R = 1 + {}_1T_0$. Next, if the RW moves from Vertex 1 to Vertex 0 (which happens with probability 1/3), then ${}_1T_0 = 1$. But if the RW moves from Vertex 1 to either Vertex 4 or Vertex 5 (which happens with probability 2/3, and which in either case gives rise to Node 2), then ${}_1T_0 = 1 + {}_2T_0$. In this latter case, without loss of generality, assume that the RW

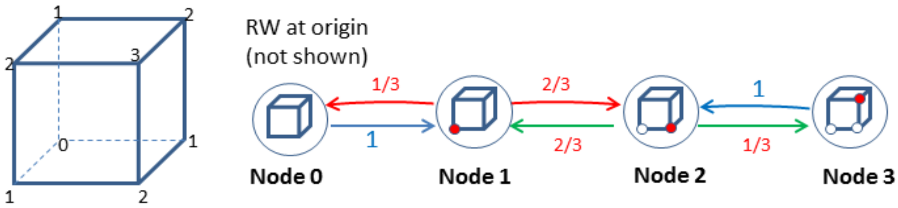


FIGURE 2. Distance of the current vertex from the origin defines these nodes. One-step transition probabilities between nodes suffice to study T_R

has moved to Vertex 4. (The case for Vertex 5 is similar.) If from Vertex 4, the RW moves to Vertex 7 (which happens with probability 1/3, and which gives rise to Node 3), then ${}_2T_0 = 1 + {}_3T_0$. Subsequently, from Node 3 the RW surely moves to Node 2, making ${}_3T_0 = 1 + {}_2T_0$ with probability 1. Hence, with probability 1/3, we have ${}_2T_0 = 2 + {}_2T_0'$, where ${}_2T_0'$ is a random variable independently and identically distributed (IID) as ${}_2T_0$. On the other hand, if from Vertex 4 the RW moves to Vertex 1 or Vertex 2 (which happens with probability 2/3), then the RW returns to Node 1; and then ${}_2T_0 = 1 + {}_1T_0$. Combining the above relations, we have the following system of equations (in distributions)

$$T_R = 1 + {}_1T_0 \tag{2.1}$$

$${}_1T_0 = \begin{cases} 1 & \text{with probability } 1/3 \\ 1 + {}_2T_0 & \text{with probability } 2/3 \end{cases} \tag{2.2}$$

$${}_2T_0 = \begin{cases} 1 + {}_1T_0 & \text{with probability } 2/3 \\ 2 + {}_2T_0{}' & \text{with probability } 1/3 \end{cases} \tag{2.3}$$

where ${}_2T_0$ and ${}_2T_0'$ are IID random variables. Substituting (2.2) in (2.3), we get

$$\begin{aligned} {}_2T_0 &= \begin{cases} 2 & \text{with probability } 2/9 \\ 2 + {}_2T_0{}' & \text{with probability } 7/9 \end{cases} \\ &= 2\,G \end{aligned} \tag{2.4}$$

where G is a geometric(2/9) random variable. Next, substituting (2.4) in (2.2), and then that composite expression in (2.1), we have

$$T_R = \begin{cases} 2 & \text{with probability } 1/3 \\ 2(1+G) & \text{with probability } 2/3 \end{cases} \tag{2.5}$$

Taking expectations in (2.5), since $E[G] = 9/2$ and $E[G^2] = 36$, we get

$$E[T_R] = (1/3) \cdot 2 + (2/3) \cdot 2 \{1 + (9/2)\} = 8$$

$$E[T_R^2] = (1/3) \cdot 4 + (2/3) \cdot 4 \{1 + 9 + 36\} = 124$$

Hence, $V(T_R) = 124 - 8^2 = 60$.

More generally, utilizing (2.5), we get the probability distribution of T_R.

Proposition 1. *The time T_R to return to origin has support $\{2, 4, 6, \ldots\}$, mean 8, mean square 124, variance 60, SD 7.7460, and probability mass function*

$$P\{T_R = 2\} = 1/3$$

$$P\{T_R = 2k + 2\} = (2/3) \cdot (2/9) \cdot (7/9)^{k-1}, \ \textit{for } k = 1, 2, 3, \ldots$$

Remark 1. It is a pleasant surprise that $E[T_R] = 8$, the number of vertices of a hexahedron. In fact, a similar result holds for all regular graphs (for which the degree of each vertex is the same). The intuition behind this phenomenon is that in the long-run (that is, as $t \to \infty$) the RW forgets its starting vertex, and it is equally likely to be at each of the vertices. This in turn means that the expected number of steps between successive visits to any particular vertex is the reciprocal of the probability that the RW is at that vertex.

2.2. **Number of Vertices Visited Before Return to Origin.** To study N_R, the required transition diagram, shown in Fig. 3, preserves all three features of the configurations—the origin, the current vertex and the pattern of visited vertices. But we treat as identical two or more configurations that are rotations and/or reflections of one another. In Fig. 3, we draw the one-step transitions until return to origin, where each arrow represents a conditional probability of 1/3. [To fit into limited page space, the diagram has been pruned when 4 (and then again when 5) non-origin vertices are visited; and the pruned portions are shown in the continuation of the figure. The originating portion and the corresponding pruned portion are similarly labeled to aid the reader glue them back.] Again, to avoid clutter, the origin is not shown; etc. We list below some illustrative (but by no means exhaustive) explanations for the figure.

(1) From the origin, the RW surely goes to Vertex 1 (after renumbering). So, we show three arrows from Node N(0) to Node N(1,1), accounting for a total probability of 1.
(2) From Node N(1,1), the RW returns to the origin [or enters Node N(1,2)] with probability 1/3; and with probability 2/3 it goes to a second non-origin vertex, renamed as Vertex 4 (by appropriate reflection about the plane through Vertices 0, 1, 7, 6). Hence, two arrows go from Node N(1,1) to Node N(2,1) accounting for a probability of 2/3.
(3) Whenever multiple arrows go from one node to another (for instance, two arrows going from Node a1 to Node a2), the reader should think about which rotation or reflection explains the multiplicity.
(4) To recognize similarity of patterns of visited vertices (temporarily ignoring which vertex is current), we may rotate or reflect a naturally occurring pattern. For example, Node b1 is rotated by 120° counterclockwise to recognize its similarity with Node b2. Likewise, Node A1 is rotated by 120° counterclockwise to recognize its similarity with Nodes A2, A3, A4.

Starting from each node, we calculate the probability of reaching either the origin or a new vertex, using Lemma 1 below. Also, we stopped the diagram of the pruned parts when 6 non-origin vertices are visited, since we can, by Lemma 2 below, calculate the conditional probability that the RW visits the origin before it visits the seventh non-origin vertex not visited yet.

Each pruned part consists of three types of nodes: (1) Source nodes that exhibit the same pattern of visited vertices and a matching origin but different location of the current vertex, with associated conditional probabilities of ever reaching these nodes from an upper stem; (2) Intermediate nodes that also exhibit the same pattern as the source nodes, but are reachable only from the source nodes (and not directly from an upper stem); and

(3) Sink nodes that are accessible from the source or the intermediate nodes when the RW reaches either the origin or a new vertex. The RW restricted to the source, the intermediate and the sink nodes, is an absorbing chain in which the source and the intermediate nodes are transient states and the sink nodes are absorbing states. For an absorbing chain, the eventual conditional probabilities of reaching an absorbing state j, starting from a transient state i is given by Lemma 1. For a proof see [4].

Lemma 1. *In an absorbing chain, consisting of t transient and $(k-t)$ absorbing states arranged in order, suppose that the one-step transition probability matrix P is partitioned as $\begin{bmatrix} Q & R \\ O & I \end{bmatrix}$, where Q is a $t \times t$ matrix, R is a $t \times (k-t)$ matrix, O is $(k-t) \times t$ matrix of zeros, and I is a $(k-t) \times (k-t)$ identity matrix. Then, writing O_t as a $t \times t$ matrix of all zeros and I_t as a $t \times t$ identity matrix, we have*

$$\lim_{n\to\infty} P^n = \begin{bmatrix} O_t & (I_t - Q)^{-1}R \\ O & I \end{bmatrix} \tag{2.6}$$

Consequently, starting from a transient state i, the probability that the RW eventually reaches an absorbing state j is given by the (i,j)-th element of $(I_t - Q)^{-1}R$.

In view of Lemma 1, starting from the nodes in the source set (with associated conditional probability row-vector $\mathbf{p}'$ of ever reaching those nodes from an upper stem), we calculate the conditional probability vector of ever reaching the nodes in the sink set, simply as $\mathbf{q}' = \mathbf{p}'\,(I_t - Q)^{-1}R$. Of course, the member of the sink set in which the current vertex is the origin (colored red in the diagrams) is a genuine absorbing state for the overall RW, while the members of the sink set in which the current vertex is a non-origin vertex form the source set for the next lower level pruned portion.

In Fig. 4, to save space, we drop the intermediate sets altogether, and draw arrows from each member of the source set to each member of the sink set; and using Lemma 1, we compute the eventual conditional probabilities of transitions and write them on each arrow. In the final stage, on the short and the long dotted arrows, we write the conditional probabilities of $\{N_R = 6\}$ and $\{N_R = 7\}$ respectively, using Lemma 2.

Lemma 2. *For a hexahedron, let V_0, V_1, V_2 denote any three distinct vertices. Let d_1, d_2, d_3 denote the distances within pairs (V_0, V_1), (V_0, V_2), (V_1, V_2).*

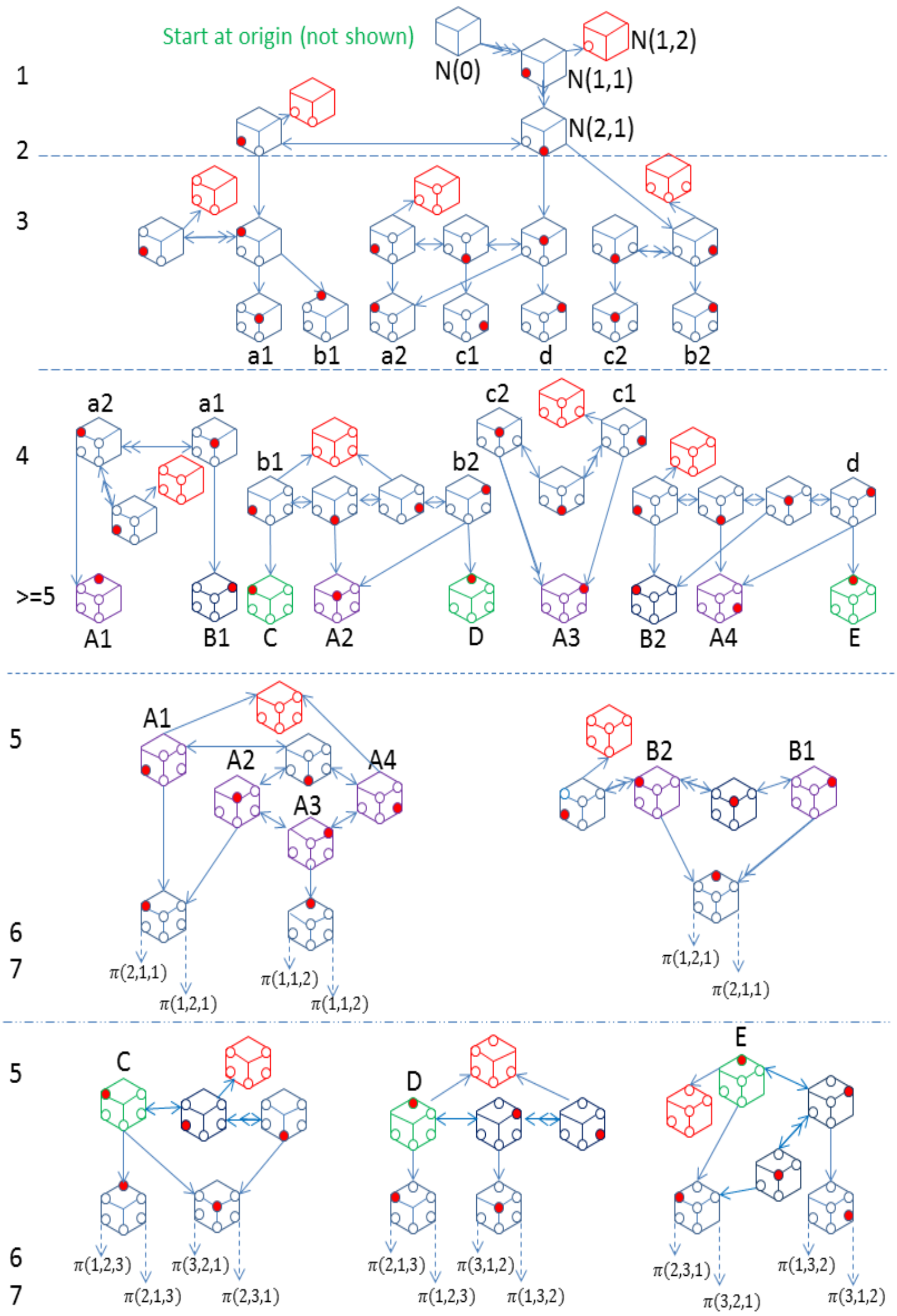


FIGURE 3. One-step transitions until return to origin, needed to study N_R. Conditional probability on each solid arrow is 1/3. The dotted short and long arrows go respectively to the origin and the only vertex not yet visited, with conditional probability written on each arrowhead.

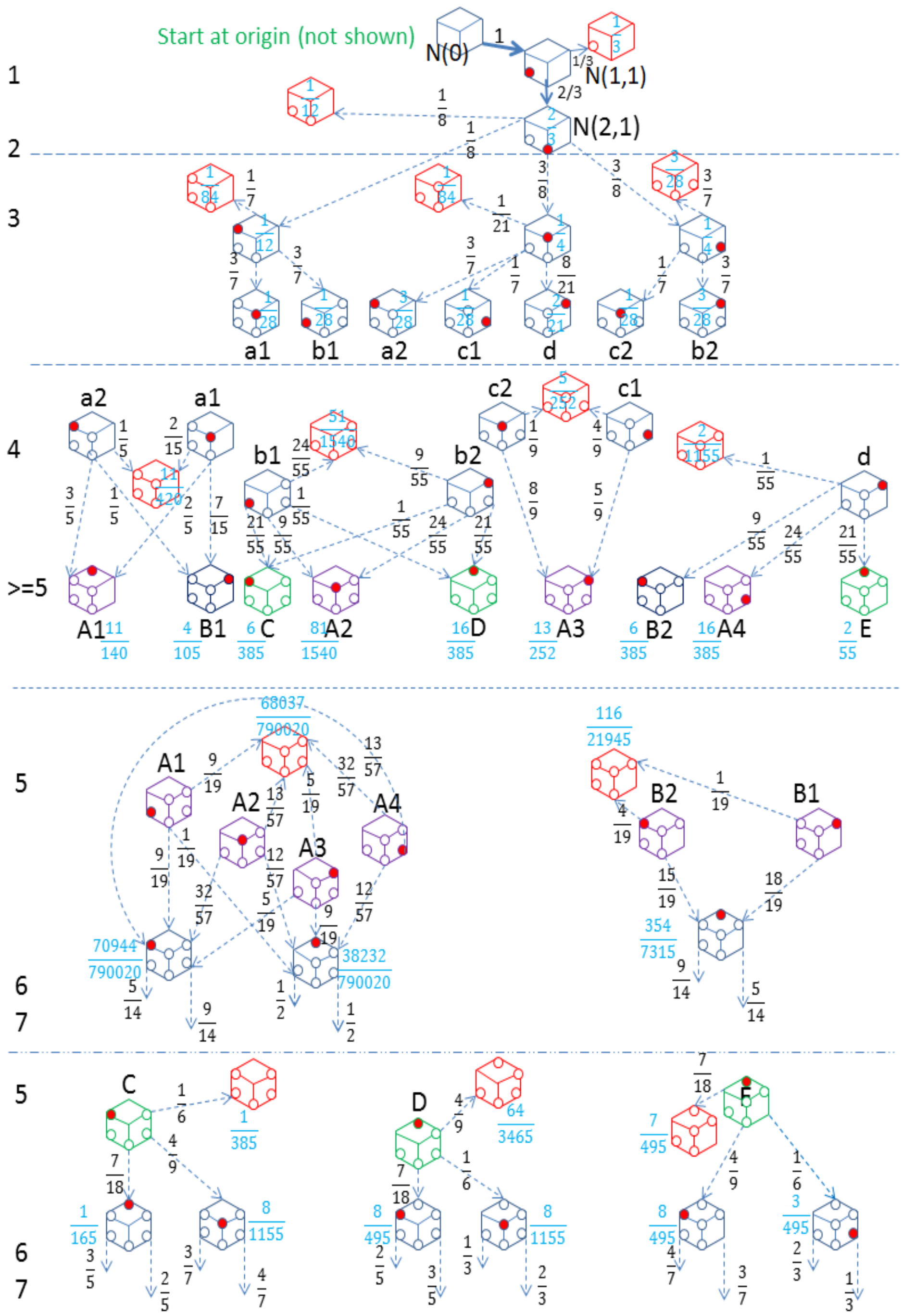


FIGURE 4. Eventual transitions until return to origin, needed to study N_R. Conditional probabilities are shown on the arrows, and arrival probabilities are shown on/near the nodes. The dotted short and long arrows respectively go to the origin and the only vertex not yet visited.

Let $\pi(d_1, d_2, d_3)$ *denote the probability that starting from Vertex* V_0*, the RW visits Vertex* V_1 *before it visits Vertex* V_2*. Then* $\pi(d_2, d_1, d_3) = 1 - \pi(d_1, d_2, d_3)$*, which implies that* $\pi(1, 1, 2) = 1/2 = \pi(2, 2, 2)$*. Also, then*

$$\pi(1,3,2) = 2/3;\ \pi(1,2,3) = 3/5;\ \pi(1,2,1) = 9/14;\ \pi(2,3,1) = 4/7 \quad (2.7)$$

Proof. By conditioning on the first move, we can see that $\pi(1, 3, 2) = (1/3) + (2/3)\pi(2, 2, 2) = 2/3$; and $\pi(1, 2, 3) = (1/3) + (2/3)[1 - \pi(1, 2, 3)]$, which when solved yields $\pi(1, 2, 3) = 3/5$. The remaining two quantities are interrelated as

$$\begin{aligned}\pi(1,2,1) &= (1/3) + (1/3)\cdot[1-\pi(1,2,1)] + (1/3)\cdot\pi(2,3,1)\\ \pi(2,3,1) &= (2/3)\cdot\pi(1,2,1) + (1/3)\cdot[1-\pi(2,3,1)]\end{aligned}$$

which we can solve simultaneously to get their values. □

Having calculated all conditional probabilities shown on the arrows in Fig. 4, we next obtain the probability of ever reaching an absorbing node (in which the RW has returned to the origin) simply by multiplying the conditional probabilities along the path from Node N(0) to that absorbing node. These probabilities are shown on/near the nodes (in blue font). By adding up the probabilities of the absorbing nodes corresponding to the same value of N_R, we obtain the probability distribution of N_R, summarized in Table 1 below.

TABLE 1. The probability distribution $p(k) = P\{N_R = k\}$; and the frequencies in a simulation (based on 10^6 iterations)

k	1	2	3	4	5	6	7	total
$p(k)$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{11}{84}$	$\frac{1121}{13860}$	$\frac{33343}{263340}$	$\frac{7639}{65835}$	$\frac{8483}{65835}$	1
	.33333	.08333	.13095	.08088	.12662	.11603	.12885	1
Simulation frequencies	333246	83856	131070	80841	126794	115838	128355	10^6

We conducted a simulation using the programming environment R (based on 10^6 iterations) until the RW returned to the origin. The frequencies of N_R are shown in the bottom row of Table 1. A statistical test of goodness of fit yields chi-square=5.9230, which on 6 degrees of freedom has a P-value of .4319. The simulated mean, variance and standard deviation are respectively 3.4450, 4.903046, 2.2143. Thus, our theoretical results are supported by the simulation outcomes.

Proposition 2. *The number N_R of vertices visited before return to origin has a probability distribution given in Table 1. In particular, N_R has mean* 3.4476*, mean square* 16.7955*, variance* 4.9096 *and SD* 2.2158*.*

Remark 2. Without even studying the cover time $\bar{T}$, we can determine which event will happen first—returning to origin or visiting all vertices. Since $T_R \neq \bar{T}$, we have $P\{T_R > \bar{T}\} = P\{N_R = 7\} = .12885$, and $P\{T_R < \bar{T}\} = P\{N_R < 7\} = .87115$.

3. Visiting All Vertices

We now study the cover time $\bar{T}$ of a hexahedron and the last vertex $L = X(\bar{T})$. The reader may imagine that there is a cookie at each vertex; and a cookie monster takes a RW and eats the cookie when it visits a vertex for the first time. During follow-up visits to the same vertex the monster eats nothing. Then $\bar{T}$ counts how many steps are needed until all 8 cookies are eaten, and $L = X(\bar{T})$ represents which cookie is eaten the last. In this section, we find the distributions of L and $\bar{T}$.

3.1. **The Vertex Visited the Last.** By symmetry, L is equally likely to be any one of Vertices 1, 2, 3 (which are adjacent to the origin), also L is equally likely to be any one of Vertices 4, 5, 6 (which are at a distance 2 from the origin). Therefore, let us define $\alpha = P\{L = 1\} = P\{L = 2\} = P\{L = 3\}$ and $\beta = P\{L = 4\} = P\{L = 5\} = P\{L = 6\}$. Also, let us define $\gamma = P\{L = 7\}$, the probability that Vertex 7 (which is diametrically opposite the origin) is the last.

To evaluate α, β, γ, let us consider a RW that starts at Vertex 1, and continues until all vertices are visited. Let L_1 denote the last vertex visited and SL_1 denotes the second last vertex visited by such a RW. Define

$$\theta_k = P\{L_1 = 0, SL_1 = k\}, \text{ for } k = 4, 5, 6, 7 \tag{3.1}$$

Again, by symmetry, we have $\theta_4 = \theta_5$. The following lemma tells us how to calculate α, β, γ based on $\theta_5, \theta_6, \theta_7$.

Lemma 3. *One can calculate* α, β, γ *based on* $\theta_5, \theta_6, \theta_7$ *as follows:*

$$1 = 3\,\alpha + 3\,\beta + \gamma \tag{3.2}$$

$$\beta = (2/3)\,(\alpha + \theta_5) + (1/3)\,(\gamma + \theta_6) \tag{3.3}$$

$$\gamma = \beta + \theta_7 \tag{3.4}$$

More explicitly, $\beta = (2 + 6\theta_5 + 3\theta_6 + \theta_7)/14$.

Proof. Equation (3.2) follows simply from the definitions of α, β, γ.

To justify (3.4), we need to show that $\{L = 7\}$ is the disjoint union of $\{L_1 = 7\}$ and $\{L_1 = 0, SL_1 = 7\}$. Here is why. Without loss of generality, the RW leaves Vertex 0 to go to Vertex 1. Thereafter, let us consider the truncated RW starting from Vertex 1 until all vertices are visited. If the last vertex is 7, (which happens with probability $P\{L_1 = 7\} = \beta$ since Vertex 7 is at a distance 2 from Vertex 1), then by augmenting Vertex 0 at the very beginning of this truncated RW, we note that the last vertex for the original RW starting at Vertex 0 is also 7. That is, $\{L_1 = 7\}$ implies $\{L = 7\}$. Alternatively, for the truncated RW starting from Vertex 1, if the second last vertex is 7 and the last vertex is 0, then also by augmenting Vertex 0 at the very beginning of this truncated RW, we note that the last vertex for the original RW starting at Vertex 0 is 7. That is, $\{L_1 = 0, SL_1 = 7\}$ also implies $\{L = 7\}$. Of course, these two possibilities are mutually exclusive and exhaustive. Hence, $P\{L = 7\} = P\{L_1 = 7\} + P\{L_1 = 0, SL_1 = 7\}$, implying that (3.4) holds.

To justify (3.3), we choose and fix any vertex at a distance 2 from the origin. Call it Vertex B. Then we rotate the cube about the axis joining Vertices 0 and 7 so that the first vertex visited is relabeled as Vertex 1. The other two vertices adjacent to Vertex 0 are arbitrarily relabeled as 2 and 3. The vertices opposite 1, 2, 3 are relabeled as 6, 5, 4 respectively. We note the new label k that Vertex B receives after relabeling: It is 4, 5 or 6 with probability 1/3 each. We consider the rest of the RW starting from Vertex 1. As in the previous paragraph, we note that $\{L = k\}$ is the disjoint union of $\{L_1 = k\}$ and $\{L_1 = 0, SL_1 = k\}$, for each equally likely case $k = 4, 5, 6$. Moreover, $\theta_4 = \theta_5$ by symmetry. Hence, (3.3) holds.

Substituting (3.2) and (3.4) in (3.3), we get the expression of β. □

It remains to evaluate θ_k for $k = 5, 6, 7$. For each k separately, we construct a transition diagram analogous to Figure 3. In each new figure, we must also declare as absorbing any node in which the RW reaches Vertex k, in addition to those nodes in which the RW reaches the origin. Consequently, all nodes previously accessible from the newly declared absorbing nodes are completely eliminated. We must also pay special attention to the cases where Figure 3 permitted a double arrow. This may or may not be permitted in the new figure. When not permitted, we must draw new nodes that are formed instead. Next, we calculate the conditional eventual transition probabilities in each absorbing chain in Figures 5–7. The details, being similar to those in Figure 3, are not shown. Finally, adding up the products of these conditional probabilities along all possible paths, we obtain $P\{L_1 = 0, SL_1 = k\}$.

In Figures 5–7, we write the probability of reaching a node on/near the node (in light blue font). This we do only for nodes that eventually lead to the terminal nodes for which $\{L_1 = 0, SL_1 = k\}$. From Figure 5, we have

$$\theta_5 = \frac{17}{1045} \cdot \frac{1}{2} + \frac{1}{330} \cdot \frac{2}{3} + \frac{1}{330} \cdot \frac{1}{3} = \frac{7}{627}$$

Similarly, from Figure 6, we have

$$\theta_6 = \frac{421}{43890} \cdot \frac{1}{2} + \frac{1}{1155} \cdot \frac{2}{3} = \frac{16}{1881}$$

and from Figure 7, we have

$$\theta_7 = \frac{1}{165} \cdot \frac{2}{5} + \frac{8}{495} \cdot \frac{3}{5} = \frac{2}{165}$$

Substituting the values of $\theta_5, \theta_6, \theta_7$ in (3.2)–(3.4), we have the following result.

Proposition 3. *Starting from Vertex 0 (origin), the Vertex L visited the last has the following distribution:*

$$P\{L = 1\} = P\{L = 2\} = P\{L = 3\} = \alpha = 8483/65835 = 0.1288524341156$$

$$P\{L = 4\} = P\{L = 5\} = P\{L = 6\} = \beta = 3299/21945 = 0.1503303713830$$

$$P\{L = 7\} = \gamma = 713/4389 = 0.1624515835042$$

We carried out a second simulation using R (based on 10^6 iterations) until the RW on a hexahedron visited all vertices. The frequencies of L corresponding to values $1, 2, 3$ were respectively $(128827, 129138, 128877)$. A statistical test of goodness of fit of equality of these three probabilities shows: chi-square=0.4326, P-value=.8055, on two degrees of freedom. Similarly, the frequencies of L corresponding to values $4, 5, 6$ were respectively

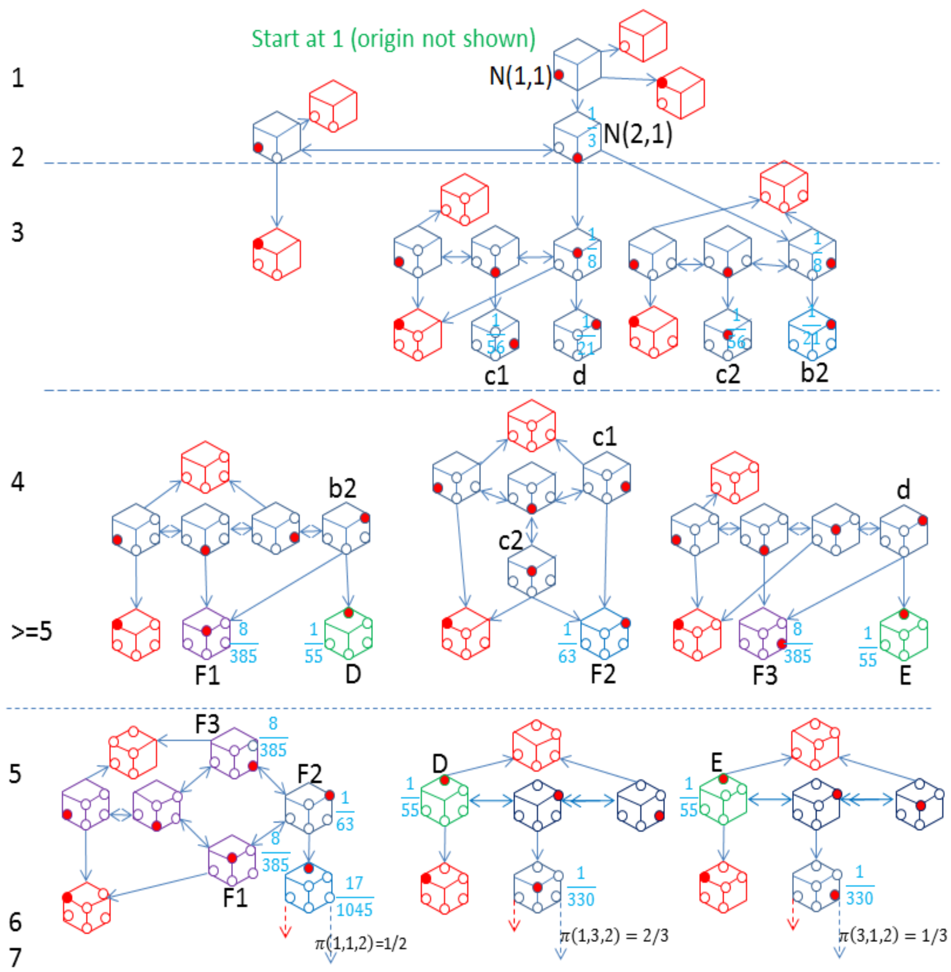


FIGURE 5. One-step transitions from Vertex 1 until the RW reaches Vertex 0 (origin) or Vertex 5, needed to compute θ_5. Probability on each arrow is 1/3, or as shown.

(150246, 150351, 150388). A statistical test of goodness of fit of equality of these three probabilities shows: chi-square=.0722, P-value=.9645, on two degrees of freedom. Lastly, and most importantly, the frequencies of L over sets $\{1,2,3\}, \{4,5,6\}, \{7\}$ were respectively (386842, 450985, 162173). A statistical test of goodness of fit of the probabilities $(3\alpha, 3\beta, \gamma)$, where α, β, γ are given in Proposition 3.1, shows: chi-square=0.6875, P-value=.7091, on two degrees of freedom. Thus, our theoretical results are supported by the simulation outcomes.

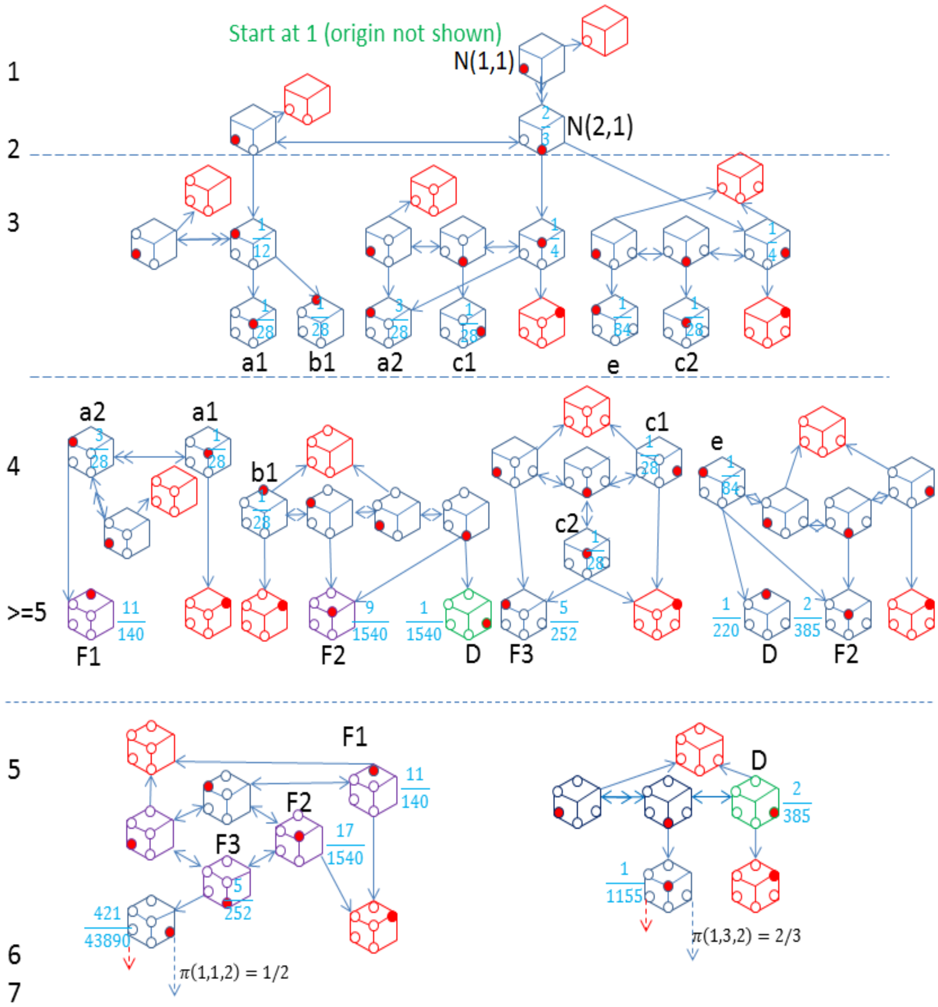


FIGURE 6. One-step transitions from Vertex 1 until the RW reaches Vertex 0 (origin) or Vertex 6, needed to compute θ_6. Probability on each arrow is 1/3, or as shown.

Remark 3. One could evaluate α from the results of Subsection 2.2 as

$$\alpha = P\{L = 1\} = P\{L_1 = 0\} = P\{T_R > \bar{T}\} = P\{N_R = 7\} = .128852$$

Our method of this subsection gives an alternative verification of the result.

3.2. **Cover Time $\bar{T}$ to Visit All Vertices.** To study $\bar{T}$ we only keep track of the pattern of visited vertices and the current vertex, and safely forget the original labeling of the vertices, including the origin. In Fig. 8,

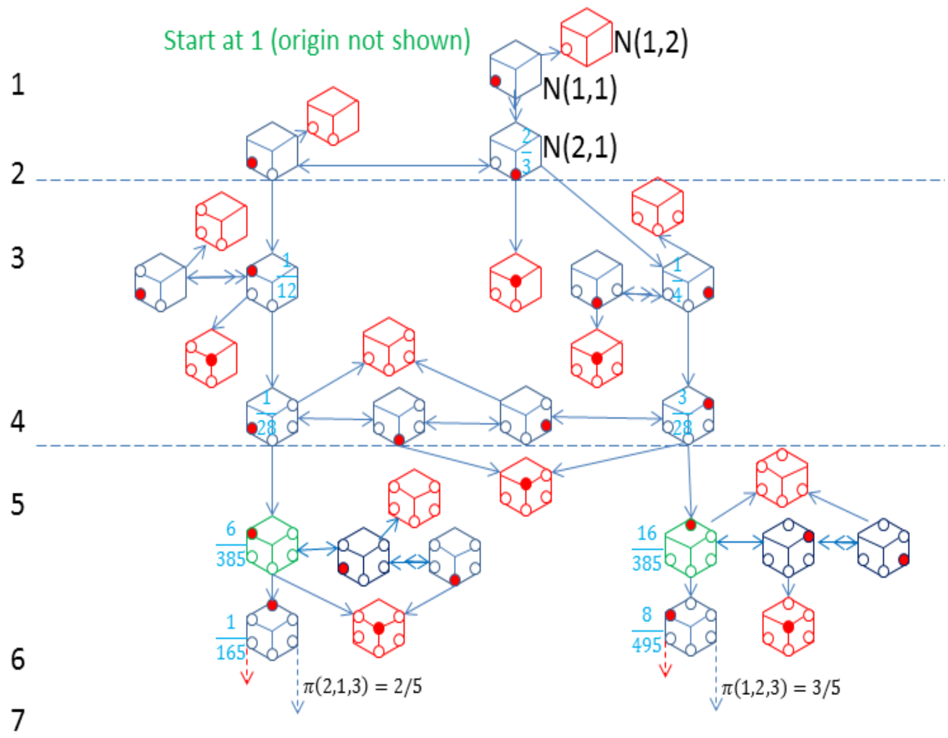


FIGURE 7. One-step transitions from Vertex 1 until the RW reaches Vertex 0 (origin) or Vertex 7, needed to compute θ_7. Probability on each arrow is 1/3, or as shown.

we draw the one-step transitions until all vertices are visited at least once, where each arrow represents a conditional probability of 1/3. We label the nodes using a distinct letter for the number of visited vertices (shown in the left margin); a first subscript to denote a subgroup whose members communicate, and a second subscript to distinguish the members within the subgroup.

The description of Fig. 8 is as follows: The RW surely leaves the origin to go to a non-origin vertex (call that Vertex 1, without loss of generality). Therefore, in one step Node A leads to Node B with probability one. Then the RW goes back to the origin in one step without visiting any other vertex with probability 1/3, in which case we can interchange the roles of the origin and the visited vertex, and imagine that the RW is still in Node B. Alternatively, the RW visits a new vertex at a distance 2 from the origin

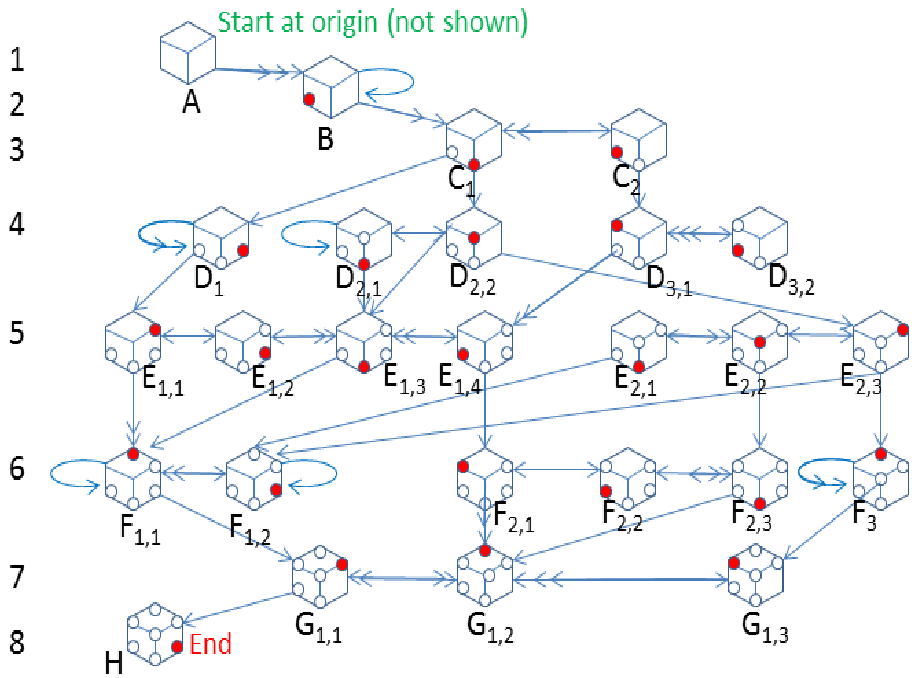


FIGURE 8. one-step transitions until all vertices are visited

with probability 2/3, leading to Node C_1. From Node C_1, with probability 1/3, the RW returns to Vertex 1 giving rise to Node C_2; with probability 1/3, the RW moves on to a vertex on the same face as the three previously visited vertices giving rise to Node D_1; and with probability 1/3, the RW moves to a vertex diametrically opposite one of the visited vertices giving rise to Node $D_{2,2}$. From node C_2, with probability 2/3 the RW moves to one of the vertices already visited; and with probability 1/3, the RW moves to a fourth vertex giving rise to Node $D_{3,1}$. In this manner, the process continues until it ends in Node H where all vertices have been visited at least once.

For any Node Z (ranging over Nodes A through H, with appropriate subscript(s)), let ${}_ZT_H$ denote the time taken by the RW to go from Node Z to Node H. In this notation, $\bar{T} = {}_AT_H$. Here we take liberty to re-use a similar notation already used earlier, but we are now referring to transition from one node to another (not from one vertex to another). Also, let us denote the mean, the mean square and the variance of ${}_ZT_H$

by ${}_Ze_H = E[{}_ZT_H]$, ${}_Zs_H = E[{}_ZT_H{}^2]$, and ${}_Zv_H = V({}_ZT_H) = {}_Zs_H - {}_Ze_H{}^2$ respectively. To avoid writing a subscript of a subscript, we may sometimes write the subscript(s) as arguments within parentheses.

Fig. 8 is the basis for constructing interrelations among the time variables ${}_ZT_H$'s, which can be solved backwards (proceeding from Node H), since, ${}_HT_H \equiv 0$. The probability distributions of ${}_{G(1,1)}T_H$, ${}_{G(1,2)}T_H$, ${}_{G(1,3)}T_H$ are interrelated as follows:

$$\begin{aligned}
{}_{G(1,1)}T_H &= \begin{cases} 1 & \text{with probability } 1/3 \\ 1 + {}_{G(1,2)}T_H & \text{with probability } 2/3 \end{cases} \\
{}_{G(1,2)}T_H &= \begin{cases} 1 + {}_{G(1,1)}T_H & \text{with probability } 2/3 \\ 1 + {}_{G(1,3)}T_H & \text{with probability } 1/3 \end{cases} \qquad (3.5) \\
{}_{G(1,3)}T_H &= 1 + {}_{G(1,2)}T_H
\end{aligned}$$

The exact distributions of ${}_{G(1,1)}T_H$, ${}_{G(1,2)}T_H$, ${}_{G(1,3)}T_H$ are difficult to derive. Moreover, the distributions of ${}_ZT_H$'s become more and more complicated as we move backwards along the nodes in Fig 8. We will not attempt to get explicit expressions for the probability distributions of any such time variables ${}_ZT_H$'s. However, Fig. 8 can also be used to construct interrelations among the means and the mean squares (or any other raw moment) of ${}_ZT_H$'s. These relations can be solved (proceeding backwards from Node H), until we get the mean and the mean square of $\bar{T} = {}_AT_H$. We will illustrate a couple of these computations and let the reader fill-in the other computations to complete the results summarized below.

(1) Taking expectations in (3.5), we get

$$\begin{aligned}
{}_{G(1,1)}e_H &= 1 + (2/3)_{G(1,2)}e_H \\
{}_{G(1,2)}e_H &= 1 + (2/3)_{G(1,1)}e_H + (1/3)_{G(1,3)}e_H \\
{}_{G(1,3)}e_H &= 1 + {}_{G(1,2)}e_H
\end{aligned}$$

which when solved yield ${}_{G(1,1)}e_H = 7$, ${}_{G(1,2)}e_H = 9$, ${}_{G(1,3)}e_H = 10$. Similarly, first squaring and then taking expectations in (3.5), we get

$$
\begin{aligned}
{}_{G(1,1)}s_H &= 1 + (2/3)_{G(1,2)}s_H + (2/3)2_{G(1,2)}e_H \\
&= 13 + (2/3)_{G(1,2)}s_H \\
{}_{G(1,2)}e_H &= 1 + (2/3)_{G(1,1)}s_H + (1/3)_{G(1,3)}s_H \\
&\quad +(2/3)2\,_{G(1,1)}e_H + (1/3)2\,_{G(1,3)}e_H \\
&= 17 + (2/3)_{G(1,1)}s_H + (1/3)_{G(1,3)}s_H \\
{}_{G(1,3)}e_H &= 1 + {}_{G(1,2)}s_H + 2_{G(1,2)}e_H = 19 + {}_{G(1,2)}s_H
\end{aligned}
$$

which when solved, yield ${}_{G(1,1)}s_H = 109$, ${}_{G(1,2)}s_H = 144$, ${}_{G(1,3)}s_H = 163$.

(2) The random time ${}_{F(3)}T_H$ equals $1 +_{G(1,3)} T_H$ with probability $1/3$ and $1 + {}_{F(3)}T'_H$ with probability $2/3$, where ${}_{F(3)}T'_H$ is IID as ${}_{F(3)}T_H$. Hence, by taking expectation, we have

$$
{}_{F(3)}e_H = 1 + (1/3)10 + (2/3)_{F(3)}T_H = 13
$$

and by first squaring and then taking expectation, we have

$$
\begin{aligned}
{}_{F(3)}s_H &= 1 + (1/3)\,_{G(1,3)}s_H + (1/3)2_{F(3)}e_H + \\
&\quad (2/3)\,_{F(3)}s_H + (2/3)\,2_{F(3)}e_H = 238
\end{aligned}
$$

In this manner, when all calculations are done, we get the results on $\bar{T} = {}_A T_H = 1 + {}_B T_H$ as summarized below.

Proposition 4. *Starting from Vertex 0 (origin), the cover time $\bar{T}$ to visit all vertices has mean* 1996/95 = 21.01*, mean square* 11468329/21660 = 529.4704*, variance* 88.02819 *and SD* 9.382334*.*

4. Return to Origin after Visiting All Vertices

Recall from Section 3.1 that the last vertex visited is adjacent to the origin with probability $3\,\alpha$, at a distance 2 from the origin with probability $3\,\beta$, and at a distance 3 from the origin with probability γ. Consequently, the time ${}_L T_R$ to return to origin has the same probability distribution as ${}_1T_0$ with probability $3\,\alpha$; as ${}_2T_0$ with probability $3\,\beta$; and as $1 + {}_2T_0$ with probability γ, where the distribution of ${}_2T_0$ is given in (2.4) and the distribution of ${}_1T_0$ is given in (2.2). Combining these results, we have the following proposition.

Proposition 5. *The random time* ${}_LT_R$ *to return to origin after visiting all vertices has a probability mass function given by*

$$\begin{aligned} P\{{}_LT_R = 1\} &= \alpha \\ P\{{}_LT_R = 2k\} &= (3\beta)\ \cdot (2/9)\cdot (7/9)^{k-1}\,,\ \text{for } k = 1,2,3,\ldots \\ P\{{}_LT_R = 2k+1\} &= (2\alpha+\gamma)\ \cdot (2/9)\cdot (7/9)^{k-1}\,,\ \text{for } k = 1,2,3,\ldots \end{aligned}$$

In particular, ${}_LT_R$ *has mean* $7+6\beta+3\gamma = 8.38930$*, mean square* $109 + 105\beta + 54\gamma = 133.5571$*, variance* 63.1761 *and SD* 7.9483*.*

The distribution of ${}_LN_R$, the number of vertices visited during the return to origin after visiting all vertices, depends on L. Given $L \in \{1,2,3\}$, the conditional distribution of ${}_LN_R$ is the same as the (unconditional) distribution of N_R given in Table 1 of Section 2. Next, the conditional distribution of ${}_LN_R$, given $L \in \{4,5,6\}$, is obtained simply by truncating the distribution of N_R to $\{k > 1\}$. Finally, to obtain the conditional distribution of ${}_LN_R$, given $L = 7$, one must carefully sort the terminal nodes in Fig. 5 according as Vertex 7 is visited or not; and only in the later case, one must increase the value of N_R by one to correctly evaluate ${}_LN_R$. Combining these conditional distributions with weights $3\alpha, 3\beta, \gamma$ respectively, one can get the (unconditional) distribution of ${}_LN_R$ as described in the next result.

Proposition 6. *Starting from the last visited Vertex* L*, the number* ${}_LN_R$ *of vertices visited while returning to origin has the probability distribution given in Table 2. In particular,* ${}_LN_R$ *has mean* 4.263337*, mean square* 22.10621*, variance* 3.930176 *and SD* 1.982467*.*

5. Conclusion

We studied RWs on the vertices of a tetrahedron and an octahedron in [6], and on a hexahedron in this paper. Recall that a dual of a planar graph is another graph where each face of the old graph becomes a vertex of the new graph (and vice versa) with the understanding that two vertices of the new graph are adjacent if and only if the corresponding faces of the old graph share a common boundary. A tetrahedron is its own dual. An octahedron and a hexahedron are duals of each other. Thus, we have also studied RWs on the faces of a tetrahedron, octahedron and hexahedron.

There are exactly five Platonic solids (named after the Greek philosopher Plato who lived about 428–347 BCE), which are three dimensional

TABLE 2. The conditional (given L) and the unconditional probability distributions of ${}_LN_R$, writing $q(k) = P\{{}_LN_R = k\}$, etc.

k	1	2	3	4	5	6	7	sum
$q(k\|L=1)$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{11}{84}$	$\frac{1121}{13860}$	$\frac{33343}{263340}$	$\frac{7639}{65835}$	$\frac{8483}{65835}$	1
$q(k\|L=4)$	0	$\frac{1}{8}$	$\frac{11}{56}$	$\frac{1121}{9240}$	$\frac{33343}{175560}$	$\frac{7639}{43890}$	$\frac{8483}{43890}$	1
$q(k\|L=7)$	0	0	$\frac{1}{7}$	$\frac{289}{1155}$	$\frac{3043}{14630}$	$\frac{2787}{14630}$	$\frac{4574}{21945}$	1
$q(k)$	$\frac{8483}{65835}$	$\frac{24499}{276553}$	$\frac{2547}{15682}$	$\frac{1887}{14902}$	$\frac{195103}{1158654}$	$\frac{1403}{9093}$	$\frac{1564}{9155}$	1
	.12885	.08859	.16242	.12663	.16839	.15429	.17084	1

convex spaces bounded by congruent regular polygonal faces with the same number of faces meeting at each vertex. See Fig. 9. In near future, we intend to study a RW on the vertices/faces of an icosahedron and a dodecahedron, which are duals of each other. Techniques developed here can be beneficial in studying RWs on other distance-regular graphs.

Tetrahedron, Hexahedron, Octahedron, Dodecahedron, Icosahedron

FIGURE 9. The five Platonic solids

We invite interested readers to study asymmetric RWs on the vertices of a Platonic solid. One form of asymmetry is to consider the edge along which the RW has reached the current vertex as a special edge labeled 0 and the other edges incident at the current vertex are labeled 1, 2, . . . , K going clockwise when looked at from outside the Platonic solid. One can then assume that starting from the origin the RW is equally likely to go to any adjacent vertex; but thereafter at each successive step, the next vertex is chosen according to a pre-specified one-step transition probability vector $\mathbf{p} = (p_0, p_1, p_2, \ldots, p_K)$, where p_h is the probability of traveling along edge h. Other forms of asymmetric RW are also possible.

Acknowledgment. The author thanks Professor Wenbo Li for suggesting the problem, and the seminar participants at his home institution for some discussions.

References

1. Göbel, F. and Jagers, A.A. (1974), Random walks on graphs, *Stochastic Processes and their Applications*, **102**, 311-336. MR0397887
2. Letac, G. and Takács, L. (1980), Random walks on a dodecahedron, *J. Appl. Prob.*, **17**, 373-384. MR0568948
3. Maiti, S.I. and Sarkar, J. (2019), Symmetric walks on paths and cycles, *Mathematics Magazine*, **92(4)**, 252-268.
4. Ross, S. (1996), *Stochastic Processes*, (2 ed.), New York: Wiley.
5. Sarkar, J. (2006), Random walk on a polygon, In: Recent Developments in Nonparametric Inference and Probability, J Sun, A DasGupta, V Melfi, C Page, Eds., *IMS Lecture Notes–Monograph Series*, **50**, 31-43, Beachwood, OH: Inst. Math. Statist. MR 2409062.
6. Sarkar, J. and Maiti, S.I. (2017), Symmetric random walks on regular tetrahedra, octahedra and hexahedra. In: Proceedings of the 9th International Triennial Calcutta Symposium on Probability and Statistics, *Calcutta Statistical Association Bulletin* 69(1) (2017), 110-128. `http://journals.sagepub.com/doi/abs/10.1177/0008068317695974`.
7. van Slijpe, A.R.D. (1984), Random walks on regular polyhedra and other distance-regular graphs, *Statistica Neerlandica*, **38**, 273-292. MR0773043
8. van Slijpe, A.R.D. (1985), Random walks on the triangular prism and other vertex-transitive graphs, *J. Comput. and Applied Math.*, **15**, 383-394. MR0851897
9. Wildberger, N.J. (2014), Hypergroups associated to random walks on the Platonic solids, `web.maths.unsw.edu.au/~norman/papers/RandomWalksPlatonicSolids.pdf`

Professor Jyotirmoy Sarkar
Department of Mathematical Sciences
Indiana University-Purdue University Indianapolis
Indianapolis, IN 46202, USA
Email: jsarkar@iupui.edu

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 87–90

GLAISHER'S PARTITION PROBLEM

ARITRO PATHAK

(Received : 26 - 06 - 2019 ; Revised : 03 - 03 - 2020)

Abstract. For natural numbers n and d, set $p(n)$ to be the number of integer partitions of n. Set $a(n)$ to be the number of integer partitions of n, all of whose parts fail to be divisible by d. Set $b(n)$ to be the number of integer partitions of n, where no part is repeated d or more times. Glaishers theorem gives a bijective proof of the fact that $a(n) = b(n)$. We present a new family of bijections to show that $p(n)-a(n) = p(n)-b(n)$, which is equivalent to Glaisher's theorem, and is a generalization of Glaisher's original argument. This provides a rich family of elementary bijections between these two sets of partitions.

1. Introduction

Glaisher's theorem states that the number of partitions of an integer N into parts not divisible by d is equal to the number of partitions where no part is repeated d or more times.

This theorem can also be proven in an elementary way using generating functions. The original combinatorial proof is well known, as can be found in [1]. It also appears as Exercise 3.2.3 in Igor Pak's survey on Partition Bijections [2]. We present an outline of Glaisher's original proof in our first proof below. In the second proof, we give a new family of bijections generalizing Glaisher's argument.

Theorem. Given integers $n \geq d \geq 1$, the number of integer partitions of n all of whose parts fail to be divisible by d, is equal to the number of integer partitions of n where no part is repeated d or more times.

Proof 1. For a partition where each part appears less than d times, split up the parts divisible by d^t into d^t parts (where $t \geq 0$ is the highest power of d that divides that part). This gives a partition with no part divisible by d.

2010 Mathematics Subject Classification: 05A17
Key words and phrases: Glaisher's Bijection, Integer Partitions

For the inverse, for any partition where no part is divisible by d, write the number of times $s^{(j)}$ appears in the base d representation: $\sum_i a_i^{(j)} d^i$. Then consider the partition where $d^i s^{(j)}$ each appear $a_i^{(j)} < d$ number of times. It can be checked that this gives a bijection. $\square$

By considering the complement of the two sets under consideration, we construct the new family of bijections in the next proof. This obviously implies Glaisher's theorem.

Proof 2. We show that the number of elements in the set A_N that are partitions of N into parts where at least one part is divisible by d, is equal to the number of elements in the set B_N that are partitions where at least one part is repeated d or more times.

Consider a partition P_A in A_N. For each $x \in \mathbb{Z}_+$ that is not divisible by d, construct a square matrix $M_{ij}^{(x)}$, $(i, j \in \mathbb{Z}_+)$, where the square at the intersection of the row i and column j contains the coefficient of d^i in the base d expansion of the integer that equals the number of times that $x \cdot d^j$ appears in P_A.

By the definition of A_N, for this partition, there exists at least one $x_{P_A} \in \mathbb{Z}_+$ not divisible by d, so that the matrix $M^{(x_{P_A})}$ has a non zero element in the columns indexed by $j \geq 2$ (i.e it's not just the first column that is non empty).

Similarly, for any given partition P_B in B_N, we can construct for each $x \in \mathbb{Z}_+$ not divisible by d, a square matrix $\tilde{M}_{ij}^{(x)}$ as above indexed by $\mathbb{Z}_+ \times \mathbb{Z}_+$ so that there exists some x_{P_B} not divisible by d, whose matrix contains a non zero element in the rows corresponding to $i \geq 2$ (i.e it's not just the first row that has non zero elements).

Now we construct the family of bijections between A_N and B_N.

For the partition P_A in A_N, looking at the matrix corresponding to x_{P_A}, we can look at the "southwest-northeast diagonals" indexed by $k \in \mathbb{Z}_+$: $D_{x_{P_A}}^{(k)} = \{(i,j) \in M_{ij}^{(x_{P_A})} | i + j = k, i, j \in \mathbb{Z}_+\}$.

In any diagonal indexed by some integer k , we can take a permutation with the restriction that the $(1, k-1)$ square is taken in the interior of this diagonal, which is the set $\text{int}(D_{x_{P_A}}^{(k)}) = \{(i,j) | (i,j) \in D_{x_{P_A}}^{(k)}, i \neq k-1,\ j \neq k-1\}$, and the $(k-1, 1)$ square is taken to the $(1, k-1)$ square.

Since we are permuting within each fixed diagonal, the contribution to the sum of N is invariant by this bijection.

It is clear that this bijectively gives us an element $\tilde{P}_A \in B_N$, and on applying the corresponding inverse permutation to all the diagonals corresponding to the matrix in $\tilde{P}_A \in B_N$, we recover $P_A \in A_N$.

Note that it is important in general, that in the previous permutation in going from $P_A \in A_N$ to $\tilde{P}_A \in B_N$, we don't take squares from $\mathrm{int}(D^{(k)}_{x_{P_A}})$ to go the upper row consisting of the squares $\{(1,k), k \geq 2\}$ by the permutation. This is to negate the cases where the only non zero elements in $M^{(x_{P_A})}$ are precisely those in the interior of the matrix, $\cup_{k=2}^{\infty}\mathrm{int}(D^{(k)}_{x_{P_A}})$, that get permuted to go into the upper row, and hence we get a matrix where all the rows corresponding to $i \geq 2$ are empty, so we do not get an element of B_N. Sending the $(k-1,1)$ square in $M_{ij}^{x_{P_A}}$ to go to $(1,k-1)$ in $\tilde{M}_{ij}^{x_{\tilde{P}_A}}$ ensures we have a well defined bijection from A_N to B_N.

This establishes the bijection between partitions in A_N and partitions in B_N. $\square$

Remark. The simplest possible permutation of any diagonal D_k is to simply interchange $(1,k-1)$ and $(k-1,1)$, while keeping all the elements of $\mathrm{int}(D^{(k)}_{x_{P_A}})$ fixed or choosing a random permutation restricted to $\mathrm{int}(D^{(k)}_{x_{P_A}})$. However, the proof above gives a bigger class of permutations.

In Glaisher's original argument for the corresponding complement sets, in one partition set only the first row of our matrix has non zero elements, while in the other partition set, only the first column has non zero elements. Glaisher's bijection involved simply swapping the $(1,j)$ th square with the $(j,1)$ square. So the simplest permutations of the previous paragraph, where the ends of the diagonal and the interior of the diagonal are permuted disjointedly, are 'Glaisher-like', while our permutations where one end of the diagonal is taken inside the interior is a variant of Glaisher's original argument.

Acknowledgement. The author is thankful to Konstantin Matveev for pointing out this problem in a Combinatorics class at Brandeis University. The author is also thankful to Dmitry Kleinbock and Ira Gessel for their suggestions upon looking at the manuscript.

References

[1] D. H. Lehmer, *Two nonexistence theorems on partitions.* Bull. Amer. Math. Soc. 52 (6)(1946), 538 -544.

[2] Igor Pak, *Partition Bijections, a survey*, The Ramanujan Journal 12 (1) (2006), 5-75.

Aritro Pathak
Department of Mathematics
Brandeis University, Waltham, MA, 02453, USA
E-mail: ap323@bradeis.edu

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 91–101

YAMABE SOLITONS ON GENERALIZED SASAKIAN-SPACE-FORMS

PRADIP MAJHI AND DEBABRATA KAR
(Received : 10 - 07 - 2019 ; Revised : 08 - 04 - 2020)

Abstract. In this paper, we study Yamabe solitons on generalized Sasakian-space-forms. We prove that the scalar curvature of such manifolds is constant as well as harmonic and the flow vector field is Killing. Moreover, we prove that either $\pounds_V \phi$ is orthogonal to the structure vector field ξ or V is an infinitesimal automorphism of the contact metric structure of (M, ϕ, ξ, η, g). Finally, we give a valuable remark.

1. Introduction

In 1988, the notion of Yamabe flow was introduced by R. S. Hamilton [13]. On a smooth Riemannian manifold (M^n, g_0), the Yamabe flow can be defined as the evolution of the metric g_0 in time t to $g = g(t)$ through the following equation:

$$\frac{\partial}{\partial t} g_t = -rg, g(0) = g_0 \tag{1.1}$$

is known as the Yamabe flow.

It is well known that a Riemannian metric g of an n-dimensional complete Riemannian manifold (M^n, g) is said to be a Yamabe soliton [9] if it satisfies

$$\pounds_V g = (\lambda - r)g, \tag{1.2}$$

for a constant $\lambda \in \mathbb{R}$ and a smooth vector field V on M^n, where r is the scalar curvature of g and $\pounds$ denotes the Lie-derivative operator. A Yamabe soliton is said to be shrinking, steady or expanding according to $\lambda > 0$, $\lambda = 0$ or $\lambda < 0$ respectively and λ is said to be the soliton constant.

2010 Mathematics Subject Classification: 53C15, 53C25
Key words and phrases: Yamabe solitons, generalized Sasakian-space-forms, scalar curvature, Killing vector field

A Yamabe soliton is a special soliton of the Yamabe flow that moves by one parameter family of diffeomorphisms ϕ_t generated by a fixed vector field V on M^n (for more details see [9], [25]).
The significance of Yamabe flow lies in the fact that it is a natural geometric deformation to metrics of constant scalar curvature. One note that Yamabe flow corresponds to the fast diffusion case of the porous medium equation (the plasma equation) in mathematical physics. Just as a Ricci soliton is a special solution of the Ricci flow, a Yamabe soliton is a special solution of the Yamabe flow that moves by one parameter family of diffeomorphisms ϕ_t generated by a fixed vector field V on M^n and homotheties, i.e. $g(.,t)) = \sigma(t)\phi_*(t)g_0$.

Given a Yamabe soliton, if $V = Df$ holds for a smooth function f on M^n, equation (1.1) becomes

$$Hessf = \frac{1}{2}(\lambda - r)g, \tag{1.3}$$

where $Hessf$ denotes the Hessian of f and D denotes the gradient operator of g on M^n. In this case f is called the potential function of the Yamabe soliton and g is said to be a gradient Yamabe soliton. A Yamabe soliton (respectively, gradient Yamabe soliton) is said to be trivial when V is Killing (respectively, f is constant).

Yamabe solitons on a three-dimensional Sasakian manifold were studied by R. Sharma [25]. Also Wang [26] studied Yamabe solitons on a three-dimensional Kenmotsu manifold. The purpose of the present paper is to study Yamabe solitons on generalized Sasakian-space-forms.

The study of curvature properties is one of the main problems in Differential geometry. As Chern said in [8] “a fundamental notion is curvature in its different forms”. Therefore, the determination of Riemann curvature tensor constitutes a very important topic.

In this sense, Blair et al. [1] introduced the notion of generalized Sasakian-space-forms.

A Sasakian manifold (M, ϕ, ξ, η, g) is said to be a Sasakian-space-form if all the ϕ-sectional curvatures $K(X \wedge \phi X)$ are equal to a constant c, where $K(X \wedge \phi X)$ denotes the sectional curvature of the section spanned by the unit vector field X, orthogonal to ξ, and ϕX. In such a case, the Riemann curvature tensor of M is given by

$$\begin{aligned} R(X,Y)Z \quad &= \frac{c+3}{4}\{g(Y,Z)X - g(X,Z)Y\} \\ &+\frac{c-1}{4}\{g(X,\phi Z)\phi Y - g(Y,\phi Z)\phi X + 2g(X,\phi Y)\phi Z\} \\ &+\frac{c-1}{4}\{\eta(X)\eta(Z)Y - \eta(Y)\eta(Z)X \\ &+g(X,Z)\eta(Y)\xi - g(Y,Z)\eta(X)\xi\}. \end{aligned} \tag{1.4}$$

These spaces can be modeled, depending on $c > -3, c = -3$ or $c < -3$.

As a natural generalization of these manifolds, Alegre, Blair and Carriazo introduced [1] the notion of generalized Sasakian-space-forms. They were defined as almost contact metric manifolds with Riemann curvature tensor satisfying an equation similar to (1.4), in which the constant quantities, $\frac{c+3}{4}$ and $\frac{c-1}{4}$ are replaced by differentiable functions, i.e., such that

$$\begin{aligned} R(X,Y)Z \quad &= f_1\{g(Y,Z)X - g(X,Z)Y\} \\ &+f_2\{g(X,\phi Z)\phi Y - g(Y,\phi Z)\phi X + 2g(X,\phi Y)\phi Z\} \\ &+f_3\{\eta(X)\eta(Z)Y - \eta(Y)\eta(Z)X \\ &+g(X,Z)\eta(Y)\xi - g(Y,Z)\eta(X)\xi\}. \end{aligned} \tag{1.5}$$

We will denote such a space by $M(f_1, f_2, f_3)$ and we write $R = f_1R_1 + f_2R_2 + f_3R_3$. Let us notice that, despites its name, a generalized Sasakian-space-form is not a Sasakian manifold in general; just an almost contact metric one. Let us assume that $f_1 \neq f_3$.

Sasakian-space-form, can be obtained as a particular case of generalized Sasakian-space-forms by taking $f_1 = \frac{c+3}{4}, f_2 = f_3 = \frac{c-1}{4}$. It is known that any three-dimensional (α, β)-trans Sasakian manifold with α, β depending on ξ is a generalized Sasakian-space-form [2]. However we can find generalized Sasakian-space-forms with non-constant functions and arbitrary dimensions. In [2], the authors cited several examples of generalized Sasakian-space-forms. In [18], Kim studied conformally flat generalized Sasakian-space-forms under the assumption that the characteristic vector field ξ is Killing and he classified locally symmetric generalized

Sasakian-space-forms. Also he proved some geometric properties of generalized Sasakian-space-forms which depend on the nature of the functions f_1, f_2 and f_3. Generalized Sasakian-space-forms have also been studied in ([2]-[7], [11], [15], [16], [17], [19]-[24], [27], [28]) and many others.

Theorem 1.1. *If (M, f_1, f_2, f_3) be a generalized Sasakian-space-form with Yamabe soliton metric, then its scalar curvature is constant and the flow vector V is Killing. Moreover, either $\pounds_V \phi$ is orthogonal to the structure field ξ or V is an infinitesimal automorphism of the contact metric structure of (M, ϕ, ξ, η, g).*

Corollary 1.1. *If (M, f_1, f_2, f_3) be a generalized Sasakian-space-form with Yamabe soliton metric, the scalar curvature is harmonic.*

According to Chow-Lu-Ni [9], Daskalopoulos-Sesum [10] and Hsu [14], the metric of any compact Yamabe soliton is a metric of constant scalar curvature. In [25] proves that the scalar curvature of a Yamabe soliton on a three-dimensional Sasakian manifold is a constant. In our paper, **Theorem 1.1** replaces the compactness with generalized Sasakian-space-form condition and induces more conclusion.
At this point of view, we are interested to point our present paper as follows:
After introduction, in section 2, we recall some well known basic formulae and properties of generalized Sasakian-space-forms. In section 3, after presenting some key we give the detailed proof of **Theorem 1.1**. In the last section, we give a valuable remark.

2. PRELIMINARIES

An odd-dimensional Riemannian manifold (M, g) is said to be an almost contact metric manifold if there exist on M a (1,1) tensor field ϕ, a vector field ξ (called the structure vector field) and a 1-form η such that ([5], [6])

$$\phi^2(X) = -X + \eta(X)\xi, \tag{2.1}$$

$$\eta(\xi) = 1, \quad g(X, \xi) = \eta(X), \tag{2.2}$$

$$g(\phi X, \phi Y) = g(X, Y) - \eta(X)\eta(Y), \tag{2.3}$$

for any vector fields X, Y on M.

In particular, in an almost contact metric manifold we also have

$$\phi\xi = 0, \quad \eta \circ \phi = 0 \tag{2.4}$$

and

$$(\nabla_X \eta)(Y) = g(\nabla_X \xi, Y). \tag{2.5}$$

Such a manifold is said to be a contact metric manifold if $d\eta = \Phi$, where

$$\Phi(X, Y) = g(X, \phi Y), \tag{2.6}$$

Φ is called the fundamental 2-form of M.

Again for a $(2n+1)$-dimensional generalized Sasakian-space-form, we have [1]

$$S(X, Y) = (2nf_1 + 3f_2 - f_3)g(X, Y) - (3f_2 + (2n-1)f_3)\eta(X)\eta(Y), \tag{2.7}$$

$$QX = (2nf_1 + 3f_2 - f_3)X - (3f_2 + (2n-1)f_3)\eta(X)\xi, \tag{2.8}$$

$$R(X, Y)\xi = (f_1 - f_3)[\eta(Y)X - \eta(X)Y], \tag{2.9}$$

$$R(\xi, X)Y = (f_1 - f_3)[g(X, Y)\xi - \eta(Y)X], \tag{2.10}$$

$$S(X, \xi) = 2n(f_1 - f_3)\eta(X), \tag{2.11}$$

$$S(\xi, \xi) = 2n(f_1 - f_3), \tag{2.12}$$

$$Q\xi = 2n(f_1 - f_3)\xi, \tag{2.13}$$

$$r = 2n(2n+1)f_1 + 6nf_2 - 4nf_3, \tag{2.14}$$

where R, S and r are the curvature tensor, Ricci tensor and scalar curvature of the space-form respectively.

Definition 2.1. *A smooth vector field V on an m-dimensional Riemannian manifold (M, g) is said to be conformal if*

$$\pounds_V g = 2\rho g, \tag{2.15}$$

for a smooth function ρ on M.

A conformal vector field V obeys the following [29]

$$(\pounds_V S)(X,Y) = -(m-2)g(\nabla_X D\rho, Y) + (\Delta\rho)g(X,Y) \tag{2.16}$$

and

$$\pounds_V r = -2\rho r + 2(m-1)\Delta\rho, \tag{2.17}$$

where D is the gradient operator and $\Delta = -\operatorname{div} D$ is the Laplacian operator of g.

Definition 2.2. [12] *An infinitesimal automorphism V is a smooth vector field such that Lie derivatives of all structure tensor along V vanishes.*
In an almost contact metric manifold (M, ϕ, ξ, η, g), the condition that a smooth vector field V is an infinitesimal automorphism is

$$\pounds_V g = \pounds_V \xi = \pounds_V \phi = \pounds_V \eta = 0. \tag{2.18}$$

3. **Proof of Theorem 1.1**

Before giving the proof of the Theorem 1.1, we state and prove the following lemmas:

Lemma 3.1. *On a generalized Sasakian-space-form M with Yamabe soliton metric, the following relations hold:*

(i) $(\pounds_V \eta)(\xi) = \frac{\lambda - r}{2}$,
(ii) $\eta(\pounds_V \xi) = \frac{r-\lambda}{2}$,
(iii) $\rho = \frac{\lambda - r}{2}$.

Proof. Since ξ is a unit vector field, we have $\eta(\xi) = 1$. Taking Lie-differentiation of this relation along V and using the first equation of (2.2) we obtain (ii).
Next taking Lie-differentiation of the second equation of (2.2) along V we get

$$\eta(\pounds_V \xi) = -(\pounds_V \eta)(\xi) = \frac{r-\lambda}{2}$$

and hence the (i).
From the soliton equation (1.2) it is quite easy to check that $\rho = \frac{\lambda - r}{2}$.

Lemma 3.2. *On a generalized Sasakian-space-form M, a conformal vector field V satisfies the following:*

(i) $(\pounds_V S)(X,Y) = \frac{2n-1}{2} g(\nabla_X Dr, Y) - \frac{1}{2}\Delta r g(X,Y)$,
(ii) $\pounds_V r = -(\lambda - r)r - 2n\Delta r$.

Proof. Putting $m = 2n+1$ in (2.16) and (2.17) and then using (iii) of the last lemma, we deduce (i) and (ii) respectively.

Proof of Theorem 1.1. Taking Lie-derivative of (2.7) along V we get

$$\begin{aligned}(\pounds_V S)(X,Y) &= (2n\pounds_V f_1 + 3\pounds_V f_2 - \pounds_V f_3)g(X,Y)\\ &\quad -(3\pounds_V f_2 + (2n-1)\pounds_V f_3)\eta(X)\eta(Y)\\ &\quad +(2nf_1 + 3f_2 - f_3)(\pounds_V g)(X,Y)\\ &\quad -(3f_2 + (2n-1)f_3)[(\pounds_V \eta)(X)\eta(Y)\\ &\quad +(\pounds_V \eta)(Y)\eta(X)]. \qquad (3.1)\end{aligned}$$

Comparing (2.16) and (3.1) we have

$$\begin{aligned}\frac{2n-1}{2} g(\nabla_X Dr, Y) &= \{(\frac{\Delta r}{2} + 2n\pounds_V f_1 + 3\pounds_V f_2 - \pounds_V f_3)\\ &\quad +(2nf_1 + 3f_2 - f_3)(\lambda - r)\}g(X,Y)\\ &\quad -(3\pounds_V f_2 + (2n-1)\pounds_V f_3)\eta(X)\eta(Y)\\ &\quad -(3f_2 + (2n-1)f_3)[(\pounds_V \eta)(X)\eta(Y)\\ &\quad +(\pounds_V \eta)(Y)\eta(X)]. \qquad (3.2)\end{aligned}$$

Putting $X = Y = \xi$ in the last equation and using $\eta(\xi) = 1$ and the first equation of lamma 3.1, we see that

$$\begin{aligned}\frac{2n-1}{2}\xi(\xi r) &= (\frac{\Delta r}{2} + 2n\pounds_V f_1\\ &\quad +3\pounds_V f_2 - \pounds_V f_3 - 3\pounds_V f_2 - (2n-1)\pounds_V f_3)\\ &\quad -(2nf_1 + 3f_2 - f_3 - 3f_2 - (2n-1)f_3)(\lambda - r), \qquad (3.3)\end{aligned}$$

from which it follows that

$$\begin{aligned}\frac{2n-1}{2}\xi(\xi r) &= \frac{\Delta r}{2} + 2n(\pounds_V f_1 - \pounds_V f_3)\\ &\quad -2n(f_1 - f_3)(\lambda - r) \qquad (3.4)\end{aligned}$$

and hence

$$\begin{aligned}\Delta r &= (2n-1)\xi(\xi r) + 4n(f_1 - f_3)(\lambda - r) \\ &\quad -4n(\pounds_V f_1 - \pounds_V f_3). \qquad (3.5)\end{aligned}$$

From (2.16) we obtain

$$(\pounds_V S)(\xi, \xi) = \frac{2n-1}{2}\xi(\xi r) - \frac{1}{2}\Delta r. \qquad (3.6)$$

Taking Lie-derivative of (2.12) along V we get

$$(\pounds_V S)(\xi, \xi) = 2n(\pounds_V f_1 - \pounds_V f_3). \qquad (3.7)$$

Equating (3.6) and (3.7) we infer

$$r = \lambda. \qquad (3.8)$$

As λ is a constant, then in light of the preceding relation, we can say that the scalar curvature r is constant.
With the help of (3.8), from (1.2) we get

$$\pounds_V g = 0, \qquad (3.9)$$

which implies that V is Killing.
Taking Lie-derivative of (2.6) along V we have

$$(\pounds_V d\eta)(X, Y) = -g(\pounds_V X, \phi Y) - g(X, (\pounds_V \phi)Y). \qquad (3.10)$$

Substituting $X = \xi$ in the above equation we find that

$$g(\xi, (\pounds_V \phi)Y) = 0, \qquad (3.11)$$

which shows that either $(\pounds_V \phi)Y = 0$ or $(\pounds_V \phi)Y \perp \xi$.
From (3.9), it follows that $\pounds_V \eta = \pounds_V \xi = 0$. Thus we see that either $(\pounds_V \phi)Y$ is perpendicular with ξ or (2.18) holds. This completes the proof.

Proof of corollary 1.1. Using the fact $r = \lambda$ and lemma 3.2 we observe that $\Delta r = 0$ and therefore r becomes harmonic and hence the proof.

4. Remark

This section dedicates a valuable remark regarding Yamabe soliton metric on generalized Sasakian-space-forms as follows:

(i) A generalized Sasakian-space-form (M, f_1, f_2, f_3) is a Sasakian manifold for $f_1 - f_3 = 1$ [1]. Thus for $f_1 - f_3 = 1$ and $n = 1, M$ becomes a Sasakian manifold of dimension 3 and hence our paper is a generalization of [25].

Acknowledgement: The authors are thankful to the referee for his/her valuable suggestions and comments towards the improvement of the paper. The author Debabrata Kar is supported by the Council of Scientific and Industrial Research, India (File no : 09/028(1007)/2017-EMR-1).

References

[1] Alegre, P., Blair, D. E., and Carriazo, A., *Generalized Sasakian-space-forms*, Israel J. Math., **14** (2004), 157-183.

[2] Alegre, P. and Carriazo, A., *Structures on generalized Sasakian-space-form*, Diff. Geom. and its appl., **26** (2008), 656-666.

[3] Alegre, P. and Carriazo, A., *Submanifolds of generalized Sasakian-space-forms*, Taiwanese J. Math., **13** (2009), 923-941.

[4] Alegre, P. and Carriazo, A., *Generalized Sasakian-space-forms and conformal changes of metric*, Results in Math., **59**(3), 485-493.

[5] Blair, D. E., *Contact Manifolds in Riemannian Geometry*, Lecture Notes in Mathematics. 509, Springer-Verlag, Berlin-New York, 1976.

[6] Blair, D. E., *Riemannian Geometry of Contact and Symplectic Manifolds*, Birkhauser, Boston, 2000.

[7] Carriazo, A., Alegre, P., *Özgür*, C. and Sular, S., *New examples of generalized Sasakian-space-forms*, Geom. Struct. on Riem. Man.-Bari, **73/1**(3-4) (2015), 63-76.

[8] Chern, S. S., *What is geometry?*, Am. Math. Monthly, **97** (8) (1990), 679-686.

[9] Chow, B., Lu, P. and Ni, L., *Hamiltons Ricci flow, Graduate studies in mathematics*, American Mathematical Society, Science Press, **77** (2006).

[10] Daskalopoulos, P. and Sesum, N., *The classification of lacally conformally flat Yamabe solitons*, Adv. Math., **240** (2013), 346-369.

[11] De, U. C. and Majhi, P., ϕ-symmetric Generalized Sasakian space forms, Published in Arab Journal of Mathematical Sciences, **21** (2015), 170-178.

[12] Erken, K., *Yamabe solitons on three-dimensional normal almost paracontact metric manifolds*, arXiv: 1708.04882v2 [math.DG], 5 September, 2017.

[13] Hamilton, R. S., *The Ricci flow on surfaces, Mathematics and General Relativity*, Contemp. Math. Santa Cruz, CA, **71** (1986), Amer. Math. Soc. Providence, RI (1988), 237-262.

[14] Hsu, S. Y., *A note on compact gradient Yamabe solitons*, J. Math. Anal. Appl., **388** (2012), 725-726.

[15] Hui, S. K. and Prakasha, D. G., *On the C-Bochner curvature tensor of generalized Sasakian-space-forms*, Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, **85**(3) (2015), 401-405.

[16] Kuhnel, W., *Conformal transformations between Einstein spaces* conform. Geomatry. Aspects Math., 12(1988), Vieweg, Braunschweig, 105-146.

[17] Kowalaski, O., *An explicit classification of 3-dimensional Riemannian spaces satisfying* $R(X,Y).R = 0$, Czechoslovak Math. J., **46**(3) (1996), 427-474.

[18] Kim, U. K., *Conformally flat generalized Sasakian-space-forms and locally symmetric Generalized Sasakian-Space-Forms*, Note Mat., **26** (2006), 55-67.

[19] Mantica, C. A. and Suh, Y. J., *Pseudo-Q-symmetric Riemannian manifolds*, Int. J. Geom. Meth. Mod. Phys., **10**(5) (2013), 1350013(25 pages).

[20] Majhi, P., and De, U. C., *The structure of a class of generalized Sasakian-space-forms*, Extracta Math., **27** (2012), 301-308.

[21] Prakasha, D. G., *On Generalized Sasakian-Space-Forms with Weyl-Conformal Curvature Tensor*, Lobachevskii J. Math., **33**(3) (2012), 223-228.

[22] Prakasha, D. G. and Chavan, V., *E-Bochner curvature tensor on generalized Sasakian space forms*, Comptes Rendus Mathematique, **354**(8) (2016), 835-841.

[23] Shukla, N. V. C. and Shah R. J., *Generalized Sasakian-space-forms with concircular curvature tensor*, J. Rajasthan Acad. Phy. Sci., **10**(1) (2011), 11-24.

[24] Sular, S. and Özgür, C., *Generalized Sasakian space forms with semi-symmetric metric connections*, An. Stiint. Univ. Al. I. Cuza Iasi. Mat. (N.S.), **60**(1) (2014), 145-156.

[25] Sharma, R., *A 3-dimensional Sasakian metric as a Yamabe Soliton*, Int. J. Geom. Methods Mod. Phys., **9**(4) (2012), 1220003 (5 pages).

[26] Wang, Y., *Yamabe solitons in three dimensional kenmotsu manifolds*, Bull. Belg. Math. Soc. Stenvin, **23** (2016), $345 - 355$.

[27] Yano, K., *Concircular geomatry I, Concircular transformation*, Proc. Imp. Acad. Tokyo, **16** (1940), 195-200.

[28] Yano, K. and Kon, M., *Structures on manifolds*, Series in Pure Math, World Scientific Publ. Co., Singapore, **3**, 1985.

[29] Yano, K., *Integral formulas in Riemannian geometry*, Dekker, New York, 1970.

Pradip Majhi and Debabrata Kar
Department of Pure Mathematics
University of Calcutta
35, Bailygumge Circular Road
Kolkata 700019, West Bengal, India
mpradipmajhi@gmail.com
debabratakar6@gmail.com

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 103–112

A MINIMAL SET OF INTEGERS SHOWING THAT NO POSITIVE-DEFINITE INTEGRAL TERNARY QUADRATIC FORM IS UNIVERSAL

GREG DOYLE AND KENNETH S. WILLIAMS
(Received : 13 - 08 - 2019 ; Revised : 02 - 01 - 2020)

Abstract. We show in an elementary way that no positive-definite integral ternary quadratic form

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

can represent all the integers $1, 2, 3, 5, 6, 10, 13, 14, 17, 21, 22, 23, 29, 31$ and so cannot be universal.

1. Introduction

An integral quadratic form $q(x_1, \ldots, x_n)$ in n variables (n a positive integer) is a homogeneous polynomial of degree 2 with integer coefficients of the form

$$q := q(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} f_{ij} x_i x_j.$$

If all the coefficients f_{ij} $(i < j)$ are even then q is said to be classically integral. If $q(x_1, \ldots, x_n) \ge 0$ for all integers $x_1, \ldots, x_n$ with equality if and only if $x_1 = \cdots = x_n = 0$, then q is said to be positive-definite. For a positive integer m, if there exist integers $x_1, \ldots, x_n$ such that

$$q(x_1, \ldots, x_n) = m$$

then q is said to represent m. The representation $(x_1, \ldots, x_n)$ is said to be primitive if $\gcd(x_1, \ldots, x_n) = 1$. If $q(x_1, \ldots, x_n)$ is a positive-definite integral quadratic form which represents all positive integers, then it is said to be universal. It has been of major interest over the years to determine those quadratic forms which are universal.

2010 Mathematics Subject Classification: 11E20, 11E04
Key words and phrases: Ternary quadratic forms, universal quadratic forms.

The complete classification of all classically integral positive-definite universal quadratic forms was carried out by Conway and Schneeberger in 1993. Their work is described in [2] and [4]. They proved but did not publish the following theorem.

15 Theorem of Conway and Schneeberger. *Let $q(x_1,\ldots,x_n)$ be a classically integral positive-definite quadratic form in n variables, for some positive integer n. If q represents all positive integers up to and including* 15*, then q is universal.*

In 2000, Bhargava [2] used escalator lattices to prove that such forms represent all positive integers under a weaker condition.

15 Theorem of Bhargava. *Let $q(x_1,\ldots,x_n)$ be a classically integral positive-definite quadratic form in n variables. If q represents the nine integers*

$$1,2,3,5,6,7,10,14,15$$

then it is universal.

It was conjectured by Conway that if an integral quadratic form represents all the positive integers up to and including 290, then it must represent all positive integers. This was proved by Bhargava and Hanke [3]. Indeed they proved what is perhaps the most important result in the study of integral quadratic forms which are universal.

290 Theorem of Bhargava and Hanke. *If a positive-definite integral quadratic form in any number of variables represents all the twenty-nine integers*

$$1,2,3,5,6,7,10,13,14,15,17,19,21,22,23,26,29,$$
$$30,31,34,35,37,42,58,93,110,145,203,290,$$

then it is universal.

At the beginning of their proof, Bhargava and Hanke establish from their computation of all 3-dimensional escalator lattices that no positive-definite integral ternary quadratic form represents all of the integers 1 through 31. This result also appears in Conway's book [5], and in particular Conway highlights the form $x^2+2y^2+yz+4z^2$ (which he calls the "Little Methuselah form") which represents every integer from 1 to 30 but not 31. Both the proofs of Bhargava-Hanke and Conway are fairly concise. It is our purpose to prove a slightly stronger result that uses nothing more than a few simple results about the equivalence of quadratic forms and a theorem of Gauss about positive-definite binary quadratic forms.

Two integral quadratic forms $q(x_1, \ldots, x_n)$ and $Q(x_1, \ldots, x_n)$ which are positive-definite are called equivalent if there is a linear change of variables

$$(x_1, \ldots, x_n) \to (X_1, \ldots, X_n)$$

such that $q(X_1, \ldots, X_n) = Q(x_1, \ldots, x_n)$. We emphasize that the $n \times n$ matrix representing this change of variables must be integral and have determinant $= \pm 1$. If the determinant is $+1$, q and Q are said to be properly equivalent. We will make considerable use of the fact that two equivalent quadratic forms represent the same positive integers, and that all forms equivalent to a positive-definite quadratic form are positive-definite.

Gauss' Theorem. *Any positive-definite integral binary quadratic form is properly equivalent to a unique form $qx^2 + rxy + sy^2$, where the integers q, r and s satisfy*

$$-q < r \leq q < s \quad \text{or} \quad 0 \leq r \leq q = s.$$

The quadratic form $qx^2 + rxy + sy^2$ is called reduced.

The proof of this classical theorem can be found in [6, p. 53] or [9, p. 159]. We will make use of the fact that the smallest integer that can be represented by a reduced positive-definite integral binary quadratic form $qx^2 + rxy + sy^2$ is q, see [9, p. 170].

It is hoped that this article provides an introduction to some of the recent amazing results about identifying those positive-definite integral quadratic forms which represent all positive integers, as well as showing the interested student how a classical theorem of Gauss can be used.

2. No positive-definite integral ternary quadratic form is universal

Theorem. *No positive-definite integral ternary quadratic form can represent all of the fourteen integers*

$$1, 2, 3, 5, 6, 10, 13, 14, 17, 21, 22, 23, 29, 31. \tag{2.1}$$

Proof. Suppose that

$$f(x,y,z) := ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

is a positive-definite integral ternary quadratic form which represents all of the integers in (2.1). In particular, $f(x,y,z)$ represents 1 and clearly any such representation must be primitive. It is known that $f(x,y,z)$ is equivalent to a positive-definite integral quadratic form $g(x,y,z)$ with leading coefficient 1, say

$$g(x,y,z) := x^2 + By^2 + Cz^2 + Dxy + Exz + Fyz,$$

see, for example, [7, p. 12]. As f and g are equivalent they represent the same integers and so g represents the integers in (2.1).

Next, we wish to complete the square in the variable x in the form $g(x,y,z)$. In order to do this in such way that the coefficients of the variables remain integers, we first multiply $g(x,y,z)$ by 4 and define

$$h(x,y,z) := 4g(x,y,z) = 4x^2 + 4By^2 + 4Cz^2 + 4Dxy + 4Exz + 4Fyz.$$

Clearly, the positive-definite integral ternary quadratic form $h(x,y,z)$ represents the fourteen integers

$$4, 8, 12, 20, 24, 40, 52, 56, 68, 84, 88, 92, 116, 124. \tag{2.2}$$

Now define the positive-definite integral ternary quadratic form $j(x,y,z)$ by

$$j(x,y,z) := h(\frac{x}{2}, y, z) = x^2 + 4By^2 + 4Cz^2 + 2Dxy + 2Exz + 4Fyz.$$

Clearly $j(x,y,z)$ represents the integers in (2.2) with x even. As the coefficients of xy and xz in $j(x,y,z)$ are both even, we can complete the square in x to obtain

$$\begin{aligned} &j(x,y,z) \\ &\quad = (x + Dy + Ez)^2 + (4B - D^2)y^2 + 2(2F - DE)yz + (4C - E^2)z^2. \end{aligned}$$

Let $k(x,y,z) := x^2 + (4B - D^2)y^2 + 2(2F - DE)yz + (4C - E^2)z^2$. The integral ternary quadratic form $k(x,y,z)$ is equivalent to $j(x,y,z)$ and so is positive-definite and represents the integers in (2.2). We let $q(y,z)$ be the integral binary quadratic form

$$q(y,z) := (4B - D^2)y^2 + 2(2F - DE)yz + (4C - E^2)z^2$$

so

$$k(x, y, z) = x^2 + q(y, z).$$

As $k(x, y, z)$ is positive-definite so is $q(y, z)$. By Gauss' theorem $q(y, z)$ is equivalent to a unique positive-definite binary quadratic form $gy^2 + 2hyz + kz^2$, where g, h and k are integers satisfying

$$-g < 2h \leq g < k \quad \text{or} \quad 0 \leq 2h \leq g = k. \tag{2.3}$$

As the two positive-definite binary quadratic forms $q(y, z)$ and $gy^2 + 2hyz + hz^2$ are equivalent, there exist integers α, β, γ and δ such that

$$\alpha\delta - \beta\gamma = 1 \tag{2.4}$$

and

$$gY^2 + 2hYZ + kZ^2 = q(y, z) \quad \text{for} \quad \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} Y \\ Z \end{bmatrix}.$$

Upon expanding this, we see that

$$\begin{aligned} &(4B - D^2)y^2 + 2(2F - DE)yz + (4C - E^2)z^2 \\ &\qquad = g(\alpha y + \beta z)^2 + 2h(\alpha y + \beta z)(\gamma y + \delta z) + k(\gamma y + \delta z)^2. \end{aligned}$$

Equating coefficients of y^2, yz and z^2, we obtain

$$\begin{cases} 4B - D^2 &= g\alpha^2 + 2h\alpha\gamma + k\gamma^2, \\ 2F - DE &= g\alpha\beta + h(\alpha\delta + \beta\gamma) + k\gamma\delta, \\ 4C - E^2 &= g\beta^2 + 2h\beta\delta + k\delta^2. \end{cases} \tag{2.5}$$

We show that these equations impose parity conditions on g, h and k. Taking the equations in (2.5) modulo 2, and noting from (2.4) that

$$\alpha\delta + \beta\gamma \equiv \alpha\delta - \beta\gamma \equiv 1 \pmod{2},$$

and recalling that $t^2 \equiv t \pmod{2}$ for any integer t, we obtain

$$\begin{cases} D & \equiv g\alpha + k\gamma \pmod{2}, \\ DE & \equiv g\alpha\beta + h + k\gamma\delta \pmod{2}, \\ E & \equiv g\beta + k\delta \pmod{2}, \end{cases}$$

so that

$$\begin{aligned} g\alpha\beta + h + k\gamma\delta \equiv DE &\equiv (g\alpha + k\gamma)(g\beta + k\delta) \\ &\equiv g\alpha\beta + gk(\alpha\delta + \beta\gamma) + k\gamma\delta \end{aligned}$$

$$\equiv g\alpha\beta + gk + k\gamma\delta \pmod 2$$

and thus $h \equiv gk \pmod 2$. Hence

$$(g, h, k) \equiv (0, 0, 0), (0, 0, 1), (1, 0, 0) \text{ or } (1, 1, 1) \pmod 2. \tag{2.6}$$

Now let $\ell(x, y, z)$ be the positive-definite integral ternary quadratic form defined by

$$\ell(x, y, z) := x^2 + gy^2 + 2hyz + kz^2.$$

As $q(y, z)$ is equivalent to $gy^2 + 2hyz + kz^2$, it follows that $k(x, y, z)$ is equivalent to $\ell(x, y, z)$. Thus $\ell(x, y, z)$ represents the integers in (2). We observe that

$$\ell(0, 0, 0) = 0, \quad \ell(\pm 1, 0, 0) = 1, \quad \ell(\pm 2, 0, 0) = 4$$

and

$$\ell(m, 0, 0) = m^2 > 8 \text{ for } |m| \geq 3.$$

Hence for $\ell(x, y, z)$ to represent 8, $gy^2+2hyz+kz^2$ must represent $4, 7$ or 8. But as $gy^2+2hyz+kz^2$ is reduced, the smallest positive integer represented by $gy^2 + 2hyz + kz^2$ is g. Thus we must have

$$1 \leq g \leq 8. \tag{2.7}$$

As $gy^2 + 2hyz + kz^2$ and $gy^2 - 2hyz + kz^2$ represent the same integers, without loss of generality we may suppose that

$$h \geq 0. \tag{2.8}$$

From (2.3), (2.6), (2.7) and (2.8) we see that the only possibilities for g, h and k are

$$\begin{aligned}(g, h, k) = &(1, 0, k(\text{even}) \geq 2), && (2, 0, k \geq 2), && (3, 0, k(\text{even}) \geq 4),\\ &(3, 1, k(\text{odd}) \geq 3), && (4, 0, k \geq 4), && (4, 2, k \geq 4),\\ &(5, 0, k(\text{even}) \geq 6), && (5, 1, k(\text{odd}) \geq 5), && (5, 2, k(\text{even}) \geq 6),\\ &(6, 0, k \geq 6), && (6, 2, k \geq 6), && (7, 0, k(\text{even}) \geq 8),\\ &(7, 1, k(\text{odd}) \geq 7), && (7, 2, k(\text{even}) \geq 8), && (7, 3, k(\text{odd}) \geq 7),\\ &(8, 0, k \geq 8), && (8, 2, k \geq 8), && (8, 4, k \geq 8).\end{aligned}$$

Table 1 shows that $\ell(x, y, z)$ fails to represent at least one of the integers in (2.2). This contradiction establishes the theorem. □

Table 1: Integers in (2.2) not represented by $\ell(x,y,z)$

ternary quadratic form $\ell(x,y,z) = x^2+gy^2+2hyz+kz^2$		positive integer from the list $8,12,20,24,40,52,56,68,84,88,92,116,124$ not represented by $\ell(x,y,z)$
$x^2+y^2+kz^2$ (k(even) ≥ 2)	12	$k=6, k(\text{even}) \geq 14$
	24	$k=10,12$
	56	$k=2,8$
	92	$k=4$
$x^2+2y^2+kz^2$ ($k \geq 2$)	20	$k=6,7,10,13,15, k\geq 21$
	40	$k=3,5,11,12,14,17,19,20$
	56	$k=4,9,16$
	92	$k=2,8,18$
$x^2+3y^2+kz^2$ (k(even) ≥ 4)	8	$k=6, k(\text{even})\geq 10$
	24	$k=4$
	40	$k=8$
$x^2+3y^2+2yz+kz^2$ (k(odd) ≥ 3)	8	$k=5, k(\text{odd})\geq 9$
	20	$k=7$
	56	$k=3$
$x^2+4y^2+kz^2$ ($k\geq 4$)	12	$k=5,6,9,10,k\geq 13$
	24	$k=12$
	56	$k=8$
	84	$k=7$
	88	$k=11$
	92	$k=4$
$x^2+4y^2+4yz+kz^2$ ($k\geq 4$)	12	$k=5,6,7,9,10,k\geq 13$
	24	$k=4,11$
	84	$k=8$
	88	$k=12$
$x^2+5y^2+kz^2$ (k(even) ≥ 6)	8	$k=6,k(\text{even})\geq 10$
	40	$k=8$
$x^2+5y^2+2yz+kz^2$ (k(odd) ≥ 5)	8	$k(\text{odd})\geq 9$
	12	$k=7$
	40	$k=5$
$x^2+5y^2+4yz+kz^2$ (k(even) ≥ 6)	8	$k(\text{even})\geq 10$
	12	$k=6$
	92	$k=8$
$x^2+6y^2+kz^2$ ($k\geq 6$)	8	$k=6,k\geq 9$
	12	$k=7$
	20	$k=8$
$x^2+6y^2+4yz+kz^2$ ($k\geq 6$)	8	$k\geq 9$
	12	$k=7$

Continued on next page

Table 1 – *Continued from previous page*

ternary quadratic form $\ell(x,y,z) = x^2 + gy^2 + 2hyz + kz^2$		positive integer from the list $8, 12, 20, 24, 40, 52, 56, 68, 84, 88, 92, 116, 124$ not represented by $\ell(x,y,z)$
	20	$k = 8$
	56	$k = 6$
$x^2 + 7y^2 + kz^2$ (k(even) ≥ 8)	12	$k = 10, k(\text{even}) \geq 14$
	20	$k = 8$
	24	$k = 12$
$x^2 + 7y^2 + 2yz + kz^2$ (k(odd) ≥ 7)	12	$k = 9, k(\text{odd}) \geq 13$
	24	$k = 7$
	40	$k = 11$
$x^2 + 7y^2 + 4yz + kz^2$ (k(even) ≥ 8)	12	$k = 10, k(\text{even}) \geq 14$
	20	$k = 12$
	52	$k = 8$
$x^2 + 7y^2 + 6yz + kz^2$ (k(odd) ≥ 7)	12	$k = 9, k(\text{odd}) \geq 13$
	40	$k = 7$
	68	$k = 11$
$x^2 + 8y^2 + kz^2$ ($k \geq 8$)	20	$k = 9, 10, 13, 14, 15, 17, 18, k \geq 21$
	40	$k = 11, 12, 19, 20$
	56	$k = 16$
	92	$k = 8$
$x^2 + 8y^2 + 4yz + kz^2$ ($k \geq 8$)	20	$k = 9, 10, 13, 14, 17, 18, k \geq 21$
	40	$k = 8$
	52	$k = 20$
	56	$k = 11$
	92	$k = 12$
	116	$k = 15, 19$
	124	$k = 16$
$x^2 + 8y^2 + 8yz + kz^2$ ($k \geq 8$)	20	$k = 8, 9, 10, 12, 13, 14, 15, 17, 18, k \geq 21$
	40	$k = 16, 19$
	56	$k = 11$
	92	$k = 20$

Table 1 was constructed using the bounds on the size of solutions (x, y, z) of $\ell(x, y, z) = n$ given in [10].

The set

$$H := \{1, 2, 3, 5, 6, 10, 13, 14, 17, 21, 22, 23, 29, 31\}$$

in the Theorem is minimal in the sense that for each $h \in H$ there is a positive-definite integral ternary quadratic form that does not represent h

but does represent all of the integers in $H\backslash\{h\}$. These forms are listed in Table 2.

Table 2: Ternaries representing the integers in $H\backslash\{h\}$ but not h

h	ternaries representing integers in $H\backslash\{h\}$ but not h
1	$2x^2 + 2y^2 + 2z^2 - 2yz - xy$
2	$x^2 + 3y^2 + 4z^2 + yz$
3	$x^2 + y^2 + 5z^2$
5	$x^2 + 2y^2 + 7z^2 + yz$
6	$x^2 + y^2 + z^2 + yz$
10	$x^2 + 2y^2 + 3z^2$
13	$2x^2 + y^2 + 2z^2 + yz + zx$
14	$x^2 + y^2 + 2z^2$
17	$2x^2 + y^2 + 5z^2 - 3yz + xy$
21	$x^2 + y^2 + 2z^2 + yz$
22	$x^2 + y^2 + 3z^2 + yz$
23	$x^2 + 2y^2 + 3z^2 + yz$
29	$2x^2 + y^2 + 4z^2 - yz + zx$
31	$x^2 + 2y^2 + 4z^2 + yz$

Table 2 was constructed by a simple search through positive-definite integral ternary quadratic forms with small coefficients using the bounds given in [10] to determine the representation or non-representation of each $h \in H$. (The "Little Methuselah Form" of Conway is the example in Table 2 for $h = 31$.) Other proofs of the non-universality of positive-definite integral ternary quadratic forms are given in [1], [5] and [8]. We conclude by remarking that everything in our proof was known to Gauss and so could have been done many years ago!

References

[1] A. A. Albert, *The integers represented by sets of ternary quadratic forms,* Amer. J. Math., **55** (1933), 274-292.

[2] M. Bhargava, *On the Conway-Schneeberger fifteen theorem,* in Quadratic Forms and their Applications, Dublin 1999, Contemp. Math., Vol. 272, American Mathematical Society, Providence, RI, 2000. pp. 27-37.

[3] M. Bhargava and J. Hanke, *Universal quadratic forms and the 290-Theorem,* preprint. `math.stanford.edu/~vakil/files/290-Theorem-preprint.pdf`.

[4] J. H. Conway, *Universal quadratic forms and the fifteen theorem,* in Quadratic Forms and their Applications, Dublin 1999, Contemp. Math., Vol. 272, American Mathematical Society, Providence, RI, 2000. pp. 23-26.

[5] J. H. Conway, *The Sensual Quadratic Form,* Carus Mathematical Monographs No. 26, Mathematical Association of America, 2005.

[6] L. E. Dickson, *Modern Elementary Theory of Numbers,* University of Chicago Press, 1947.

[7] L. E. Dickson, *Studies in the Theory of Numbers,* originally published by the University of Chicago Press, 1930, republished by Chelsea Publ. Co., New York, 1961.

[8] G. Doyle and K. S. Williams, *A positive-definite ternary quadratic form does not represent all positive integers,* Integers **17** (2017), #A41, 19pp.

[9] I. Niven, H. S. Zuckerman and H. L. Montgomery, *An Introduction to the Theory of Numbers, Fifth Edition,* John Wiley & Sons, Inc., 1991.

[10] K. S. Williams, *Bounds for the representations of integers by positive quadratic forms,* Math. Mag. **89** (2016), 122-131.

Greg Doyle and Kenneth S. Williams,
School of Mathematics and Statistics,
Carleton University,
1125 Colonel By Dr, Ottawa, ON, CANADA,
K1S 5B6
gdoyle@math.carleton.ca
kenneth.williams@carleton.ca

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 113–120

BERNOULLI NUMBERS AND POLYNOMIALS

M. NARAYAN MURTY AND BINAYAK PADHY
(Received : 09 - 10 - 2019 ; Revised : 02 - 03 - 2020)

Abstract. In this article, Bernoulli numbers and polynomials are discussed. Bernoulli numbers are rational numbers. Bernoulli numbers and polynomials can be calculated using their generating functions. The values of Riemann zeta functions can be found out from Bernoulli numbers.

1. INTRODUCTION

The sequence of rational numbers B_0, B_1, B_2, B_3,..diccovered by the mathematician Jacob Bernoulli ($27^{th}December1654$ - $16^{th}August1705$) of Switzerland is called Bernoulli numbers. These numbers were discussed in his posthumous work *ArtisConjectandi* (1713). Jacob was one of the many prominent mathematicians in Bernoulli family.
Taking $B_0 = 1$, the Bernoulli numbers $B_k$are given by

$$B_k = -\frac{1}{k+1}\sum_{i=0}^{k-1}\left(\begin{array}{c} k+1 \\ i \end{array}\right)\mathrm{B}_i \text{ for } k \geq 1 \tag{1.1}$$

For any positive integer n and any integer k satisfying $0 \leq k \leq n$, the binomial coefficients $\left(\begin{array}{c} n \\ k \end{array}\right)$ are defined by

$$\left(\begin{array}{c} n \\ k \end{array}\right) = \frac{n!}{k!(n-k)!} \tag{1.2}$$

Using the factorial notation (1.2) in (1.1), the following linear equations are obtained.

$$\begin{array}{ll} 1+2B_1=0 & \text{for } k=1 \\ 1+3B_1+3B_2=0 & \text{for } k=2 \\ 1+4B_1+6B_2+4B_3=0 & \text{for } k=3 \\ 1+5B_1+10B_2+10B_3+5B_4=0 & \text{for } k=4 \\ \text{........ ....} & \end{array} \tag{1.3}$$

One can find out the Bernoulli numbers solving above equations. The first few Bernoulli numbersB_k are
$B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, $B_4 = -1/30$, $B_5 = 0$,$B_6 = 1/42$, $B_7 = 0$, $B_8 = -1/30$, $B_9 = 0$, $B_{10} = 5/66$, $B_{11} = 0$, $B_{12} = -691/2730$, $B_{13} = 0$, $B_{14} = 7/6$, $B_{15} = 0$, $B_{16} = -3617/510$, $B_{17} = 0$, $B_{18} = 43867/798$,

2010 Mathematics Subject Classification: 11B68, 11B65
Key words and phrases: Bernoulli numbers, Generating function, Bernoulli polynomials, Riemann zeta function

$B_{19} = 0,\ B_{20} = -174611/330,\ B_{21} = 0$

The Bernoulli numbers $B_{(2n+1)}(n = 0, 1, 2,)$ beyond the first are all equal to zero, while all of the numbers $B_{2n}(n = 0, 1, 2,)$ are rational numbers, which after the first, alternate in sign. That is, the Bernoulli number $B_{(4n)}(n = 1, 2,)$ is negative rational and $B_{(4n-2)}$ is positive rational. In 1734, Leonhard Euler (1707 - 1783) calculated the values of Bernoulli numbers up to

$$B_{30} = 8615841276005/14322$$

Alternative Bernoulli numbers: The Bernoulli numbers can also be taken with the second Bernoulli number B_1 as 1/2 and all other numbers remaining same. In this case, the alternate expression for Bernoulli numbers is

$$B_k = 1 - \frac{1}{k+1} \sum_{i=0}^{k-1} \binom{k+1}{i} \mathrm{B}_i \text{ for } k \geq 1 \tag{1.4}$$

Either of the expressions (1.1) and (1.4) is used for finding the values of Bernoulli numbers. The rest of the article is organized as follows. Proof of Bernoulli formula is given in Section 2. Generating function for Bernoulli numbers is mentioned in Section 3. In Section-4, Bernoulli polynomials are discussed. Some facts about Bernoulli numbers and polynomials are given in Section 5.

2. PROOF OF BERNOULLI FORMULA

The following is the proof of Bernoulli formula and the expression (1.4) for calculating the values of Bernoulli numbers.

The sum of the k^{th} power of n natural numbers is denoted as

$$S_k(n) = 1^k + 2^k + 3^k + ... + n^k = \sum_{m=1}^{n} m^k \tag{2.1}$$

For $k = 0$, we can write $S_0(n) = 1^0 + 2^0 + 3^0 + ... + n^0 \Rightarrow \mathrm{S}_0(n) = n$ (2.2)

For $k = 1$, the sum of the first n natural numbers is

$S_1(n) = 1 + 2 + 3 + ... + n$

Reversing the order of terms in RHS, the above expression can be written as

$S_1(n) = n + (n - 1) + (n - 2) + ... + 1$

Summing the above two expressions term-by-term, we get

$$2S_1(n) = (n + 1) + (n + 1) + (n + 1) + ... + (n + 1) = n(n + 1)$$

$$\Rightarrow S_1(n) = \frac{n^2}{2} + \frac{n}{2} \tag{2.3}$$

One can find out $S_k(n)$ of (2.1) using the identity

$$[(m + 1)^{(k+1)} - m^{(k+1)}] = \sum_{i=0}^{k} \binom{k+1}{i} \mathrm{m}^i$$

$$= \sum_{i=0}^{k-1} \binom{k+1}{i} \mathrm{m}^i + \binom{k+1}{k} \mathrm{m}^k$$

$$= \sum_{i=0}^{k-1} \binom{k+1}{i} \mathrm{m}^i + \frac{k+1!}{k!1!} m^k$$

$$= \sum_{i=0}^{k-1} \binom{k+1}{i} \mathrm{m}^i + (k + 1) m^k$$

$$\Rightarrow (k+1)m^k=[(m+1)^{(k+1)}-m^{(k+1)}] - \sum_{i=0}^{k-1}\binom{k+1}{i}m^i$$

Taking the summation over m from 1 to n, we have

$$(k+1)\sum_{m=1}^{n} m^k=\sum_{m=1}^{n}[(m+1)^{(k+1)}-m^{(k+1)}] - \sum_{i=0}^{k-1}\sum_{m=1}^{n}\binom{k+1}{i}m^i \quad (2.4)$$

Let us find out the value of the 1^{st} summation in RHS of the above expression.

$$\sum_{m=1}^{n}[(m+1)^2-m^2]=(2^2-1^2)+(3^2-2^2)+....+[n^2-(n-1)^2]+[(n+1)^2-n^2]$$
$$=(n+1)^2-1 \quad (2.5)$$

Replacing the power 2 of the terms in the above expression by k+1, we obtain

$$\sum_{m=1}^{n}[(m+1)^{(k+1)}-m^{(k+1)}] =(n+1)^{k+1}-1 \quad (2.6)$$

Using (2.6) in (2.4), we get

$$(k+1)\sum_{m=1}^{n} m^k=(n+1)^{k+1}-1-\sum_{i=0}^{k-1}\binom{k+1}{i}\sum_{m=1}^{n}m^i$$

Using (2.1) in the above expression, we have

$$(k+1)S_k(n)=(n+1)^{k+1}-1-\sum_{i=0}^{k-1}\binom{k+1}{i}S_i(n)$$
$$=(n+1)^{k+1}-1-\binom{k+1}{0}S_0(n)-\sum_{i=1}^{k-1}\binom{k+1}{i}S_i(n)$$

Since $\binom{k+1}{0}=1$ and $S_0(n)=n$ as per (2.2), we get

$$(k+1)S_k(n)=(n+1)^{k+1}-1-n-\sum_{i=1}^{k-1}\binom{k+1}{i}S_i(n)$$

$$\Rightarrow S_k(n)=\frac{1}{(k+1)}[(n+1)^{k+1}-1-n-\sum_{i=1}^{k-1}\binom{k+1}{i}S_i(n)] \quad (2.7)$$

Using the above expression, one can easily calculate

$$S_2(n)=\frac{n^3}{3}+\frac{n^2}{2}+\frac{n}{6} \quad (2.8)$$
$$S_3(n)=\frac{n^4}{4}+\frac{n^3}{2}+\frac{n^2}{4} \quad (2.9)$$
$$S_4(n)=\frac{n^5}{5}+\frac{n^4}{2}+\frac{n^3}{3}-\frac{n}{30} \quad (2.10)$$
$$S_5(n)=\frac{n^6}{6}+\frac{n^5}{2}+\frac{5n^4}{12}-\frac{n^2}{12} \quad (2.11)$$
$$S_6(n)=\frac{n^7}{7}+\frac{n^6}{2}+\frac{n^5}{2}-\frac{n^3}{6}+\frac{n}{42} \quad (2.12)$$
$$S_7(n)=\frac{n^8}{8}+\frac{n^7}{2}+\frac{7n^6}{12}-\frac{7n^4}{24}+\frac{n^2}{12} \quad (2.13)$$

....

Hence in general, the sum of the first n whole numbers raised to the k^{th} power taken in (2.1) can be written as

$$S_k(n)=\frac{n^{k+1}}{k+1}+\frac{1}{2}n^k+\frac{1}{6}\frac{k}{2}n^{k-1}+0(n^{k-2})-\frac{1}{30}\frac{(k-2)(k-1)k}{4!}n^{k-3}+ \ldots\ldots$$

$$\Rightarrow S_k(n)=\frac{1}{(k+1)}\sum_{i=0}^{k}\binom{k+1}{i}B_i n^{k+1-i} \quad (2.14)$$

Where B_i is the i^{th} Bernoulli number and the above formula (2.14) is called Bernoulli formula. From (2.3),(2.8), (2.9) and (2.10), we get $S_1(1)=S_2(1)=S_3(1)=S_4(1)=1$. Hence, in general we have

$$S_k(1)=1 \quad (2.15)$$

For n=1, the Bernoulli formula (2.14) can be written as

$$S_k(1)=\frac{1}{(k+1)}\sum_{i=0}^{k}\binom{k+1}{i}\mathrm{B}_i \tag{2.16}$$

Using (2.15) in (2.16), we get

$$1=\frac{1}{(k+1)}\sum_{i=0}^{k}\binom{k+1}{i}\mathrm{B}_i$$

$$\Rightarrow 1=\frac{1}{(k+1)}\sum_{i=0}^{k-1}\binom{k+1}{i}\mathrm{B}_i+\frac{1}{k+1}\binom{k+1}{k}B_k$$

$$\Rightarrow 1=\frac{1}{(k+1)}\sum_{i=0}^{k-1}\binom{k+1}{i}\mathrm{B}_i+B_k$$

$$\Rightarrow B_k=1-\frac{1}{(k+1)}\sum_{i=0}^{k-1}\binom{k+1}{i}\mathrm{B}_i \tag{2.17}$$

This expression is same as (1.4). Hence the expression (1.4) for calculating Bernoulli numbers is proved from Bernoulli formula. Taking $B_0 = 1$, one can find out the alternate Bernoulli numbers $B_1 = 1/2$, $B_2 = 1/6$ $B_3 = 0, B_4 = -1/30, B_5 = 0....$ from the above expression.

3. GENERATING FUNCTION FOR BERNOULLI NUMBERS

The generating function for Bernoulli numbers was given by Euler. Let the function $f(x)$ be such that $f^{(k)}(0) = B_k$, where $f^{(k)}$is the k^{th} derivative of f with the convention that $f^{(0)} = f$. The Taylor series expansion of the function $f(x)$ around 0 is

$$f(x)=\sum_{k=0}^{\infty}f^{(k)}(0)\frac{x^k}{k!}$$

Since $f^{(k)}(0) = B_k$, we have

$$f(x)=\sum_{k=0}^{\infty}B_k\frac{x^k}{k!} \tag{3.1}$$

Where B_k is the k^{th} Bernoulli number.
The series expansion of e^x is

$$e^x=\sum_{k=0}^{\infty}\frac{x^k}{k!} \tag{3.2}$$

Multiplying (3.1) with (3.2)

$$\begin{aligned} f(x)e^x &= \sum_{k=0}^{\infty}B_k\frac{x^k}{k!}\sum_{k=0}^{\infty}\frac{x^k}{k!}\\ &=\sum_{k=0}^{\infty}[\sum_{i=0}^{k}B_i\frac{x^i}{i!}\frac{x^{k-i}}{(k-i)!}]\\ &=\sum_{k=0}^{\infty}[\sum_{i=0}^{k}\frac{B_i}{i!(k-1)!}]x^k\\ &=\sum_{k=0}^{\infty}[\sum_{i=0}^{k}\binom{k}{i}\mathrm{B}_i]\frac{x^k}{k!}\\ &=\sum_{k=0}^{\infty}[\sum_{i=0}^{k-1}\binom{k}{i}B_i+\binom{k}{k}B_k]\frac{x^k}{k!}\\ &=\sum_{k=0}^{\infty}[\sum_{i=0}^{k-1}\binom{k}{i}B_i+B_k]\frac{x^k}{k!} \end{aligned} \tag{3.3}$$

One can be easily prove that for any value of $k \geq 1$ with Bernoulli numbers $B_0 = 1, B_1 = 1/2, B_2 = 1/6, ...$

$$\sum_{i=0}^{k-1}\binom{k}{i}B_i=k \tag{3.4}$$

Substituting (3.4) in (3.3)

$$f(x)e^x = \sum_{k=0}^{\infty}[k + B_k]\frac{x^k}{k!}$$
$$= \sum_{k=0}^{\infty}\frac{kx^k}{k!} + \sum_{k=0}^{\infty}B_k\frac{x^k}{k!}$$

Using (3.1) in RHS of the above expression

$$f(x)e^x = \sum_{k=0}^{\infty}\frac{kx^k}{k!} + f(x)$$
$$= \sum_{k=1}^{\infty}\frac{kx^k}{k!} + f(x)$$
$$= x\sum_{k=1}^{\infty}\frac{x^{k-1}}{(k-1)!} + f(x)$$
$$= xe^x + f(x)$$
$$\Rightarrow f(x) = \frac{xe^x}{e^x - 1} \tag{3.5}$$

The function f(x) defined by the above expression (3.5) is called generating function. Using (3.1) in (3.5), the generating function can be written as

$$f(x) = \frac{xe^x}{e^x - 1} = \sum_{k=0}^{\infty}B_k\frac{x^k}{k!} \tag{3.6}$$

Substituting the series expansion of e^x in the above expression $x + x^2 + \frac{x^3}{2!} + \frac{x^4}{3!} + ... = (x + x^2/2! + x^3/3! + ...)(B_0 + B_1x + B_2x^2/2! + B_3x^3/3! + ...)$ Comparing the coefficients of $x, x^2, x^3, ...$on both sides we can find the values of Bernoulli numbers as $B_0 = 1, B_1 = 1/2, B_2 = 1/6, B_3 = 0, ...$

For finding the Bernoulli numbers $B_0 = 1, B_1 = -1/2, B_2 = 1/6, B_3 = 0, ...$ a separate generating function as defined below is used.

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty}B_k\frac{x^k}{k!} \tag{3.7}$$

Multiplying both sides by $(e^x - 1)$ and using the series expansion of e^x, we obtain

$$x = (e^x - 1)\sum_{k=0}^{\infty}B_k\frac{x^k}{k!}$$
$$= \sum_{n=1}^{\infty}\frac{x^n}{n!}\sum_{k=0}^{\infty}B_k\frac{x^k}{k!}$$
$$= \sum_{n=1}^{\infty}(B_0\frac{x^n}{n!} + B_1\frac{x^{n+1}}{n!1!} + B_2\frac{x^{n+2}}{n!2!} + ...)$$

Equating the coefficients of $x, x^2, x^3, ...$on both sides, we get the following linear equations.

$B_0 = 1, 1 + 2B_1 = 0, 1 + 3B_1 + 3B_2 = 0, ...$

These are the same equations given in (1.3). So, one can find the above Bernoulli numbers from these equations.

4. BERNOULLI POLYNOMIALS

The Bernoulli polynomials $B_n(x)$ are defined by

$$B_0(x) = 1, \frac{dB_n(x)}{dx} = nB_{n-1}(x) \text{ and } \int_0^1 B_n(x)dx = 0, n \geq 1 \tag{4.1}$$

Thus the first few Bernoulli polynomials satisfying (4.1) are given by

$$B_1(x) = x - \frac{1}{2}, \qquad B_2(x) = x^2 - x + \frac{1}{6},$$
$$B_3(x) = x^3 - \frac{3x^2}{2} + \frac{x}{2}, \qquad B_4(x) = x^4 - 2x^3 + x^2 - \frac{1}{30} \tag{4.2}$$
$$B_5(x) = x^5 - \frac{5x^4}{2} + \frac{5x^3}{3} - \frac{x}{6}, \quad B_6(x) = x^6 - 3x^5 + \frac{5x^4}{2} - \frac{x^2}{2} + \frac{1}{42},$$
$$........ \qquad ....$$

These polynomials satisfy the integration in (4.1). The constant term in Bernoulli polynomial gives the corresponding Bernoulli number. Hence, the n^{th}Bernoulli number B_n is equal to the Bernoulli polynomial $B_n(x)$ for $x = 0$.

$$B_n = B_n(0) \tag{4.3}$$

Putting $x = 0$ in Bernoulli polynomials, one can find out the Bernoulli numbers $B_0 = 1, B_1 = -1/2, B_2 = 1/6, B_3 = 0, B_4 = -1/30, B_5 = 0, B_6 = 1/42, ...$ Bernoulli polynomials can also be defined by

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_k x^{n-k} \tag{4.4}$$

Thus, we have $B_1(x) = B_0(x) + B_1 = x - 1/2, B_2(x) = B_0x^2 + 2B_1x + B_2 = x^2 - x + 1/6$,etc., as given in (4.2).

Alternative Bernoulli polynomials: The alternative Bernoulli polynomials $\hat{B}_n(x)$ are defined by

$$\hat{B}_0(x) = 1, \frac{d\hat{B}_n(x)}{dx} = n\hat{B}_{(n-1)}(x) \text{ and } \int_0^1 \hat{B}_n(x)dx = 1, n \geq 1 \tag{4.5}$$

The relation between alternative Bernoulli polynomials and Bernoulli polynomials is

$$\hat{B}_n(x) = B_n(x) + nx^{(n-1)} = (-1)^n B_n(-x) \tag{4.6}$$

The above relation satisfies the integral in (4.5). Using (4.2) in (4.5), the alternate Bernoulli polynomials can be written as

$$\begin{array}{ll} \hat{B}_1(x) = x + \frac{1}{2}\ , & \hat{B}_2(x) = x^2 + x + \frac{1}{6}\ , \\ \hat{B}_3(x) = x^3 + \frac{3x^2}{2} + \frac{x}{2}\ , & \hat{B}_4(x) = x^4 + 2x^3 + x^2 - \frac{1}{30} \\ \ldots\ldots\ldots & \ldots\ldots\ldots \end{array} \tag{4.7}$$

These polynomials satisfy the integration in (4.5). The alternative Bernoulli numbers $\hat{B}_n$ can be calculated from their Bernoulli polynomials (4.7) using the relation

$$\hat{B}_n = \hat{B}_n(0) = B_n(1) \tag{4.8}$$

The alternative Bernoulli numbers are

$$\hat{B}_1 = 1/2 = -B_1, \hat{B}_n = B_n, \text{ for } n = 0, 2, 3, ...$$

The expressions for Bernoulli numbers and polynomials differ from those of their alternatives.

5. SOME FACTS ABOUT BERNOULLI NUMBERS AND POLYNOMIALS

The Bernoulli polynomials $B_n(x)$ can also be defined by the generating function

$$F(t, x) = \frac{te^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{t^n}{n!} \tag{5.1}$$

Using the generating function $F(t, x)$, we can deduce some properties of these polynomials

(a) For $x = 0$, the generating function of Bernoulli polynomials (5.1) is

$$F(t, 0) = \frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n(0) \frac{t^n}{n!} \tag{5.2}$$

Taking the generating function (3.7) for Bernoulli numbers

$$\frac{t}{e^t-1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!} \tag{5.3}$$

Comparing (5) with (5.3), we get

$$B_n(0) = B_n \tag{5.4}$$

The property (5.4) was already given in (4.3).

(b) For $x = 1$, the generating function (5.1) for Bernoulli polynomials is

$$F(t,1) = \frac{te^t}{e^t-1} = \frac{t}{1-e^{-t}} = \frac{-t}{e^{-t}-1}$$
$$\Rightarrow \mathrm{F(t,1)=F(-t,0)}$$
$$\Rightarrow \sum_{n=0}^{\infty} B_n(1)\frac{t^n}{n!} = \sum_{n=0}^{\infty} B_n(0)\frac{(-t)^n}{n!} \text{ [using (5.1)]}$$
$$=\sum_{n=0}^{\infty}(-1)^n B_n(0)\frac{t^n}{n!}$$

Comparing both sides of the above expression and using (5.4), we get

$$B_n(1) = (-1)^n B_n(0) = (-1)^n B_n \tag{5.5}$$

(c) Differentiating the generating function (5.1) with respect to x, we obtain

$$\frac{t^2 e^{xt}}{e^t-1} = \sum_{n=0}^{\infty} B_n'(x)\frac{t^n}{n!}$$
$$\Rightarrow \mathrm{tF}(\mathrm{t},x) = \sum_{n=0}^{\infty} B_n'(x)\frac{t^n}{n!}$$
$$\Rightarrow \sum_{n=0}^{\infty} B_n(x)\frac{t^{n+1}}{n!} = \sum_{n=0}^{\infty} B_n'(x)\frac{t^n}{n!}$$

Equating the coefficients of t^n on both sides

$$\frac{B_{n-1}(x)}{(n-1)!} = \frac{B_n'(x)}{n!} \Rightarrow B_n'(x) = nB_{n-1}(x) \tag{5.6}$$

The above recursive relation (5.6) is used to find higher polynomials from lower ones. This property of Bernoulli polynomials is mentioned in (4.1)

(d) One can find out Bernoulli polynomials from the generating function. Taking the series expansions of e^{xt} and e^t, the generating function (5.1) is written as

$$\frac{t[1+xt+\frac{(xt)^2}{2!}+....]}{t+\frac{t^2}{2!}+\frac{t^3}{3!}+....} = \sum_{n=0}^{\infty} B_n(x)\frac{t^n}{n!}$$
$$\Rightarrow [1 + xt + \frac{(xt)^2}{2!} + ...][1 + (\frac{t}{2} + \frac{t^2}{6} + ...)]^{-1} = \sum_{n=0}^{\infty} B_n(x)\frac{t^n}{n!}$$
$$\Rightarrow [1+xt+\frac{(xt)^2}{2!}+....][1-(\frac{t}{2}+\frac{t^2}{6}+....)+(\frac{t}{2}+\frac{t^2}{6}+....)^2....] = \sum_{n=0}^{\infty} B_n(x)\frac{t^n}{n!}$$
$$\Rightarrow (1+xt+\frac{x^2t^2}{2}+....)-(\frac{t}{2}+\frac{xt^2}{2}+....)-(\frac{t^2}{6}+\frac{xt^3}{6}+....)-...+(\frac{t^2}{4}+...).......$$
$$= B_0(x) + B_1(x)t + B_2(x)\frac{t^2}{2} + ...$$

Equating the coefficients $t^0, t, t^2, ...$ on both sides of the above expression,

we will obtain the Bernoulli polynomials as given in (4.2).

$$B_0(x) = 1; B_1(x) = x - \frac{1}{2}; B_2(x) = x^2 - x + \frac{1}{6}; ....$$

(e) Bernoulli polynomials (4.2) satisfy the following integrals.

(i)$\int_0^1 B_n(x)B_m(x)dx = (-1)^{n-1}\frac{m!n!}{(m+n)!}B_{n+m}$for $m, n \geq 1$ (5.7)
Since, $B_{n+m} = 0$, when (n+m) is odd, except B_1, we have

$$\int_0^1 B_n(x)B_m(x)dx = 0 \text{ for } m, n \geq 1 \text{ when n+m is odd} \quad (5.8)$$

For example, $\int_0^1 B_n(x)B_m(x)dx = -1/120$ using (4.2). Then (5.7) can be verified as$B_4 = -1/30$.

(ii)$\int_\alpha^\beta B_n(x)dx = \frac{B_{(n+1)}(\beta)-B_{(n+1)}(\alpha)}{(n+1)}$ (5.9)

For example, let $n = 1, \alpha = 1$ and $\beta = 2$ Then, one can easily prove that LHS=1=RHS.

Since, according to (4.3),$B_n = B_n(0)$, the above integral can be written as

$$\int_0^\beta B_n(x)dx = \frac{(B_{(n+1)}(\beta)-B_{(n+1)}}{(n+1)} \quad (5.10)$$

(iii) $\int_t^{t+1} B_n(x)dx = t^n$ (5.11)

(f) Bernoulli polynomials (4.2) obey the following relations

$$B_n(x+1)-B_n(x) = nx^{n-1} \quad (5.12)$$
$$B_n(1-x) = (-1)^n B_n(x) \quad (5.13)$$
$$(-1)^n B_n(-x) = B_n(x)+nx^{n-1} \quad (5.14)$$
$$B_n(\tfrac{1}{2}) = (\tfrac{1}{2^{n-1}} - 1)B_n \quad (5.15)$$

(g) The sum of natural numbers (2.3) can be expressed by two Bernoulli numbers.

$$1+2+3+....+n = \frac{n^2}{2}+\frac{n}{2} = \frac{1}{2}B_0n^2+B_1n \quad (5.16)$$

(h) Bernoulli numbers grow rapidly with n and the value of Bernoulli number for large n is given by

$$|B_{2n}| \approx \frac{(2n)!}{(2\pi)^{2n}}, n \to \infty \quad (5.17)$$

The true asymptotic behaviour of B_{2n} is

$$B_{2n} \approx 2(-1)^{n+1}\frac{(2n)!}{(2\pi)^{2n}} \quad (5.18)$$

(i) The Bernoulli formula (2.14) can be stated as given below using (4.4).

$$B_{k+1}(n) = \sum_{i=0}^{k+1}\binom{k+1}{i}B_in^{k+1-i} = \sum_{i=0}^{k}\binom{k+1}{i}B_in^{k+1-i} + B_{k+1}$$

$$\Rightarrow B_{k+1}(n) - B_{k+1} = \sum_{i=0}^{k}\binom{k+1}{i}\mathrm{B}_in^{k+1-i} \quad (5.19)$$

Using the above (5.19) in (2.14), the Bernoulli formula can be written as

$$S_k(n) = \frac{1}{k+1}[B_{k+1}(n) - B_{k+1}] \tag{5.20}$$

(j) The Riemann zeta function is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{5.21}$$

$$\text{Thus,}\zeta(2) = 1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+... \tag{5.22}$$

The value of this series was calculated by Euler. Slightly after 1734 Euler derived the following formula between Riemann zeta function and Bernoulli numbers.

$$\zeta(2n) = (-1)^{n+1}\frac{(2\pi)^{2n}B_{2n}}{2(2n)!}, n \geq 1 \tag{5.23}$$

Taking $n = 1$, we get $\zeta(2) = \frac{(2\pi)^2 B_2}{2(2)!}$. Since, $B_2 = 1/6$, we find $\zeta(2) = \pi^2/6$, which is the value of the sum (5.22) Similarly taking $n = 2, 3$ and 4, we get $\zeta(4) = \pi^4/90$,$\zeta(6) = \pi^6/945$ and $\zeta(8) = \pi^8/9450$, since $B_4 = (-1)/30, B_6 = 1/42$ and $B_8 = -1/30$. Thus the values of Riemann zeta function can be evaluated using Bernoulli numbers. Further, the relation $\zeta(-2n) = -\frac{B_{2n+1}}{2n+1}$ gives the trivial zeros of Riemann zeta function, since $B_{2n+1} = 0$ for $n \geq 1$.

References

[1] Burton, David M. : *Elementary Number Theory,* McGraw Hill Education (India) Private Limited (2007).

[2] Ernst, E.S. : *Bernoulli polynomials, Fourier series and zeta functions,* International Journal of Pure and Applied mathematics, 88 No.1 (2013), 65-75.

[3] Kenneth, I., Michael, R. : *A Classical Introduction to Modern Number Theory,* Springer (1990).

[4] Muthukumar, T. : *Bernoulli numbers and polynomials,* (2014), 1-7.

[5] Tony, F. : *Bernoulli polynomials,* Talks given at LSBU, (2015).

M. Narayan Murty,
Visiting Faculty, Department of Physics,
Sri Sathya Sai Institute of Higher Learning,
Prasanthi Nilayam 515134, Andhra Pradesh, India
Email: mnarayanamurty@rediffmail.com

Binayak Padhy,
Department of Physics,
Khallikote Autonomous College,
Berhampur -760001, Odisha, India

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 121–139

A FORBIDDEN-MINOR CHARACTERIZATION FOR THE CLASS OF GRAPHIC MATROIDS WHICH YIELD THE CO-GRAPHIC ELEMENT-SPLITTING MATROIDS

S. B. DHOTRE

(Received : 17 - 10 - 2019 ; Revised : 05 - 01 - 2020)

Abstract. The element splitting operation on a graphic matroid, in general may not yield a cographic matroid. In this paper, we give a necessary and sufficient condition for the graphic matroid to yield cographic matroid under the element splitting operation.

1. Introduction

Fleischner [5] introduced the idea of splitting a vertex of degree at least three in a connected graph and used the operation to characterize Eulerian graphs. Raghunathan, Shikare and Waphare [11] extended the splitting operation from graphs to binary matroids. $M_{x,y}$ denotes the splitting matroid obtained by applying splitting operation on a binary matroid M, by a pair of elements $\{x, y\}$ of M.

Slater [14] specified the *n-point splitting operation* on a graph in the following way: Let G be a graph and u be a vertex of degree at least $2n-2$ in G. Let H be the graph obtained from G by replacing u by two adjacent vertices u_1, u_2 such that each point formerly joined to u is joined to exactly one of u_1 and u_2 so that in H, $\deg(u_1) \geq n$ and $\deg(u_2) \geq n$. We say that H arises from G by n-point splitting operation. We illustrate this construction with the help of Figure 1.

2010 Mathematics Subject Classification: 05B35
Key words and phrases: binary matroids, splitting operation, element splitting operation

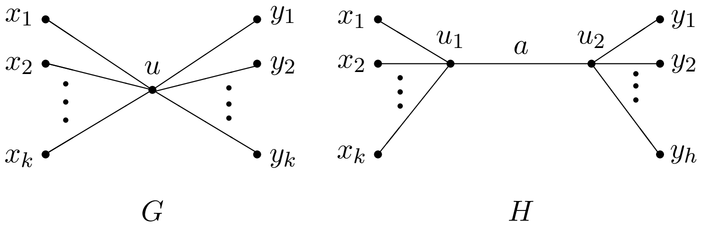


Figure 1

If $X = \{x_1, x_2, ...x_k\}$ be the set of edges incident at u_1 then we denote H by G'_X. Tutte [15] characterized 3-connected graphs in terms of edge addition and 3-point splitting. Slater [14] proved that the class of 2-connected graphs is the class of graphs obtained from K_3 by finite sequence of edge addition and 2-point splitting; and if G is n-connected and H arise from G by n-point splitting, then H is n-connected. Further, he classified 4-connected graphs using n-point splitting operation (see [14]).

Shikare and Azadi [13, 1] extended the notion of n-point splitting operation on graphs to binary matroids as follows.

Definition 1.1. Let M be a binary matroid on a set E and A be a matrix over $GF(2)$ that represents the matroid M. Suppose that X is a subset of $E(M)$. Let A'_X be the matrix that is obtained by adjoining an extra row to A with this row being zero everywhere except in the columns corresponding to the elements of X where it takes the value 1 and then adjoining an extra column (corresponding to a) with this column being zero everywhere except in the last row where it takes the value 1. Suppose M'_X be the vector matroid of the matrix A'_X. The transition from M to M'_X is called the element splitting operation.

We call the matroid M'_X as the element splitting matroid. If $|X| = 2$ and $X = \{x, y\}$. We denote the matroid M'_X by $M'_{x,y}$. Azadi [1], characterized circuits of the element splitting matroid in terms of circuits of the binary matroids as follows.

Proposition 1.1. Let $M = (E, C)$ be a binary matroid together with the collection of circuits C. Suppose $X \subseteq E$ and $a \notin E$. Then $M'_X = (E \cup \{a\}, C')$ where $C' = C_0 \cup C_1 \cup C_2$ and

$C_0 = \{C \in C \mid C$ contains an even number of elements of X $\}$;

$C_1 =$ The set of minimal members of $\{C_1 \cup C_2 \mid C_1, C_2 \in C, C_1 \cap C_2 \neq \phi$ and each of C_1 and C_2 contains an odd number of elements of X such that $C_1 \cup C_2$ contains no member of C_0 $\}$;

$C_2 = \{C \cup \{a\} \mid C \in C$ and C contains an odd number of elements of X };

Various properties concerning the element splitting matroids have been studied in [1, 6, 13].

The element splitting operation on a graphic (cographic) matroid may not yield a graphic (cographic) matroid. Dalvi, Borse and Shikare [3, 4] characterized graphic (cographic) matroids whose element splitting matroids are graphic (cographic) when $|X| = 2$. In fact, they proved the following result.

Theorem 1.2. *The element splitting operation, by any pair of elements, on a graphic (cographic) matroid yields a graphic (cographic) matroid if and only if it has no minor isomorphic to $M(K_4)$, where K_4 is the complete graph on 4 vertices.*

In this paper, we obtain a forbidden-minor characterization for graphic matroids whose element splitting matroid is cographic when $|X| = 2$. The main result in this paper is the following theorem.

Theorem 1.3. *The element splitting operation, by any pair of elements, on a graphic matroid yields a cographic matroid if and only if it has no minor isomorphic to $M(K_4)$, where K_4 is the complete graph on 4 vertices.*

2. Properties of Element Splitting Operation

In this section, we provide necessary Lemmas which are used in the proof of Theorem 2.4. Dalvi, Borse and Shikare [3] proved the following useful Lemma.

Lemma 2.1. *Let x and y be distinct elements of a binary matroid M and let $r(M)$ denote the rank of M. Then, using the notations introduced in Section 1,*

(i) $M_{x,y} = M'_{x,y} \setminus \{a\}$;
(ii) $M = M'_{x,y}/\{a\}$;
(iii) $r(M'_{x,y}) = r(M) + 1$;
(iv) Every cocircuit of M is a cocircuit of the matroid $M'_{x,y}$;
(v) if $\{x, y\}$ is a cocircuit of M then $\{a\}$ and $\{x, y\}$ are cocircuits of $M'_{x,y}$;
(vi) If $\{x, y\}$ does not contain a cocircuit, then $\{x, y, a\}$ is a cocircuit

of $M'_{x,y}$;
(vii) $M'_{x,y} \setminus x/y \cong M \setminus x$;
(viii) If M *is graphic and* x, y *are adjacent edges in a corresponding graph, then* $M'_{x,y}$ *is graphic;*
(ix) $M'_{x,y}$ *is not Eulerian.*

The following two results are well known minor based characterizations of graphic and cographic matroids (see [10]).

Theorem 2.2. *A binary matroid is graphic if and only if it has no minor isomorphic to* $F_7, F_7^*, M^*(K_{3,3})$ *or* $M^*(K_5)$.

Theorem 2.3. *A binary matroid is cographic if and only if it has no minor isomorphic to* $F_7, F_7^*, M(K_{3,3})$ *or* $M(K_5)$.

Notation. For the sake of convenience, let
$\mathcal{F} = \{F_7,\ F_7^*,\ M(K_5),\ M(K_{3,3})\}$.

Theorem 2.4. *The element splitting operation, by any pair of elements, on a graphic matroid yields a cographic matroid if and only if it has no minor isomorphic to* $M(K_4)$, *where* K_4 *is the complete graph on 4 vertices.*

3. Properties of Element Splitting Operation

In this section, we provide necessary Lemmas which are used in the proof of Theorem 2.4. Dalvi, Borse and Shikare [3] proved the following useful Lemma.

Lemma 3.1. *Let* x *and* y *be distinct elements of a binary matroid* M *and let* $r(M)$ *denote the rank of* M. *Then, using the notations introduced in Section 1,*
(i) $M_{x,y} = M'_{x,y} \setminus \{a\}$;
(ii) $M = M'_{x,y}/\{a\}$;
(iii) $r(M'_{x,y}) = r(M) + 1$;
(iv) Every cocircuit of M *is a cocircuit of the matroid* $M'_{x,y}$;
(v) if $\{x, y\}$ *is a cocircuit of* M *then* $\{a\}$ *and* $\{x, y\}$ *are cocircuits of* $M'_{x,y}$;
(vi) If $\{x, y\}$ *does not contain a cocircuit, then* $\{x, y, a\}$ *is a cocircuit of* $M'_{x,y}$;
(vii) $M'_{x,y} \setminus x/y \cong M \setminus x$;
(viii) If M *is graphic and* x, y *are adjacent edges in a corresponding*

graph, then $M'_{x,y}$ is graphic;
(ix) $M'_{x,y}$ is not Eulerian.

The following two results are well known minor based characterizations of graphic and cographic matroids (see [10]).

Theorem 3.2. *A binary matroid is graphic if and only if it has no minor isomorphic to $F_7, F_7^*, M^*(K_{3,3})$ or $M^*(K_5)$.*

Theorem 3.3. *A binary matroid is cographic if and only if it has no minor isomorphic to F_7, F_7^*, $M(K_{3,3})$ or $M(K_5)$.*

Notation. For the sake of convenience, let
$\mathcal{F} = \{F_7,\ F_7^*,\ M(K_5),\ M(K_{3,3})\}$.

In the following Lemma, we provide a necessary condition for a graphic matroid whose element splitting matroid is not cographic.

Lemma 3.4. *Let M be a graphic matroid and let $x, y \in E(M)$ such that $M'_{x,y}$ is not cographic. Then M has a minor isomorphic to $M(K_4)$ or there is a minor N of M such that no two elements of N are in series and $N'_{x,y} \setminus \{a\}/\{x\} \cong F$ or $N'_{x,y} \setminus \{a\}/\{x, y\} \cong F$ or $N'_{x,y} \cong F$ or $N'_{x,y}/\{x\} \cong F$ or $N'_{x,y}/\{y\} \cong F$ or $N'_{x,y}/\{x, y\} \cong F$ for some $F \in \mathcal{F}$.*

Proof. Suppose that $M'_{x,y}$ is not cographic and M has no minor isomorphic to $M(K_4)$. Since $M'_{x,y}$ is not cographic $M'_{x,y} \setminus T_1/T_2 \cong F$ for some $T_1, T_2 \subseteq E(M'_{x,y})$. Let $T'_i = T_i - \{a, x, y\}$ for $i = 1, 2$. Then $T'_i \subseteq E(M)$ for each i. Let $N = M \setminus T'_1/T'_2$. Then $N'_{x,y} = M'_{x,y} \setminus T'_1/T'_2$. Let $T''_i = T_i - T'_i$ for $i = 1, 2$. Then $N'_{x,y} \setminus T''_1/T''_2 \cong F$. If $a \in T''_2$, then F is a minor of $M'_{x,y}/a$ and hence, by Lemma 3.1(ii), F is a minor of M. Since M is graphic F_7, F_7^* can not be the minors of M. So $F = M(K_5)$ or $F = M(K_{3,3})$, but both $M(K_5)$ and $M(K_{3,3})$ have minor isomorphic to $M(K_4)$. Consequently M has a minor isomorphic to $M(K_4)$. Which is a contradiction. Suppose $a \in T''_1$. By Lemma 3.1 (i), $M_{x,y} = M'_{x,y} \setminus a$. Hence F is a minor of $M_{x,y}$. It follows from Theorem 2.3 of [12] that N does not contain a 2-cocircuit and further, $N_{x,y}/x \cong F$ or $N_{x,y}/\{x, y\} \cong F$. This implies that $N'_{x,y} \setminus \{a\}/x \cong F$ or $N'_{x,y} \setminus \{a\}/\{x, y\} \cong F$. Suppose that $a \notin T''_1 \cup T''_2$. Hence $a \notin T_1 \cup T_2$. If $T''_1 \cup T''_2 = \phi$, then $N'_{x,y} \cong F$. If $T''_2 = \phi$, then $N'_{x,y} \setminus x \cong F$ or $N'_{x,y} \setminus y \cong F$ or $N'_{x,y} \setminus \{x, y\} \cong F$. In the first case, a forms a 2-cocircuit with x or y which ever is remained, and in the later case, a is a coloop, both are contradictions. Hence $T''_2 \neq \phi$. If $T''_1 \neq \phi$ then, by Lemma 3.1(vii), F is minor of M, which is a contradiction. Hence $T_1'' = \phi$ and $N'_{x,y}/x \cong F$ or $N'_{x,y}/y \cong F$ or $N'_{x,y}/x, y \cong F$. Assume that N contains a 2-cocircuit Q. By Lemma 3.1(iv), Q is a 2-cocircuit in $N'_{x,y}$. Since F is 3-connected, it does not contain a 2-cocircuit. It follows that $N'_{x,y}$ is not isomorphic to F. Hence $N'_{x,y} \setminus \{a\}/x \cong F$ or $N'_{x,y} \setminus \{a\}/\{x, y\} \cong F$ or $N'_{x,y}/\{x\} \cong F$ or $N'_{x,y}/\{y\} \cong F$ or $N'_{x,y}/\{x, y\} \cong F$. If $Q \cap \{x, y\} = \phi$, then it is retained in all these cases and thus F has a 2-cocircuit, which is a contradiction. If $Q = \{x, y\}$, a contradiction follows from Lemma 3.1(v). Hence Q contains exactly one of x and y. Suppose that $x \in Q$. Then $N'_{x,y}/y \ncong F$. Let x_1 be the other element of Q. Let $L = N/x_1$. Then L is a minor of M in which no pair of elements is in series. Further, $L'_{x,y} = N'_{x,y}/x_1 \cong N'_{x,y}/x$. Thus we have $L'_{x,y} \setminus \{a\} \cong F$ or $L'_{x,y} \setminus \{a\}/y \cong F$ or $L'_{x,y} \cong F$ or $L'_{x,y}/y \cong F$. Since $L_{x,y} \cong L'_{x,y} \setminus \{a\}$, and x, y are in series in $L_{x,y}$, it follows that $L'_{x,y} \setminus \{a\} \ncong F$ and also $L'_{x,y} \setminus \{a\}/y \cong L'_{x,y} \setminus \{a\}/x$. If $y \in Q$, then $N'_{x,y}/x \ncong F$. Also, $L'_{x,y} \cong N'_{x,y}/y$. In this case we get $L'_{x,y} \setminus \{a\}/x \cong F$ or $L'_{x,y} \cong F$ or $L'_{x,y}/x \cong F$. □

Definition 3.5. Let M be a graphic matroid in which no two elements are in series and let $F \in \mathcal{F}$. We say that M is minimal with respect to F if there exist two elements x and y of M such that $M'_{x,y} \setminus \{a\}/\{x\} \cong F$ or $M'_{x,y} \setminus \{a\}/\{x, y\} \cong F$ or $M'_{x,y} \cong F$ or $M'_{x,y}/\{x\} \cong F$ or $M'_{x,y}/\{x, y\} \cong F$.

Corollary 3.6. *Let M be a graphic matroid. For $x, y \in E(M)$, the matroid $M'_{x,y}$ is cographic if and only if M has no minor isomorphic to a minimal matroid with respect to any $F \in \mathcal{F}$.*

Proof. If $M'_{x,y}$ is not cographic for some x, y then, by Lemma 3.4, M has a minor N in which no two elements are in series and $N'_{x,y} \setminus \{a\}/\{x\} \cong F$ or $N'_{x,y} \setminus \{a\}/\{x, y\} \cong F$ or $N'_{x,y} \cong F$ or $N'_{x,y}/\{x\} \cong F$ or $N'_{x,y}/\{y\} \cong F$ or $N'_{x,y}/\{x, y\} \cong F$ for some $F \in \mathcal{F}$. If $N'_{x,y}/y \cong F$ but $N'_{x,y}/x \ncong F$, then interchange roles of x and y.

Conversely, suppose that M has a minor N isomorphic to a minimal matroid with respect to some $F \in \mathcal{F}$. Then $N'_{x,y} \setminus \{a\}$ or $N'_{x,y}/\{x\}$ or $N'_{x,y}/\{x, y\}$ or $N'_{x,y} \cong F$, for some $x, y \in E(M)$. We conclude that $M'_{x,y}$ has a minor isomorphic to F and hence it is not cographic. □

In the following Lemma, we prove some basic properties of graphic minimal matroids.

Lemma 3.7. *Let M be a graphic matroid corresponding to a graph G. If M is minimal with respect to some $F \in \mathcal{F}$, then*
(i) M has neither loops nor coloops;
(ii) every pair of parallel elements of G must contain either x or y;
(iii) x and y cannot be parallel in G;
(iv) if $M'_{x,y} \cong F_7^*$ or $M(K_{3,3})$ then G is simple, and there is no odd circuit of M containing both x and y, and also there is no even circuit of M containing precisely one of x and y;
(v) if $M'_{x,y}/\{x\} \cong F_7^*$ or $M(K_{3,3})$ then G is simple and there is no 3-circuit of M containing both x and y;
(vi) if $M'_{x,y}/\{x\} \cong F_7$ or $M(K_5)$, then G has exactly one pair of parallel elements and there is no 3-circuit of M containing both x and y;
(vii) if $M'_{x,y}/\{x, y\} \cong F$ then G is simple and there is no 3 or 4-circuit of G containing both x and y; and
(viii) $M'_{x,y}$ is not isomorphic to F_7 or $M(K_5)$.

Proof. The proof is straightforward. □

Lemma 3.8. [8] *Let $F \in \mathcal{F}$ and let M be a binary matroid such that either $M_{x,y}/\{x\} \cong F$ or $M_{x,y}/\{x, y\} \cong F$ for some pair $x, y \in E(M)$. Then the following statements hold.*
(i) M has neither loops nor coloops;
(ii) x and y can not be parallel in M;
(iii) if x_1 and x_2 are parallel elements of M , then one of them is either x or y ;
(iv) if $M_{x,y}/\{x, y\} \cong F$, then M has at most one pair of parallel elements;
(v) if $M_{x,y}/\{x\} \cong M(K_{3,3})$ or $M_{x,y}/\{x, y\} \cong M(K_{3,3})$, then every odd circuit of M contains x or y; and
(vi) if $M_{x,y}/\{x\} \cong M(K_5)$ or $M_{x,y}/\{x, y\} \cong M(K_5)$, then every odd cocircuit of M contains x or y.

A matroid is said to be Eulerian if its ground set can be expressed as a union of circuits [16].

Lemma 3.9. [12] *Suppose x and y are non adjacent edges of a graph G and $M = M(G)$. If $M_{x,y}/\{x, y\}$ is Eulerian, then either G is Eulerian or the end vertices of x and y are precisely the vertices of odd degree.*

4. MAIN RESULTS

In this section, we obtain the minimal cographic matroids corresponding to each of the four matroids $F_7, F_7^*, M(K_{3,3})$ and $M(K_5)$ and use them to give a proof of Theorem 2.4.

The minimal graphic matroids corresponding to the matroid F_7 and F_7^* are characterized by Shikare and Dalvi [3] in the following Lemma 4.1 and Lemma 4.2

Lemma 4.1. *Let M be a graphic matroid. Then M is minimal with respect to the matroid F_7 if and only if M is isomorphic to one of the cycle matroids $M(G_1)$, $M(G_2)$ or $M(G_3)$, where G_1, G_2 and G_3 are the graphs of figure 2.*

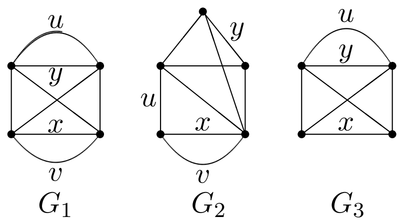


Figure 2

Lemma 4.2. *Let M be a graphic matroid. Then M is minimal with respect to the matroid F_7^* if and only if M is isomorphic the cycle matroid $M(G_4)$ or $M(G_5)$, where G_4 and G_5 are the graphs of figure 3.*

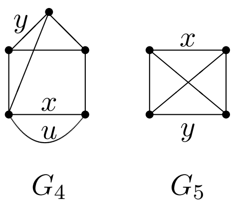


Figure 3

In the following Lemma, the minimal matroids corresponding to the matroid $M(K_{3,3})$ are characterized.

Lemma 4.3. *Let M be a graphic matroid. Then M is minimal with respect to the matroid $M(K_{3,3})$ if and only if M is isomorphic to $M(G_6)$, $M(G_7)$, $M(G_8)$, $M(G_9)$, or $M(G_{10})$, where G_6, G_7, G_8, G_9, G_{10} are the graphs of Figure 4.*

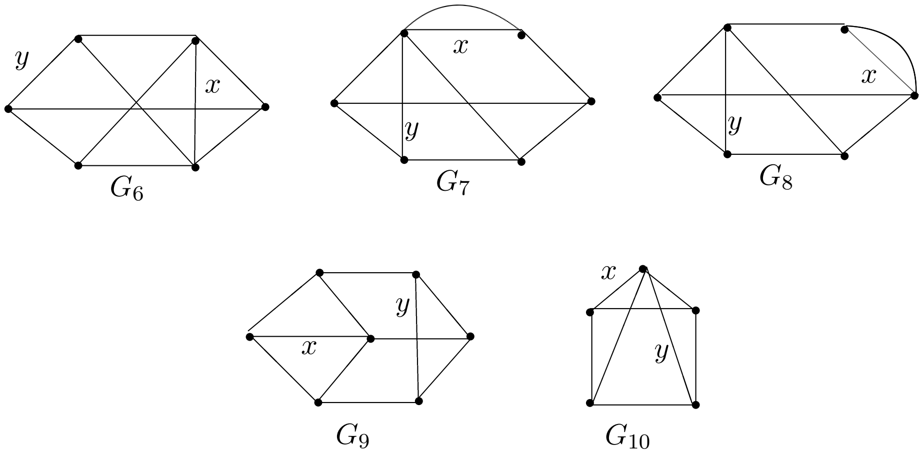


Figure 4

Proof. We have $M'(G_6)_{x,y} \setminus \{a\}/\{x\} \cong M(K_{3,3})$, $M'(G_7)_{x,y} \setminus \{a\}/\{x\} \cong M(K_{3,3})$, $M'(G_8)_{x,y} \setminus \{a\}/\{x\} \cong M(K_{3,3})$, $M'(G_9)_{x,y} \setminus \{a\}/\{x,y\} \cong M(K_{3,3})$, $M'(G_{10})_{x,y} \cong M(K_{3,3})$. Therefore $M(G_6), M(G_7), M(G_8), M(G_9)$, $M(G_{10})$ are minimal matroids with respect to the matroid $M(K_{3,3})$.

Conversely, suppose that M is a minimal matroid with respect to the matroid $M(K_{3,3})$. Then there exist elements x and y of M such that $M'_{x,y} \setminus \{a\}/\{x\} \cong M(K_{3,3})$ or $M'_{x,y} \setminus \{a\}/\{x,y\} \cong M(K_{3,3})$ or $M'_{x,y} \cong M(K_{3,3})$ or $M'_{x,y}/\{x\} \cong M(K_{3,3})$ or $M'_{x,y}/\{x,y\} \cong M(K_{3,3})$ and further, no two elements of M are in series.

Case (i) $M'_{x,y} \setminus \{a\}/\{x\} \cong M(K_{3,3})$.

By Lemma 3.1 (i), $M_{x,y}/\{x\} \cong M(K_{3,3})$. Since $r(M_{x,y}/\{x\}) = r(M(K_{3,3})) = 5$. $M_{x,y}$ is a matroid of rank 6 and $|E(M)| = 10$. In the light of the Lemma 3.1(iii), the matroid M has rank 5 and its ground set has 10 elements. Let G be a connected graph corresponding to M. Then G has 6 vertices, 10 edges, and has no vertex of degree 2. Hence, by Lemma 3.7, G has minimum degree at least 3 since no two elements are in series.Thus the degree sequence of G is (5,3,3,3,3,3) or (4,4,3,3,3,3). By Harary [[7], p 223], each simple connected graph with these degree sequences is isomorphic to one of the graphs of Figure 5 below.

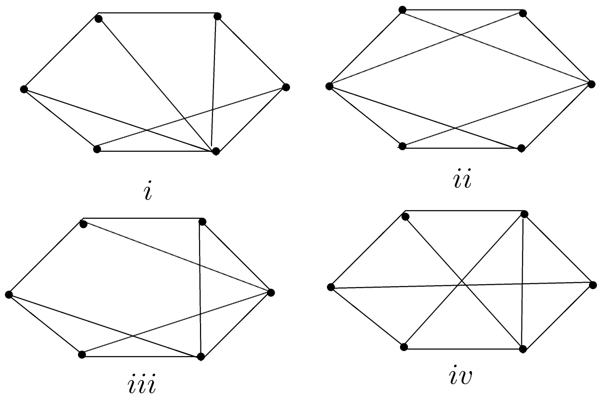


Figure 5

By the nature of the circuits of $M(K_{3,3})$ or $M_{x,y}$ and by Lemma 3.7, it follows that G can not have, (i) two or more edge disjoint triangles, (ii) a circuit of size 3 or 4 or 6 containing both x and y. Since each of the graphs $(i), (ii)$ and (iii) of Figure 5 contains two or more edge disjoint triangles, we discard them.The graph (iv) of Figure 5, is isomorphic to the graph G_6 in the statement of the Lemma.

Suppose G is a multigraph. Then by Lemma 3.3 of [2], G is isomorphic to G_7 or G_8 of Figure 4.

Case (ii). $M'_{x,y} \setminus \{a\}/\{x, y\} \cong M(K_{3,3})$.

By Lemma 3.1(i), $M_{x,y}/\{x, y\} \cong M(K_{3,3})$. As $r(M(K_{3,3})) = 5$, $r(M_{x,y}) = 7$. Hence $r(M) = 6$ and $|E(M)| = 11$. Let G be connected graph corresponding to M. Then G has 7 vertices, 11 edges and has minimum degree at least 3. Therefore the degree sequence of G is (4,3,3,3,3,3,3). It follows from Lemma 3.7 that G can not have (i) more than two edge disjoint triangles; (ii) a cycle of size other than 6 which contains both x and y; and (iii) a triangle and a 2-circuit which are edge disjoint.

Then, by case (ii) of Lemma 3.3 of [2], G is isomorphic to G_9 of Figure 4.

Case (iii). $M'_{x,y} \cong M(K_{3,3})$.

Since $r(M(K_{3,3})) = 5$ and $|E(M(K_{3,3}))| = 9$, $r(M) = 4$ and $|E(M)| = 8$. Therefore M cannot have $M(K_{3,3})$ and $M(K_5)$ as a minor. We conclude that M is cographic. Let G be a graph which corresponds to the matroid M. Then G has 5 vertices and 8 edges. As M is graphic and cographic,

G is planar. By Lemma 3.7(iv), G is simple. Since M has no coloop and no two elements are in series, minimum degree in G is at least 3. There is only one non-isomorphic simple graph with 5 vertices and 8 edges [7], see Figure 6.

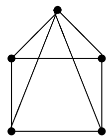


Figure 6 : Simple graph with 5 vertices and 8 edges

Thus, G is isomorphic to this graph, which is nothing but the graph G_{10} of Figure 4.

Case (iv). $M'_{x,y}/\{x\} \cong M(K_{3,3})$.

As in the above cases, considering the rank of $M(K_{3,3})$ we have $r(M) = 5$ and $|E(M)| = 9$. If M is not cographic then M has minor a isomorphic to $M(K_{3,3})$ or $M(K_5)$. Suppose M has a minor isomorphic to $M(K_{3,3})$ and $M'_{x,y}/\{x\} \cong M(K_{3,3})$. Then $M \cong M(K_{3,3})$.
Now $x, y \in M$ and since M is isomorphic to $M(K_{3,3})$, x, y lie in a 4-circuit say C but then $C - \{x\}$ is a triangle in $M'_{x,y}/\{x\} \cong M(K_{3,3})$, a contradiction to the fact that $M'_{x,y}/\{x\}$ is bipartite and does not contain any odd circuit. Thus M does not contain $M(K_{3,3})$ as a minor.
Let M have minor a isomorphic to $M(K_5)$. Since $r(M) = 5$, $|E(M)| = 9$ and $|E(M(K_5)| = 10$, so M does not contain $M(K_5)$ as a minor. Hence M is cographic.
Let $M = M(G)$ be graphic matroid. Then G has 6 vertices and 9 edges. Further, G is simple and planar. Also, minimum degree in G is at least 3.

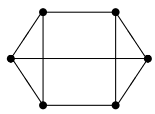


Figure 7

Thus G is isomorphic to the graph of Figure 7 (see [7]). If a triangle of G contains neither x nor y then it is preserved in $M'_{x,y}/\{x\}$, a contradiction. Hence x belongs to a triangle of G. This gives rise to a 4-circuit in $M'_{x,y}$

containing x and a. Hence, we get a 3-circuit in $M'_{x,y}/\{x\}$, a contradiction. Thus M is not isomorphic to the graph in Figure 7.

Case (v). $M'_{x,y}/\{x,y\} \cong M(K_{3,3})$.

Then $r(M) = 6$ and $|E(M)| = 10$. If M is not cographic then, M has minor isomorphic to $M(K_{3,3})$ or $M(K_5)$.
Suppose M has a minor isomorphic to $M(K_{3,3})$ then, since M is graphic, G must be the graph in Figure 8 (see [7]).

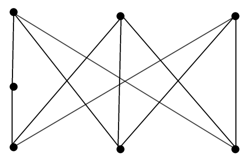


Figure 8

Any two edges in the graph of Figure 8 are in a 4-cycle or in a 5-cycle so are x and y also. Then these circuits (cycles in G) will be preserved in $M'_{x,y}$ and hence a 2-circuit or a triangle is formed in $M'_{x,y}/\{x,y\} \cong M(K_{3,3})$, a contradiction to the fact that $M'_{x,y}/\{x,y\} \cong M(K_{3,3})$ is a simple bipartite matroid. Thus M has no minor isomorphic to $M(K_{3,3})$.
Suppose that M has a minor isomorphic to $M(K_5)$ but then $|E(M)| = 10$, hence $M = M(K_5)$. Consequently $r(M) = 4$ and this is a contradiction to the fact that $r(M) = 6$. Thus M does not have a minor isomorphic to $M(K_5)$. Thus M is cographic. We conclude that M is graphic as well as cographic. Suppose G is a graph corresponding to M, then G has 7 vertices and 10 edges. This implies that G has at least one vertex of degree 2, which is a contradiction to the fact that M has no 2-cocircuit. Therefore, the situation $M'_{x,y}/\{x,y\} \cong M(K_{3,3})$ does not occur. □

Finally, we characterize minimal matroids corresponding to the matroid $M(K_5)$ in the following Lemma.

Lemma 4.4. *Let M be a graphic matroid. Then M is minimal with respect to the matroid $M(K_5)$ if and only if M is isomorphic to one of the seven matroids $M(G_{11})$, $M(G_{12})$, $M(G_{13})$, $M(G_{14})$, $M(G_{15})$, $M(G_{16})$ and $M(G_{17})$, where G_{11}, G_{12}, G_{13}, G_{14}, G_{15}, G_{16} and G_{17} are the graphs of Figure 9.*

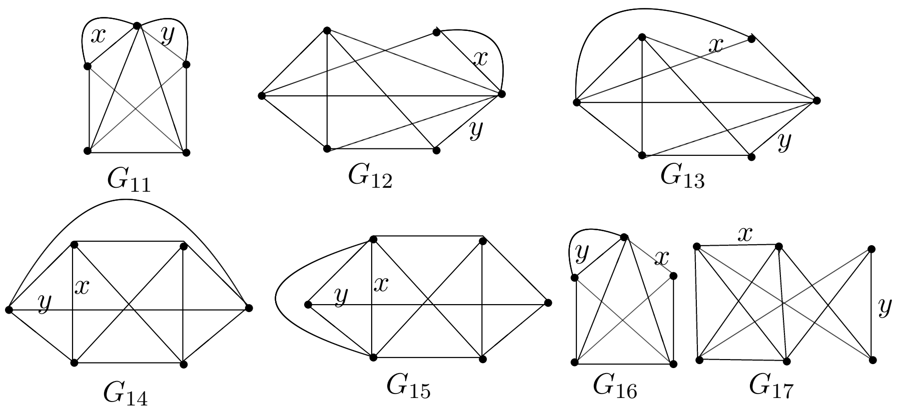


Figure 9

Proof. We have $M'(G_{11})_{x,y} \setminus \{a\}/\{x\} \cong M(K_5)$, $M'(G_{12})_{x,y} \setminus \{a\}/\{x,y\} \cong M(K_5)$, $M'(G_{13})_{x,y} \setminus \{a\}/\{x,y\} \cong M(K_5)$, $M'(G_{16})_{x,y}/\{x\} \cong M(K_5)$, $M'(G_{14})_{x,y} \setminus \{a\}/\{x,y\} \cong M(K_5)$, $M'(G_{15})_{x,y} \setminus \{a\}/\{x,y\} \cong M(K_5)$ $M'(G_{17})_{x,y}/\{x,y\} \cong M(K_5)$. Therefore $M(G_{11})$, $M(G_{12})$, $M(G_{13})$, $M(G_{14})$, $M(G_{15})$, $M(G_{16})$ and $M(G_{17})$ are minimal matroids with respect to the matroid $M(K_5)$.

Conversely, suppose that M is a minimal matroid with respect to the matroid $M(K_5)$. Then there exist elements x and y of M such that $M'_{x,y} \setminus \{a\}/\{x\} \cong M(K_5)$ or $M'_{x,y} \setminus \{a\}/\{x,y\} \cong M(K_5)$ or $M'_{x,y} \cong M(K_5)$ or $M'_{x,y}/\{x\} \cong M(K_5)$ or $M'_{x,y}/\{x,y\} \cong M(K_5)$ and also, M does not contain a 2-cocircuit.

Case (i). $M'_{x,y} \setminus \{a\}/\{x\} \cong M(K_5)$.

By Lemma 3.1(i), $M(G)_{x,y}/\{x\} \cong M(K_5)$. Hence, by similar argument as in Lemma 3.4, in [4], M is isomorphic to the cycle matroid $M(G_{11})$, where G_{11} is the graph of Figure 9.

Case (ii). $M'_{x,y} \setminus \{a\}/\{x,y\} \cong M(K_5)$.

By Lemma 3.1(i), $M(G)_{x,y}/\{x,y\} \cong M(K_5)$. Then $r(M(K_5)) = 4$, $r(M_{x,y}) = 6$ and $|E(M)| = 12$. Let G be a connected graph corresponding to M. Then G has 6 vertices, 12 edges and has minimum degree at least 3. Suppose that G is simple. By Lemma 4.4 of [2], there are 5 non isomorphic simple graphs each with 6 vertices and 12 edges, out of which, two graphs are discarded in case (*ii*) of Lemma 4.4 of [2]. So, only three graphs are remaining and these graphs are not planar. These graphs are given in Figure 10.

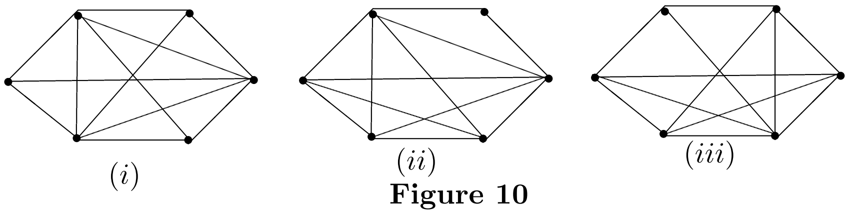


Figure 10

In graph (*i*) of Figure 10, not every odd cocircuit of M contains x or y, a contradiction to the fact that if $M(G)_{x,y}/\{x, y\} \cong M(K_5)$, then every odd cocircuit contain x or y otherwise if both of them are absent then that odd cocircuit of M is the odd cocircuit in $M(G)_{x,y}/\{x, y\} \cong M(K_5)$. Consequently $M(G)_{x,y}/\{x, y\} \cong M(K_5)$ becomes non Eulerian. In each of the graphs (*ii*) and (*iii*) of Figure 10, x and y together belong to a 3-cycle or a 4-cycle, a contradiction to Lemma 3.8

Suppose that G is not simple. Then, by Lemma 3.8 (*iv*), G has exactly one pair of parallel edges. Then G can be obtained from a simple graph on 6 ver-

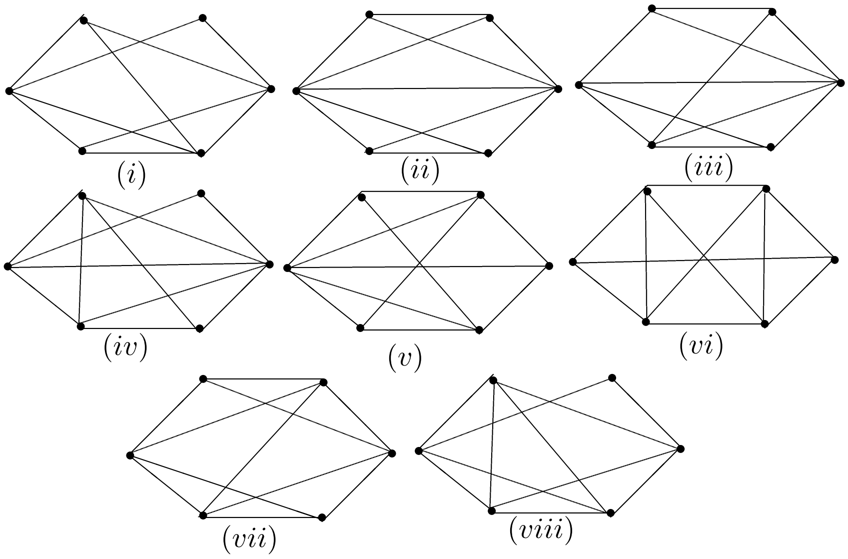


Figure 11

tices and 11 edges by adding a parallel edge.

There are 8 non isomorphic connected simple graphs, each with 6 vertices and 11 edges ([7] pp. 223) as shown in Figure 11. It follows that by Lemma 3.8(*i*) and (*iv*) and Lemma 3.9 that G can not be obtained from the graphs

$(ii), (iii)$ and (vii) of Figure 11. Suppose that G is obtained from the graphs (i) or (iv). Then G is isomorphic to one of the four graphs of Figure 12. By Lemma 3.8, G is not isomorphic to each of the two graphs (i) and (ii) of Figure 12. Hence G is isomorphic to graphs (iii) and (iv) of Figure 12, which are nothing but the graphs G_{12} and G_{13} of Figure 9.

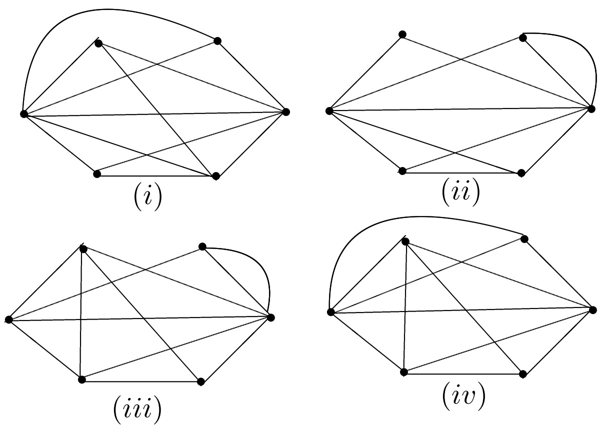


Figure 12

By Lemma 3.8, G can not be obtained from graph (v) of Figure 11. Suppose that G is obtained from graph $(viii)$ of Figure 11. Then G is isomorphic to one of the two graphs of Figure 13. By Lemma 3.8 (iv) and 3.9, G is not isomorphic to graph (i) of Figure 13. By Lemma 3.8 (ii) and (iv) and the fact that $M(K_5)$ does not contain odd cocircuit, G can not be isomorphic to the graph (ii) of Figure 13.

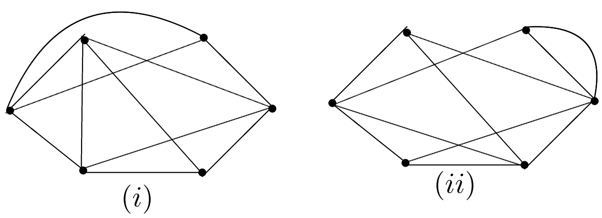


Figure 13

Suppose that G is obtained from graph (vi) of Figure 11. Then G is isomorphic to one of the two graphs G_{14} and G_{15} of Figure 9.

Case (iii). $M'_{x,y} \cong M(K_5)$.

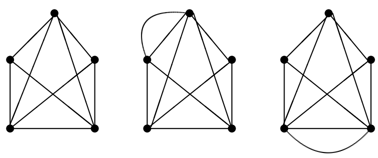
Then a contradiction follows from Lemma 3.7 $(viii)$.

Case(iv). $M'_{x,y}/\{x\} \cong M(K_5)$.

Subcase (i). Suppose that M is not cographic.

Let G be a graph that corresponds to the matroid M. Since $r(M(K_5)) = 4$, $r(M'_{x,y}) = 5$. Further, $r(M) = 4$ and $|E(M)| = 10$. Therefore, G has 5 vertihapterces and 10 edges. M has no minor isomorphic to $M(K_{3,3})$ as $K_{3,3}$ has 6 vertices. Suppose that M has a minor isomorphic to $M(K_5)$ then $M \cong M(K_5)$. By Lemma 3.7, x, y can not both be in a triangle. Hence x and y are not adjacent. Let C^* ba cocircuit of M containing y but not x such that $|C^*| = 4$, since we can always find a set of 4 edges containing y incident to some vertex in K_5, that set of edges is a cocircuit of $M(K_5)$. Then $C^* \cup \{a\}$ becomes a cocircuit of $M'_{x,y}/\{x\} \cong M(K_5)$, a contradiction to the fact that $M(K_5)$ is Eulerian and $|C^* \cup \{a\}| = 5$. Thus M is cographic.

Since M is graphic and cographic, G is planar. By Harary [7], there is no simple planar graph with 5 vertices and 10 edges. Hence G must be non-simple. Then, by Lemma 3.7(vi), G has exactly one pair of parallel edges. G can be obtained from a simple planar graph with 5 vertices and 9 edges by adding an edge in parallel. By Harary [7], every simple planar graph with 5 vertices and 9 edges is isomorphic to graph (i) of Figure 14. Therefore, G is isomorphic to the graph (ii) or (iii) of Figure 14. In graph (iii), there are two edge-disjoint 3-cutsets (i.e. 3-cocircuits in $M(G)$). Hence, one of them is preserved in $M'_{x,y}/\{x\}$, and this is a contradiction. Thus, G is isomorphic to graph (ii) of Figure 16, which is nothing but the graph G_{16} of Figure 9.



(i) (ii) (iii)

Figure 14

Case (v). $M'_{x,y}/\{x, y\} \cong M(K_5)$.

Subcase (i). Suppose that M is not cographic.

Let G be a graph which corresponds to the matroid M. Since $r(M(K_5)) = 4$, $r(M'_{x,y}) = 6$. So, $r(M) = 5$. Further, $|E(M)| = 11$.
Suppose that M has a minor isomorphic to $M(K_{3,3})$. Let G be the connected graph corresponding to M. Then G is the graph with 6 vertices and 11 edges. By Lemma 3.7 (*vii*), G has to be simple. Consequently, G is isomorphic to one of the following two graphs in Figure 15 (see Harary [7]).

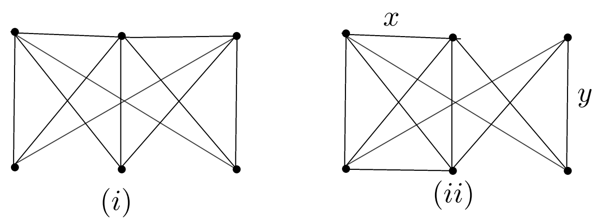


Figure 15

In the graph (*i*) of Figure 15, any two edges are either in a 3-cycle or a 4-cycle. The elements x and y can not be in a 3-circuit or a 4-circuit because such circuit becomes a loop or a 2-circuit in $M'_{x,y}/\{x,y\} \cong M(K_5)$. This is a contradiction to the fact that $M(K_5)$ is simple.
In the graph (*ii*) of Figure 15, $M'_{x,y}/\{x,y\} \cong M(K_5)$ for the edges x, y shown in the graph (*ii*) of Figure 15. Hence the graph (*ii*) which is G_{17} of Figure 9, is a minimal graph with respect to $M(K_5)$. Suppose that M has a minor isomorphic to $M(K_5)$. Since $r(M(K_5)) = 4$, $r(M'_{x,y}) = 6$. So, $r(M) = 5$. Further, $|E(M)| = 11$. Let G be a graph corresponding to the matroid M. Thus G is the graph with 6 vertices and 11 edges with $M(K_5)$ as a minor. In fact G is the graph shown in Figure 16.

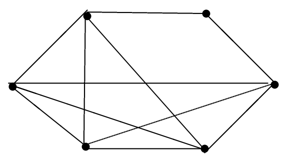


Figure 16

Observe that the graph of Figure 16, contains a vertex of degree 2, i.e. there is a 2-cocircuit in M. This is a contradiction to the fact that M does not have any pair of elements in series (i.e. 2-cocircuit). Hence M has no minor isomorphic to $M(K_5)$.

Thus M is cographic. Consequently, G is the planar graph with 6 vertices and 11 edges and has minimum degree at least 3. By Lemma 3.7 (vii), G is simple.

There are in all 9 non-isomorphic simple graphs with 6 vertices and 11 edges (see [7]). Out of which, four graphs are non-planar and two graphs contain a degree two vertex, remaining 3 graphs are shown in Figure 17. Here, G cannot have a 3 or a 4-cycle containing both x and y. Also, each

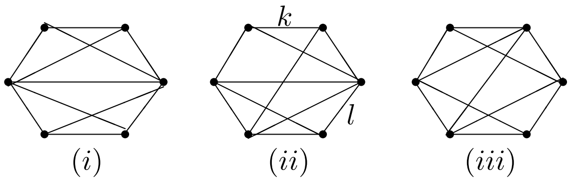


Figure 17

3-cocycle and a 5-cocycle of G must contain x or y. These conditions are not satisfied for the graphs (i) and (iii) for any pair of edges x, y. The choice for $(x,\ y)$ in graph (ii) is $(k,\ l)$. But then $M'_{x,y}/\{x, y\}$ is not eulerian. Hence the cycle matroids of the graphs in Figure 17 are not minimal with respect to $M(K_5)$. □

Now, we use Lemmas 4.1, 4.2, 4.3 and 4.4 to prove Theorem 2.4.

Proof of Theorem 2.4. Let M be a graphic matroid. On combining Corollary 3.6 and Lemmas 4.1, 4.2, 4.3 and 4.4, it follows that $M_{x,y}$ is cographic for every pair $\{x, y\}$ of elements of M if and only if M has no minor isomorphic to any of the matroids $M(G_i),\ i = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17$, where the graphs G_i are shown in the statements of the Lemmas 4.1, 4.2, 4.3 and 4.4. However, one can check that each of the matroids $M(G_i),\ i = 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17$ has matroid $M(G_5)$ of Figure 5 as a minor. This completes the proof of the theorem. □

Acknowledgement. I thank the anonymous referee for his valuable suggestions which have helped to improve the presentation of the paper.

References

1. Azadi G., *Generalized splitting operation for binary matroids and related results*, Ph.D. Thesis, University of Pune (2001).
2. Borse Y. M., Shikare M. M. and Dalvi Kiran, Excluded-Minor characterization for the class of cographic splitting matroids, *Ars Combin.* 105(2014), 219-237.

3. Dalvi K.V., Borse Y.M., Shikare M.M., Forbidden-minors for the class of graphic element-splitting matroids. *Discussiones Mathematicae Graph Theory,* 29 (2009), 629-644.
4. Dalvi K.V., Borse Y.M., Shikare M.M.,Forbidden-minors for the class of cographic element-splitting matroids. *Discussiones Mathematicae Graph Theory*, 31(2011), 601-606.
5. Fleischner H., *Eulerian Graphs and Related Topics*, Part 1, Vol 1, North Holland, Amsterdam (1990).
6. Habib Azanchiler, *Some new operations on matroids and related results* , Ph. D. Thesis, University of Pune (2005).
7. Harary F., *Graph Theory*, Addison-Wesley, Reading, 1969.
8. Malavadkar P.P., Dhotre S. B. and Shikare M. M.,Forbidden-Minors for the Class of Cographic Matroids Which Yield the Graphic Element-Splitting Matroids *South East Asian Bulletin of Mathematics* 43 (2019), 105 -119.
9. Naiyer P., *The Splitting operation for binary matroids and excuded minors for certain classes of splitting Matroids,* Ph.D. Thesis, University of Pune (2011).
10. Oxley G., *Matroid Theory*, Oxford University Press, Oxford, 1992.
11. Raghunathan T. T., Shikare M. M. and Waphare B. N., Splitting in a binary matroid, *Discrete Math.* 184 (1998), 267-271.
12. Shikare M. M. and Waphare B. N., Excluded-Minors for the class of graphic splitting matroids, *Ars Combinatoria* 97 (2010), 111.
13. Shikare M. M., The Element Splitting Operation for Graphs, Binary Matroids and Its Applications, *The Mathematics Student*, 80(2010), 85-90.
14. Slater P. J., A Classification of 4-connected graphs, *J.Combin. Theory*, 17 (1974), 282-298.
15. Tutte W. T., A theory of 3-connected graphs, *Indag. Math.* 23 (1961), 441-455.
16. Welsh D. J. A., *Matroid Theory*, Academic Press, London, 1976.

S. B. Dhotre
Department of Mathematics
Savitribai Phule Pune University, Pune-411007 (India)
E-mail: dsantosh2@yahoo.co.in.

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 141–152

COMMON FIXED POINT THEOREMS IN GENERALIZED METRIC SPACES IN ANY NUMBER OF ARGUMENTS

BHUMI AMIN AND RAJENDRA G. VYAS

(Received : 11 - 10 - 2019 ; Revised : 29 - 01 - 2020)

Abstract. Inspired by Roldáns' notion of G-metric in any number of arguments (namely, G_n-metric space) and fixed point theorems on G_n-metric space, we prove common fixed point theorem on G_n-metric space. Further, we prove common fixed point theorem on a preordered G_n-metric space that extend the results in this field.

1. Introduction

Taking the inspiration by the immense work done in fixed point theory (see ([1]), and references therein) and the applications of fixed point theorems in various fields, in this paper we discuss about common fixed point theorems in G_n-metric space. In 2003, Mustafa and Sims ([3]) introduced the concept of G-metric. In 2014, Roldán ([5]) introduced G-metric space in any number of arguments, namely G_n-metric space and later, derived fixed point theorems on the introduced preordered metric space. We have made natural extensions of common fixed point results for G_n-metric space as well as for the preordered G_n-metric space.

Notations.

(1) Φ is the family of all mappings $\varphi : [0, \infty) \to [0, \infty)$ verifying: if $\{t_m\}_{m \in \mathbb{N}} \subset [0, \infty)$ and $\varphi(t_m) \to 0$ then $t_m \to 0$.
(2) Ψ is the family of all altering distance functions.
(3) FixT is the collection of all fixed points of map T.
(4) Co(T,g) is the collection of all coincidence points of maps T and g.

2010 Mathematics Subject Classification: 47H10, 47H09, 46T99, 54H25
Key words and phrases: G-metric space, fixed point, contractive map, common fixed point

Definition 1.1 (**Altering distance function**). An altering distance function is a continuous, non-decreasing mapping $\psi : [0, \infty) \to [0, \infty)$ such that $\psi^{-1}(0) = 0$.

Definition 1.2 (**Picard sequence**). Let T and g : $X \to X$ be self maps defined on a non-empty set X.

(1) A sequence $\{x_n\}$ is a Picard sequence of T if $x_{n+1} = Tx_n$, for all $n \in \mathbb{N}$.
(2) A sequence $\{x_n\}$ is a Picard sequence of (T, g) if $gx_{n+1} = Tx_n$, for all $n \in \mathbb{N}$. Some authors say this sequence is based on initial point x_0.
(3) $T : X \to X$ is said to be a Picard operator if for all initial points $x_0 \in X$, the Picard sequence of T based on x_0 converges to a fixed point of T.

Definition 1.3 (**Preordered set**). We say that $\preceq$ is a partial preorder on X (or (X,$\preceq$) is a preordered set) if the following properties hold:

- Reflexivity: x $\preceq$ x, for all x $\in$ X.
- Transitivity: If x, y, z $\in$ X verify x $\preceq$ y and y $\preceq$ z, then x $\preceq$ z.

Definition 1.4. Let (X,$\preceq$) be a preordered set and let T, g : X $\to$ X be two mappings. We say that T is:

- (g,$\preceq$)-non-decreasing if Tx $\preceq$ Ty for all x, y $\in$ X such that gx $\preceq$ gy;
- (g,$\preceq$)-non-increasing if Tx $\succeq$ Ty for all x, y $\in$ X such that gx $\preceq$ gy;

Definition 1.5 (**Mustafa and Sims(**[4]**)**). A generalized metric (or G-metric) space on X is a mapping G : $X^3 \to [0, \infty)$ verifying, for all x, y, z, a $\in$ X;

(G_1) G(x,x,x) = 0.
(G_2) G(x,x,y) > 0 if x $\neq$ y.
(G_3) G(x,x,y) $\leq$ G(x,y,z) if y$\neq$ z.
(G_4) G(x,y,z) = G(y,z,x) = G(z,x,y)=...(symmetry in all three variables).
(G_5) G(x,y,z) $\leq$ G(x,a,a) + G(a,y,z) (rectangle inequality).

Definition 1.6 ([5]). A G_n-metric on X is a mapping G : $X^n \to [0, \infty)$ verifying, for all $x_1, x_2, ..., x_n$, x, y, a $\in$ X;

(A_1) G(x,x,...,x) = 0.
(A_2) If $x_1 \neq x_2$ then $G(x_1, x_2, x_3, ..., x_n) > 0$.

(A_3) If $\sigma : \Lambda_n \to \Lambda_n$ is a permutation, then $G(x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(n)}) = G(x_1, x_2, ..., x_n)$ (symmetry in all its variables).

(A_4) $G(x_1, x_2, ..., x_n) \leq G(x_1$,a,...,a$) + G(a, x_2, ..., x_n)$ (multidimensional inequality).

(A_5) If x $\neq$ y $\neq x_3 \neq ... \neq x_{n-1} \neq x_n$, then
G(x,x,...,x,y) $\leq$ G(x,y,$x_3, x_4, ..., x_n$).

Definition 1.7. Let (X,G) be a G_n-metric space, let $\{x_m\} \subseteq X$ be a sequence and let x $\in X$ be a point.
We will say that:

- $\{x_m\}$ G-converges to x ($x_m \underset{\to}{G}$ x) if
$$\lim_{m_1, m_2, ..., m_{n-1} \to \infty} G(x_{m_1}, x_{m_2}, ..., x_{m_{n-1}}, x) = 0,$$
that is, for all $\epsilon > 0$ there exists $m_0 \in \mathbb{N}$ such that
$$G(x_{m_1}, x_{m_2}, ..., x_{m_{n-1}}, x) < \epsilon,$$
for all $m_1, m_2, ..., m_{n-1} \geq m_0$.
- $\{x_m\}$ is a G-Cauchy sequence if
$$\lim_{m_1, m_2, ..., m_{n-1}, m_n \to \infty} G(x_{m_1}, x_{m_2}, ..., x_{m_{n-1}}, x_{m_n}) = 0,$$
that is, for all $\epsilon > 0$ there exists $m_0 \in \mathbb{N}$ such that
$$G(x_{m_1}, x_{m_2}, ..., x_{m_{n-1}}, x_{m_n}) < \epsilon,$$
for all $m_1, m_2, ..., m_{n-1}, m_n \geq m_0$.
- A subset $A \subseteq X$ is said to be G-complete if every G-Cauchy sequence in A is convergent in A.
- A mapping F : $X^N \to X$ is G-continuous if for all N sequences $\{{x_m}^1\}, \{{x_m}^2\}, ..., \{{x_m}^N\} \subseteq$ X such that ${x_m}^i \underset{\to}{G} z^i$ for all i $\in$ 1,2,...N, we have,
$$F({x_m}^1, {x_m}^2, ..., {x_m}^N) \underset{\to}{G} F(z^1, z^2, ..., z^N).$$

Remark. $\Psi \subset \Phi$

Remark. For convenience, sometimes, we write G(x,y,...y)=G(x, $[y]^{n-1}$).

Lemma 1.1 (([1]), Lemma, p.52)**.** *If $T(X) \subseteq g(X)$, then there exists a Picard sequence of (T, g) based on any x_0.*

Lemma 1.2. *Let there be two mappings $\psi \in \Psi$ and $\varphi \in \Phi$. Let $\{t_n\} \subset [0, \infty)$ be a sequence such that $\psi(t_{n+1}) \leq \psi(t_n) - \varphi(t_n)$, for all $n \in \mathbb{N}$, then $t_n \to 0$.*

Proof. Case(i): Assume that there exists some $n_0 \in \mathbb{N}$ such that $t_{n_0} \leq t_{n_0+1}$. Since ψ is non-decreasing, $\psi(t_{n_0+1}) \leq \psi(t_{n_0}) - \varphi(t_{n_0}) \leq \psi(t_{n_0}) \leq$

$\psi(t_{n_0+1})$. Therefore, $\psi(t_{n_0}) = \psi(t_{n_0+1})$ and $\varphi(t_{n_0}) = 0$. Since $\varphi \in \Phi$, $t_{n_0} = 0$.
Therefore, $\psi(t_{n_0+1}) \leq \psi(0) - \varphi(0) = 0$. So $\psi(t_{n_0+1}) = 0$ and also $t_{n_0+1} = 0$. Repeating this argument, $t_n = 0$ for all n $\geq n_0$ that is, $t_n \to 0$.

Case(ii): Assume $t_{n+1} < t_n$, for all n $\in \mathbb{N}$. Therefore $\{t_n\}$ is strictly decreasing sequence of non-negative real numbers. Then, there exists L $\geq$ 0 such that $t_n \to L$. Since $0 \leq \varphi(t_n) \leq \psi(t_n) - \psi(t_{n+1})$ for all n $\in \mathbb{N}$ and ψ is continuous, then $\varphi(t_n) \to 0$ and $\phi \in \Phi$ implies $t_n \to 0$. □

In our main results, we use the indirect method to prove that the iterative sequence is Cauchy. The following lemma states the result about non-Cauchy sequences which helps us to prove common fixed point theorems.

Lemma 1.3 (([5]), Lemma, p.6). *Suppose that a sequence $\{x_m\}_{m\in\mathbb{N}}$ in a G_n-metric space (X,G) is not G-Cauchy and verifies $G(x_{m+1}, [x_m]^{n-1}) \to 0$. Then there exist $\epsilon_0 > 0$ and two subsequences $\{x_{m(k)}\}_{k\in\mathbb{N}}$ and $\{x_{n(k)}\}_{k\in\mathbb{N}}$ such that, for all $k \in \mathbb{N}$,*

$$k \leq n(k) < m(k) < n(k+1),$$
$$G(x_{m(k)}, [x_{n(k)}]^{n-1}) > \epsilon_0 \text{ and } G(x_{m(k)-1}, [x_{n(k)}]^{n-1}) \leq \epsilon_0.$$

Furthermore, for all $p_1, p_2, ..., p_n \geq 0$, we have

$$\lim_{k\to\infty} G(x_{n(k)+p_1}, x_{n(k)+p_2}, ..., x_{n(k)+p_{n-1}}, x_{m(k)+p_n}) = \epsilon_0.$$

Moreover,

$$\lim_{k\to\infty} G(x_{n(k)-1}, x_{n(k)}, x_{n(k)+1}, ..., x_{n(k)+n-3}, x_{m(k)-1}) = \epsilon_0.$$

In particular,

$$\lim_{k\to\infty} G(x_{m(k)}, [x_{n(k)}]^{n-1}) = \epsilon_0.$$

2. Main Results.

In this section, we present theorems concerning existence and uniqueness of common fixed points for T, g : X $\to$ X using the following contractive condition on (X,G), where (X,G) is the G_n-metric.
For all $gx_1, gx_2, ..., gx_n \in$ X,

$$\psi(G(Tx_1, Tx_2, ..., Tx_n)) \leq \psi(G(gx_1, gx_2, ..., gx_n)) - \phi(G(gx_1, gx_2, ..., gx_n)), \tag{2.1}$$

where $\psi \in \Psi$, $\phi \in \Phi$.

We prove common fixed point theorems in context of preordered G_n-metric spaces taking inspiration from Theorem 5.3.1 in [1].

Theorem 2.1. *Let (X,G,$\preceq$) be a preordered G_n-metric space and let $T : X \to X$ be a self mapping. Suppose that the following conditions hold:*

(1.) *(X,G) is complete;*
(2.) *there exists $x_0 \in X$ such that $gx_0 \preceq Tx_0$;*
(3.) *T(X) $\subseteq$ g(X);*
(4.) *T is (g,$\preceq$)-non-decreasing;*
(5.) *g is continuous and commutes with T;*
(6.) *T is G-continuous;*
(7.) *there exists $\psi \in \Psi$ and $\phi \in \Phi$ such that for all x_1, x_2,..., $x_n \in X$,*
$\psi(G(Tx_1, Tx_2, ..., Tx_n)) \leq \psi(G(gx_1, gx_2, ..., gx_n))$ - $\phi(G(gx_1, gx_2, ..., gx_n))$, *for* $gx_1 \preceq gx_2 \preceq ... \preceq gx_n$.

Then T and g have, at least, a coincidence point, that is, there exists z $\in$ X such that Tz = gz.
Furthermore, assume that for all x, y $\in$ Co(T,g), there exists w $\in$ X such that gx $\preceq$ gw and gy $\preceq$ gw. Then:

- *gx = gy for all x, y $\in$ Co(T,g), and*
- *T and g have a unique common fixed point a, which is a = Tx, where x $\in$ Co(T,g).*

Proof. Let $x_0 \in$ X such that $gx_0 \preceq Tx_0$. Since T(X) $\subseteq$ g(X), by Lemma 1.1, there exists Picard sequence $\{x_n\}$ of (T,g) based on x_0, that is,

$$gx_n = Tx_n, \text{ for all n} \geq 0.$$

Also, T is a (g,$\preceq$)-non-decreasing mapping so we observe that

$$gx_0 \preceq Tx_0 = gx_1 \text{ implies } gx_1 = Tx_0 \preceq Tx_1 = gx_2.$$

Inductively, we obtain

$$gx_0 \preceq gx_1 \preceq gx_2 \preceq ... \preceq gx_n \preceq ...$$

If there exists $n_0 \in \mathbb{N}$ such that $gx_{n_0} = gx_{n_{0+1}} = Tx_{n_0}$ means T and g have a coincidence point, which completes the existence part in the proof. So we assume the contrary that is $gx_n \neq gx_{n+1}$, for all n $\in \mathbb{N}$ which implies $G(gx_{n+1}, gx_{n+2}, ..., gx_{2n}) > 0$.

$$\psi(G(gx_{n+1}, gx_{n+2}, ..., gx_{2n})) = \psi(G(Tx_n, Tx_{n+1}, ..., Tx_{2n-1}))$$
$$\leq \psi(G(gx_n, gx_{n+1}, ..., gx_{2n-1})) - \phi(G(gx_n, gx_{n+1}, ..., gx_{2n-1})),$$

as $gx_n \preceq gx_{n+1} \preceq ... \; gx_{2n-1}$.

From Lemma 1.2, we deduce that

$$G(gx_n, gx_{n+1}, ..., gx_{2n-1}) \to 0 \text{ as n} \to \infty.$$

Since $G((gx_{n+1}, [gx_n]^{n-1})) \leq G(gx_n, gx_{n+1}, ..., gx_{2n-1})$, it follows that $G((gx_{n+1}, [gx_n]^{n-1})) \to 0$ as n $\to \infty$.

Next, we prove that $\{gx_n\}$ is G-Cauchy in X.
Assume that $\{gx_n\}$ is not Cauchy in (X,G). Then by Lemma 1.3, there exists $\epsilon_0 > 0$ and two subsequences $\{gx_{m(k)}\}_{k\in\mathbb{N}}$ and $\{gx_{n(k)}\}_{k\in\mathbb{N}}$ such that, for all k $\in \mathbb{N}$,

$$k \leq n(k) < m(k) < n(k+1),$$
$$G(gx_{m(k)}, [gx_{n(k)}]^{n-1}) > \epsilon_0 \text{ and } G(gx_{m(k)-1}, [gx_{n(k)}]^{n-1}) \leq \epsilon_0.$$

Furthermore, for all $p_1, p_2, ..., p_n \geq 0$, we have

$$\lim_{k\to\infty} G(gx_{n(k)+p_1}, gx_{n(k)+p_2}, ..., gx_{n(k)+p_{n-1}}, gx_{m(k)+p_n}) = \epsilon_0.$$

Take $p_1 = p_2 = ... = p_n =$ -1 $\in \mathbb{Z}$,
$lim_{n\to\infty} \; G([gx_{n(k)}]^{n-1}, gx_{m(k)}) = lim_{n\to\infty} \; G([gx_{n(k)-1}]^{n-1}, gx_{m(k)}) = \epsilon_0$.
Now,

$$\begin{aligned} \psi(G([gx_{n(k)}]^{n-1}, gx_{m(k)})) &= \psi(G([Tx_{n(k)-1}]^{n-1}, Tx_{m(k)-1})) \\ &\leq (\psi - \phi)(G([gx_{n(k)-1}]^{n-1}, gx_{m(k)-1})). \end{aligned} \tag{2.2}$$

Take $\{t_k = G([gx_{n(k)}]^{n-1}, gx_{m(k)})\}$, $\{s_k = G([gx_{n(k)-1}]^{n-1}, gx_{m(k)})\}$. Then $\{t_k\}$ and $\{s_k\}$ are sequences converging to the same limit ϵ_0 and they satisfy $\psi(t_k) \leq \psi(s_k)$ - $\phi(s_k)$, for all k.
Therefore, $\phi(s_k) \leq \psi(s_k)$ - $\psi(t_k)$.
Applying limit as k $\to \infty$, since $\psi \in \Psi$, we have

$$\lim_{k\to\infty} \phi(s_k) \leq \psi(\epsilon_0) \text{ - } \psi(\epsilon_0) = 0.$$

Since $\phi \in \Phi$, $\lim_{k\to\infty} s_k = 0$. This implies $\epsilon_0 = 0$ which is not true.
As a consequence, $\{gx_n\}$ is Cauchy in (X,G).

Taking into account that (X,G) is complete, there exists z $\in$ X such that $gx_n \to$ z. As g is G-continuous $ggx_n \to$ gz. Therefore, $gTx_{n-1} \to$ gz and since T and g commute it implies that $Tgx_{n-1} \to$ gz.
Similarly, using G-continuity of T, we have $Tgx_n \to$ Tz. By uniqueness of

limit, we have, Tz = gz.
Therefore, z $\in$ Co(T,g).

Furthermore, assume that for all x, y $\in$ Co(T,g), there exists w $\in$ X such that gx $\preceq$ gw and gy $\preceq$ gw.
Let $\{w_n\}$ be a Picard sequence of (T,g) based on w. As gx $\preceq$ gw and gy $\preceq$ gw and T is (g,$\preceq$)-non-decreasing, we have,

$$gx = Tx \preceq Tw = gw_1 \text{ and } gy = Ty \preceq Tw = gw_1.$$

Similarly, by induction, we observe that,

$$gx \preceq gw_n \text{ and } gy \preceq gw_n, \text{ for all } n \in \mathbb{N}.$$

For all k $\in \mathbb{N}$,

$$\begin{aligned}\psi(G(gx,[gw_{n+1}]^{n-1})) &= \psi(G(Tx,[Tw_n]^{n-1})) \\ &\leq (\psi-\phi)(G(gx,[gw_n]^{n-1}))\end{aligned} \tag{2.3}$$

By Lemma 1.2, $G(gx,[gw_n]^{n-1}) \to 0$.
Therefore, $gw_n \to$ gx. Similarly, $gw_n \to$ gy.
Uniqueness of limit implies gx = gy, for all x, y $\in$ Co(T,g).

Next, we show that for any coincidence point x of T and g, the point a = Tx is a common fixed point of T and g.
Let x $\in$ Co(T,g), arbitrarily chosen. Let a = Tx = gx. Ta = Tgx = gTx = ga as T and g commute. Therefore, a $\in$ Co(T,g). But since a, x $\in$ Co(T,g), as shown before, gx = ga.
And, Ta = ga = gx = Tx = a. Therefore, a is a common fixed point of T and g.

Finally we show that T and g have a unique common fixed point.
Let s, r be two common fixed points of T and g. Therefore, s = Ts = gs and r = Tr = gr. So s, r $\in$ Co(T,g) which implies gs = gr.
Therefore s = r and hence we obtain the uniqueness. □

The following corollary is a restriction of only two variables in the contractive condition of Theorem 2.1.

Corollary 2.2. *Let (X,G,$\preceq$) be a preordered G_n-metric space and let $T : X \to X$ be a self mapping. Suppose that the following conditions hold:*

(1.) *(X,G) is complete;*
(2.) *there exists $x_0 \in X$ such that $gx_0 \preceq Tx_0$;*
(3.) *T(X) $\subseteq$ g(X);*
(4.) *T is (g,$\preceq$)-non-decreasing;*
(5.) *g is continuous and commutes with T;*
(6.) *T is G-continuous;*
(7.) *there exists $\psi \in \Psi$ and $\phi \in \Phi$ such that for all x_1, x_2,..., $x_n \in X$,*
$\psi(G(Tx_1, [Tx_2]^{n-1})) \leq \psi(G(gx_1, [gx_2]^{n-1}))$ - $\phi(G(gx_1, [gx_2]^{n-1}))$, *for $gx_1 \preceq gx_2$.*

Then T and g have, at least, a coincidence point, that is, there exists z $\in X$ such that $Tz = gz$.
Furthermore, assume that for all x, y $\in$ Co(T,g), there exists $w \in X$ such that $gx \preceq gw$ and $gy \preceq gw$. Then:

- *$gx = gy$ for all x, y $\in$ Co(T,g), and*
- *T and g have a unique common fixed point a, which is $a = Tx$, where $x \in$ Co(T,g).*

Example 2.1. Let X = $\mathbb{R}$ and define G_n- metric on X as G : $X^n \to [0,\infty)$ where $G(x_1, x_2, ..., x_n) = \max\{|x_1 - x_2|, |x_2 - x_3|, ..., |x_n - x_{n-1}|, |x_1 - x_n|\}$, for all x_1, x_2, ... , $x_n \in$ X. Let $'\preceq'$ be a preorder defined by
x $\preceq$ y $\Leftrightarrow$ (x = y or x < y $\leq$ 0).
Define T : $\mathbb{R} \to \mathbb{R}$ as

$$Tx = \begin{cases} \frac{x}{3}, & if\, x \leq 0; \\ 2x, & if\, x > 0. \end{cases}$$

Let g be the identity mapping on X. Hence, T and g are continuous. If $gx \preceq gy$ then $x \preceq y$ which implies $Tx \preceq Ty$. Therefore, T is (g,$\preceq$) non-decreasing. We now show that contractive condition (7.) stated in Theorem 2.1 holds with $\psi(t)$ = t and $\phi(t) = \frac{t}{2}$. If $x_1 = x_2 = ... = x_n$, the contractive condition holds trivially.
Without loss of generality, assume $x_1 \neq x_2 \neq ... \neq x_n$ then $x_1 < x_2 < ... < x_n \leq 0$.

$$\psi(G(Tx_1, Tx_2, ..., Tx_n)) = G(\frac{x_1}{3}, \frac{x_2}{3}, ..., \frac{x_n}{3})$$
$$= max\{|\frac{x_1}{3} - \frac{x_2}{3}|, |\frac{x_2}{3} - \frac{x_3}{3}|, ..., |\frac{x_n}{3} - \frac{x_{n-1}}{3}|, |\frac{x_n}{3} - \frac{x_1}{3}|\}$$

$$= \frac{1}{3}G(x_1, x_2, ..., x_n)$$
$$\leq (\psi - \phi)G(gx_1, gx_2, ..., gx_n)$$

Hence, it satisfies the hypothesis of Theorem 2.1 and thus T and g have unique common fixed point, which in this case is x = 0.

Remark.: If the contractive condition holds for all x_1, x_2 ,..., $x_n \in$ X then continuity of g implies continuity of T but it is not so in the setting of preordered G_n- metric space X and that is why we include condition (6.) stated in Theorem 2.1.

Example 2.2. Under the same setting of preordered G_n-metric space as in Example 2.1, define T : $\mathbb{R} \to \mathbb{R}$ as

$$Tx = \begin{cases} \frac{x}{3}, & if\, x \leq 0; \\ 2x+1, & if\, x > 0. \end{cases}$$

and g as the identity mapping on $\mathbb{R}$. Then g is continuous.
But $\lim_{n\to\infty} \frac{1}{n} = 0$ and $\lim_{n\to\infty} T(\frac{1}{n}) = 1 \neq$ T(0).

Lemma 2.3. *Let T, g : $X \to X$ be self maps on (X,G) such that (2.1) holds. Then T is G-continuous at every point in which g is G-continuous.*

Proof. Assume g is G-continuous at $\omega \in X$. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in X such that $x_n \to \omega$.
Therefore $gx_n \to g\omega$, that is,

$$\lim_{n\to\infty} G(gx_n, [g\omega]^{n-1}) = 0.$$

By (2.1),

$$\begin{aligned}\psi(G(Tx_n, [T\omega]^{n-1})) &\leq \psi(G(gx_n, [g\omega]^{n-1})) - \phi(G(gx_n, [g\omega]^{n-1})) \\ &\leq \psi(G(gx_n, [g\omega]^{n-1})).\end{aligned} \tag{2.4}$$

Taking limit as n $\to \infty$ in (2.4) and since $\psi \in \Psi$, we have,

$$\lim_{n\to\infty} G(Tx_n, [T\omega]^{n-1}) = 0.$$

Therefore $Tx_n \to T\omega$ and hence T is continuous at ω. □

As a consequence of Theorem 2.1, the following theorem is a natural extension of Theorem 4.3.2 in [1] to the case of G_n-metric space.

Theorem 2.4. *Let (X,G) be a G_n-metric space and let T, g : $X \to X$. Assume the following conditions:*

(a) *(X,G) is complete.*
(b) *T(X)* $\subseteq$ *g(X).*
(c) *There exists* $\psi \in \Psi$ *and* $\phi \in \Phi$ *such that, for all* $x_1, x_2, .., x_n \in X$,
$\psi(G(Tx_1, Tx_2, .., Tx_n)) \leq \psi(G(gx_1, gx_2, .., gx_n)) -$
$\phi(G(gx_1, gx_2, .., gx_n))$.
(d) *g is continuous and commutes with T.*

Then T and g have unique common fixed point w, that is, w = Tw = gw. In fact, for any x $\in$ *Co(T,g), Tx = w. In particular, gx = gy, for all x, y* $\in$ *Co(T,g).*

The following contractive condition, which is weaker than the one in Theorem 2.4, leads to similar conclusion.

Corollary 2.5. *Let (X,G) be a* G_n*-metric space and let T, g : X* $\to$*X. Assume the following conditions:*

(a) *(X,G) is complete.*
(b) *T(X)* $\subseteq$ *g(X).*
(c) *There exists* $\psi \in \Psi$ *and* $\phi \in \Phi$ *such that, for all* $x_1, x_2 \in X$,
$\psi(G(Tx_1, [Tx_2]^{n-1})) \leq \psi(G(gx_1, [gx_2]^{n-1})) - \phi(G(gx_1, [gx_2]^{n-1}))$.
(d) *g is continuous and commutes with T.*

Then T and g have unique common fixed point w, that is, w = Tw = gw. Infact, for any x $\in$ *Co(T,g), Tx = w. In particular, gx = gy, for all x, y* $\in$ *Co(T,g).*

If we take ψ to be the identity mapping, Theorem 2.4 reduces to the following:

Corollary 2.6. *Let (X,G) be a* G_n*-metric space and let T, g : X* $\to$*X. Assume the following conditions:*

(a) *(X,G) is complete.*
(b) *T(X)* $\subseteq$ *g(X).*
(c) *There exists* $\psi \in \Psi$ *and* $\phi \in \Phi$ *such that, for all* $x_1, x_2, ..., x_n \in X$,
$G(Tx_1, Tx_2, ..., Tx_n) \leq G(gx_1, gx_2, ..., gx_n) - \phi(G(gx_1, gx_2, ..., gx_n))$.
(d) *g is continuous and commutes with T.*

Then T and g have unique common fixed point w, that is, w = Tw = gw. Infact, for any x $\in$ *Co(T,g), Tx = w. In particular, gx = gy, for all x, y* $\in$ *Co(T,g).*

In the previous result, if we take ϕ(t)=(1-k)t for all t≥ 0, k$\in [0,1)$, the particular case can be stated as follows:

Corollary 2.7. *Let (X,G) be a G_n-metric space and let T, g : $X \to X$. Assume the following conditions:*

(a) *(X,G) is complete.*
(b) *T(X) $\subseteq$ g(X).*
(c) *There exists $\psi \in \Psi$ and $\phi \in \Phi$ such that, for all $x_1, x_2, ..., x_n \in X$,* $G(Tx_1, Tx_2, ..., Tx_n) \leq kG(gx_1, gx_2, ..., gx_n)$.
(d) *g is continuous and commutes with T.*

Then T and g have unique common fixed point w, that is, $w = Tw = gw$. Infact, for any $x \in$ Co(T,g), $Tx = w$. In particular, $gx = gy$, for all x, $y \in$ Co(T,g).

Note that, the previous result is a natural extension of Theorem 4.3.1 in [1] of unique existence of common fixed points in context of G-metric space to the case of G_n-metric space.
Similarly, we can restrict the previous two Corollaries to the case of only two distinct variables.

References

(1) R.P. Agarwal, Karapınar, O'Regan, *Fixed point Theory in Metric Type Spaces*, Springer International Publishing Switzerland, (2015).
(2) Mustafa, Z, Sims, B.: Some remarks concerning D-metric spaces, *In: Proceedings of the International Conferences on Fixed Point Theory and Applications*, Valencia(Spain), July, (2003), 189–198.
(3) Mustafa, Z., Sims, B.:A new approach to generalized metric spaces, *J. Nonlinear Convex Anal.*, **7(2)** (2006), 289–297.
(4) Mustafa, Z.: *A new structure for generalized metric spaces with applications to fixed point theory*, PhD thesis, The University of Newcastle, Australia, (2005).
(5) Roldán, Karapınar, Kumam: G-metric spaces in any number of arguments and related fixed point theorems, *Fixed point Theory and applications* (2014), 1–10.

Bhumi Amin
Department of Mathematics, Faculty of Science
The Maharaja Sayajirao University of Baroda
Vadodara, Gujarat 390002, India.
E-mail: *bhumi95amin@gmail.com*

Rajendra G. Vyas
Department of Mathematics, Faculty of Science
The Maharaja Sayajirao University of Baroda
Vadodara, Gujarat 390002, India.
E-mail: *vyas.rajendra@gmail.com*

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 153–167

DETERMINATION AND GRAPHICAL ANALYSIS OF MTFS AND AVAILABILITY OF A TWO UNIT COLD STANDBY STOCHASTIC SYSTEM

MANJU DEVI, V. K. GUPTA AND VINOD KUMAR
(Received : 10 - 11 - 2019 ; Revised : 24 - 01 - 2020)

Abstract. Two and three unit cold standby systems have widely been studied by a large number of researchers in the field of reliability, under various assumptions and concepts by using the Regenerative Point Technique and other methodologies, for doing the reliability and availability analysis of the stochastic systems. In this paper, the path analysis of a two unit cold standby stochastic system with a perfect switching over device and a single server/repairmen, has been done to determine the MTSF and availability of the system using RPGT and to do the graphical analysis of the same. Initially, one unit is in operative mode and a similar unit in the cold standby mode which is automatically switched on after the failure of the operating unit and it needs activation time before it becomes operative. Immediately after the standby unit is activated, the inspection of the failed unit starts (if the repairman is available) to ascertain whether the failed unit is repairable or needs replacement. The distributions of the failure times are exponential and that of the inspection times, waiting times, repair times and replacement times may be the general distributions whereas the graphical study has been made for MTSF and the availability of the system by considering these as exponential distributions.

1. Introduction

The researchers including Taneja, Anita et al. [1], Shakeel and Vinod [6], Kumar, Ashok et al. [3] and many others have used the Regenerative Point Technique and other methodologies, for doing the reliability and availability analysis of various stochastic systems. They discussed two and three unit

2010 Mathematics Subject Classification: 93E10
Key words and phrases: Reachable state, Regenerative Point, Graphical Technique, Primary Circuit, Path Analysis

cold standby stochastic systems in the field of reliability, under various assumptions and concepts. With the increase in the number of states to which a system can transit and increase in the number of transitions from any state to the others states, the Regenerative Point Technique and other methodologies, become very time consuming and cumbersome for doing the analysis for finding the key the parameters of a stochastic system. To overcome this difficulty, Gupta [4] introduced a new approach known as Regenerative Point Graphical Technique (RPGT), for finding the MTSF and Availability and other key parameters Gupta et al. [5, 7] (under steady state conditions). All the simple paths from the initial state to the other states and the various circuits with their levels (primary, secondary or of higher order), along the different paths, are to be located in totality, for doing the analysis using RPGT of a semi Markov process. In this paper, path analysis of a two unit cold standby stochastic system with an automatic perfect switching over device, analyzed by Taneja and Tuteja [1], has been done to find the MTSF and availability by using RPGT and graphical analysis has been done by using exponential distributions for the inspection times, waiting times, repair times and replacement times.

2. The System

The system consists of an operating unit and a cold standby unit along with an automatic perfect switching over device and a single server/repairman facility available. On the failure of the operating unit, the cold standby unit is automatically switched on instantaneously with probability 1 and it needs some time (activation-time) to be in the operative mode. Until the activation time of the standby unit is over, the failed unit is kept waiting for inspection and the system is treated in the down state. After the standby unit is activated, it becomes operational and the inspection of the failed unit is carried starts immediately, to ascertain whether the failed unit is repairable or it needs replacement. The system is in up-state/available only if a unit is in operative mode otherwise it is in down/or failed state.

3. Assumptions, Notations & Symbols

The various assumptions, notations and the symbols used are given as under:

1. The system starts from the good state '0' at time $t = 0$, with probability 1.
2. The switching over device is 100% perfect.
3. The standby and the operating units are identical and have same failure rate.
4. On the failure of a unit, the cold standby unit is automatically switched on with probability 1 and the cold standby unit needs some activation time.
5. The inspection is carried out immediately after the standby unit is activated.
6. Repairs/replacement start immediately after the inspection finishes. During the inspection/repairs/replacement, the other operational unit may fail and the system transits into a failed state. After repairs, the unit works as a new one.
7. There is a single server/repairman available on demand for repairs and replacements etc. and the server do not leave the system while doing inspection, repairs/replacement.
8. All the random variables are independent and are un-correlated.
9. The distributions of the failure times are exponential and that of the inspection times, waiting times, repair times and replacement times are general distributions.

$pr/*/©$	:	Probability/Laplace transformation/Laplace Convolution.
$\underline{\text{K}}$	:	Non regenerative state 'k'
$\{a_0, a_1, \ldots, a_{n-1}, a_n\}$	:	A directed path from the state a_0 to the reachable state a_n, through the states $a_1, a_2, \ldots, a_{n-1}$.
$p.d.f./f_i$	:	Probability density function/Fuzziness measure of the i-state; $f_i = 0$, if 'i' is a failed state and $f_i = 1$, if 'i' is an up state.
$(i, j)/p(i, j)$	:	Steady state transition probability from the regenerative state i to the regenerative state j, without visiting any other states.

	$(i,j) = p(i,j) = \lim_{s \to 0} q^*_{i,j}(s)$;
	$(i,j,k) = (i,j)(j,k) = p(i,j).p(j,k)$
$(i,\underline{\text{k}},j)/p(i,\underline{\text{k}},j)$	: Steady state transition probability from the regenerative state i to the regenerative state j, visiting non regenerative state $\underline{\text{k}}$. $(i,\underline{\text{k}},j) = p(i,\underline{k},j) = \lim_{s \to 0} q^*_{i,\underline{k},j}(s)$
$\overline{cycle}$	: A circuit formed through un-failed states.
$k\text{-}\overline{cycle}$	: A circuit formed through un-failed states, with terminals at the regenerative state k.
$k\text{-}cycle$	: A circuit with terminals at the regenerative state k.
$V\overline{(k,k)}$	: Transition probability factor of the reachable state k of the $k\text{-}\overline{cycle}$ formed through un-failed states.
$V(k,k)$	: Transition probability factor of the reachable state k of the $k\text{-}cycle$.
$V(i,j)$	: Transition probability factor of reachable state 'j' from state 'i'.
$(i \xrightarrow{s_r} j)$	: r-th directed simple path from state 'i' to state 'j'; r takes positive integral values for different paths from state 'i' to state 'j'.
$(i \xrightarrow{sff} j)$	: Directed simple failure free path from state 'i' to state 'j'.
$q_{i,j}(t)/q_{i,\underline{k},j}(t)$	: p.d.f. of the first passage time from a regenerative state i to a regenerative state j or to a failed state j without visiting any other regenerative state in $(0,t]$/while visiting $\underline{\text{k}}$ only once in $(0,t]$, given that the system entered regenerative state i at $t = 0$.
$R_i(t)$	: Reliability of the system at time t, given the system is initially in the regenerative state 'i'.
μ_i/μ'_i	: Mean sojourn time of the state 'i'/total un-conditional time spent before transiting to any other regenerative state(s), given that the system entered regenerative state 'i' at $t = 0$.
λ	: Constant failure rate of operating unit.
p_1/p_2	: Probability that unit is repairable/needs replacement. $p_1 + p_2 = 1$.
$w(t)/W(t)/\overline{W}(t)$	: p.d.f./c.d.f. of the activation time of the standby unit.

		$\overline{W}(t) = 1 - W(t)$.
$h_1(t)/H_1(t)/\overline{H_1}(t)$	:	p.d.f./c.d.f. of the inspection time. $\overline{H_1}(t) = 1 - H_1(t)$.
$g(t)/G(t)/\overline{G}(t)$	:	p.d.f./c.d.f. of the repair time. $\overline{G}(t) = 1 - G(t)$.
$h_2(t)/H_2(t)/\overline{H_2}(t)$	:	p.d.f./c.d.f. of the replacement time. $\overline{H_2(t)} = 1 - H_2(t)$
$C_O/C_S/C_{AO}$	:	Unit is in operating/cold standby/activation mode.
Fwi/Fui	:	Failed unit is waiting for inspection/under inspection.
$Fur/Frep$	:	Failed unit is under repairs/replacement.
$F_{UI}/F_{UR}/F_{REP}$	:	Failed unit under inspection/repairs/Replacement continuing from the previous state.

4. State Transition Diagram

Under the given assumptions and conditions, the State Space of the stochastic system is $S = \{S0, S1, S2, S3, S4, S5, S6, S7, S8, S9\}$ and there are a total of 15 transitions that can take place amongst these 10 transition states. All the states are reachable and aperiodic states and $S0$, $S1$, $S2$, $S3$, $S4$, $S8$, and $S9$ are the regenerative states and $S5$, $S6$, $S7$ are the failed and non-regenerative states whereas $S8$, $S9$ are the failed regenerative states. The state '$S1$' is a down and un-available state as shown in Table 1 and the transition diagram is shown in Figure 1.

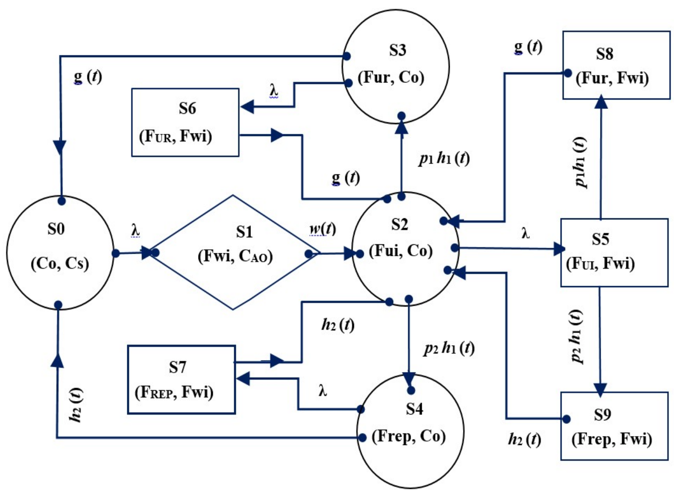


FIGURE 1. Transition diagram

TABLE 1.

Type of State	Symbol	States
Regenerative State	●	$S0$, $S1$, $S2$, $S3$, $S4$, $S8$, $S9$
Down State	◇	$S1$
Up-State	○	$S0$, $S2$, $S3$, $S4$
Failed State	□	$S5$, $S6$, $S7$, $S8$ and $S9$

5. Transition Probabilities and Mean Sojourn Times

5.1. **Transition Probabilities.** The following transition probabilities p_{ij} are determined:
$p(0,1) = 1$; $p(1,2) = 1$; $p(2,3) = p_1 h_1^*(\lambda)$; $p(2,4) = p_2 h_1^*(\lambda)$; $p(2,5) = 1 - h_1^*(\lambda)$; $p(2,\underline{5},8) = p_1\{1 - h_1^*(\lambda)\}$; $p(2,\underline{5},9) = p_2\{1 - h_2^*(\lambda)\}$; $p(3,0) = g^*(\lambda)$; $p(3,6) = 1 - g^*(\lambda)$; $p(3,\underline{6},2) = 1 - g^*(\lambda)$; $p(4,0) = h_2^*(\lambda)$; $p(4,7) = 1 - h_2^*(\lambda)$; $p(4,\underline{7},2) = 1 - h_2^*(\lambda)$; $p(8,2) = 1$; $p(9,2) = 1$.

5.2. **Mean Sojourn Times.** The mean sojourn times (μ_i) for the different regenerative states are:
$\mu_0 = \frac{1}{\lambda}$; $\mu_1 = \int_0^\infty t.d\{W(t)\} = -w^{*'}(0)$; $\mu_2 = \{1 - h_1^*(\lambda)\}/\lambda$; $\mu_3 = \{1 - g^*(\lambda)\}/\lambda$; $\mu_4 = \{1 - h_2^*(\lambda)\}/\lambda$; $\mu_8 = \int_0^\infty t.d\{G(t)\} = -g^{*'}(0)$; $\mu_9 = \int_0^\infty t.d\{H_2(t)\} = -h_2^{*'}(0)$.

5.3. **Total Un-Conditional Times.** The total unconditional mean times (μ_i') spent before transiting to any other regenerative state are:
$\mu_0' = \mu_0$; $\mu_1' = \mu_1$; $\mu_2' = -h_1^{*'}(0)$; $\mu_3' = -g^{*'}(0)$; $\mu_4' = -h_2^{*'}(0)$; $\mu_8' = -g^{*'}(0) = \mu_8$; $\mu_9' = -h_2^{*'}(0) = \mu_9$.

6. Path Analysis of the System

All the directed simple paths from each state to the other reachable states and the various circuits at all the vertices are shown in Tables 2-3. On taking $S_{22} = \{2,3,0,1,2\}$, $\{2,3,6,2\}$, $\{2,4,0,1,2\}$, $\{2,4,7,2\}$, $\{2,5,8,2\}$, $\{2,5,9,2\}$ are the simple circuits at the vertex '2'.

TABLE 2. Paths from a State 'i' to the Reachable State 'j': P0

i	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$
0	$\{0,1,2,3,0\}$, $\{0,1,2,4,0\}$	$\{0,1\}$	$\{0,1,2\}$	$\{0,1,2,3\}$	$\{0,1,2,4\}$
1	$\{1,2,3,0\}$, $\{1,2,4,0\}$	$\{1,2,3,0,1\}$, $\{1,2,4,0,1\}$	$\{1,2\}$	$\{1,2,3\}$	$\{1,2,4\}$
2	$\{2,3,0\}$, $\{2,4,0\}$	$\{2,3,0,1\}$, $\{2,4,0,1\}$	S_{22}	$\{2,3\}$	$\{2,4\}$
3	$\{3,0\}$, $\{3,6,2,4,0\}$	$\{3,0,1\}$, $\{3,6,2,4,0,1\}$	$\{3,0,1,2\}$ $\{3,6,2\}$	$\{3,0,1,2,3\}$ $\{3,6,2,3\}$	$\{3,0,1,2,4\}$ $\{3,6,2,4\}$
4	$\{4,0\}$, $\{4,7,2,3,0\}$	$\{4,0,1\}$, $\{4,7,2,3,0,1\}$	$\{4,0,1,2\}$ $\{4,7,2\}$	$\{4,0,1,2,3\}$ $\{4,7,2,3\}$	$\{4,0,1,2,4\}$ $\{4,7,2,4\}$
5	$\{5,8,2,3,0\}$, $\{5,8,2,4,0\}$ $\{5,9,2,3,0\}$, $\{5,9,2,4,0\}$	$\{5,8,2,3,0,1\}$, $\{5,8,2,4,0,1\}$ $\{5,9,2,3,0,1\}$, $\{5,9,2,4,0,1\}$	$\{5,8,2\}$ $\{5,9,2\}$	$\{5,8,2,3\}$ $\{5,9,2,3\}$	$\{5,8,2,4\}$ $\{5,9,2,4\}$
6	$\{6,2,3,0\}$, $\{6,2,4,0\}$	$\{6,2,3,0,1\}$, $\{6,2,4,0,1\}$	$\{6,2\}$	$\{6,2,3\}$	$\{6,2,4\}$
7	$\{7,2,3,0\}$, $\{7,2,4,0\}$	$\{7,2,3,0,1\}$, $\{7,2,4,0,1\}$	$\{7,2\}$	$\{7,2,3\}$	$\{7,2,4\}$
8	$\{8,2,3,0\}$, $\{8,2,4,0\}$	$\{8,2,3,0,1\}$, $\{8,2,4,0,1\}$	$\{8,2\}$	$\{8,2,3\}$	$\{8,2,4\}$
9	$\{9,2,3,0\}$, $\{9,2,4,0\}$	$\{9,2,3,0,1\}$, $\{9,2,4,0,1\}$	$\{9,2\}$	$\{9,2,3\}$	$\{9,2,4\}$

TABLE 3. Paths from State 'i' to the Reachable State 'j': P0

i	$j=5$	$j=6$	$j=7$	$j=8$	$j=9$
0	$\{0,1,2,5\}$	$\{0,1,2,3,6\}$	$\{0,1,2,4,7\}$	$\{0,1,2,5,8\}$	$\{0,1,2,5,9\}$
1	$\{1,2,5\}$	$\{1,2,3,6\}$	$\{1,2,4,7\}$	$\{1,2,5,8\}$	$\{1,2,5,9\}$
2	$\{2,5\}$	$\{2,3,6\}$	$\{2,4,7\}$	$\{2,5,8\}$	$\{2,5,9\}$
3	$\{3,0,1,2,5\}$ $\{3,6,2,5\}$	$\{3,6\}$	$\{3,0,1,2,4,7\}$ $\{3,6,2,4,7\}$	$\{3,0,1,2,5,8\}$ $\{3,6,2,5,8\}$	$\{3,0,1,2,5,9\}$, $\{3,6,2,5,9\}$
4	$\{4,0,1,2,5\}$ $\{4,7,2,5\}$	$\{4,0,1,2,3,6\}$ $\{4,7,2,3,6\}$	$\{4,7\}$	$\{4,0,1,2,5,8\}$ $\{4,7,2,5,8\}$	$\{4,0,1,2,5,9\}$, $\{4,7,2,5,9\}$
5	$\{5,8,2,5\}$ $\{5,9,2,5\}$	$\{5,8,2,3,6\}$ $\{5,9,2,3,6\}$	$\{5,8,2,4,7\}$ $\{5,9,2,4,7\}$	$\{5,8\}$	$\{5,9\}$
6	$\{6,2,5\}$	$\{6,2,3,6\}$	$\{6,2,4,7\}$	$\{6,2,5,8\}$	$\{6,2,5,9\}$
7	$\{7,2,5\}$	$\{7,2,3,6\}$	$\{7,2,4,7\}$	$\{7,2,5,8\}$	$\{7,2,5,9\}$
8	$\{8,2,5\}$	$\{8,2,3,6\}$	$\{8,2,4,7\}$	$\{8,2,5,8\}$	$\{8,2,5,9\}$
9	$\{9,2,5\}$	$\{9,2,3,6\}$	$\{9,2,4,7\}$	$\{9,2,5,8\}$	$\{9,2,5,9\}$

7. EVALUATION OF MTSF & AVAILABILITY OF THE SYSTEM

The mean time to the system failure (MTSF) and availability of the system (under steady state conditions) are evaluated, by applying RPGT as under:

7.1. Mean Time to System Failure (MTSF). From Figure 1, the regenerative un-failed/down states to which the system can transit (with initial state '0' at time't' $= 0$), before transiting to any failed states are: $i = 0, 1, 2, 3$ and 4 where state 1 is a down state. The mean time to system failure is obtained by using RPGT as under:

$$MTSF = \left[\sum_{i,s_r}\left\{\frac{\{pr(0 \xrightarrow{s_r(sff)} i)\}\cdot\mu_i}{\prod_{k_1\neq 0}\{1-V_{\overline{k_1,k_1}}\}}\right\}\right] \div \left[1-\sum_{s_r}\left\{\frac{\{pr(0 \xrightarrow{s_r(sff)} 0)\}}{\prod_{k_2\neq 0}\{1-V_{\overline{k_2,k_2}}\}}\right\}\right]$$

The failure free paths from initial state '0' to the un-failed/down states 'i' and circuits through un-failed states are shown in Table 4. Here k_1 = Nil and k_2 = Nil.

TABLE 4.

Vertex 'i'	Failure Free Paths from initial state '0' to un-failed/down state 'i'		Failure Free Circuits
	$(0 \overset{s_{r(sff)}}{\longrightarrow} j)$	P_0	P_1
0	$(0 \overset{s_1(sff)}{\longrightarrow} 0)$	$\{0, 1, 2, 3, 0\}$	–
	$(0 \overset{s_2(sff)}{\longrightarrow} 0)$	$\{0, 1, 2, 4, 0\}$	–
1	$(0 \overset{s_1(sff)}{\longrightarrow} 1)$	$\{0, 1\}$	–
2	$(0 \overset{s_1(sff)}{\longrightarrow} 2)$	$\{0, 1, 2\}$	–
3	$(0 \overset{s_1(sff)}{\longrightarrow} 3)$	$\{0, 1, 2, 3\}$	–
4	$(0 \overset{s_1(sff)}{\longrightarrow} 4)$	$\{0, 1, 2, 4\}$	–

$$\begin{aligned}
MTSF &= [(0,0)\mu_0 + (0,1)\mu_1 + (0,1,2)\mu_2 + (0,1,2,3)\mu_3 + (0,1,2,4)\mu_4] \\
&\quad \div [1 - (0,1,2,3,0) - (0,1,2,4,0)] \\
&= [\mu_0 + p(0,1)\mu_1 + p(0,1)\cdot p(1,2)\mu_2 + p(0,1)\cdot p(1,2)\cdot p(2,3)\mu_3 \\
&\quad + p(0,1)\cdot p(1,2)\cdot p(2,4)\mu_4] \div [1 - p(0,1)\cdot p(1,2)\cdot p(2,3)\cdot p(3,0) \\
&\quad - p(0,1)\cdot p(1,2)\cdot p(2,4)\cdot p(4,0)] \\
&= N_0/D_0; \quad p(0,0) = 1
\end{aligned}$$

where $N_0 = \mu_0 + \mu_1 + \mu_2 + p(2,3)\mu_3 + p(2,4)\mu_4$ and $D_0 = 1 - p(2,3).p(3,0) - p(2,4).p(4,0)$.

7.2. **Availability of the System.** From Figure 1, the regenerative available un-failed/up-states to which the system can transit, are: $j = 0, 2, 3$ and 4. All the regenerative states to which the system can transit, are: $i = 0, 1, 2, 3, 8$ and 9. The steady state availability of the system (with initial state '0' at time $t = 0$), using RPGT, is given by:

$$A_0 = \left[\sum_{j,s_r}\left\{\frac{\{pr(0 \xrightarrow{s_r} j)\}f_j \cdot \mu_j}{\prod_{k_1\neq 0}\{1 - V_{k_1,k_1}\}}\right\}\right] \div \left[\sum_{i,s_r}\left\{\frac{\{pr(0 \xrightarrow{s_r} i)\} \cdot \mu_i'}{\prod_{k_2\neq 0}\{1 - V_{k_2,k_2}\}}\right\}\right]$$

Here $k_1 = 2$ and $k_2 = 2$ and on taking $f_j = 1$, for all j,

$$\begin{aligned} A_0 &= [(0,0)\mu_0 + \{(0,1,2)/(1-L_1)\}\mu_2 + (0,1,2,3)/(1-L_1)\mu_3 \\ &\quad + \{(0,1,2,4)/(1-L_1)\}\mu_4] \div [(0,0)\mu_0' + (0,1)\mu_1' + \{(0,1,2)/(1-L_1)\}\mu_2' \\ &\quad + \{(0,1,2,3)/(1-L_1)\}\mu_3' + \{(0,1,2,4)/(1-L_1)\}\mu_4' \\ &\quad + \{(0,1,2,\underline{5},8)/(1-L_1)\}\mu_8' + \{(0,1,2,\underline{5},9)/(1-L_1)\}\mu_9'] \\ &= N_1 \div D_1 \end{aligned}$$

where

$$\begin{aligned} L_1 &= (2,3,6,2) + (2,4,\underline{7},2) + (2,\underline{5},8,2) + (2,\underline{5},9,2) \\ &= 1 - \{(2,3,0) + (2,4,0)\} \end{aligned}$$

and

$$\begin{aligned} N_1 &= \{p(2,3).p(3,0) + p(2,4).p(4,0)\}\mu_0 + \mu_2 + p(2,3)\mu_3 + p(2,4)\mu_4; \\ D_1 &= \{p(2,3).p(3,0) + p(2,4).p(4,0)\}(\mu_0 + \mu_1) + \mu_2' \\ &\quad + \{p(2,3) + p(2,\underline{5,}8)\}\mu_8 + \{p(2,4) + p(2,\underline{5,}9)\}\mu_9 \end{aligned}$$

8. Observation

On taking $f_j = 1$ for all the up-states namely 'j' $= 0, 2, 3$ and 4, the mean time to system failure and the availability of the two unit cold standby stochastic system, are the same as by Taneja and Tuteja [1], but these are obtained very easily and quickly by doing the path analysis and using RPGT, thus it is more economical.

9. Graphical Analysis For MTSF and Availability of the System

Here the graphical analysis has been done by using the exponential distributions for the inspection times, waiting times, repair times and replacement times. By considering the following particular cases: $h_1(t) = \Theta_1 e^{-\Theta_1^t}$; $h_2(t) = \Theta_2 e^{-\Theta_2^t}$; $w(t) = \alpha e^{-\alpha t}$; $g(t) = \beta e^{-\beta t}$ and on taking $\beta = 0.90$; $\Theta_1 = 0.3661$ and $\Theta_2 = 0.2031$:

9.1. **MTSF vs. Failure Rate of the Main Unit.** On taking $\alpha = 0.96$, and considering the different cases (i) $p_1 > p_2$ (ii) $p_1 = p_2$ (iii) $p_1 < p_2$, the MTSF has been determined for different values the failure rate (λ) of the main/operative unit and the corresponding graphs have been shown in Figure 2. It has been observed that the MTSF (T_0) of the system decreases rapidly with the increase in failure rate (λ) of the operating unit for any fixed value of p_1/p_2. For $p_1 = 0.75$, as λ varies from 0.001 to 0.010, it has been observed that % decrease in MTSF varies from 75% to 18% approximately and in all the three cases this % decrease in MTSF is almost the same, whereas T_0 decreases as p_1 decreases (or p_2 increases) for a given value λ. As p_1 decreases from 0.75 to 0.25, the value of T_0 decreases by 28.22% for $\lambda = 0.001$ and it decreases by 26.34% for $\lambda = 0.010$ which shows that for higher value of λ, the variations in MTSF (T_0) is smaller for given variations in p_1.

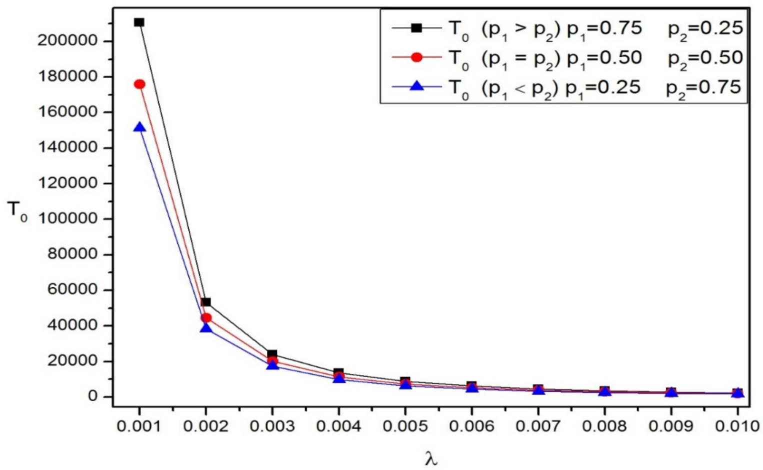


FIGURE 2.

9.2. **MTSF vs. Activation Time Rate of the Cold Standby Unit.** On taking $\lambda = 0.005$, and considering the different cases (i) $p_1 > p_2$ (ii) $p_1 = p_2$ (iii) $p_1 < p_2$, the MTSF has been determined for different values the activation rate (α) of the cold standby unit and the corresponding graphs have been shown in Figure 3. It has been observed that the MTSF (T_0) of the system decreases very slowly with the increase in the activation time rate (α) of the standby unit for any fixed value of p_1/p_2. Further, as α varies from 0.25 to 3.00, it has been observed that % decrease in T_0 varies from 0.958% to 0.034% for $p_1 = 0.75$; 0.954% to 0.034% for $p_1 = 0.50$ and from 0.950% to 0.034% for $p_1 = 0.25$ and the % decrease in T_0 is almost the same in all the three cases and when p_1 decreases from 0.75 to 0.25, the value of T_0 decreases by 27.35% (for $\alpha = 0.25$) and 27.34% (for $\alpha = 3.0$) which shows that the % variation in T_0 is negligible for higher values of α, for a given variations in p_1.

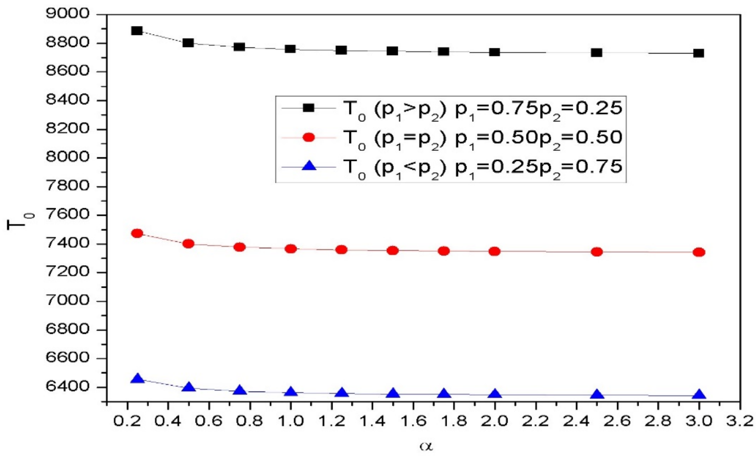


FIGURE 3.

9.3. **Availability vs. Failure Rate of the Main Unit.** On taking $\alpha = 0.96$, considering the different cases (i) $p_1 > p_2$ (ii) $p_1 = p_2$ (iii) $p_1 < p_2$, the availability of the system (A_0) has been determined for different values the failure rate (λ) of the operative unit and the corresponding graphs have been shown in Figure 4. It has been observed that A_0 decreases rapidly with the increase in failure rate (λ) of operating unit for any fixed value of p_1/p_2, whereas A_0 decreases as p_1 decreases (or p_2 increases) for a given value of λ. Further, it has been observed that as p_1 decreases from 0.75 to 0.25, A_0 decreases by 0.002% for $\lambda = 0.001$ and it decreases by 0.188% for $\lambda = 0.010$ which shows that for higher failure rate (λ) the variation in A_0 is more for a given variation in p_1.

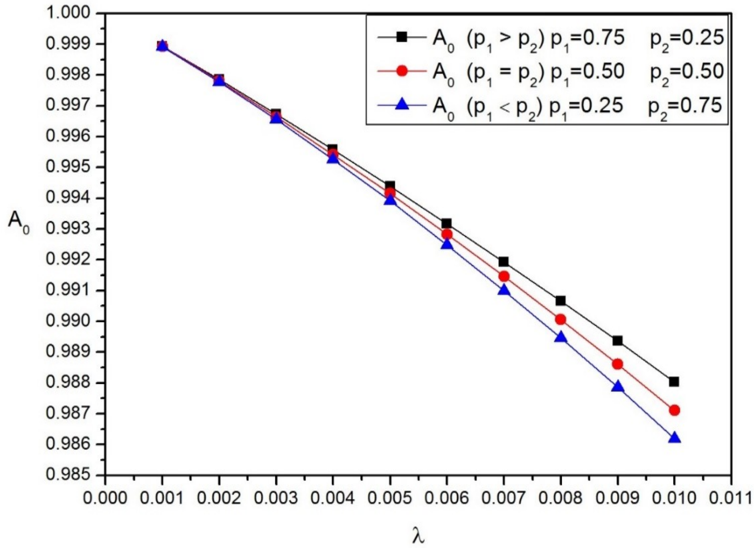


Figure 4.

9.4. **Availability vs. Activation Time Rate of the Cold Standby Unit.** On taking $\lambda = 0.005$, and considering the different cases (i) $p_1 > p_2$ (ii) $p_1 = p_2$ (iii) $p_1 < p_2$, the availability (A_0) of the system has been determined for different values the activation time rate (α) of the cold standby unit and the corresponding graphs have been shown in Figure 5. It has been observed that A_0 increases slowly with the increase in the value of α, for any fixed value of p_1/p_2. Further, as α varies from 0.25 to 3.00, it has been observed that % increase in A_0 varies from 0.967% to 0.0347% for $p_1 = 0.75$; 0.962% to 0.346% for $p_1 = 0.50$ and from 0.957% to 0.344% and this variation in A_0 is almost the same in all the three cases and as p_1 decreases from 0.75 to 0.25, the value of A_0 decreases by 0.035% for $\alpha = 0.25$ and it decreases by 0.52% for $\alpha = 3.00$ which shows that for higher activation time rate (α) the variations in availability (A_0) is more, for a given variation in p_1.

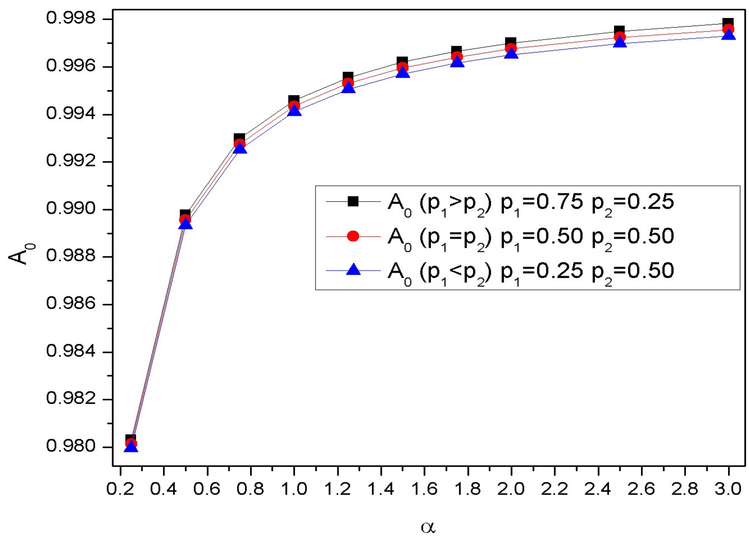


FIGURE 5.

REFERENCES

[1] Taneja, A. and Tuteja, R.K., Reliability and profit evaluation of a two-unit cold standby system with inspection and chances of replacement, *Proc. National Conference on 'Information Technology: Setting Trends in Modern Era' and 8th Annual conference of Indian Society of Information Theory & Applications (RGN Publications)* held at N.C. College of Engineering, Israna (Panipat), India, March 18-20, 2006, pp. 193-197.

[2] Gupta, V.K., Singh, J., Behaviour and profit analysis of a redundant system with imperfect switch, *Journal of Mathematics and Systems Sciences* **3** (2007), no. 1, 16–35.

[3] Kumar, A. and Sanjeev, R.K., Behavior of a complex sub-system with two kinds of failure, *Proc. National Conference on Application of Mathematics in Engineering & Technology*, MIMI, Malout (Punjab), 26-28 March, 2005, 106 – 110.

[4] Gupta, V.K., *Behaviour and Profit Analysis of Some Process Industries*, Ph.D. Thesis, National Institute of Technology, Kurukshetra (India), (2008).

[5] Gupta, V.K., Kumar, K. and Singh, J, Analysis of a PLC's system — A new approach, *Proc. of National Conference held at MMPG College*, Fatehabad, *Journal of Mathematics and Systems Sciences* **5** (2009), no. 2, 39–57.

[6] Ahmad, S. and Kumar, V., Availability analysis of a two unit centrifuge system considering the halt state on occurrence of minor/major fault, *International Journal of Engineering Sciences & Research Technology* **4** (2015), no. 5; 200 – 208.

[7] Gupta, V.K., Analysis of a single unit shock model by using regenerative point graphical technique, *International Journal of Computer Science & Management Studies* **11** (2011), no. 1, 27–33.

[8] Kadyan, M.S., Kumar, J. and Malik, S.C., Stochastic analysis of a two-unit parallel system subject to degradation and inspection for feasibility of repair, *JMASS* **6** (2010), no. 1, 5–13.

[9] Satpal and Goel, P., Availability analysis of a dairy plant with priority in repair using RPGT, *International Journal of Informative & Futuristic Research* **2** (2014), no. 2, 298–306.

[10] Devi, M., Gupta, V.K. and Kumar, V., Path analysis for MTSF & availability of a three unit standby stochastic system, *Aryabhata Journal of Mathematics & Informatics* **10** (2018), no. 2, 405–412.

Manju Devi
Govt. College Behrampur, Bapauli (Haryana), India
E-mail : ghalyan.manju@gmail.com
V. K. Gupta
NCCE, Israna (Haryana), India
E-mail : vijaygupta 1953@yahoo.co.in
Vinod Kumar
Baba Mastnath University, Rohtak (Haryana), India
E-mail : kakoriavinod@gmail.com

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 169–174

IRREDUCIBILITY OF $X^N - A$

BISWAJIT KOLEY AND A. SATYANARAYANA REDDY*

(Received : 14 - 12 - 2019 ; Revised : 12 - 04 - 2020)

Abstract. A. Capelli gave a necessary and sufficient condition for the reducibility of $x^n - a$ over $\mathbb{Q}$. In this article, we are providing an alternate elementary proof for the same.

In this article, we present an elementary proof of a theorem about the irreducibility of $x^n - a$ over $\mathbb{Q}$. Vahlen [4] is the first mathematician who characterized the irreducibility conditions of $x^n - a$ over $\mathbb{Q}$. A. Capelli [1] extended this result to all fields of characteristic zero. Later L. Rédei [5] proved this result for all fields of positive characteristic. But this theorem referred to as Capelli's theorem.

Theorem 1 ([1], [4], [5]). *Let $n \geq 2$. A polynomial $x^n - a \in \mathbb{Q}[x]$ is reducible over $\mathbb{Q}$ if and only if either $a = b^t$ for some $t|n, t > 1$, or $4|n$ and $a = -4b^4$, for some $b \in \mathbb{Q}$.*

Since Theorem 1 is true for arbitrary fields, all of the proofs are proved by using field extensions except the proof given by Vahlen [4]. Vahlen assumes that the binomial $x^n - a$ is reducible and proves Theorem 1 by using the properties of n^{th} roots of unity and by comparing the coefficients on both sides of the following equation

$$x^n - a = (x^m + a_{m-1}x^{m-1} + \cdots + a_0)(x^{n-m} + b_{n-m-1}x^{n-m-1} + \cdots + b_0)$$

for some m, $0 < m < n$. Reader can consult ([3], p.425) for a proof using field theory. We give a proof particularly over $\mathbb{Q}$ by using very little machinery.

2010 Mathematics Subject Classification: 11R09, 12D05

Key words and phrases: Irreducible polynomials, cyclotomic polynomials

* The research of this author is supported by Matrics MTR/2019/001206 of SERB, India

Let $f(x) = x^n - a$, $a = \frac{b}{c} \in \mathbb{Q}$ and $(b, c) = 1$. Then $c^n f(x) = (cx)^n - c^{n-1}b \in \mathbb{Z}[x]$. Hence $x^n - a$ is reducible over $\mathbb{Q}$ if and only if $y^n - c^{n-1}b$ is reducible over $\mathbb{Z}$. It is, therefore, sufficient to consider $a \in \mathbb{Z}$ and throughout the article, by reducibility, we will mean reducible over $\mathbb{Z}$.

Theorem 2. *Let $n \geq 2$. A polynomial $x^n - a \in \mathbb{Z}[x]$ is reducible over $\mathbb{Z}$ if and only if either $a = b^t$ for some $t|n, t > 1$, or $4|n$ and $a = -4b^4$, for some $b \in \mathbb{Z}$.*

The polynomial $x^n - 1$ is a product of cyclotomic polynomials and if $n = 2^{n_1}u$ with $2 \nmid u$, then

$$x^n + 1 = \prod_{d|u} \Phi_{2^{n_1+1}d}(x).$$

Therefore, from now onwards we assume that $a > 1$, if not specified, and check the reducibility of the polynomial $x^n \pm a$ for $n \geq 2$. If there exists a prime p such that $p|a$ but $p^2 \nmid a$, then $x^n \pm a$ is irreducible by Eisenstein's criterion. In other words, if $a = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ is the prime factorization of a and $x^n \pm a$ is reducible, then $a_i \geq 2$ for every $i \in \{1, 2, \ldots, k\}$. More generally,

Lemma 3. *Let $n \geq 2$, $a = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ be the prime factorization of a and let $x^n \pm a$ be reducible. Then $\gcd(a_1, a_2, \ldots, a_k) \geq 2$ and $\gcd(\gcd(a_1, a_2, \ldots, a_k), n) > 1$.*

Proof. We prove the result by induction on $k = \omega(a)$, the number of distinct prime divisors of a. The roots of $x^n \pm a$ are of the form $a^{1/n}(\mp\zeta_n)^e$, where $e \in \mathbb{Z}$ and ζ_n is a primitive n^{th} root of unity. Since the proof barely depends upon the sign of roots, we restrict to the case $x^n - a$. Let $f(x)$ be a proper factor of $x^n - a$, where $\deg(f) = s < n$. If $f(0) = \pm d$, then $\pm d = a^{s/n}\zeta_n^w$ for some $w \in \mathbb{Z}$.

Let $k = 1$ and $a = p_1^{a_1}$. From Eisenstein's criterion, $a_1 \geq 2$. If $d = p_1^{\alpha}$, then $d^n = a^s$ gives, $\alpha n = a_1 s$. Since $a_1 \geq 2$ and $s < n$, we deduce that $(a_1, n) > 1$.

Let $k = 2$ and $a = p_1^{a_1} p_2^{a_2}$. From $d^n = a^s$, let $d = p_1^{d_1} p_2^{d_2}$ be the prime factorization of d. Then $nd_1 = a_1 s, nd_2 = a_2 s$ would give $d_1 a_2 = d_2 a_1$. If $(a_1, a_2) = 1$, then $d_1 = a_1 c$ for some $c|d_2$ and $nc = s < n$ is a contradicton. Thus, $m = (a_1, a_2) \geq 2$. Next we need to show that $(n, m) > 1$. Suppose $a_1 = mb_1, a_2 = mb_2$ with $(b_1, b_2) = 1$. From $d_1 a_2 = d_2 a_1$, we deduce that

$d_1 b_2 = d_2 b_1$. Then $d_1 = b_1 r$ for some $r|d_2$. If $(n,m) = 1$, then $nd_1 = a_1 s = mb_1 s$ will give $n|s$, a contradiction. Hence, $(n,m) > 1$.

Suppose the result is true for some $k \geq 2$. Thus, if $a = p_1^{a_1} \cdots p_k^{a_k}$, then $(a_1, \ldots, a_k) = u > 1$ and $(u,n) > 1$. To show that the result is true for $k+1$. Let $a = p_1^{a_1} \cdots p_k^{a_k} p_{k+1}^{a_{k+1}} = a_1^u p_{k+1}^{a_{k+1}}$, where $a_1 = p_1^{w_1} \cdots p_k^{w_k}$ and $(w_1, \ldots, w_k) = 1$. From $d^n = a^s$, we can write d as $d = b_1^v p_{k+1}^{d_{k+1}}$, where $b_1 = p_1^{v_1} \cdots p_k^{v_k}$, $(v_1, \ldots, v_k) = 1$. From the fundamental theorem of arithmetic, $d_{k+1} n = a_{k+1} s$ and $b_1 = a_1$, $vn = us$. That is $d_{k+1} u = a_{k+1} v$. If $(u, a_{k+1}) = 1$, then $d_{k+1} = a_{k+1} h$ for some $h|v$, and $na_{k+1} h = a_{k+1} s$ implies $n|s$. This contradicts the fact that $s < n$. Thus, $(u, a_{k+1}) = m > 1$.

To show that $(n,m) > 1$. Let $u = mu_1, a_{k+1} = ma'_{k+1}$, where $(u_1, a'_{k+1}) = 1$. From $d_{k+1} u = a_{k+1} v$, we get $u_1 d_{k+1} = va'_{k+1}$. Since $(u_1, a'_{k+1}) = 1$, $d_{k+1} = a'_{k+1} t$ for some $t|v$. On the other hand, $nd_1 = a_1 s, nd_{k+1} = a_{k+1} s$ would imply $a_1 d_{k+1} = d_1 a_{k+1}$. If $a_1 = ua'_1 = mu_1 a'_1$, then $a_1 d_{k+1} = d_1 a_{k+1}$ implies $d_1 = u_1 a'_1 t$. Using this in $nd_1 = a_1 s$, we have $nt = ms$. If $(n,m) = 1$, then $n|s$ is a contradiction. Thus, $(n,m) > 1$. By induction principle, the result is true for every $k \geq 1$. □

In other words, if $x^n \pm a$ is reducible, then a has to be of the form b^m, where $(n,m) > 1$ and $m \geq 2$. With a rearrangement in powers, we can say

Corollary 4. *Let $n \geq 2$ and $x^n \pm a$ be reducible over $\mathbb{Z}$. Then $a = b^m$ for some $m \geq 2, m|n$, and b is either a prime number or $b = (p_1^{b_1} p_2^{b_2} \cdots p_k^{b_k})^d$, where $k \geq 2, (b_1, b_2, \ldots, b_k) = 1$ and $(d,n) = 1$.*

Suppose $f(x) = x^{25} \pm 6^8$. Then $b = 6$, $m = 8$, and $(8, 25) = 1$ implies that the polynomial $x^{25} \pm 6^8$ is irreducible by Corollary 4. Let $g(x) = x^{25} \pm (243)^2$. If we consider $b = 243$, then $(2, 25) = 1$ imply that $x^{25} \pm (243)^2$ is irreducible. But $x^5 \pm 9 | g(x)$. The reason is, $b = 243$ is not as in Corollary 4. Since $243 = 3^5$, b will be 3^2 and $m = 5$ so that $m|n$. Because of this reason, we will say

A positive integer '*b has the property $\mathcal{P}$*' if b is in the form as given in Corollary 4.

Lemma 5. *Let $m \geq 2, m|n$, and b has the property $\mathcal{P}$. Then $x^n \pm b^m$ is reducible except possibly for $x^n + b^{2^r}, r \geq 1$.*

Proof. If $m|n$, then

$$x^n - b^m = \left(x^{n/m} - b\right)\left(x^{n(m-1)/m} + x^{n(m-2)/m} b + \cdots + x^{n/m} b^{m-2} + b^{m-1}\right).$$

Let $m = 2^r m_1$, where $2 \nmid m_1$ and $r \geq 0$. Then

$$x^n + b^m = \prod_{d|m_1} b^{2^r \varphi(d)} \Phi_{2^{r+1}d}\left(\frac{x^{n/m}}{b}\right),$$

where $b^{2^r\varphi(d)}\Phi_{2^{r+1}d}\left(\frac{x^{n/m}}{b}\right) \in \mathbb{Z}[x]$ and φ is the Euler totient function. □

Lemma 5 is true even if b does not have the property $\mathcal{P}$. If $m = 2^r \geq 2, m|n$, and b has the property $\mathcal{P}$, then the reducibility condition of $x^n + b^m$ completes the proof of Theorem 2.

Selmer ([2], p.298) made the following observation. Let $g(x) \in \mathbb{Z}[x]$ be an arbitrary irreducible polynomial of degree n. If $g(x^2)$ is reducible, then, using the fact that $\mathbb{Z}[x]$ is a unique factorization domain, we get

$$g(x^2) = (-1)^n g_1(x) g_1(-x),$$

where $g_1(x)$ is an irreducible polynomial in $\mathbb{Z}[x]$. Thus, if $g_1(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, then

$$g(x^2) = (a_n x^n + a_{n-2}x^{n-2} + \cdots + a_0)^2 - (a_{n-1}x^{n-1} + \cdots + a_1 x)^2.$$

Let k be an odd integer. Then $g(k^2) \equiv g(1) \pmod 4$. Since the right hand side of the last equation is the difference between the two squares, $g(k^2) \equiv 0, \pm 1 \pmod 4$. Combining all of these, one can conclude that

Lemma 6. *Let $g(x) \in \mathbb{Z}[x]$ be an irreducible polynomial.*

(a) If $g(k^2) \equiv 2 \pmod 4$ for an odd integer k, then $g(x^2)$ is irreducible over $\mathbb{Z}$.

(b) If $g(x^2)$ is reducible, then there are unique (up to sign) polynomials $f_1(x)$ and $f_2(x)$ such that $g(x^2) = f_1(x)^2 - f_2(x)^2$. Furthermore, in this case, we can write $f_1(x) = h_1(x^2)$ and $f_2(x) = x h_2(x^2)$, where $h_1(x), h_2(x) \in \mathbb{Z}[x]$.

The proof of (b) follows from the fact that $\mathbb{Z}[x]$ is a unique factorization domain.

Lemma 7. *Let $m = 2^r \geq 2$ and n be an odd positive integer. If b has the property $\mathcal{P}$, then $x^{2^i n} + b^m$ is irreducible for every i, $0 \leq i \leq r$.*

Proof. We proceed by induction on i. If $i = 0$ and b has the property $\mathcal{P}$, then $f(x) = x^n + b^m$ is irreducible by Lemma 3. If $i = 1$, then $f(x) = x^n + b^m$ is irreducible, and if $f(x^2) = x^{2n} + b^m = (x^n + b^{m/2})^2 - 2b^{m/2}x^n$ is reducible, from Lemma 6, $2b^{m/2}x^n$ has to be of the form $x^2h(x^2)^2$ for some $h(x) \in \mathbb{Z}[x]$. Since n is odd, this is not possible and hence $f(x^2)$ is irreducible.

Suppose the result is true for some i, $0 \le i \le r$ and we will show that it is true for $i+1 \le r$. So, $f(x) = x^{2^i n} + b^m$ is irreducible for some $i \le r$. From Lemma 6, if

$$f(x^2) = x^{2^{i+1}n} + b^m = (x^{2^i n} + b^{m/2})^2 - 2b^{m/2}x^{2^i n}$$

is reducible, then $2b^{m/2}x^{2^i n}$ has to be of the form $(xg(x^2))^2$ for some $g \in \mathbb{Z}[x]$. This is possible only when $m = 2$ and $b = 2^\alpha b_1^2$, where α, b_1 are odd positive integers. That is $r = 1$ and hence $i = 0$. We have already seen that $f(x^2)$ is irreducible in this case. Therefore, by the induction principle, $x^{2^i n} + b^{2^r}$ is irreducible for every $i \le r$. □

Lemma 8. *Let $m = 2^r \ge 2$ and let b be an odd integer which has the property $\mathcal{P}$. If $m|n$, then $x^n + b^m$ is irreducible.*

Proof. Let $n = mt$. If t is odd, then by Lemma 7, $x^n + b^m$ is irreducible. Let $t = 2^{t_1}u$, $t_1 \ge 1$ and u is odd. From Lemma 7, $g(x) = x^{mu} + b^m$ is irreducible. Since b is odd, $g(k^2) \equiv 2 \pmod 4$ for any odd integer k. Applying Lemma 6 repeatedly to $g(x)$, the result follows. □

Corollary 9. *Let $m = 2^r \ge 2$, $m|n$, and b has the property $\mathcal{P}$. If $x^n + b^m$ is reducible, then both b and $\frac{n}{m}$ are even integers.*

Lemma 10. *Let $t, r \in \mathbb{N}$ and b has the property $\mathcal{P}$. Then $x^{2^r t} + b^{2^r}$ is reducible if and only if t is even, $r = 1$, and $b = 2d^2$ for some $d \in \mathbb{N}$.*

Proof. If t is even, $r = 1$,and $b = 2d^2$, then

$$x^{2t} + 4d^4 = (x^t + 2d^2)^2 - 4d^2x^t = (x^t - 2dx^{t/2} + 2d^2)(x^t + 2dx^{t/2} + 2d^2).$$

Conversely, let $g(x) = x^{2^r t} + b^{2^r}$ is reducible. By Corollary 9, both t and b are even integers. Let $t = 2^{t_1}u, b = 2^{b_1}v$, where u, v are odd integers and $t_1, b_1 \ge 1$. By Lemma 7, the polynomial $h(x) = x^{2^r u} + b^{2^r}$ is irreducible. Since $g(x) = h(x^{2^{t_1}})$ is reducible, there is some i, $1 \le i \le t_1$ such that $h(x^{2^{i-1}})$ is irreducible and

$$h(x^{2^i}) = x^{2^{r+i}u} + 2^{2^r b_1}v^{2^r} = (x^{2^{r+i-1}u} + 2^{2^{r-1}b_1}v^{2^{r-1}})^2 - 2^{2^{r-1}b_1+1}v^{2^{r-1}}x^{2^{r+i-1}u}$$

is reducible. From uniqueness property of Lemma 6, $2^{2^{r-1}b_1+1}v^{2^{r-1}}x^{2^{r+i-1}u}$ has to be of the form $(xl(x^2))^2$ for some $l(x) \in \mathbb{Z}[x]$. This is possible only when $r = 1$ and $2^{2^{r-1}b_1+1}v^{2^{r-1}}$ is a perfect square. Hence, $2^{b_1+1}v = c^2$ would imply $b = 2\left(\frac{c}{2}\right)^2$ with $\frac{c}{2} \in \mathbb{N}$. $\square$

Proof of Theorem 2 and hence Theorem 1 follows from Lemmas 3, 5 and 10.

Acknowledgement. We would like to thank the referee for valuable comments.

References

[1] A. Capelli, *Sulla riduttibilita delle equazioni algebriche*, Nota prima, Red. Accad. Fis. Mat. Soc. Napoli(3), 3(1897), 243–252.

[2] E. S. Selmer, *On the irreducibility on certain trinomials*, Math. Scand., 4 (1956), 287–302.

[3] G. Karpilovsky, *Topics in field theory*, ISBN: 0444872973, North-Holland, 1989.

[4] K. Th. Vahlen, *Über reductible Binome*, Acta Math., 19(1)(1895), 195–198.

[5] L. Rédei, *Algebra*, Erster Teil, Akademische Verlaggesellschaft, Leipzig, 1959.

Biswajit Koley and A.Satyanarayana Reddy
Department of Mathematics
Shiv Nadar University, India-201314
E-mail: bk140@snu.edu.in
E-mail: satyanarayana.reddy@snu.edu.in

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020), 175–180

A NEW GENERALIZATION OF STANLEY'S THEOREM

GEORGE E. ANDREWS AND MIRCEA MERCA
(Received : 27 - 12 - 2019 ; Revised : 13 - 02 - 2020)

Abstract. We consider parts counted without multiplicity in all the partitions of n and obtain a generalization of Stanley's theorem.

1. Introduction

The following result in the theory of integer partitions is widely known as Stanley's theorem. Richard Stanley did independently discover this theorem. However, in an article in the American Mathematical Monthly [2], R. Gilbert, gave the history of this result and traced it back to early work by Nathan Fine. Her article bears the ironic title, *A Fine Rediscovery.*

Theorem 1.1 (Stanley)**.** *The number of* 1*'s in the partitions of* n *equals the number of parts that appear at least once in a given partition of* n*, summed over all the partitions of* n*.*

For example, the partitions of 6 are: 6, $5+1$, $4+2$, $4+1+1$, $3+3$, $3+2+1$, $3+1+1+1$, $2+2+2$, $2+2+1+1$, $2+1+1+1+1$ and $1+1+1+1+1+1$. As we can see, the number of 1's in the partitions of 6 is equal to 19. The number of parts, counted without multiplicity, in all the partitions of 6 is $1+2+2+2+1+3+2+1+2+2+1=19$.

There is a generalization of Stanley's theorem known as Elder's theorem. R. Gilbert [2] also found that Nathan Fine had proved this as well.

Theorem 1.2 (Elder)**.** *The number of* k*'s used in all the partitions of* n *equals the number of parts that appear at least* k *times in a given partition of* n*, summed over all the partitions of* n*.*

We have eight 2's in the partitions of 6, while the number of times parts appear at least twice is $0+0+0+1+1+0+1+1+2+1+1=8$.

2010 Mathematics Subject Classification: 05A17, 11P81.
Key words and phrases: Partition, Stanley's theorem

On the other hand, the number of even parts, counted without multiplicity, in all the partitions of 6 is $1+0+2+1+0+1+0+1+1+1+0 = 8$. Moreover, the partitions of 5 are: 5, $4+1$, $3+2$, $3+1+1$, $2+2+1$, $2+1+1+1$ and $1+1+1+1+1$. The number of odd parts, counted without multiplicity, in all the partitions of 5 is $1+1+1+2+1+1+1 = 8$. Are these identities just a coincidence?

In light of the charming simplicity of the Stanley and Elder theorems along with their surprising history, it is tremendously appealing to explore these topics in greater depth. In this paper, we shall first generalize Stanley's theorem. Subsequently we consider the sum of parts, counted without multiplicity, in all partitions of n.

Theorem 1.3. *Let k, n and r be nonnegative integers such that $0 \leq r < k$. The number of parts equal to k in all the partitions of n is equal to the number of parts $\equiv -r \pmod k$ that appear at least once in a given partition of $n-r$, summed over all the partitions of $n-r$.*

For example, the number of 3's in the partitions of 6 equals 4. The number of parts congruent to $0 \pmod 3$, counted without multiplicity, in all the partitions of 6 is $1+0+0+0+1+1+1+0+0+0+0 = 4$. The number of parts congruent to $-1 \pmod 3$, counted without multiplicity, in all the partitions of 5 is $1+0+1+0+1+1+0 = 4$. The partitions of 4 are: 4, $3+1$, $2+2$, $2+1+1$ and $1+1+1+1$. The number of parts congruent to $-2 \pmod 3$, counted without multiplicity, in all partitions of 4 is $1+1+0+1+1 = 4$.

Theorem 1.4. *The number of 1's in all the partitions of n equals the difference between the sum of parts, counted without multiplicity, in all the partitions of n and the sum of parts, counted without multiplicity, in all the partitions of $n-1$.*

We see that the sum of parts, counted without multiplicity, in all the partitions of 6 is $6+6+6+5+3+6+4+2+3+3+1 = 45$, while the sum of parts, counted without multiplicity, in all the partitions of 5 is $5+5+5+4+3+3+1 = 26$. The number of 1's in the partitions of 6 equals the difference $45-26$.

Upon reflection, motivated by Theorem 1.3, one expects that there might be a more general result where Theorem 1.4 is a very special case.

Theorem 1.5. *The number of k's in the all partitions of n multiplied by k equals the difference between the sum of parts divisible by k, counted without multiplicity, in all the partitions of n and the sums of parts divisible by k, counted without multiplicity, in all the partitions of $n-k$.*

The sum of even parts, counted without multiplicity, in the partitions of 6 is: $6+0+6+4+0+2+0+2+2+2+0=24$, while the sums of even parts counted without multiplicity, in all the partitions of 4 is: $4+0+2+2+0=8$. We have $(24-8)/2=8$, the number of 2's in the partitions of 6.

The rest of this paper is organized as follows. We will first prove Theorem 1.3 in Section 2. In Section 3, we will provide a proof of Theorem 1.5. In the last section, we remark two connections between partitions and divisors.

2. Proof of Theorem 1.3.

First we want the generating function for partitions where z keeps track of the number of parts equal to k. This is

$$\frac{1}{1-zq^k}\prod_{\substack{n=1\\n\neq k}}^{\infty}\frac{1}{1-q^n}=\frac{1}{(q;q)_\infty}\cdot\frac{1-q^k}{1-zq^k}.$$

Let $c_k(n)$ denote the total number of k's in all the partitions of n. Hence

$$\sum_{n=0}^{\infty}c_k(n)q^n=\frac{d}{dz}\Big|_{z=1}\frac{(1-q^k)}{(q;q)_\infty(1-zq^k)}=\frac{q^k}{1-q^k}\cdot\frac{1}{(q;q)_\infty}.$$

Second we consider the generating function for the partitions where z keeps track of the number of time that parts congruent to $-r \pmod k$ appear at least once. This is

$$\begin{aligned}
&\prod_{j=1}^{\infty}\left(1+z(q^{kj-r}+q^{2kj-r}+q^{3kj-r})+\cdots\right)\prod_{\substack{n=1\\n\not\equiv -r \pmod k}}^{\infty}\frac{1}{1-q^n}\\
&=\prod_{j=1}^{\infty}\left(1+\frac{zq^{kj-r}}{1-q^{kj-r}}\right)\prod_{\substack{n=1\\n\not\equiv -r \pmod k}}^{\infty}\frac{1}{1-q^n}\\
&=\frac{1}{(q;q)}\prod_{j=1}^{\infty}\left(1+(z-1)q^{jk-r}\right).
\end{aligned}$$

Let $d_{k,r}(n)$ denote the total number of times parts congruent to $-r$ $\pmod{k}$ appear at least once in all the partitions of n. We have

$$\begin{aligned}\sum_{n=0}^{\infty} d_{k,r}(n)q^n &= \frac{d}{dz}\Big|_{z=1} \frac{1}{(q;q)_\infty} \prod_{j=1}^{\infty} \left(1+(z-1)q^{jk-r}\right) \\ &= \frac{1}{(q;q)_\infty} \sum_{j=1}^{\infty} q^{jk-r} \\ &= \frac{q^{k-r}}{(q;q)_\infty(1-q^k)}.\end{aligned}$$

Hence

$$q^r \sum_{n=0}^{\infty} d_{k,r}(n)q^n = \sum_{n=0}^{\infty} c_k(n)q^n.$$

Therefore $d_{k,r}(n-r) = c_k(n)$.

3. Proof of Theorem 1.5.

As noted before, the generating function for the first class of partitions is

$$\frac{q^k}{1-q^k} \cdot \frac{1}{(q;q)_\infty}.$$

The generating function where z keeps track of parts divisible by k without multiplicity is

$$\prod_{\substack{n=1\\ n\neq k}}^{\infty} \frac{1}{1-q^n} \prod_{j=1}^{\infty} \left(1+z^{jk}(q^{jk}+q^{2jk}+\cdots)\right) = \frac{1}{(q;q)_\infty} \prod_{j=1}^{\infty} \left(1+(z^{jk}-1)q^{jk}\right).$$

If we apply $\frac{d}{dz}\Big|_{z=1}$ to this expression, we obtain

$$\frac{1}{(q;q)_\infty} \sum_{j=1}^{\infty} jkq^{jk} = \frac{kq^k}{(q;q)_\infty(1-q^k)^2}.$$

Thus

$$\frac{kq^k}{(q;q)_\infty(1-q^k)^2} \cdot (1-q^k) = \frac{kq^k}{(q;q)_\infty(1-q^k)},$$

and this confirms the theorem.

4. Concluding remarks.

It is well-known (cf. [3, p. 231]) that logarithmic differentiation of the generating function for $p(n)$ yields

$$\sum_{n=1}^{\infty} np(n)q^n = \frac{1}{(q;q)_\infty} \sum_{n=1}^{\infty} \frac{nq^n}{1-q^n} = \frac{1}{(q;q)_\infty} \sum_{n=1}^{\infty} \sigma(n)q^n,$$

where $\sigma(n)$ is the sum of the positive divisors of n, and this in turn via the Euler's pentagonal number theorem

$$\sum_{n=-\infty}^{\infty} (-1)^n q^{n(3n-1)/2} = (q;q)_\infty$$

that

$$\sum_{j=-\infty}^{\infty} (-1)^j \big(n - j(3j-1)/2\big) p\big(n - j(3j-1)/2\big) = \sigma(n).$$

Nothing that $np(n)$ is clearly the sum of all the parts in the partitions of n, it seems appropriate to conclude this paper with an analogue observation about $C(n)$, the total number of parts in the partitions of n. Here logarithmic differentiation of

$$\prod_{n=1}^{\infty} \frac{1}{1-zq^n}$$

in $z = 1$ yields

$$\sum_{n=1}^{\infty} C(n)q^n = \frac{1}{(q;q)_\infty} \sum_{n=1}^{\infty} \frac{q^n}{1-q^n} = \frac{1}{(q;q)_\infty} \sum_{n=1}^{\infty} d(n)q^n,$$

where $d(n)$ is the number of the positive divisors of n. Consequently

$$\sum_{j=-\infty}^{\infty} (-1)^j C\big(n - j(3j-1)/2\big) = d(n).$$

References

[1] G. E. Andrews, *The Theory of Partitions*, Addison-Wesley Publishing, Reading, MA, 1976, reissued by Cambridge University Press, Cambridge, 1998.

[2] R. A. Gilbert, A Fine rediscovery, *Amer. Math. Monthly*, **122** (2015) 322–331.

[3] H. Rademacher, Topics in Analytic Number Theory, Springer, Berlin, 1973.

George E. Andrews
Department of Mathematics
The Pennsylvania State University
University Park, PA 16802 USA
E-mail address: gea1@psu.edu

Mircea Merca
Academy of Romanian Scientists
Ilfov 3, Sector 5, Bucharest, Romania
E-mail address: mircea.merca@profinfo.edu.ro

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January- June (2020), 181–192

PROBLEM SECTION

In the Math. Student Vol. 88 (3-4) 2019, we had invited solutions from the floor to the remaining problems *1, 3, 5, 6, 7* and *8* mentioned in the Math. Student Vol. 88 (1-2) 2019 as well as solutions to the eight new problems proposed in the Math. Student Vol. 88 (3-4) 2019, till April 30, 2020.

We received one correct solution to Problem 3 [MS 88(1-20) 2019] and one correct solutions to Problem 6 [MS 88 (3-4) 2019] from the floor and we publish these solutions here. One solution to Problem 7 [MS 88 (3-4) 2019] was received but it was not correct. We did not receive solutions to other problems so we are printing here the solutions provided by the proposers of the Problems.

We first present seven new problems. We invite Solutions to these problems and solutions to the remaining problems of MS 88 (3-4) 2019 from the researchers till September 30, 2020. Correct solutions received from the researchers by September 30, 2020 will be published in the MS 89 (3-4) 2020.

The following two problems have been proposed by Prof. Ram Murty, Department of Mathematics, Queen's University, Ontario, Canada.

MS 89 (1-2) 2020: Problem 1. Let p be a prime number and suppose that $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{Q}$ is a non-zero function. We may view f as a function on the natural numbers by setting $f(n)$ to be $f(a)$ where $n \equiv a \pmod{p}$. Suppose that $\sum_{a=1}^{p} f(a) = 0$. Show that

$$\sum_{n=0}^{\infty} \frac{f(n)}{n!}$$

is transcendental.

MS 89 (1-2) 2020: Problem 2. Prove that for $n \geq 1$,

$$\sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} \frac{1}{k^2} = \sum_{k=1}^{n} \frac{H_k}{k}$$

where $H_n = 1 + 1/2 + \cdots + 1/n$.

The following problem is proposed by Dr. Siddhi Pathak, Department of Mathematics, Pennsylvania State University, State College, PA 16802, USA.

MS 89 (1-2) 2020: Problem 3. Let $f : \mathbb{R}\backslash\{1\} \to \mathbb{R}$ be an integrable function satisfying the condition that

$$f(x)\, f\left(1 - \frac{1}{x}\right) = -x \quad \text{for all } x \in \mathbb{R}\backslash\{0,\, 1\}.$$

Suppose that $f(x) > 0$ for $x > 1$. Then show that

$$\int_2^{\infty} (f(x) - 1)^2\, dx = 1.$$

The following two problems are proposed by Dr. Anup Dixit, Department of Mathematics, Queen's University, Ontario, Canada.

MS 89 (1-2) 2020: Problem 4. Let a, b, c be positive real numbers. Show that

$$\frac{3a}{4b + 5c} + \frac{3b}{4c + 5a} + \frac{3c}{4a + 5b} \geq 1.$$

MS 89 (1-2) 2020: Problem 5. Find all natural numbers $n > 1$ such that $n \mid 2^n + 1$.

Prof. M. M. Shikare, the Editor of the Mathematics Student has proposed the following two problems.

MS 89 (1-2) 2020: Problem 6. Let G be a graph with 10 vertices. Among any three vertices of G, at least two are adjacent. Find the least number of edges that G can have. Find a graph with this property.

MS 89 (1-2) 2020: Problem 7. At the end of a birthday party, the hostess wants to give away candies. She has 6 types of cookies. Each child is given a gift packet which contains two types of cookies. Each type of cookie is used in combination with at least three others. Prove there are three children, who between them, have all the six types of cookies.

Solutions to the Old Problems

MS 88 (1-2) 2019, Problem- 3 (Proposed by Prof. George E. Andrews, The Pennsylvania State University, PA, USA).

Prove that the number of partitions of n with no part congruent to 3 mod 4 equal the number of partitions of n in which the difference between any two parts cannot be 1 if the larger part is even.

For example, for $n = 7$, the nine partitions of 7 with no part congruent to 3 mod 4 are: $6+1, 5+2, 5+1+1, 4+2+1, 4+1+1+1, 2+2+2+1, 2+2+1+1+1, 2+1+1+1+1+1, 1+1+1+1+1+1+1$. The nine partitions in which no even part is followed by a part that is 1 less are: $7, 6+1, 5+2, 5+1+1, 4+1+1+1, 3+3+1, 3+2+2, 3+1+1+1+1, 1+1+1+1+1+1+1$.

Solution (Provided by Subhas Chand Bhoria, Village Bakali, Kurukshetra, Haryana; the solution was received by the Editor on April 23, 2020).

Let $P(n)$ denotes the number of partitions of n with no part congruent to 3 mod 4. Also let $Q(n)$ denotes the number of partitions of n in which difference between any two parts cannot be 1 if the larger part is even. We begin by writing generating function for the partitions of $Q(n)$ type and show that these are equinumerous to the partitions of $P(n)$ type.

$$\begin{aligned}
\sum_{n=0}^{\infty} Q(n)q^n &= \prod_{n=1}^{\infty}\left(1+\frac{q^{2n-1}}{1-q^{2n-1}}+\frac{q^{2n}}{1-q^{2n}}\right)\\
&= \prod_{n=1}^{\infty}\left(\frac{1-q^{2n-1}-q^{2n}+q^{4n-1}+q^{2n-1}-q^{4n-1}+q^{2n}-q^{4n-1}}{(1-q^{2n-1})(1-q^{2n})}\right)\\
&= \prod_{n=1}^{\infty}\left(\frac{1-q^{4n-1}}{(1-q^{2n-1})(1-q^{2n})}\right)\\
&= \prod_{n=1}^{\infty}\left(\frac{1-q^{4n-1}}{(1-q^{n})}\right)\\
&= \prod_{n=1}^{\infty}\left(\frac{1-q^{4n-1}}{(1-q^{4n})(1-q^{4n-1})(1-q^{4n-2})(1-q^{4n-3})}\right)\\
&= \prod_{n=1}^{\infty}\left(\frac{1}{(1-q^{4n})(1-q^{4n-2})(1-q^{4n-3})}\right)\\
&= \sum_{n=0}^{\infty} P(n)q^n.
\end{aligned}$$

This completes the proof.

Remarks:

A similar result can be observed as number of partitions of n with no part congruent to 1 mod 4 (but greater than 1) equals the number of partitions of n in which the difference between any two parts cannot be 1 if the larger part is odd. For example, for $n = 5$ there are six partitions of

5 with no part congruent to 1 mod 4 (but greater than 1): $4+1, 3+2, 3+1+1, 2+2+1, 2+1+1+1, 1+1+1+1+1$.

The six partitions in which no odd part is followed by a part that is 1 less are: $5, 4+1, 3+1+1, 2+2+1, 2+1+1+1, 1+1+1+1+1$.

To prove this remark the generating function identity will be:

$$\frac{1}{1-q}\prod_{n=1}^{\infty}\left(1+\frac{q^{2n}}{1-q^{2n}}+\frac{q^{2n+1}}{1-q^{2n+1}}\right)=\prod_{n=1}^{\infty}\left(\frac{1-q^{4n+1}}{1-q^{n}}\right).$$

MS 88 (3-4) 2019, Problem- 6 (Proposed by Prof. Ram Murty, Queen's University, Onterio, Canada).
Prove that $\int_0^1 \frac{x-1}{\log x} dx$ is transcendental.

Solution (Provided by Subhas Chand Bhoria, Village Bakali, Kurukshetra, Haryana; the solution was received by the Editor on April 23, 2020).

First, we derive a more general result, namely, for $w > 0$,

$$I(w) := \int_0^1 \frac{x^w - 1}{\log x} dx = \log(1+w). \tag{1}$$

Thus, in particular, for $w = 1$, the value of the required integral will be $\log(2)$. This is a transcendental number since $\log(n)$ is transcendental for any natural number $n > 1$.

Now we shall give a proof of the equation (1).

It is easy to verify that $\frac{d}{dt}(\frac{x^t-1}{\log x}) = x^t$, where x and t are independent variable. Now we integrate both sides from 0 to w with respect to the variable t, to get

$$\int_0^w \frac{d}{dt}\left(\frac{x^t-1}{\log x}\right) dt = \int_0^w x^t dt$$
$$\Rightarrow \frac{x^w-1}{\log x} = \int_0^w x^t dt.$$

Again we integrate from 0 to 1 with respect to the variable x, to obtain

$$\int_0^1 \frac{x^w-1}{\log x} dx = \int_0^1 \int_0^w x^t \, dt \, dx,$$

now we will interchange the integral order in the right hand side, this is possible because of the Fubini's theorem, since the function $f(x,t) = x^t$

is a continuous function in the domain $[0,1] \times [0,w]$ and $\int_0^w \int_0^1 x^t \, dx \, dt$ is finite. (This integral we evaluated below). Thus we will have

$$\begin{aligned}\int_0^1 \frac{x^w - 1}{\log x} dx &= \int_0^w \int_0^1 x^t \, dx \, dt \\ &= \int_0^w \frac{1}{t+1} dt \\ &= \log(1+w).\end{aligned}$$

This completes the proof of (1). Therefore, in general, $I(n) = \log(1+n)$ is transcendental for any natural number n.

MS 88 (1-2) 2019, Problem 1 (Proposed by Prof. R. P. Pakshirajan, Mysore). Let g be an arbitrary real continuous function defined on $[0,1]$. Let μ be the standard Wiener measure on the Borel σ-field $\mathscr{C}$ of the space $\mathbf{C}$ of real continuous functions on $[0,1]$ all vanishing at zero. Let $\alpha_n = \sum_{j=1}^{2^n} g(t_{j,n})(x(\frac{j}{2^n}) - x(\frac{j-1}{2^n}))$, $t_{j,n}$ being an arbitrary point in $(\frac{j-1}{2^n}, \frac{j}{2^n})$. Show that α_n is a Cauchy sequesnce in $L^2(\mu)$.

Desribe the distribution of the limit function α (=$\lim_n \alpha_n$) under μ.

Solution (by Prof. R. P. Pakshirajan).

Step 1. Since g is a continuous function defined on $[0,\ 1]$, it will be automatically uniformly continuous on $[0,\ 1]$. Let $t_{j,n}$ be an arbitrary point in the interval $[\frac{j-1}{2^n},\ \frac{j}{2^n}]$, $j = 1,\ 2, \ldots,\ 2^n$, $n = 1,\ 2, \ldots.$. Let $g_n = \frac{1}{2^n} \sum_{j=1}^{2^n} g(t_{j,n})$. Then (g_n) is a Cauchy sequence of real numbers and hence is convergent. *Proof.* Given $\varepsilon > 0$, choose $\mathbf{m}$ such that $|g(t) - g(t')| < \varepsilon$ whenever $|t - t'| < \frac{1}{2^{\mathbf{m}}}$. This is possible since g is uniformly continuous. Let $\mathbf{m} < m < n$. Now

$$\begin{aligned} g_n - g_m &= \frac{1}{2^n} \sum_{j=1}^{2^n} g(t_{j,n}) - \frac{1}{2^m} \sum_{j=1}^{2^m} g(t_{j,m}) \\ &= \frac{1}{2^m} \sum_{j=1}^{2^m} \frac{1}{2^{n-m}} \sum_{r=1}^{2^{n-m}} g(t_{2^{n-m}(j-1)+r,\ n}) - \frac{1}{2^m} \sum_{j=1}^{2^m} g(t_{j,m}) \\ & \frac{1}{2^m} \sum_{j=1}^{2^m} \frac{1}{2^{n-m}} \sum_{r=1}^{2^{n-m}} \Big\{ g(t_{2^{n-m}(j-1)+r,\ n}) - g(t_{j,m}) \Big\}. \end{aligned}$$

All the numbers $t_{2^{n-m}(j-1)+r,\ n}$, $1 \le r \le 2^{n-m}$ lie in the interval $[\frac{j-1}{2^m},\ \frac{j}{2^m}]$. Hence

$$|g_n - g_m| \le \frac{1}{2^m}\sum_{j=1}^{2^m}\frac{1}{2^{n-m}}\sum_{r=1}^{2^{n-m}}\varepsilon = \varepsilon.$$

That (g_n) is a Cauchy sequence, and hence a convergent sequence of numbers, follows from this.

Note. Since $|g_n| \le \|g\|$ and since g is a continuous function, there exists $t \in [0,\ 1]$ such that $g(t) = \lim_{n\to\infty} g_n$.

Step 2.

Let g denote an arbitrary continuous function defined on $[0,\ 1]$. When the underlying measure on $\mathscr{C}$ is μ, consider the sequence of *rv*s (α_n) where $\alpha_n(x) = \sum_{j=1}^{2^n} g(t_{j,n})\Big(x(\frac{j}{2^n}) - x(\frac{j-1}{2^n})\Big)$ ($t_{j,n} \in (\frac{j-1}{2^n},\ \frac{j}{2^n})$) being an arbitrary point. We note $\alpha_n \in L^2(\mu)$. This is because, under μ, α_n is a normal variable defined on $(\mathbf{C},\ \mathscr{C},\ \mu)$. Let us examine if this sequence has a limit in the norm topology of $L^2(\mu)$.

$$\mathbb{E}\alpha_n^2 = \sum_{j=1}^{2^n}\big(g(t_{j,n})\big)^2\frac{1}{2^n} \to \int_0^1 \big(g(t)\big)^2\,\mathrm{d}\,t \tag{2}$$

since g^2, being a continuous function, is Riemann integrable. Let $1 \le m < n < \infty$.

$$\begin{aligned}\mathbb{E}\alpha_m\alpha_n &= \mathbb{E}\big[\sum_{j=1}^{2^m}\{(x(\frac{j}{2^m}) - x(\frac{j-1}{2^m}))g(t_{j,m})\}\times \\ &\quad \{\sum_{j=1}^{2^n}(x(\frac{j}{2^n}) - x(\frac{j-1}{2^n}))g(t_{j,n})\}\big] \\ &= \mathbb{E}\Big[\sum_{j=1}^{2^m}\{(x(\frac{j}{2^m}) - x(\frac{(j-1)}{2^m}))g(t_{j,m})\}\times \\ &\quad \{\sum_{j=1}^{2^m}(\sum_{r=1}^{2^{n-m}}(x(\frac{(j-1)}{2^m}+\frac{r}{2^n}) - x(\frac{(j-1)}{2^m}+\frac{r-1}{2^n}))\times \\ &\quad g(t_{2^{n-m}(j-1)+r,n})\}\Big] \\ &= \sum_{j=1}^{2^m} g(t_{j,m})\times\Lambda_j,\ \text{say, where}\end{aligned}$$

$$\begin{aligned}\Lambda_j = \mathbb{E}\Big[&\big(x(\frac{j}{2^m}) - x(\frac{(j-1)}{2^m})\big)\times \\ &\sum_{r=1}^{2^{n-m}} \Big\{\big(x(\frac{2^{n-m}(j-1)+r}{2^n}) - x(\frac{2^{n-m}(j-1)+r-1}{2^n})\big)\times \\ &\big(g(t_{2^{n-m}(j-1)+r,\ n})\big)\Big\}\Big] \\ = &\sum_{r=1}^{2^{n-m}} g(t_{2^{n-m}(j-1)+r,\ n})\times \\ &\times \mathbb{E}\Big\{\big(x(\frac{2^{n-m}j}{2^n}) - x(\frac{2^{n-m}(j-1)}{2^n})\big) \times \big(x(\frac{2^{n-m}(j-1)+r}{2^n}) \\ &\quad - x(\frac{2^{n-m}(j-1)+r-1}{2^n})\big)\Big\} \\ = &\sum_{r=1}^{2^{n-m}} g(t_{2^{n-m}(j-1)+r,\ n})\times \\ &\quad \times \mathbb{E}\big(x(\frac{2^{n-m}(j-1)+r}{2^n}) - x(\frac{2^{n-m}(j-1)+r-1}{2^n})\big)^2 \\ = &\sum_{r=1}^{2^{n-m}} g(t_{2^{n-m}(j-1)+r,\ n})\frac{1}{2^n}\end{aligned}$$

(the other terms in the product being zero each)

Hence

$$\begin{aligned}&\lim_{m\to\infty}\lim_{n\to\infty}\mathbb{E}\alpha_m\alpha_n \\ &\quad = \lim_{m\to\infty}\lim_{n\to\infty}\sum_{j=1}^{2^m} g(t_{j,m})\frac{1}{2^m}\Big\{\sum_{r=1}^{2^{n-m}} g(t_{2^{n-m}(j-1)+r,\ n})\frac{1}{2^{n-m}}\Big\} \\ &\quad = \lim_{m\to\infty}\frac{1}{2^m}\sum_{j=1}^{2^m} g(t_{j,m})\lim_{n\to\infty}\Big\{\frac{1}{2^{n-m}}\sum_{r=1}^{2^{n-m}} g(t_{2^{n-m}(j-1)+r,\ n})\Big\}.\end{aligned}$$

All the points $t_{2^{n-m}(j-1)+r,n}$, $r = 1, 2, \ldots, 2^{n-m}$ lie in the interval $[\frac{j-1}{2^m}, \frac{j}{2^m}]$. By Step 1 above, there exists $t'_{j,m}$ such that

$$\lim_{n\to\infty}\Big\{\frac{1}{2^{n-m}}\sum_{r=1}^{2^{n-m}} g(t_{2^{n-m}(j-1)+r,\ n})\Big\} = g(t'_{j,m}).$$

We have

$$\lim_{m\to\infty}\lim_{n\to\infty}\mathbb{E}\alpha_m\alpha_n = \lim_{m\to\infty}\frac{1}{2^m}\sum_{j=1}^{2^m} g(t_{j,m})g(t'_{j,m}) = \int_0^1 g^2(t)\,\mathrm{d}\,t \qquad (3)$$

Results (1) and (2) together imply (α_n) is a Cauchy sequence. Since $L^2(\mu)$ is a complete metric space, there exists $\alpha \in L^2(\mu)$ to which the sequence (α_n) converges in $L^2(\mu)$ norm.

Distribution of α.
It is obvious that under μ, α_n is a zero mean normal variable. Since sequence (α_n) converges in L^2 norm, it converges in distribution. Thus necessarily, the limit variable α is normally distributed with zero mean and variance $\int_0^1 g^2(t)dt \; \big(= \lim_{n\to\infty}\mathbb{E}\alpha_n^2\big)$.

MS 88 (1-2) 2019, Problem 5 (Proposed by Prof. B. S. Sury, ISI, Bangalore). Let θ be a non-zero real number and let N be a positive integer. Consider the fractional parts of $n\theta$ for $0 \le n < N$. This creats a maximum of N possible intervals in [0, 1]. Prove that there are only three possible lengths for these intervals.

Solution (by Prof. B. S. Sury).

There were some individual queries made to the proposer that indicated that some people found it difficult to understand the question. Therefore, we give an example first to show what is meant.
Suppose $\theta = \sqrt{2}$ and $N = 6$. Then, the fractional parts $\{n\theta\}$ for $n = 1, 2, 3, 4, 5$ are, respectively,

$\{\sqrt{2}\} = \theta - 1, \{2\sqrt{2}\} = 2\theta - 2, \{3\sqrt{2}\} = 3\theta - 4, \{4\sqrt{2}\} = 4\theta - 5, \{5\sqrt{2}\} = 5\theta - 7.$

Arrange them in order to get

$$0 < 5\theta - 7 < 3\theta - 4 < \theta - 1 < 4\theta - 5 < 2\theta - 2 < 1.$$

The lengths of these six intervals are then

$$5\theta - 7, 3 - 2\theta, 3 - 2\theta, 3\theta - 4, 3 - 2\theta, 3 - 2\theta$$

which shows there are just three lengths $5\theta - 7, 3 - 2\theta$ and their sum $3\theta - 4$.

Let us now solve the problem in general.
Let θ and N be given and we consider the fractional parts of $n\theta$ for $n = 1, 2, \cdots, N$ (the notation here is chosen for convenience and is slightly

different from the problem where I took $n < N$).
Call $\{a\theta\}$ and $\{b\theta\}$ be the largest among these N fractional parts. Put

$$\alpha = \{a\theta\}, \beta = 1 - \{b\theta\}.$$

Now, clearly $\{(a+b)\theta\}$ equals either $\{a\theta\} + \{b\theta\}$ or $\{a\theta\} + \{b\theta\} - 1$.
In either case, it is either less than $\{a\theta\}$ or larger than $\{b\theta\}$. This means therefore that $a + b \notin \{1, 2, \cdots, N\}$. That is, $N \leq a + b - 1$.
Mark on the interval $[0, 1]$, the points $P_r = \{r\theta\}$. We consider the 'positive' segments P_rP_s of length $\{(s-r)\theta\}$ if $s > r$ and of length $1 - \{(r-s)\theta\}$ in case $r > s$. Hence, by choice of α, β, we have that the length P_rP_s is $\geq \alpha$ if $s > r$ and $\geq \beta$ if $r > s$; call these inequalities as $(*)$.
Moreover, equality holds in the first inequality above iff $s - r = a$ and in the second one iff $r - s = b$.
When the P_r's are arranged in increasing order from left to right on $[0, 1]$, we look at 'close' segments P_rP_s where P_r immediately precedes P_s and write $P_r \to P_s$.
By $(*)$, we have that $P_0 \to P_a$, $P_1 \to P_{1+a}$ etc. by steps of length $\{a\theta\}$. These $P_r \to P_{r+a}$ for $0 \leq r \leq N - a$ are $N + 1 - a$ such close segments, each of length α.
In the same manner, using $(*)$ again, we have $N + 1 - b$ close segments $P_s \to P_{s-b}$ for $b \leq s \leq N$ each of length $\beta = 1 - \{b\theta\}$.
We claim that the rest of the close segments are of the form $P_t \to P_{t+a-b}$ where $N - a < t < b$; and that the common length of all these close segments is $\alpha + \beta$.
To see this, first consider the case $\{v\theta\} > \{t\theta\}$ with $v > t$ and t in the range $N - a < t < b$. Then, $v - t \leq N - t < a$ and so $\{v\theta\} - \{t\theta\} > \alpha$. So

$$\{v\theta\} - \{t\theta\} = \{a\theta - (a - v + t)\theta\} = \alpha + (1 - \{(a - v + t)\theta\}.$$

If $a - v + t \neq b$ (that is, if $v \neq t + a - b$), the second term $(1 - \{(a - v + t)\theta\}$ above is $> \beta$ and, so, we will have $\{v\theta\} - \{t\theta\} > \alpha + \beta$ when $v \neq t + a - b$. If $v = t + a - b$, we get equality with $\alpha + \beta$. The above argument gives the asserted expression if $a > b$. If $a < b$, the asserted expression comes from the case $\{v\theta\} > \{t\theta\}$ but $v < t$.
Therefore, we have shown that the $N + 1$ close segments (including the last one covering the end point 1) are of three types:
(I) $N + 1 - a$ of the type $P_r \to P_{r+a}$ $(0 \leq r \leq N - a)$ each of length α;

(II) $N+1-b$ of the type $P_s \to P_{s-b}$ $(b \leq s \leq N)$ each of length β; and
(III) $a+b-N-1$ of the type $P_t \to P_{t+a-b}$ $(N-a < t < b)$ each of length $\alpha+\beta$.
Therefore, there are only three possible lengths and one of them is the sum of the other two.
This property has been generalized to higher dimensions.

MS 88 (1-2) 2019, Problem 6(Proposed by Prof. B. S. Sury, ISI, Bangalore). If V is a vector space of finite dimension over a finite field F, find the minimum number of proper subspaces of V whose union is the whole of V. Answer the same question when the proper subspaces have dimension at most $dim\ V - k$ for some $k < dim\ V$

Solution (by Prof. B. S. Sury).

The idea is already contained in the solution of the case when $dim(V) = 2$. Indeed, if $|F| = q$ and $dim(V) = 2$, then we claim that the $q+1$ "lines" (one-dimensional subspaces) spanned by $(1,0)$ and $(0,t)$ as t varies in F cover V. Indeed, any two of these lines intersect only in $(0,0)$ (any two points will determine a line) and each line has exactly q points. Therefore, counting all the non-zero points in this union, we have the number $(q+1)(q-1) = q^2-1$ which is the number of points in $V \setminus \{0\}$. Thus, the union of these $q+1$ lines is V. Also, the argument used shows that q lines cannot cover V as the union can have at the most $q(q-1)$ non-zero points. Thus, the least number is $q+1$.
Let us prove in general now that $q+1$ is the least number.
Let $dim(V) = n \geq 2$ and let $\{v_1, v_2\}$ be a linearly independent set in V. Extend it to a basis $\{v_1, \cdots, v_n\}$ of V.
For any fixed $(c_1, c_2) \neq (0,0) \in F^2$, consider the $(n-1)$-dimensional space

$$W(c_1, c_2) := \{\sum_{i=1}^{n} c_i v_i : c_i \in F\ \forall\ i > 2\}.$$

Now, $W(c_1, c_2) = W(d_1, d_2)$ if and only if $(c_1, c_2) = t(d_1, d_2)$ for some $t \in F^*$. Thus, these are $(q^2-1)/(q-1) = q+1$ proper subspaces. It is obvious that they cover V (the vectors $v = \sum_{i=3}^{n} c_i v_i$ are in every $W(c_1, c_2)$ and the others are in a unique $W(c_1, c_2)$).
To answer the question about the least number of proper subspaces of codimension at least k, we observe that the problem solved above is the case $k = 1$ and also observe that while expressing V as a smallest union of

proper subspaces, we may assume without loss of generality that all the proper subspaces have codimension 1 (simply by expanding them if they are not). Thus, the least number of proper subspaces (corresponding to $k = 1$) is the smallest number that is at least $|P(V)|/|P(V/F)| = \frac{q^n-1}{q^{n-1}-1}$; this is $q + 1$. For general codimension k, we may work with the projective space of V and V/F^k which makes it into the previous problem of codimension ≥ 1 spaces. The required number is the smallest integer which is $\geq \frac{q^n-1}{q^{n-k}-1}$. Thus, the answer to the second part is $\left\lceil \frac{q^n-1}{q^{n-k}-1} \right\rceil$.
See A. Khare's paper 'Vector spaces as unions of proper subspaces' in 'Linear Algebra and its Applications, Vol. 431 (2009), P. 1681-1686' for a vast generalization of this problem.

MS 88 (1-2) 2019, Problem 7 (Proposed by Prof. B. S. Sury, ISI, Bangalore). For a positive integer n, a prime $p \leq n$, and any $0 \leq k < p$, show that the power of p dividing $\sum_{r \equiv k \ mod \ p} \binom{n}{r}(-1)^r$ is at least the greatest integer $\leq (n-1)/(p-1)$.

Solution (by Prof. B. S. Sury).

We shall prove that for each $j < p$, p^q divides $\sum_{r \equiv j \ mod \ p} \binom{n}{r}(-1)^r$ where $q = [\frac{n-1}{p-1}]$.
For this, we will realize the binomial sum as an entry of a permutation matrix. Indeed, let P be the $p \times p$ matrix whose i-th columns is the $(i-1)$-th column of the identity matrix (where the numbers i are considered modulo p). The permutation matrix P corresponds to the p-cycle $(1, 2, \cdots, p)$; so, $P^p = I$. Hence, $p_{ij} = 1$ if $j \equiv i + 1 \bmod p$ and 0 otherwise, Hence, P^r has (i, j)-th entry 1 for $j \equiv i + r \bmod p$, and 0 otherwise. In other words,

$$\sum_{r \equiv j \ mod \ p} \binom{n}{r}(-1)^r = ((I - P)^n)_{1,j+1}.$$

We will prove in fact that p^q divides the entries of $(I - P)^n$. Now the polynomial $(1-X)^p = (1-X^p)+p(1-X)f$ for some polynomial $f \in \mathbb{Z}[X]$ since p divides $\binom{p}{k}$ for $0 < k < p$. Now, since $P^p = I$, we have

$$(I - P)^p = p(I - P)f(P).$$

An easy induction shows that for all $q \geq 1$, we have

$$(I - P)^{q(p-1)+1} = p^q(I - P)f(P)^q.$$

Now, take $q = [\frac{n-1}{p-1}]$. Then, $q(p-1) \leq n-1$. So, writing $k = (n-1) - q(p-1) \geq 0$ we have $n = k + q(p-1) + 1$. Hence

$$(I-P)^n = (I-P)^k (I-P)^{q(p-1)+1} = p^q (I-P)^{k+1} f(P)^q.$$

Therefore, we clearly have each entry $(I-P)^n$ to be a multiple of p^q.

MS 88 (1-2) 2019, Problem 8 (Proposed by Prof. B. S. Sury, ISI, Bangalore). Given any positive integer n, characterize all positive integers m that there is an invertible $n \times n$ matrix with rational entries whose order is m.

Solution (by Prof. B. S. Sury).

As any finite subgroup of $GL_n(\mathbb{Q})$ is conjugate to a matrix in $GL_n(\mathbb{Z})$, we need consider only the latter group. We will show that the group $GL_n(\mathbb{Z})$ has an element of order $m = \prod_{i=1}^t p_i^{e_i}$ if, and only if, $\sum_i p_i^{e_i-1}(p_i - 1) \leq n+1$ (respectively $\leq n$) when $n \equiv 2 \bmod 4$ (respectively, $n \not\equiv 2 \bmod 4$).
We will first show the sufficiency of the condition.
To see this, observe that the companion matrix C_d of the cyclotomic polynomial $\Phi_d(X)$, then C_d has order d.
Let $m = \prod_{i=1}^t p_i^{e_i}$, where firstly suppose $p_1^{e_1} \neq 2$. Then, the block matrix whose blocks are $C_{p_i^{e_i}}$ of sizes $\phi(p_i^{e_i})$ with orders $p_i^{e_i}$ has size $\sum_{i=1}^t \phi(p_i^{e_i})$. If this sum equals n, this block matrix has order $m = LCM(p_i^{e_i})$. If this sum is less than n, add a block of identity matrix to make an $n \times n$ matrix of order m.
Now, let $p_1^{e_1} = 2$. Then, $\sum_{i=1}^t \phi(p_i^{e_i}) - 1 = \sum_{i=2}^t \phi(p_i^{e_i})$.
As $m/2$ is odd, and $m/2 = \prod_{i=2}^t p_i^{e_i}$, the previous case shows there is an $n \times n$ matrix B of order $m/2$ if $\prod_{i=2}^t p_i^{e_i} \leq n$. So, $-B$ has order m since $m/2$ is odd.
Conversely, if $GL_n(\mathbb{Z})$ has an element M of order m, then the minimal polynomial of M is the product of the cyclotomic polynomials $\Phi_{d_1}, \cdots, \Phi_{d_r}$ where $LCM(d_1, \cdots, d_r) = m$. Then the matrix M is conjugate to the block matrix of blocks $M_1, \cdots, M_r$ with minimal polynomial of M_i is Φ_{d_i}. Let M_i have size $k_i \times k_i$. Then, $n = k_1 + \cdots + k_r$ and also $k_i \geq \phi(d_i)$ as the minimal polynomial of M_i has degree $\phi(d_i)$.
But, an easy elementary argument shows easily that if $m = \prod_{i=1}^t p_i^{e_i}$, then $\sum_{i=1}^t \phi(p_i^{e_i}) \leq \sum_{i=1}^r \phi(d_i)$ because $m = LCM(d_1, \cdots, d_r)$.
Therefore, $\sum_{i=1}^t \phi(p_i^{e_i}) \leq \sum_{i=1}^r \phi(d_i) \leq n$ which shows the necessity of the condition.

FORM IV
(See Rule 8)

1.	Place of Publication:	PUNE
2.	Periodicity of publication:	QUARTERLY
3.	Printer's Name: Nationality: Address:	DINESH BARVE INDIAN PARASURAM PROCESS 38/8, ERANDWANE PUNE-411 004, INDIA
4.	Publisher's Name: Nationality: Address:	SATYA DEO INDIAN GENERAL SECRETARY THE INDIAN MATHEMATICAL SOCIETY HARISH CHANDRA RESEARCH INSTITUTE CHHATNAG ROAD, JHUNSI ALLAHABAD-211 019, UP, INDIA
5.	Editor's Name: Nationality: Address:	M. M. SHIKARE INDIAN CENTER FOR ADVANCED STUDY IN MATHEMATICS, S. P. PUNE UNIVERSITY PUNE-411 007, MAHARASHTRA, INDIA
6.	Names and addresses of individuals who own the newspaper and partners or shareholders holding more than 1% of the total capital:	THE INDIAN MATHEMATICAL SOCIETY

I, Satya Deo, the General Secretary of the IMS, hereby declare that the particulars given above are true to the best of my knowledge and belief.

SATYA DEO

Dated: May 6, 2020 Signature of the Publisher

Published by Prof. Satya Deo for the Indian Mathematical Society, type set by M. M. Shikare, "Krushnakali", Servey No. 73/6/1, Gulmohar Colony, Jagtap Patil Estate, Pimple Gurav, Pune 411061 and printed by Dinesh Barve at Parashuram Process, Shed No. 1246/3, S. No. 129/5/2, Dalviwadi Road, Barangani Mala, Wadgaon Dhayari, Pune 411 041 (India). Printed in India

The Mathematics Student ISSN: 0025-5742
Vol. 89, Nos. 1-2, January-June (2020)

Back volumes of our periodicals, except for a few numbers out of stock, are available.

Edited by M. M. Shikare and published by Satya Deo, the General
Secretary of the IMS, for the Indian Mathematical Society.

Type set by M. M. Shikare, “Krushnakali”, Survey No. 73/6/1, Gulmohar Colony, Jagtap Patil Estate, Pimple Gurav, Pune 411061 and printed by Dinesh Barve at Parashuram Process, Shed No. 1246/3, S. No.129/5/2, Dalviwadi Road, Barangani Mala, Wadgaon Dhayari, Pune - 411 041, Maharashtra, India. Printed in India