ISSN 1592-7415 eISSN 2282-8214

Volume n. 39 – 2020

RATIO MATHEMATICA

Journal of Mathematics, Statistics, and Applications

Publisher
A.P.A.V.
Accademia Piceno – Aprutina dei Velati in Teramo
www.eiris.it – www.apav.it
apavsegreteria@gmail.com

RATIO MATHEMATICA
Journal of Mathematics, Statistics, and Applications

ISSN 1592-7415 - eISSN 2282-8214

Ratio Mathematica is an International, double peer-reviewed, open access journal, published every six months (June-December).

Ratio Mathematica publishes original research articles on theoretical Mathematics and Statistics. However, contributions with applications to Social Science, Engineering, and Economics are welcome. Only English-language publications are accepted.

The main topics of interest for Ratio Mathematica are:

-Advances in theoretical mathematics and statistics;
-Applications of mathematics and statistics to social science, engineering, ecology, and economics;
-Decision-making in conditions of uncertainty;
-Fuzzy logic;
-Probability theory;
-New theories for dissemination and communication of mathematics and statistics.

Note on Peer-Review
All manuscripts are subjected to a double-blind review process. The reviewers are selected from the editorial board, but they also can be external subjects. The journal’s policies are described at:
http://eiris.it/ojs/index.php/ratiomathematica/about/submissions#authorGuidelines

Publisher: **APAV** - Accademia Piceno-Aprutina dei Velati in Teramo
Tax Code 92036140678, **Vat Id. Number** 02184450688
Registered Office Address: via del Concilio, 24 - 65121 Pescara
Operational Office: via Chiarini, 191 - 65126 Pescara

Dass, Bal Khishan, Department of Mathematics, University of Delhi, Delhi, India.

De Sanctis, Angela, Department of Business Administration, University of Chieti-Pescara, Pescara, Italy.

Di Battista, Tonio, Department of Philosophical, Pedagogical and Economic-Quantitative Sciences University of Chieti-Pescara, Chieti, Italy.

Dramalidis, Achilles, Department of Education Sciences in Pre-School Age, Democritus University of Thrace, Alexandroupolis, Greece.

Eugeni, Franco, Department of Communication Science, University of Teramo, Teramo, Italy.

Fensore, Stefania, Department of Philosophical, Pedagogical and Economic-Quantitative Sciences University of Chieti-Pescara, Chieti, Italy.

Gattone, Stefano Antonio, Department of Philosophical, Pedagogical and Economic-Quantitative Sciences University of Chieti-Pescara, Chieti, Italy.

Jánský, Jiří, Department of Mathematics and Physics, University of Defence, Brno, Czech Republic, Brno, Czech Republic.

Kacprzyk, Janusz, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland.

Mari, Carlo, Department of Economics, University of Chieti-Pescara, Chieti, Italy.

Rahman, Bootan, Mathematics Unit, School of Science and Engineering, University of Kurdistan Hewlêr (UKH), Erbil, Kurdistan Region, Iraq.

Rosenberg, Ivo, Université de Montreal, QC, Canada, Montreal, Canada.

Sakonidis, Haralambos, Democritus University of Thrace, Komotiní (DUTH), Alexandroupolis Greece.

Scafati, Maria, Department of Mathematics "Guido Catelnuovo", University of Rome "La Sapienza", Rome, Italy.

Sessa, Salvatore, Department of Architecture, University of Naples Federico II, Naples, Italy.

Participation in the editorial board is renewed annually on the basis of the actual contribution, participation in the journal, and direction strategies defined by the chief editors. No communication is expected for the annual changes to the content of the board.

Ratio Mathematica Volume 39, 2020, pp. 7-32

A simple goodness-of-fit test for continuous conditional distributions

Peter Veazie[*]
Zhiqiu Ye[†]

Abstract

This paper presents a pragmatic specification test for conditional continuous distributions with uncensored data. We employ Monte Carlo (MC) experiments and the 2011 Medical Expenditure Panel Survey data to examine coverage and the power to discern deviations from the correct model specification in distribution and parameterization. We carry out MC experiments using 2000 runs for sample sizes 500 and 1000. The experiments show that the test has accurate coverage under correct specification, and that the test can discern deviations from the correct specification in both the distributional family and parameterization. The power increases as sample size increases. The empirical example shows the test's ability to identify specific distributions from other candidates using real cost data. Although the test can be used as a goodness-of-fit test for marginal distributions, it is particularly useful as an easy-to-use test for conditional continuous distributions, even those with one observation per pattern of explanatory variables.

Keywords: Goodness-of-fit test; model specification test; conditional continuous distributions.

2010 AMS subject classification: 62F03[‡]

---

[*] University of Rochester, Rochester New York, USA; peter_veazie@urmc.rochester.edu.
[†] University of Rochester, Rochester New York, USA; sophieye999@gmail.com.
[‡] Received on June 10th, 2020. Accepted on December 17th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.524. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

To determine whether a probability model is statistically adequate for representing a data generating process (DGP), it is common to test whether the model fits with a data set produced by that process. The investigation into the model specification of a conditional distribution is fundamental for methods such as Maximum Likelihood Estimation (MLE), which is consistent and asymptotically efficient only if the distribution is correctly specified (Amemiya, 1985). However, there are two key challenges for a general test of continuous conditional distribution models, if it is to be broadly adopted in applied sciences such as social and health sciences: First, is the sparse empirical information regarding the conditional distribution when patterns of the explanatory variables have few corresponding observations. Second, is the ease of use: many researchers do not have the background, time, or inclination to engage in complicated programming in order to implement a statistical test—to be useful to such researchers a test must be easily implemented.

Regarding sparse information, consider the data shown in Figure 1: although some data points appear close to each other, for most of the data there is no more than one observation at each value of x. Consequently, the empirical distribution of random variable *Y* conditioned on such a value for variable *X* is based on a trivial point mass. How then can we test a model of the conditional distribution of *Y* for such sparse data?

Regarding ease of use, existing tests for conditional distributions require more mathematical and computational skill than many applied researchers may have to make their implementation generally accepted. Some of these tests require the use of kernel or local polynomial functions with arbitrary smoothing parameters (Zheng, 2000, Fan et al., 2006). Others, such as the Conditional Kolmogorov Test, compare model and distribution functions additionally incorporating the empirical distribution functions of the conditioning set of variables (Andrews, 1997). Transformations to the unit interval have been applied to construct tests for goodness of fit such as the Rincon-Gallardo et al. test for multivariate normality (Rincon-Gallardo et al., 1979). However, their method is also technically difficult and computationally intense in general applications due to procedures involved in the transformation (O'Reilly and Quesenberry, 1973). Additionally, some are dependent on the order of the data being transformed (O'Reilly and Stephens, 1982); therefore, researchers may obtain variant test results if the same data were ordered differently. What is needed for the applied researcher who does not have the mathematical or programming skills to meaningfully implement complex algorithms is a simple pragmatic test. This paper presents a pragmatic general goodness-of-fit statistic for continuous conditional models using uncensored data.

In the next section, we introduce the goodness-of-fit test and the rationale behind it. We then evaluate the performance of the goodness-of-fit statistic in Section 3 using two groups of Monte Carlo experiments. The first group of experiments focuses on discerning deviations from correct specification in the distributional family; the second group focuses on discerning deviations in parameterization. We choose these investigations because they represent the two misspecification issues in estimating conditional probability models. In Section 4, we apply the goodness-of-fit test to the 2011 Medical Expenditures Panel Survey (MEPS) dataset, modeling three health care expenditure outcomes as functions of patient characteristics. Finally, in Section 5, we conclude our paper with a summary of the findings and discussions about the applications of the goodness-of-fit test. The Appendix provides the expected value of the statistic and the procedure for the calculation of the degrees of freedom for the test statistics, the data generating process for each Monte Carlo experiment, and the analyses modelling cost data from MEPS.

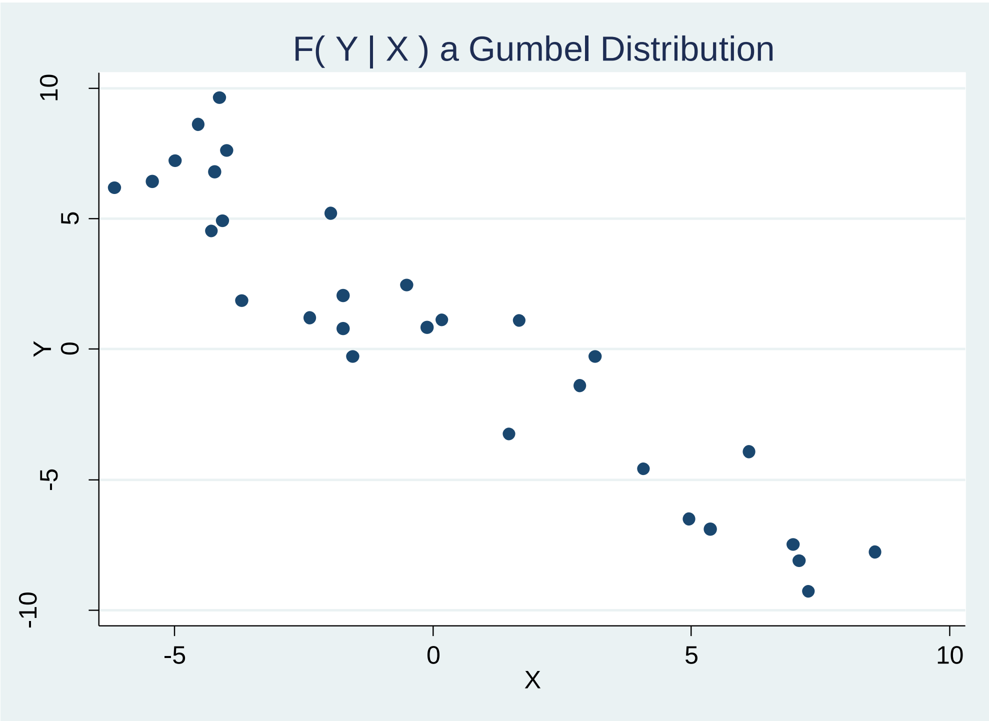


Figure 1. A typical conditional Gumbel distribution with sparse observations for each conditional value of observed X.

2 A proposed goodness-of-fit test

The Pearson Chi-square goodness-of-fit statistic is based on comparing the number of actual observations within each set of a partition of the random variable's range to the number of observations that would be expected to show up in those sets if the model correctly represents the DGP (Schervish, 1995). If the model is correct, then the expected number of observations is the expected number for the DGP; consequently, the observed and predicted number in each set should be different merely by random variation.

The Chi-square goodness-of-fit statistic for continuous distributions is created by partitioning the range of a continuous random variable Y into K regions. Denote each region $k\in\{1, 2, \ldots K\}$ as

$$R_k = \{y: y_{k-1} < y \leq y_k\},$$

the number of observations with values of y in region R_k as N_k, and the total sample size as N. The probability of an observation with y in region R_k is then

$$P_k = \int_{y_{k-1}}^{y_k} f_Y(y)\,dy,$$

in which $f_Y(y)$ is the probability density for y associated with a cumulative distribution function (CDF) $F_Y(y)$. The Chi-square statistic is defined as

$$C = \sum_k \frac{(N_k - N\cdot P_k)^2}{N\cdot P_k}.$$

The corresponding sample statistic is

$$C_N = \sum_k \frac{\left(N_k - N\cdot \hat{P}_{k,N}\right)^2}{N\cdot \hat{P}_{k,N}}$$

in which $\hat{P}_{k,N}$ is a consistent estimator of P_k. If $F_Y(y;\theta)$ accurately represents the data generating process, C_N converges to C with increasing N and the corresponding asymptotic distribution of C_N is a Chi-square with degrees of freedom equal to the number of groups in the partition minus the number of estimated parameters plus one (Schervish, 1995).

If we are interested in a model of the conditional distribution $F(y|x;\phi)$, the preceding statistic is not generally applicable because N_k can contain insufficient observations to inform the conditional distribution. Indeed, with x

containing precisely measured continuous variables, there may be only one value y for some observed x values (see Figure 1 as an example). However, we can take advantage of the probability integral transform and consequent fact that the CDF of a continuous random variable is itself a random variable with a uniform distribution on the unit interval. Because the uniform distribution is the same regardless of underlying CDF, a set of random variables from independent observations with different conditional distributions can all be converted by their CDFs to the same uniform distribution. We can use this fact to construct a test of the conditional distribution; even if each observation has a different conditioning value (i.e. the data in Figure 1 will pose no problem for this test).

Because the CDF for each random variable has a uniform distribution, the CDF values of sample results from a correctly specified model for each random variable will produce a single realization from a uniform distribution. Therefore, the full sample results should together provide a histogram that deviates only by chance from a uniform distribution. We can use a Pearson Chi-square type statistic applied to the uniform distribution to test of the specification for the conditional distributions.

The process is quite simple. For each observation i we have a model specification for the distribution $F(y_i|x_i)$ and therefore can obtain from the estimated model the sample quantity $u_i = F(y_i|x_i)$ for which (x_i, y_i) are the observed values for observation i. The random variable underlying u_i has a uniform distribution on the unit interval if F is correctly specified. We can construct a goodness-of-fit test by partitioning the unit interval into K subintervals defined by equally spaced boundary points, which for $K = 10$ is

$$R_k = \left\{u: \tfrac{k-1}{10} < u \leq \tfrac{k}{10}\right\}, \text{ s.t. } k \in \{1, 2, \ldots 10\}.$$

The statistic is then

$$U = \sum_k \frac{(N_k - N \cdot P_k)^2}{N \cdot P_k}$$

for which N is the total sample size, and N_k is the observed number of u values in the R_k interval. This statistic can alternatively be written as

$$U = N \cdot \sum_k \frac{\left(\hat{P}_k - P_k\right)^2}{P_k}.$$

in which $\hat{P}_k$ is the observed proportion in interval R_k. Because the statistic is based on a partition of the uniform into K equal sized intervals, $P_k = {}^{1}\!/_{K}$; therefore,

$$U = K \cdot N \cdot \sum (\hat{P}_k - \tfrac{1}{K})^2 .$$

As shown in the Appendix, the expected value of U, which is the degrees of freedom for its approximating Chi-square distribution, is equal to the degrees of freedom for the usual Pearson Chi-square test (i.e. $K - 1$) minus a factor due to the estimation of model parameters.

Since P_k is known, which in the case of $K = 10$ intervals is 0.1, we can simply state the statistic for $K = 10$ as

$$U = 10 \cdot N \cdot \sum_k \left(\hat{P}_k - 0.1\right)^2.$$

The selection of $K = 10$ is arbitrary, as it is with the Hosmer-Lemeshow test for logistic regression (Hosmer and Lemeshow, 1980). For other values of K, the degrees of freedom can be directly estimated as shown in the Appendix or determined by Monte Carlo simulation (see Box 2).

The U statistic has a distribution proportional to the sum of gamma random variables with different parameters. Specifically, denoting $\hat{P}_k - \tfrac{1}{K}$ as z_k, as shown in the appendix z_k is asymptotically normally distributed with mean 0 and variance σ_k^2. Consequently, the ratio of z_k squared to σ_k^2 has an asymptotic Chi-square distribution with degrees of freedom 1, which is a Gamma distribution with parameters 0.5 and 2 (i.e. $\Gamma(0.5, 2)$). Therefore, z_k^2 has a distribution $\sigma_k^2 \cdot \Gamma(0.5, 2)$, which is $\Gamma(0.5, 2 \cdot \sigma_k^2)$. U is therefore proportional to the sum of K differently scaled gamma random variables. Moschopoulos shows that sum of such variates can be express as a gamma series in which the series coefficients can be recursively determined (Moschopoulos, 1985). The use of this recursive coefficient determination and gamma series is overly complex for the practical application of this statistic among many applied researchers. However, ease of use is the purpose of this goodness-of-fit statistic. Fortunately, the Monte Carlo experiments presented below indicate that for a correct specification the statistic is approximately Chi-square in distribution with degrees of freedom 7.5 when $K = 10$ and calculated as shown in the Appendix or as shown in Box 2 if K is not 10.

3 Simulation experiments

3.1 Methods

We investigated finite sample performance of the proposed statistic using Monte Carlo experiments of conditional Normal, Gumbel, Gamma, and Weibull models, each applied to data generating processes based on the same set of distributions. The first set of experiments comprised a total of sixteen model/DGP comparisons. We evaluated each model/DGP pair for sample sizes 500 and 1000, each using 2000 Monte Carlo samples from the DGP (see Appendix Table A1 for parameter specifications). We inspected rejection rates for significance levels spanning between 0 and 0.2 for each comparison. For each correct model/DGP pair (i.e. Normal/Normal, Gumbel/Gumbel, Gamma/Gamma, and Weibull/Weibull), the plot of the empirical cumulative distribution function (eCDF) of the calculated p values, across the 2000 MC samples, should approximately match the significance level (i.e. this plot should be approximately a straight line). For example, the use of a significance level of 0.01 should reject the model for approximately 1 per cent of the 2000 samples; using a significant level of 0.05 should reject approximately 5 per cent of the samples; and a 0.1 significance level should result in approximately 10 per cent rejections. For mismatched pairs (e.g. Weibull/Gumbel), if the fit test is useful it should produce rejection rates that are higher than the significance levels and increase with sample size; consequently, the eCDF of the test's p-value should be above the significance level.

The second set of experiments compared models in which parameters are specified as linear in conditioning variables to the DGP having the same distributional family but with parameters quadratic in the conditioning variables (see Appendix Table A2 for parameter specifications). For the normal distribution, we estimated models with homoscedasticity and heteroscedasticity. In the case of heteroscedasticity both the mean and variance were generated as quadratic in X in the DGP, but they were modeled as linear in the misspecified model. Similarly, we carried out experiments for sample sizes 500 and 1000. These experiments provide evidence regarding whether the test can identify deviations in parameterization as well as distributional family. In the Monte Carlo experiments reported below, we applied the steps presented in Box 1 to obtain p-values for each of 2000 data sets generated for each model/DGP being considered. We calculated both a p-value using degrees of freedom equal to 7.5 and also using the mean of the 2000 calculated U values for each DGP considered when using the correct model (remember that the degrees of freedom are associated with the distribution of U given the model is correct).

3.2 Results

Because we tested continuous conditional distributions, it is difficult to see the differences between the model and DGP for all patterns of explanatory variables. However, Table 1 shows the probability density functions for the true DGP (in the solid line) and the estimation model (the dotted line, using the average parameter values across the 2000 estimated models) evaluated at the mean of *X*. This gives some sense of the differences between the distributions being tested in the first set of experiments; however, the deviation of the model from the underlying distribution that drives larger values of U may be from other regions of the conditioning set than at the mean of *X*.

BOX 1. How to calculate U and its p-value using K = 10

Step 1. For a candidate model $F(y_i \mid x_i \; ; \theta)$, estimate the parameters, obtaining $\hat{\theta}$.

Step 2. Calculate the CDF value of $u_i = F(y_i \mid x_i \; ; \hat{\theta})$ for each observation (y_i , x_i) in the data.

Step 3. Calculate the proportion ($\hat{P}_k$) of u_i in each of the ten intervals R_k for $k \in \{1, 2, \ldots 10\}$.

Step 4. Calculate the statistic U using the equation

$$U = 10 * N \cdot \sum_k \left(\hat{P}_k - 0.1\right)^2.$$

Step 5. Calculate the p-value as the upper tail area of a Chi-square distribution with degrees of freedom set to 7.5 or set to the estimated value determined by the equations presented in the Appendix or the model-specific Monte Carlo determined empirical degrees of freedom (see Box 2 for the algorithm).

Tables 2 and 3 present the eCDFs of the statistic's p-values for each indicated model applied to the indicated DGP plotted for significance levels up to 0.2. Table 2 presents results for sample sizes of 500; Table 3 presents results for sample sizes of 1000. The thin straight lines show the points where the eCDFs would be if it corresponded to the significance level. The thick dark lines (or curves) are the eCDFs associated with p values based on degrees of freedom set to 7.5. The thick light lines (or curves) are the eCDFs associated with the Monte Carlo based empirical degrees of freedom. We determined the 7.5 degrees of freedom approximation by the average of the four empirical degrees of freedom across the DGPs using sample sizes of 1000. We also ran Monte Carlo experiments for correct model specifications using 10 correlated explanatory variables (results not presented); these experiments showed that the empirical degrees of freedom remained around 7.5 in multivariable models. Specifically, the means of the *U* statistics, and therefore the degrees of freedom, in these experiments for the Normal, Gamma, Weibull, and Gumbel were 7.52, 7.32, 7.54, and 7.27 respectively.

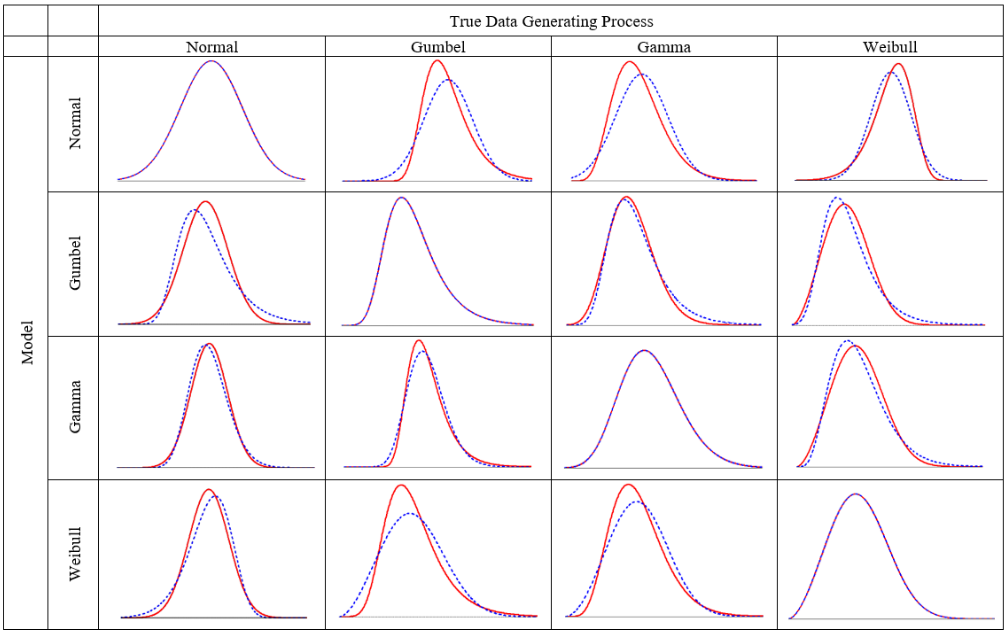


Table 1. Probability density functions of the true data generating process (solid curve) and the estimated model (dashed curve) evaluated at X=0 for the Monte Carlo simulations.

The figures on the diagonals of Tables 2 and 3 show the coverage of the test when the model is correctly specified. The results fell along the line representing accurate coverage: the eCDF corresponds to the significance level. Not surprisingly, the empirical degrees of freedom (the thick lighter line) were more accurate than using the approximate degrees of freedom of 7.5; however, the differences were slight, particularly up to the 0.1 significance level.

The off-diagonal figures in Tables 2 and 3 show the rejection rate for the test of misspecified models across significance levels. The test was sufficiently powerful for some of the model/DGP combinations to reject the model for all 2000 samples at all significance levels greater than 0.001. Results for these combinations are simply indicated by the phrase 'ALL DATA SETS REJECTED AT SIGNIFICANCE LEVEL 0.001'. Not surprisingly, comparing Table 2 to Table 3, the curve has a greater departure from the straight line in Table 3; it is evident that the power of the test increases with sample size. It is also clear that using the approximate 7.5

degrees of freedom provide similar results to that of using the Monte Carlo determined empirical degrees of freedom.

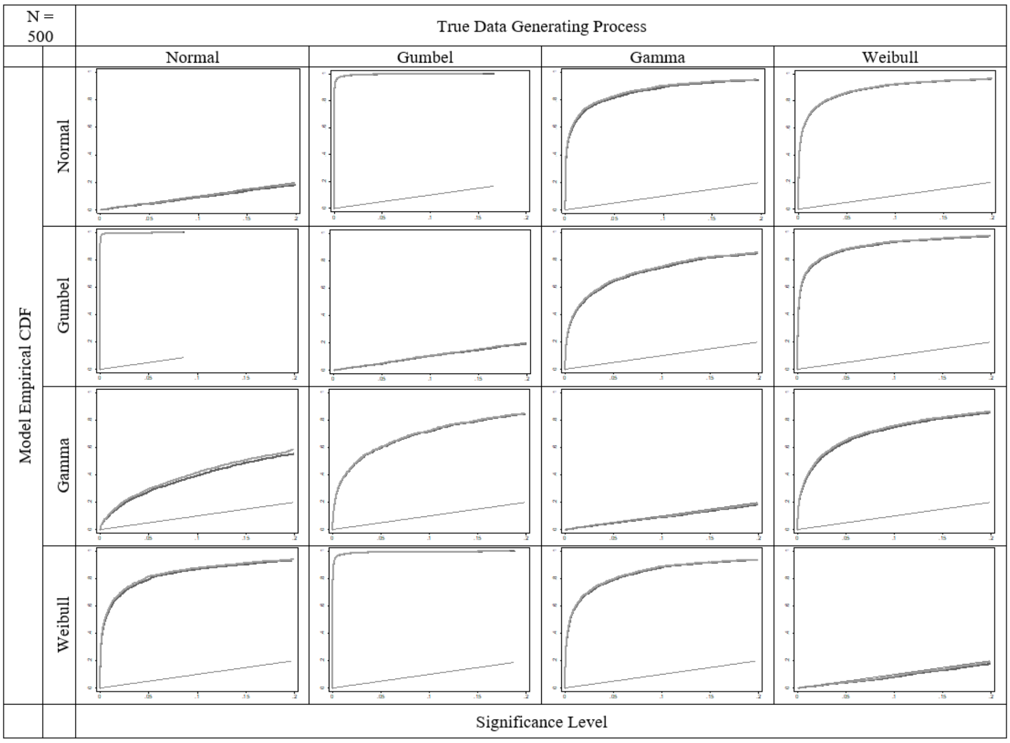


Table 2. Monte Carlo Simulation: Empirical CDFs of Experiments on Distribution Specifications (N=500).

Table 4 presents results for the second set of experiments, which tested deviations from correct specification in the parameterization. The upper two rows show results for sample sizes of 500; the lower two rows show results for sample sizes of 1000. Similar to the first set of experiments, results showed accurate coverage for the test when the model was correctly specified and the ability to discern deviations from correct specification in parameterization. As the sample size went up, the power of the test to discern such deviations increased. The approximate 7.5 degrees of freedom yields results that were similar to the Monte Carlo calculated empirical degrees of freedom.

4 Example

To present an example with real data, we used a random sample of 2000 individuals from the Household Component of the 2011 Medical Expenditure Panel Survey data file (MEPS). As one of the largest national health survey, MEPS has been widely used to study the patterns of health care access, utilization and expenditures in the United States (Cohen et al., 2009). We modeled each of the three outcomes – annual total health care expenditure, total office-based visits expenditure, and total dental care expenditure – as a function of individual demographics, socioeconomic status, self-rated health status, common chronic conditions, presence of usual source of care provider, and health insurance coverage. These covariates were selected in accordance with prior studies focusing on modeling health care costs using MEPS survey data (Fenton et al., 2012, Fleishman and Cohen, 2010).

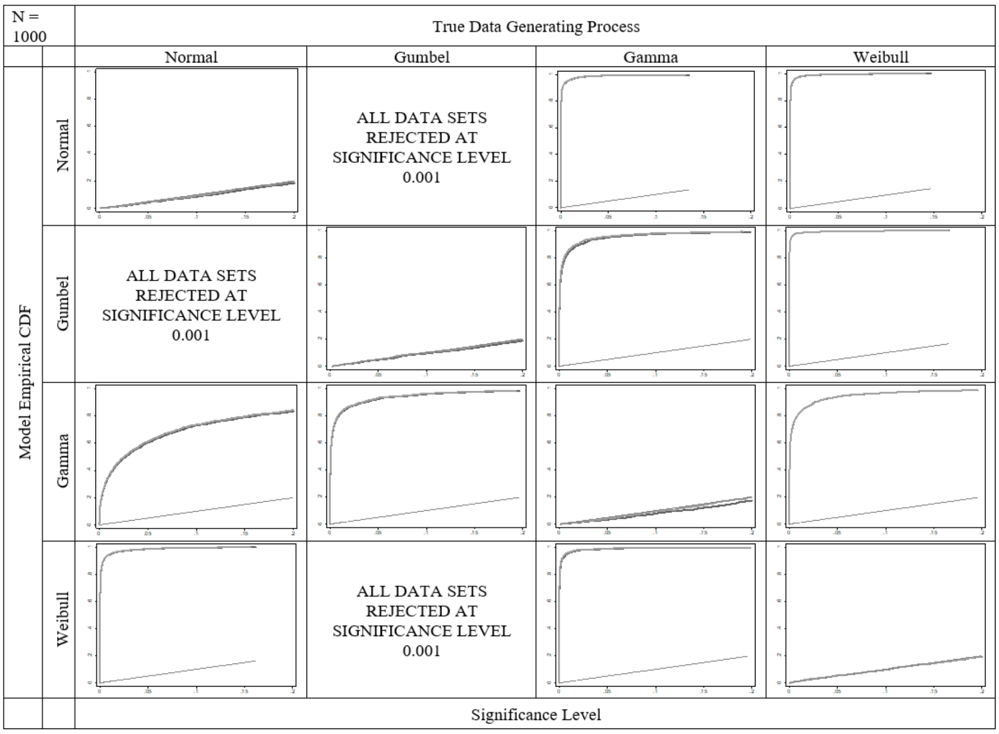


Table 3. Monte Carlo Simulation: Empirical CDFs of Experiments on Distribution Specifications (N=1000).

For each model, we included all individuals who reported an expense on the outcome of interest and took the log of the expenditure as the dependent variable. There were 1527 and 1215 individuals reporting expenses on health care and office-based services, which represented 76.4% and 60.8% of the total sample, respectively. Much fewer individuals reported any expenses on dental care (N= 724, 36.2%). Appendix Table A3 presents the descriptive statistics and distribution of the outcome variables and the covariates that we employed in the model.

We used Pregibon's link test (Pregibon, 1980) to identify a statistically adequate specification of the explanatory variables for each model. We then computed U to test the hypothesis that the specified distribution was correct. This allows us to use the test to focus on testing deviations in the distributional family. We calculated the p-value based on the approximate degrees of freedom of 7.5 and the empirical degrees of freedom calculated from the parameter estimates of the specified model, based on 500 Monte Carlo samples. The algorithm for computing the empirical degrees of freedom is shown in Box 2. Table 5 presents the results from the empirical example for the three health care expenditure outcomes. The test clearly discerns the goodness-of-fit performance of different distributions. Results for the model of the logarithm of total health care expenditure strongly rejected the hypothesis

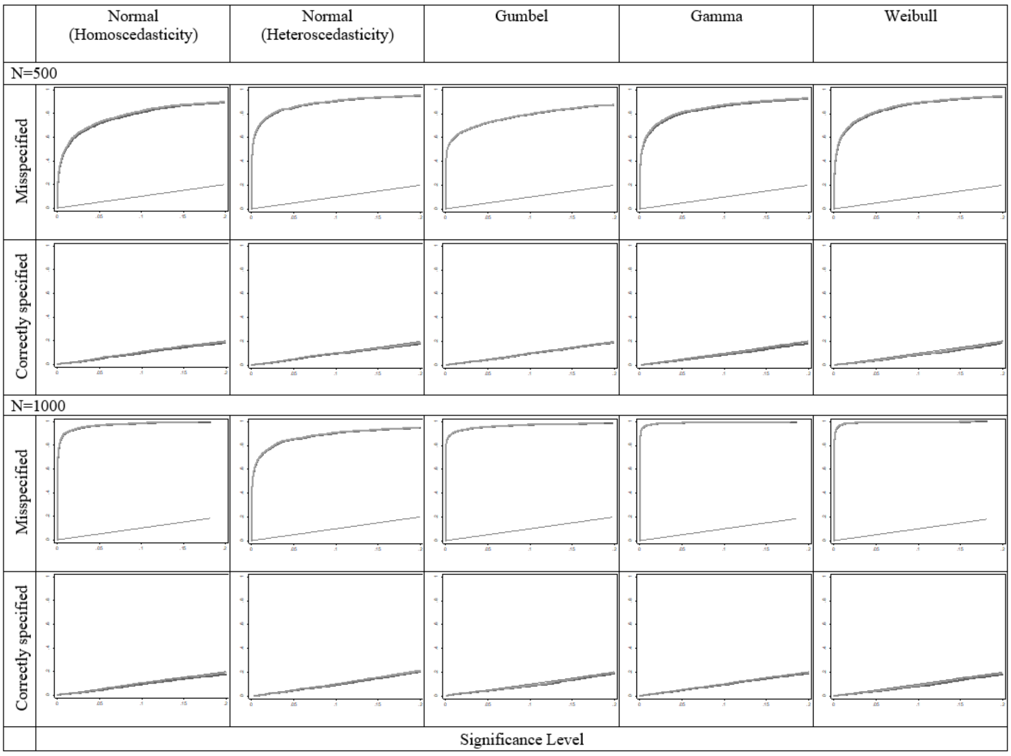


Table 4. Monte Carlo Simulation: Empirical CDFs of Experiments on Parameter Specifications.

that the conditional distribution follows a Gamma, Weibull or Gumbel distribution (*U* ranges from 21.985 to 116.578, p-value < 0.001 for all), and unequivocally failed to reject the hypothesis for normal (*U* = 5.304, p-value = 0.676 with approximate degrees of freedom of 7.5). For the model of office-based visits expenditure, we strongly rejected the hypotheses for the Gumbel and Weibull distribution (U equals to 59.626 and 33.776, respectively, p-value

< 0.001 for both) and fail to reject the Normal (U=12.467, p-value = 0.107) or Gamma (U = 8.897, p-value = 0.305). For the model of dental care expenditure, we rejected all distributions except for the Gumbel (U = 14.333, p-value = 0.058 with degrees of freedom of 7.5). Figures A1-A3, in the Appendix, show the histograms of the residuals obtained from these models, standardized by the estimated standard deviations. Figure A1 shows the symmetry expected of a Normal distribution, which was not rejected by the test that unambiguously rejected the other distributions. Figure A2 shows a right-skewedness characteristic of a Gamma distribution (Model 1), but it is insufficiently skewed to reject the Normal at a significance level of 0.05. However, U is smaller in the Gamma indicating a better fit to the data. Under certain circumstances (i.e. shape parameter sufficiently large, >15), the Gamma distribution is approximately a Normal distribution (Rothschild and Logothetis, 1986). In this real-data example, the estimated shape parameter equaled to 35 in the model assuming Gamma distribution. It is therefore not surprising the test did not reject either the Gamma or the Normal distributions. Figure A3 demonstrates the clear right-skewedness of the residual from the model of dental care expenditure, which is expected of a Gumbel distribution. The calculated Monte Carlo empirical degrees of freedoms were approximately 7.5 for all three outcomes and therefore yielded similar results. As there were 21 variables in the empirical model, these results again show that the degrees of freedom for the statistic distribution based on 10 categories is approximately 7.5 in multivariable models.

	Total Health Care Expenditure (N=1527)				***Total Office-Based Visits Expenditure (N=1215)***				***Total Dental Care Expenditure (N=724)***			
	U	p-value (DF=7.5)	p-value (DF=edf)	edf	U	p-value (DF=7.5)	p-value (DF=edf)	edf	U	p-value (DF=7.5)	p-value (DF=edf)	edf
Normal	5.30	0.676	0.686	7.60	12.47	0.107	0.099	7.30	89.27	<0.001	<0.001	7.19
Gumbel	116.58	<0.001	<0.001	7.79	59.63	<0.001	<0.001	7.71	14.33	0.058	0.054	7.36
Gamma	23.76	0.002	0.002	7.40	8.90	0.305	0.307	7.53	43.36	<0.001	<0.001	7.38
Weibull	21.99	0.004	0.004	7.59	33.78	<0.001	<0.001	7.41	90.94	<0.001	<0.001	7.57

Note. DF, degrees of freedom; edf, empirical degrees of freedom.

Table 5. Empirical Example: Goodness-of-Fit Tests on Conditional Probability Models for Log-Transformed Health Expenditures from MEPS.

5 Conclusion

In this paper, we presented a simple specification test for conditional continuous distributions using uncensored data (see Box 1). We showed, using simulation experiments, that the test has accurate coverage under correct specification, and that the test can discern deviations from correct specification in both the distributional family as well as parameterization. The empirical example shows its ability to distinguish specific distributions from other candidates using real data.

The results of our analysis indicate that U is approximately distributed Chi-square with degrees of freedom 7.5. We also provide a Monte Carlo method for an empirical determination of degrees of freedom in Box 2 and a direct estimator in the Appendix should the researcher not wish to use the approximating 7.5, for example when the p-value using the approximating 7.5 degrees of freedom is close to the test's designated significance level. However, comparing the empirical degrees of freedom to 7.5 across all Monte Carlo experiments and real-data analyses of our study, the differences were slight and not likely to impact inferences. If a researcher does not wish to approximate the distribution using a Chi-square, a p-value based on the Monte Carlo distribution of statistic values generated in the process of Box 2 can be used as a parametric bootstrap test (Davison et al., 2003).

Because the test discerns deviations in parameterization as well as the distributional family, an extra step is required to investigate the distributional

BOX 2. How to calculate the empirical degrees of freedom

Step 1. Obtain the parameter estimates predicted from the estimated model ($\hat{\theta}$).

Step 2. Generate outcome values as random draws from the distribution defined by the estimated parameters $\hat{y}_i \sim F(Y \mid X = x_i\,;\,\hat{\theta})$ for all x_i in the data.

Step 3. Re-estimate the model using the generated outcomes.

Step 4. Obtain the predicted parameter estimates ($\tilde{\theta}$) from using the 'correctly' specified model in Step 3.

Step 5. Calculate the value of $\hat{u}_i = F(\hat{Y}_i | X_i; \tilde{\theta})$ for each observation.

Step 6. Calculate U.

Step 7. Repeat the steps 2 through 6 multiple times (e.g. we repeated 500 times in the empirical example), saving the statistic values.

Step 8. Set the degrees of freedom to the mean of the calculated U values.

family alone. Specifically, the researcher should engage in standard tests to identify the best parameter specification within each proposed model (e.g. we used Pregibon's link test in the preceding example). Using the best within-family model specification, the test will then primarily be identifying deviations in the distributional family.

It is important to note that our results using multiple explanatory variables in the models indicate the degrees of freedom for the statistic's distribution is not a function of the number of estimated parameters. This is different from the direct application of the Pearson Chi-square test to distributions with multiple parameters in which the degrees of freedom depend on the number of parameters m. This is an advantage since the degrees of freedom in the latter case is typically $K–m−1$, which implies m must be less than $K −1$ for those applications (Schervish, 1995): our test does not have this constraint.

Although our test can be used as a goodness-of-fit test for marginal distributions, it is particularly useful as an easy-to-use model fit test of continuous conditional distributions for uncensored data, particularly in the case of few observations, indeed even one observation per pattern of explanatory variables, such as a time-series.

References

AMEMIYA, T. 1985. *Advanced Econometrics*, Cambridge, MA, Harvard University Press.

ANDREWS, D. W. K. 1997. A Conditional Kolmogorov Test. *Econometrica,* 65**,** 1097-1128.

COHEN, J. W., COHEN, S. B. & BANTHIN, J. S. 2009. The medical expenditure panel survey: a national information resource to support healthcare cost research and inform policy and practice. *Med. Care.,* 47**,** S44-50.

DAVISON, A. C., HINKLEY, D. V. & YOUNG, G. A. 2003. Recent developments in bootstrap methodology. *Statistical Science,* 18**,** 141-157.

FAN, Y. Q., LI, Q. & MIN, I. 2006. A Nonparametric bootstrap test of conditional distributions. *Economet. Theor.,* 22**,** 587-613.

FENTON, J. J., JERANT, A. F., BERTAKIS, K. D. & FRANKS, P. 2012. The cost of satisfaction: a national study of patient satisfaction, health care utilization, expenditures, and mortality. *Arch. Intern. Med.,* 172**,** 405-11.

FLEISHMAN, J. A. & COHEN, J. W. 2010. Using Information on Clinical Conditions to Predict High-Cost Patients. *Health. Serv. Res.,* 45**,** 532-552.

HOSMER, D. W. & LEMESHOW, S. 1980. A goodness-of-fit test for the multiple logistic regression model. *Communications in Statistics Part A-Theory and Methods,* 10**,** 1043-1069.

MOSCHOPOULOS, P. G. 1985. The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics,* 37**,** 541-544.

O'REILLY, F. J. & QUESENBERRY, C. P. 1973. The Conditional Probability Integral Transformation and Applications to Obtain Composite Chi-Square Goodness-of-Fit Tests. *Ann. Statist.,* 1**,** 74-83.

O'REILLY, F. J. & STEPHENS, M. A. 1982. Characterizations and Goodness of Fit Tests. *J. Roy. Statist. Soc. Ser. B,* 44**,** 353-360.

PREGIBON, D. 1980. Goodness of Link Tests for Generalized Linear Models. *J. Roy. Statist. Soc. Ser. C,* 29**,** 15-24.

RINCON-GALLARDO, S., QUESENBERRY, C. P. & O'REILLY, F. J. 1979. Conditional Probability Integral Transformations and Goodness-of-Fit Tests for Multivariate Normal Distributions. *Ann. Statist.,* 7**,** 1052-1057.

ROTHSCHILD, V. & LOGOTHETIS, N. 1986. *Probability distributions*, Wiley.

SCHERVISH, M. J. 1995. *Theory of Statistics,* New York, Springer-Verlag.

ZHENG, J. X. 2000. A consistent test of conditional parametric distributions *Economet. Theor.,* 16**,** 667-691.

Appendix

A1 The Expected Value of the U-Statistic

The expected value of U is the expected value associated with the distribution of the standard Pearson Chi-square goodness-of-fit statistic minus a factor due to estimating the parameters of the model. In this appendix we provide the determination of the expected value, and we provide an estimator for the adjustment factor and thereby an estimator of the expected value for the proposed statistic.

The expected value of U is proportional to the sum of expected values across the K equal-length regions of the partition of the unit interval being considered:

$$\begin{aligned} E(U) &= E[K \cdot N \cdot \sum_k (\hat{P}_k - P_k)^2] \\ &= K \cdot N \cdot \sum_k E[(\hat{P}_k - P_k)^2] \end{aligned}$$

The expected values under the summation sign on the right-hand side of this equation are variances. This is seen by denoting an indicator of whether observation i falls in region k as

$$I_{k,i} = \begin{cases} 1 & u_i \in ((k-1)\cdot 0.1, k \cdot 0.1) \\ 0 & \mathit{Otherwise} \end{cases}$$

and noting that the expected value of the estimated proportion in category k is

$$\begin{aligned} E[\hat{P}_k] &= \int E[\hat{P}_k \mid \hat{\theta}] dF(\hat{\theta}) \\ &= \int E[\tfrac{1}{N} \sum_{i=1}^{N} I_{k,i} \mid \hat{\theta}] dF(\hat{\theta}) \\ &= \tfrac{1}{N} \cdot \int \sum_{i=1}^{N} E[I_{k,i} \mid \hat{\theta}] dF(\hat{\theta}) \\ &= \tfrac{1}{N} \cdot \int \sum_{i=1}^{N} P_{k,i}(\hat{\theta}) dF(\hat{\theta}) \\ &= \int P_k(\hat{\theta}) dF(\hat{\theta}), \text{ for } P_{k,i} = P_k \text{ for all } i \\ &= E(P_k(\hat{\theta})) \end{aligned}$$

To determine $E(P_k(\hat{\theta}))$, consider a first order Taylor series approximation around the true value θ

$$P_k(\hat{\theta}) = P_k(\theta) + \frac{\partial P_k}{\partial \theta'} \cdot (\hat{\theta} - \theta),$$

which yields

$$\sqrt{N} \cdot (P_k(\hat{\theta}) - P_k(\theta)) = \frac{\partial P_k}{\partial \theta'} \cdot \sqrt{N} \cdot (\hat{\theta} - \theta).$$

For an estimator, such as the maximum likelihood estimator, for which $\sqrt{N} \cdot (\hat{\theta} - \theta)$ converges to a normal distribution $N(0,\Sigma)$ by a central limit theorem, the left-hand side converges in distribution to a normal as well:

$$\sqrt{N} \cdot (P_k(\hat{\theta}) - P_k(\theta)) \xrightarrow{d} N(0, \frac{\partial P_k}{\partial \theta'} \Sigma \frac{\partial P_k}{\partial \theta}).$$

Therefore, $P_k(\hat{\theta})$ has an asymptotic distribution with expected value of $E[P_k(\hat{\theta})] = P_k(\theta)$ and variance of $V[P_k(\hat{\theta})] = \frac{1}{N} \cdot \left(\frac{\partial P_k}{\partial \theta'} \Sigma \frac{\partial P_k}{\partial \theta} \right)$. Consequently, since $E[P_k(\hat{\theta})] = P_k(\theta)$,

$$E[(\hat{P}_k - P_k)^2] = V[\hat{P}_k].$$

The expected value of the U is then proportional to the sum of variances:

$$E[K \cdot N \cdot \sum_k (\hat{P}_k - P_k)^2] = K \cdot N \cdot \sum_k V[\hat{P}_k] \cdot$$

The variance terms under the summation sign on the right-hand side are

$$\begin{aligned}
V[\hat{P}_k] &= \int V[\hat{P}_k \mid \hat{\theta}] dF(\hat{\theta}) \\
&= \int V[\tfrac{1}{N} \cdot \sum_{i=1}^{N} I_{k,i} \mid \hat{\theta}] dF(\hat{\theta}) \\
&= \tfrac{1}{N^2} \cdot \int \sum_{i=1}^{N} V[I_{k,i} \mid \hat{\theta}] dF(\hat{\theta}), \text{ for independent observations} \\
&= \tfrac{1}{N} \cdot \int P_k(\hat{\theta})(1 - P_k(\hat{\theta})) dF(\hat{\theta}) \\
&= \tfrac{1}{N} \cdot \int (P_k(\hat{\theta}) - P_k(\hat{\theta})^2) dF(\hat{\theta}) \\
&= \tfrac{1}{N} \cdot [E(P_k(\hat{\theta})) - E(P_k(\hat{\theta}))^2 - V(P_k(\hat{\theta}))) \\
&= \tfrac{1}{N} \cdot [P_k(\theta) - P_k(\theta)^2 - V(P_k(\hat{\theta})]
\end{aligned}$$

Therefore,

$$E[K \cdot N \cdot \sum_k (\hat{P}_k - P_k)^2] = K \cdot N \cdot \sum_k \tfrac{1}{N} \cdot [(P_k - P_k^2) - V(P_k(\hat{\theta}))]$$

$$= K \cdot N \cdot \sum_k \tfrac{1}{N} \cdot [(\frac{1}{K} - \frac{1}{K^2}) - V(P_k(\hat{\theta}))]$$

$$= (K-1) - K \cdot \sum_k V(P_k(\hat{\theta}))$$

The expected value of U is the degrees of freedom for a common Pearson Chi-Square test statistic (i.e. $K - 1$) minus a factor due to estimation of the distribution parameters. For $K = 10$, the expected value of U is then $9 - 10 \cdot \sum_k V(P_k(\hat{\theta}))$.

A2 Estimation of the Shrinkage Factor

The variance terms in the shrinkage factor can be estimated by using consistent estimators for the derivatives $\frac{\partial P_k}{\partial \theta}$ and the covariance matrix Σ. The derivative of P_k is determined by noting that

$$P_k = \int [F(y_k^*(x) \mid x;\theta) - F(y_{k-1}^*(x) \mid x;\theta)] dF_x(x) ,$$

for which y_k^* are the critical values

$$y_k^*(x) = F^{-1}(\frac{k}{K} \mid x) .$$

Therefore, assuming we can interchange the order of integration and differentiation,

$$\frac{\partial P_k}{\partial \theta} = \int \frac{\partial}{\partial \theta} F(y_k^*(x) \mid x;\theta) dF_x(x) - \int \frac{\partial}{\partial \theta} F(y_{k-1}^*(x) \mid x;\theta) dF_x(x) .$$

Estimating the integrals on the right-hand side of the equation by sample means yields the estimator

$$\frac{\partial P_k}{\partial \theta} = \frac{1}{N} \cdot \sum_{i=1}^{N} \frac{\partial}{\partial \theta} F(y_k^*(x_i) \mid x_i;\hat{\theta}) - \frac{1}{N} \cdot \sum_{i=1}^{N} \frac{\partial}{\partial \theta} F(y_{k-1}^*(x_i) \mid x_i;\hat{\theta}) .$$

The estimator for the variances in the shrinkage factor is therefore

$$\hat{V}[P_k(\hat{\theta})] = \frac{1}{N} \cdot \left(\frac{\partial P_k}{\partial \theta'} \cdot \hat{\Sigma} \cdot \frac{\partial P_k}{\partial \theta} \right).$$

For the maximum likelihood estimator, note that the scaled deviation of the estimator converges in distribution to a normal:

$$\sqrt{N}\cdot(\hat{\theta}-\theta)\xrightarrow{d}N(0,[-E(\tfrac{1}{N}H(\theta))]^{-1}),$$

for H denoting the matrix of second derivatives of the log-likelihood with respect to the parameters. Therefore,

$$\begin{aligned}\Sigma &= [-E(\tfrac{1}{N}H(\theta))]^{-1}\\ &= N\cdot[-E(H(\theta))]^{-1}\end{aligned}$$

Using the sample mean for the expectation of the Hessian, evaluated at the estimated parameter values, yields the estimator

$$\hat{\Sigma}=N^2\cdot[-H(\hat{\theta})]^{-1}.$$

The estimated variance of $P_k(\hat{\theta})$ is then

$$\hat{V}[P_k(\hat{\theta})]=N\cdot\left(\frac{\partial P_k}{\partial\theta'}\cdot[-H(\hat{\theta})]^{-1}\cdot\frac{\partial P_k}{\partial\theta}\right).$$

For example, consider the Weibull distribution specified in Table A1. The Weibull CDF is

$$F(y\,|\,x)=1-e^{-(e^{a_0+a_1\cdot x})\cdot y^{e^{b_0+b_1\cdot x}}}.$$

The derivatives with respect to the parameters are

$$\begin{aligned}\frac{\partial F}{\partial a_0}&=D\\ \frac{\partial F}{\partial a_1}&=D\cdot x\\ \frac{\partial F}{\partial b_0}&=D\cdot(e^{b_0+b_1\cdot x})\cdot\ln(y)\\ \frac{\partial F}{\partial b_1}&=D\cdot(e^{b_0+b_1\cdot x})\cdot\ln(y)\cdot x\end{aligned}$$

where,

$$D=y^{e^{b_0+b_1\cdot x}}\cdot e^{-y^{e^{b_0+b_1\cdot x}}\cdot e^{a_0+a_1\cdot x}}\cdot e^{a_0+a_1\cdot x}.$$

Evaluating each of these derivatives and each observation in the sample $i \in \{1, \ldots N\}$ at the estimated parameter values, data values x_i, and the corresponding critical values $y_0^*(x_i)$, and $y_k^*(x_i)$ for each $k \in \{1,\ldots 10\}$ creates variables for which the sample means can be used to determine $\frac{\partial P_k}{\partial \theta}$. These estimated derivatives combined with the estimated parameter covariance matrix $\hat{\Sigma}$ provide the information to calculate the shrinkage factor as shown above.

Table A0 presents the means of the estimated expected value of *U* using the above equations and means of the calculated *U* values across 100,000 data sets of sample sizes 100, 1000, and 10,000. The mean estimated E(u) was very similar to the mean of *U* values, rounding to 7.37 for each. An alternative for estimating the expected value of *U* (i.e. degrees of freedom for an approximating Chi square distribution) is the Monte Carlo method shown in Box 2 of the main text.

Sample Size	Mean Estimated E(u)	Mean U-statistic
100	7.367	7.369
1000	7.373	7.367
10000	7.374	7.374

Table A0. Mean estimated E(u) and mean *U* across 100,000 samples.

A3 Additional Tables and Figures

		True Data Generating Process			
		Normal	Gumbel	Gamma	Weibull
Model Empirical CDF	Normal	$\mu = e^{(2+0.1x)}$ $\sigma^2 = \mu^2$	Location $\mu = 10 + x$ Scale $\beta = e^{(0.1x)}$	Shape $\alpha = e^{(2+0.2x)}$ Scale $\beta = e^{(-0.2+0.1x)}$	Shape $\alpha = e^{(2.5+0.1x)}$ Scale $\beta = e^{(0.1+0.2x)}$
	Gumbel	$\mu = 10 + x$ $\sigma^2 = e^{(0.5+0.1x)}$	Location $\mu = 5 + x$ Scale $\beta = e^{(0.6+0.1x)}$	Shape $\alpha = e^{(3+0.2x)}$ Scale $\beta = e^{(-0.6+0.1x)}$	Shape $\alpha = e^{(1+0.1x)}$ Scale $\beta = e^{(0.1+0.2x)}$
	Gamma	$\mu = e^{(1.03+0.001x)}$ $\sigma^2 = 0.04\mu^2$	Location $\mu = 8 + x$ Scale $\beta = e^{(0.5+0.1x)}$	Shape $\alpha = e^{(3+0.2x)}$ Scale $\beta = e^{(-0.6+0.1x)}$	Shape $\alpha = e^{(1+0.1x)}$ Scale $\beta = e^{(0.2+0.2x)}$
	Weibull	$\mu = 8 + 0.1x$ $\sigma^2 = e^{(0.05+0.1x)}$	Location $\mu = 5 + x$ Scale $\beta = e^{(0.6+0.1x)}$	Shape $\alpha = e^{(2+0.2x)}$ Scale $\beta = e^{(0.1x)}$	Shape $\alpha = e^{(1+0.1x)}$ Scale $\beta = e^{(0.1+0.2x)}$

Table A1. Simulation process: conditional distribution of the data for the test of incorrect distributional family.

True Data Generating Process
Normal/Normal (homoscedasticity)
$\mu = 6 + 0.1x + 0.65x^2$
$\sigma^2 = 1$
Normal/Normal (heteroscedasticity)
$\mu = 6 + 0.1x + 0.35x^2$
$\sigma^2 = e^{(1+0.2x+0.35x^2)}$
Gumbel/Gumbel
Location $\mu = 5 + x + 0.24x^2$
Scale $\beta = e^{(0.6+0.1x+0.24x^2)}$
Gamma/Gamma
Shape $\alpha = e^{(2+0.2x+0.1x^2)}$
Scale $\beta = e^{(0.1x+0.1x^2)}$
Weibull/Weibull
Shape $\alpha = e^{(1+0.1x+0.47x^2)}$
Scale $\beta = e^{(0.2+0.2x+0.47x^2)}$

Table A2. Simulation process: conditional distribution of the data for the test of incorrect parameterization

Variables	**Sample for positive total health care expenditure (N=1527)**	**Sample for positive total office-based visits expenditure (N=1215)**	**Sample for positive total dental care expenditure (N=724)**
*Cost-related outcome variables**			
Total health care expenditure, median (IQR), $	947 (291-3359)	---	---
Total office-based visits expenditure, median (IQR), $	---	371 (145-1053)	---
Total dental care expenditure, median(IQR), $	---	---	225 (113-501)
Explanatory variables			
Age, median (IQR), y	35 (16-55)	39 (17-58)	33 (14-55)
Female sex	849 (55.60)	686 (56.46)	388 (53.6)
Race/Ethnicity			
White	709 (46.43)	584 (48.07)	384 (53.04)
African American	307 (20.10)	232 (19.09)	114 (15.75)
Hispanic	390 (25.54)	301 (24.77)	170 (23.48)
Other	121 (7.92)	98 (8.07)	56 (7.73)
Education			
<High school	359 (26.03)	265 (24.40)	197 (29.27)
Some high school	107 (7.76)	89 (8.20)	32 (4.75)
High school graduate	349 (25.31)	279 (25.69)	125 (18.57)
Some college	254 (18.42)	207 (19.06)	117 (17.38)
College graduate and above	310 (22.48)	246 (22.65)	202 (30.01)
Self-rated health status			
Excellent	480 (31.54)	365 (30.17)	258 (35.73)
Very good	467 (30.68)	356 (29.42)	230 (31.86)
Good	383 (25.16)	318 (26.28)	168 (23.27)
Fair	148 (9.72)	132 (10.91)	60 (8.31)
Poor	44 (2.89)	39 (3.22)	6 (0.83)
Chronic diseases (≥3)	140 (9.17)	129 (10.62)	676 (6.63)
Usual source of care	1235 (82.44)	1037 (86.34)	595 (83.45)
Household income relative to percentage of FPL			
<100	307 (20.10)	237 (19.51)	128 (17.68)
100-124	92 (6.02)	71 (5.84)	33 (4.56)
125-199	251 (16.44)	194 (15.97)	95 (13.12)
200-399	450 (29.47)	359 (29.55)	212 (29.28)
>400	427 (27.96)	354 (29.14)	256 (35.36)
Health insurance coverage			
Private	926 (60.64)	751 (61.81)	473 (65.33)
Public	466 (30.52)	377 (31.03)	215 (29.70)
None	135 (8.84)	87 (7.16)	36 (4.97)

Note. Numbers are number (%) unless otherwise indicated. * The cost variables are log-transformed before entered into the MLE model. IQR, interquartile range. FPL, federal poverty level.

Table A3. Distribution of the cost-related outcome variables and patient characteristics.

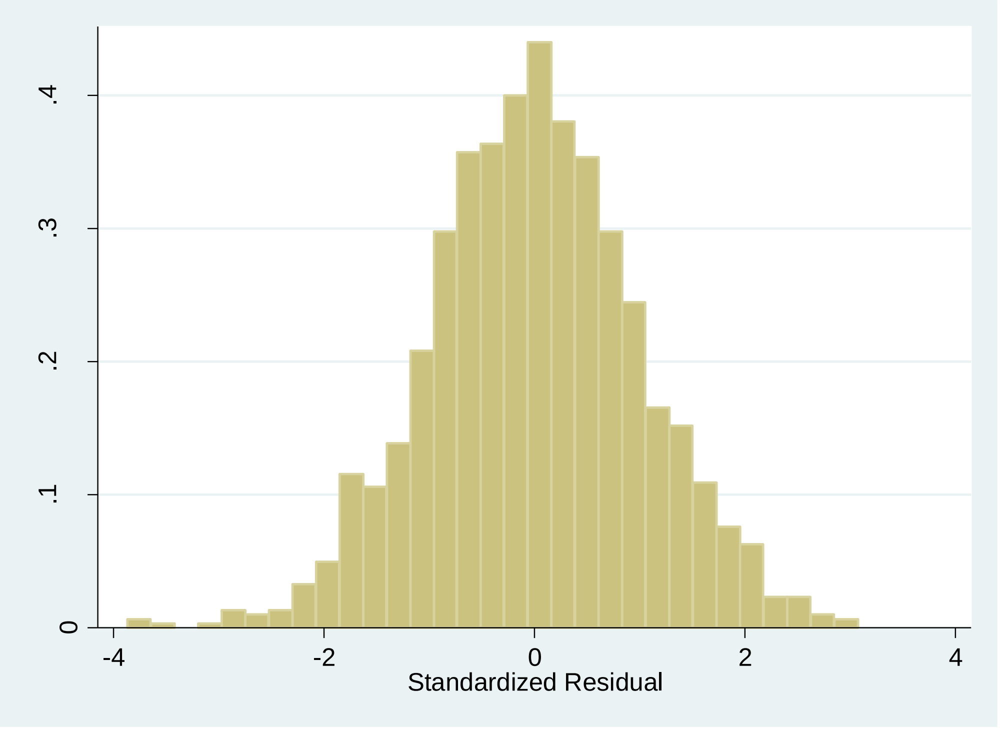
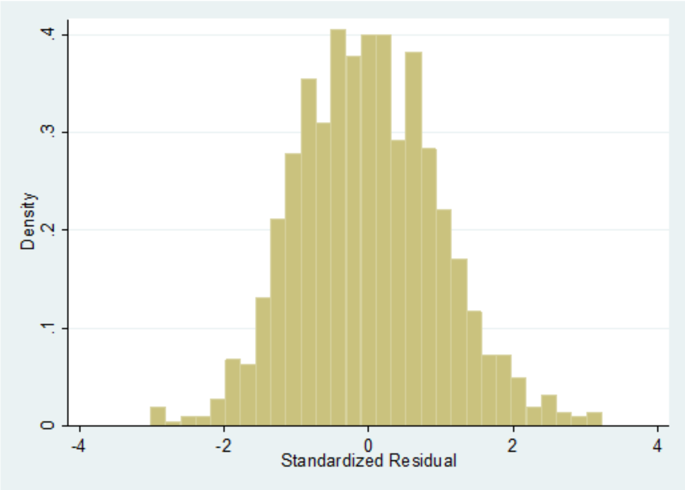


Figure A1. Histogram of the Standardized Residual from the Model for Annual Total Health Care Expenditure. Model: MLE assuming Normal distribution with heteroskedasticity.

Model 1: MLE assuming Gamma distribution:



Model 2: MLE assuming Normal distribution with heteroskedasticity:

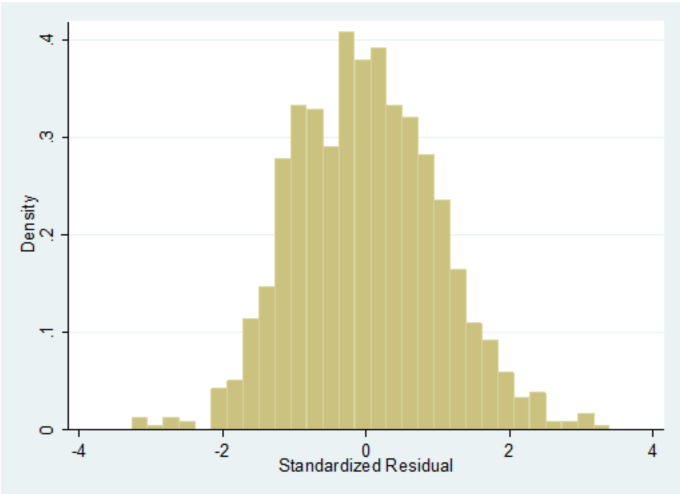


Figure A2. Histogram of the Standardized Residual from the Model for Annual Total Expenditures on Office-Based Visits.

Model: MLE assuming Gumbel distribution:

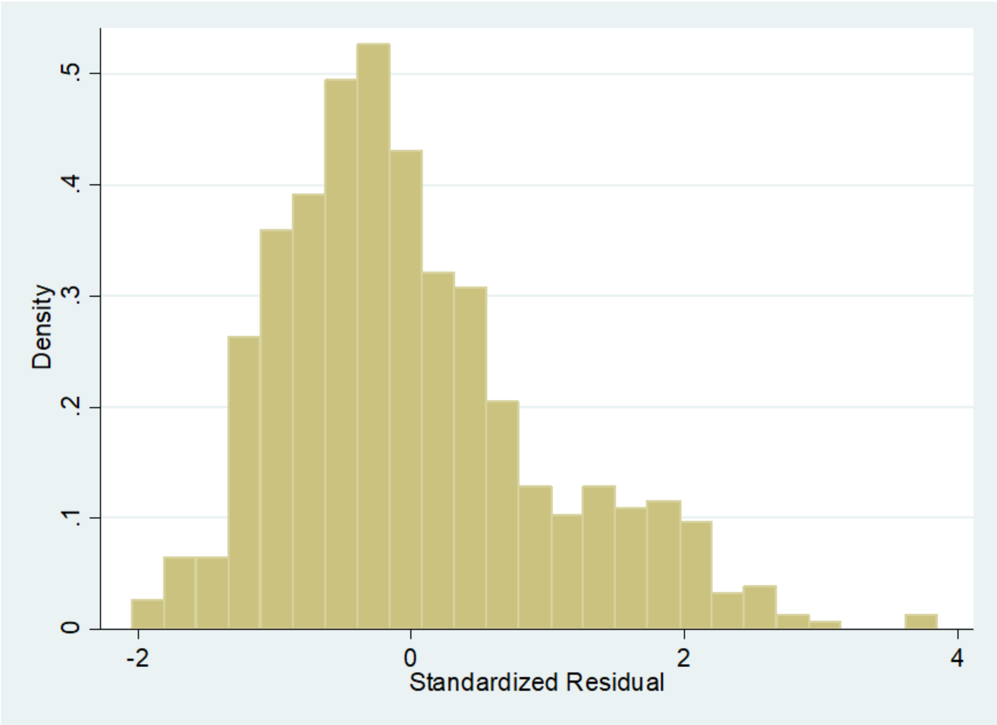


Figure A3. Histogram of the Standardized Residual from the Model for Annual Total Expenditures on Dental Care.

Ratio Mathematica Volume 39, 2020, pp. 33-54

Modelling the shape of sunspot cycle using a modified Maxwell-Boltzmann probability distribution function

Amaranathan Sabarinath [*]
Girija Puthumana Beena [†]
Ajimandiram Krishnankuttynair Anilkumar[‡]

Abstract

The 11-year sunspot number cycle has been a fascinating phenomenon for many scientists in the last three centuries. Various mathematical functions have been used for modelling the 11-year sunspot number cycles. In this paper, we present a new model, which is derived from the well known Maxwell-Boltzmann probability distribution function. A modification has been carried out by introducing a new parameter, called area parameter to model sunspot number cycle using Maxwell-Boltzmann probability distribution function. This parameter removes the normality condition possessed by probability density function, and fits an arbitrary sunspot cycle of any magnitude. The new model has been fitted in the actual monthly averaged sunspot cycles and it is found that, the Hathaway, Wilson and Reichmann measure, the goodness of fit is high. The estimated parameters of the sunspot number cycles 1 to 24 have been presented in this paper. A Monte Carlo based simple random search is used for nonlinear parameter estimation. The Prediction has been carried out for the next sunspot number cycle 25 through a model by averaging of recent cycle's model parameters. This prediction can be used for simulating a more realistic sunspot cycle profile. Through extensive Monte Carlo simulations, a large number of sunspot cycle profiles could be generated and these can be used in the studies of the orbital dynamics.

Keywords: solar cycle; modelling; sunspot number
2010 AMS subject classification: 70F15; 97M10.[§]

[*] Vikram Sarabhai Space Centre, Thiruvananthapuram, Kerala, India; a_sabarinath@yahoo.co.in.
[†] Assistant Professor, St.Gregorios College, Kottarakkara, Kerala, India; beenamabhi@gmail.com.
[‡] Director, DSSAM, ISRO Head Quarters, Bangalore, Karnataka, India; ak_anilkumar@isro.gov.in.

1 Introduction

To know in advance the multitude of atmospheric processes that cause concern to mankind, in particular phenomena occurring in the solar plasma receive great consideration from the scientific world. Since the 18th century, scientists are conducting systematic research on a multitude of processes caused by solar activity. Solar Activity forecasting is crucial in scientific and technological fields such as spacecraft orbital life time prediction, airline communications and geophysical applications, mainly it is the energy source behind all phenomena driving space weather. The low Earth orbiting satellites are also influenced by solar activity (Seeds,M.A,Backman,D,[2015]; Hathaway,D.H., [2010]). However, predicting the solar cycle is challenging on the basis of time series of various proposed indicators, due to the high frequency contents, noise contamination, high dispersion level and high variability both in phase and amplitude.

The prediction of solar activity is complicated by the lack of a quantitative theoretical model of the sun's magnetic cycle. The effect of solar activity is greater on space activities especially on the operations of low Earth orbiting satellites which provide significant contribution in communication, national defence and Earth mapping. Such satellites also handle a large quantity of scientific data. During higher solar activity, the maximum ultraviolet rays are emitted from the sun that heat up Earth's upper atmosphere, and expands the atmosphere. This affects the life time of operational space crafts in the low earth orbits (Whitlock,D, [2006]). Therefore better predictions of solar activity are essential to help spacecraft mission planning and design.

2 Satellite life time estimation and re-entry prediction

In spacecraft mission design, orbital life time estimation is a critical activity (Whitlock, D, [2006]). Many uncertain parameters need to be considered while doing orbital life time estimation. The upper atmospheric density variation is the primary factor which is so difficult to predict. Many studies have been taken place to model the atmospheric density accurately. Orbital life time estimation community has always been looking up for better models of atmospheric density. Atmospheric models generally use parameters such as ap or Kp, and F10.7. Solar flux receives a lot of attention because it is an important parameter in determining atmospheric density. Most predictions rely

§ Received on October 24th, 2020. Accepted on December 17th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.532. ISSN: 1592-7415. eISSN: 2282-8214.

on the sunspot activity happening in the sun. This has been monitored since the 17th century regularly. An empirical relationship exists between the sunspot number, R, averaged over a month, and F10.7 (David A.Vallado et.al ,[2014]).

$$\mathrm{F10.7} = 63.7 + 0.728\,\mathrm{R} + 0.000\,89\,\mathrm{R}^2, \qquad (1)$$

From the above equation, we can see that 10.7 cm radio flux has a base level of about 63.7 solar flux units. To understand and estimate the radio emissions effectively we can use the following equation (David A.Vallado et.al, [2014])

$$\mathrm{F10.7} = 145 + 75\,\mathrm{COS}\left(0.001696\,\mathrm{t} + 0.35\,\mathrm{SIN}\,(0.00001695)\right), \qquad (2)$$

where t is the number of days from January 1, 1981.

We can summarise it as, atmospheric density is directly related to the solar flux, which in turn can be related to the solar activity. Studies done by different scientists and academicians shows that solar activity and solar flux have affirmed relation, a monthly estimate of F10.7 and sunspot number has been well established. Predicting the solar flux accurately can generate more accurate atmospheric density models that will help in fine tuning the fuel budget for longer satellite life.

The discussion went so far reminds that the accurate prediction of the life time requires a very good predicted solar flux profile. In turn, it is sufficient to have a predicted sunspot number cycle. Since, via equation (1) one can transform sunspot numbers into solar flux. In this paper we try to predict sunspot number cycle in a simple and powerful technique. Initially, we model the sunspot cycle using a skew-symmetric probability distribution. The Maxwell-Boltzmann distribution is considered for this purpose. Then a preliminary level prediction is proposed as an average (mean) cycle of some recent cycles. Then a varying error band is derived from the past cycles. Within this error profiles, via Monte Carlo sampling, the predicted averaged cycle is transformed into many profiles. Sample profiles are taken and plotted. Before venturing into the details, a brief review of sunspot data and review some of the recent models are provided.

3 Sunspot number cycles and sunspot number data

In 1848 the Swiss astronomer Johann Rudolph Wolff introduced actual measurements of sunspot number. His method uses still today. Total number of spots visible on the face of the sun is 'n' and the number of groups into which they cluster is 'g' then the sunspot number R_n is defined as

$$R_n = 10\,g + n. \tag{3}$$

To compensate the observational limitations like Earth's atmosphere variability above the observing site and sun's rotation, each daily sunspot number is computed as a weighted average of measurements made from a network of observatories. The 11-year cyclic variation in the sunspot numbers was first noted by Schwabe, M., [1844]. In 1848 Rudolf Wolf at Swiss Federal Observatory in Zurich, Switzerland devised his measure of sunspot numbers that continues to this day as the International sunspot number. Wolf recognised that it is far easier to identify sunspot groups than to identify each individual sunspot. This relative sunspot number,R_z with emphasis on sunspot groups is defined as,

$$R_z = k\,(10\,g + n), \tag{4}$$

Where k the correction factor for the observer, g is the number of identified sunspot groups, and n is the number of individual sunspots. These sunspot numbers are called the Zurich or International sunspot numbers have been obtained daily since 1848.

Sunspot cycle time series is one of the longest time series which was studied by many experts for various reasons. First of all, this time series is non-stationary, cyclic and highly nonlinear in the time domain. In the present study, the prediction of sunspot cycles is carried out with the monthly averaged sunspot number values. The monthly averaged sunspot data were available from, http://www.sidc.be/silso/versionarchive at the royal observatory, Belgium is being used for the present study. It may be noted that, the scientific community recently recalibrated the entire historical sunspot number record and that SILSO (Sunspot Index and Long-term Solar Observations) maintains this new definitive record as well as the original version of sunspot numbers.

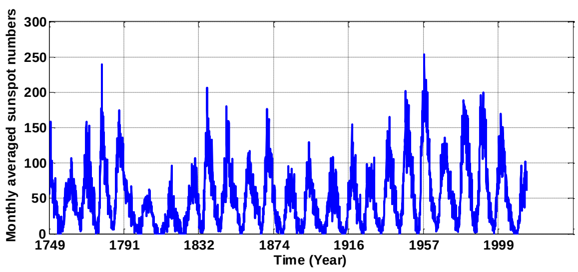


Figure 1: Sunspot cycle evolution-Monthly averaged sunspot numbers from the year 1749 to December 2016.

4 Existing models of sunspot number cycles

Several mathematical functions were introduced to model the shape of the sunspot number cycle. Due to the exponential rise and decay, the exponential function was used by Nordemann, [1992], Nordemann, et.al.,[1992]. The bell shaped nature of the sunspot cycle was explored by Hathaway et.al.,[1994]. Few statistical probability distribution functions were also proposed for the shape modelling by various authors. De Mayer, F.,[1981], proposed a model using periodic functions. In prediction, averaged models are used as an initial estimate of the future cycle.

We have an exhaustive list of details and voluminous data literature available at hand pertaining to the attempts to predict the future behaviour of solar activity (Hathaway, et.al., [1999]). It can be categorised under five heads, based on the nature of the prediction methods. They are: 1) Curve fitting, 2) Precursor, 3) Spectral, 4) Neural Networks and 5) Climatology (Sello, S.,[2001]). McNish-Lincoln curve fitting was the first attempt on the methodology of curve fitting (de Meyer,[1981], McNish, A.G., Lincoln, J.V.,[1949]). Over the years, various techniques and models have been proposed by several authors working in the field for the prediction of the nonlinear behaviour of sunspot cycles. The first breakthrough in the field of modelling the shape of the sunspot cycles by fitting an exponential function over the sunspot number cycle time series was due to Nordemann,[1992]. In this method, fitting the rise to maximum and the fall to minimum were fitted with a function of exponential function demanding six free parameters. Later a modified version of F-distribution density function with five parameters was proposed by Elling and Schwentek[1992]. Nordemann's[1992] method suggests exponential fitting and explain the solar behaviour. Hathaway, Wilson, and Reichmann[1994] substantiated the superiority of a new model along with a measure for the goodness of fit. Number of free parameters in this model is reduced to four. All these models introduce high amount of error in the prediction, due to the incompetence to fit the peak locations of the sunspot cycle. The continuous nature of the model at the high solar activity period contributes a large amount of uncertainty and hence in the applications such as the orbital re-entry predictions these models are not suitable. The next subsection surveys the literature pertaining to some models, especially on the shape of sunspot cycles.

4.1 Stewart and Panofsky model

Stewart and Panofsky [1938] proposed a function for the shape of the cycle with the form

$$R(t) = a(t - t_0)^b e^{-c(t-t_0)}, \quad (5)$$

where $a, b, c,$ and t_0 are parameters that vary from cycle to cycle. The important thing to be noticed is that, this model gives a power law for the rising phase of a cycle and an exponential for the declining phase of a cycle. The model parameters for cycle 1to 16 were computed and there by the maximum amplitude, the epoch of the peak sunspot number, etc. was predicted.

4.2 Nordemann model

Nordemann used the solution of the differential equation $\frac{dN}{dt} = KN,$ in analogy with the nuclear decay process. Thus the declining phase of a sunspot cycle is represented by:

$$N = N_0 e^{Kt} \qquad K < 0 \quad (6)$$

and the solution of $\frac{dN}{dt} = A + KN,$ is used to represent the ascent phase of a sunspot cycle. Thus the model for the ascent phase is:

$$N = \frac{A}{K}(1 - e^{Kt}) \qquad K < 0 \quad (7)$$

Where N represents sunspot numbers, K decay constant and A a production parameter. The estimated values of the parameters N_0, K and A for all the 22 sunspot cycles were given in Nordemann [1992].

4.3 Elling and Schwentek model

Instead of using yearly means, quarterly averages of sunspot numbers were utilised by Elling and Schwentek[1992] for optimal fitting of each cycle. They used a modified F-distribution density function that required five free parameters. This approach is much more worth than the previous models. In this model fitting concluded only for modern era of sunspot cycles (10 to 21).

By considering the maxima and minima of mean sunspot number as a function of time, affinity can be observed in each cycles. While considering different sunspot cycles the ascending phase take dwindle time than the descending phase, that means Starting from a minimum, time taken for reaching the maximum is always shorter as compared to the time from maximum down to minimum .

They explained very effectively, ascending and descending branches of the various cycle curves have curvatures which are rather similar to those of the F-distribution curves. For this reason, each sunspot cycles from cycle 10 to cycle 21 has been approximated by a modified F-distribution, $f(t)$ which is defined by:

$$f(t) = P_4 \frac{\Gamma\left[\frac{P_2 + P_3}{2}\right]}{\Gamma\left(\frac{P_2}{2}\right)\Gamma\left(\frac{P_3}{2}\right)} P_2^{\frac{P_2}{2}} P_3^{\frac{P_3}{2}} \frac{[P_1(t + P_5)]^{\frac{P_2}{2}-1}}{[P_3 + P_2 P_1(t + P_5)]^{\frac{(P_2+P_3)}{2}}}, \tag{8}$$

where t is the time and $\Gamma(x)$ is the gamma function. P_1 is the length or duration of the sunspot cycle, that is, the time interval from one minimum to the next, P_2 to the curvature of the ascending branch of $f(t)$, P_3 to the curvature of the descending branch of $f(t)$, P_4 to the amplitude of the maximum of $f(t)$, P_5 to the time shift of the $f(t)$ curve. Through least square fit all the five parameters are estimated.

4.4 Hathaway, Wilson, and Reichmann model

Hathaway et.al [1994] suggested a model with free parameters fewer than the models which we had come across. They utilised a four-parameter quasi-Planck function to fit the monthly mean sunspot numbers of a solar cycle, similar to that of Stewart and Panofsky[1938]. But the only difference we can see that a fixed power law for the initial rise of the sunspot cycle and the phase starting from maximum down to minimum can be well represented by a function that decreases as e^{-t^2}. By combining these, the model as a function of time can be written as:

$$f(t) = \frac{a(t - t_0)^3}{e^{\left[\frac{(t-t_0)^2}{b^2}\right]} - c}, \tag{9}$$

This model has four parameters. a represents the amplitude and is directly related to the rate of rise from minimum; b is related to the time in months from minimum to maximum; c gives the asymmetry of the cycle; and a starting time t_0 . Along with the early detection of parameters to predict the solar activity they examine the relationship between the parameters. It is similar to the Plank function but contains four free parameters and has a more rapid decrease after maximum, but causes lack of accuracy. The estimation of these parameters was obtained through Levenberg-Marquardt methods (Press, W, H., [1992]).

4.5 Volobuev’s one-parameter fit

In 2009, Volobuev introduced a function of two-parameters and he refers to this as a one parameter fit. We can see that the parameters are correlated (r = 0.88) for all the 23 solar cycles. The correlation between the parameters provides the possibility of a one-parameter fit by neglecting the need to determine the best starting time. He showed that a one-parameter fit can also be derived from truncated dynamo models. Due to the unavoidable uncertainty of starting time goodness of fit value is not better as compared to the empirical fit.

We can see that this model is also similar to that of Stewart and Panofsky [1938] proposed Pearson's type III curves by putting b =2 and modifying the growth multiplier and decay multiplier properly by introducing the new parameters T_s and T_d.

The empirical model used is written as:

$$R = \left(\frac{t - t_0}{T_s}\right)^2 e^{-\left(\frac{t-t_0}{T_d}\right)^2}, \tag{10}$$

4.6 Sabarinath and Anilkumar model

Sabarinath and Anilkumar[2008] proposed a model consist of a mixture of Laplace distribution with six parameters (later reduced to two). This model fits the multiple sharp peaks in a solar cycle. The model for a generic cycle is:

$$F = \frac{A_1}{33.2} \exp\left(\frac{-|t - 41.7|}{16.6}\right) + \frac{A_2}{46} \exp\left(\frac{-|t - 67.3|}{23}\right), \tag{11}$$

where t is the time.

5 Skew symmetrical distributions

Sunspot cycles are asymmetric with respect to their maxima (Hathaway, D.H., [2010]). Starting from minimum the time taken to reach maximum is 48 months and 84 months to fall back to minimum again. An average cycle can be constructed by stretching and contracting each cycle to the average length and normalising each to the average amplitude.

In general, if we survey any model of the shape of the sunspot cycle, it is evident that, all functions are a product of a polynomial and a negative exponential function. Then the goodness of fit solely depends on how the model parameters are chosen in the model. In this context, we propose a skew symmetrical function from the class of skew symmetrical probability functions.

6 Maxwell-Boltzmann probability distribution function

In statistical physics, Maxwell-Boltzmann distribution is a probability distribution named after the famous Scottish physicist James Clerk Maxwell and Ludwig Boltzmann. It is used in atomic physics for describing particle speeds in idealised gas. The Maxwell-Boltzmann distribution function is given as (Balakrishnan, N., Nevzorov, V.B., [2003])

$$f(v) = \sqrt{\left(\frac{m}{2\pi kT}\right)^3}\, 4\pi v^2 e^{-\frac{mv^2}{2kT}}\,, \tag{12}$$

where m the particle mass and kT is the product of Boltzmann's constant and thermodynamic temperature. From Equation (12), if we put $\alpha = \sqrt{\frac{kT}{m}}$, then the Maxwell-Boltzmann probability distribution function can be simplified as

$$f(x;\alpha) = \frac{1}{\alpha^3}\sqrt{\frac{2}{\pi}}\; x^2 e^{-\frac{x^2}{2\alpha^2}}, \tag{13}$$

where the variable v is replaced with a generic random variable x with $x \geq 0$ and it can be noted that the parameter $\alpha \geq 0$ is a real quantity.

Typical shape of Maxwell-Boltzmann distribution is given in Figure-2, for a value of α=30. One can clearly see from Figure-2 that the ascend phase is of 47 units and the descent phase is 85 units. There by, a skew symmetrical process or phenomenal could be modelled by the Maxwell-Boltzmann distribution. Our interest is in modelling the sunspot cycle. By observing all the cycles individually one can easily see that the rise time (starting minimum to maximum sunspot number) and fall time (maximum sunspot number to cycle end) are not equal or not symmetrical about the peak sunspot number occurring epoch during the 11 year sunspot cycle period.

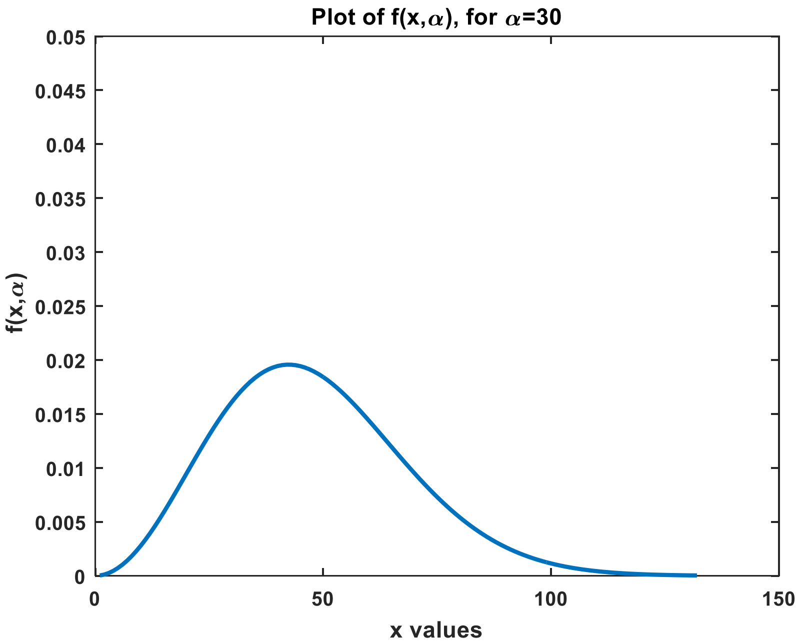


Figure-2. Maxwell-Boltzmann distribution for a value of α=30.0

7 Modified Maxwell-Boltzman probability distribution function (MMPDF)

Since equation (13) being a probability density function, we know that, mathematically the area under the probability density function is 1, that is,

$$\int_{-\infty}^{\infty} f(x)dx = 1, \tag{14}$$

So, if we want to fit this equation (13) into an arbitrary set S of N data points, $S = \{(x_i, y_i);\ x_i \in R, y_i \in R,\ i = 1,2,\dots,N\}$, where R is the set of real numbers, we need to de-normalise the property of f(x) given by equation (14). This is because; the area under the curve determined by the set of points in S need not be equal to one. That is,

$$\sum_{i=2}^{N}\left[(x_i - x_{i-1})\frac{(y_i + y_{i-1})}{2}\right] = A, \tag{15}$$

where A need not be equal to 1. In this case we can modify equation (13) to fit into any arbitrary set as equation (16) by introducing a new parameter called area parameter A.

$$f(x;\alpha;A) = \frac{A}{\alpha^3}\sqrt{\frac{2}{\pi}}\ x^2 e^{-\frac{x^2}{2\alpha^2}}, \tag{16}$$

Now, it may be noted that,

$$\int_{-\infty}^{\infty} f(x)dx = A, \tag{17}$$

Modified model for the sunspot cycles is

$$f(x; \alpha; A) = \frac{A}{\alpha^3}\sqrt{\frac{2}{\pi}}\ x^2 e^{-\frac{x^2}{2\alpha^2}}, \tag{18}$$

where A is the area parameter.

Modified Maxwell-Boltzmann distribution with a value of α=30 and A=6000 is given in Figure-3.

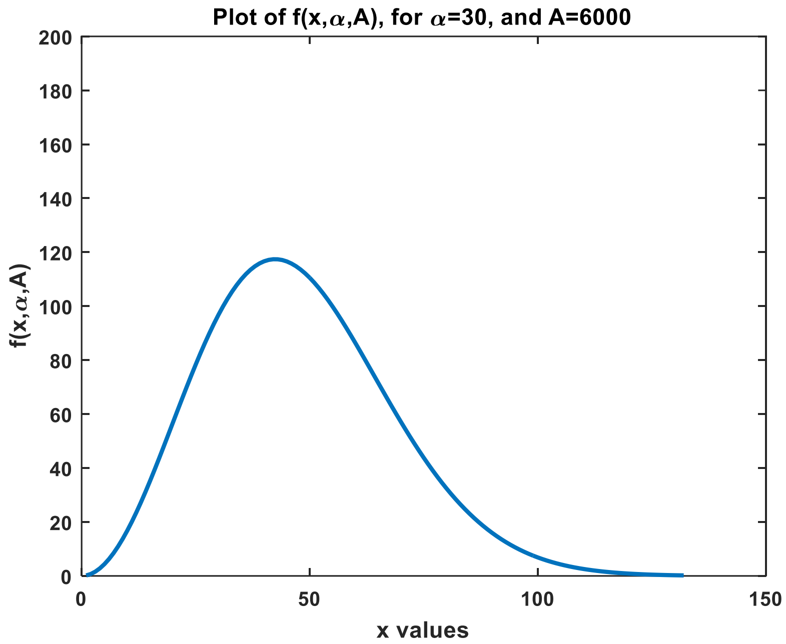


Figure-3. Modified Maxwell-Boltzmann distribution for a value of α=30 and A=6000.

8 Estimation of model parameters

The function in which parameters to be estimated is,

$$f(x; \alpha; A) = \frac{A}{\alpha^3}\sqrt{\frac{2}{\pi}}\ x^2 e^{-\frac{x^2}{2\alpha^2}}. \tag{19}$$

The maximum likelihood estimate of the parameters α and A are considered to be the best unbiased, consistent and sufficient estimate of the parameters

(Sorenson, H.W., [1980]). Practically, the least square estimate is considered to be the maximum likelihood estimate. The simple mathematical procedure to estimate the parameters is to minimise the sum of squared error function J,

$$J = \sum_{r} e_r^2 \,, \tag{20}$$

Where e_r is the error.

The minimum of J can be found by differentiating J with respect to the parameters α and A.

In the present study, if we consider without loss of generality, a sunspot cycle having a length of 132 months(≈ 11 year), and if we assume $\{s_n : n = 1,2,\dots,132\}$ as the realised sunspot number values, then the J function can be written as,

$$J = \sum_{n=1}^{132} [s_n - f(x_n, \alpha, A)]^2, \tag{21}$$

where, $x_n = 1,2,\dots,132$, represents the months for each $n = 1,2,\dots.,132$. Then our objective is to compute and solve α and A from

$$\frac{\partial J}{\partial \alpha} = 0, \tag{22}$$

$$\frac{\partial J}{\partial A} = 0, \tag{23}$$

Analytically solving the equations (22) and (23) for α and A is not possible due to the nonlinear terms involved in the equations. Hence we go with numerical procedures for estimating the parameters. Monte Carlo based simple random search based procedure is considered here to estimate the parameters. This procedure is described below as an algorithm.

Step-1. Start with a search region α and A. Let S_α and S_A are the bounded search regions of α and A. Our objective is to find an $\alpha_0 \in S_\alpha$ and $A_0 \in S_A$, such that,

$$J_{\alpha_0, A_0} = \sum_{n=1}^{132} [s_n - f(x_n, \alpha_0, A_0)]^2, \tag{24}$$

is minimum or

$$J_{\alpha_0, A_0} \le J_{\alpha, A} \tag{25}$$

for any $\alpha \in S_\alpha$ and $A \in S_A$.

Step-2. Start with a random initial value of α in S_α and A in S_A. Compute J and in each iteration keep the minimum value of J, α and A. After a very large number of iterations take the value of, α and A corresponds to the global minimum value of J.

9 Fitting of MMPDF on sunspot cycles

Using the method described in section 8, the model parameters are estimated for all the past 24 cycles. It is noticed that the fit is very much close to the actual sunspot numbers. This is evident in the goodness of fit computed for each of the 24 cycles, which is discussed in the next section in detail. Figure 4 and 5 shows the model and actual data of sunspot cycles 20 and 22.

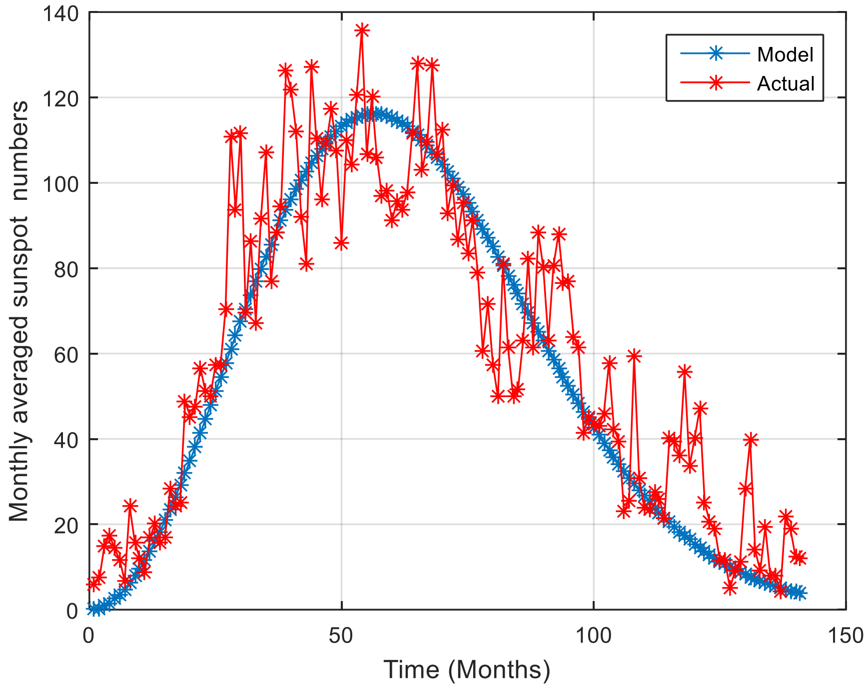


Figure-4. Fitting of sunspot cycle 20 by the model

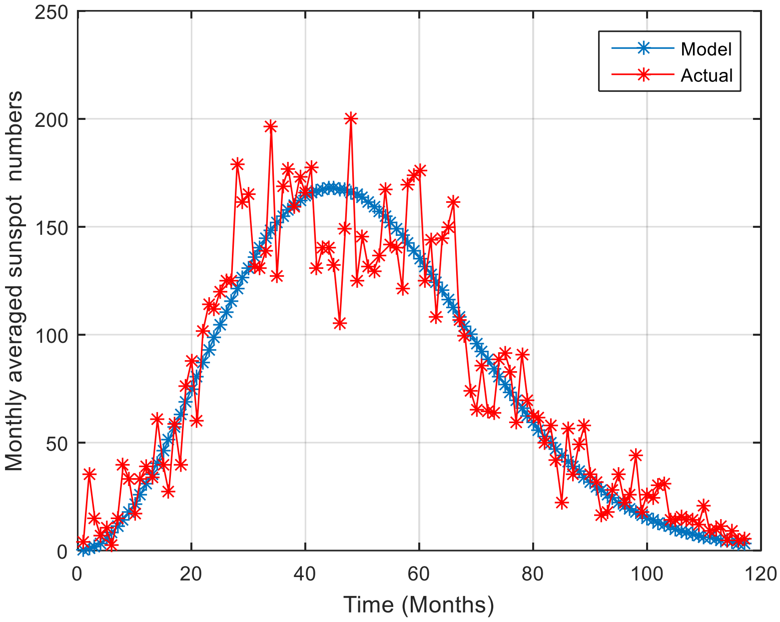


Figure-5. Fitting of sunspot cycle 22 by the model

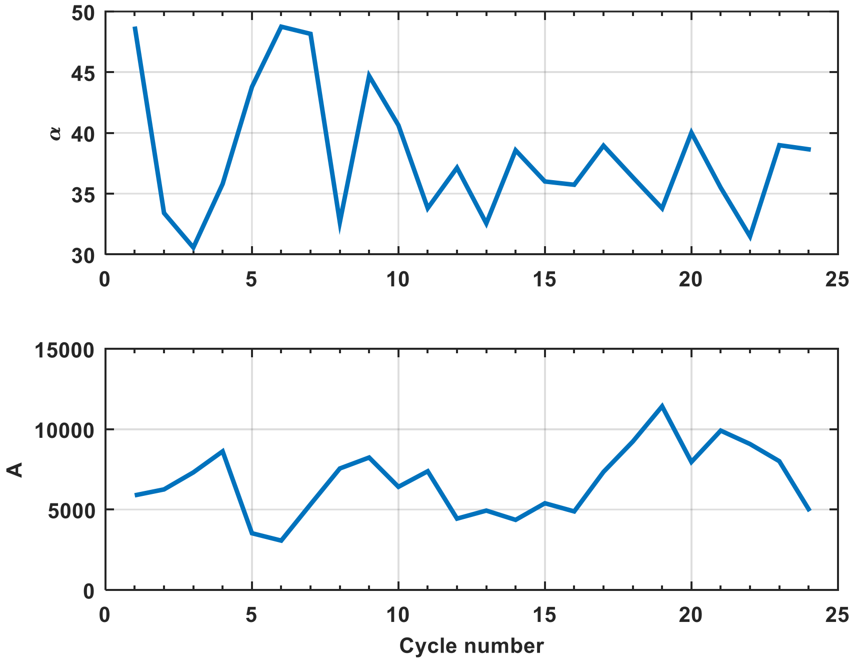


Figure-6. The parameters α and A for all the 24 cycles.

10 Models of sunspot cycles 1 to 24

The estimated model for all the past 24 cycles is given in Table-1. In Figure-6, the variation trends of the parameters α and A for all the 24 cycles are plotted. It may be noted that, the average of the parameters are 36.25 units of α and 7095.76 of A.

Table-1. Estimated parameters of cycles 1 to 24

Cycle No	α	A
1	48.76	5883.33
2	33.40	6251.65
3	30.56	7309.46
4	35.80	8619.41
5	43.79	3525.58
6	48.75	3067.09
7	48.16	5322.72
8	32.65	7552.73
9	44.70	8234.25
10	40.63	6410.82
11	33.80	7381.50
12	37.13	4433.49
13	32.55	4933.80
14	38.57	4356.00
15	36.00	5390.27
16	35.73	4882.42
17	38.95	7341.83
18	36.35	9228.41
19	33.79	11420.62
20	40.02	7959.33
21	35.49	9907.72
22	31.48	9075.22
23	38.99	8006.39
24	38.65	5023.60
Mean 1 to 24	38.11	6729.90
Mean 11 to 24	36.25	7095.76

It may be noted that variation in α is less and variation in A is more. So A is a more sensitive parameter than α. Variation in A is not much significant as its sensitivity is less.

10.1 Goodness of fit

Goodness of fit by Hathaway, Wilson, and Reichmann [1994] is measured by the following function

$$\chi = \sqrt{\frac{\left(\sum_{i=1}^{N}(R_i - f_i)^2 \Big/ s_i^2\right)}{N}}, \qquad (26)$$

where, R_i and s_i is the monthly averaged sunspot number and its standard deviation respectively , f_i gives the functional fit value, N is the number of months in the cycle. Using this equation, computed χ value for all the 23 cycles. For Checking the Goodness of fit of the proposed model we have to consider other popular methods available in the literature. The second column of Table 2 gives the goodness of fit of the proposed Modified Maxwell-Boltzmann distribution function; the third and the fourth column gives the goodness of fit by three and two parameter fit of Hathaway, Wilson, and Reichmann [1994], respectively; the fifth column gives the goodness of fit by the five parameter function of Elling and Schwentek[1992] who considered cycles 10 to 21 for their study. Figure-7, shows the goodness of fit of 3 different models along with the Modified Maxwell-Boltzmann distribution function model.

It may be observed from the goodness of fit value, that the present model proposed in this study has a very good fitness compared with other models. Especially the modern cycles (cycles 11 to 24) shows very good fitness for the Modified Maxwell-Boltzmann distribution function model.

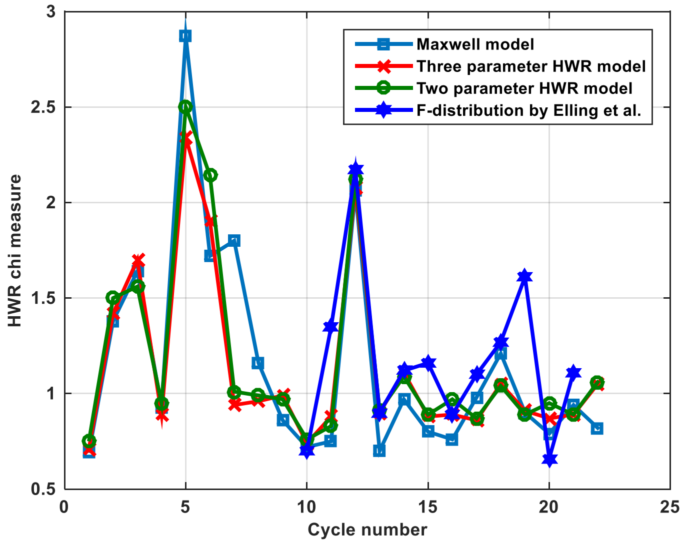


Figure-7. The goodness of fit of 3 different models and MMPDF model

Table 2: Hathaway, Wilson and Reichmann χ -measure of the goodness of fit value computed for all the 22 sunspot cycles with different models. MMPDF shows good fit compared with other models.

Cycle Number	MMPDF model	Three-parameter fit by Hathaway et.al.	Two-parameter fit by Hathaway et.al.	Elling-Schwentek F-distribution fit
1	0.69	0.71	0.75	
2	1.38	1.42	1.50	
3	1.64	1.70	1.56	
4	0.93	0.89	0.95	
5	2.87	2.34	2.50	
6	1.72	1.90	2.14	
7	1.80	0.94	1.01	
8	1.16	0.96	0.99	
9	0.86	0.99	0.97	
10	0.72	0.74	0.76	0.70
11	0.75	0.88	0.83	1.35
12	2.06	2.08	2.12	2.17
13	0.70	0.90	0.91	0.90
14	0.97	1.11	1.09	1.12
15	0.80	0.88	0.89	1.16
16	0.76	0.89	0.97	0.89
17	0.98	0.86	0.87	1.10
18	1.21	1.05	1.04	1.27
19	0.90	0.91	0.89	1.61
20	0.79	0.87	0.95	0.66
21	0.94	0.89	0.89	1.11
22	0.82	1.05	1.06	
23	0.79			

11 Prediction of sunspot cycle 25

As an attempt to predict the sunspot cycle 25, we consider the average of the model parameters by considering cycles-11 to 24. This computed average is given in Table-1. Thus, the parameter values of cycle 25 are: α = 36.25, and A = 7095.76. Hence the model is,

$$f(x; \alpha; A) = \frac{A}{\alpha^3}\sqrt{\frac{2}{\pi}}\ x^2 e^{-\frac{x^2}{2\alpha^2}}, \tag{27}$$

where, α = 36.25, and A = 7095.76. That is,

$$f(x; 36.74; 6608.04) = 0.119\ x^2 e^{-0.00038x^2}, \tag{28}$$

is the model for cycle-25. Figure-8 shows the shape of cycle 25 in an average sense. It may be observed that cycle 25 may peak up to 105 units and it is also fairly a slow cycle as cycle 24.

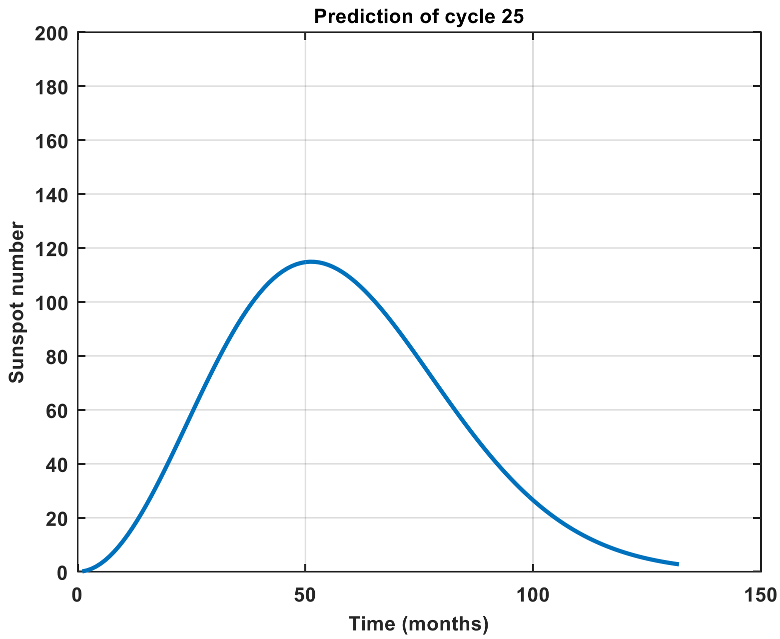


Figure-8. Preliminary level prediction of sunspot cycle 25

.

12 Prediction error and simulated sunspot cycles

Any prediction or forecast is partial, if it is not supplemented with a prediction error. Here, for our study we propose a prediction error band based on the statistical variation of all the cycles. For this, consider all the monthly averaged cycles. We propose the error band each month data as ±s, where s is the standard deviation of the sunspot numbers for that month. Figure-9 shows the mean along with the mean+s, the upper bound, and mean-s, the lower bound profile.

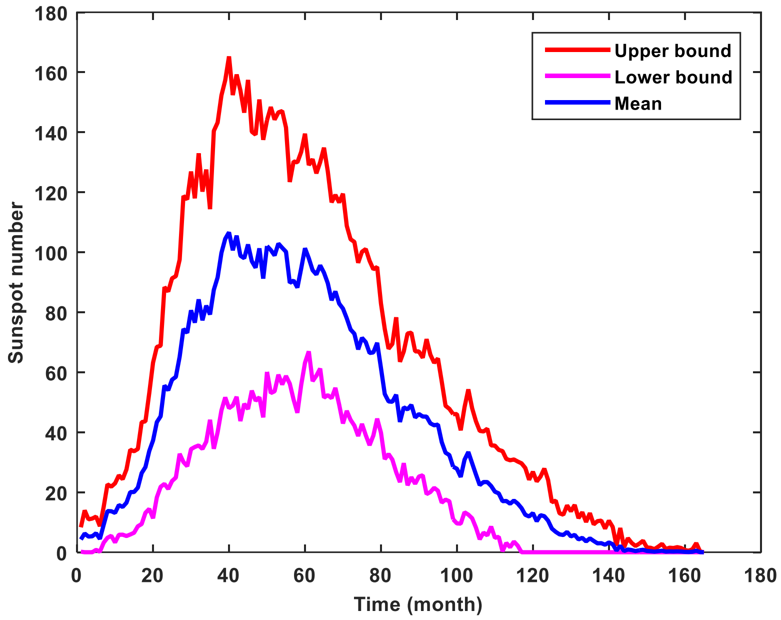


Figure-9. Mean cycle from the actual monthly sunspot cycle along with the mean+s, the upper bound and mean-s, the lower bound profile

Once we are having a prediction error and a prediction model, we can generate any number of forecast profile based on simple Monte Carlo method. Here we consider the envelop derived above as the envelope with 99.7% confidence or 3sigma confidence level, since all the realised cycles falls inside the proposed confidence interval band. Hence in the Monte Carlo simulation a typical profile will be generated using equation (29).

$$s_n'(i) = m_n(i) + \mathrm{rand}(i) \times \left(\frac{\mathrm{env}(i)}{3}\right), \tag{29}$$

where $s_n'(i)$, is the simulated n-th sunspot cycle, $i = 1,2, \dots,$ Cycle length, $m_n(i)$ is the model value, $\mathrm{rand}(i)$ is the random number and $\mathrm{env}(i)$ is the envelop value given in Figure-10.

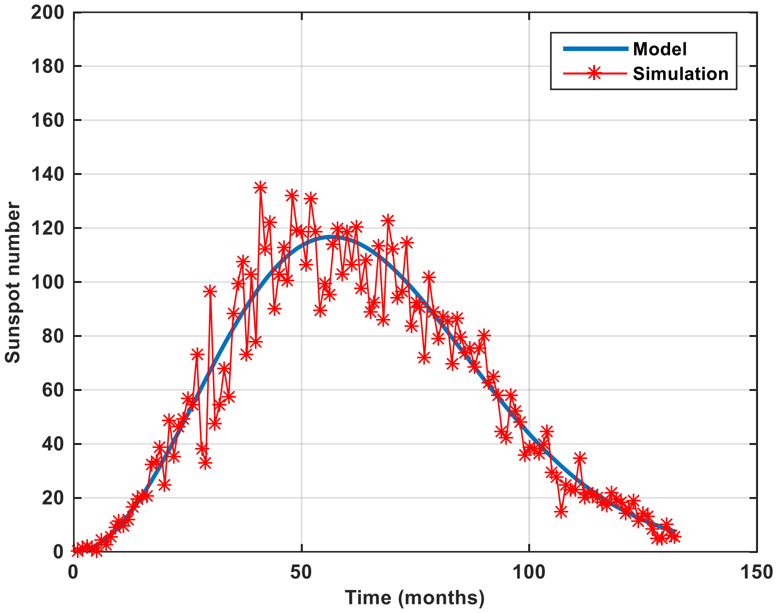


Figure-10. Simulated sunspot cycle 20 by the model

The same methodology proposed in the study can be implemented to the F10.7 cm solar flux value and one can easiliy forecast an entire cycle and subsequently it can be applied in the life time computation of satellites.

13 Conclusions

The 11-year sunspot number cycles have been a fascinating phenomenon for many in the last three centuries. Different mathematical models have been derived for modelling the shape of the 11-year sunspot number cycles. In the present study, we introduced a new model which is derived from the well known Maxwell-Boltzmann probability distribution function. The modification has been carried out by introducing a new parameter, called area parameter. The new model has been fitted in the original monthly averaged sunspot cycles data and it is found that a very high goodness of fit through the Hathaway, Wilson and Reichmann measure. The models estimated for all the sunspot cycles from 1 to 24 have been presented. Detailed discussion on the nonlinear parameter estimation carried out for fitting the function in the original data is also summarised. An attempt has been carried out for predicting the next sunspot cycles 25. The sunspot cycle 25 may peak up to 105 units and it is also fairly a slow cycle as the previous cycle 24.

References

[1].Balakrishnan, N., Nevzorov, V.B., A primer on statistical distributions, John Wiley & Sons, Inc., Hoboken, New Jersey. 2003.

[2].De Mayer, F., Mathematical Modelling of the sunspot cycle, Solar Phys., Vol. 70, 259-272,1981.

[3].David, A. Vallado, David Finkleman, A Critical Assessment of Satellite Drag and Atmospheric Density Modeling. 2014.

[4].Elling, W., Schwentek, H., Fitting the sunspot cycles 10-21 by a modified F-distribution density function, Solar Phys., Vol. 137, 155-165, 1992.

[5].Hathaway, D.H., The solar cycle, Living reviews in Solar Physics, 7, 1. 2010.

[6].Hathaway, D.H., Wilson, R.M., Reichmann, E.J., J. Geophys. Res., Vol. 104, No. A10, 1999.

[7].Hathaway,D.H., Wilson, R.M., Reichmann, E.J., The shape of the sunspot cycle, Solar Phys., April 1994, Vol. 151, No.1, pp 177-190.1994.

[8].McNish, A.G., Lincoln, J.V., Prediction of sunspot numbers, Eos Trans. AGU, 30, 673–685.1949.

[9].Nordemann, D.J.R., Sunspot number time series: exponential fitting and solar behavior, Solar Phys., Vol.141, 199-202.1992.

[10].Nordemann, D.J.R., Trivedi, N.B., Sunspot number time series: exponential fitting and solar behavior, Solar Phys., Vol.142, 411-414.1992.

[11].Petrovay, K., Solar cycle prediction. Living reviews in solar physics, 7(1):6.2010.

[12].Press, W, H., The art of scientific computing. Cambridge university press.1992.

[13].Sabarinath, A., Anilkumar, A.K., Modelling of sunspot numbers by a modified binary mixture of Laplace distribution functions, Solar Phys., Vol. 250, 183-197.2008.

[14].Seeds, M.A, Backman, D, The solar system. Nelson Education.2015.

[15].Sello, S., Solar cycle forecasting: a nonlinear dynamics approach. Astronomy & Astrophysics, 377(1):312–320. 2001.

[16].Schwabe, H. von Herrn Hofrath Schwabe, Sonnen -Beobachtungen im jahre 1843. in dessau. Astronomische Nachrichten, 21:233. 1844.

[17].Sorenson, H.W., Parameter Estimation-Principles and Problems, Marcel Dekker, New York.1980.

[18].Stewart, J.Q., Panofsky, H.A.A., The Mathematical Characteristics of Sunspot Variations, Astrophys. J., Vol. 88, 385–407. 1938.

[19].Volobuev, D.M., The shape of the sunspot cycle: A one parameter fit, Solar Phys., Vol. 258, 319-330.2009.

[20].Whitlock, D, Modeling the effect of high solar activity on the orbital debris environment, Orbital Debris Quarterly News, 10(2):4.2006.

Ratio Mathematica Volume 39, 2020, pp. 55-67

New structure of norms on $\mathbb{R}^n$ and their relations with the curvature of the plane curves

Amir Veisi*
Ali Delbaznasab†

Abstract

Let $f_1, f_2, \ldots, f_n$ be fixed nonzero real-valued functions on $\mathbb{R}$, the real numbers. Let $\varphi_n(X_n) = \left(x_1^2 f_1^2 + x_2^2 f_2^2 + \ldots + x_n^2 f_n^2\right)^{\frac{1}{2}}$, where $X_n = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. We show that φ_n has properties similar to a norm function on the normed linear space. Although φ_n is not a norm on $\mathbb{R}^n$ in general, it induces a norm on $\mathbb{R}^n$. For the nonzero function $F : \mathbb{R}^2 \to \mathbb{R}$, a curvature formula for the implicit curve $G(x, y) = F^2(x, y) = c \neq 0$ at any regular point is given. A similar result is presented when F is a nonzero function from $\mathbb{R}^3$ to $\mathbb{R}$. In continued, we concentrate on $F(x, y) = \int_a^b \varphi_2(x, y)dt$. It is shown that the curvature of $F(x, y) = c$, where $c > 0$ is a positive multiple of c^2. Particularly, we observe that $F(x, y) = \int_0^{\frac{\pi}{2}} \sqrt{x^2 \cos^2 t + y^2 \sin^2 t}dt$ is an elliptic integral of the second kind.

Keywords: norm; curvature; homogeneous function; elliptic integral.
2010 AMS subject classifications: 53A10. **2010 AMS subject classifications**: 53A10. [1]

*Faculty of Petroleum and Gas, Yasouj University, Gachsaran, Iran; aveisi@yu.ac.ir
†Farhangian University, Kohgiluyeh and Boyer-Ahmad Province, Yasouj, Iran; delbaznasab@gmail.com

[1]Received on October 31st, 2020. Accepted on December 17th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.552. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

A *normed linear space* is a real linear space X such that a number $\|x\|$, the *norm* of x, is associated with each $x \in X$, satisfying: $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$; $\|\lambda x\| = |\lambda|\|x\|$ for all $\lambda \in \mathbb{R}$ and $\|x + y\| \leq \|x\| + \|y\|$.
For example, let X be a Tychonoff space, $C^*(X)$ the ring of all bounded real-valued continuous functions on X. Then $C^*(X)$ is a normed linear space with the norm $\|f\| = \sup\{|f(x)| : x \in X\}$ and pointwise addition and scalar multiplication. This is called the *supremum-norm* on $C^*(X)$. The associated metric is defined by $d(f, g) = \|f - g\|$. A non-empty set $C \subseteq \mathbb{R}^n$ is called a *convex* set if whenever P and Q belong to C, the segment joining P and Q belongs to C. Analytically the definition can be formulated in this way: if P is represented by the vector x, and Q by the vector y, then C is a convex set if with P and Q it contains also every point with a vector of form $\lambda x + (1 - \lambda)y$, where $0 \leq \lambda \leq 1$. A point P is an *interior point* of a set S contained in $\mathbb{R}^n$, if there exists an n-dimensional ball, with center at P, all of whose points lie in S. An *open set* is a set containing only interior points. A subset $C \subseteq \mathbb{R}^n$ is *centrally symmetric* (or 0-*symmetric*) if for every point $Q \in \mathbb{R}^n$ contained in C, $-Q \in C$, where $-Q$ is the reflection of Q through the origin, that is $C = -C$.

Definition 1.1. *([Siegel, 1989, page 5]) A convex body is a bounded, centrally symmetric convex open set in $\mathbb{R}^n$.*

Example 1.1. *The interior of an n-dimensional ball, defined by $x_1^2 + x_2^2 + \cdots + x_n^2 < a^2$ provides an example of a convex body.*

One of the many important ideas introduced by Minkowski into the study of convex bodies was that of gauge function. Roughly, the gauge function is the equation of a convex body. Minkowski showed that the gauge function could be defined in a purely geometric way and that it must have certain properties analogous to those possessed by the distance of a point from the origin. He also showed that conversely given any function possessing these properties, there exists a convex body with the given function as its gauge function.

Definition 1.2. *([Siegel, 1989, page 6]) Given a convex body $\mathcal{B} \subseteq \mathbb{R}^n$ containing the origin O, we define a function $f : \mathbb{R}^n \to [0, \infty)$ as follows.*

$$f(x) = \begin{cases} 1 & \text{if } x \in \partial\mathcal{B}, \\ 0 & \text{if } x = 0, \\ \lambda & \text{if } 0 \neq x = \lambda y, \end{cases}$$

where λ is the unique positive real number such that the ray through O and the point (whose vector is) x intersects the surface $\partial\mathcal{B}$ (the boundary of $\mathcal{B}$) in a point y. The function f so defined is the gauge function of the convex body $\mathcal{B}$.

Example 1.2. *Let $f : \mathbb{R} \to [0, \infty)$ defined by*

$$f(x) = \max\{|x_1|, |x_2|, \dots, |x_n|\},$$

where $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Then int$\mathcal{B}$, the interior of the cubic $\mathcal{B} = \{(x_1, x_2, \dots, x_n) : |x_i| \leq 1\}$ is a convex body and f is a gauge function of it.

It is shown in [Siegel, 1989, Theorems 4-7] that a function $f : \mathbb{R} \to [0, \infty)$ is a gauge function if and only if the following conditions hold: $f(x) \geq 0$ for $x \neq 0$, $f(0) = 0$; $f(\lambda x) = \lambda f(x)$, for $0 \leq \lambda \in \mathbb{R}$; and $f(x + y) \leq f(x) + f(y)$. Moreover, f is continuous and the convex body of f is $\mathcal{B} = \{x : f(x) < 1\}$.

A brief outline of this paper is as follows. In section 2, we introduce a function φ_n on $\mathbb{R}^n$, by the formula

$$\varphi_n(X_n) = \sqrt{x_1^2 f_1^2 + x_2^2 f_2^2 + \cdots + x_n^2 f_n^2},$$

when n fixed nonzero real-valued functions $f_1, f_2, \dots, f_n$ on $\mathbb{R}$ are given. We show that the mappings φ_n have similar properties such as norm functions within difference the ranges of these functions lie in $\mathbb{R}^{\mathbb{R}}$ while the range of a norm function is in the $[0, \infty)$. This definition allows us to define a norm and hence a gauge function on $\mathbb{R}^n$. So it turns $\mathbb{R}^n$ into a metric space. In Section 3, we focus on $n = 2$, φ_2 and the induced norm on $\mathbb{R}^2$. First, we show that if $F : \mathbb{R}^2 \to \mathbb{R}$ is a nonzero function, then k, the curvature of the implicit $G(x, y) = F^2(x, y) = c \neq 0$ at every regular point is calculated by this formula:

$$k = \frac{|\mathbf{H}G| - 4F^2|\mathbf{H}F|}{4F\left(F_x^2 + F_y^2\right)^{\frac{3}{2}}},$$

where $\mathbf{H}F$ and $\mathbf{H}G$ are the Hessian matrices of F and G respectively. It is also shown if $F(x, y) = \int_a^b \sqrt{x^2 f^2(t) + y^2 g^2(t)}dt$, then $|\mathbf{H}F| = 0$ and the eigenvalues of $\mathbf{H}F$ and $\mathbf{H}G$, where $G = F^2$ are nonnegative. Particularly, when $f(t) = \cos t$ and $g(t) = \sin t$, we prove that $\int_0^{\frac{\pi}{2}} \sqrt{x^2 f^2(t) + y^2 g^2(t)}dt$ is an elliptical integral of the second type.

2 A norm on $\mathbb{R}^n$ made by the real valued functions on $\mathbb{R}$

We begin with the following notation.

Notation 2.1. *Suppose that $f_1, f_2, \dots, f_n$ are nonzero real-valued functions on $\mathbb{R}$ and define $\varphi_n : \mathbb{R}^n \to \mathbb{R}^{\mathbb{R}}$ with*

$$\varphi_n(X_n) = \sqrt{x_1^2 f_1^2 + x_2^2 f_2^2 + \cdots + x_n^2 f_n^2}, \qquad (*)$$

where $X_n = (x_1, x_2, \ldots, x_n)$ *and* $\mathbb{R}^{\mathbb{R}}$ *is the set (in fact, ring) of all real-valued functions on* $\mathbb{R}$.

The following statement is a key lemma. However, its proof is straightforward and elementary, it will be used in the proof of the triangle inequality in the next results.

Lemma 2.1. *Let* a, b, c *and* d *are nonnegative real numbers. Then*

$$\sqrt{ac} + \sqrt{bd} \leq \sqrt{(a+b)(c+d)}.$$

Proposition 2.1. *Let* $X_n, Y_n \in \mathbb{R}^n$, $n = 1, 2$ *or* 3. *Then* $\varphi_n(X_n + Y_n) \leq \varphi_n(X_n) + \varphi_n(Y_n)$.

Proof. The inequality clearly holds when $n = 1$. Next, we do the proof for $n = 2$. Take $X_2 = (x_1, y_1)$, $Y_2 = (x_2, y_2) \in \mathbb{R}^2$ and suppose that f and g are nonzero elements of $\mathbb{R}^{\mathbb{R}}$. Then

$$\begin{aligned}
\varphi_2(X_2 + Y_2) &= \sqrt{(x_1 + x_2)^2 f^2 + (y_1 + y_2)^2 g^2} \\
&\leq \sqrt{x_1^2 f^2 + y_1^2 g^2} + \sqrt{x_2^2 f^2 + y_2^2 g^2} \\
&= \varphi_2(X_2) + \varphi_2(Y_2)
\end{aligned}$$

if and only if

$$\begin{aligned}
x_1 x_2 f^2 + y_1 y_2 g^2 &\leq \sqrt{\left[x_1^2 f^2 + y_1^2 g^2\right]\left[x_2^2 f^2 + y_2^2 g^2\right]} \\
&= \varphi_2(X_2)\varphi_2(Y_2). \qquad (\star)
\end{aligned}$$

Now, if we let $B := x_1 x_2 f^2 + y_1 y_2 g^2$ and suppose that $B \geq 0$, then $(\star)$ holds if and only if

$$f^2 g^2 (x_1 y_2 - x_2 y_1)^2 \geq 0,$$

which is always true (note, $(\star)$ trivially holds if $B \leq 0$). Hence, in this case, the proof is complete.
Here, we prove the proposition for $n = 3$. Let $X_3 = (x_1, y_1, z_1) = (X_2, z_1)$ and $Y_3 = (x_2, y_2, z_2) = (Y_2, z_2)$, where $X_2 = (x_1, y_1)$, $Y_2 = (x_2, y_2)$ and let f, g, h be nonzero elements of $\mathbb{R}^{\mathbb{R}}$. Then

$$\begin{aligned}
\varphi_3(X_3 + Y_3) &= \sqrt{(x_1 + x_2)^2 f^2 + (y_1 + y_2)^2 g^2 + (z_1 + z_2)^2 h^2} \\
&\leq \sqrt{x_1^2 f^2 + y_1^2 g^2 + z_1^2 h^2} + \sqrt{x_2^2 f^2 + y_2^2 g^2 + z_2^2 h^2} \\
&= \varphi_3(X_3) + \varphi_3(Y_3)
\end{aligned}$$

if and only if

$$\begin{aligned} x_1x_2f^2 + y_1y_2g^2 + z_1z_2h^2 &\leq \sqrt{[x_1^2f^2 + y_1^2g^2 + z_1^2h^2][x_2^2f^2 + y_2^2g^2 + z_2^2h^2]} \\ &= \sqrt{\left[\varphi_2^2(X_2) + z_1^2h^2\right]\left[\varphi_2^2(Y_2) + z_2^2h^2\right]} \end{aligned}$$

Now, if we let $a = \varphi_2^2(X_2), b = z_1^2h^2, c = \varphi_2^2(Y_2)$ and $d = z_2^2h^2$, then by $(*)$ in Notation 2.1, we have

$$x_1x_2f^2 + y_1y_2g^2 \leq \sqrt{ac}.$$

Moreover, it is clear that $z_1z_2h^2 \leq \sqrt{bd}$. Therefore,

$$x_1x_2f^2 + y_1y_2g^2 + z_1z_2h^2 \leq \sqrt{ac} + \sqrt{bd}.$$

In view of Lemma 2.1, the proof is now complete. $\square$

Next, we state the general case of Proposition 2.1.

Theorem 2.1. *Let $X_n = (x_1, x_2, \dots, x_n), Y_n = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n, \lambda \in \mathbb{R}$ and φ_n be as defined in Notation 2.1. Then the following statements hold.*

(i) $\varphi_n(X_n) = 0$ *if and only if* $X_n = 0$,

(ii) $\varphi_n(\lambda X_n) = |\lambda|\varphi_n(X_n)$,

(iii) $\varphi_n(X_n + Y_n) \leq \varphi_n(X_n) + \varphi_n(Y_n)$ *(triangle inequality).*

Proof. (i) and (ii) are evident. (iii). The proof is done by induction on n, see Proposition 2.1. If we set $X_{n-1} = (x_1, x_2, \dots, x_{n-1})$ and $Y_{n-1} = (y_1, y_2, \dots, y_{n-1})$ then X_n and Y_n can be substituted by (X_{n-1}, x_n) and (Y_{n-1}, y_n) respectively. Therefore,

$$\varphi_n(X_n + Y_n) \leq \varphi_n(X_n) + \varphi_n(Y_n)$$

if and only if

$$\begin{aligned} x_1y_1f_1^2 + \cdots + x_ny_nf_n^2 &\leq \varphi_n(X_n)\varphi_n(Y_n) \\ &= \sqrt{\left[\varphi_{n-1}^2(X_{n-1}) + x_n^2f_n^2\right]\left[\varphi_{n-1}^2(Y_{n-1}) + y_n^2f_n^2\right]}. \end{aligned}$$

Now, let $a = \varphi_{n-1}^2(X_{n-1}), b = x_n^2f_n^2, c = \varphi_{n-1}^2(Y_{n-1})$ and $d = y_n^2f_n^2$ plus the assumption of induction, we have

$$x_1y_1f_1^2 + \cdots + x_{n-1}y_{n-1}f_{n-1}^2 \leq \sqrt{ac}.$$

Moreover, it is obvious that $x_ny_nf_n^2 \leq \sqrt{bd}$. Thus, $x_1y_1f_1^2 + \cdots + x_ny_nf_n^2 \leq \sqrt{ac} + \sqrt{bd}$. Lemma 2.1 now yields the result. $\square$

Corollary 2.1. *If $f_1, f_2, \ldots, f_n$ are nonzero constant functions, then φ_n is a norm (and hence a gauge function) on $\mathbb{R}^n$.*

By Theorem 2.1, we obtain the following result.

Proposition 2.2. *Let a, b be real numbers, $f_1, f_2, \ldots,$ and f_n the restrictions of some non-zero elements of $\mathbb{R}^{\mathbb{R}}$ on $[a, b]$ such that each of them is nonzero on this set, and let φ_n be as defined in the previous parts (Notation 2.1). Then the mapping $\psi_n : \mathbb{R}^n \to [0, \infty)$ defined by*

$$\psi_n(X_n) = \int_a^b \varphi_n(X_n)dt$$

is a norm on $\mathbb{R}^n$, and hence $d(X_n, Y_n) = \psi(X_n - Y_n)$ turns $\mathbb{R}^n$ into a metric space.

Corollary 2.2. *The mapping ψ_n is a gauge function on $\mathbb{R}^n$ with the convex body $C_n = \{X_n \in \mathbb{R}^n : \psi_n(X_n) < 1\}$.*

3 $F(x, y) = \int_a^b \varphi_2(x, y)dt$ as a norm on $\mathbb{R}^2$ and the curvature in the plane

Proposition 3.1. *([Goldman, 2005, Proposition 3.1]) For a curve defined by the implicit equation $F(x, y) = 0$, the curvature of F (denoted by κ) at a regular point (x_0, y_0) (i.e., the first partial derivatives F_x and F_y at this point are not both equal to 0) is given by the formula*

$$\kappa = \frac{|F_y^2 F_{xx} - 2F_x F_y F_{xy} + F_x^2 F_{yy}|}{\left(F_x^2 + F_y^2\right)^{\frac{3}{2}}},$$

where F_x denotes the first partial derivative with respect to x, F_y, F_{xx} denotes the second partial derivative with respect to x, F_{yy}, and F_{xy} denotes the mixed second partial derivative (for readability of the above formulas, the argument (x_0, y_0) has been omitted).

We recall that the *Hessian matrix* of $z = F(x, y)$ and $w = F(x, y, z)$ are defined to be $\mathbf{H}z = \begin{bmatrix} F_{xx} & F_{xy} \\ F_{yx} & F_{yy} \end{bmatrix}$ and $\mathbf{H}w = \begin{bmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{yx} & F_{yy} & F_{yz} \\ F_{zx} & F_{zy} & F_{zz} \end{bmatrix}$ at any point at which all the second partial derivatives of F exist.

Theorem 3.1. *Let $F : \mathbb{R}^2 \to \mathbb{R}$ be a nonzero function and $(x_0, y_0) \in \mathbb{R}^2$ a regular point. Suppose that the second partial derivatives of F at (x_0, y_0) exist and further $F_{xy} = F_{yx}$ at this point. Let $\mathbf{H}F$ and $\mathbf{H}G$ be the Hessian matrices of F and F^2 respectively (we assume that $G = F^2$) and let k be the curvature of $G(x, y) = F^2(x, y) = c \neq 0$ at (x_0, y_0). Then we have*

$$k = \frac{|\mathbf{H}G| - 4F^2|\mathbf{H}F|}{4F\big(F_x^2 + F_y^2\big)^{\frac{3}{2}}}.$$

Proof. For simplicity, we do the proof without (x_0, y_0). The partial derivatives of $G = F^2$ are as follows:

$$\begin{aligned} &G_x = 2FF_x, \qquad G_{xx} = 2({F_x}^2 + FF_{xx}),\\ &G_y = 2FF_y, \qquad G_{yy} = 2(F_y^2 + FF_{yy}), \text{ and } G_{xy}^2 = 4(F_xF_y + FF_{xy})^2. \end{aligned}$$

Therefore,

$$\begin{aligned} |\mathbf{H}G| &= G_{xx}G_{yy} - G_{xy}^2 = 4\big({F_x}^2 + FF_{xx}\big)\big(F_y^2 + FF_{yy}\big) - 4\big(F_xF_y + FF_{xy}\big)^2\\ &= 4\Big[F_x^2F_y^2 + FF_x^2F_{yy} + FF_y^2F_{xx} + F^2F_{xx}F_{yy} - F_x^2F_y^2 - 2FF_xF_yF_{xy}\\ &\quad - F^2F_{xy}^2\Big]\\ &= 4\Big[F^2\big(F_{xx}F_{yy} - F_{xy}^2\big) + F\big(F_x^2F_{yy} - 2F_xF_yF_{xy} + F^2yF_{xx}\big)\Big]\\ &= 4\Big[F^2|\mathbf{H}F| + F\big(F_x^2F_{yy} - 2F_xF_yF_{xy} + F^2yF_{xx}\big)\Big]. \end{aligned}$$

In view of Proposition 3.1, we have

$$\begin{aligned} |\mathbf{H}G| &= 4\Big[F^2|\mathbf{H}F| + F\big(F_x^2F_{yy} - 2F_xF_yF_{xy} + F^2yF_{xx}\big)\Big]\\ &= 4\Big[F^2|\mathbf{H}F| + Fk\big(F_x^2 + F_y^2\big)^{\frac{3}{2}}\Big] \end{aligned}$$

Therefore,

$$k = \frac{|\mathbf{H}G| - 4F^2|\mathbf{H}F|}{4F\big(F_x^2 + F_y^2\big)^{\frac{3}{2}}},$$

and we are done. □

The next result is a similar consequence for the implicit surface.

Theorem 3.2. *Let $F : \mathbb{R}^3 \to \mathbb{R}$ be a nonzero function and $(x_0, y_0, z_0) \in \mathbb{R}^3$ a regular point. Suppose that the second partial derivatives of F at (x_0, y_0, z_0) exist* *and further the mixed partial derivatives at this point are equivalent. If k is the curvature of $G(x, y, z) = F^2(x, y, z) = c \neq 0$ at (x_0, y_0, z_0), then we have*

$$k = \frac{|\mathbf{H}G| - 8F^3|\mathbf{H}F|}{8F^2\big(F_x^2 + F_y^2 + F_z^2\big)^{\frac{3}{2}}},$$

where $\mathbf{H}F$ and $\mathbf{H}G$ are the Hessian matrices of F and F^2 respectively (we assume that $G = F^2$).

Proof. As we did in the previous theorem, the proof is done without (x_0, y_0, z_0). Let $K = \begin{bmatrix} F_{xx} & F_{xy} & F_{xz} & F_x \\ F_{xy} & F_{yy} & F_{yz} & F_y \\ F_{xz} & F_{yz} & F_{zz} & F_z \\ F_x & F_y & F_z & 0 \end{bmatrix}$. It is known that the curvature k of the implicit surface $F(x, y, z) = 0$ is $k = |K|$ at every regular point in which the second partial derivatives of F exist. We first calculate the partial derivatives of G and in continued we obtain determinant of $\mathbf{H}G$.

$$\begin{aligned}
G_x &= 2FF_x, & G_{xx} &= 2({F_x}^2 + FF_{xx}), & G_{xy}^2 &= 4(F_xF_y + FF_{xy})^2 \\
G_y &= 2FF_y, & G_{yy} &= 2(F_y^2 + FF_{yy}), & G_{xz}^2 &= 4(F_xF_z + FF_{xz})^2 \\
G_z &= 2FF_z, & G_{zz} &= 2(F_z^2 + FF_{zz}), & G_{yz}^2 &= 4(F_yF_z + FF_{yz})^2.
\end{aligned}$$

Recall that the Hessian matrices of F and G are

$$\mathbf{H}F = \begin{bmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{xy} & F_{yy} & F_{yz} \\ F_{xz} & F_{yz} & F_{zz} \end{bmatrix}, \text{ and } \mathbf{H}G = \begin{bmatrix} G_{xx} & G_{xy} & G_{xz} \\ G_{xy} & G_{yy} & G_{yz} \\ G_{xz} & G_{yz} & G_{zz} \end{bmatrix}.$$

Here, we compute the determinant of $\mathbf{H}G$.

$$\begin{aligned}
1/8|\mathbf{H}G| &= F_{xx}\big(F_{yy}F_{zz} - F_{yz}^2\big) - F_{xy}\big(F_{xy}F_{zz} - F_{xz}F_{yz}\big) \\
&+ F_{xz}\big(F_{xy}F_{yz} - F_{xz}F_{yy}\big) \\
&= F_{xx}F_{yy}F_{zz} - F_{xx}F_{yz}^2 - F_{yy}F_{xz}^2 - F_{zz}F_{xy}^2 + 2F_{xy}F_{yz}F_{xz} \\
&= \big(F_x^2 + FF_{xx}\big)\big(F_y^2 + FF_{yy}\big)\big(F_z^2 + FF_{zz}\big) \\
&- \big(F_x^2 + FF_{xx}\big)\big(F_yF_z + FF_{yz}\big)^2 \\
&- \big(F_y^2 + FF_{yy}\big)\big(F_xF_z + FF_{xz}\big)^2 - \big(F_z^2 + FF_{zz}\big)\big(F_xF_y + FF_{xy}\big)^2 \\
&+ \big(F_xF_z + FF_{xz}\big)\big(F_yF_z + FF_{yz}\big)\big(F_xF_y + FF_{xy}\big) \\
&= F^3\Big[F_{xx}F_{yy}F_{zz} - F_{xx}F_{yz}^2 - F_{yy}F_{xz}^2 - F_{xx}F_{xy}^2 + 2F_{xy}F_{yz}F_{xz}\Big] \\
&+ F^2\Big[F_{xx}F_{yy}F_z^2 + F_{xx}F_{zz}F_y^2 + F_{yy}F_{zz}F_x^2 - 2F_{xy}F_{xz}F_yF_z \\
&- 2F_{xy}F_{yz}F_xF_z - 2F_{xz}F_{yz}F_xF_y + F_{xy}^2F_z^2 + F_{xz}^2F_y^2 + F_{yz}^2F_x^2\Big] + F\big[0\big].
\end{aligned}$$

Therefore, we have $1/8|\mathbf{H}G| = F^3|\mathbf{H}F| + F^2k(F_x^2 + F_y^2 + F_z^2)^{\frac{3}{2}}$. So the result is obtained, i.e.,

$$k = \frac{|\mathbf{H}G| - 8F^3|\mathbf{H}F|}{8F^2\big(F_x^2 + F_y^2 + F_z^2\big)^{\frac{3}{2}}}.$$

□

Theorem 3.3. *Let f, g be nonzero real-valued functions on $\mathbb{R}$, $a, b \in \mathbb{R}$ and $F : \mathbb{R}^2 \to \mathbb{R}$ defined by $F(x,y) = \int_a^b \sqrt{x^2 f^2(t) + y^2 g^2(t)}d(t)$. Then*

(i) The curvature of $F(x,y) = c$, where $c > 0$ at any point of the curve is positive multiple of c^2.

(ii) $tr(\mathbf{H}F) = F_{xx} + F_{yy} \geq 0$.

Proof. (i). First, we note that $F \geq 0$. The surface F meets the plane $z = 0$ at the origin only. But the intersection of F with the plane $z = c$ (where $c > 0$) is the curve $F(x,y) = c$. Here the partial derivatives of F are calculated (see [Rudin, 1976, Theorem 9.42]).

$$F_x = \int_a^b \frac{xf^2(t)}{\sqrt{x^2f^2(t) + y^2g^2(t)}}d(t), \qquad F_y = \int_a^b \frac{yg^2(t)}{\sqrt{x^2f^2(t) + y^2g^2(t)}}d(t),$$

$$F_{xx} = \int_a^b \frac{y^2f^2(t)g^2(t)}{\big(x^2f^2(t) + y^2g^2(t)\big)^{\frac{3}{2}}}d(t), \qquad F_{yy} = \int_a^b \frac{x^2f^2(t)g^2(t)}{\big(x^2f^2(t) + y^2g^2(t)\big)^{\frac{3}{2}}}d(t),$$

and

$$F_{xy} = -\int_a^b \frac{xyf^2(t)g^2(t)}{\big(x^2f^2(t) + y^2g^2(t)\big)^{\frac{3}{2}}}d(t) = F_{yx}.$$

Let us put $\varphi := \sqrt{x^2f^2(t) + y^2g^2(t)}$. For the simplicity, we set

$$F_x = \int \frac{xf^2}{\varphi}, \quad F_y = \int \frac{yg^2}{\varphi}, \text{ and so on}\ldots$$

By formula of the curvature k in Proposition 3.1, we obtain

$$\begin{aligned}
k &= \frac{1}{\left(F_x^2+F_y^2\right)^{\frac{3}{2}}}\Big[\big(y^2\int\frac{f^2g^2}{\varphi^3}\big)\big(y\int\frac{g^2}{\varphi}\big)^2+2\int\frac{xyf^2g^2}{\varphi^3}\int\frac{xf^2}{\varphi}\int\frac{yg^2}{\varphi} \\
&\quad+\big(x^2\int\frac{f^2g^2}{\varphi^3}\big)\big(x\int\frac{f^2}{\varphi}\big)^2\Big] \\
&= \frac{\int\frac{f^2g^2}{\varphi^3}}{\left(F_x^2+F_y^2\right)^{\frac{3}{2}}}\Big[y^4\big(\int\frac{g^2}{\varphi}\big)^2+2x^2y^2\int\frac{f^2}{\varphi}\int\frac{g^2}{\varphi}+x^4\big(\int\frac{f^2}{\varphi}\big)^2\Big] \\
&= \frac{\int\frac{f^2g^2}{\varphi^3}}{\left(F_x^2+F_y^2\right)^{\frac{3}{2}}}\Big[\int\frac{x^2f^2}{\varphi}+\int\frac{y^2g^2}{\varphi}\Big]^2 \\
&= \frac{\int\frac{f^2g^2}{\varphi^3}}{\left(F_x^2+F_y^2\right)^{\frac{3}{2}}}\Big[\int\frac{x^2f^2+y^2g^2}{\varphi}\Big]^2 \\
&= \frac{\int\frac{f^2g^2}{\varphi^3}}{\left(F_x^2+F_y^2\right)^{\frac{3}{2}}}\Big[\int\varphi\Big]^2 \\
&= \frac{\int\frac{f^2g^2}{\varphi^3}}{\left(F_x^2+F_y^2\right)^{\frac{3}{2}}}F^2(x,y).
\end{aligned}$$

Hence, we observe that the curvature of $F(x,y)=c$ at (x_0,y_0) is a positive multiple of $F^2(x_0,y_0)=c^2$, and we are done.
(ii). Since

$$\frac{f^2g^2(x^2+y^2)}{\varphi^3}\geq 0,$$

it is clear that $F_{xx}+F_{yy}\geq 0$. So the result holds. □

Lemma 3.1. *Let $F:\mathbb{R}^2\to\mathbb{R}$ be a homogeneous function of degree one. Suppose that the second derivatives of F at $(a,b)\in\mathbb{R}^2$ exist. Moreover, $F_{xy}=F_{yx}$ at this point. Then*

(i) $|\mathbf{H}F|_{(a,b)}=0$.

(ii) The eigenvalues of $\mathbf{H}F$ *are* 0 *and tr*$(\mathbf{H}F)$ *at* (a,b).

Proof. (i). First, we note that $F(\lambda x,\lambda y)=\lambda F(x,y)$, for all $(x,y)\in\mathbb{R}^2$ and $\lambda\in\mathbb{R}$. Also, we remind the reader of the following fact, which is known as *Euler's property*,

$$xF_x+yF_y=F(x,y).$$

Therefore,

$$xF_{xx} + F_x + yF_{xy} = F_x, \text{ and } xF_{xy} + F_y + yF_{yy} = F_y.$$

Consequently, $xF_{xx} = -yF_{xy}$ and $xF_{xy} = -yF_{yy}$. Now, consider the Hessian matrix $\mathbf{H}F = \begin{bmatrix} Fxx & F_{xy} \\ F_{xy} & F_{yy} \end{bmatrix}$ of F. For the point $(0, b)$, where $b \neq 0$, we have $F_{yy}(0, b) = 0 = F_{xy}(0, b)$. This implies that $|\mathbf{H}F| = 0$. Also, considering the point $(a, 0)$, where $a \neq 0$ gives $F_{xy}(a, 0) = 0 = F_{xx}(a, 0)$, this again yields $|\mathbf{H}F| = 0$. Now, let (a, b) such that $a \neq 0$ and $b \neq 0$. Then $F_{xx}(a, b) = \frac{-b}{a}F_{xy}(a, b)$ and $F_{yy}(a, b) = \frac{-a}{b}F_{xy}(a, b)$. Hence, $|\mathbf{H}F| = 0$. So we always have $|\mathbf{H}F| = 0$. The proof of (i) is now complete. (ii). Recall that the characteristic equation of $\mathbf{H}F$ is

$$\lambda^2 - (tr(\mathbf{H}F) = F_{xx} + F_{yy})\lambda + (|\mathbf{H}F| = F_{xx}F_{yy} - F_{xy}^2) = 0.$$

So $\lambda^2 - (F_{xx} + F_{yy})\lambda = 0$. Therefore, $\lambda = 0$ or $\lambda = \text{tr}(\mathbf{H}F)$, and we are done. □

Proposition 3.2. *Let f, g be nonzero real-valued functions on $\mathbb{R}$ and $F : \mathbb{R}^2 \to \mathbb{R}$ defined by $F(x, y) = \int_a^b \sqrt{x^2 f^2(t) + y^2 g^2(t)}dt$ and let $G(x, y) = F^2(x, y)$. Then the eigenvalues of $\mathbf{H}F$ and $\mathbf{H}G$ at any point except the origin are nonnegative. (In fact, the eigenvalues of $\mathbf{H}F$ are zero and tr$(\mathbf{H}F)$ at that point).*

Proof. We observe that F is a homogeneous function of degree one. So Lemma 3.1 and Theorem 3.3 (ii) yield the result. For the matrix $\mathbf{H}G$, we look to the Theorem 3.1. Since, $F^2|\mathbf{H}F| = 0$, we have

$$|\mathbf{H}G| = 4Fk\big(F_x^2 + F_y^2\big)^{\frac{3}{2}}.$$

We notice that $F, k \geq 0$ gives $|\mathbf{H}G| \geq 0$. On the other hand, tr$(\mathbf{H}G) = G_{xx} + G_{yy} \geq 0$. Therefore, the roots of $\lambda^2 - \text{tr}(\mathbf{H}G)\lambda + |\mathbf{H}G| = 0$, which are the eigenvalues of $\mathbf{H}G$, are nonnegative. The proof is finished. □

In the following result, we present a norm on $\mathbb{R}^2$ which is an elliptic integral of the second kind.

Corollary 3.1. *Let $f(t) = \cos t$, $g(t) = \sin t$ and let $F : \mathbb{R}^2 \to \mathbb{R}$ is given by*

$$F(x, y) = \int_0^{\frac{\pi}{2}} \sqrt{x^2 \cos^2 t + y^2 \sin^2 t}dt.$$

Then the following statements hold.

(i) The eigenvalues of $\mathbf{H}F$ and $\mathbf{H}G$, where $G = F^2$ at every point except the origin are nonnegative.

(ii) $F(x,y)$ is an elliptic integral of the second kind.

Proof. (i). It follows from Proposition 3.2. (ii). Notice that

$$F(x,y)=\int_0^{\frac{\pi}{2}}\sqrt{x^2(1-\sin^2\theta)+y^2\sin^2\theta}d\theta=|x|\int_0^{\frac{\pi}{2}}\sqrt{1-k^2\sin^2\theta}d\theta,$$

where $k=\frac{\sqrt{x^2-y^2}}{|x|}$ and $|x|\geq|y|$. So this gives $F(x,y)$ is an elliptic integral of the second kind and we are done. □

Corollary 3.2. *There are ordered pairs (x,y) with rational coordinates (other than the origin) which satisfy the inequality $\int_0^{\frac{\pi}{2}}\sqrt{x^2\cos^2\theta+y^2\sin^2\theta}d\theta\leq r$, when $0<r\in\mathbb{Q}$. Also, if $r\notin\mathbb{Q}$ then (x,y) has irrational coordinates.*

Proof. It is sufficient to take the pairs $(r,0),(0,r),(-r,0)$ and $(0,-r)$. □

We end this article with the next results.

Proposition 3.3. *Let $0\leq x,y\in\mathbb{R}$. Then*

$$\int_0^{\frac{\pi}{2}}\sqrt{x^2\cos^2 t+y^2\sin^2 t}dt\leq x+y.$$

Proof. First, note that

$$x^2\cos^2 t+y^2\sin^2 t=(x\cos t+y\sin t)^2-2xy\sin t\cos t,$$

and take $0\leq\phi\leq\frac{\pi}{2}$ such that $\tan\phi=\frac{y}{x}$ (if $x>0$). Now,

$$\begin{aligned}(x\cos t+y\sin t)^2&=x^2(\cos t+\frac{y}{x}\sin t)^2=x^2(\cos t+\frac{\sin\phi}{\cos\phi}\sin t)^2\\&=\frac{x^2(\cos t\cos\phi+\sin t\sin\phi)^2}{\cos^2\phi}=\frac{x^2\cos^2(t-\phi)}{\cos^2\phi}\\&=(x^2+y^2)\cos^2(t-\phi)\ (\text{note, }\cos^2\phi=\frac{x^2}{x^2+y^2}).\end{aligned}$$

Hence, $x^2\cos^2 t+y^2\sin^2 t\leq(x^2+y^2)\cos^2(t-\phi)$. Therefore,

$$\begin{aligned}\int_0^{\frac{\pi}{2}}\sqrt{x^2\cos^2 t+y^2\sin^2 t}dt&\leq\int_0^{\frac{\pi}{2}}\sqrt{(x^2+y^2)\cos^2(t-\phi)}dt\\&=\sqrt{x^2+y^2}\int_0^{\frac{\pi}{2}}|\cos(t-\phi)|dt\\&=\sqrt{x^2+y^2}\int_{-\phi}^{\frac{\pi}{2}-\phi}\cos Tdt\quad(T=t-\phi)\\&=x+y.\end{aligned}$$

□

Remark 3.1. *We find* $4\int_0^{\frac{\pi}{2}} \sqrt{x^2\cos^2 t + y^2\sin^2 t}dt \leq 2(2x+2y)$. *The left phrase is the length of the ellipse* $x' = x\cos t$ *and* $y' = y\sin t$*, while* $2x$ *and* $2y$ *are the major axis and minor axis of this ellipse.*

References

R. Goldman. Curvature formulas for implicit curves and surfaces. *Computer Aided Geometric Design*, 22(7):632–658, 2005.

C. G. Lekkerkerker. *Geometry of numbers* North-Holland Publishing Company, Amsterdam, 1969.

W. Rudin. *Principles of Mathematical Analysis*, McGraw-Hill, International Book Company, Ltd, 1976.

C. L. Siegel. *Lecturcs on the geometry of numbers* Springer-Verlag Berlin Heidelberg, 1989.

S. Willard. *General Topology*, Addison-Wesley, 1970.

Ratio Mathematica Volume 39, 2020, pp. 69-77

Parameter estimation of p-dimensional Rayleigh distribution under weighted loss function

Arun Kumar Rao*
Himanshu Pandey †

Abstract

In this paper, p-dimensional Rayleigh distribution is considered. The classical maximum likelihood estimator has been obtained. Bayesian method of estimation is employed in order to estimate the scale parameter of p-dimensional Rayleigh distribution by using quasi and inverted gamma priors. The Bayes estimators of the scale parameter have been obtained under squared error and weighted loss functions.

Keywords: Bayesian method, p-dimensional Rayleigh distribution, quasi and inverted gamma priors, squared error and weighted loss functions.

2010 AMS subject classification: 60E05, 62E15, 62H10, 62H12.‡

---

* Department of Statistics, Maharana Pratap P.G. College, Jungle Dhusan, Gorakhpur, INDIA
E-mail: arunrao1972@gmail.com

† Department of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur, INDIA
E-mail: himanshu_pandey62@yahoo.com

‡ Received on October 26th, 2020. Accepted on December 17th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.561. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

The probability density function (pdf) of p-dimensional Rayleigh distribution is given by

$$f\left(x;\theta\right)=\frac{2}{\Gamma\left(p/2\right)}\frac{\theta^{-p/2}}{x^{p-1}}e^{-\left(x^{2}\right)/\theta} \quad ; x\geq 0,\ \theta>0. \tag{1}$$

(Cohen and Whitten [1]).

The distribution with pdf (1), in which p=1, sometimes called the folded Gaussian, the folded normal, or the half normal distribution. With p=2, the pdf of (1) is reduced to two-dimensional Rayleigh distribution. With p=3, the pdf of (1) is reduced to Maxwell-Boltzmann distribution. Let $x_1, x_2, \ldots\ldots\ldots, x_n$ be a random sample of size n having probability density function (1), then the likelihood function of (1) is given by (Rao and Pandey [2])

$$f\left(\underline{x};\theta\right)=\left(\frac{2}{\Gamma\left(p/2\right)}\right)^{n}\theta^{-np/2}\left(\prod_{i=1}^{n}x_i^{p-1}\right)e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i^{2}}$$

(2)

The log likelihood function is given by

$$log\ f\left(\underline{x};\theta\right)=n\,log\,2-n\,log\,\Gamma\left(p/2\right)-\frac{np}{2}log\,\theta+log\prod_{i=1}^{n}x_i^{p-1}-\frac{1}{\theta}\sum_{i=1}^{n}x_i^{2} \tag{3}$$

Differentiating (3) with respect to θ and equating to zero, we get

$$\hat{\theta}=\frac{2\sum_{i=1}^{n}x_i^{2}}{np} \tag{4}$$

2 Bayesian Method of Estimation

In Bayesian analysis the fundamental problem is that of the choice of prior distribution g (θ) and a loss function $L\left(\hat{\theta},\theta\right)$. The squared error loss function for the scale parameter θ are defined as

$$L\left(\hat{\theta},\theta\right)=\left(\hat{\theta}-\theta\right)^{2} \tag{5}$$

The Bayes estimator under the above loss function, say, $\hat{\theta}_s$ is the posterior mean, i.e,

$$\hat{\theta}_S = E(\theta) \tag{6}$$

This loss function is often used because it does not lead to extensive numerical computations but several authors (Zellner [3], Basu and Ebrahimi [4]) have recognized that the inappropriateness of using symmetric loss function. J.G.Norstrom [5] introduced an alternative asymmetric precautionary loss function. and also presented a general class of precautionary loss functions with quadratic loss function as a special case. Weighted loss function (Ahamad et al. [6]) is given a

$$L\left(\hat{\theta},\theta\right) = \frac{\left(\hat{\theta}-\theta\right)^2}{\theta} \tag{7}$$

The Bayes estimator under weighted loss function is denoted by $\hat{\theta}_W$ and is obtained as

$$\hat{\theta}_W = \left[E\left(\frac{1}{\theta}\right)\right]^{-1} \tag{8}$$

Let us consider two prior distributions of θ to obtain the Bayes estimators.

(i) Quasi-prior: For the situation where the experimenter has no prior information about the parameter θ, one may use the quasi density as given by

$$g_1(\theta) = \frac{1}{\theta^d} \ ; \theta > 0,\ d \geq 0, \tag{9}$$

where $d = 0$ leads to a diffuse prior and $d = 1$, a non-informative prior.

(ii) Inverted gamma prior: The most widely used prior distribution of θ is the inverted gamma distribution with parameters α and $\beta(>0)$ with probability density function given by

$$g_2(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-(\alpha+1)}e^{-\beta/\theta} \ ; \theta > 0. \tag{10}$$

3 Bayes Estimators under $g_1(\theta)$

The posterior density of θ under $g_1(\theta)$, on using (2), is given by

$$f(\theta/\underline{x}) = \frac{\left(\frac{2}{\Gamma(p/2)}\right)^n \theta^{-np/2}\left(\prod_{i=1}^n x_i^{p-1}\right) e^{-\frac{1}{\theta}\sum_{i=1}^n x_i^2}\,\theta^{-d}}{\int_0^\infty \left(\frac{2}{\Gamma(p/2)}\right)^n \theta^{-np/2}\left(\prod_{i=1}^n x_i^{p-1}\right) e^{-\frac{1}{\theta}\sum_{i=1}^n x_i^2}\,\theta^{-d}\,d\theta}$$

$$= \frac{\theta^{-\left(\frac{np}{2}+d\right)} e^{-\frac{1}{\theta}\sum_{i=1}^n x_i^2}}{\int_0^\infty \theta^{-\left(\frac{np}{2}+d\right)} e^{-\frac{1}{\theta}\sum_{i=1}^n x_i^2}\,d\theta}$$

$$= \frac{\left(\sum_{i=1}^n x_i^2\right)^{\left(\frac{np}{2}+d-1\right)}}{\Gamma\left(\frac{np}{2}+d-1\right)}\,\theta^{-\left(\frac{np}{2}+d\right)} e^{-\frac{1}{\theta}\sum_{i=1}^n x_i^2} \tag{11}$$

Theorem 1. Assuming the squared error loss function, the Bayes estimate of the scale parameter θ, is of the form

$$\hat{\theta}_S = \frac{\sum_{i=1}^n x_i^2}{\left(\frac{np}{2}+d-2\right)} \tag{12}$$

Proof. From equation (6), on using (11),

$$\hat{\theta}_S = E(\theta) = \int \theta f(\theta/\underline{x})\,d\theta$$

$$= \frac{\left(\sum_{i=1}^n x_i^2\right)^{\left(\frac{np}{2}+d-1\right)}}{\Gamma\left(\frac{np}{2}+d-1\right)}\int_0^\infty \theta^{-\left(\frac{np}{2}+d-1\right)} e^{-\frac{1}{\theta}\sum_{i=1}^n x_i^2}\,d\theta$$

$$= \frac{\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+d-1}}{\Gamma\left(\frac{np}{2}+d-1\right)} \frac{\Gamma\left(\frac{np}{2}+d-2\right)}{\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+d-2}}$$

or, $\hat{\theta}_S = \dfrac{\sum_{i=1}^{n} x_i^2}{\left(\frac{np}{2}+d-2\right)}$.

Theorem 2. Assuming the weighted loss function, the Bayes estimate of the scale parameter θ, is of the form

$$\hat{\theta}_W = \frac{\sum_{i=1}^{n} x_i^2}{\left(\frac{np}{2}+d-1\right)} \tag{13}$$

Proof. From equation (8), on using (11),

$$\hat{\theta}_W = \left[E\left(\frac{1}{\theta}\right)\right]^{-1} = \left[\int \frac{1}{\theta} f\left(\theta/\underline{x}\right) d\theta\right]^{-1}$$

$$= \left[\frac{\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+d-1}}{\Gamma\left(\frac{np}{2}+d-1\right)} \int_0^{\infty} \theta^{-\left(\frac{np}{2}+d+1\right)} e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i^2} d\theta\right]^{-1}$$

$$= \left[\frac{\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+d-1}}{\Gamma\left(\frac{np}{2}+d-1\right)} \frac{\Gamma\left(\frac{np}{2}+d\right)}{\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+d}}\right]^{-1}$$

$$= \left[\frac{\frac{np}{2} + d - 1}{\sum_{i=1}^{n} x_i^2} \right]^{-1}$$

$$\Rightarrow \qquad \hat{\theta}_W = \frac{\sum_{i=1}^{n} x_i^2}{\left(\frac{np}{2} + d - 1 \right)}$$

4 Bayes Estimators under $g_2(\theta)$

Under $g_2(\theta)$, the posterior density of θ, using equation (2), is obtained as

$$f(\theta/\underline{x}) = \frac{\left(\frac{2}{\Gamma(p/2)} \right)^n \theta^{-np/2} \left(\prod_{i=1}^{n} x_i^{p-1} \right) e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i^2} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} e^{-\beta/\theta}}{\int_0^\infty \left(\frac{2}{\Gamma(p/2)} \right)^n \theta^{-np/2} \left(\prod_{i=1}^{n} x_i^{p-1} \right) e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i^2} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} e^{-\beta/\theta} d\theta}$$

$$= \frac{\theta^{-\left(\frac{np}{2}+\alpha+1\right)} e^{-\frac{1}{\theta}\left(\beta + \sum_{i=1}^{n} x_i^2\right)}}{\int_0^\infty \theta^{-\left(\frac{np}{2}+\alpha+1\right)} e^{-\frac{1}{\theta}\left(\beta + \sum_{i=1}^{n} x_i^2\right)} d\theta}$$

$$= \frac{\theta^{-\left(\frac{np}{2}+\alpha+1\right)} e^{-\frac{1}{\theta}\left(\beta + \sum_{i=1}^{n} x_i^2\right)}}{\Gamma\left(\frac{np}{2} + \alpha \right) \Big/ \left(\beta + \sum_{i=1}^{n} x_i^2 \right)^{\frac{np}{2}+\alpha}}$$

$$= \frac{\left(\beta + \sum_{i=1}^{n} x_i^2 \right)^{\frac{np}{2}+\alpha}}{\Gamma\left(\frac{np}{2} + \alpha \right)} \theta^{-\left(\frac{np}{2}+\alpha+1\right)} e^{-\frac{1}{\theta}\left(\beta + \sum_{i=1}^{n} x_i^2\right)} \qquad (14)$$

Theorem 3. Assuming the squared error loss function, the Bayes estimate of the scale parameter θ, is of the form

$$\hat{\theta}_S = \frac{\beta + \sum_{i=1}^{n} x_i^2}{\frac{np}{2} + \alpha - 1} \tag{15}$$

Proof. From equation (6), on using (14),

$$\hat{\theta}_S = E(\theta) = \int \theta f(\theta / \underline{x})\, d\theta$$

$$= \frac{\left(\beta + \sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+\alpha}}{\Gamma\left(\frac{np}{2}+\alpha\right)} \int_0^{\infty} \theta^{-\left(\frac{np}{2}+\alpha\right)} e^{-\frac{1}{\theta}\left(\beta + \sum_{i=1}^{n} x_i^2\right)} d\theta$$

$$= \frac{\left(\beta + \sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+\alpha}}{\Gamma\left(\frac{np}{2}+\alpha\right)} \frac{\Gamma\left(\frac{np}{2}+\alpha-1\right)}{\left(\beta + \sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+\alpha-1}}$$

or, $$\hat{\theta}_S = \frac{\beta + \sum_{i=1}^{n} x_i^2}{\frac{np}{2} + \alpha - 1}.$$

Theorem 4. Assuming the weighted loss function, the Bayes estimate of the scale parameter θ, is of the form

$$\hat{\theta}_W = \frac{\beta + \sum_{i=1}^{n} x_i^2}{\frac{np}{2} + \alpha} \tag{16}$$

Proof. From equation (8), on using (14),

$$\hat{\theta}_W = \left[E\left(\frac{1}{\theta}\right)\right]^{-1} = \left[\int \frac{1}{\theta} f(\theta / \underline{x})\, d\theta\right]^{-1}$$

$$= \left[\frac{\left(\beta + \sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+\alpha}}{\Gamma\left(\frac{np}{2}+\alpha\right)} \int_0^\infty \theta^{-\left(\frac{np}{2}+\alpha+2\right)} e^{-\frac{1}{\theta}\left(\beta+\sum_{i=1}^{n} x_i^2\right)} d\theta \right]^{-1}$$

$$= \left[\frac{\left(\beta + \sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+\alpha}}{\Gamma\left(\frac{np}{2}+\alpha\right)} \frac{\Gamma\left(\frac{np}{2}+\alpha+1\right)}{\left(\beta+\sum_{i=1}^{n} x_i^2\right)^{\frac{np}{2}+\alpha+1}} \right]^{-1}$$

$$= \left[\frac{\frac{np}{2}+\alpha}{\left(\beta+\sum_{i=1}^{n} x_i^2\right)} \right]^{-1}$$

or, $$\hat{\theta}_W = \frac{\beta+\sum_{i=1}^{n} x_i^2}{\frac{np}{2}+\alpha} .$$

5 Conclusion

In this paper, we have obtained a number of estimators of parameter. In equation (4) we have obtained the maximum likelihood estimator of the parameter. In equation (12) and (13) we have obtained the Bayes estimators under squared error and weighted loss function using quasi prior. In equation (15) and (16) we have obtained the Bayes estimators under squared error and weighted loss function using inverted gamma prior. In the above equation, it is clear that the Bayes estimators depend upon the parameters of the prior distribution.

References

[1] Cohen,A.C.,Whitten,B.J., "*Parameter Estimation in Reliability and Life Span Models",* Dekker, New York, ISBN 0-8247-7980-0, 182. 1988.

[2] Rao, A.K. and Pandey, H., "Bayesian estimation of scale parameterof p-dimensional Rayleigh distribution using precautionary loss function", Varahmihir Journal of Mathematical Sciences, Vol. 8, No. 1, 1-8. 2008.

[3] Zellner, A., "Bayesian estimation and prediction using asymmetric loss functions," Jour. Amer. Stat. Assoc., 91, 446-451. 1986.

[4] Basu, A. P. and Ebrahimi, N., "Bayesian approach to life testing and reliability estimation using asymmetric loss function," Jour. Stat. Plann. Infer., 29, 21-31. 1991.

[5] Norstrom, J. G., "The use of precautionary loss functions in Risk Analysis," IEEE Trans. Reliab., 45(3), 400-403. 1996.

[6] Ahmad, S. P., et al., "Bayesian Estimation for the Class of Life-Time Distributions under Different Loss Functions," Statistics and Applications, Vol. 14, Nos. 1&2, pp. 75-91. 2016.

Ratio Mathematica Volume 39, 2020, pp. 79-110

On some computational and applications of finite fields

Jean Pierre Muhirwa*

Abstract

Finite field is a wide topic in mathematics. Consequently, none can talk about the whole contents of finite fields. That is why this research focuses on small content of finite fields such as polynomials computational, ring of integers modulo p where p is prime or a power of prime. Most of the times, books which talk about finite fields are rarely to be found, therefore one can know how arithmetic computational on small finite fields works and be able to extend to the higher order. This means how integer and polynomial arithmetic operations are done for $\mathbb{Z}_p$ such as addition, subtraction, division and multiplication in $\mathbb{Z}_p$ followed by reduction of p (modulo p). Only addition and multiplication arithmetic operations are considered for a small range of finite fields ($\mathbb{Z}_2 - \mathbb{Z}_{17}$). With polynomials, one can learn how arithmetic computational through polynomials over finite fields are performed as their coefficients are drawn from finite fields. The paper includes also construction of polynomials over finite fields as an extension of finite fields with polynomials i.e $F_q[x]/f(x)$, where $f(x)$ is irreducible over F_q. From the past decades, many researchers complained about the applications of some topics in pure mathematics and therefore the finite fields play more important role in coding theory, such as error-coding detection and error-correction as well as cyclic codes. Hence, this paper shows these applications.

Keywords: Finite Fields; Error-detection; Error-correction; Coding; Decoding; Codewords; Cosets; Syndromes.[1]

*University of Rwanda, College of Science and Technology, School of Science, Department of Mathematics, Kigali, Rwanda; muhijeapi@gmail.com.

[1]Received on January 20th, 2020. Accepted on June 19th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.521. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

The structure of this research paper includes the introductory part where some preliminary properties of set theory, group theory, ring theory and fields theory are discussed. In reality we can not know what is a field without defining a group and a ring since the field is a special case of the ring. Apart from introductory, the second section consist of computational in the first seven finite fields. The third, the fourth and the fifth parts of this paper discuss and compare the usual polynomial arithmetic computational and the finite field polynomial computational. The sixth part of this paper explains some of the applications of finite fields with the typical examples in coding and decoding theories, the seventh section gives the conclusion of the research paper while the last part acknowledges the financial support received from the Eastern Africa Universities Mathematics Programme-International Science Programme, University of Rwanda Node (EAUMP-ISP, UR-Node).

1.1 Preliminaries

Definition 1.1. *A set is a collection of distinct objects, considered as an object in it own rights. Sets are the one of the most fundamental concepts of mathematics.*

Example 1.1. *The set $\mathbb{R}$, denote the set of all real numbers, and this set includes rational numbers and irrational numbers (example π, $\sqrt{2}$, and e) $\mathbb{Z}$, denote the set of all integers for both sign (negative and positive).*

Definition 1.2. ***Group Theory**, a set $\mathbb{R}$ together with a binary operation is called a group if it satisfies the conditions such that closure, associative, admits identity element and inverse element under the operation within the elements of $\mathbb{R}$.*

Definition 1.3. ***Abelian Group**, a set $\mathbb{R}$ is an abelian group if it is a group for which commutative law within an operation together with $\mathbb{R}$ to the elements of $\mathbb{R}$ is verified.*

Definition 1.4. ***Ring Theory**, a set $\mathbb{R}$ together with two binary operations (addition and multiplication) on the elements of $\mathbb{R}$ is called a ring if the following conditions are satisfied:*

1. *($\mathbb{R}$, +) is an abelian group.*
2. *Associative law for multiplication and distributive law are also satisfied.*

Definition 1.5. ***Commutative Ring**, a commutative ring is a ring for which the multiplication is commutative.*

Definition 1.6. ***Commutative Ring with Unity***, *a commutative ring with unity is a ring for which there exists a non-zero multiplicative identity element.*

Example 1.2. *The set of integers $\mathbb{Z}$ is commutative ring with 1 as a multiplicative identity element.*

Definition 1.7. ***Field***, *a field is a commutative ring with unity and for which every non-zero element of that commutative ring is invertible.*

Example 1.3. *In the set of rational numbers, $\mathbb{Q}$, every non-zero element has its inverse i.e (Every non-zero element is invertible).*

Definition 1.8. ***Finite Field***, *a finite field is a field with a finite number of elements.*

Example 1.4. *Consider the set of integers modulo p ($\mathbb{Z}_p$), where p is prime integers). This set consists of $p - 1$ elements and all non-zero elements of this set are invertible.*

Definition 1.9. ***Galois Group***, *the Galois group of an extension of fields F/K, is the set of all automorphisms obtained by fixing the elements of K.*

Definition 1.10. ***Codewords***, *codewords are string of digits that can be interpreted by any machine as words or characters.*

Example 1.5. *The string 100110 is a codeword of the vector space $V(6, 2)$ of the length 6 over the finite field F_2.*

Definition 1.11. ***Prime Number***, *a prime number is a natural number that can be divisible only by 1 and itself (i.e, a prime number has two divisors namely 1 and the number itself).*

Example 1.6. *The first ten prime numbers are $2, 3, 5, 7, 11, 13, 17, 19, 23, 29$.*

Definition 1.12. ***Algorithm***, *an algorithm is a scientific term for solving an instance or a set of instructions that can be followed for solving a problem.*

Example 1.7. *To find the greatest common divisor (GCD) of two numbers a and b, we can apply division algorithm, and the GCD is the last non-zero remainder. All steps that are followed to determine the GCD will make an algorithm.*

1.2 Mathematical Definition of a Group

A set R together with a binary operation $(*)$ is said to be a group if it satisfies the following properties:

For a, b and $c \in R$,

1. $a * b \in R$ (closure)

2. $(a * b) * c = a * (b * c)$ (associativity)

3. There exists additive identity element e of R such that $a * e = e * a = a$, for all $a \in R$ (for (*) operation, identity is always e (identity element))

4. There exists inverse element a^{-1} of R such that $a * a^{-1} = a^{-1} * a = e$ (inverse element)

5. Furthermore if $a * b = b * a$, then R is said to be a commutative group or an abelian Group.

Note: this operation is not always $(*)$ it can be also addition, and it may be another operation defined on a set R.
However, in this research paper we are restricted on the usual addition and multiplication operators.

1.3 Mathematical Definition of a Ring

A set R together with two binary operations namely addition $(+)$ and multiplication $(*)$ is said to be a ring if the following 3 conditions are satisfied:
Fora, b and $c \in R$,

1. (R, +) must be an abelian group

2. $a * (b * c) = (a * b) * c$: associativity law for multiplication

3. $a * (b + c) = a * b + a * c$ (left distributive law) $(a + b) * c = a * c + b * c$ (right distributive law)

Note: The above two operations $(+)$ and $(*)$ are not necessarily the ordinary addition and multiplication operations, reason why the definition of these operations may be needed in mathematical expressions. But this paper considers them as ordinary addition and multiplication.

If there exists multiplicative identity element of R for each every non-zero element of R, always denoted 1 such that $a * 1 = 1 * a = a$, then we can call the ring R to be the ring with unity.

The inverse of an element a for the abelian group $(R, +)$ is denoted $(-a)$.

In addition if $a * b = b * a$, then R is called a commutative ring with unity. If every non- zero element of a commutative ring R with unity is invertible, then R

becomes a field.

1.4 Classification of fields

Fields can be classified by size or by the number of elements that a field possesses. If a field contains a finite number of elements then that field is called finite field, otherwise it is an infinite field. For the rest of the work we will proceed with the finite field only.

For example consider the commutative ring, $\mathbb{Z}_p$, where p is a prime number, is a commutative ring with unity which is the field hence finite field because it possesses finite number of element. This is the most popular example of finite field.

Then, definition of this topic as the name indicated above, a finite field is a field with a finite order (i.e number of elements is finite). It is also called Galois field (so named in honor of Evariste Galois). The order of a finite field is always a prime number or a power of a prime number. A finite field of order p^n is denoted $GF(p^n)$, often written as $F(p^n)$ in current usage.

$GF(p^n)$ is called the prime field of order p, where the p elements are denoted $0, 1, 2, 3, ..., p-1$. In the finite field $GF(p)$ if two elements are written as $a = b$ this is the same as $a \equiv b(mod \quad p)$. Finite fields are therefore denoted by $GF(p^n)$ instead of $GF(k)$ where $k = p^n$, for clarity. The finite field $GF(2)$ consists of elements $0, 1$ which satisfy the addition and multiplication modulo 2. Let us first consider the addition and multiplication of elements in $GF(2)$ as shown in following two tables below:

+	0	1
0	0	1
1	1	0

Table 1: The table shows the addition in $GF(2)$

*	0	1
0	0	0
1	0	1

Table 2: The table describes the multiplication in $GF(2)$

Clearly $GF(2)$ is finite field since it contains two elements 0 and 1 which is a finite number of elements and also by the rule that every non-zero element is invertible, in the table it is clear that 1 is the only non-zero element and it is invertible. The finite fields are classified by size, as follows:

1. The order or number of elements of finite fields is of the form p^n, where p is a prime number called the characteristic of the field, and n is a positive integer.
2. For every prime number p and a positive integer n, there exists a finite field with p^n elements.
3. Any two finite fields with the same number of elements are isomorphic. For example $\mathbb{Z}/(3)$ is isomorphic to F_3. That is under some renaming of the elements of one of these two fields, its addition and multiplication tables become identical to the corresponding tables of the other one. This classification is justified by using a naming scheme for finite fields that specifies only the order of the field.

Note: Finite fields are important and very useful in number theory, algebraic geometry, Galois Theory, cryptography, coding theory and quantum error correction. Its applications may also be appearing in the electrical circuits.

2 Computational Over Finite Fields with First seven Rings ($\mathbb{Z}_p$, where $p = 2, 3, 5, 7, 11, 13, 17$)

Arithmetic in a finite field is different from standard integers arithmetic. There are a limited number of elements in the finite field; all operations performed in the finite field result in an element within that field.

While each finite field is itself not infinite, there are infinitely many different finite fields; their number of elements (which is also called cardinality) is necessarily of the form p^n, where p is a prime number and n is a positive integer, and two finite fields of the same size are isomorphic. Consider $\mathbb{Z}/(3)$ is isomorphic to $\mathbb{Z}_3$. The prime p is called the characteristic of the finite field, and the positive integer n is called the dimension of the field over its prime field.

The finite field with p^n elements is denoted $GF(p^n)$ and is also called the Galois Field, in honor of the founder of finite field theory, Evariste Galois [Cox, 2011]. $GF(p)$, where p is a prime number, is simply the ring of integers modulo p. That is, one can perform operations (addition, subtraction, division and multiplication) by using the usual operation on integers, followed by reduction modulo p. For instance, in $GF(5)$, $4+3=7$ is reduced to 2 modulo 5. Division is multiplication by the inverse modulo p, which may be computed using the extended Euclidean algorithm.

A particular case is $GF(2)$, as addition and multiplication have been shown above in Table 1 and Table 2 respectively, and the only invertible element is 1. Now arithmetic operations in this paper are done on the first seven rings of integers modulo p ($\mathbb{Z}_p$), where p is a prime number, and those are $\mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_5, \mathbb{Z}_7, \mathbb{Z}_{11}, \mathbb{Z}_{13}$ and Z_{17}.

2.1 Arthmetic Operation in the Ring of Integers ($\mathbb{Z}_3$)

The class of residues in $\mathbb{Z}_3$ are $0, 1, 2$

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Table 3: This is a table that shows the addition in $\mathbb{Z}_3$

*	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Table 4: This table describes the multiplication in $\mathbb{Z}_3$

2.2 Arthmetic Operation in the Ring of Integers ($\mathbb{Z}_5$)

The class residues in $\mathbb{Z}_5$ are $0, 1, 2, 3, 4$

*	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Table 6: This a multiplication table in $\mathbb{Z}_5$

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Table 5: This is a table that illustrates how an addition is done in $\mathbb{Z}_5$

2.3 Arthmetic Operation in the Ring of Integers ($\mathbb{Z}_7$)

The class residues of $\mathbb{Z}_7$ are $0, 1, 2, 3, 4, 5, 6$

+	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

Table 7: An addition table in $\mathbb{Z}_7$

*	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

Table 8: Multiplication table in $\mathbb{Z}_7$

2.4 Arthmetic Operation in the Ring of Integers ($\mathbb{Z}_{11}$)

The class residues of $\mathbb{Z}_{11}$ are $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$

+	0	1	2	3	4	5	6	7	8	9	10
0	0	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10	0
2	2	3	4	5	6	7	8	9	10	0	1
3	3	4	5	6	7	8	9	10	0	1	2
4	4	5	6	7	8	9	10	0	1	2	3
5	5	6	7	8	9	10	0	1	2	3	4
6	6	7	8	9	10	0	1	2	3	4	5
7	7	8	9	10	0	1	2	3	4	5	6
8	8	9	10	0	1	2	3	4	5	6	7
9	9	10	0	1	2	3	4	5	6	7	8
10	10	0	1	2	3	4	5	6	7	8	9

Table 9: This table demonstrates the addition in $\mathbb{Z}_{11}$

*	0	1	2	3	4	5	6	7	8	9	10
0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9	10
2	0	2	4	6	8	10	1	3	5	7	9
3	0	3	6	9	1	4	7	10	2	5	8
4	0	4	8	1	5	9	2	6	10	3	7
5	0	5	10	4	9	3	8	2	7	1	6
6	0	6	1	7	2	8	3	9	4	10	5
7	0	7	3	10	6	2	9	5	1	8	4
8	0	8	5	2	10	7	4	1	9	6	3
9	0	9	7	5	3	1	10	8	6	4	2
10	0	10	9	8	7	6	5	4	3	2	1

Table 10: Multiplication table in $\mathbb{Z}_{11}$

2.5 Arthmetic Operation in the Ring of Integers ($\mathbb{Z}_{13}$)

The class residues of $\mathbb{Z}_{13}$ are $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$

+	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	1	2	3	4	5	6	7	8	9	10	11	12
1	1	2	3	4	5	6	7	8	9	10	11	12	0
2	2	3	4	5	6	7	8	9	10	11	12	0	1
3	3	4	5	6	7	8	9	10	11	12	0	1	2
4	4	5	6	7	8	9	10	11	12	0	1	2	3
5	5	6	7	8	9	10	11	12	0	1	2	3	4
6	6	7	8	9	10	11	12	0	1	2	3	4	5
7	7	8	9	10	11	12	0	1	2	3	4	5	6
8	8	9	10	11	12	0	1	2	3	4	5	6	7
9	9	10	11	12	0	1	2	3	4	5	6	7	8
10	10	11	12	0	1	2	3	4	5	6	7	8	9
11	11	12	0	1	2	3	4	5	6	7	8	9	10
12	12	0	1	2	3	4	5	6	7	8	9	10	11

Table 11: This is an addition table in $\mathbb{Z}_{13}$

*	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9	10	11	12
2	0	2	4	6	8	10	12	1	3	5	7	9	11
3	0	3	6	9	12	2	5	8	11	1	4	7	10
4	0	4	8	12	3	7	11	2	6	10	1	5	9
5	0	5	10	2	7	12	4	9	1	6	11	3	8
6	0	6	12	5	11	4	10	3	9	2	8	1	7
7	0	7	1	8	2	12	3	10	4	11	5	12	6
8	0	8	3	11	6	1	8	4	12	7	2	10	5
9	0	9	5	1	10	6	2	11	7	3	12	8	4
10	0	10	7	4	1	11	8	5	2	12	9	6	3
11	0	11	9	7	5	3	1	12	10	8	6	4	2
12	0	12	11	10	9	8	7	6	5	4	3	2	1

Table 12: Multiplication Table in $\mathbb{Z}_{13}$

From this table, each non-zero element has its multiplicative inverse, the multiplicative inverse of 8 for example is 5, the multiplicative inverse of 11 is 6, and so on.

2.6 Arithmetic Operation in the Ring of Integers ($\mathbb{Z}_{17}$)

The class residues of $\mathbb{Z}_{17}$ are $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16$

+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	0
2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	0	1
3	3	4	5	6	7	8	9	10	11	12	13	14	15	16	0	1	2
4	4	5	6	7	8	9	10	11	12	13	14	15	16	0	1	2	3
5	5	6	7	8	9	10	11	12	13	14	15	16	0	1	2	3	4
6	6	7	8	9	10	11	12	13	14	15	16	0	1	2	3	4	5
7	7	8	9	10	11	12	13	14	15	16	0	1	2	3	4	5	6
8	8	9	10	11	12	13	14	15	16	0	1	2	3	4	5	6	7
9	9	10	11	12	13	14	15	16	0	1	2	3	4	5	6	7	8
10	10	11	12	13	14	15	16	0	1	2	3	4	5	6	7	8	9
11	11	12	13	14	15	16	0	1	2	3	4	5	6	7	8	9	10
12	12	13	14	15	16	0	1	2	3	4	5	6	7	8	9	10	11
13	13	14	15	16	0	1	2	3	4	5	6	7	8	9	10	11	12
14	14	15	16	0	1	2	3	4	5	6	7	8	9	10	11	12	13
15	15	16	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
16	16	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Table 13: This table points out how to perform an addition in $\mathbb{Z}_{17}$

*	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	0	2	4	6	8	10	12	14	16	1	3	5	7	9	11	13	15
3	0	3	6	9	12	15	1	4	7	10	13	16	2	5	8	11	14
4	0	4	8	12	16	3	7	11	15	2	6	10	14	1	5	9	13
5	0	5	10	15	3	8	13	1	6	11	16	4	9	14	2	7	12
6	0	6	12	1	7	13	2	8	14	3	9	15	4	10	16	5	11
7	0	7	14	4	11	1	8	15	5	12	2	9	16	6	13	3	10
8	0	8	16	7	15	6	14	5	13	4	12	3	11	2	10	1	9
9	0	9	1	10	2	11	3	12	4	13	5	14	6	15	7	16	8
10	0	10	3	13	6	16	9	2	12	5	15	8	1	11	4	14	7
11	0	11	5	16	10	4	15	9	3	14	8	2	13	7	1	12	6
12	0	12	7	2	14	9	4	16	11	6	1	13	8	3	15	10	5
13	0	13	9	5	1	14	10	6	2	15	11	7	3	16	12	8	4
14	0	14	11	8	5	2	16	13	10	7	4	1	15	12	9	6	3
15	0	15	13	11	9	7	5	3	1	16	14	12	10	8	6	4	2
16	0	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1

Table 14: This a multiplication table in $\mathbb{Z}_{17}$

Apart from the 14 tables represented above, one may proceed in the same way up to the finite fields of $p - 1$ class residues with p being a prime number or a power of a prime number.

3 Arithmetic Computational of Polynomials over Finite Fields

The theory of polynomials over finite fields is important for investigating the algebraic structure of finite fields as well as for many applications. Above all, irreducible polynomials, the prime elements of polynomial rings over finite fields are indispensable for constructing finite fields and computing with the elements of finite fields [Rónyai, 1992].
A polynomial is an expression of the form $a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0$, for some non-negative integer n and where the coefficients $a_0, a_1 ..., a_n$ are drawn from some designated set S, which is in particular finite field and called the coefficient set.

Polynomial arithmetic deals with the addition, subtraction, multiplication, and division of polynomials.

3.1 What Problems Does Polynomial Arithmetic Adress?

Given two polynomials whose coefficients are derived from a set S, what can we say about the coefficients of the polynomial that results from an arithmetic operation on the two polynomials? If we insist that the polynomial coefficient all come from a particular S, then which arithmetic operations are permitted and which prohibited? Let us say that the coefficient set is a finite field F with its own rules for addition, subtraction, multiplication and division, and let us further say that when we carry out an arithmetic operation on two polynomials, we subject the operations on the coefficients to those that apply to the finite field F. Now what can be said about the set of such polynomials? All these questions will have their answers as we move on in this paper.

3.2 Ordinary Addition and Subtraction of Polynomials

Let $f(x) = a_2x^2 + a_1x + a_0$ and $g(x) = b_1x + b_0$ Then $f(x) + g(x) = a_2x^2 + (b_1 + a_1)x + (a_0 + b_0)$
Let $f(x) = a_2x^2 + a_1x + a_0$ and $g(x) = b_3x^3 + b_0$, Then $f(x) - g(x) = -b_3x^3 + a_2x^2 + a_1x + (a_0 - b_0)$

3.3 Ordinary Multiplication of Polynomials

Let $f(x) = a_2x^2 + a_1x + a_0$ and $g(x) = b_1x + b_0$, Then $f(x) * g(x) = a_2b_1x^3 + (a_2b_0 + a_1b_1)x^2 + (a_1b_0 + a_0b_1)x + a_0b_0$.

3.4 Ordinary Division of Polynomials

3.4.1 When is Division of Polynomials Permited?

Polynomial division is obviously not allowed for polynomials that are not defined over certain fields. For example, for polynomials defined over the set of all integers, you cannot divide $4x^2 + 5$ by the polynomial $5x$. If you tried, the first term of the quotient would be $(\frac{4}{5})x$ where the coefficient of x is not an integer. You can always divide polynomials defined over a certain field. What that means is that the operation of division is legal when the coefficients are drawn from a finite field. Note that, in general, when you divide such polynomial by another, you will end up with a remainder, and when, in general you divide one integer by another integer it is possible in purely integer arithmetic.

Therefore, in general, for polynomials defined over a field, the division of a polynomial $f(x)$ of a degree m by another polynomial $g(x)$ of a degree $n - m$ can be expressed by $f(x)/g(x) = q(x)g(x) + r(x)$, where $q(x)$ is the quotient and, $r(x)$ the remainder, so we can write for any two polynomials defined over a field, $f(x) = q(x) * g(x) + r(x)$ assuming that the degree of $f(x)$ is not less than that of $g(x)$. When $r(x)$ is zero, we say that $g(x)$ divides $f(x)$. This fact can also be expressed by saying that $g(x)$ is a divisor of $f(x)$ and by notation, $g(x)|f(x)$.

3.4.2 Division of a Polynomial by Another Upon Using Long Division

Let us divide the polynomial $8x^2 + 3x + 2$ by the polynomial $2x + 1$:

In this example, our dividend is $8x^2 + 3x + 2$ and the divisor is $2x + 1$. We now need to find the quotient.
Long division for polynomials consists of the following steps:

Step 1: Arrange both the dividend and the divisor in the descending powers of the variable.
Step 2: Divide the first term of the dividend by the first term of the divisor and write the result as the first term of the quotient.
In our example, the first term of the dividend is $8x^2$ and the first term of the divisor is $2x$, so the first term of the quotient is $4x$.
Step 3: Multiply the divisor with the quotient term just obtained and arranges the result under the dividend so that the same powers of x match up. Subtract the expression just laid out from the dividend. In our example, $4x$ times $2x + 1$ is equal to $8x^2 + 4x$. Subtracting this from the dividend yields $-x + 2$. consider the result of the above subtraction as the new dividend and go back to the first step. (The new dividend in our case is $(-x + 2)$. In our example, dividing the polynomial $8x^2 + 3x + 2$ by the polynomial $2x + 1$, yield quotient of $4x - 0.5$ and a remainder of 2.5.

3.5 Arithmetic Operations on Polynomials whose Coefficients Belong to a Defined Finite Fields

The arithmetic operations on polynomials whose coefficients are drawn from finite fields is not the same as the usual operations of polynomials. To see this, Let us consider the set of all polynomials whose coefficients belong to the finite field $\mathbb{Z}_7$ (which is the same as GF(7)). Here is an example of adding two such polynomials: $f(x) = 5x^2 + 4x + 6$, $g(x) = 5x + 6$ we get $f(x) + g(x) = 5x^2 + 9x + 12 =$

$5x^2 + 2x + 5$

If we perform the difference of both polynomials, $f(x) = 5x^2 + 4x + 6$ and $g(x) = 5x + 6$ then $f(x) - g(x) = 5x^2 - x = 5x^2 + 6x$ since the additive inverse of 5 in $\mathbb{Z}_7$ is 2 and that of 6 is 1. So $4x - 5x$ is the same as $4x + 2x$ and $6 - 6$ is the same as $6 + 1$, with both additions modulo 7.
The multiplication of polynomials $f(x) = 5x^2 + 4x + 6$, and $g(x) = 5x + 6$ is given by $f(x) * g(x) = 4x^3 + x^2 + 5x + 1$
Lastly the divison of polynomials $f(x) = 5x^2 + 4x + 6$, $g(x) = 2x + 1$ is given by $f(x)/g(x) = 6x + 6$. If you multiply the divisor $2x + 1$ with the quotient $6x + 6$, you get the dividend $5x^2 + 4x + 6$.

Let consider also the polynomials defined over $GF(2)$. Recall that the notation $GF(2)$ means the same thing as $\mathbb{Z}_2$. We are obviously talking about arithmetic modulo 2. First of all, $GF(2)$ is a sweet basic finite field. Recall that the number 2 is the first prime. (A prime has exactly two distinct divisors, 1 and itself). $GF(2)$ consists of the set $0, 1$. The two elements of this set obey the following addition and multiplication rules:

0 + 0 = 0
0 x 0 = 0
0 + 1 = 1
0 x 1 = 0
1 + 0 = 1
1 x 0 = 0
1 + 1 = 0
1 x 1 = 1

So the addition over $GF(2)$ is equivalent to the logical XOR operation, and multiplication to the logical AND operation. Some examples of polynomials defined over $GF(2)$: are $x^3 + x^2 - 1; -x^5 + x^4 - x^2 + 1; x + 1$, etc.

3.5.1 Arithmetic Computational of Polynomials Defined Over $GF(2)$

Here is an example of adding two such polynomials: $f(x) = x^2 + x + 1, g(x) = x + 1$, therefore $f(x) + g(x) = x^2 + 2x + 2 = x^2$

- Here is an example of subtracting two such polynomials, $f(x) = x^2 + x + 1, g(x) = x + 1$, then $f(x) - g(x) = x^2$
- Here is an example of multiplying two such polynomials, $f(x) = x^2 + x + 1$, and $g(x) = x + 1$, then $f(x) \times g(x) = x^3 + 1$

- Here is an example of dividing two such polynomials,$f(x) = x^2 + x + 1, g(x) = x + 1$, then $f(x)/g(x) = x$.

  If you multiply the divisor, $x + 1$ with the quotient x, you get $x^2 + x$. That when added to the remainder 1 gives us back the dividend $x^2 + x + 1$

3.6 Division of Polynomials Defined Over Finite Fileds

First, note that a polynomial is defined over a field if all its coefficients are drawn from that field. Dividing polynomials defined over a finite field is a little bit more frustrating than performing other arithmetic operations on such polynomials. Now your mental gymnastics must include both additive inverses and multiplicative inverses.Consider again the polynomials defined over $GF(7)$. Let's say we want to divide $5x^2 + 4x + 6$ by $2x + 1$. In a long division, we must start by dividing $5x^2$ by $2x$. This requires that we divide 5 by 2 in $GF(7)$. Dividing 5 by 2 is the same as multiplying 5 by the multiplicative inverse of 2. Multiplicative inverse of 2 is 4 since $2 \equiv 4$ mod 7 is 1. So we have $5 \equiv 2^{-1} = 5 \equiv 4 = 20$ mod $7 = 6$. Therefore, the first term of the quotient is 6x. Since the product of $6x$ and $2x + 1$ is $5x^2 + 6x$, we need to subtract $5x^2 + 6x$ from the dividend $5x^2 + 4x + 6$. The result is $(4 - 6)x + 6$, which (since the additive inverse of 6 is 1) is the same as $(4 + 1)x + 6$, and that is the same as $5x + 6$.

Our new dividend for the next round of long division is therefore $5x + 6$. To find the next quotient term, we need to divide $5x$ by the first term of the divisor, that is by $2x$. Reasoning as before, we see that the next quotient term is again 6. The final result is that when the coefficients are drawn from the set $GF(7)$), $5x^2 + 4x + 6$ divided by $2x + 1$ yields a quotient of $6x + 6$ with the remainder zero.
So we can say that as a polynomial defined over the field, $GF(7)$, $5x^2 + 4x + 6$ is a product of two factors, $2x+1$ and $6x+6$. We can therefore write $5x^2+4x+6 = (2x + 1) \equiv (6x + 6)$

4 Irreducible Polynomials or Prime Polynomials

Definition 4.1. *According to [Rónyai, 1992], a polynomial $f \in F[x]$ is said to be irreducible over F (or irreducible in $F[x]$, or prime in $F[x]$) if f has positive degree and $f = g * h$, with $g, h \in F[x]$ implies that either g or f is a constant polynomial, otherwise it is reducible over F. The reducibility or irreducibility of a given polynomial depends heavily on the field under considerations. For instance, the polynomial $x^2 - 2 \in Q(x)$ is irreducible over the field Q of rational numbers, but $x^2 - 2 = (x + \sqrt{2})(x - \sqrt{2})$ but reducible over the field of real numbers*

$(\mathbb{R})$. *For polynomials over finite fields, the same argument hold except that the coefficients are reduced in mod* p.

Example 4.1. $f(x) = x^2 + x + 1$ *is irreducible over* F_2 *but* $g(x) = x^2 + 1$ *is reducible over* F_2 *to see this* $g(x) = x^2 + 1 = (x + 1)(x + 1) = x^2 + 2x + 1$, *since* $2 \equiv 0 \quad mod(2)$, *and then* $2x \equiv 0 \quad mod(2)$. *In few words we can say, when* $g(x)$ *divides* $f(x)$ *without leaving a remainder, we say* $g(x)$ *is a factor of* $f(x)$. *A polynomial* $f(x)$ *over a field* F *is called irreducible, if* $f(x)$ *cannot be expressed as a product of two polynomials, both over* F *and both of degree lower than that of* $f(x)$. *An irreducible polynomial is also referred to as a prime polynomial.*

5 Some Computational Tables of Quotient Polynomials Over Finite Fields

To represent the elements of an extension fields over finite fields in a computational table, we must have the quotient? $F_q[x]/f(x)$, where $f(x)$ is irreducible over $F_q[x]$. This form of polynomials are looked like powers of prime [Lidl and Niederreiter, 1994].

Example 5.1. *Let* $f(x) = x^2 + 1 \in F_3[x]$. *Thus to find the computational tables of* $F_3[x]/(f(x))$, *we need to find the residue class ring as* p^n *where* n *is the degree of polynomial* $f(x)$, *and then we have a set of residue class ring of* $3^2 = 9$ *elements, as it looks like a representation of* $F(9)$, *such as* $0, 1, 2, x, 1 + x, 2 + x, 1 + 2x, 2x, 2 + 2x$, *these are precisely the polynomials of degree less than 2 over* F_3 *by equating* $x^2 + 1 = 0$ *and this implies that* $x^2 = -1 = 2$, *but remember that computational in finite fields are followed by mod* p *[Gong et al., 2013]*

.

+	0	1	2	x	1+x	2+x	1+2x	2x	2+2x
0	0	1	2	x	1+x	2+x	1+2x	2x	2+2x
1	1	2	0	1+x	2+x	x	2+2x	1+2x	2x
2	2	0	1	2+x	x	1+x	2x	2+2x	1+2x
x	x	1+x	2+x	2x	1+2x	2+2x	1	0	2
1+x	1+x	2+x	x	1+2x	2+2x	2x	2	1	0
2+x	2+x	x	1+x	2+2x	2x	1+2x	0	2	1
1+2x	1+2x	2+2x	2x	1	2	0	2+x	1+x	x
2x	2x	2x+1	2+2x	0	1	2	1+x	x	2+x
2+2x	2+2x	2x	1+2x	2	0	1	x	2+x	1+x

Table 15: Addition table for $F_3[x]/(f(x))$

*	0	1	2	x	1+x	2+x	1+2x	2x	2+2x
0	0	0	0	0	0	0	0	0	0
1	0	1	2	x	1+x	2+x	1+2x	2x	2+2x
2	0	2	1	2x	2+2x	1+2x	2+x	x	1+x
x	0	x	2x	2	x+2	2x+2	x+1	1	2x+1
1+x	0	1+x	2+2x	x+2	2x	1	2	2x+1	x
2+x	0	2+x	1+2x	2x+2	1	x	2x	x+1	2
1+2x	0	1+2x	2+x	x+1	2	2x	x	2x+2	1
2x	0	2x	x	-2	2x+1	x+1	2x+2	x	x+2
2+2x	0	2+2x	1+x	2x+1	x	2	1	x+2	2x

Table 16: Multiplication Table for $F_3[x]/(f(x))$

6 Applications of Finite Fields

6.1 Algebraic Coding Theory

It is one of the major applications of finite field. This theory has its origin in famous theorem of Shannon that guarantees the existence of codes that can transmit information at rates close to the capacity of a communication channel with an arbitrary small probability of error. One of the purposes of algebraic coding theory, the theory of error-correcting and error-detecting codes is to devise methods for construction of such codes [von zur Gathen et al.]. During the last two decades more and more abstract algebraic tools such as the theory of finite fields and the theory of polynomials over finite fields have influenced coding. In particular, the description of redundant codes by polynomials over F_q is a milestone in this development. The fact that one can use shift registers for coding and decoding establishes a connection with linear recurring sequences. In our discussion of algebraic coding theory we do not consider any of the problems of the implementation or technical realization of the codes. We restrict ourselves to the study of basic properties of block codes and the description of some interesting classes of block codes.

6.1.1 Linear coding

The problem of communicating the information, in particular the coding and decoding of information for the reliable transmission over a "noisy" channel is of great importance today. Typically, one has to transmit a message which consists of finite string of symbols that are elements of some finite alphabet. For instance, if this alphabet consists of simply 0 and 1, the message can be described as binary number.

Generally the alphabet is assumed to be finite fields. Now the transmission of finite string of elements of the alphabet over a communication channel need not to be perfect in the sense that each bit of information is transmitted unaltered over this channel. As there is no ideal channel without "noise" the receiver of the transmitted message may obtain distorted information and may make errors in interpreting the transmitted signal.

One of the main problems of coding theory is to make the errors, which occur for instance because of noisy channel, extremely improbable.

The methods of improve the reliability of transmission depend on properties of finite fields. A basic idea in algebraic coding theory is to transmit redundant information together with the message one wants to communicate; that is, one extends the string of message symbols to a longer string in a systematic way.

A simple model of communication system is shown in the figure bellow:

We assume that the symbols of the message and the coded message are elements of the same finite field F_q. Coding means to encode a block of k message symbols $a_1, a_2, ..., a_k$ where $a_i \in F_q$ into a code word $c_1, c_2, ..., c_n$ of n symbols, where $c_j \in F_q$, with $n > k$. We regard the code word as an n-dimensional row vector $c \in F_q^n$. Thus f in the Figure below is a function from F_q^k into F_q^n, called a coding scheme, and $g : F_q^n \to F_q^k$ is a decoding scheme.

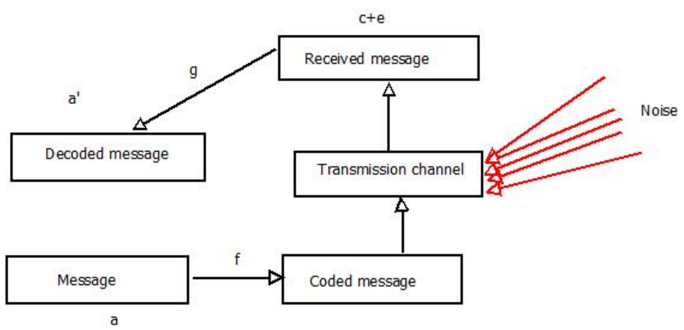


Figure 1: Communication figure that shows how a message is coded, transmitted and decoded

A simple type of coding scheme arises when each block $a_1 a_2 ... a_k$ of message symbols is encoded into a code word of the form $a_1 a_2 ... a_k c_{k+1} ... c_n$, where the first k

symbols are the original message symbols and the additional $n-k$ symbols in F_q are control symbols. Such coding schemes are often presented in the following way. Let H be a given $(n-k) \times n$ matrix with entries in F_q that is of the special form $H = (A, I_{n-k})$, where A is $(n-k) \times k$ matrix and I_{n-k} is the identity matrix of order $n-k$. The control symbols $c_{k+1}, ..., c_n$ can then be calculated from the system of the equations $HC^T = 0$, for code word c. The equations of this system are called parity-check equations. The examples of this theory will be given later.

6.2 Error-Correcting Codes (Practice of Linear Code)

Since the theory of codes was developed in order to ensure reliability of transmitted information, as an example, consider the ISBN (International Standard Book Number) of published book. This number usually appears on the back of the book in the bottom right-hand corner. The ISBN consists of a nine-digits $0, 1, ..., 9$ or the symbol X (standing for 10). This final symbol may be calculated from the other nine as follows:
From an integer N by adding together the first digit, twice the second digit, three times the third and so on. The check digit is the remainder when N is divided by 11. For example, a book with first 9 digits 019853453 will have $N = 0 + 2 + 27 + 32 + 25 + 18 + 28 + 40 + 27 = 199$, and so the check digit should be 1, giving ISBN 01953453 1. The point about such a number is that if it is inaccurately copied, and an error is made in any of the digits in the first nine locations (such as the last "5" being copied as a "3"), then the resulting number will not have "1" as its check digit. This is an example of error-detecting code: the ISBN detects when a single error is made after transcribing the number. Another example of finding check digit is that of 102463798, then $N = 1\times1+0\times2+2\times3+4\times4+6\times5+3\times6+7\times7+9\times8+8\times9 = 264$, and divide this number by 11 to get the check digit which is 0, and hence giving ISBN 102463798 0 In this part we shall explain methods which not only detect errors, but also enables us to correct it.

Definition 6.1. *Let p be prime integer. Denote by $V(n, p)$ the set of all sequences of length n of the elements from the set $\mathbb{Z}_p$ of congruence classes modulo p, so that $V(n, p)$ has p^n elements. We will usually omit the commas and brackets commonly used to denote elements of the vector spaces, so that* $(1, 0, 1)$*, will be written as* 101*. Thus $V(3, 2)$ consists of the eight sequences* $000, 001, 010, 011, 100, 101, 110, 111$ *while $V(2, 3)$ consists of the nine sequences* $00, 01, 02, 10, 11, 12, 20, 21, 22$*. We add sequences by adding the corresponding terms, by just remembering that we are adding congruence classes. Thus, for example in $V(3, 2)$,* $110 + 011 = 101$

while in $V(2,3)$*,* $12+11 = 20$*. We can also multiply an element in* $V(n,p)$ *by a congruence class by multiplying each term in the sequence by the representative for the congruence class and reducing modulo* p*. For example, in the space* $V(3,3)$ *we see that 2(102) =201. In fact* $V(n,p)$ *is a vector space of dimension n over the field* $\mathbb{Z}_p$

Definition 6.2. *A linear* (n,k)*-code is any* k*-dimensional subspace* C *of the vector space* $V(n,p)$*. Thus* C *satisfies the following two conditions:*
The difference of any two elements of C *is an element of* C*, and the product of any element of* C *with an element of* $\mathbb{Z}_p$ *is also an element of* C*. The elements of* C *are called codewords.*

Note: A subspace of a vector space is necessary non-empty, so condition (1) ensures that the zero element of the vector space is in the subspace C. It then follows by the additive version that C is a group under addition.

Example 6.1. *Consider the four elements* $000, 001, 010, 011$ *of* $V(3,2)$*. These are precisely the four sequences which start with 0. This subspace of* $V(3,2)$ *satisfies condition one, that subtracting any two of these gives a sequence starting with* 0*. Also condition (2) holds, since 0 and 1 are the only elements of* $\mathbb{Z}_2$ *and then multiply each sequence by any of these two elements we get an element starting with* 0*. Therefore the four elements form a linear* $(3,2)$*-code.*

Definition 6.3. *Let* v *be any element of* $V(n,p)$*. The weight of* v *is the number of non-zero terms in the sequence* v*. If* v *and* w *are two elements of* $V(n,p)$*, the distance* $d(v,w)$ *is the number of places at which* v *and* w *differ.*

Example 6.2. *In* $V(4,3)$ *the weight of* 1201 *is three, since there are three non-zero entries. The distance from* 1201 *to* 2211 *is two, since these two vectors differ in two places. In* $V(5,5)$ *the weight of* 13402 *is four and so on.*

Proposition 6.1. *Let* u, v *and* w *be any elements of* $V(n,p)$*. Then*

1. $d(u,v) \geq 0$ with equality if and only if $u = v$;
2. $d(u,v) = d(v,u)$; and
3. $d(u,v) + d(v,w) \geq d(u,w)$.

Proof.

1. It follows directly from the definition that $d(u, v)$ is positive except u and v do not differ anywhere.

2. This is always true for u and v.

3. In each location at which u and w differ, v cannot agree with both u and w. Thus every contribution to the value of $d(u, w)$ provides a contribution to either $d(u, v)$ or to $d(v, w)$. $\square$

Definition 6.4. *Let C be subspace of $V(n, p)$. The minimum distance d of C is the least distance between different codewords: $d = \min_{u,v}\{d(u, v)\}$. The next result shows that for a linear code, the minimum distance d can be calculated from the code words.*

Proposition 6.2. *Let C be a linear $(n - k)$-code. Then the minimum distance of C is equal to the smallest possible weight of any non-zero codeword.*

Proof.

Let f be the smallest possible weight of any non-zero codeword, and let 0 denote the sequence consisting entirely of zeros. Suppose that w is a codeword of weight f. Then $d(w, 0)$ if and only if so $f \geq d$. Now let u and v be pair of codewords with $d(u, v) = d$. Since C is a linear code, the word $u - v$ is a codeword of weight d, so $d \geq f$. It follows that $d = f$.

The importance of the minimum distance lies in the detecting the errors and correction of those errors. To see this, consider the following proposition. $\square$

Proposition 6.3. *Let C be linear code with minimum distance d. Then C detects $d - 1$ or fewer errors, and corrects e errors for any e with $2e + 1 \leq d$.*

Proof.

Let v be a vector which has distance f from a codeword c, where $f \leq d - 1$. We think of c as the transmitted word and v as the received word, so that there are f errors in transmission. Since d is the minimum distance for C, the received v cannot be a codeword. We express this by saying that the code C detects d or fewer errors. Suppose now that v has distance e from a codeword c and also that $2e + 1 \leq d$. Then there can be no other codeword near to v: If c_1 was in C and $d(v, c_1) \leq e$, then by property of triangle inequality $d(c, c_1) \leq d(c, v) + d(v, c_1) \leq e + e < d$, which contradicts the definition. Thus there is a unique nearest codeword to v, and we say that C corrects e errors in this case. □

Definition 6.5. *Let n and k be any positive integers with $n > k$. Let p be a prime number. A (standard) generator matrix G over $\mathbb{Z}_p$ is a $k \times n$ matrix with entries in $\mathbb{Z}_p$, in which the first k columns form an identity $k \times k$ matrix. Given such a matrix, we obtain a linear code by regarding the rows as sequences and taking all possible linear combinations of these. Alternatively, we can consider the code as consisting of all sequences obtained from matrix multiplications of the form $u.G$ as u varies over all sequences of length k over $\mathbb{Z}_p$.*

Example 6.3. *Consider the generator matrix over $\mathbb{Z}_2$*

$$G = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

The corresponding code consists of the combinations of the rows and so has four elements: $000; 101; 011$ *and* 110. *The codewords can also be described as the vectors of the form uG, as u varies over the four vectors* $00; 01; 10; 11$. *Every non-zero codeword has weight* 2, *so the codeword detects one error, but does not correct errors. For example, 111 is not among codewords (so it is detected) but it is of equal distance from the two codewords 101 and 011 in G, so it cannot be corrected.*

Example 6.4. *Another example of a binary code (code over $\mathbb{Z}_2$) is provided by the matrix*

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

There are 8 code words obtained from the rows of this matrix:
$000000; 100110; 010101; 110011; 001011; 101101; 011110; 111000.$

There are four code words of weight 3, three code words of weight 4 and one of weight 0. The minimum distance (d) of this code is therefore 3, so the code detects $d - 1$ errors means two errors and corrects one error. For example, 100111 lies at distance one from a unique codeword, 100110 and so there is unique way to correct one error. The vector 100001, however has distance two from 000000 and 110011, so cannot be corrected.

Example 6.5. *Consider the following generator matrix over $\mathbb{Z}_3$:*

$$G = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$$

In this case, the codeword consists of the linear combinations of the rows of the matrix, including multiplication by 1 and 2 since $p = 3$. There are 9 code words: $0000; 1021; 2012; 0112; 1100; 2121; 0221; 1212$ *and* 2200.

Since there is a codeword of weight 2, this code detects one error. Note that the minimum distance is 2 despite the fact that each row of the generator matrix has weight 3.

Example 6.6. *Consider also the following important code over $\mathbb{Z}_3$*

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 2 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

By considering this matrix, the minimum distance of this code is at most 5 since there is a row of the generator matrix of weight 5. It can be shown that the minimum distance is exactly 5, so that code corrects two errors. This is the Ternary Golay code and is one of the most important code. More details and its descriptions are found in [Cohen et al., 2013].

We now consider the problem of decoding a linear (n, k)-code C. This is done by listing the left cosets of the subgroup C of $V(n, p)$ in a table known as the cosets decoding table. The table is organized by writing the codewords as its first row with the zero codeword first. Each subsequent row is a left coset of C. The entries in the first column are the coset representatives, now called cosets leaders. The algorithm for choosing the r^{th}-coset leader is to choose any word of minimum weight not already included in the first $(r - 1)$ rows. Then to decode a given vector, locate it in the table, and correct it to the codeword standing in the same column of the coset decoding table.

Example 6.7. *Consider Example 6.6, above there are eight code words which form a subgroup C of the vector space $V(6, 2)$. Since V has $2^6 = 64$ elements,* *this subgroup has index 64/8=8. To form a complete coset decoding table, we list the elements of C in a row. We then choose any element v_2 which is of smallest weight among those not in the first row and write this at the left hand end of the second row. The second row is obtained by adding each element of C in turn to this. Thus the second row is just the coset of C with respect to v_2. Continue this process by choosing v_3 to be of the smallest weight among the elements not in the first two rows, and so on. This process is not unique, but depends upon the choice of coset representatives [Pless, 1998]. One example of these choices is given in the following table*

000000, 100110, 010101, 110011, 001011, 101101, 011110, 111000
100000, 000110, 110101, 010011, 101011, 001101, 111110, 011000
010000, 110110, 000101, 100011, 011011, 111101, 001110, 101000
001000, 101110, 011101, 111011, 000011, 100101, 010110, 110000
000100, 100010, 010001, 110111, 001111, 101001, 011010, 111100
000010, 100100, 010111, 110001, 001001, 101111, 011100, 111010
000001, 100111, 010100, 110010, 001010, 101100, 011111, 111001
100001, 000111, 110100, 010010, 101010, 001100, 111111, 011001

To decode any element v of $V(6, 2)$, we locate v in the table and then correct it to the element in the first row of the column containing v. Thus to use the table to decode 011010, *we need to locate it (it is in the fifth row and seventh column) and correct it to the element in the first row and the same column, giving* 011110. *Note that the cosets representative for the last row is not easy to find. According to the algorithm, we need a word of weight* 2 *not in the first seven rows. The representative we choose,* 100001, *is not unique. This is actually a somewhat cumbersome way to arrange the decoding, since an exhaustive search is required. The calculation can be made more systematic for codes given by (standard) generator matrices using (standard) parity check matrices [Sayed, 2011].*

Definition 6.6. *Let C be an (n, k)-linear code over $\mathbb{Z}_p$ defined using $k \times n$ generator matrix G of the form*

$$G = \begin{pmatrix} 1 & 0 & 0 & ... & 0 & \\ 0 & 1 & 0 & ... & 0 & A \\ \vdots & & & & & \\ 0 & 0 & 0 & ... & 1 & \end{pmatrix}$$

where, A is $k \times (n-k)$ matrix. The parity check matrix associated with G is the $(n-k) \times n$ matrix

$$P = \begin{pmatrix} & 1 & 0 & 0 & ... & 0 \\ -A^T & 0 & 1 & 0 & ... & 0 \\ & \vdots & & & & \\ & 0 & 0 & 0 & ... & 1 \end{pmatrix}$$

Note: The generator matrix G above is often written, in a block matrix form as $G = (I_k|A)$. similarly, the parity check matrix is written as $P = (-A^T|I_{(n-k)})$ [kar, 2012].

Example 6.8. *The parity check matrix of the generator matrix over $\mathbb{Z}_2$. The parity check of the matrix of*

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix},$$

is the matrix

$$P = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

[kar, 2012]. This matrix is obtained by considering the matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

and then after computing $-A^T$ get the above matrix P given by $-A^T|I_{(n-k)}$ [kar, 2012].

Definition 6.7. *Let C be a linear (n,k)-code with generator matrix G and associated parity matrix P. For any v in $V(n,p)$, let v^T denote the transpose of v, the column vector obtained by writing the members of the sequence v vertically. Then the syndrome of v is the element of $V(n-k,p)$ given by Pv^T. Thus in the above example, the syndrome of $v = 100000$ is 110 and the syndrome of $v = 110011$ is 000.*

Note: If C is a code with standard parity check matrix P, then an element v in $V(n,p)$ is a codeword if and only if the syndrome of v is the zero sequence.

Example 6.9. *We need not store or the complete coset decoding table, but merely a table of two columns, the coset representatives and their syndromes. In our previous example in which P was*

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

This table would be as the following,

Coset representatives	Syndromes
000000	000
100000	110
010000	101
001000	011
000100	100
000010	010
000001	001
100001	111

Table 17: This a table of syndromes and cosets representatives

Thus to decode a given vector such as 100111, calculate its syndrome to obtain 001. This is the syndrome for the seventh row, so this vector is not a codeword, but the word 100110 obtained by subtracting 000001 is a codeword. The advantage of listing coset representatives together with syndromes is that, it is much easier to find any missing coset representatives, since each sequence in $V(n-k,p)$ occurs as syndrome. Thus in this above example, the syndrome for the last row must be 111 *because the other seven sequences of length 3 have already been used as syndromes. This enables us to find a representative relatively easily (compared with searching through the first seven rows), by seeing how to combine known coset leaders and their syndromes to obtain 111.*

6.3 Cyclic Codes

Definition 6.8. *In the paper of [Peterson and Brown, 1961], a linear code C is called a cyclic code if it has the following property:*

If $(c_0, c_1, c_2, ..., c_{n-1}) \in C$, then it is also reality that $(c_1, c_2, ..., c_{n-1}, c_0) \in C$. From this definition the automorphism group Aut(C) of a code C is the set of permutations $\delta \in S_n$ such that $\delta(c) \in C$ for all $c \in C$, where $\delta(c_0, c_1, c_2, ..., c_{n-1}) =$

$(c_{\delta(0)}, ..., c_{\delta(n-1)})$. *In other words, the code, C is cyclic if and only if the permutation $\delta = (0, 1, 2, ..., n-1)$ is in Aut (C)[Roberts and Vivaldi, 2005].*

Example 6.10. *Let C be a subspace of a vector space $V(6,7)$ and consider the code words $v = (345601)$ of C, then C is cyclic code if $(4,5,6,0,1,3)$; $(5,6,0,1,3,4)$; $(6,0,1,3,4,5)$; $(0,1,3,4,5,6)$; $(1,3,4,5,6,0)$, all are elements of C. We can define an algebraic structure by looking at cyclic code if we let C to be a cyclic code over the field F_q and we set $R_n \doteqdot F_q[x]/(x^n-1)$. We can take the elements of R_n as polynomials of degree at most $n-1$ over F_q, where multiplication can be happen except that $x^n = 1, x^{n+1} = x$, and so on. From this, we can deduce one to one correspondence between polynomials and the code words of cyclic code as can be seen in [Sziklai, 2013].*

Example 6.11. *Let C be a subspace of a vector space $V(5,7)$ over $F_7 = \mathbb{Z}/7\mathbb{Z}$ and let consider the code word $(1,2,3,5,6)$. Then we can find the polynomial of degree less than 5 correspond to this code word which is given by $1 + 2x + 3x^2 + 5x^3 + 6x^4$. To find the elements of $R_n \doteqdot F_q[x]/(x^n-1)$, we do it as found for the previous case of quotient finite fields, and these are precisely the polynomials of degree at most $n-1$, hence the total number of the elements of $R_n \doteqdot F_q[x]/(x^n-1)$, are q^n elements.*

Example 6.12. *Let find the elements of $R_3 \doteqdot F_2[x]/(x^3-1)$, here our $q = 2$ and $n = 3$, therefore the total number of the elements of this polynomial field are $q^n = 2^3 = 8$ polynomials of degree less than 3 whose coefficients are in F_2. So the elements $R_3 \doteqdot F_2[x]/(x^3-1)$ are $0, 1, x, 1+x, x^2, x^2+1, 1+x+x^2, x+x^2$.*

Theorem 6.1. *From this kind of cyclic codes we define also an ideal of R_n given by $I_C \doteqdot (c(x) \doteqdot c_0 + c_1x + ... + c_{n-1}x^{n-1}) \in R_n c \doteqdot (c_0, c_1, ..., c_{n-1}) \in C$*

Proof.

Let $c, d \in I_C$, $a \in R_n$, then we want to show that $c - d \in I_C$ and $ac \in I_C$, therefore $c(x) = c_0 + c_1x + ... + c_{n-1}x^{(n-1)}, d(x) = d_0 + d_1x + ... + d_{(n-1)}x^{(n-1)}$ and $a(x) = a_0 + a_1x + ... + a_{n-1}x^{(n-1)}$. So, $c(x) - d(x) = c_0 - d_0 + (c_1 - d_1)x + ... + (c_{n-1} - d_{n-1})x^{(n-1)} \in I_C \Rightarrow (c_0 - d_0, c_1 - d_1, ..., c_{n-1} - d_{n-1}) \in C$.
$C \in I_C \Leftrightarrow (c_0, c_1, ..., c_{n-1}) \in C$
$d \in I_C \Leftrightarrow (d_0, d_1, ..., d_{n-1}) \in C$.
$(c_0 - d_0, c_1 - d_1, ..., c_{n-1} - d_{n-1}) \in C$, is a code word of cyclic code C (since C is a vector space of $V(n,q)$ over F_q. It remains to show that $a(x)c(x)$ is an element of I_C. Then $a(x)c(x) = (a_0 + a_1x + ... + a_{n-1}x^{(n-1)})(c_0 + c_1x + ... + c_{n-1}x^{(n-1)}) = a_0c_0 + a_0c_1 + a_0c_2 + ... a_1c_0 + a_1c_1 + ... + a_2c_0 + ...$, is also a code word of length $n-1$.

Let illustrate by using example, let $a(x) \in R_3 \doteqdot F_2[x]/(x^3 - 1)$, and $c(x) \in I_C$, we have $a(x) = a_0 + a_1x + a_2x^2$ and $c(x) = c_0 + c_1x + c_2x^2$, where $a_i \in F_q$ for $i = 0, 1, 2$ and $c_i \in C$ for $i = 1, 2, 3$. Then $a(x)c(x) = (a_0+a_1x+a_2x^2)(c_0+c_1x+c_2x^2) = a_0c_0+(a_0c_1+a_1c_0)x+(a_0c_2+a_1c_1+a_2c_0)x^2+(a_1c_2+a_2c_1)x^3+(a_2c_2)x^4$. But $x^3 = 1$ and $x^4 = x$, then we have $a_0c_0+a_1c_2+a_2c_1+(a_0c_1+a_1c_0+a_2c_2)x+(a_0c_2 + a_1c_1 + a_2c_0)x^2 \in I_C$.

$\Rightarrow (a_0c_0 + a_1c_2 + a_2c_1; a_0c_1 + a_1c_0 + a_2c_2; a_0c_2 + a_1c_1 + a_2c_0) \in C$.
$\Rightarrow (a_0c_0, a_0c_1, a_0c_2) + (a_1c_2, a_1c_0, a_1c_1) + (a_2c_1, a_2c_2, a_2c_0)$.
$\Rightarrow a_0(c_0, c_1, c_2) + a_1(c_2, c_0, c_1) + a_2(c_1, c_2, c_0)$.
But $(c_0, c_1, c_2), (c_2, c_0, c_1), (c_1, c_2, c_0) \in C$ since C is cyclic code. Therefore I_C is an ideal of R_n. □

Theorem 6.2. *Let I_C be an ideal of $R_(n)$ and let $g(x) \in C$ be monic polynomial of minimal degree $l = deg(g(x))$. Then*

a. *$g(x)$ is the only monic polynomial of degree l in I_C.*

b. *$g(x)$ generates I_C as an ideal of R_n.*

Proof.

Let f be any other non- zero monic polynomial of minimal of I with degree less than l then $f - g \in I$, but $f \neq g \Rightarrow f - g \neq 0$, $f(x) - g(x) = c_kx^k + ... + c_1x + c_0$ and this polynomial is not monic, it becomes monic if we divide it by c_k^{-1} with $c_k \neq 0$, and then we get $1/c_k(f(x)-g(x)) = x^k+...+d_1x+d_0$, where $d = c_i/c_k$ for $i = 0, 1, ..., k$. Hence $k < l$ which contradicts that l is the minimal degree. Therefore, $g(x)$ is unique monic polynomial of the minimal degree. $g(x)$ generates I means that $I =< g >= gh, h \in R_n$, this also means if $f \in I$, then $f = gh$ for some $h \in R_n$. Let $f \in I \subset R_n = F_q[x]/(x^n - 1)$, write $f(x) = g(x)q(x) + r(x) \in I$ with $deg(r(x)) < deg(g(x)) = l$.
$\Leftrightarrow f(x) - g(x)q(x) = r(x) \in I$ (since $q(x), g(x) \in I$).
$\Rightarrow r(x) = 0$
$\Rightarrow f(x) = g(x)q(x)$
$\Rightarrow f \in< g >$ and $I \in< g >$ But $g \in I$, so $< g >\in I$. Hence $I =< g >$. □

7 Conclusion

This paper has discussed about finite fields whereby some important definitions, propositions, theorems and their proofs have been given in order to capture what finite fields are and how finite fields deal with operations in different ways from usual known operations that may be performed for a set of integers. The operations procedure required any arithmetic followed by reduction of p, and this is the reason why several tables from finite fields $\mathbb{Z}_2$ to $\mathbb{Z}_{17}$ are computed to highlight how one may compute in finite fields. It includes polynomials arithmetic operations over finite fields such as addition, subtraction, multiplication, and division. The arithmetic polynomials over finite fields are computed by using the reduction of p to its coefficients, because their coefficients are drawn from finite fields that are taken into consideration. Besides polynomials computational over finite fields, this paper also explains what are cyclic codes and their applications. This research paper has further shown the applications of finite fields in the most important domain of communication regarding algebraic coding theory, code error-detection and error-correction, whereby coding and decoding schemes using cosets representative and syndromes table are discussed by using tangible examples. From this paper one may learn about finite fields and its applications and be able to extend up to $p-1$ class residues with p being any prime number or any power of a prime number.

8 Acknowledgement

I would like to acknowledge the financial support from the Eastern Africa Universities Mathematics Programme-International Science Programme, University of Rwanda Node (EAUMP-ISP, UR-Node).

References

Parity check matrix recognition from noisy codewords. *arXiv preprint arXiv:1205.4641*, 2012.

Arjeh M Cohen, Hans Cuypers, and Hans Sterk. *Some tapas of computer algebra*, volume 4. Springer Science & Business Media, 2013.

David A Cox. *Galois Theory.*, volume 61. John Wiley & Sons, 2011.

Guang Gong, Katalin Gyarmati, Fernando Hernando, Sophie Huczynska, Dieter Jungnickel, Gohar M Kyureghyan, Gary McGuire, Harald Niederreiter, Alina Ostafe, and Igor E Shparlinski. *Finite fields and their applications: character sums and polynomials*, volume 11. Walter de Gruyter, 2013.

Rudolf Lidl and Harald Niederreiter. *Introduction to Finite Fields and their Applications*. Cambridge University Press, 2 edition, 1994. doi: 10.1017/CBO9781139172769.

William Wesley Peterson and Daniel T Brown. Cyclic codes for error detection. *Proceedings of the IRE*, 49(1):228–235, 1961.

Vera Pless. *Introduction to the theory of error-correcting codes*, volume 48. John Wiley & Sons, 1998.

John AG Roberts and Franco Vivaldi. Signature of time-reversal symmetry in polynomial automorphisms over finite fields. *Nonlinearity*, 18(5):2171, 2005.

Lajos Rónyai. Galois groups and factoring polynomials over finite fields. *SIAM Journal on Discrete Mathematics*, 5(3):345–365, 1992.

Mohamed Sayed. Coset decomposition method for decoding linear codes. *International Journal of Algebra*, 5(28):1395–1404, 2011.

Péter Sziklai. *Applications of polynomials over finite fields*. PhD thesis, ELTE TTK, 2013.

Joachim von zur Gathen, Igor E Shparlinski, and Henning Stichtenoth. Finite fields: Theory and applications.

Ratio Mathematica Volume 39, 2020, pp. 111-136

Teaching as a decision-making model: strategies in mathematics from a practical requirement

Viviana Ventre[*]
Eva Ferrara Dentice
Roberta Martino[†]

Abstract

The need in the current social context to adopt teaching methods that can stimulate students and lead them towards autonomy, awareness and independence in studying could conflict with the needs of students with specific learning disorders, especially in higher education, where self-learning and self-orientation are required. In this sense, the choice of effective teaching strategies becomes a decision-making problem and must, therefore, be addressed as such. This article discusses some mathematical models for choosing effective methods in mathematics education for students with specific learning disorders. It moves from the case study of a student with specific reading and writing disorders enrolled in the mathematical analysis course 1 of the degree course in architecture and describes the personalised teaching strategy created for him.

Keywords: decision-making; inquiry model; social skills; personalised didactic strategy; Analytic Hierarchy Process.
2010 AMS subject classification: 91B06, 97D60.[‡]

---

[*]. Corresponding author: viviana.ventre@unicampania.it
†. All the authors belong to the Department of Mathematics and Physics of the University of the Study of Campania "L. Vanvitelli", viale A. Lincoln n.5, I-81100, Caserta (CE), Italy.
‡ Received on November 20th, 2020. Accepted on December 22th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.559. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

In recent years the institutions have increased their interest in a very worrying phenomenon that concerns Italy and generally speaking the countries in the Western world: the growth of 'disaffection' towards mathematics due to a traditional didactic approach to the subject (Piochi, 2008). Young people coming out of secondary schools often have the idea that mathematics consists of mechanical processes, seeing it as an arid and pre-packaged discipline whose understanding and description seem impersonal. The mathematics one learns at school is very often a set of basic notions, axioms and definitions given by the teacher and practically impossible to discuss, causing the view of a subject that is "*already done*" and immutable (Castelnuovo, 1963). The experience of mathematicians, on the other hand, is very different: mathematics is something extremely changeable whose results are the result of hard work, debate and controversy. So, axioms and definitions first presented in textbooks come into reality only at the end, when the whole structure of the problem is understood. Then, the following question arises: what is mathematics? Definitions such as "*mathematics is the science of numbers and forms*" accepted 200 years ago is now reductive and ineffective because mathematics has developed so rapidly and intensely that no definition can take into account all the multiple aspects (Baccaglini-Frank, Di Martino, Natalini, Rosolini, 2018 (A)). The list of applications of this discipline in daily life could be endless, and so could the list of motivations that could be given to pupils to convince them to study.

About this matter it is really important the following statement: "*No doubt, mathematical knowledge is crucial to produce and maintain the most important aspects of our present life. This does not imply that the majority of people should know mathematics.*" (Vinner, 2000). Mathematics can also cause terror in students (the phenomenon of "*fear for mathematics*" (Bartilomo and Favilli, 2005)) or a state of dissatisfaction with the common conception that "*you have to be made for it*" so much that even great professionals boast that they have never understood anything about mathematics. So, is it necessary to teach mathematics to everyone? The answer is simple: apart from the fact that having a basis in mathematics is a cultural question regardless of the future job, mathematics teaches to evaluate multiple aspects of a question, and provides knowledge and skills in order to consciously face a discussion defending one's own positions with responsibility and respect for the arguments of others (National Indications, 2007).

The key role of mathematics education in the development of rational thinking and with it the responsibilities of mathematics teachers at all levels is therefore underlined. Already in 1958, the theme of the congress of the Belgian

Mathematical Society was entitled “*The human responsibility of the mathematics teacher*” (Castelnuovo, 1963).

The most effective way to bring students closer to mathematics is, therefore, the image of a “*method for dealing with problems, a language, a box of tools that allows us to strengthen our intuition*” (Baccaglini-Frank, Di Martino, Natalini, Rosolini, 2018 (A)).

2 Mathematical education: theories and models

2.1 The concept of error and the inquiry model: mathematics as a humanistic discipline

Mathematics is one of the disciplines in which many 'students' manifest difficulties that compromise the relationship with the subject. A student who comes out of secondary school has a long series of 'failures' accompanied by the conviction that she can never do mathematics because she is not good at it. The problem lies in identifying errors and difficulties in mathematical learning with the conviction that the absence of errors certifies the absence of difficulties and on the other hand the absence of difficulties guarantees the absence of errors (Zan, 2007 (A)).

This identification leads to the didactic objective of obtaining the greatest number of correct answers by nourishing the "*compromise of correct answers*" (Gardner, 2002): on the one hand the teacher chooses activities that are not "*too*" difficult and on the other hand the students elaborate the answers expected by the teacher in a reproductive way. Of course, this method does not guarantee any learning, revealing itself dangerous and counterproductive (Di Martino, 2017). Moreover, with it the fear of making mistakes arise and also the conviction that mathematics is not for everyone (for instance: you can't study mathematics if you do not have a good memory!) (Zan, Di Martino, 2004). In order to face the identification of difficulty-error, there is, therefore, a need to revolutionise the conception of error and to convey to students that *"making a mistake at school may not be perceived as something negative to avoid at all, because it could be an opportunity for new learning (and teaching) opportunities to be exploited*" (Borasi,1996).

The Inquiry model is a teaching-learning model that proposes a positive and fundamental role of errors in mathematics teaching. This model sees knowledge as a dynamic process of investigation where cognitive conflict and doubt represent the motivations to continuously search for a more and more refined understanding. Therefore, instead of eliminating ambiguities and contradictions to avoid confusion or errors, these elements must be highlighted to stimulate and give shape to ideas and discussions. Questions such as "*what would happen if this result were true?"* or "*under what circumstances could this error be corrected?*" lead to a reformulation of the problem where the error is only the starting point for a deeper understanding. Communication in the classroom plays a fundamental role, and so does the conception of mathematics as a humanistic discipline: the teacher provides the necessary support for the student's autonomous search for understanding, who in turn is an active member of a research community (Tematico, Pasucci, 2014). Learning turns out to be a process of constructing meanings, and in this way, the students also understand that what is written in textbooks is the result of debates and arguments and not simply something for its own sake ("*falling from the sky*").

2.2 Cooperative learning: development of disciplinary and social skills

Some studies highlight the need to build learning teaching models that take into account students' emotions, perceptions and culture based on the idea that human learning has a specific social character (Radford, 2006). The collaborative group and peer tutoring are two models that take on both the disciplinary dimension and the affective and social dimension and facilitate discussion in the classroom. In fact, in most cases, the teacher cannot give everyone the opportunity to express themselves, nor is he able to solicit the interventions of those who are not used to intervene.

Collaborative learning, instead, sees the involvement of all the students in two successive moments: first within the individual group and then in the final discussion in class. The necessary conditions for such learning are positive interdependence and the assignment of roles: the first is reached when the members of the group understand that there can be no individual success without collective success; the second condition allows the distribution of social and disciplinary competences among the various members of the group favouring collaboration and interdependence. The recognition of roles also helps to overcome problems such as low self-esteem or a sense of ineffectiveness, allowing social skills to grow: knowing how to make decisions, how to express one's own opinions and listen to those of others, how to mediate and share, how to encourage, help and resolve conflicts are skills that the school must teach with the same care with which disciplinary skills are taught.

Dialogue among peers guarantees greater freedom and spontaneity: the majority of students identify that among peers there is no fear of expressing doubts and perplexities, the main motivation that justifies the effectiveness of such models (Baldrighi, Pesci, Torresani, 2003; Pesci, 2011).

2.3 Recovery and enhancement interventions: breaking the educational contract

The variety of possible processes, the fact that behind correct answers there can be difficulties and that some mistakes can come out of significant thought processes, brings important elements to support the criticism of the identification between mistakes and difficulties. For example, the incorrect resolution of a problem is not necessarily due to the inability to manage the mathematical structure of the proposed situation but is probably due to a lack of understanding of the problem itself. The understanding of a text is not always immediate because it involves the student's personal knowledge of common words and scripts. Understanding is, therefore reduced to a selective reading that aims to identify the numerical data and the right operations suggested by keywords (Zan, 2012).

Recovery interventions must therefore be based on the analysis of the processes that led the student to make mistakes, shifting the attention from the observation of errors to the observation of failed behaviour with the sole objective of change. The student, in turn, must take responsibility for her own recovery and therefore there is a need for teaching that makes her feel that she is the protagonist of new situations and not simply the executor of procedures to be applied to repetitive exercises (Zan, 2007 (B)).

The teacher must propose exercises and problems that do not favour a mechanical approach but question the rules that pupils are used to use and that form part of the so-called *teaching contract* (D'Amore, 2007; D'Amore, Gagatsis,1997). The idea of a didactic contract was born to explain the causes of elective failure in mathematics, that is, the kind of failure reserved only for mathematics by students who instead do well in other subjects. The didactic contract holds the interactions between student and teacher and is made up of "*the set of teacher's behaviours expected by the student and the set of student's behaviours expected by the teacher*" (Brousseau, 1986).

This explains the students' belief that a problem or exercise always has a solution because it is the teacher's job to make sure that there is only one answer to the proposed question and that all the data is necessary (Baruk, 1985).

In Bagni (1997) the following goniometry test is proposed to fourth-year students in three classes of scientific high school (students aged 17 to 18).
Determine the values of x belonging to ℝ for which it results:

a) sinx = $1/2$	b) cosx = $1/2$
c) sinx = $1/3$	d) tgx = 2
e) sinx = $\pi/3$	f) cosx = $\pi/2$
g) sinx = $\sqrt{3}$	h) cosx = $\sqrt{3}/3$

Table 2.1 Experiment in Bagni, 1997.

Remember that the goniometric functions are often introduced by making initial reference to the values they assume in correspondence to relatively common angles of use, so we have the well-known table shown in the next page (Table 2.2.)

x	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	$2\pi/3$	$3\pi/4$	$5\pi/6$	π	…
sinx	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0	…
cosx	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0	$-1/2$	$-\sqrt{2}/2$	$-\sqrt{3}/2$	-1	…
tgx	0	$\sqrt{3}/3$	1	$\sqrt{3}$	*n.d.*	$-\sqrt{3}$	-1	$-\sqrt{3}/3$	0	…
…	…	…	…	…	…	…	…	…	…	…

Table 2.2 Values assumed by common angles where "*n.d.*" is "*not define*".

Let us now examine the test: agreed time 30 minutes and pupils were not allowed to use protractor tables nor scientific calculator. It has been conceived with:

- two "*traditional*" questions (a), (b);
- two possible questions, but with the results not included between the values of x "*of common use*" (c), (d);
- two impossible questions (e), (f) but with values (of *sin*x and *cos*x) that recall the measurements in radians of "common use" angles ($\pi/3$, $\pi/2$);
- two questions (g), (h) where the first impossible and the second possible. They propose instead values (of sin*x*, cos*x*) that are included in the table referred to the angles "*of common use*" but in relation to other goniometric functions (tg*x*, cotg*x*).

Well, as far as the answers to the questions (e), (f) are concerned, the didactic contract has led some pupils to look for solutions anyway; and the "*solutions*" that most spontaneously presented themselves to their mind are the ones that they see associated, in the case of the sinus function, the two values $\pi/3$ and $\sqrt{3}/2$ and, in the case of the cosine function, the two values $\pi/2$ and 0. So we have, for instance, the following errors:

$$\sin x = \pi/3 \quad \text{so} \quad x = \sqrt{3}/2$$
$$\cos x = \pi/2 \quad \text{so} \quad x = 0$$

As far as the answers to questions (g), (h) are concerned, the reference to the tangent function was clearly expressed in the answers of some students: also in this case, some students, not finding the proposed values among those corresponding to the most frequently used x values (for the sine and cosine functions, in the table above), were induced to look for another correspondence in which the proposed values are involved. We then find errors such as:

$$\text{if } \sin x = \sqrt{3}\text{, then } x = \pi/3$$
$$\text{if } \sin x = \sqrt{3}\text{, then } x = \pi/3 + k\pi$$

What has now been pointed out obliges us to conclude that the need that leads the student to always and in any case look for a result for each proposed exercise is unstoppable: *breaking the teaching contract can be used as a* *teaching strategy* to overcome the mechanical approach used by the students and enhance knowledge (Bagni, 1997).

3 Concept image ad concept definition

These notions were developed to analyse the learning processes of mathematical definitions (Tall and Vinner, 1981). Concept image is the whole cognitive structure related to the concept and includes all mental images, the properties and processes of recall and manipulation associated with a concept, bringing into play its meaning and use. It is built through years of experience of all kinds, changing with the encounter of new stimuli and the growth of the individual. The concept definition is the set of words used to specify a concept and turns out to be personal and can often differ from the formal definition because it represents the reconstruction made by the student and the form of the words he uses to explain his concept image. It can change from time to time and for each individual the concept definition can generate its own concept image which can be called in this case concept definition image. The acquisition of a concept occurs when a good relationship is developed between the concept name, the concept image and the concept definition. Students tend to learn definitions in a mechanical way and this can lead to conflict factors when concept image or concept definition are invoked at the same time which conflict with another part of the concept image or concept definition acquired on the same concept. To explore this topic a questionnaire was administered to 41 students with an A or B grade in mathematics. They were asked: "*Which of the following functions are continuous? If possible, give your reason for your answer*."

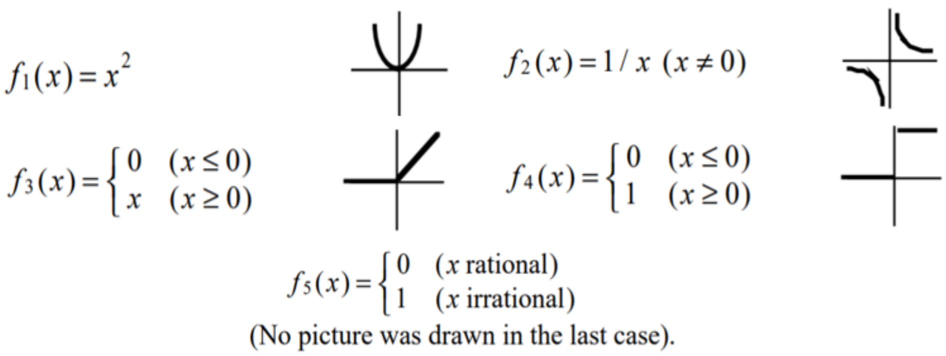


Figure 3.1 Images from Tall and Vinner, 1981.

We see that the concept image of this topic comes from a variety of resources such as the colloquial use of the term "*continuous*" in phrases such as "*It rained all day long*". So, often the use of the term "*continuous function*" implies the idea that the graph of the function can be drawn continuously. The answers are summarised in the table shown in Figure 3.2.

	1	2	3	4	5
continuous	41	6	27	1	8
discontinuous	0	35	12	28	26
no answer	0	0	2	2	7

Figure 3.2 Tables from Tall and Vinner, 1981. It Summarises the results of the experiment.

The reasons given to justify the discontinuity of $f_2(x)$ and $f_4(x)$ are of the type: "*The graph is not in a single piece*", "*There is no single formu*la". In these answers, we see that many students invoked a concept image including a graph without any interruption or a function defined by a "*single formula*". Instead, there are many continuous functions that conflict with the concept images just mentioned as the following:

$$f(x) = \begin{cases} 0\ (x < 0\ or\ x^2 < 2) \\ 1(x > 0\ or\ x^2 > 2) \end{cases}$$

whose graph is:

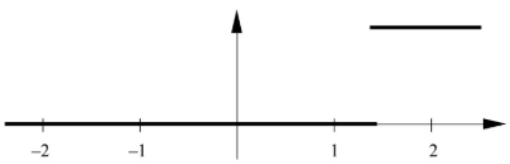


Figure 3.3 Image from Tall and Vinner, 1981. It represents the function defined above.

The idea that emerges from similar issues is that *mathematical concepts should be learned in the everyday, not technical, way of thinking*, starting with many examples and non-examples through which the concept image is formed and then arriving at a formal definition. Students should use the formal definition, but in order to internalise the concept it is necessary to aim at cognitive conflicts between concept image and concept definition. To do this it is necessary to give tasks that do not refer only to the concept image for a correct resolution, inducing the students to use the definition (Baccaglini-Frank, Di Martino, Natalini and Rosolini , 2018 (B)).

4 Teaching as a decision problem

Today more than ever, the world of education has to work on the construction of personalities that can favour to all the students with freedom of choice and reactivity. The social context in which we live is complex because it comprehends factors of unpredictability and uncertainty: the educational systems have the job to provide a path that aims to thought and action autonomy.

The school, therefore, has an orientation character, where the term "*orientation*" indicates a continuous and personal process that involves awareness, learning and education in choice (Biagioli, 2003). In particular, placing orientation as the main purpose of teaching "*means developing strategies, methodologies and contents aimed from the acquisition of awareness to understanding the complex society and the mechanisms that govern the world of studies and work*" (Guerrini, 2017). To give the proper and necessary instruments to the student, in order to activate the auto-orientation processes, the teacher has to chose what the best didactic strategy is. Therefore, on an operational point of view, the teaching is a decisional problem and has to be faced as it is. Indeed, we can speak of decision when in a situation there are: *alternatives* (being able to act in several different ways), *probability* (the possibility that the results relating to each alternative will be achieved) and the *consequences* associated with the results. Such factors are characteristic of the school world. So, to realise the best didactic strategy it is necessary to start with a representation of the problem: only through the calculation of the expectations and the evaluation of the results, it is possible to choose the right option.

Decisions can be studied in terms of absolute rationality or limited rationality. The first model ideally combines rationality and information by preferring the best alternative; the second recognises the objective narrowness of the human mind by proposing the selection of the most satisfactory alternative (Lanciano, 2019-20).

It is important to emphasise that the consequences of a decision are determined also by the context in which the decision-making process is developed. On the basis of the decision maker's knowledge of the state of nature. we distinguish various types of decisions:

- decisions in a situation of certainty: when the decision-maker knows the state of nature;
- decisions in risk situations: when the decision-maker does not directly know each state of nature, but has a probability measure for them;
- decisions in situations of uncertainty: when the decision maker has neither information on the state of nature nor the probability associated with it.

The decision maker can adopt two kinds of approaches:

- Normative approach. which bases the choice with reference to rational decision-making ideals;
- Descriptive approach which analyses how to make a decision based on the context.

So, the teacher has to consider on the basis of the objectives and the context the various alternatives, and for each one of them, the possible consequences. For each pair (alternative, circumstance) the teacher obtains a result according to a utility function.

However, the decision is subjective: it is based on the criterion of obtaining a maximum value for the utility function. Moreover, even if the choice is rational, it is made in terms of limited rationality because, in general, there are few alternatives, but it increases as the teacher expands his/her culture and experience (Delli Rocili, Maturo, 2013; Maturo, Zappacosta, 2017).

4.1 A model for evaluating educational alternatives

Multi-criteria decision analysis (MCDA) provides support to the decision maker, or a group of decision makers, when many conflicting assessments have to be considered, especially in data synthesis phase while working with complex and heterogeneous pieces of information.

Let $A = \{A_1, A_2, ..., A_m\}$ be the set of the alternatives, i. e. the possible educational strategies. Let $O = \{O_1, O_2, ..., O_n\}$ be the set of the objectives that we want to achieve. Let $D = \{D_1, D_2, ..., D_k\}$ be the set of the decision making processes. The first phase consists of the establishment of a procedure that is able to assign to each couple (alternative A_i, objective O_j) a p_{ij} score. In this way, the responsible for the decision measure the grade in which the alternative A_i satisfies the objective O_j. Assume that p_{ij} is in [0, 1], where:

- $p_{ij} = 0$ if the objective O_j is not at all satisfied by A_i;
- $p_{ij} = 1$ if the objective O_j is completely satisfied by A_i.

At the end of the procedure we obtain a matrix $P = [p_{ij}]$ of the scores which is the starting point of the elaborations that lead to the choice of the alternative, or at least to their ordering, possibly even partial (Maturo, Ventre, 2009a, 2009b). There may be constraints: it could be necessary to establish for each objective O_j a threshold $j > 0$, with the constraint $p_{ij} \geq j$, for each i.
Furthermore, through a convex linear combinations of alternatives A_i it is possible to take into consideration mixed strategies that will have the following form:

$$A(h_1, h_2, ..., h_m) = h_1 A_1 + h_2 A_2 + ... + h_m A_m$$

with:

- $h_1, h_2, ..., h_m$ non-negative real numbers;
- the h_i's are such that $h_1 + h_2 + ... + h_m = 1$;

The number h_i can represent the fraction of time in which the teaching strategy A_i is adopted. If we consider also the mixed strategies, then the single alternatives A_i are called pure strategies. The mixed strategies are particularly considered in presence of "*at risk*" alternatives: these situations have high scores for certain objectives and low for others (possibly below the threshold).

It is appropriate to construct a ranking of the alternative educational plans, i. e., a linear ordering of the alternatives that takes into account the objectives which contribute to the most suitable formation of the student. Such a ranking can be usefully obtained by means of the application of the Analytic hierarchy process, a procedure due to T. L. Saaty (1980, 2008).

4.2 The Analytic Hierarchy Process: attributions of weights and scores

The Analytic Hierarchy Process (AHP) is both a method and a technique that allows to compare alternatives of different qualitative and quantitative nature, not easily comparable in a direct way, through the assignment of numerical values that specify their priority. The first thing to do is represent the elements of the decision problem through the construction a hierarchical structure. Indeed, the Analytic Hierarchy Process is based on the representation of the problem in terms of a directed graph G = (V, A). Let us recall that (Knuth, 1973):

- a *directed graph*, or *digraph*, is a pair G = (V, A), where V is a non-empty set whose elements are called *vertices* and A is a set of ordered pairs of vertices, called *arcs*;
- a vertex is indicated with a Latin letter; for every arc (u, v) u is called the *initial vertex* and v the *final vertex* or *end vertex*;
- an ordered n-tuple of vertices $(v_1, v_2, ..., v_n)$, n > 1, is called a *path* with length n -1, if, and only if, every pair (v_i, v_{i+1}), i = 1, 2, …, n-1, is an arc of G.

Furthermore, in our context, we assume the following conditions be satisfied from a directed graph:

- the vertices are distributed in a fixed integer number $n \geq 2$ of levels; each level is indexed from 1 to n;
- there is only one vertex of level 1, called the *root* of the directed graph;
- for every vertex v different from the root there is at least one path having the root as the initial vertex and v as the final vertex;
- every vertex u of level i < n is the initial vertex of at least one arc and there are no arcs with the initial vertex of level n;
- if an arc has the initial vertex of level i<n, then it has the end vertex in the level i+1.

Let us describe, considering for example n=3, functional aspects of each level:

- the level 1 vertex is called the *general objective* and denoted GO. It indicates the objective of the entire decision making process;
- the vertices of level 2 are called *criteria.* With this level we indicate the parameters used to evaluate the alternatives;
- vertices of level 3 are called *alternatives* that represent the various ways of reaching the GO.

So we have the structure shown in the next page (Figure 4.1).

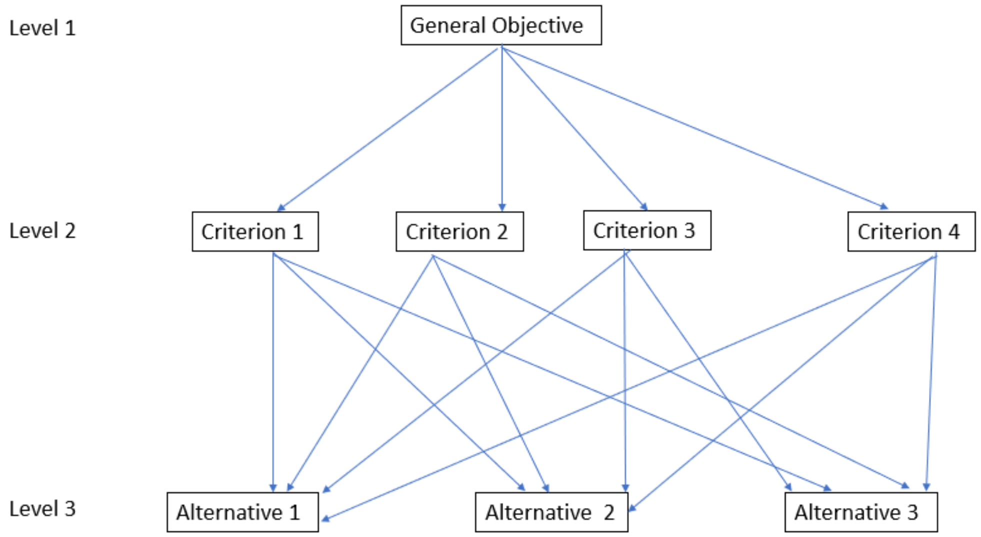


Figure 4.1 The Digraph of the Functional Aspects.

For instance, to build up a decisional model for the mathematical didactic the elements of the hierarchy might be the following:

- General objective: to guarantee a route that, through education, ensures to all students the formation of knowledge and the development of action and communication skills;
- Criterion 1: acquisition of the objectives set out in the educational plan;
- Criterion 2: ability to measure oneself against peers in a safe and rational way while respecting the ideas of others;
- Criterion 3: internalisation of the objectives set by the didactic plan. With the term "*internalisation*" we indicate the ability to re-elaborate knowledge from a critical and personal point of view;
- Criterion 4: ability to cope with trials, planned or not, without negative moods, anxiety and terror of judgement;
- Alternative 1: new didactics, inspired by the inquiry model that puts the pupil and her emotions at the center of the context with cooperative learning experiences;
- Alternative 2: traditional didactic, that is a model of theaching-learning that prefers frontal lessons. In terms of learning this alternative hypothesises that the acquisition and internalisation take place at the same time and that the error is the manifestation of the failure to complete one of the two processes.
- Alternative 3: distance learning, inspired by the inquiry model mediated through the use of technological tools characterised by a total absence of sharing the same physical space between student and teacher.

A decision-maker assigns a score to each arc following the AHP procedure (Saaty, 1980, 2008; see also: Maturo, Ventre, 2009a, 2009b). So, the second step consists in determining the ratios of preference of the elements of a level over any of the higher, i.e. previous, level: we, therefore, compare alternatives A1, A2 and A3 with respect to each criterion and the individual criteria with respect to the general objective. In order to determine the value of each comparison, the following scale of evaluations is used:

Value a_{ij}	**Interpretation**
1	*i* and *j* are equally important
3	*i* is a little more important than *j*
5	*i* is quite more important than *j*
7	*i* is definitely more important than *j*
9	*i* is absolutely more important than *j*
1/3	*i* is a little less important than *j*
1/5	*i* is quite less important than *j*
1/7	*i* is definitely less important than *j*
1/9	*i* is absolutely less important than *j*

Table 4.1 Scale of evaluations.

Matrices are called *pairwise comparison matrices* and they represent quantitative preferences between criteria or between alternatives and satisfy the followings:

- if alternative *i* assumes the value x in comparison with alternative *j* with respect to a criterion, then the comparison of alternative *j* with alternative *i* with respect to the same criterion assumes the value 1/x. Analogous is the procedure to assign values when comparing couples of criteria;
- since equally important alternatives correspond to value 1, the diagonal of the matrices are composed entirely of unit values.

In our case, we obtain the matrices shown in the next page (from Table 4.2 to Table 4.6) whose values have been assigned due to the following considerations:

- acquisition and internalisation are two different processes and only through a mutual combination of them the full formation of the student can be guaranteed;
- traditional didactic is far from the social character of human learning;
- the relationship with others is a fundamental space for personal and social development, necessary for the student to learn to respect rules and roles;
- distance learning offers insufficient physical interaction between student-teacher and student-student: expressions and gestures make the difference in the learning process;
- exercising young people to face tests in a lucid way is a fundamental aspect for the construction of a personality that faces in a competitive way the working challenges of a competitive society.

O	C1	C2	C3	C4
C1	*1*	*5*	*1*	*1*
C2	*1/5*	*1*	*1/5*	*1/7*
C3	*1*	*5*	*1*	*1*
C4	*1*	*7*	*1*	*1*

Table 4.2 Matrix M1: comparison of criteria by the ratio of preference with respect to the general objective

C1	A1	A2	A3
A1	*1*	*5*	*1*
A2	*1/5*	*1*	*1/3*
A3	*1*	*3*	*1*

Table 4.3 Matrix M2: comparison of alternatives by the ratio of preference with respect to the criterion 1

C2	A1	A2	A3
A1	*1*	*5*	*7*
A2	*1/5*	*1*	*3*
A3	*1/7*	*1/3*	*1*

Table 4.4 Matrix M3: comparison of alternatives by the ratio of preference with respect to the criterion 2

C3	A1	A2	A3
A1	*1*	*7*	*5*
A2	*1/5*	*1*	*3*
A3	*1/7*	*1/3*	*1*

Table 4.5 Matrix M4: comparison of alternatives by the ratio of preference with respect to the criterion 3

C4	**A1**	**A2**	**A3**
A1	*1*	*3*	*7*
A2	*1/3*	*1*	*3*
A3	*1/7*	*1/3*	*1*

Table 4.6 Matrix M5: comparison of alternatives by the ratio of preference with respect to the criterion 4

After constructing the pairwise comparison matrices, we have to fix the weights of the elements in each level. This step is fundamental to determine whether the matrices are relevant through a scale of values ranging from 0 to 1. These weights must meet the normality condition:

$$w_1+w_2+\ldots+w_n=1$$

This procedure supposes that, if the decision maker knew all the actual weights of the elements of the pairwise comparison matrix, then it would be:

$$A=\left({}^{w_i}/_{w_j}\right)=\begin{pmatrix} {}^{w_1}/_{w_1} & \cdots & {}^{w_1}/_{w_n} \\ \vdots & \ddots & \vdots \\ {}^{w_n}/_{w_1} & \cdots & {}^{w_n}/_{w_n} \end{pmatrix}$$

In this case the weights would be obtained from any of the rows which are all multiple of the same row $\left({}^{1}/_{w_1}, {}^{1}/_{w_2}, \ldots, {}^{1}/_{w_n}\right)$. It follows that the matrix A has rank 1. Being $w=(w_1, w_2, \ldots, w_n)^T$ we get:

$$Aw = nw.$$

Thus, from the equation above, n is an eigenvalue of A and w is one of the eigenvectors associated with n. Since the elements on the diagonal are all 1, denoted with $\lambda_1, \lambda_2, \ldots, \lambda_n = n$ the eigenvalues of A, the value of the trace of A is:

$$tr(A)=\lambda_1+\lambda_2+\ldots+\lambda_n=n$$

As *n* is an eigenvalue of A the other *n-1* eigenvalues of A must be zero. The matrix A satisfies the condition:

$$\mathbf{a}_{ij}\mathbf{a}_{jk}=\mathbf{a}_{ik}$$

for every i, j, k, called *consistency condition*, that implies transitivity of the preferences, and A is said to be *consistent* is said to be A.

In practice the decision maker does not know the vector *w*: the a_{ij} values that he assigns according to his judgement may deviate from the unknown w_i/w_j. So the decision maker may produce inconsistent pairwise comparison matrices. However the closer the a_{ij} values are to w_i/w_j, the closer the maximum eigenvalue is to n and the closer the other eigenvalues are to zero.

Therefore the vector of the weights $w'^T=(w'_1, w'_2, \ldots, w'_n)$ associated to the maximum eigenvalue (among the infinite w'^T we choose the one for which $w'_1+w'_2+\ldots+w'_n=1$) will be an estimate of the vector $w^T=(w_1, w_2, \ldots, w_n)$ the more precise the more the maximum eigenvalue λ_{max} of A is close to n, what is due to the continuity of the involved operations (Ventre, 2019). Where:

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

is defined as the *consistency index* of A, and n is the order of the matrix itself. The consistency index reveals how far from consistency the matrix is. In our case with the help of MATLAB we obtain:

	M1	M2	M3	M4	M5
λ_{max}	4,0142	3,0291	3,0649	3,0649	3,0070
CI	0,0047	0,0145	0,0324	0,0342	0,0035

Table 4.7 CI Values for each matrix.

We can therefore write the local priorities obtained by proceeding with the last step of the AHP method which, through the aggregation of the relative weights of each level, provides a weighted ranking of the alternatives.

The third step implements the estimation of local assessments, i.e. the weightings that express the relative importance of the elements of a hierarchical level over any element of the next higher level.

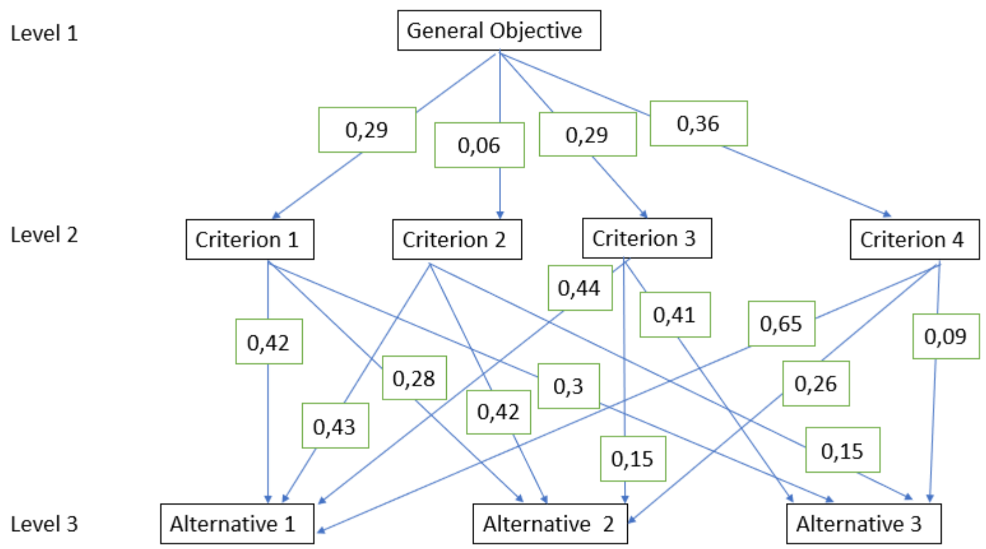


Figure 4.2 The Digraph of the Functional Aspects, with weights.

We observe that scores are nonnegative real numbers and such that the sum of the scores of the arcs coming out of the same vertex u is equal to 1. The score assigned to an arc (u, v) indicates the extent to which the final vertex v meets the initial vertex u: the *score of a path* is the product of the scores of the arcs that form the path.

- For every vertex v different from GO the score p(v) of v is the sum of the scores of all the paths that start from GO and arrive in v.

- For every level, the sum of the points of the vertices of level i is equal to 1.

The global properties determined will then be:

- A1=0,42*0,29+0,43*0,06+0,44*0,29+0,65*0,36=0,51;
- A2=0,28*0,29+0,42*0,06+0,15*0,29+0,26*0,36=0,24;
- A3=0,3*0,29+0,15*0,06+0,41*0,29+0,09*0,36=0,25.

We can conclude that, in order to achieve the objective, a new teaching strategy is preferred to a traditional one and direct communication and human contact are factors not to be overlooked. Therefore, teaching cannot be reduced to an online practice because the presence of the student is not attributable to a "*virtual presence*".

5 Mathematics and specific learning disorders

5.1 Background

Students with specific learning disorders (SLD) need a personalised learning plan that appropriately accommodates their difficulties. At the university a student with SLD attending the course of mathematical analysis has succeeded, thanks to the semester tutoring activity dedicated to the subject, to face his difficulties by passing the exam at the first useful date. The course has evolved through a personalised didactic strategy mostly based on the use of mind maps and peer comparison activities. Specific learning disorders SLD usually involve reading, writing and calculation skills. In Italy, two students with dyslexia out of three do not receive an adequate diagnosis of the disorder and therefore SLD are one of the main causes of school difficulties with important negative repercussions also in the personal sphere of the individual.

It is therefore clear the importance of a conscious environment able to respect the "*different*" way of learning of a student with SLD whose cognitive abilities and physical characteristics are in the norm. In fact, although they have an intelligence appropriate to their age, unlike their peers, these subjects learn at a slower pace because during the study they dissipate most of their energy to compensate for their disorders. Initiatives promoted by the MIUR, accompanied by individual schooling, support the right to study of students with SLD, which since the beginning of the year two thousand is also protected at the legislative level (MIUR law no.170/2010). To this end, the Regional School Offices are committed to promote the issue of detailed certifications that allow as much as possible, together with parents and the figures who follow the student in school activities, a Personalised Educational Plan that aims to achieve the same objectives of peers through compensatory tools and dispensative measures (MIUR, 2011).

5.2 SLDs in mathematics

Specific reading disorder

On the whole, this disorder leads to difficulties in reading, which is slow and incorrect, making it difficult to understand the text and to distinguish useful parts from those containing additional information (ICD F81.0, 2007). In mathematics, these deficits make even the simplest problems complex as the student may not be able to distinguish between hypotheses and data. In addition, to maximise learning, rather than studying theory from the book, we suggest the constant use of mind maps that allow you to correct and rework ideas slowly until the subject is mastered. This type of maps is the one that best suits the way of learning of students with specific reading disorders because, through images and colors, it aims to stimulate visual memory through the insertion of mental associations.

Specific writing disorder

The specific writing disorder is called dysgraphia if it affects writing and dysorthography if it affects spelling. The former involves a deficit of a motor nature and refers to the graphic aspects of handwriting, the latter is a text encoding disorder that therefore involves the linguistic component. Disgraphers therefore produce poorly readable texts (even by themselves) with words that are often misaligned and characterised by letters of different sizes, while disorthographers manifest errors such as inversion of syllables, arbitrary cuts of words and omissions of letters in words making the content unclear. In both cases the elaboration of a written text is a difficult and long process with serious repercussions also in the mathematical field. In fact, mathematics has its own language, characterised by symbols, signs and letters of the greek and latin alphabets: just think of the use of lowercase greek letters in the geometric field to indicate angles and uppercase latin letters to indicate vertices. In the set of symbols, the symbol of belonging "$\in$" may not be decoded, leading to confusion with "*E*" even though it does not denote, from a didactic point of view, a lack of understanding of the set meaning itself.

Inaccurate writing also causes errors in the resolution of algebraic expressions, for example by confusing the letter "s" and the number "5", or in the resolution of a linear system which requires many transcription steps. Another difficulty due to alignment is found in the case of powers where base and exponent are confused with a multi-digit number (34 instead of 3^4). Along this line, the teacher should take into account the content rather than the form, in the process of evaluating the written texts. Errors should not be penalised when the concept expressed is clear, whereas oral verifications should acquire more weight in the final evaluations. In order to deal with these problems, the teacher, in the process of evaluating the written tests, could take into account the content rather than the form, not penalising errors when the concept expressed is clear and giving more weight to oral verifications.

Specific disturbance of arithmetic activities

The specific disturbance of arithmetic activities disorder implies a deficit that concerns the mastery of fundamental calculation skills which generally can be divided into two different profiles according to the type of error. The first profile is defined as "*Weakness in the cognitive structuring of numerical cognition components*" and summarises both the difficulty in comparing/quantifying the elements of one or more sets and the reduced counting capacity. This type of disorder is inspired by Butterworth's studies (1999; 2005): he hypothesised the existence of a "*mathematical brain*" specialised in classifying the world in terms of numbers. The second profile is renamed with "*Difficulty in acquiring calculation procedures and algorithms*" and includes the following three cases (Temple, 1991): dyslexia for digits indicating an incorrect reading/writing of the numbers (the student sees the number 3 and pronounces 6); procedural dyscalculia indicating the difficulty in the choice, application and maintenance of procedures leading to errors in borrowing, carry-over or sticking; dyscalculia for arithmetic facts which leads to confusion between the rules of rapid access with the consequent compromise of the acquisition of numerical facts within the calculation system. For both profiles examined the tools to compensate could be tables, diagrams, calculators and an extensive collection of procedural examples.

Objective

Learning disorders are, therefore varied and create different deficits that require different compensatory instruments and dispensation measures. For this reason, it is important to experiment the different strategies for teaching in order to identify a scheme which, although it can never be universal, can be taken as a canvas and then refined according to the personal limits and objectives of the student. The aim is, therefore to encourage learning with the aim of making the student as autonomous as possible, also increasing the level of self-esteem and personal gratification. Below is the strategy used for a student with specific reading and writing disorders during the tutoring activity of the mathematical analysis course 1.

The case study

The student, after presenting his certificate at the beginning of the course, immediately showed interest in possible remedial activities. Although the student was initially autonomous, after almost a month, the first difficulties began to emerge. This situation prompted the student to make constant use of the tutoring hours in which a personalised teaching strategy was constructed. The construction of such a strategy was obtained through the procedure previously shown: alternatives, objectives and criteria were modified considering the specific difficulties. The method proved to be successful: indeed, the student was able to acquire his study methodology, which proved to be effective.

Stuff and methods

From a didactical point of view, it is particularly interesting that problems related to short-term memory are a common feature of the different SLDs. Baddeley and Hitch (1974) extended this concept, in the field of cognitive psychology, to the "*working memory*" which specifically indicates that set of notions necessary for written and oral productions that remain short in the student's mind. If this capacity is reduced, then temporary archiving and the first management/manipulation of data will be compromised with the following consequences: difficulty in taking notes, difficulty in maintaining attention, need for longer periods. To cope with this situation, it is necessary to reduce the information load by giving priority to fundamental concepts and using support tools that favour direct observation and experimentation.
Some indications for the didactic strategy are:

- Use of schemes and concept maps;
- Dispense with reading aloud and mnemonic study;
- Privileging learning from experience and laboratory teaching;
- Encourage students to self-assess their learning processes;
- Encourage peer tutoring and promote collaborative learning;
- Guarantee longer times for written tests and study;
- Take an encouraging attitude to improve self-esteem;
- Evaluate according to progress and difficulties;
- Use of calculator and digital devices.

In our case, the implementation of the didactic strategy took place through afternoon meetings, lasting two hours, usually held after the lesson held by the teacher in the same morning. At first, the meeting was based on the study of the last topics explained by the teacher and already at this stage it was evident to what extent the characteristic features of dysgraphia hindered learning: the bare and confused notes, characterised by large empty spaces, were practically impossible to read and to study. In order to try to provide constructivist learning and to avoid the use of the book, since reading was slow and incorrect due to dyslexia, the theory was flanked by practice or questions to answer. Once the new subject was finished, work was done on the previous ones.

In order to prevent the pupil from distracting himself, while following me with interest, I often drew his attention by naming him and repeating the concept in different ways. Difficulties began to come up from the first topics when, in the numerical set exercises, there was a lot of confusion between the parentheses used to describe the intervals. For example:

- interval from 'a' to 'b' not including the extremes,]a, b[;

- interval from 'a' to 'b' including the extremes, [a, b];

To cope with this difficulty, with a bit of imagination, I made him imagine drawing two arms 'embracing' the number if the number had to be present in the interval or, otherwise, they refused to do so. In this way, together with the use of graphic representations, the pupil acquired the competence to distinguish between the two types of parentheses and rarely found confusion until the end of the course. The mind maps, which we built together, were of great help especially in solving equations and inequalities with absolute values. Initially, the student's difficulty consisted in not being able to “*visualize*” the writing of the systems that came out of the procedure. With the use of a map, similar to the one below, he was able, after several lessons, to carry out the simplest exercises correctly but still presenting difficulties for the more complex ones. The result was satisfactory, however, because I believe that these difficulties were linked not so much to a lack of understanding of the absolute value function, but rather to a lack of ability to concentrate for so long on the same procedure. Example of a map for the absolute value function:

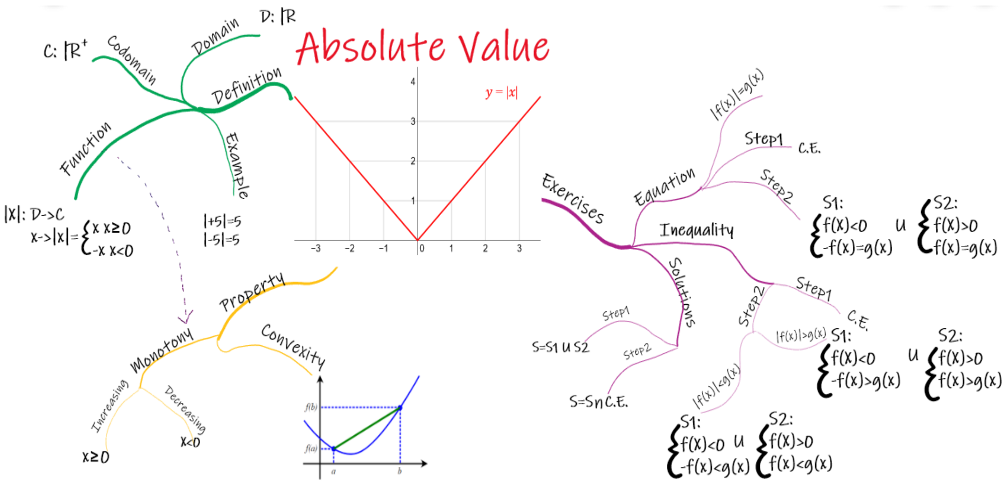


Figure 5.1 Example of a map built during the activity.

Towards the end of the course, as the exam date was approaching, other students also started to attend tutoring lessons on a regular basis. For this motive, the last topics of the course, derived and function study, were addressed through a collective study during which the student with specific learning disorders discussed with his classmates to the point of realising, under my guidance, his mistakes. During the lessons we made extensive use of the calculator, reducing the material to be memorised as much as possible. For example, to study the domain of a function, initially the student made extensive use of the tables summarising the domains of elementary functions but, subsequently, using the calculator (we get "*ERR*" if the function is not applied to an elementary of the domain) he learned what this procedure really represented by obtaining, on his own, the conditions for simple functions such as roots and logarithms.

Results

The boy achieved the minimum objectives, and at the end of the course, he passed the exam. The fundamental concepts were acquired and therefore, the test was approached independently. The improvements were gradual: as the tutoring lessons increased, the student was able to follow more and more. The shared study experience with peers was certainly helpful as, through peer comparison, he gained such confidence that he could guide his peers in difficulty during the exercises. In this regard, it is important to stress that the objectives achieved are not only didactic in nature but also personal: the ability to compare oneself with peers while respecting the ideas of others has increased the level of self-esteem and gratification.

6 Conclusion

The purpose of this document is, therefore, to underline that by using Saaty's hierarchical structure it is possible to choose the didactic strategies that best suit the various situations. In fact, all students and in particular students with SLD can maximise their learning if followed correctly. In our case, the didactic plan allowed for peer learning. This strategy, having been applied to only one student, cannot be generalised because each case requires different objectives and tools depending on the disturbance and the problems that this entails. To conclude, we stress that the affective dimension has played a fundamental role: a serene and stimulating environment has been a necessary condition to motivate students to improve themselves. However, the suggested indications are a good start to deal with different cases.

References

[1] Baccaglini-Frank, A. E., Di Martino, P., Natalini, R. and Rosolini, G., (2018). *Didattica della matematica*. Mondadori Università. A) Ch. 1.1 (pp.1-3). B) Ch. 3.4.2.

[2] Baddeley, A.D. (1986). *Working Memory*. Oxford University Press. Clarendon Press. p. 46.

[3] Baddeley A.D. and Hitch G. J. L. (1974). *Working Memory*, in Q. J. Exp. Psychology, vol. 18, n. 4, pp. 302-9, DOI:10.1080/14640746608400047, PMID 5956072.

[4] Baldrighi, A., Pesci, A., and Torresani, M. (2003). *Relazioni disciplinari e sociali nell'apprendimento cooperativo. Esperienze didattiche e spunti di riflessione*. Atti Matematica e Difficoltà n. 12 "Osservare, valutare, orientare gli alunni in difficoltà", pp. 170-178.

[4] Bagni, G.T. (1997). *Trigonometric functions: learning and didactical contract.* In D'Amore, B. & Gagatsis, A. (Eds.), Didactics of Mathematics–Technology in Education (pp. 3-10). Erasmus ICP–96–G–2011/11, Thessaloniki.

[5] Bartilomo, R. and Favilli, F. (2005). *Chi ha paura della matematica? Alla ricerca delle origini del problema*. Università di Pisa. Corso di Laurea in Matematica.

[6] Baruk, S. (1985), *L'âge du capitain*, Seuil, Paris.

[7] Biagioli, R. (2003). *L'orientamento formativo*, Pisa: ETS.

[8] Borasi, R. (1996). *Reconceiving mathematics instruction. A focus on errors.* Norwood, NJ: Ablex.

[9] Brousseau, G. (1986), *Fondaments et méthodes de la didactique des mathématiques.* Recherches en didactique des mathématiques, 7, 2, 33-115.

[10] Butterworth, B. (1999). *The mathematical brain*. London: Macmillan (trad. it. L'intelligenza matematica. Milano: Rizzoli).

[11] Butterworth, B. (2005), *The development of arithmetical abilities*. Journal of Child Psychology and Psychiatry, 46, 3–18.

[12] Castelnuovo, E. (1963), *Didattica della matematica.* Firenze, La Nuova Italia.

[13] Cecchi, F. (2018). *Didattica della matematica con studenti che presentano disturbi specifici dell'apprendimento*. Università di Bologna, Corso di Studio in Matematica.

[14] Coppola, C. and Di Martino, P. (2018). *Il problem solving come strategia per una diversa gestione dell'errore nell'educazione matematica al primo ciclo*. Annali online della Didattica e della Formazione Docente, 9(14), 76-90.

[15] D'Amore, B. (2007). *La didattica della matematica, oggi.* (2007). La matematica e la sua didattica. Atti del I Convegno Nazionale, Giulianova.

[16] D'Amore, B. and Gagatsis, A. (1997). *Didactics of mathematics*. Technology in education: Erasmus ICP, Thessaloniki.

[17] Delli Rocili, L. and Maturo, A. (2013). *Teaching mathematics to children: social aspects, psychological problems and decision making models*. Interdisciplinary approaches in social sciences, Editura Universitatii A. I. Cuza, Iasi, Romania.

[18] Di Martino, P. (2017). Problem solving e argomentazione matematica. *Didattica della matematica. Dalla ricerca alle pratiche d'aula, 1*, 23-37. DOI:

[19] Gardner, H. (2002). Educare al comprendere. Stereotipi infantili e apprendimento scolastico. Feltrinelli, Milano.

[20] Guerrini, V. (2017). La dimensione orientante nell'insegnamento. Una competenza chiave per la scuola del XXI secolo. *Formazione e Insegnamento.* Rivista internazionale di Scienze dell'educazione e della formazione, 15(2), 165-174.

[21] Knuth, D.E. (1973). The art of computer programming. London: Addison-Wesley

[22] Lanciano, T. (A.A 19-20). Decision making e Processi motivazionali, LM – Psicologia

[23] Maturo, A. and Ventre, A.G.S. (2009a). "An Application of the Analytic Hierarchy Process to Enhancing Consensus in Multiagent Decision Making", Proceeding of the International Symposium on the Analytic Hierarchy Process for Multicriteria Decision Making, July 29- August 1, 2009, paper 48, 1-12. University of Pittsburg.

[24] Maturo, A. and Ventre, A.G.S. (2009b). "Aggregation and consensus in multiobjective and multi person decision making""""." International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems vol.17, no. 4, 491-499.

[25] Maturo, A. and Zappacosta, M. G. (2017). Mathematical models for the comparison of teaching strategies in primary school. Science and Philosophy, 5(2), 25-38.

[26] Pesci, A. (2011). Sollecitare la riflessione metacognitiva in attività di tutoraggio per valorizzare le risorse di tutti gli studenti. Atti del XVII Convegno Nazionale Matematica & Difficoltà, Il senso dell'educazione matematica. Valorizzare valutando. Matematica e Difficoltà, 16, 69-78.

[27] Piochi, B. (2008). I Problemi e l'apprendimento della Matematica. CIDI Brescia, 5 dicembre 2008

[28] Radford, L. (2006, November). Comunicazione, apprendimento e formazione dell'io comunitario. In Proceedings of the 20th National Italian Conference «Incontri con la Matematica» (Bologna, 2006) (pp. 65-72).

[29] Saaty, T.L. (1980). The Analytic Hierarchy Process, New York: McGraw-Hill.

[30] Saaty, T.L., (2008), "Relative Measurement and Its Generalisation in Decision Making, Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors, The Analytic Hierarchy/Network Process". Rev. R. Acad. Cien. Ser. A. Mat., Vol. 102 (2), 251–318.

[31] Tall, D. and Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. Educational studies in mathematics, 12(2), 151-169.

[32] Tematico, N. and Pascucci, A. (2014). L'Inquiry Based Science Education IBSE nella formazione docenti e nella pratica didattica. forum.indire.it.

[33] Temple, C. M. (1991). Procedural dyscalculia and number fact dyscalculia: Double dissociation in developmental dyscalculia. *Cognitive neuropsychology*, *8*(2), 155-176.

[34] Ventre, A. (2019). *Algebra lineare e algoritmi*, Zanichelli, Bologna.

[35] Vinner, S. (2000). Procedures, rituals and ' 'man's search for meaning. Lecture given at ICME, 9.

[36] Zan, R. (2007 (A)). Difficoltà in matematica. Osservare, interpretare, intervenire. Convergenze, Springer Italia.

[37] Zan, R. and Di Martino, P., Io e la matematica: una, cento, mille storie. *La didattica della matematica: una scienza per la scuola* (2004).

[38] Zan, R. (2012). La dimensione narrativa di un problema: il modello C&D per l'analisi e la (ri)- formulazione del testo. *L'insegnamento della matematica e delle scienze integrate*, 35(2), 107-126.

[39] Zan, R. (2007 (B)). L'interpretazione dei comportamenti fallimentari. Difficoltà in matematica: Osservare, interpretare, intervenire, 197-225.

[40] Zan, R. (2001). I danni del bravo insegnante. Atti del Convegno Le difficoltà in matematica: da problema di pochi a risorsa per tutti, 135-141, Castel San Pietro Terme, Italia.

Documents

[1] Indicazioni per il curricolo. Roma, September 2009.

[2] International Classification of Diseases, tenth printing (2007).

[3] MIUR, Ministero dell'Istruzione, dell'Università e della Ricerca (2010), Legge 170/2010, Nuove norme in materia di disturbi specifici dell'apprendimento.

[4] MIUR, Ministero dell'Istruzione, dell'Università e della Ricerca (2011), Linee guida per il diritto allo studio degli alunni e degli studenti con disturbi specifici di apprendimento.

Ratio Mathematica Volume 39, 2020, pp. 137-145

On Homomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$

Vinod S*
Biju G.S†

Abstract

In this paper, using elementary algebra and analysis, we characterize and compute all continuous homomorphism from $\mathbb{C}^n$ to $\mathbb{C}^m$. Also we prove that the cardinality of the set of all non-continous group homomorphism from $\mathbb{C}^n$ to $\mathbb{C}^m$ is at least the cardinality of the continuum.

Keywords: homomorphism; continuous function; Hamel basis;
2010 AMS subject classifications: 97U99. [1]

*Department of Mathematics, Government College for Women, Thiruvananthapuram, Kerala, India; wenod76@gmail.com.
†Department of Mathematics, College of Engineering, Thiruvananthapuram-695016, Kerala, India; gsbiju@cet.ac.in.
[1]Received on June 27th, 2020. Accepted on December 15th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.562. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

Hamel [1905] introduced the concept of basis for real numbers and proved its existence in 1905 by exploring functions which satisfy Cauchy's functional equation $f(x+y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$. Using the existence of such a basis, he described all solutions of Cauchy's functional equation and established the existence of discontinuous solutions. Cauchy demonstrated that any additive function is rationally homogeneous. He also proved that the only continuous additive functions are real homogeneous and thus linear, and that an additive function with a discontinuity is discontinuous throughout. Further restrictions were placed on a non-linear additive function by Darboux [1875] who showed in 1875 that an additive function bounded above or below on some interval is continuous, hence linear. A survey of the research concerning additive functions can be found in Green and Gustin [1950]

The continuous ring homomorphisms from $\mathbb{C}$ to $\mathbb{C}$ are trivial map, identity map and complex conjugation. Since $\mathbb{C}$ is a field, all non-trivial ring homomorphisms are automorphisms on $\mathbb{C}$. Thus identity map and complex conjugation are the only continuous automorphisms on $\mathbb{C}$. Any automorphisms on $\mathbb{C}$ other than identity and complex conjugation is called a "wild" automorphism on $\mathbb{C}$. Kestelman [1951] proved the existence of so-called wild automorphism on $\mathbb{C}$ and the showed that the set of such automorphisms on $\mathbb{C}$ has cardinality 2^c. Many properties of wild automorphism on $\mathbb{C}$ are still open.

Calculating the number of homomorphisms between two groups or two rings is a fundamental problem in abstract algebra. It is not easy to determine the number of distinct homomorphism between any two given groups or rings. Most of the current results in this area are limited to groups or specific types of rings. For example, Chigira et al. [2000] studied the number of homomorphisms from a finite group to a general linear group over a finite field. In a later study Bate [2007] furnished the upper and lower limits for the number of completely reducible homomorphisms from a finite group Γ to general linear and unitary groups over arbitrary finite fields and to orthogonal and symplectic groups over finite fields of odd characteristics. Matei and Suciu [2005] discusses a method for calculating the number of epimorphisms from a finitely presented group G to a finite solvable group Γ. Further discussion on homomorphisms on certain finite groups can be found in Mal'cev [1983], Riley [1971], Hyers and Rassias [1992], but the solution to the general problem is still elusive. Hence the purpose of the paper is to characterize and compute all continuous group homomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$.

2 Notations and Basic Results

Most of the notations, functions and terms we mentioned in this paper can be find in Jacobson [2013], Gallian [1994] and Kestelman [1951].

We can interpret Hamel's concept as follows. The set $\mathbb{R}$ of real numbers is a linear space over the field $\mathbb{Q}$ of rational numbers. This linear space has a basis. Namely, there exists a subset $H \subset \mathbb{R}$ such that every non-zero $x \in \mathbb{R}$ can uniquely be written as a linear combination of the elements of H with rational coefficients. That is, there exist distinct elements $h_1, h_2, \ldots, h_k$ of H and non-zero rational numbers $w_{h_1}(x), w_{h_2}(x), \ldots, w_{h_k}(x)$ such that

$$x = \sum_{i=1}^{k} w_{h_i}(x) h_i \tag{1}$$

Thus for $x \in \mathbb{R}$, by adding the terms of the form $0 \cdot h_j$ in the representation (1), we can write

$$x = \sum_{h \in H} w_h(x) h \tag{2}$$

where $w_h(x) \in \mathbb{Q}$ and $w_h(x) = 0$ for all h except for a finite number of values of h. Hamel based his argument on Zermelo's fundamental result which states that every set can be well ordered. Hamel's argument is valid for an arbitrary linear space $L \neq \{0\}$ over a field. For this reason, recently such a basis is called a Hamel basis(see also Cohn and Cohn [1981], Halpern [1966], Jacobson [2013], Kharazishvili [2017]).

If $f : \mathbb{R} \to \mathbb{R}$ is additive, then it is easy to derive

$$f(rx) = rf(x)$$

for every $r \in \mathbb{Q}$ and $x \in \mathbb{R}$. Thus, if $H \subset \mathbb{R}$ is a Hamel basis and x is a real number, we obtain

$$f(x) = f\left(\sum_{h \in H} w_h(x) h\right) = \sum_{h \in H} f\Big(w_h(x) h\Big) = \sum_{h \in H} w_h(x) f(h). \tag{3}$$

Observing that the Hamel bases of a linear space L coincide with the maximal linearly independent subsets of L the existence of a Hamel basis is established with the aid of Zorn's maximum principle.

Theorem 2.1. *Let L be a vector space over the field F. Then L has a Hamel basis.*

Theorem 2.2. *Any continuous function $f : \mathbb{C}^n \to \mathbb{C}$ which assume only rational values is constant.*

Halbeisen and Hungerbühler [2000] showed that in an infinite dimensional Banach space, every Hamel base has the cardinality of the Banach space, which is at least the cardinality of the continuum.

Theorem 2.3. *If $K \subset \mathbb{C}$ is a field and E is a Banach space over K such that dim$(E) = \infty$, then every Hamel base of E has cardinality $|E|$.*

3 Homomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$

First we will characterize all continuous group homomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$.

Theorem 3.1. *The cardinality of the set of continuous group homomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$ is equal to the cardinality of the continuum.*

Proof. Let $\phi : \mathbb{C}^n \to \mathbb{C}^m$ be a continuous group homomorphism. For $1 \leq j \leq n$; denote e_j for the n-tuple whose j^{th} component is 1 and 0's elsewhere, and denote $\hat{e}_j$ for the n-tuple whose j^{th} component is i and 0's elsewhere.

We will complete the proof by the following steps.

Step 1: $\phi(ne_j) = n\phi(e_j)$ and $\phi(n\hat{e}_j) = n\phi(\hat{e}_j)$ for all $n \in \mathbb{Z}$ and for all $j \ (1 \leq j \leq n)$.

For $n \in \mathbb{N}$, the argument is clear since ϕ is a group homomorphism.

Since ϕ is a group homomorphis,

$$\phi(-ne_j) = -\phi(ne_j) = -n\phi(e_j) \quad \text{and} \quad \phi(0e_j) = 0\phi(e_j)$$

Therefore $\phi(ne_j) = n\phi(e_j)$ for all $n \in \mathbb{Z}$ and for all $j \ (1 \leq j \leq n)$. Similarly we can prove $\phi(n\hat{e}_j) = n\phi(\hat{e}_j)$ for all $n \in \mathbb{Z}$ and for all $j \ (1 \leq j \leq n)$.

Step 2: $\phi(re_j) = r\phi(e_j)$ and $\phi(r\hat{e}_j) = r\phi(\hat{e}_j)$ for all $r \in \mathbb{Q}$ and for all $j \ (1 \leq j \leq n)$.

Let $r = \dfrac{p}{q}$, where $p \in \mathbb{Z}$, $q \in \mathbb{N}$. Then $rq = p$ and hence $rqe_j = pe_j$. So

$$\begin{aligned}
&\phi(rqe_j) = \phi(pe_j) \\
\implies\ & q\phi(re_j) = p\phi(e_j) \\
\implies\ & \phi(re_j) = \frac{p}{q}\phi(e_j) \\
\implies\ & \phi(re_j) = r\phi(e_j), \text{ for all } r \in \mathbb{Q} \text{ and for all } j\ (1 \le j \le n).
\end{aligned}$$

Similarly, $\phi(r\hat{e}_j) = r\phi(\hat{e}_j)$ for all $r \in \mathbb{Q}$ and for all j $(1 \le j \le n)$.

Step 3: $\phi(xe_j) = x\phi(e_j)$ and $\phi(x\hat{e}_j) = x\phi(\hat{e}_j)$ for all $x \in \mathbb{R}$ and for all j $(1 \le j \le n)$.

Let $x \in \mathbb{R}$ and $1 \le j \le n$. Then there is a sequence (r_m) of rational numbers such that $r_m \to x$ in $\mathbb{R}$. Then $r_m e_j \to xe_j$ as $m \to \infty$. Since ϕ is continuous at xe_j, we have

$$\begin{aligned}
\phi(xe_j) &= \lim_{m\to\infty} \phi(r_m e_j) \\
&= (\lim_{m\to\infty} r_m)\phi(e_j) \qquad ; \text{ by step 2} \\
&= x\phi(e_j)
\end{aligned}$$

Similarly, $\phi(x\hat{e}_j) = x\phi(\hat{e}_j)$ for all $x \in \mathbb{R}$ and for all j $(1 \le j \le n)$.

Step 4: Characterization of continuous homomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$.

Let $z = (z_1, z_2, \ldots, z_n) \in \mathbb{C}^n$. For $1 \le j \le n$, let $x_j = \text{Re}(z_j)$ and $y_j = \text{Im}(z_j)$. Then

$$\begin{aligned}
z &= (x_1, x_2, \ldots, x_n) + (iy_1, iy_2, \ldots, iy_n) \\
&= \sum_{j=1}^{n} x_j e_j + \sum_{j=1}^{n} y_j \hat{e}_j
\end{aligned}$$

So

$$\begin{aligned}\phi(z) &= \phi\bigg(\sum_{j=1}^{n} x_j e_j + \sum_{j=1}^{n} y_j \hat{e}_j\bigg)\\ &= \sum_{j=1}^{n}\phi(x_j e_j) + \sum_{j=1}^{n}\phi(y_j\hat{e}_j)\\ &= \sum_{j=1}^{n} x_j\phi(e_j) + \sum_{j=1}^{n} y_j\phi(\hat{e}_j)\\ &= \sum_{j=1}^{n}\mathrm{Re}(z_j)\phi(e_j) + \sum_{j=1}^{n}\mathrm{Im}(z_j)\phi(\hat{e}_j).\end{aligned}$$

Conversely, if $a_j (1 \leq j \leq n)$ and $b_j (1 \leq j \leq n)$ be $2n$ complex numbers , then the map ϕ given by

$$\phi(z_1, z_2, \dots, z_n) = \sum_{j=1}^{n}\mathrm{Re}(z_j)a_j + \sum_{j=1}^{n}\mathrm{Im}(z_j)b_j$$

is a continuous group homomorphism from $\mathbb{C}^n$ to $\mathbb{C}^m$. Hence the cardinality of the set of continuous group homomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$ is same as the cardinality of $\mathbb{C}^{2nm}$, which is the cardinality of the continuum. $\square$

Now, we provide a proof to the existence of non-continuous group homomorphism from $\mathbb{C}^n$ to $\mathbb{C}^m$.

Theorem 3.2. *The cardinality of the set of all non-continous group homomorphism from $\mathbb{C}^n$ to $\mathbb{C}^m$ is at least the cardinality of the continuum.*

Proof. Consider $\mathbb{C}^n$ as a vector space over the field $\mathbb{Q}$ of rational numbers and H be a Hamel basis of $\mathbb{C}^n$ over $\mathbb{Q}$. Then every vector $z \in \mathbb{C}^n$ can be uniquely expressed

$$z = \sum_{h\in H} w_h(z)h \tag{4}$$

where $w_h(z) \in \mathbb{Q}$ and $w_h(x) = 0$ for all h except for a finite number of values of h. Let e_0 and $\hat{e}_0$ are the zero elements in $\mathbb{C}^n$ and $\mathbb{C}^m$ respectively. Let e_1 and $\hat{e}_1$ are the $n-$tuple and $m-$tuple respectively such that first component is 1 and all other components are 0. Let h' be a fixed element in H. Define a map $\psi_{h'} : \mathbb{C}^n \to \mathbb{C}^m$ by

$$\psi_{h'}(z) = \psi_{h'}\bigg(\sum_{h\in H} w_h(z)h\bigg) = w_{h'}(z)\hat{e}_1.$$

Let $z = \sum_{h \in H} w_h(z)h$ and $z' = \sum_{h \in H} w_h(z')h$ be two elements in $\mathbb{C}^n$. Then

$$\begin{aligned}
\psi_{h'}(z+z') &= \psi_{h'}\left(\sum_{h \in H} w_h(z)h + \sum_{h \in H} w_h(z')h\right) \\
&= \psi_{h'}\left(\sum_{h \in H} [w_h(z) + w_h(z')]h\right) \\
&= [w_{h'}(z) + w_{h'}(z')]\hat{e}_1 \\
&= w_{h'}(z)\hat{e}_1 + w_{h'}(z')\hat{e}_1 \\
&= \psi_{h'}(z) + \psi_{h'}(z').
\end{aligned}$$

Hence $\psi_{h'} : \mathbb{C}^n \to \mathbb{C}^m$ is a group homomorphism.

For $z = (z_1, z_2, \ldots, z_m) \in \mathbb{C}^m$, define $\phi : \mathbb{C}^m \to \mathbb{C}$ by $\phi(z) = z_1$. Then ϕ is a continuous function. Define $g : \mathbb{C}^n \to \mathbb{C}$ by $g(z) = \phi \circ \psi_{h'}(z)$ for all $z \in \mathbb{C}^n$. Then for $z = \sum_{h \in H} w_h(z)h \in \mathbb{C}^n$,

$$g(z) = \phi \circ \psi_{h'}\left(\sum_{h \in H} w_h(z)h\right) = \phi(w_{h'}(z)\hat{e}_1) = w_{h'}(z) \in \mathbb{Q},$$

$$g(h') = g\left(1 \cdot h' + \sum_{h \in H, h \neq h'} 0h\right) = \phi \circ \psi_{h'}\left(1 \cdot h' + \sum_{h \in H, h \neq h'} 0h\right) = \phi(1 \cdot \hat{e}_1) = 1$$

and

$$g(e_0) = \phi \circ \psi_{h'}(e_0) = \phi \circ \psi_{h'}\left(0 \cdot h' + \sum_{h \in H, h \neq h'} 0h\right) = \phi(0 \cdot \hat{e}_1) = 0.$$

Hence g is a non-constant function from $\mathbb{C}^n$ to $\mathbb{C}$ which assumes only rational values. Therefore g is not continuous and which gives the function $\psi_{h'}$ is discontinuous.

Let h' and h'' be two distinct elements in H. Then

$$\psi_{h'}(h') = \psi_{h'}(1 \cdot h') = 1 \cdot \hat{e}_1 = \hat{e}_1$$

and

$$\psi_{h''}(h') = \psi_{h''}(0 \cdot h'' + 1 \cdot h') = 0 \cdot \hat{e}_1 = \hat{e}_0.$$

Therefore $\psi_{h'}$ and $\psi_{h''}$ are distinct. Then the cardinality of set of all non-continuous group homomorphism from $\mathbb{C}^n$ to $\mathbb{C}^m$ is at least $|H| = |\mathbb{C}^n|$ = the cardinality of the continuum. $\square$

4 Conclusions

In this paper, we characterized all continuous group homomorphisms from $\mathbb{C}^n$ to $\mathbb{C}^m$. Also we proved that the cardinality of the set of all non-continous group homomorphism from $\mathbb{C}^n$ to $\mathbb{C}^m$ is at least the cardinality of the continuum.

References

Michael Bate. The number of homomorphisms from finite groups to classical groups. *Journal of Algebra*, 308(2):612–628, 2007.

Naoki Chigira, Yugen Takegahara, and Tomoyuki Yoshida. On the number of homomorphisms from a finite group to a general linear group. *Journal of Algebra*, 232(1):236–254, 2000.

Paul Moritz Cohn and Paul Moritz Cohn. *Universal algebra*, volume 159. Reidel Dordrecht, 1981.

Gaston Darboux. Sur la composition des forces en statique. *Bulletin des sciences mathématiques et astronomiques*, 9:281–288, 1875.

Joseph A Gallian. *Contemporary Abstract Algebra*. D.C.Heath and Company, 1994.

JW Green and W Gustin. Quasiconvex sets. *Canadian Journal of Mathematics*, 2:489–507, 1950.

Lorenz Halbeisen and Norbert Hungerbühler. The cardinality of hamel bases of banach spaces. *East-West Journal of Mathematics*, 2:153–159, 2000.

James D Halpern. Bases in vector spaces and the axiom of choice. *Proceedings of the American Mathematical Society*, 17(3):670–673, 1966.

Georg Hamel. Eine basis aller zahlen und die unstetigen lösungen der funktionalgleichung: f (x+ y)= f (x)+ f (y). *Mathematische Annalen*, 60(3):459–462, 1905.

Donald H Hyers and Themistocles M Rassias. Approximate homomorphisms. *Aequationes mathematicae*, 44(2-3):125–153, 1992.

Nathan Jacobson. *Lectures in Abstract Algebra: II. Linear Algebra*, volume 31. Springer Science & Business Media, 2013.

Hyman Kestelman. Automorphisms of the field of complex numbers. *Proceedings of the London Mathematical Society*, 2(1):1–12, 1951.

Alexander Kharazishvili. *Strange functions in real analysis*. CRC Press, 2017.

Anatoly I Mal'cev. On homomorphisms onto finite groups. *American Mathematical Society Translations, Series*, 2(119):67–79, 1983.

Daniel Matei and Alexander I Suciu. Counting homomorphisms onto finite solvable groups. *Journal of Algebra*, 286(1):161–186, 2005.

Robert Riley. Homomorphisms of knot groups on finite groups. *mathematics of computation*, 25(115):603–619, 1971.

Ratio Mathematica Volume 39, 2020, pp. 147-163

Frattini submultigroups of multigroups

Joseph Achile Otuwe*
Musa Adeku Ibrahim†

Abstract

In this paper, we introduce and study maximal submultigroups and present some of its algebraic properties. Frattini submultigroups as an extension of Frattini subgroups is considered. A few submultigroups results on the new concepts in connection to normal, characteristic, commutator, abelian and center of a multigroup are established and the ideas of generating sets, fully and non-fully Frattini multigroups are presented with some significant results.

Keywords: Maximal, Cyclic Multigroup, Commutator and Generating Set.

2010 AMS subject classification: 55U10.‡

---

* Department of Mathematics, College of Science (Ahmadu Bello University, Zaria, Nigeria), talk2josephotuwe2016@gmail.com.

† Department of Mathematics, College of Science (Ahmadu Bello University, Zaria, Nigeria), amibrahim@abu.edu.ng.

‡ Received on October 13th, 2020. Accepted on December 19th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.534. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

The term multigroup was first mentioned in [7] as an algebraic structure that satisfied all the axioms of group except that the binary operation is multivalued. This concept was later redefined in [19] via count function of multisets and some of its properties were vividly discussed. The idea of submultigroup and its classes were established in [14]. Concept of maximal subgroups is established in [3], [5] and [17] and some of its properties were investigated. Also, normal and characteristic submultigroup were introduced in [8] and [13] respectively, and some of its properties were presented. Frattini in [11], introduced a special subgroup named Frattini subgroup and some results were obtained. Other related work on Frattini subgroup can be found in [1], [2], [6], [12], [15], [16], [18], [20] and [21]. Furthermore, in [10] Frattini subgroup was represented but in fuzzy environment called Frattini fuzzy subgroup. In this paper, we focus on multiset setting to obtain Frattini submultigroups and finally establish some related results.

In general, the union of submultigroups of a multigroup may not be a multigroup, we therefore establish some conditions under which the union of all maximal submultigroups is a multigroup. When this occur, the Frattini submultigroup obtained from such maximal submultigroups is called "fully Frattini" otherwise it is called "non-fully Frattini". Furthermore, other relevant concepts such as; cyclic multigroup, minimal generating set of a multigroup, generator and non-generator of a multigroup are introduced with reference to Frattini submultigroups. Finally, we study some properties of center of a multigroup, normal, commutator, minimal and characteristic submultigroups.

2 Preliminaries

Definition 2.1 (|23|). Let X be a set. A multiset A over X is just a pair $\langle X,\ C_A\rangle$, where X is a set and $C_A\colon X \to \mathbb{N}$ is a function. Any ordinary set B is actually a multiset $\langle B, \chi_B\rangle$, where χ_B is its characteristic function.

The set X is called the ground or generic set of the class of all multisets containing objects from X.

Definition 2.2 (|22|). Let A and B be two multisets over X, A is called a submultiset of B written as $A \subseteq B$ if $C_A(x) \leq C_B(x)$ for all $x \in X$. Also, if $A \subseteq B$ and $A \neq B$, then A is called a proper submultiset of B and denoted as $A \subset B$.

Definition 2.3 ([22]). Let A and B be two multisets over X, then A and B are equal if and only if $C_A(x) = C_B(x)$ for all $x \in X$. Two multisets A and B are comparable to each other if $A \subseteq B$ or $B \subseteq A$.

Definition 2.4 ([23]). Suppose that A, $B \in MS(X)$ such that $A = \langle X, C_A \rangle$ and $B = \langle X, C_B \rangle$.

i. Their intersection denoted by $A \cap B$ is the multiset $C = \langle X, C_C \rangle$, where $x \in X$, $C_C(x) = C_A(x) \wedge C_B(x)$.

ii. Their union denoted by $A \cup B$ is the multiset $C = \langle X, C_C \rangle$, where $x \in X$, $C_C(x) = C_A(x) \vee C_B(x)$.

iii. Their sum denoted by $A \bigoplus B$ is the multiset $C = A = \langle X, C_C \rangle$, where $x \in X$, $C_C(x) = C_A(x) + C_B(x)$.

Definition 2.5 ([19]). Let X be a group and $A \in MS(X)$. A is said to be a multigroup of X if the count function of A or C_A satisfies the following two conditions:

i. $C_A(xy) \geq [C_A(x) \wedge C_A(y)]$, $\forall x, y \in X$.

ii. $C_A(x^{-1}) \geq C_A(x)$, $\forall\, x \in X$,

Where C_A is a function that takes X to a natural number, and $\wedge$ denotes minimum operation.

The set of all multigroups defined over X is denoted by $MG(X)$.

Definition 2.6 ([19]). Let $A \in MG(X)$. Then A^{-1} is defined by $C_A(x) = C_A(x^{-1})$ $\forall x \in X$.

Thus, $A \in MG(X) \Leftrightarrow A^{-1} \in MG(X)$.

Definition 2.7 ([19]). Let $A \in MG(X)$. Then A is said to be abelian or commutative if $C_A(xy) = C_A(yx)$ $\forall\, x, y \in X$.

Definition 2.8 ([19]). Let $A \in MG(X)$. Then the sets A_* and A^* are defined as

$$A_* = \{x \in X | C_A(x) > 0\} \quad \text{and} \quad A^* = \{x \in X | C_A(x) = C_A(e)\},$$

where e is the identity element of X.

Definition 2.9 ([19]). Let $\{A_i\}_{i \in I}$, $I = 1, 2, \dots, n$ be an arbitrary family of multigroups of a group X. Then

$$C_{\cap A_i}(x) = \bigwedge C_{A_i}(x) \quad \forall\, x \in X.$$

$$C_{\cup A_i}(x) = \bigvee C_{A_i}(x) \ \ \forall\, x \in X.$$

Definition 2.10 (|14|). Let $A \in MG(X)$. Then the center of A is defined as $C(A) = \{x \in A | C_A([x,y]) = C_A(e) \forall\ y \in X\}$.

Definition 2.11 (|9|) Commutator of two Submultigroup: Let A and B be submultigroups of $C \in MG(X)$. Then the commutator of A and B is the multiset (A, B) of X defined as follows:

$$C_{(A,\ B)}(x) = \begin{cases} \bigvee_{x=[a,b]} [C_A(a) \wedge C_B(b)] & \\ 0 & otherwise \end{cases}$$

That is, $C_{(A,\ B)}(x) = \vee_{x=aba^{-1}b^{-1}}[C_A(a) \wedge C_B(b)]$. Since the supremum of an empty set is zero.

$C_{(A,\ B)}(x) = 0$ if x is not a commutator.

Definition 2.11 (|4|). Let $A \in MG(X)$. Then the order of A denoted by $O(A)$ is defined as $O(A) = \sum_{x \in X} C_A(x)$. i.e., the total numbers of all multiplicities of its element.

Definition 2.12 (|14|). Let $A \in MG(X)$. A submultiset B of A is called a submultigroup of A denoted by $B \subseteq A$ if B is a multigroup. A submultigroup B of A is a proper submultigroup denoted by $B \subset A$, if $B \subseteq A$ and $A \neq B$.

Definition 2.13 (|14|). Let $A \in MG(X)$. Then a submultigroup B of A is said to be complete if $B_* = A_*$, incomplete if $B_* \neq A_*$, regular complete if B is complete and $C_B(x) = C_B(y) \forall\ x, y\ \in X$ and regular incomplete if B is incomplete and $C_B(x) = C_B(y)\ \forall\, x, y\ \in X$.

Definition 2.14 (|8|). Let $A, B \in MG(X)$ such that $A \subseteq B$. Then A is called a normal submultigroup of B if $C_A(xyx^{-1}) \geq\ C_A(y) \forall x\ y, \in X$.

Definition 2.15 (|10|). Let X and Y be two groups and let $f: X \to Y$ be a homomorphism. Suppose A and B are multigroups of X and Y respectively, then f induces a homomorphism from A to B which satisfies

i. $C_{f(A)}(y_1 y_2) \geq C_{f(A)}(y_1) \wedge C_{f(A)}(y_2)\ \ \forall y_1, y_2 \in Y.$

ii. $C_{f(B)}(f(x_1x_2)) \geq C_{(B)}(f(x_1)) \wedge C_{(B)}(f(x_2)) \ \forall x_1, x_2 \in X$ where

i. the image of A under f denoted by $f(A)$, is a multiset of Y defined by

$$C_{f(A)}(y) = \begin{cases} \vee_{x \in f^{-1}(y)} C_A(x), & f^{-1}(y) \neq \emptyset. \\ 0, & otherwise. \end{cases}$$

for each $y \in Y$.

ii. the inverse image of B under f denoted by $f^{-1}(B)$, is a multiset of X defined by

$C_{f^{-1}(B)}(x) = C_B(f(x)) \forall x \in X$.

Definition 2.16 (|10|). Let X and Y be groups and let $A \in MG(X)$ and $B \in MG(Y)$ respectively. Then a homomorphism f from X to X is called an automorphism of A onto A if f is both injective and surjective, that is, bijective.

Definition 2.17 (|13|). Let $A, B \in MG(X)$ such that $A \subseteq B$. Then A is called a characteristic (fully invariant) submultigroup of B if

$C_{A^\theta}(x) = C_A(x) \ \forall x \in X$ for every automorphism, θ of X.

That is, $\theta(A) \subseteq A$ for every $\theta \in Aut(X)$.

3 Frattini Submultigroups and their Properties

In this section we propose the concept of minimal, maximal, Frattini, commutator submultigroups, cyclic, fully and non-fully Frattini multigroup and generating set of a multigroup with some illustrative examples.

Definition 3.1

a. **Minimal Submultigroup:** Let G be a group and $A \in MG(X)$. Then a non trivial proper submultigroup denoted by $\underline{B}$ of A is said to be minimal if there exists no other non-trivial submultigroup C of A such that $C \subseteq \underline{B}$.

Remark 3.1. Every minimal complete submultigroup of a multigroup is unique.

b. **Maximal Submultigroup:** Let X be a group and $A \in MG(X)$. Then a proper normal submultigroup denoted by $\bar{B}$ of A is said to be maximal if there exists no other proper submultigroup C of A such that $\bar{B} \subseteq C$.

c. Frattini Submultigroup: Let X be a group. Suppose A is a multigroup of X and, A_1, A_2, ..., A_n (or simply A_i for $i = 1,2,..,n$) are maximal submultigroups of A. Then the Frattini submultigroup of A denoted by $\Phi(A)$ is the intersection of A_i defined by

$$C_{\Phi(A)}(x) = C_A(x) \wedge C_{A_2}(x) \wedge \ldots \wedge C_{A_n}(x)\ \forall x \in X$$

or simply by

$$C_{\Phi(A)}(x) = \wedge_{i=1}^{n} C_{A_i}(x)\ \forall x\ \in X.$$

Remark 3.2.

i. Let X be a non abelian group and $A \in MG(X)$. If K is a normal submultigroup of A with an incomplete maximal submultigroups and M_i for each i are the maximal subgroups of A, then the maximal submultigroups of K are submultigroups of M_i.

ii. Let $A \in MG(X)$. If K is a submultigroup of A and $\Phi(K)$ is the Frattini submultigroup of K then, $\Phi(K) \sqsubseteq A$.

d. Commutator Submultigroup of a Multigroup: Let $A \in MG(X)$ such that the commutator subgroup of X is given as $X' = \{[a,b]: a,b \in X\}$. Then the commutator submultigroup of A denoted by A' is defined as

$$C_{(A,A)}(x) = \begin{cases} \max\ min\ \{\{C_A(a), C_A(b)\} : x = [a,b], \forall\ a,b \in X\} \\ \qquad\qquad 0 \qquad\qquad otherwise\ x \in A \end{cases}$$

e. Let $A \in MG(X)$. Then the sets $[\Phi(A)]_*$ and $[\Phi(A)]^*$ are defined as

$[\Phi(A)]_* = \{x \in X | C_{\Phi(A)}(x) > 0\}$ and

$[\Phi(A)]^* = \{x \in X | C_{\Phi(A)}(x) = C_{\Phi(A)}(e)\}$, where e is the identity element of X.

Remark 3.3.

i. The commutator submultigroup of every abelian multigroup is $\{e\}_n$

ii. Let $A \in MG(X)$ and A' be the commutator submultigroup of A. Then $A' \trianglelefteq A$.

Remark 3.4. Let $A, B, C \in MG(X)$ such that $B, C \sqsubseteq A$ and D' be the commutator submultigroup of B and C. Then $D' \trianglelefteq \Phi(A)$.

f. Cyclic Multigroup: Let $X = \langle a \rangle$ be a group generated by a. Then a multigroup A over X is said to be a cyclic multigroup if $\exists\ n \in \mathbb{N}$ such that $C_A(na) = C_A(x)\ \forall\ x \in X$. The element a is then called the generator of A otherwise, a non generator of A.

g. **Generating Set of a Multigroup:** Let X be a cyclic group and $A \in MG(X)$. A subset S of X is said to be a generating set for A if all elements of A and its inverses can be expressed as a finite product of elements in S with $C_A(nS) = C_A(x)\ \forall\ x \in X$ and for some $n \in \mathbb{N}$.

h. **Minimal Generating Set of a Multigroup:** Let X be a cyclic group and $A \in MG(X)$. A subset B of X is termed minimal generating set of A if $\exists\ n \in \mathbb{N}$ such that $C_A(nB) = C_A(x)$ and there is no proper subset C of B with $C_A(nC) = C_A(x), \forall\ x \in X$.

i. **Fully Frattini Multigroup:** Let $A \in MG(X)$. Then A is called fully Frattini if the union of the maximal submultigroups equals A. Otherwise, it is called non-fully Frattini. In addition, every multigroup without incomplete maximal submultigroup is called trivial fully Frattini.

4. Some Results on Frattini Submultigroups

In this section, we present some results on Frattini submultigroup of multigroups.

Theorem 4.1. Let $A \in MG(X)$ with complete maximal submultigroups. Then every minimal submultigroup of A is a submultigroup of $\Phi(A)$.

Proof. Suppose $\Phi(A)$ is the Frattini submultigroup of A then A has maximal submultigroup say $A_i\ \forall\ i$ such that $\Phi(A) = \wedge_{i=1}^{n} A_i$. Since A is multigroup over X, A has a minimal submultigroup say $\underline{A}$. If $\underline{A}$ is not a submultigroup of $\Phi(A)$ then there exist at least an element $x \in \Phi(A)$ such that $C_{\Phi(A)}(x) < C_{\underline{A}}(x)\ \ \forall\ x \in X$ which contradicts the fact that $\underline{A}$ is a minimal submultigroup of A . Hence $\underline{A}$ is a submultigroup of $\Phi(A)$.

Theorem 4.2 If $\theta(\Phi(A)) \subset \Phi(A)$ for all $\theta \in Aut(A)$, then $\Phi(A)$ is characteristic in A.

Proof. Since θ is an automorphism, the inverse θ^{-1} is also an automorphism of A. Hence we have $\theta^{-1}\,\theta\big(\Phi(A)\big) \subset \theta\big(\Phi(A)\big)$.

Applying θ, we have $\theta\theta^{-1}\,\big(\Phi(A)\big) \subset \theta\big(\Phi(A)\big)$. Then we obtain

$\Phi(A) = \theta\theta^{-1}\,\big(\Phi(A)\big) \subset \theta\big(\Phi(A)\big) \subset \Phi(A)$. By this fact, equality holds and so $\theta\big(\Phi(A)\big) = \Phi(A)$. Hence the Frattini submultigroup is characteristic in A.

Theorem 4.3 Every Frattini submultigroup of a multigroup is characteristic.

Proof. By Theorem 4.2, it suffices to proof that $\theta(\Phi(A)) \subset \Phi(A)$ for every automorphism $\theta \in Aut(A)$.

Let $x \in \theta(\Phi(A))$. Then there exists $y \in \Phi(A)$ such that $x = \theta(y)$.

To show that $x \in \Phi(A)$, we consider an arbitrary element $g \in A$. Then since θ is an automorphism, we have $A = \theta(A)$. Thus there exists g' in A such that $g = \theta(g')$.

We have $C_A(xg) = C_A(\theta(y)\theta(g'))$

$= C_A(\theta(yg'))$ (Since θ is a homomorphism)

$= C_A(\theta(g'y))$ (Since $y \in \Phi(A)$)

$= C_A(\theta(g')\theta(y))$ (Since θ is a homomorphism)

$= C_A(gx)$

Since this is true for all $g \in A$ it follows that $x \in \Phi(A)$, and thus $\theta(\Phi(A)) \subset \Phi(A)$. Hence the result.

Theorem 4.4 Every Frattini submultigroup of a multigroup is abelian.

Proof. Let $A \in MG(X)$ and $\Phi(A)$ be the Frattini submultigroup of A. It follows that $\Phi(A)$ is a normal submultigroup of A by definition 2.14 Consequently,

$$C_{\Phi(A)}(xyx^{-1}) = C_{\Phi(A)}(y) \forall\, x, y \in X.$$

Thus, $C_{\Phi(A)}(xy) = C_{\Phi(A)}(yx) \forall\, x, y \in X$.

Hence, the result follows by Definition *2.7*

Theorem 4.5 Every $\Phi(A)$ is a normal submultigroup of A.

Proof. Let $A \in MG(X)$ and $\Phi(A)$ be the Frattini submultigroup of A. Then

$C_{\Phi(A)}(e) \geq C_{\Phi(A)}(x)\ \forall\ x \in X$, since $C_{A_i}(e) \geq C_{A_i}(x)\ \forall\ x \in X$.

Now, let $x, y \in X$, then since $\Phi(A)$ is a multigroup over X by *definition 2.14*, we get

$C_{\Phi(A)}(xy^{-1}) \geq C_{\Phi(A)}(x) \wedge C_{\Phi(A)}(y)$

Now we proof that $\Phi(A)$ is a normal submultigroup of A. Let $x, y \in X$, then it follows that

$$\begin{aligned} C_{\Phi(A)}(yxy^{-1}) &= C_{\Phi(A)}((yx)y^{-1}) \\ &= C_{\Phi(A)}(x(yy^{-1})) \\ &= C_{\Phi(A)}(xe) \geq C_{\Phi(A)}(x). \end{aligned}$$

Hence, the result by Definition *2.14*

Theorem 4.6 Let A be a multigroup over a non-Abelian group X, then $C(A) \trianglelefteq [\Phi(A)]_*$.

Proof. $C(A) \neq \emptyset$, since at least $e \in C(A)$. Let $x, y \in C(A)$ then for all $z \in [\Phi(A)]_*$, $C_A([x,z]) = C_A(e)$ and $C_A([y,z]) = C_A(e)$. Consequently,

$$\begin{aligned} C_A([xy,\ z]) &= C_A([x,z]^y[y,z]) \text{ where } [x,z]^y = yx^{-1}z^{-1}xzy^{-1} \\ &\geq C_A([x,z]^y) \wedge C_A([y,z]) \\ &\geq C_A([x,z]^y) \text{ since } C_A([y,z]) = C_A(e) \\ &= C_A(y[x,z]y^{-1}) = C_A([x,z]) = C_A(e). \end{aligned}$$

Thus $xy \in C(A)$.

Now, let $x \in C(A)$. Then $C_A([x,z]) = C_A(e) \forall\ z \in [\Phi(A)]_*$.

Hence, $C_A([x^{-1},z]) = C_A(xz^{-1}x^{-1}z) = C_A(xz^{-1}x^{-1}zxx^{-1})$

$$\begin{aligned} &= C_A(z^{-1}x^{-1}zxx^{-1}x) = C_A([z,x]) \\ &= C_A([x,z]^{-1}) = C_A([x,z]) = C_A(e). \end{aligned}$$

Thus, $x^{-1} \in C(A)$ therefore, $C(A)$ is a subgroup of $\Phi(A)_*$. To show that $C(A)$ is a normal subgroup of $[\Phi(A)]_*$. Let $x \in [\Phi(A)]_*$ and $y \in \mathrm{C(A)}$

Then $xyx^{-1} = (xy)x^{-1} = (yx)x^{-1} = y \in \mathrm{C(A)}$.

Thus, $x \in [\Phi(A)]_*$ and $y \in C(A) \Rightarrow xyx^{-1} \in C(A)$. Hence, $C(A) \trianglelefteq [\Phi(A)]_*$

Remark 4.1 If A is a multigroup over an abelian group X, $[\Phi(A)]_*$ is the root set of Frattini submultigroup of A and $C(A)$ is the center of A then $[\Phi(A)]_*$ is a normal subgroup of $C(A)$.

Theorem 4.7. If A is a multigroup over a non-Abelian group X and K is a normal submultigroup of A, then $\Phi(K) \sqsubseteq \Phi(A)$.

Proof. Clearly, $\Phi(K)$ and $\Phi(A)$ are submultigroups of A.

Let M_i be the maximal submultigroups of K and M_j be the maximal submultigroups of A for each i and j. Then by *Remark* 3.2 (i) we have

$$M_i \sqsubseteq M_j \text{ for each } i \text{ and } j.$$

$$|(\cap M_i)| \leq |(\cap M_j)|$$

$\Longrightarrow |\Phi(K)| \leq |\Phi(A)|$

Therefore, $\Phi(K) \subseteq \Phi(A)$.

To show that $\Phi(K) \sqsubseteq \Phi(A)$, suppose $M_i \cap M_j = \{e\}_n$ then the result holds trivially. But if $M_i \cap M_j \neq \{e\}_n$ then for any element $a \in M_i$, $C_{M_i}(a) \leq C_{M_j}(a)$ for each i and j. Therefore, $\Phi(K) \sqsubseteq \Phi(A)$.

Theorem 4.8 If A is a regular multigroup over a group X. Then $A' \trianglelefteq \Phi(A)$.

Proof. Since A' is a multigroup over X, then $C_{A'}(xy^{-1}) \geq [C_{A'}(x) \wedge C_{A'}(y)]$ $\forall\, x, y \in X$.

Let $a \in A'$ and $b \in \Phi(A)$ then $C_{A'}(ab^{-1}) \geq [C_{A'}(a) \wedge C_{A'}(b)]$. Thus $ab^{-1} \in A'$.

Therefore A' is a submultigroup of $\Phi(A)$.

Now by Theorem 4.5, $\Phi(A) \trianglelefteq A$ and clearly $A' \trianglelefteq A$. So let $a \in A'$ and $g \in \Phi(A)$, then $C_{A'}(ag) = C_{A'}(ga)$ implies $C_{A'}(gag^{-1}) = C_{A'}(a)$. Hence $A' \trianglelefteq \Phi(A)$.

Theorem 4.9 If $A \in MG(X)$, A' is the commutator submultigroup of A and $\Phi(A)$ is the Frattini submultigroup of A. Then $\Phi(A) \trianglelefteq A'$.

Proof. Since $\Phi(A)$ is a multigroup over X, then $C_{\Phi(A)}(xy^{-1}) \geq [C_{\Phi(A)}(x) \wedge C_{\Phi(A)}(y)] \; \forall \; x, y \in X$.

Let $a \in \Phi(A)$ and $b \in A'$, then $C_{\Phi(A)}(ab^{-1}) \geq [C_{\Phi(A)}(a) \wedge C_{\Phi(A)}(b)]$. Thus $ab^{-1} \in \Phi(A)$. Therefore $\Phi(A)$ is a submultigroup of A'.

Now, by Theorem 4.5, $\Phi(A) \trianglelefteq A$. Let $a \in \Phi(A)$ and $g \in A'$, then $C_{\Phi(A)}(ag) = C_{\Phi(A)}(ga)$ implies $C_{\Phi(A)}(gag^{-1}) = C_{\Phi(A)}(a)$. Hence, $\Phi(A) \trianglelefteq A'$.

Theorem 4.10. Every Frattini submultigroup of a cyclic multigroup is abelian.
Proof. Let $\Phi(A)$ be the Frattini submultigroup of a cyclic multigroup A over a cyclic group X, then there exists $a \in X$ such that $\forall x, y \in X$ we have $C_{\Phi(A)}(na) = C_{\Phi(A)}(x)$ and $C_{\Phi(A)}(ma) = C_{\Phi(A)}(y)$ for $n, m \in \mathbb{N}$. It now follows that $C_{\Phi(A)}(x + y) = C_{\Phi(A)}(na + ma) = C_{\Phi(A)}\, a(n + m)$

$$= C_{\Phi(A)}\, a(m + n) = C_{\Phi(A)}(ma + na) = C_{\Phi(A)}(y + x).$$

Theorem 4.11. If A is a regular multigroup with an incomplete maximal submultigroups over a cyclic group X. Then $\Phi(A)$ is contained in the set of all non-generators of A. In particular, $[\Phi(A)]^*$ coincide with the set of all non-generators if A has only one maximal submultigroup.

Proof. Let X be a cyclic group and $A \in MG(X)$ and $\Phi(A)$ denotes the Frattini submultigroup of A. Let $Gen(A)$ be the set of all generators of A and $M_{i=1,2,\ldots,n}$ be the incomplete maximal submultigroups of A, then for all $x \in Gen(A)$, $x \notin M_{i=1,2,\ldots,n}$. In fact, all $y \in M_i$ is a non-generator.

Further, $A^* = Gen(A) \cup M_1{}^* \cup M_2{}^* \cup \ldots \cup M_n{}^*$ and $y \in A^* \backslash Gen(A) = \cup M_i{}^*$. Since $C_{\Phi(A)}(x) = \bigwedge_{i=1}^{n} C_{A_i}(x) \; \forall x \in X$, we have that for all $y \in \Phi(A)$, $y \in \cup A_i$. This implies that $\Phi(A) \subseteq \cup M_i$. But $\cup M_i$ is the largest set containing all non generators. Hence $\Phi(A)$ is contained in the set of all non-generators. Suppose A has only one nontrivial maximal submultigroup say B then, $A^* = Gen(A) \cup B^*$ and $y \in B$ $(y \notin Gen(A))$. Since $[\Phi(A)]^* = B^*$, therefore $y \in \Phi(A)$ for all non-generators y. Hence, $[\Phi(A)]^*$ is indeed the set of all non-generators.

Theorem 4.12. If a regular multigroup A over a cyclic group X has two incomplete maximal submultigroups, then the union of its generators coincide with the non-generating set of A^*.

Proof. Let M_1 and M_2 be the maximal submultigroups of A and $Gen(A)$ be the collection of all generators of A. Clearly, $Gen(A) \not\subseteq M_1$ and M_2 (since M_1 and M_2 does not contain any generator). Now, $[n(Gen(A))]$ can be expressed as $[n(Gen(A))] = M^* \, Gen(A)$ if n is odd and $[n(Gen(A))] = M^*$ if n is even with $M^* \cap Gen(A) = \emptyset$ for any maximal submultigroup of A.

Also, $M_1{}^* Gen(A) = A^* \backslash M_2{}^*$. That is, $[n(Gen(A))] = A^* \backslash M_2{}^*$ for odd values of n and for any M but since $A^* = Gen(A) \cup M_1{}^* \cup M_2{}^*$ for even n we have $[n(Gen(A))] = M_1{}^* = A^* \backslash M_2{}^* \neq A^*$ for any M. Hence the result.

Theorem 4.13. If a regular multigroup A over a cyclic group X has two maximal submultigroups, then the union of the non-generators coincide with the generating set of A^*.

Proof. Let M_1 and M_2 be the maximal submultigroups of A and $NGen(A)$ be the collection of all non- generators of A. Clearly, $NGen(A) \subseteq M_1{}^* \cup M_2{}^*$ and so $NGen(A)$ generates A^*.

$$NGen(A) = \{i, j, k, \dots, u\}, \text{ where } i, j, k, \dots, u \in A.$$

$$[2(NGen(A))] = \{i, j, k, \dots, u\}\{i, j, k, \dots, u\}$$

Taking every $i, j \in NGen(A)$, $ij = k \in A$ (for some $k \in A$) if $i, k \neq \{e\}_n$ (i.e., ij generates distinct elements in A).

Since $e \in NGen(A)$, we have that $[n(NGen(A))] = NGen(A) \cup (A \backslash NGen(A))$ for some $n \in \mathbb{N}$. More explicitly,

$[n(NGen(A))] = \{i, j, k, \dots, u\}\{i, j, k, \dots, u\} \dots for\ n\ time$

$= A^* = \{i, j, k, \dots u, p, r, q, \dots, t\}$ where $p, r, q, \dots, t \in Gen(A)$.

This yields $[n(NGen(A))] = NGen(A) \cup Gen(A) = A^*$ for some $n \in \mathbb{N}$.

Theorem 4.14. If a regular multigroup A over a cyclic group X has two incomplete maximal submultigroups and $Gen(A)$ is the set of generators of A, then $[n(NGen(A))]$ form one of the root set of the maximal submultigroup of A for some $n \in \mathbb{N}$.

Proof. Let A be a multigroup over a cyclic group, M_1, M_2 be the maximal submultigroups of A and $i, j, k, \dots, u$ be the generators of A.

Then, $[n(Gen(A))] = \{i, j, k, \dots, u\}\{i, j, k, \dots, u\} \dots for\ n\ time$.

Since $e \notin Gen(A)$, for all $s, t \in Gen(A)$, $s \cdot t \notin Gen(A)$, $s \cdot t \in [2(Gen(A))]$ and $s, t \notin M_{i=1,2}$.

But, $A^* = Gen(A) \cup {M_1}^* \cup {M_2}^*$. Therefore, $s \cdot t \in A^* \backslash Gen(A)$.

In particular, $[2(Gen(A)] \subseteq A^* \backslash Gen(A)$ and $[2(Gen(A)] = {M_i}^*$ for any i.

Remark 4.2

a. A generator of any multigroup over a cyclic group is not contained in any of its maximal submultigroups.

b. The set of non-generators of any multigroup may not be a submultigroup.

Theorem 4.15. If β is a minimal generating set of a multigroup A over a cyclic group X, then $\beta \not\subseteq \Phi(A)$.

Proof. Suppose $\beta \subseteq \Phi(A)$, then $\beta \subseteq M_i$, $\forall\, i$ where M_i is a maximal submultigroup of A. Now, since β contains at least one generator of A, then every M_i contains at least one generator of A which is a contradiction. Hence, $\beta \not\subseteq \Phi(A)$.

Remarks 4.3

i. If A is a multigroup over a cyclic group X. Then the union of all the minimal generating sets of A is equal to A^*.

ii. Every minimal generating set contains a non-generator.

iii. Given a multigroup A over a cyclic group X with order k, If X is a minimal generating set of A then X^m gives $m+1$ elements of A^*.

Theorem 4.16 Every irregular multigroup with complete maximal submultigroups over a group is fully Frattini.

Proof. For multigroup A to be irregular implies $\forall\, x, y \in X$, $C_A(x) \neq C_A(y)$. Now let M_i for each i be the complete maximal submultigroups of A. For M_i to be complete in A implies $M_i{}^* = A^*$. Since M_i is complete in A, then there exists $x \in M_i$ such that $C_{M_i}(x) = C_A(x)$ $\forall\, x \in X$. Therefore $\cup M_i = A$ for each i.

Theorem 4.17 Every irregular multigroup with an incomplete maximal submultigroups over a non-cyclic group is fully Frattini.

Proof. X is a non-cyclic group, implies it has no generator and $C_A(x) \neq C_A(y)$ $\forall\, x, y \in X$. Now let M_i for each i be the incomplete maximal submultigroups of A. Then for each $x \in X$, x is contained in at least one of the M_i with $C_{M_i}(x) = C_A(x)\ \forall\, x \in X$. Therefore $\cup$ for each i.

$M_i = A$

Theorem 4.18. Every cyclic multigroup with incomplete maximal submultigroups is not fully Frattini.

Proof. Suppose X is a cyclic group and A is a multigroup with incomplete maximal submultigroups over X. Where A has set of generators $Gen(A)$. Now let $Gen(A) = \{x_j\}$ for some finite j then $A^* = Gen(A) \cup M_i{}^*$, where $M_i{}^*$are the root sets of all the maximal submultigroups of A for each i and $x_j \notin M_i$.by remark 4.2a, $\cup M_i = A \backslash Gen(A) \neq A$. Hence, A is not fully Frattini.

Theorem 4.19 Every regular multigroup over a group is non-fully Frattini.
Proof. Let X be a group and A be a multigroup over X. For A to be a regular multigroup implies $\forall\, x, y \in X$, $C_A(x) = C_A(y)$. Let M_i for each i be the maximal submultigroup of A. Since A is regular and by Definition $3.1b$, there exists at least an element $x \in X$ such that $C_{M_i}(x) < C_A(x)$ for each i and so $\vee_{i=1}^{n} M_i \neq A$.

Remark 4.4. Let A be a nontrivial fully Frattini multigroup over a group X then A has at least three maximalsubmultigroups if $C_A(x) < C_A(e)\ \forall\, x \in X$ where e is the identity element of X.

5 Conclusions

Most results in Frattini subgroup are extended to multigroup. A number of new results were obtained. Notion of cyclic, generators, non-generators, minimal generating sets were introduced and results with reference to Frattini submultigroups were established. Notwithstanding, more properties of maximal and Frattini submultigroups, fully and non-fully Frattini submultigroups and also cyclic multigroups are amenable for further investigation in multigroup framework.

References

[1] W. Alice. On the Frattini Subgroup, Transactions of the American Mathematical Society, 141, 323-333. 1969.

[2] R. B. J. T. Allenby. Normal Subgroups Contained in Frattini Subgroups are Frattini Subgroups, Proceedings of the American Mathematical Society, 78(3), 315-318. 1980.

[3] M. Aschbacher, and L. Scott. Maximal Subgroup of a Finite Group, Journal of Algebra. 92(1), 44-80. 1985.

[4] J. A. Awolola, and P. A. Ejegwa. On Some Algebraic Properties of Order of an Element of a Multigroup. Quasigroups and Related Systems 25:21-26. 2017.

[5] B. A. Ballester, and L. M. Ezquerro. On Maximal Subgroups of Finite Groups, Comms in Algebra, **19**(8), 373-2394. 1991.

[6] J. C. Beidleman and T. K. Seo., Generalized Frattini Subgroups of Finite Groups, Pacific Journal of Mathematics, 23, No. 3, 441-450. 1967.

[7] M. Dresher, and O. Ore. Theory of Multigroups, American J. Math. 60. 1930.

[8] P. A. Ejegwa, and A. M. Ibrahim. Normal Submultigroups and Comultisets of a Multigroup, Quasigroups and Related Systems,25: 231-244. 2017a.

[9] P. A. Ejegwa, and A. M. Ibrahim. Some Groups Analogous Results in Multigroup Setting. Annals of Fuzzy Mathematics and Informatics, 17(3): 231-245. 2019.

[10] P. A. Ejegwa, and J. A. Otuwe. Frattini fuzzy subgroups of fuzzy groups. Annals of Communications in Mathematics, **2**(1), 24-31. 2019.

[11] F. Giovanni. Intornoalla Generazioni Dei Gruppi di Operazioni, Rend. Accad Lincei 4(1), 281-285. 1885.

[12] A. M. Ibrahim, and P. A. Ejegwa. A Survey on the Concept of Multigroups, Journal of the Nigerian Association of Mathematical Physics, 38: 1-8. 2016.

[13] A. M. Ibrahim, and P. A. Ejegwa. Characteristic Submultigroups of Multigroups, Gulf Journal of Mathematics, 5(4): 1-8. 2017a.

[14] A. M. Ibrahim, and P. A. Ejegwa. Multigroup Actions on Multiset, Annals of Fuzzy Mathematics and Informatics, 14(5): 515-526. 2017b.

[15] C. B. James, and K. S. Tae. Generalized Frattini Subgroups of a Finite Groups, Pacific Journal of Mathematics, 23 (3), 44-450. 1967.

[16] O. Javier. The Frattini Subgroup of a Group, Margarita Mathematica en Memoria de Jose Javier (CHICHO) Guadalupe Hernandes, 1-12. 2001.

[17] G. A. Miller. Maximal Subgroup of a given Group, Proc. N. A. S. Department of Mathematics, University of Illinois, 68-70. 1940.

[18] J. Mohit, and P. Ajit. Nilpotency in Frattini Subgroups, International Journal of Mathematics Trends and Technology, 44, (3),121-122. 2017.

[19] S. K. Nazmul, P. Majumdar, and S. K. Samanta. On Multisets and Multigroups. Annals of Fuzzy Mathematics and Informatics, 6(3): 643-656. 2013.

[20] B. Prabir, and N. P. Mukherjee. A Generalized Frattini Subgroup of a Finite Group, International Journal of Mathematical Science, 12 (2), 263-266. 1989.

[21] A. Stefanos, and B. B. Adolfo. On the Frattini Subgroup of a Finite Group. Journal of Algebra, (16), 1-8. 2016.

[22] D. Singh, A. M. Ibrahim, T. Yohanna and J. N. Singh. An Overview of the Application of Multisets, Novi Sad Journal of Mathematics, 37(2), 73-92. 2007.

[23] A. Syropoulos, Mathematics of Multisets, Springer-verlac Berlin Heidelberg. 2001.

Ratio Mathematica Volume 39, 2020, pp. 165-186

KCD indices and coindices of graphs

Keerthi G. Mirajkar[*]
Akshata Morajkar[†]

Abstract

The relationship between vertices of a graph is always an interesting fact, but the relations of vertices and edges also seeks attention. Motivation of the relationship between vertices and edges of a graph has helped to arise with a set of new degree based topological indices and coindices named KCD indices and coindices. These indices and coindices are elaborated by establishing a set of properties. More fascinating results of some graph operations using KCD indices are developed in this article.
Keywords: KCD indices, KCD coindices, graph operations.
2010 AMS subject classifications: 05C07, 05C76. [1]

[*]Department of Mathematics, Karnatak University's Karnatak Arts College, Dharwad - 580 001, Karnataka, INDIA; keerthi.mirajkar@gmail.com

[†]Department of Mathematics, Karnatak University's Karnatak Arts College, Dharwad - 580 001, Karnataka, INDIA; akmorajkar@gmail.com

[1]Received on October 26th, 2020. Accepted on December 17th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.550. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

Graph theory plays a vital role in the quantification of chemical structures through topological indices. Topological indices are molecular descriptors which characterize the topology of a graph through numerical parameters. Abundant number of topological indices are identified these days. Amongst these the first degree based topological indices are Zagreb indices [Gutman and Trinajstić, 1972]. Recently along with Zagreb indices Zagreb coindices is also gaining much attention for research. This has put forward versitile forms of Zagreb indices of graphs. The present work aims to establish some new form of topological indices of graphs.

This paper considers the graph to be simple, finite and undirected. The graph is denoted as $G = (V, E)$ with $|V(G)| = n$ as the vertex set and $|E(G)| = m$ as the edge set. The set of vertices are also referred to as the order of the graph G and the edge set as the size of the graph G. The edge connecting the two vertices u and v is denoted as $e = uv$. The degree of the vertex u in a graph G is denoted as $d_G(u)$ and defined as the number of edges of a graph G incident with the vertex u. The degree of edge $d_G(e)$ of a graph G is defined as $d_G(e) = d_G(u) + d_G(v) - 2$. The complement $\overline{G}$ of a graph G is one in which two vertices are adjacent if and only if they are not adjacent in G. For $\overline{G}$, $|V(\overline{G})| = n$, $|E(\overline{G})| = \overline{m} = \binom{n}{2} - m$ [Alwardi et al., 2018]. Also $uv \in E(\overline{G}) \iff uv \notin E(G)$. The degree of a vertex u in $\overline{G}$ is denoted as $d_{\overline{G}}(u)$ and defined as $d_{\overline{G}}(u) = n - 1 - d_G(u)$ [Alwardi et al., 2018]. The degree of edge of $\overline{G}$ is represented as $d_{\overline{G}}(e)$, defined as $d_{\overline{G}}(e) = d_{\overline{G}}(u) + d_{\overline{G}}(v) - 2$. For undefined terminologies refer [Harary, 1969]. The Zagreb indices were defined by Gutman and Trinajstić [Gutman and Trinajstić, 1972] as

$$M_1(G) = \sum_{u \in V(G)} d_G(u)^2 \quad (1)$$

$$M_2(G) = \sum_{uv \in E(G)} d_G(u) d_G(v) \,. \quad (2)$$

Here $M_1(G)$ refers first Zagreb index and $M_2(G)$ refers second Zagreb index. First Zagreb index is also expressed as [Došlić, 2008, Došlic et al., 2011]

$$M_1(G) = \sum_{uv \in E(G)} \Big(d_G(u) + d_G(v) \Big). \quad (3)$$

For properties and information on Zagreb indices refer [Gutman and Das, 2004, Zhou and Gutman, 2005, Zhou, 2004].

Further, Zagreb coindices were introduced by Došlić [Došlić, 2008] as

$$\overline{M_1}(G) = \sum_{uv \notin E(G)} \Big(d_G(u) + d_G(v)\Big) \tag{4}$$

$$\overline{M_2}(G) = \sum_{uv \notin E(G)} d_G(u) d_G(v). \tag{5}$$

The detailed study on Zagreb coindices is reported in [Ashrafi et al., 2010, 2011], the association between Zagreb indices and coindices is encountered in [Das et al., 2012, Gutman et al., 2015].
Shirdel et al. [Shirdel et al., 2013] defined hyper Zagreb index as

$$HM(G) = \sum_{uv \in E(G)} \Big(d_G(u) + d_G(v)\Big)^2. \tag{6}$$

Further, hyper Zagreb coindex was introduced as

$$\overline{HM}(G) = \sum_{uv \notin E(G)} \Big(d_G(u) + d_G(v)\Big)^2. \tag{7}$$

These graph invariants were studied in [Pattabiraman and Vijayaragavan, 2017, Veylaki et al., 2016]. Relationship between hyper Zagreb index and coindex is established in [Gutman, 2017].

Now, we introduce a set of new degree-based topological indices and coindices named as *Karnatak College Dharwad* indices and coindices or KCD indices and coindices in short, which is dedicated to Karnatak College Dharwad as the college has completed hundred years of its service in education to the society in the year 2017. Further the research Supervisior and research scholar belong to the same college.
i.e., The first and second KCD indices of a graph G are respectively

$$KCD_1(G) = \sum_{e=uv \in E(G)} \Bigg(\Big(d_G(u) + d_G(v)\Big) + d_G(e)\Bigg) \tag{8}$$

$$KCD_2(G) = \sum_{e=uv \in E(G)} \Big(d_G(u) + d_G(v)\Big) d_G(e). \tag{9}$$

We proceed further to define KCD coindices as follows

$$\overline{KCD_1}(G) = \sum_{e=uv \notin E(G)} \Bigg(\Big(d_G(u) + d_G(v)\Big) + d_G(e)\Bigg) \tag{10}$$

$$\overline{KCD_2}(G) = \sum_{e=uv \notin E(G)} \Big(d_G(u) + d_G(v)\Big) d_G(e). \tag{11}$$

Here $\overline{KCD_1}(G)$ and $\overline{KCD_2}(G)$ are first and second KCD coindices of a graph G respectively.

The remaining paper is distributed as follows. Section 2 expresses the properties of first KCD indices and coindices of a graph and its complement. Section 3 concentrates on properties of second KCD indices and coindices of a graph and its complement, while Section 4 is devoted for the study of KCD indices of certain graph operations.
The following previously known results are considered for present investigation.

Theorem 1.1. *[Gutman et al., 2015] Let G be a graph with n vertices and m edges. Then,*

$$\begin{aligned} M_1(\overline{G}) &= M_1(G) + n(n-1)^2 - 4m(n-1) && (12)\\ \overline{M_1}(G) &= 2m(n-1) - M_1(G). && (13)\end{aligned}$$

Corollary 1.2. *[Gutman et al., 2015] Let G be any graph and $\overline{G}$ its complement. Then*

$$\overline{M_1}(G) = \overline{M_1}(\overline{G}). \tag{14}$$

Theorem 1.3. *[Gutman, 2017] Let G be a graph with n vertices and m edges. Then,*

$$\begin{aligned} \overline{HM}(G) &= 4m^2 + (n-2)M_1(G) - HM(G) && (15)\\ HM(\overline{G}) &= 2n(n-1)^3 - 12m(n-1)^2 + 4m^2 && (16)\\ &\quad +(5n-6)M_1(G) - HM(G) && \\ \overline{HM}(\overline{G}) &= 4m(n-1)^2 + 4(n-1)M_1(G) + HM(G). && (17)\end{aligned}$$

2 Basic properties of first KCD indices and coindices

Theorem 2.1. *Let G be a graph with n vertices and m edges. Then,*

$$\begin{aligned} KCD_1(G) &= (4n-6)m - 4m(n-1) + 2M_1(G) && (18)\\ KCD_1(\overline{G}) &= 4n(\overline{m} - m) - 6\overline{m} + 4m + 2M_1(G) && (19)\\ \overline{KCD_1}(G) &= 4m(n-1) - 2\Big(\overline{m} + M_1(G)\Big) && (20)\\ \overline{KCD_1}(\overline{G}) &= (4n-6)m - 2M_1(G). && (21)\end{aligned}$$

Proof.
Proof of Eq. (18):

For any vertex u of G,

$$d_G(u) = n - 1 - d_{\overline{G}}(u). \tag{22}$$

and for any edge $e = uv$ of G,

$$d_G(e) = 2n - 4 - \Big(d_{\overline{G}}(u) + d_{\overline{G}}(v)\Big). \tag{23}$$

Thus by Eqs. (8), (22) and (23), we have

$$\begin{aligned}
KCD_1(G) &= \sum_{e=uv\in E(G)} \Bigg(\Big(d_G(u) + d_G(v)\Big) + d_G(e)\Bigg) \\
&= \sum_{e=uv\notin E(\overline{G})} \Bigg(\Big(n - 1 - d_{\overline{G}}(u) + n - 1 - d_{\overline{G}}(v)\Big) \\
&\quad + \Big(2n - 4 - (d_{\overline{G}}(u) + d_{\overline{G}}(v)\Big)\Bigg) \\
&= \sum_{e=uv\notin E(\overline{G})} \Bigg(4n - 6 - 2\Big(d_{\overline{G}}(u) + d_{\overline{G}}(v)\Big)\Bigg) \\
&= (4n-6)m - 2\sum_{e=uv\notin E(\overline{G})} \Big(d_{\overline{G}}(u) + d_{\overline{G}}(v)\Big).
\end{aligned}$$

According To Eq. (4)

$$\overline{M_1}(\overline{G}) = \sum_{uv\notin E(\overline{G})} \Big(d_{\overline{G}}(u) + d_{\overline{G}}(v)\Big).$$

Hence,

$$KCD_1(G) = (4n-6)m - 2\overline{M_1}(\overline{G}) \tag{24}$$

Substitution of Eqs. (13) and (14) in (24) results into Eq. (18).

Proof of Eq. (19):

For any vertex u of the complement $\overline{G}$,

$$d_{\overline{G}}(u) = n - 1 - d_G(u). \tag{25}$$

and for any edge $e = uv$ of the complement $\overline{G}$,

$$d_{\overline{G}}(e) = 2n - 4 - \Big(d_G(u) + d_G(v)\Big). \tag{26}$$

Bearing in mind Eqs. (8), (25) and (26), we get

$$\begin{aligned}
KCD_1(\overline{G}) &= \sum_{e=uv\in E(\overline{G})} \left(\Big(d_{\overline{G}}(u) + d_{\overline{G}}(v)\Big) + d_{\overline{G}}(e)\right) \\
&= \sum_{e=uv\notin E(G)} \left(\Big(n - 1 - d_G(u) + n - 1 - d_G(v)\Big)\right. \\
&\quad \left. + \Big(2n - 4 - (d_G(u) + d_G(v))\Big)\right) \\
&= \sum_{e=uv\notin E(G)} \left(4n - 6 - 2\Big(d_G(u) + d_G(v)\Big)\right) \\
&= (4n-6)\overline{m} - 2\sum_{e=uv\notin E(G)} \Big(d_G(u) + d_G(v)\Big).
\end{aligned}$$

Thus by Eq. (4),

$$KCD_1(\overline{G}) = (4n-6)\overline{m} - 2\overline{M_1}(G) \tag{27}$$

Employing Eq. (13) in (27) generates Eq. (19).

Proof of Eq. (20):

Using Eqs. (10), (22) and (23), we have

$$\begin{aligned}
\overline{KCD_1}(G) &= \sum_{e=uv\notin E(G)} \left(\Big(d_G(u) + d_G(v)\Big) + d_G(e)\right) \\
&= \sum_{e=uv\in E(\overline{G})} \left(\Big(n - 1 - d_{\overline{G}}(u) + n - 1 - d_{\overline{G}}(v)\Big)\right. \\
&\quad \left. + \Big(2n - 4 - (d_{\overline{G}}(u) + d_{\overline{G}}(v))\Big)\right) \\
&= \sum_{e=uv\in E(\overline{G})} \left(4n - 6 - 2\Big(d_{\overline{G}}(u) + d_{\overline{G}}(v)\Big)\right).
\end{aligned}$$

By Eq. (3)

$$M_1(\overline{G}) \;=\; \sum_{uv\in E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big).$$

Thus,

$$\overline{KCD_1}(G) \;=\; (4n-6)\overline{m}-2M_1(\overline{G}) \tag{28}$$

Substitution of Eq. (12) in (28) gives Eq. (20).

Proof of Eq. (21):

In view of Eq. (10), (25) and (26), we get

$$\begin{aligned}
\overline{KCD_1}(\overline{G}) &= \sum_{e=uv\notin E(\overline{G})}\Bigg(\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big)+d_{\overline{G}}(e)\Bigg)\\
&= \sum_{e=uv\in E(G)}\Bigg(\Big(n-1-d_G(u)+n-1-d_G(v)\Big)\\
&\quad +\Big(2n-4-(d_G(u)+d_G(v)\Big)\Bigg)\\
&= \sum_{e=uv\in E(G)}\Bigg(4n-6-2\Big(d_G(u)+d_G(v)\Big)\Bigg)\\
&= (4n-6)m-2\sum_{e=uv\in E(G)}\Big(d_G(u)+d_G(v)\Big)
\end{aligned}$$

Considering Eq. (3) we directly arrive at Eq. (21).

□

3 Basic properties of second KCD indices and coindices

Theorem 3.1. *Let G be a graph with n vertices and m edges. Then,*

$$\begin{aligned} KCD_2(G) &= HM(G) - 2M_1(G) && (29)\\ KCD_2(\overline{G}) &= 4(n-1)\Big(\overline{m}(n-2) - m(2n-3)\Big) + 4m^2 && (30)\\ &\quad +(5n-8)M_1(G) - HM(G) \\ \overline{KCD_2}(G) &= 4(n-1)(n-2)\overline{m} - (4n-6)(n-1)\Big(n(n-1) - 4m\Big) && (31)\\ &\quad +2(n-1)^2\Big(n(n-1)-6m\Big) + 4m^2 + nM_1(G) - HM(G) \\ \overline{KCD_2}(\overline{G}) &= 4(n-1)(n-2)m - (4n-6)M_1(G) + HM(G). && (32) \end{aligned}$$

Proof.

Proof of Eq. (29):

Considering Eqs. (9), (22) and (23), we have

$$\begin{aligned} KCD_2(G) &= \sum_{e=uv\in E(G)} \Big(d_G(u) + d_G(v)\Big) d_G(e) \\ &= \sum_{e=uv\notin E(\overline{G})} \Big(n-1-d_{\overline{G}}(u) + n-1-d_{\overline{G}}(v)\Big)\Big(2n-4-(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big) \\ &= \sum_{e=uv\notin E(\overline{G})} 4(n-1)(n-2) - (4n-6)\sum_{e=uv\notin E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big) \\ &\quad + \sum_{e=uv\notin E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big)^2. \end{aligned}$$

By an analogous reasoning,

$$\overline{M_1}(\overline{G}) = \sum_{uv\notin E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big) \ \ and \ \ \overline{HM}(\overline{G}) = \sum_{uv\notin E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big)^2.$$

Thus,

$$KCD_2(G) \ = \ 4m(n-1)(n-2) - (4n-6)\overline{M_1}(\overline{G}) + \overline{HM}(\overline{G}).$$

In view of Eq. (14)

$$KCD_2(G) \ = \ 4m(n-1)(n-2) - (4n-6)\overline{M_1}(G) + \overline{HM}(\overline{G}) \qquad (33)$$

Taking into account Eqs. (13) and (17), Eq. (33) results into Eq. (29).

Proof of Eq. (30):

In view of Eqs. (9), (25) and (26), we get

$$\begin{aligned} KCD_2(\overline{G}) &= \sum_{e=uv\in E(\overline{G})} \Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big)d_{\overline{G}}(e) \\ &= \sum_{e=uv\notin E(G)} \Big(n-1-d_G(u)+n-1-d_G(v)\Big)\Big(2n-4-(d_G(u)+d_G(v)\Big) \\ &= \sum_{e=uv\notin E(G)} 4(n-1)(n-2)-(4n-6)\sum_{e=uv\notin E(G)}\Big(d_G(u)+d_G(v)\Big) \\ &\quad + \sum_{e=uv\notin E(G)}\Big(d_G(u)+d_G(v)\Big)^2. \end{aligned}$$

By Eqs. (4) and (7), it directly follows

$$KCD_2(\overline{G}) = 4\overline{m}(n-1)(n-2)-(4n-6)\overline{M_1}(G)+\overline{HM}(G) \qquad (34)$$

Application of Eqs. (13) and (15) to Eq. (34) yields Eq. (30).

Proof of Eq. (31):

Using Eqs. (11), (22) and (23), we have

$$\begin{aligned} \overline{KCD_2}(G) &= \sum_{e=uv\notin E(G)} \Big(d_G(u)+d_G(v)\Big)d_G(e) \\ &= \sum_{e=uv\in E(\overline{G})} \Big(n-1-d_{\overline{G}}(u)+n-1-d_{\overline{G}}(v)\Big) \\ &\quad \Big(2n-4-(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big) \\ &= \sum_{e=uv\in E(\overline{G})} 4(n-1)(n-2)-(4n-6)\sum_{e=uv\in E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big) \\ &\quad + \sum_{e=uv\in E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big)^2. \end{aligned}$$

By reasoning,

$$M_1(\overline{G}) = \sum_{uv\in E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big) \ \ and \ \ HM(\overline{G}) = \sum_{uv\in E(\overline{G})}\Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big)^2.$$

Hence

$$\overline{KCD_2}(G) = 4\overline{m}(n-1)(n-2)-(4n-6)M_1(\overline{G})+HM(\overline{G}) \tag{35}$$

Substituting Eqs. (12) and (16) in Eq. (35), simple calculation yields Eq. (31).

Proof of Eq. (32):

With the help of Eqs. (11), (25) and (26), we get

$$\begin{aligned}
\overline{KCD_2}(\overline{G}) &= \sum_{e=uv\notin E(\overline{G})} \Big(d_{\overline{G}}(u)+d_{\overline{G}}(v)\Big)d_{\overline{G}}(e) \\
&= \sum_{e=uv\in E(G)} \Big(n-1-d_G(u)+n-1-d_G(v)\Big)\Big(2n-4-(d_G(u)+d_G(v)\Big) \\
&= \sum_{e=uv\in E(G)} 4(n-1)(n-2)-(4n-6)\sum_{e=uv\in E(G)}\Big(d_G(u)+d_G(v)\Big) \\
&\quad + \sum_{e=uv\in E(G)}\Big(d_G(u)+d_G(v)\Big)^2
\end{aligned}$$

Eq. (32) immediately follows.

□

4 KCD indices of some graph operations

In this section, we study the graph operations using KCD indices. The well-known graph operations sum(join), cartesian product and composition of graphs are considered. All operations considered under the context are binary, with finite and simple graphs G and H. For the graphs G and H vertex and edge sets are denoted by $V(G)$ and $V(H)$, $E(G)$ and $E(H)$ respectively. The detailed information on sum(join) of graphs is refered in[Khalifeh et al., 2008a], cartesian product of graphs studied in[Khalifeh et al., 2008b] and composition of graphs is reported in [Imrich and Klavzar, 2000, Khalifeh et al., 2008a]. We refer [Khalifeh et al., 2009] for detailed information about graph operations.

Sum(join):

The sum(join) $G+H$ of two graphs G and H with disjoint vertex sets $|V(G)|$ and $|V(H)|$ is the graph on the vertex set $V(G)\cup V(H)$ and the edge set $E(G)\cup E(H)\cup\{uv : u\in V(G) \;\; and \;\; v\in V(H)\}$. For the graph $G+H$, $|V(G+H)| = |V(G)|+V(H)|$, $|E(G+H)| = |E(G)|+|E(H)|+|V(G)||V(H)|$, the degree of any vertex $u \in G+H$ is

$$d_{G+H}(u) = \begin{cases} d_G(u) + |V(H)| & u \in V(G) \\ d_H(u) + |V(G)| & u \in V(H). \end{cases}$$

Theorem 4.1. *Let G and H be graphs. Then*

$$\begin{aligned} KCD_1(G+H) &= 2\Bigg(M_1(G) + M_1(H) + |E(H)|\Big(4|V(G)| - 1\Big) \\ &\quad + |E(G)|\Big(4|V(H)| - 1\Big) \\ &\quad + |V(G)||V(H)|\Big(|V(G)| + |V(H)| - 1\Big)\Bigg). \end{aligned}$$

Proof:
By definition of sum(join) $G+H$ of two graphs G, H and Eq. (8), we have

$$KCD_1(G+H) = \sum_{e=uv \in E(G+H)} \Bigg(\Big(d_{G+H}(u) + d_{G+H}(v)\Big) + d_{G+H}(e)\Bigg).$$

Since,

$$d_{G+H}(e) = d_{G+H}(u) + d_{G+H}(v) - 2.$$

$$KCD_1(G+H) = 2 \sum_{e=uv \in E(G+H)} \Big(d_{G+H}(u) + d_{G+H}(v) - 1\Big).$$

$$\begin{aligned} KCD_1(G+H) &= 2 \sum_{e=uv \in E(H)} \Big(d_{G+H}(u) + d_{G+H}(v) - 1\Big) \\ &\quad + 2 \sum_{e=uv \in E(G)} \Big(d_{G+H}(u) + d_{G+H}(v) - 1\Big) \\ &\quad + 2 \sum_{e=uv \in \{uv : u \in V(G), v \in V(H)\}} \Big(d_{G+H}(u) + d_{G+H}(v) - 1\Big). \end{aligned} \tag{36}$$

Observe that,

$$\begin{aligned} \sum_{e=uv \in E(H)} \Big(d_{G+H}(u) + d_{G+H}(v) - 1\Big) &= \sum_{e=uv \in E(H)} \Big(d_H(u) + |V(G)| \\ &\quad + d_H(v) + |V(G)| - 1\Big) \\ &= \sum_{e=uv \in E(H)} \Big(d_H(u) + d_H(v) + 2|V(G)| - 1\Big). \end{aligned}$$

Thus,

$$\sum_{e=uv\in E(H)}\Big(d_{G+H}(u)+d_{G+H}(v)-1\Big) = M_1(H)+2|V(G)||E(H)| - |E(H)|. \quad (37)$$

Similarly,

$$\sum_{e=uv\in E(G)}\Big(d_{G+H}(u)+d_{G+H}(v)-1\Big) = M_1(G)+2|V(H)||E(G)| - |E(G)|. \quad (38)$$

In the same way,

$$\sum_{u\in V(G),v\in V(H)}\Big(d_{G+H}(u)+d_{G+H}(v)-1\Big) = 2|V(H)||E(G)|+|V(H)|^2|V(G)| +2|E(H)||V(G)|+|V(G)|^2|V(H)|-|V(G)||V(H)|. \quad (39)$$

Substituting Eqs. (37), (38) and (39) in Eq. (36) completes the proof.

□

Theorem 4.2. *Let G and H be graphs. Then*

$$KCD_2(G+H) = HM(G)+HM(H)+\Big(5|V(H)|-2\Big)M_1(G)+\Big(5|V(G)|-2\Big)M_1(H) +8\Bigg(|V(G)||E(H)|\Big(|V(G)|-1\Big)+|V(H)||E(G)|\Big(|V(H)|-1\Big)+|E(G)||E(H)|\Bigg) +|V(G)||V(H)|\Bigg(\Big(|V(G)|+|V(H)|\Big)^2+4\Big(|E(G)|+|E(H)|\Big)-2\Big(|V(G)|+|V(H)|\Big)\Bigg).$$

Proof.
With the knowledge of sum(join) $G+H$ of two graphs G, H and Eq. (9), we have

$$KCD_2(G+H) = \sum_{e=uv\in E(G+H)}\Big(d_{G+H}(u)+d_{G+H}(v)\Big)d_{G+H}(e).$$

As,

$$d_{G+H}(e) = d_{G+H}(u)+d_{G+H}(v)-2.$$

This implies,

$$\begin{aligned}
KCD_2(G+H) &= \sum_{e=uv\in E(G+H)} \Big(d_{G+H}(u)+d_{G+H}(v)\Big)^2 - 2\Big(d_{G+H}(u)+d_{G+H}(v)\Big) \\
&= \sum_{e=uv\in E(H)} \Big(d_{G+H}(u)+d_{G+H}(v)\Big)^2 - 2\Big(d_{G+H}(u)+d_{G+H}(v)\Big) \\
&\quad + \sum_{e=uv\in E(G)} \Big(d_{G+H}(u)+d_{G+H}(v)\Big)^2 - 2\Big(d_{G+H}(u)+d_{G+H}(v)\Big) \\
&\quad + \sum_{e=uv\in \{uv: u\in V(G), v\in V(H)\}} \Big(d_{G+H}(u)+d_{G+H}(v)\Big)^2 \\
&\quad -2\Big(d_{G+H}(u)+d_{G+H}(v)\Big).
\end{aligned}$$

It follows that,

$$\begin{aligned}
\sum_{e=uv\in E(H)} \Big(d_{G+H}(u)+d_{G+H}(v)\Big)^2 - 2\Big(d_{G+H}(u)+d_{G+H}(v)\Big) &= \sum_{e=uv\in E(H)} \Bigg(\Big(d_H(u) \\
+|V(G)| + d_H(v) + |V(G)|\Big)^2 - 2\Big(d_H(u)+|V(G)|+d_H(v)+|V(G)|\Big)\Bigg)
\end{aligned}$$

$$\begin{aligned}
= \sum_{e=uv\in E(H)} \Bigg(\Big(d_H(u)+d_H(v)\Big)^2 + 4|V(G)|^2 + 4|V(G)|\Big(d_H(u)+d_H(v)\Big) - 2\Big(d_H(u) \\
+d_H(v)\Big) - 4|V(G)|\Bigg).
\end{aligned}$$

$$\begin{aligned}
\sum_{e=uv\in E(H)} \Big(d_{G+H}(u)+d_{G+H}(v)\Big)^2 - 2\Big(d_{G+H}(u)+d_{G+H}(v)\Big) &= HM(H) \\
+4|V(G)|^2|E(H)| + 4|V(G)|M_1(H) - 2M_1(H) - 4|V(G)||E(H)|.
\end{aligned} \tag{40}$$

Similarly,

$$\begin{aligned}
\sum_{e=uv\in E(G)} \Big(d_{G+H}(u)+d_{G+H}(v)\Big)^2 - 2\Big(d_{G+H}(u)+d_{G+H}(v)\Big) &= HM(G) \\
+4|V(H)|^2|E(G)| + 4|V(H)|M_1(G) - 2M_1(G) - 4|V(H)||E(G)|.
\end{aligned} \tag{41}$$

In the same way

$$
\begin{aligned}
\sum_{u\in V(G),v\in V(H)} \Big(d_{G+H}(u)+d_{G+H}(v)\Big)^2 - 2\Big(d_{G+H}(u)+d_{G+H}(v)\Big) &= M_1(G)|V(H)| \\
+M_1(H)|V(G)| + 8|E(G)||E(H)| + |V(G)||V(H)|\Big(|V(G)|+|V(H)|\Big)^2 \\
+4|E(G)||V(H)|^2 + 4|E(G)||V(G)||V(H)| + 4|E(H)||V(G)||V(H)| \\
+4|E(H)||V(G)|^2 - 4|E(G)||V(H)| - 4|E(H)||V(G)| \\
-2|V(G)||V(H)|\Big(|V(G)|+|V(H)|\Big).
\end{aligned} \tag{42}
$$

Finally, the summaton of Eqs. (40), (41) and (42) gives the desired result.

□

Cartesian Product:

The cartesian product $G \times H$ of two graphs G and H has the vertex set $V(G \times H) = V(G) \times V(H)$ and $e = (a,x)(b,y)$ is an edge of $G \times H$ if $a = b$ and $xy \in E(H)$, or $ab \in E(H)$ and $x = y$. For the graph $G \times H$, $|V(G \times H)| = |V(G)|V(H)|$, $|E(G\times H)| = |E(G)||V(H)| + |V(G)||E(H)|$, The degree of any vertex $(a,x) \in G \times H$ is $d_{G\times H}((a,x)) = d_G(a) + d_H(x)$.

Theorem 4.3. *Let G and H be graphs. Then*

$$
\begin{aligned}
KCD_1(G\times H) &= 2\Bigg(|V(G)|M_1(H) + |V(H)|M_1(G) + 8|E(G)||E(H)| - \\
&\quad \Big(|V(G)||E(H)| + |V(H)||E(G)|\Big)\Bigg).
\end{aligned}
$$

Proof.
In the view of definition of cartesian product $G \times H$ of two graphs G, H and Eq. (8), we have

$$
KCD_1(G\times H) = \sum_{e=(a,x)(b,y)\in E(G\times H)} \Bigg(\Big(d_{G\times H}((a,x)) + d_{G\times H}((b,y))\Big) + d_{G\times H}((e))\Bigg).
$$

It is known that,

$$
d_{G\times H}((e)) = d_{G\times H}((a,x)) + d_{G\times H}((b,y)) - 2.
$$

Thus,

$$
\begin{aligned}
KCD_1(G\times H) &= 2\sum_{e=(a,x)(b,y)\in E(G\times H)}\Big(d_{G\times H}((a,x))+d_{G\times H}((b,y))-1\Big)\\
&= 2\sum_{a\in V(G)}\sum_{xy\in E(H)}\Big(d_G(a)+d_H(x)+d_G(a)+d_H(y)-1\Big)\\
&\quad +2\sum_{x\in V(H)}\sum_{ab\in E(G)}\Big(d_H(x)+d_G(a)+d_H(x)+d_G(b)-1\Big)\\
&= 2\sum_{a\in V(G)}\sum_{xy\in E(H)}\left(2d_G(a)+\Big(d_H(x)+d_H(y)\Big)-1\right)\\
&\quad +2\sum_{x\in V(H)}\sum_{ab\in E(G)}\left(2d_H(x)+\Big(d_G(a)+d_G(b)\Big)-1\right)
\end{aligned}
$$

By simple reasoning we straightforwardly obtain the required result.

□

Theorem 4.4. *Let G and H be graphs. Then*

$$
\begin{aligned}
KCD_2(G\times H) &= |V(G)|HM(H)+|V(H)|HM(G)+\Big(12|E(H)|-2|V(H)|\Big)M_1(G)\\
&\quad +\Big(12|E(G)|-2|V(G)|\Big)M_1(H)-16|E(G)||E(H)|.
\end{aligned}
$$

Proof.
Taking into account the definition of cartesian product $G\times H$ of two graphs G and H, start with Eq. (9) as

$$
KCD_2(G\times H) = \sum_{e=(a,x)(b,y)\in E(G\times H)}\Big(d_{G\times H}((a,x))+d_{G\times H}((b,y))\Big)d_{G\times H}((e)).
$$

Since

$$
d_{G\times H}((e)) = d_{G\times H}((a,x))+d_{G\times H}((b,y))-2.
$$

We have

$$
\begin{aligned}
KCD_2(G\times H) &= \sum_{e=(a,x)(b,y)\in E(G\times H)} \Bigg(\Big(d_{G\times H}((a,x))+d_{G\times H}((b,y))\Big)^2 \\
&\quad -2\Big(d_{G\times H}((a,x))+d_{G\times H}((b,y))\Big)\Bigg) \\
&= \sum_{a\in V(G)}\sum_{xy\in E(H)} \Bigg(\Big(d_G(a)+d_H(x)+d_G(a)+d_H(y)\Big)^2 \\
&\quad -2\Big(d_G(a)+d_H(x)+d_G(a)+d_H(y)\Big)\Bigg) + \sum_{x\in V(H)}\sum_{ab\in E(G)} \Bigg(\Big(d_H(x) \\
&\quad +d_G(a)+d_H(x)+d_G(b)\Big)^2 - 2\Big(d_H(x)+d_G(a)+d_H(x)+d_G(b)\Big)\Bigg) \\
&= \sum_{a\in V(G)}\sum_{xy\in E(H)} \Bigg(\Big(2d_G(a)+d_H(x)+d_H(y)\Big)^2 - 2\Big(2d_G(a)+d_H(x) \\
&\quad +d_H(y)\Big)\Bigg) + \sum_{x\in V(H)}\sum_{ab\in E(G)} \Bigg(\Big(2d_H(x)+d_G(a)+d_G(b)\Big)^2 \\
&\quad -2\Big(2d_H(x)+d_G(a)+d_G(b)\Big)\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
KCD_2(G\times H) &= \sum_{a\in V(G)}\sum_{xy\in E(H)} \Bigg(4\Big(d_G(a)\Big)^2 + \Big(d_H(x)+d_H(y)\Big)^2 \\
&\quad +4d_G(a)\Big(d_H(x)+d_H(y)\Big) - 2\Big(2d_G(a)+\Big(d_H(x)+d_H(y)\Big)\Big)\Bigg) \\
&\quad + \sum_{x\in V(H)}\sum_{ab\in E(G)} \Bigg(4\Big(d_H(x)\Big)^2 + \Big(d_G(a)+d_G(b)\Big)^2 \\
&\quad +4d_H(x)\Big(d_G(a)+d_G(b)\Big) - 2\Big(2d_H(x)+\Big(d_G(a)+d_G(b)\Big)\Big)\Bigg)
\end{aligned}
$$

and the required result immediately follows.

□

Composition:

The composition $G[H]$ of two graphs G and H with disjoint vertex sets $V(G)$ and $V(H)$, edge sets E(G) and E(H) is the graph with vertex set $V(G)\times V(H)$ and (a,x) is adjacent to (b,y) whenever a is adjacent to b, or $a=b$ and x is adjacent to y. For the graph $G[H]$, $|V(G[H])|=|V(G)||V(H)|$, $|E(G[H])|=|E(G)||V(H)|^2+|E(H)||V(G)|$, The degree of any vertex $(a,x)\in G[H]$ is $d_{G[H]}((a,x))=|V(H)|d_G(a)+d_H(x)$.

Theorem 4.5. *Let G and H be graphs. Then*

$$\begin{aligned} KCD_1(G[H]) &= 2\Big(|V(H)|^3M_1(G)+|V(G)|M_1(H)+8|V(H)||E(G)||E(H)| \\ &\quad -|V(H)|^2|E(G)|-|E(H)||V(G)|\Big). \end{aligned}$$

Proof.
Using the definition of composition $G[H]$ of two graphs G, H and Eq. (8), we have

$$KCD_1(G[H]) = \sum_{e=(a,x)(b,y)\in E(G[H])}\left(\Big(d_{G[H]}((a,x))+d_{G[H]}((b,y))\Big)+d_{G[H]}((e))\right).$$

But

$$d_{G[H]}((e)) = d_{G[H]}((a,x))+d_{G[H]}((b,y))-2.$$

This implies,

$$KCD_1(G[H]) = 2\sum_{e=(a,x)(b,y)\in E(G[H])}\Big(d_{G[H]}((a,x))+d_{G[H]}((b,y))-1\Big).$$

$$\begin{aligned} KCD_1(G[H]) = {}& 2\sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)}\Big(|V(H)|d_G(a)+d_H(x)+|V(H)|d_G(b) \\ & +d_H(y)-1\Big)+2\sum_{a\in V(G)}\sum_{xy\in E(H)}\Big(|V(H)|d_G(a)+d_H(x) \\ & +|V(H)|d_G(a)+d_H(y)-1\Big). \end{aligned} \tag{43}$$

We start with

$$\sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)}\Big(|V(H)|d_G(a)+d_H(x)+|V(H)|d_G(b)+d_H(y)-1\Big)=$$
$$\sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)}\Bigg(|V(H)|\Big(d_G(a)+d_G(b)\Big)+\Big(d_H(x)+d_H(y)\Big)-1\Bigg)$$

Thus,

$$\begin{aligned}\sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)}\Big(|V(H)|d_G(a)+d_H(x)+|V(H)|d_G(b)+d_H(y)-1\Big)=\\ |V(H)|^3M_1(G)+4|V(H)||E(G)||E(H)|-|V(H)|^2|E(G)|.\end{aligned}\tag{44}$$

Similarly,

$$\begin{aligned}\sum_{a\in V(G)}\sum_{xy\in E(H)}\Big(|V(H)|d_G(a)+d_H(x)+|V(H)|d_G(a)+d_H(y)-1\Big)=\\ 4|V(H)||E(G)||E(H)|+|V(G)|M_1(H)-|V(G)||E(H)|.\end{aligned}\tag{45}$$

Substituting Eqs. (44) and (45) in Eq. (43) generates the desired result.

□

Theorem 4.6. *Let G and H be graphs. Then*

$$\begin{aligned}KCD_2(G[H]) &= |V(H)|^4HM(G)+|V(G)|HM(H)\\ &\quad+2|V(H)|^2M_1(G)\Big(6|E(H)|-|V(H)|\Big)\\ &\quad+2M_1(H)\Big(5|V(H)||E(G)|-|V(G)|\Big)\\ &\quad+8|E(G)||E(H)|\Big(|E(H)|-2|V(H)|\Big).\end{aligned}$$

Proof.
In view of definition of composition $G[H]$ of two graphs G, H and Eq. (9), we start with

$$KCD_2(G[H]) = \sum_{e=(a,x)(b,y)\in E(G[H])}\Big(d_{G[H]}((a,x))+d_{G[H]}((b,y))\Big)d_{G[H]}((e)).$$

It is known that,

$$d_{G[H]}((e)) = d_{G[H]}((a,x))+d_{G[H]}((b,y))-2.$$

We get,

$$\begin{aligned} KCD_2(G[H]) \;&=\; \sum_{e=(a,x)(b,y)\in E(G[H])} \bigg(\Big(d_{G[H]}((a,x))+d_{G[H]}((b,y))\Big)^2 \\ &\qquad -2\Big(d_{G[H]}((a,x))+d_{G[H]}((b,y))\Big)\bigg). \end{aligned}$$

$$\begin{aligned} KCD_2(G[H]) = \; & \sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)} \bigg(\Big(|V(H)|d_G(a)+d_H(x)+|V(H)|d_G(b)+d_H(y)\Big)^2 \\ & \qquad -2\Big(|V(H)|d_G(a)+d_H(x)+|V(H)|d_G(b)+d_H(y)\Big)\bigg) \\ & + \sum_{a\in V(G)}\sum_{xy\in E(H)} \bigg(\Big(|V(H)|d_G(a)+d_H(x)+|V(H)|d_G(a)+d_H(y)\Big)^2 \\ & \qquad -2\Big(|V(H)|d_G(a)+d_H(x)+|V(H)|d_G(a)+d_H(y)\Big)\bigg). \end{aligned}$$

Thus,

$$\begin{aligned} KCD_2(G[H]) = \; & \sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)} \bigg(\Big(|V(H)|\Big(d_G(a)+d_G(b)\Big)+d_H(x)+d_H(y)\Big)^2 \\ & \qquad -2\Big(|V(H)|\Big(d_G(a)+d_G(b)\Big)+d_H(x)+d_H(y)\Big)\bigg) \\ & + \sum_{a\in V(G)}\sum_{xy\in E(H)} \bigg(\Big(2|V(H)|d_G(a)+d_H(x)+d_H(y)\Big)^2 \\ & \qquad -2\Big(2|V(H)|d_G(a)+d_H(x)+d_H(y)\Big)\bigg). \end{aligned}$$

It follows that,

$$\sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)}\Bigg(\Big(|V(H)|\Big(d_G(a)+d_G(b)\Big)+d_H(x)+d_H(y)\Big)^2$$
$$-2\Big(|V(H)|\Big(d_G(a)+d_G(b)\Big)+d_H(x)+d_H(y)\Big)\Bigg)$$
$$=\sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)}\Bigg(|V(H)|^2\Big(d_G(a)+d_G(b)\Big)^2+\Big(d_H(x)+d_H(y)\Big)^2$$
$$+2|V(H)|\Big(d_G(a)+d_G(b)\Big)\Big(d_H(x)+d_H(y)\Big)-2|V(H)|\Big(d_G(a)+d_G(b)\Big)$$
$$-2|V(H)|d_H(x)-2|V(H)|d_H(y)\Bigg)$$

Hence,

$$\sum_{x\in V(H)}\sum_{y\in V(H)}\sum_{ab\in E(G)}\Bigg(\Big(|V(H)|\Big(d_G(a)+d_G(b)\Big)+d_H(x)+d_H(y)\Big)^2$$
$$-2\Big(|V(H)|\Big(d_G(a)+d_G(b)\Big)+d_H(x)+d_H(y)\Big)\Bigg)=|V(H)|^4HM(G) \tag{46}$$
$$+2|V(H)||E(G)|M_1(H)+8|E(H)|^2|E(G)|+8|V(H)|^2|E(H)|M_1(G)$$
$$-2|V(H)|^3M_1(G)-8|E(H)||E(G)||V(H)|.$$

Similarly,

$$\sum_{a\in V(G)}\sum_{xy\in E(H)}\Bigg(\Big(2|V(H)|d_G(a)+d_H(x)+d_H(y)\Big)^2-2\Big(2|V(H)|d_G(a)$$
$$+d_H(x)+d_H(y)\Big)\Bigg)=4|V(H)|^2|E(H)|M_1(G)+|V(G)|HM(H) \tag{47}$$
$$+8|V(H)||E(G)|M_1(H)-8|V(H)||E(G)||E(H)|-2|V(G)|M_1(H).$$

Finally, summation of Eqs. (46) and (47) gives the required result.

□

5 Conclusion

In this paper, we have introduced few new degree based topological indices and coindices named KCD indices and coindices. A set of properties of these indices and coindices are obtained. Finally, some graph operations are studied using KCD indices. These results have scope for further development using remaining graph operations.

References

Anwar Alwardi, Akram Alqesmah, R Rangarajan, and Ismail Naci Cangul. Entire zagreb indices of graphs. *Discrete Mathematics, Algorithms and Applications*, 10(03):1850037, 2018.

Ali Reza Ashrafi, Tomislav Došlić, and Asma Hamzeh. The zagreb coindices of graph operations. *Discrete Applied Mathematics*, 158(15):1571–1578, 2010.

Ali Reza Ashrafi, Tomislav Došlić, and Asma Hamzeh. Extremal graphs with respect to the zagreb coindices. *MATCH-Communications in Mathematical and in Computer Chemistry*, 65(1):85–92, 2011.

Kinkar Ch Das, Ivan Gutman, and Batmend Horoldagva. Comparing zagreb indices and coindices of trees. *MATCH-Communications in Mathematical and in Computer Chemistry*, 68:189–198, 2012.

Tomislav Došlić. Vertex-weighted wiener polynomials for composite graphs. *Ars Mathematica Contemporanea*, 1(1):66–80, 2008.

Tomislav Došlic, Boris Furtula, Ante Graovac, Ivan Gutman, Sirous Moradi, and Zahra Yarahmadi. On vertex–degree–based molecular structure descriptors. *MATCH-Communications in Mathematical and in Computer Chemistry*, 66(2): 613–626, 2011.

Ivan Gutman. On hyper–zagreb index and coindex. *Bulletin (Académie serbe des sciences et des arts. Classe des sciences mathématiques et naturelles. Sciences mathématiques)*, (42):1–8, 2017.

Ivan Gutman and Kinkar Ch Das. The first zagreb index 30 years after. *MATCH-Communications in Mathematical and in Computer Chemistry*, 50(1):83–92, 2004.

Ivan Gutman and Nenad Trinajstić. Graph theory and molecular orbitals. total φ-electron energy of alternant hydrocarbons. *Chemical Physics Letters*, 17(4): 535–538, 1972.

Ivan Gutman, Boris Furtula, Zana Kovijanić Vukićević, and Goran Popivoda. On zagreb indices and coindices. *MATCH-Communications in Mathematical and in Computer Chemistry*, 74(1):5–16, 2015.

Frank Harary. *Graph theory*. Addison–Wesely, Reading, Mass, 1969.

Wilfried Imrich and Sandi Klavzar. *Product graphs: structure and recognition*. Wiley, 2000.

M H Khalifeh, Hassan Yousefi-Azari, and Ali Reza Ashrafi. A matrix method for computing szeged and vertex pi indices of join and composition of graphs. *Linear algebra and its applications*, 429(11-12):2702–2709, 2008a.

M H Khalifeh, Hassan Yousefi-Azari, and Ali Reza Ashrafi. Vertex and edge pi indices of cartesian product graphs. *Discrete Applied Mathematics*, 156(10): 1780–1789, 2008b.

M H Khalifeh, Hassan Yousefi-Azari, and Ali Reza Ashrafi. The first and second zagreb indices of some graph operations. *Discrete Applied Mathematics*, 157 (4):804–811, 2009.

K Pattabiraman and M Vijayaragavan. Hyper zagreb indices and its coindices of graphs. *Bulletin of International Mathematical Virtual Institute*, 7(1):31–41, 2017.

G H Shirdel, H Rezapour, and A M Sayadi. The hyper-zagreb index of graph operations. *Iranian Journal of Mathematical Chemistry*, 4(2):213–220, 2013.

Maryam Veylaki, Mohammad Javad Nikmehr, and Hamid Agha Tavallaee. The third and hyper-zagreb coindices of some graph operations. *Journal of Applied Mathematics and Computing*, 50(1-2):315–325, 2016.

Bo Zhou. Zagreb indices. *MATCH-Communications in Mathematical and in Computer Chemistry*, (52):113–118, 2004.

Bo Zhou and Ivan Gutman. Further properties of zagreb indices. *MATCH-Communications in Mathematical and in Computer Chemistry*, 54(1):233–239, 2005.

Ratio Mathematica Volume 39, 2020, pp. 187-212

CAS wavelet approximation of functions of Hölder's class $H^{\alpha}[0,1)$ and Solution of Fredholm Integral Equations

Shyam Lal*
Satish Kumar†

Abstract

In this paper, cosine and sine wavelet is considered. Two new CAS wavelet estimators $E^{(1)}_{2^k,2M+1}(f)$ and $E^{(2)}_{2^k,2M+1}(f)$ for the approximation of a function f whose first derivative f' and second derivative f'' belong to Hölder's class $H^{\alpha}[0,1)$ of order $0 < \alpha \leqslant 1$, have been obtained. These estimators are sharper and best in wavelet analysis. Using CAS wavelet, a computational method has been developed to solve Fredholm integral equation of second kind. In this process, Fredholm integral equations are reduced into a system of linear equations. Approximation of functions by CAS wavelet method is applied in obtaining the solution of Fredholm integral equation of second kind. CAS wavelet coefficient matrices are prepared using the properties of CAS wavelets. Two examples are illustrated to show the validity and efficiency of the technique discussed in this paper.

Keywords: CAS wavelet, CAS Wavelet Approximation, Function of Hölder's class, Orthonormal basis, Fredholm integral equation.
Mathematics Subject Classification:42C40, 65T60, 45G10,45B05. [1]

*Department of Mathematics, Institute of Science, Banaras Hindu University, Varanasi-221005, India; shyam_lal@rediffmail.com.
†Department of Mathematics, Institute of Science, Banaras Hindu University, Varanasi-221005, India); satishkumar3102@gmail.com.
[1]Received on June 21st, 2020. Accepted on December 15th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.549. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

Wavelet is a very recent and powerful tool in pure as well as applied mathematical research area. It has wide range applications in engineering, science and technology, signal analysis, time-frequency analysis, fast numerical algorithm. Several problems of Physics, Engineering, science and Technology are found in the form of integral equations. In some cases, integral equations are reformulated into ordinary differential equations and partial differential equations. In many cases, it is very difficult to solve integral equations analytically and hence there is a need of approximate solution of integral equations. In recents years, the approximate solutions of integral equations have been obtained by orthogonal basis functions as well as orthogonal wavelets. The main advantage of using orthonormal basis is that it converts the mathematical problems to a system of algebraic equations. Working in same direction, several researchers like [2], Sahu [3] etc. have been solved integral equations. It is known that wavelets are considerably useful in the solution of integral equations. In science and Technology, some problems are available in the form of Fredholm integral equations of second kind:

$$u(x) = f(x) + \int_0^1 K(x,y)u(y)dy \tag{1}$$

where $f \in L^2[0,1)$ and $K \in L^2[0,1) \times L^2[0,1)$ are known functions and u is unknown function to be determined (Ray and Sahu [3]).

In best of our knowledge, there is no work associated with the solution of Fredholm integral eq$^{\text{n}}$ (1) by CAS wavelet method. The main objectives of the research paper are as follows:

1. To estimate the approximation of functions belonging to Hölder's class $H^\alpha[0,1)$ of order $0 < \alpha \leqslant 1$ by CAS wavelet method.

2. To develop a procedure to solve Fredholm integral equation of second kind by using CAS wavelet approximation.

3. To compare the solutions of Fredholm integral eq$^{\text{n}}$ (1) obtained by CAS wavelet, Legendre wavelet and Haar wavelet method with their exact solutions.

It is remarkable to note that the solution of Fredholm integral eq$^{\text{n}}$ (1) obtained by CAS wavelet method and its exact solution are almost same. The solution of Fredholm integral eq$^{\text{n}}$ (1) obtained by CAS wavelet method is better and more closed to its exact solution than the solutions obtained by Legendre wavelet and Haar wavelet method. It is observed in numerical comparison of these solutions. It is a significant achievement of the proposed method.

2 Definitions and Preliminaries

2.1 Basic Wavelets And CAS Wavelets

Let $\psi \in L^2(\mathbb{R})$. ψ is called a basic wavelet if it satisfies the admissibility condition:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}|^2}{|w|} dw < \infty \text{ (Chui [1])} \tag{2}$$

The integral wavelet transform, relative to a basic wavelet ψ, is defined by

$$(W_\psi f)(b,a) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t)\overline{\psi(\frac{b-a}{a})} dt \ , f \in L^2(\mathbb{R}) \tag{3}$$

where $a, b \in \mathbb{R}, a \neq 0$. Set

$$\psi_{b,a}(t) = |a|^{-1/2}\psi(\frac{b-a}{a}). \tag{4}$$

This is a family of wavelets. If we restrict the parameters a and b to discrete values

$$a = a_0^{-k}, b = nb_0a_0^{-k}, a_0 > 1, b_0 > 0$$

where n and k are positive integers, then

$$\psi_{b,a}(t) = \psi_{n,k}(t) = |a_0|^{k/2}\psi(a_0^k t - nb_0). \tag{5}$$

Taking $a_0 = 2, b_0 = 1$ in eq^n (5),

$$\psi_{n,k}(t) = 2^{k/2}\psi(2^k t - n). \tag{6}$$

If

$$\psi(2^k t - n) = \cos(2m\pi(2^k t - n + 1)) + \sin(2m\pi(2^k t - n + 1)) \tag{7}$$
$$= CAS_m(2^k t - n + 1). \tag{8}$$

Using eq^n(7), eq^n (6) becomes

$$\psi_{n,m}(t) = \begin{cases} 2^{\frac{k}{2}}\{\cos(2m\pi(2^k t - n + 1)) + \sin(2m\pi(2^k t - n + 1))\}, & if \ \frac{n-1}{2^k} \leqslant t < \frac{n}{2^k}, \\ 0, & otherwise. \end{cases}$$

$\{\psi_{n,m}\}_{n,m\in\mathbb{Z}}$ are orthonormal CAS wavelets defined on [0,1) .

3 Function belonging to Hölder's class $H^{\alpha}[0,1)$

A function f is said to belong to Hölder's class $H^{\alpha}[0,1)$ of order $0<\alpha\leqslant 1$ if f satifies the following condition :

$$|f(x)-f(y)|\leqslant A|x-y|^{\alpha},\ \ \forall x,y\in \mathbb{R} \tag{9}$$

for some positive constant A (Zheng, Wei [4]).

3.1 Proposition

Let f be a function such that its second derivative $f^{''}$ is in $H^{\alpha}[0,1)$, then its first derivative $f^{'}$ is in $H^{\alpha}[0,1)$.
Proof : Let $\phi^{''}\in H^{\alpha}[0,1)$.

$$\begin{aligned}
f(x) &= \int_0^{x^{\alpha}} \phi^{'}(t)\,dt\\
f^{'}(x) &= \int_0^{x^{\alpha}} \phi^{''}(t)\,dt \ \text{ and } \ f^{'}(y)=\int_0^{y^{\alpha}} \phi^{''}(t)\,dt\\
|f^{'}(x)-f^{'}(y)| &= |\int_0^{x^{\alpha}} \phi^{''}(t)\,dt-\int_0^{y^{\alpha}} \phi^{''}(t)\,dt|=|\int_{y^{\alpha}}^{x^{\alpha}} \phi^{''}(t)\,dt|\\
&\leq M|x^{\alpha}-y^{\alpha}|\leq M|x-y|^{\alpha}, \qquad M=\underset{t\in[0,1)}{sup}\{\phi^{''}(t)\}
\end{aligned}$$

Converse is not true. Consider the example $f(x)=\frac{x^{\alpha+1}}{\alpha+1}$ $0<\alpha<1$.Then, $f^{'}(x)=x^{\alpha}$ and $f^{''}(x)=\alpha x^{\alpha-1}$. For $x=\frac{1}{N^{\frac{1}{1-\alpha}}}$, $y=\frac{1}{(1+N)^{\frac{1}{1-\alpha}}}$, we have $|x-y|\leq \frac{1}{N^{\frac{1}{1-\alpha}}}-\frac{1}{(1+N)^{\frac{1}{1-\alpha}}}\leq \frac{1}{N^{\frac{1}{1-\alpha}}}=\delta$.
And $|f^{''}(x)-f''(y)|=\alpha(1+N-N)=\alpha$
If $0<\epsilon<\alpha$, then $|f^{''}(x)-f^{''}(y)|\nleq\epsilon$ whenever $|x-y|\leq\delta=\frac{1}{N^{\frac{1}{1-\alpha}}}$. Hence, $f^{'}\in H^{\alpha}[0,1)$ but $f^{''}\notin H^{\alpha}[0,1)$.

3.2 Difference between Hölder's class and Lipschitz class

1. Consider the function $f(x)=\sqrt{x^2+5}$ $\qquad \forall x\in[0,1]$. Then

$$|f(x)-f(y)|\leq|\sqrt{x^2+5}-\sqrt{y^2+5}|\leq|\sqrt{x^2-y^2}|\leq\sqrt{2}|x-y|^{\frac{1}{2}} \tag{10}$$

Eq^n(10) shows that $f\in H^{\frac{1}{2}}[0,1)$. And also, we have

$$|f^{'}(x)|\leq|\frac{x}{x^2+5}|\leq 1, \qquad \forall\, x\in[0,1] \tag{11}$$

Eq^n(10) and Eq^n(11) shows that $f \in Lip_{\frac{1}{2}}[0,1)$.

2. Define the function $f(x) = \sqrt{x} \qquad \forall x \in [0,1]$, then we have

$$|f(x) - f(y)| \leq |\sqrt{x} - \sqrt{y}| \leq |x-y|^{\frac{1}{2}} \implies f \in H^{\frac{1}{2}}[0,1).$$

And since, $f^{'}(x) = \frac{1}{2\sqrt{x}} \to \infty$ as $x \to 0^{+}$. Hence, f is not bounded.
$\therefore f \notin Lip_{\frac{1}{2}}[0,1)$. Hence, we conclude that $Lip_{\alpha}[0,1] \subset H^{\alpha}[0,1]$.

4 Approximation of function

Since $\{\psi_{n,m}\}_{n,m\in\mathbb{Z}}$ forms an orthonormal basis for $L^2[0,1]$, therefore a function $f \in L^2[0,1)$ can be expressed into CAS wavelet series as:

$$f(t) = \sum_{n=1}^{\infty} \sum_{m=-\infty}^{\infty} c_{n,m}\psi_{n,m}(t) \tag{12}$$

where the coefficients $c_{n,m}$ are given by

$$c_{n,m} =< f, \psi_{n,m} > \tag{13}$$

$(2^k, 2M+1)^{\text{th}}$ partial sum $S_{2^k,2M+1}(f)(t)$ of (12) is given by

$$S_{2^k,2M+1}(f)(t) = \sum_{n=1}^{2^k} \sum_{m=-M}^{M} c_{n,m}\psi_{n,m}(t) = C^T\Psi(t) \tag{14}$$

where C and $\Psi(t)$ are given by
$C = [c_{1,(-M)}, c_{1,(-M+1)}, ..., c_{1,M}, c_{2,(-M)}, ..., c_{2,M}, ..., c_{2^k,(-M)}, ..., c_{2^k,M}]^T$
and

$$\begin{aligned}\Psi(t) \quad = \quad & [\psi_{1,(-M)}(t), \psi_{1,(-M+1)}(t), ..., \psi_{1,M}(t), \psi_{2,(-M)}(t), ..., \psi_{2,M}(t), ..., \\ & \psi_{2^k,(-M)}(t), ..., \psi_{2^k,M}(t)]^T.\end{aligned}$$

Extended Legendre Wavelet expansion of function $f \in L^2[0,1)$ is

$$f(x) = \sum_{n=1}^{\infty}\sum_{m=0}^{\infty} c_{n,m}\psi_{n,m}^{(\mu)}(x),$$

and its $(\mu^k, M)^{th}$ partial sum is

$$S_{\mu^k,M}(f)(x) = \sum_{n=1}^{\mu^k}\sum_{m=0}^{M} c_{n,m}\psi_{n,m}^{(\mu)}(x).$$

The extended Legendre wavelet approximation $E_{\mu^k,M}(f)$ of f by $(\mu^k, M)^{th}$partial sum $S_{\mu^k,M}(f)$ is defined by

$$E_{\mu^k,M}(f) = \min_{S_{\mu^k,M}(f)} ||f - S_{\mu^k,M}(f)||_2 \ .$$

In our case, the CAS wavelet approximation $E_{2^k,2M+1}(f)$ of f by $(2^k, 2M+1)^{\text{th}}$ partial sum $S_{2^k,2M+1}(f)$ of series (12) is defined by

$$E_{2^k,2M+1}(f) = \min_{S_{2^k,2M+1}(f)} ||f - S_{2^k,2M+1}(f)||_2 \ . \tag{15}$$

5 Theorems

In this paper, we prove the following theorems:

Theorem 5.1. *If $f \in L^2[0,1)$ is a function such that $f' \in H^\alpha[0,1)$ and its CAS wavelet expansion is*

$$f(t) = \sum_{n=1}^{\infty} \sum_{m=-\infty}^{\infty} c_{n,m}\psi_{n,m}(t) \tag{16}$$

then the approximation error $E^{(1)}_{2^k,2M+1}(f)$ of f by $(2^k, 2M+1)^{th}$ partial sum

$$S_{2^k,2M+1}(f)(t) = \sum_{n=1}^{2^k} \sum_{m=-M}^{M} c_{n,m}\psi_{n,m}(t) \tag{17}$$

of expansion 16 is given by

$$E^{(1)}_{2^k,2M+1}(f) = \min_{S_{2^k,2M+1}(f)} ||f - (S_{2^k,2M+1}f)||_2 = O(\frac{1}{\sqrt{M+1}\ 2^{k(\alpha+1)}}) \tag{18}$$

Theorem 5.2. *If $f \in L^2[0,1)$ is a function such that $f'' \in H^\alpha[0,1)$ and its CAS wavelet expansion is given by the series (16) , then the approximation error $E^{(2)}_{2^k,2M+1}(f)$ of f by $(2^k, 2M+1)^{th}$ partial sum $S_{2^k,2M+1}(f)(t)$ of series (16) is given by*

$$E^{(2)}_{2^k,2M+1}(f) = \min_{S_{2^k,2M+1}(f)} ||f - (S_{2^k,2M+1}f)||_2 = O(\frac{1}{(M+1)^{\frac{3}{2}}\ 2^{k(\alpha+2)}}) \tag{19}$$

Proof of theorem (5.1) Since

$$f(t) = \sum_{n=1}^{\infty} \sum_{m=-\infty}^{\infty} c_{n,m}\psi_{n,m}(t)$$

and

$$S_{2^k,2M+1}(f)(t) = \sum_{n=1}^{2^k} \sum_{m=-M}^{M} c_{n,m}\psi_{n,m}(t)$$

$$\begin{aligned}
\therefore f(t) - S_{2^k,2M+1}(f)(t) &= \sum_{n=1}^{\infty} \sum_{m=-\infty}^{\infty} c_{n,m}\psi_{n,m}(t) - \sum_{n=1}^{2^k} \sum_{m=-M}^{M} c_{n,m}\psi_{n,m}(t) \\
&= (\sum_{n=1}^{2^k} + \sum_{n=2^k+1}^{\infty})(\sum_{m=-\infty}^{-M-1} + \sum_{m=-M}^{M} + \sum_{m=M+1}^{\infty})c_{n,m}\psi_{n,m}(t) \\
&\quad -\sum_{n=1}^{2^k} \sum_{m=-M}^{M} c_{n,m}\psi_{n,m}(t) \\
&= \sum_{n=1}^{2^k} \sum_{m=-\infty}^{-M-1} c_{n,m}\psi_{n,m}(t) + \sum_{n=1}^{2^k} \sum_{m=M+1}^{\infty} c_{n,m}\psi_{n,m}(t) \\
(f(t) - S_{2^k,2M+1}(f)(t))^2 &= \sum_{n=1}^{2^k} \sum_{m=-\infty}^{-M-1} c^2_{n,m}\psi^2_{n,m}(t) + \sum_{n=1}^{2^k} \sum_{m=M+1}^{\infty} c^2_{n,m}\psi^2_{n,m}(t) \\
&\quad +2 \sum\sum_{1\leqslant\ n\neq n'\leq\ 2^k} \sum\sum_{-\infty\leq m\neq m'\leq -M-1} c_{n,m}c_{n',m'}\psi^T_{n,m}(t)\psi_{n',m'}(t) \\
&\quad +2 \sum\sum_{1\leqslant\ n\neq n'\leq\ 2^k} \sum\sum_{M+1\leq m\neq m'\leq\infty} c_{n,m}c_{n',m'}\psi^T_{n,m}(t)\psi_{n',m'}(t) \\
||f - S_{2^k,2M+1}(f)||_2^2 &= \int_0^1 |f(t) - S_{2^k,2M+1}(f)(t)|^2 dt \\
&\leqslant \sum_{n=1}^{2^k} \sum_{m=-\infty}^{-M-1} |c_{n,m}|^2 \int_0^1 |\psi_{n,m}(t)|^2 dt \\
&+ \sum_{n=1}^{2^k} \sum_{m=M+1}^{\infty} |c_{n,m}|^2 \int_0^1 |\psi_{n,m}(t)|^2 dt \\
&+ 2 \sum\sum_{1\leqslant\ n\neq n'\leq\ 2^k} \sum\sum_{-\infty\leq m\neq m'\leq -M-1} |c_{n,m}||c_{n',m'}| \\
&\quad \int_0^1 |\psi^T_{n,m}(t)\psi_{n',m'}(t)|dt
\end{aligned}$$

$$+2\sum\sum_{1\leqslant\ n\neq n'\leq\ 2^k}\ \sum\sum_{M+1\leq m\neq m'\leq\infty}|c_{n,m}||c_{n',m'}|\int_0^1|\psi_{n,m}^T(t)\psi_{n',m'}(t)|dt$$

$$=\ \sum_{n=1}^{2^k}\sum_{m=-\infty}^{-M-1}|c_{n,m}|^2+\sum_{n=1}^{2^k}\sum_{m=M+1}^{\infty}|c_{n,m}|^2\ ,\ \text{by orthonormality of }\{\psi_{n,m}\}_{n,m\in\mathbb{Z}}$$

$$||f-S_{2^k,2M+1}(f)||_2^2\ \leq\ \sum_{n=1}^{2^k}(\sum_{m=-\infty}^{-M-1}+\sum_{m=M+1}^{\infty})|c_{n,m}|^2 \tag{20}$$

$$\begin{aligned}
c_{n,m} &= <f,\psi_{n,m}> \\
&= \int_{\frac{n-1}{2^k}}^{\frac{n}{2^k}} f(t)\ 2^{\frac{k}{2}}\{\cos(2m\pi(2^kt-n+1))+\sin(2m\pi(2^kt-n+1))\}\ dt \\
&= \frac{1}{2^{\frac{k}{2}}}\int_0^1 f(\frac{x+n-1}{2^k})\ (\cos(2m\pi x)+\sin(2m\pi x))\ dx,\ \ 2^kt-n+1=x \\
&= \frac{1}{(2m\pi)\ 2^{\frac{3k}{2}}}\int_0^1 f^{'}(\frac{x+n-1}{2^k})(\cos(2m\pi x)-\sin(2m\pi x))dx,\ \ \text{integrating by part} \\
&= \frac{1}{(2m\pi)\ 2^{\frac{3k}{2}}}[\int_0^1\{f^{'}(\frac{x+n-1}{2^k})-f^{'}(\frac{n-1}{2^k})\}(\cos(2m\pi x)-\sin(2m\pi x))dx \\
&\quad -f^{'}(\frac{n-1}{2^k})\ \int_0^1(\cos(2m\pi x)-\sin(2m\pi x))dx] \\
&= \frac{1}{(2m\pi)\ 2^{\frac{3k}{2}}}\int_0^1\{f^{'}(\frac{x+n-1}{2^k})-f^{'}(\frac{n-1}{2^k})\}(\cos(2m\pi x)-\sin(2m\pi x))dx \\
|c_{n,m}| &\leqslant \frac{1}{(2m\pi)\ 2^{\frac{3k}{2}}}\int_0^1|f^{'}(\frac{x+n-1}{2^k})-f^{'}(\frac{n-1}{2^k})|\ |\cos(2m\pi x)-\sin(2m\pi x)|dx \\
&\leqslant \frac{A}{(2m\pi)\ 2^{\frac{3k}{2}}}\int_0^1\ |\frac{x}{2^k}|^\alpha\ |\cos(2m\pi x)-\sin(2m\pi x)|\ dx,\ \ \text{since}\ \ f^{'}\in H^\alpha[0,1)
\end{aligned}$$

Now by Cauchy Schwarz inequality, we have

$$\begin{aligned}
|c_{n,m}| &\leqslant \frac{A}{(2m\pi)\ 2^{\frac{3k}{2}}}\{\int_0^1\ |\frac{x}{2^k}|^{2\alpha}\ dx\}^{\frac{1}{2}}\ \{\int_0^1|\cos(2m\pi x)-\sin(2m\pi x)|^2dx\}^{\frac{1}{2}} \\
&= \frac{A}{(2m\pi)\ 2^{(\frac{3k}{2}+k\alpha)}}\{\int_0^1|x|^{2\alpha}\ dx\}^{\frac{1}{2}} \\
&= \frac{A}{(2m\pi)\ 2^{(\frac{3}{2}+\alpha)k}}\ \frac{1}{\sqrt{2\alpha+1}} \\
|c_{n,m}| &\leqslant \frac{A}{2m\pi\ \sqrt{2\alpha+1}\ 2^{(\frac{3}{2}+\alpha)k}}
\end{aligned}$$

By eq^n (20) and (21) , we have

$$\begin{aligned}
||f - S_{2^k,2M+1}(f)||_2^2 &\leq \sum_{n=1}^{2^k}(\sum_{m=-\infty}^{-M-1} + \sum_{m=M+1}^{\infty})\frac{A^2}{4m^2\pi^2(2\alpha+1)\,2^{(3+2\alpha)k}}\,,\\
&= \frac{A^2}{4\pi^2(2\alpha+1)}(\sum_{m=-\infty}^{-M-1} + \sum_{m=M+1}^{\infty})\frac{2^k}{2^{(3+2\alpha)k}\,m^2}\\
&= \frac{A^2}{4\pi^2(2\alpha+1)}\,\frac{1}{2^{(2+2\alpha)k}}(\sum_{m=-\infty}^{-M-1}\frac{1}{m^2} + \sum_{m=M+1}^{\infty}\frac{1}{m^2})\\
&= \frac{A^2}{4\pi^2(2\alpha+1)}\,\frac{1}{2^{(1+\alpha)2k}}\,(\frac{1}{M+1} + \frac{1}{M+1})\\
&= \frac{A^2}{2\pi^2(2\alpha+1)}\,\frac{1}{2^{(1+\alpha)2k}}\frac{1}{M+1}\\
\therefore \min_{S_{2^k,2M+1}(f)}||f - S_{2^k,2M+1}(f)||_2 &\leqslant \frac{A}{\pi\sqrt{2(2\alpha+1)}}\,\frac{1}{2^{k(\alpha+1)}}\frac{1}{\sqrt{M+1}}
\end{aligned}$$

$$\therefore E^{(1)}_{2^k,2M+1}(f) = \min_{S_{2^k,2M+1}(f)}||f - S_{2^k,2M+1}(f)||_2 = O(\frac{1}{\sqrt{M+1}\,2^{k(\alpha+1)}})$$

Thus, theorem (5.1) is completely established.

Proof of theorem (5.2) Following the steps of the proof of theorem (5.1)

$$\begin{aligned}
c_{n,m} &= \frac{1}{(2m\pi)\,2^{\frac{3k}{2}}}\int_0^1 f'(\frac{x+n-1}{2^k})(\cos(2m\pi x) - \sin(2m\pi x))dx\\
&= \frac{-1}{(4m^2\pi^2)\,2^{\frac{5k}{2}}}\int_0^1 f''(\frac{x+n-1}{2^k})(\cos(2m\pi x) + \sin(2m\pi x))dx,\\
&= \frac{-1}{(4m^2\pi^2)\,2^{\frac{5k}{2}}}[\int_0^1 \{f''(\frac{x+n-1}{2^k}) - f''(\frac{n-1}{2^k})\}(\cos(2m\pi x) + \sin(2m\pi x))dx\\
&\quad -f''(\frac{n-1}{2^k})\int_0^1(\cos(2m\pi x) - \sin(2m\pi x))dx]\\
|c_{n,m}| &\leqslant \frac{1}{(4m^2\pi^2)\,2^{\frac{5k}{2}}}\int_0^1 |f''(\frac{x+n-1}{2^k}) - f''(\frac{n-1}{2^k})|\;|\cos(2m\pi x) + \sin(2m\pi x)|dx\\
&\leqslant \frac{B}{(4m^2\pi^2)\,2^{\frac{5k}{2}}}\int_0^1 |\frac{x}{2^k}|^{\alpha}\;|\cos(2m\pi x) + \sin(2m\pi x)|dx, \;\; \text{since} f'' \in H^{\alpha}[0,1)
\end{aligned}$$

Now by Cauchy Schwarz inequality, we have

$$|c_{n,m}| \leqslant \frac{B}{(4m^2\pi^2)\,2^{(\frac{5k}{2}+k\alpha)}}\{\int_0^1 |\frac{x}{2^k}|^{2\alpha}dx\}^{\frac{1}{2}}\{\int_0^1 |\cos(2m\pi x) + \sin(2m\pi x)|^2dx\}^{\frac{1}{2}}$$

$$
\begin{aligned}
|c_{n,m}| &\leqslant \frac{B}{(4m^2\pi^2)\ 2^{(\frac{5}{2}+\alpha)k}}\ (\frac{1}{\sqrt{2\alpha+1}}) \\
|c_{n,m}| &\leqslant \frac{B}{4m^2\pi^2\ \sqrt{2\alpha+1}\ 2^{(\frac{5}{2}+\alpha)k}} \qquad (21)
\end{aligned}
$$

From eq^n (20) and (21), we have

$$
\begin{aligned}
||f-S_{2^k,2M+1}(f)||_2^2 &= \sum_{n=1}^{2^k}(\sum_{m=-\infty}^{-M-1}+\sum_{m=M+1}^{\infty})|c_{n,m}|^2 \\
&\leqslant \sum_{n=1}^{2^k}(\sum_{m=-\infty}^{-M-1}+\sum_{m=M+1}^{\infty})\frac{B^2}{16m^4\pi^4\ (2\alpha+1)\ 2^{(5+2\alpha)k}}\ , \\
&= \frac{B^2}{16\pi^4\ (2\alpha+1)\ 2^{(5+2\alpha)k}}(\sum_{m=-\infty}^{-M-1}+\sum_{m=M+1}^{\infty})\frac{2^k}{m^4} \\
&= \frac{B^2}{16\pi^4\ (2\alpha+1)\ 2^{(4+2\alpha)k}}(\sum_{m=-\infty}^{-M-1}\frac{1}{m^4}+\sum_{m=M+1}^{\infty}\frac{1}{m^4}) \\
&= \frac{B^2}{16\pi^4\ (2\alpha+1)\ 2^{(4+2\alpha)k}}(\frac{1}{3(M+1)^3}+\frac{1}{3(M+1)^3}) \\
&= \frac{B^2}{24\pi^4\ (2\alpha+1)\ 2^{2k(\alpha+2)}}\frac{1}{(M+1)^3} \\
\therefore \min_{S_{2^k,M}(f)}||f-S_{2^k,2M+1}(f)||_2 &\leqslant \frac{B}{2\sqrt{6}\pi^2\ \sqrt{(2\alpha+1)}\ 2^{k(\alpha+2)}}\frac{1}{(M+1)^{\frac{3}{2}}}
\end{aligned}
$$

$$
\therefore E^{(2)}_{2^k,2M+1}(f) = \min_{S_{2^k,2M+1}(f)}||f-S_{2^k,2M+1}(f)||_2 = O(\frac{1}{(M+1)^{\frac{3}{2}}\ 2^{k(\alpha+2)}})
$$

Hence, theorem (5.2) has been proved.

6 Solution of the Fredholm integral equation of second kind

Consider the Fredholm integral equation of second kind given by eq^n (1).
Using CAS wavelet approximations,

$$
\begin{aligned}
u(x) &= U^T\Psi(x) = \Psi^T(x)U\ , \qquad (22) \\
f(x) &= F^T\Psi(x) = \Psi^T(x)F\ , \\
\text{and}\ \ K(x,y) &= \Psi^T(x)\mathbf{K}\Psi(y)\ ,
\end{aligned}
$$

where **K** is a square matrix of order $2^k(2M+1)$, which is calculated as follows

$$\int_0^1 \int_0^1 \psi_{n,m}(x)\psi_{n',m'}(y)K(x,y)dxdy\ , \tag{23}$$

where $1 \leqslant n, n^{'} \leqslant 2^k$ and $-M \leqslant m, m^{'} \leqslant M$, equation (1) becomes

$$\Psi^t(x)U = \Psi^t(x)F + \Psi^t(x)\mathbf{K}\int_0^1 \Psi(y)\Psi^t(y)Udy \tag{24}$$

By orthonormality of CAS wavelets, equation (24) reduces to

$$U = (I - \mathbf{K})^{-1}F \tag{25}$$

where I is identity matrix of order $2^k(2M+1)$. Subtituting the value of U from eq^{n} (25) in eq^{n} (22) , the solution u(x) of Fredholm integral equation of second kind (1) can be obtained.

6.1 Solution of integral eq^{n} (1) by Haar wavelet method

Let Haar wavelet solution of intgral eq^{n} (1) be of the form

$$u(x) = \sum_{i=1}^{2M} a_i h_i(x) \tag{26}$$

Subtituting the eq^{n} (26) in eq^{n} (1) , we have

$$\sum_{i=1}^{2M} a_i(h_i(x) - g_i(x)) = f(x) \tag{27}$$

where

$$g_i(x) = \int_0^1 k(x,y)h_i(y)dy \tag{28}$$

Taking the collocation points $x_k = \frac{k-\frac{1}{2}}{2M},\ k = 1, 2, ..., 2M$, in eq^{ns} (27) and (26), we obtain

$$\sum_{i=1}^{2M} a_i(h_i(x_k) - g_i(x_k)) = f(x_k) \tag{29}$$

$$\text{and } u(x_k) = \sum_{i=1}^{2M} a_i h_i(x_k) \tag{30}$$

The wavelet coefficients $a_i, i = 1, 2, ..., 2M$ are obtained by solving $2M$ system of equations in (29). Subtituting these coefficients in the eq^n(30) we can obtain the Haar wavelet solution of the integral eq^n (1).

7 Illustrated Numerical Examples

Two Fredholm integral equations have been solved by proposed method ie. CAS wavelet method discussed in this paper. Exact solutions of considered integral eq^n are compared with their approximate solutions obtained by CAS wavelet, Legendre wavelet and Haar wavelet method. The graphs of these solutions are plotted. It is observed that exact solution and approximate solutions of Fredholm integral equations obtained by CAS wavelet method are almost equal. The solutions of Fredholm integral equation derived by the help of CAS wavelet method are more closed than the solutions of this integral equation obtained by Legendre wavelet and Haar wavelet method. This comparison shows the advantages of proposed method of this paper. This is illustrated in following two examples.

Example 1

Subtituting $f(x) = \sin(8\pi x)$ and $K(x, y) = y^2$, in the Fredholm integral equation (1), it reduces to

$$u(x) = \sin(8\pi x) + \int_0^1 y^2 u(y)dy \tag{31}$$

The exact solution of integral eq^n (31) is given by

$$u(x) = \sin(8\pi x) - \frac{3}{16\pi} \tag{32}$$

CAS wavelet solution

For CAS wavelet solution, take $k = 2, M = 1$ in the eq^n (14) . In this case,

$$\begin{aligned}\Psi(x) &= [\psi_{1,-1}(x), \psi_{1,0}(x), \psi_{1,1}(x), \psi_{2,-1}(x), \psi_{2,0}(x), \psi_{2,1}(x),\\ &\quad \psi_{3,-1}(x), \psi_{3,0}(x), \psi_{3,1}(x), \psi_{4,-1}(x), \psi_{4,0}(x), \psi_{4,1}(x)]^T\end{aligned} \tag{33}$$

where

$$\left.\begin{aligned} \psi_{1,-1}(x) &= 2(\cos(8\pi x) - \sin(8\pi x)) \\ \psi_{1,0}(x) &= 2 \\ \psi_{1,1}(x) &= 2(\cos(8\pi x) + \sin(8\pi x)) \end{aligned}\right\} \quad 0 \leqslant x < \frac{1}{4}\,,$$

$$\left.\begin{aligned} \psi_{2,-1}(x) &= 2(\cos(8\pi x) - \sin(8\pi x)) \\ \psi_{2,0}(x) &= 2 \\ \psi_{2,1}(x) &= 2(\cos(8\pi x) + \sin(8\pi x)) \end{aligned}\right\} \quad \frac{1}{4} \leqslant x < \frac{1}{2}\,,$$

$$\left.\begin{aligned} \psi_{3,-1}(x) &= 2(\cos(8\pi x) - \sin(8\pi x)) \\ \psi_{3,0}(x) &= 2 \\ \psi_{3,1}(x) &= 2(\cos(8\pi x) + \sin(8\pi x)) \end{aligned}\right\} \quad \frac{1}{2} \leqslant x < \frac{3}{4}\,,$$

and

$$\left.\begin{aligned} \psi_{4,-1}(x) &= 2(\cos(8\pi x) - \sin(8\pi x)) \\ \psi_{4,0}(x) &= 2 \\ \psi_{4,1}(x) &= 2(\cos(8\pi x) + \sin(8\pi x)) \end{aligned}\right\} \quad \frac{3}{4} \leqslant x < 1\,.$$

$$F \;=\; [\frac{-1}{4}, 0, \frac{1}{4}, \frac{-1}{4}, 0, \frac{1}{4}, \frac{-1}{4}, 0, \frac{1}{4}, \frac{-1}{4}, 0, \frac{1}{4}]^{T},$$

The matrix **K** is calculated as follows:

$$
\begin{aligned}
K_{i,j} &= \int_0^1 \int_0^1 \psi_i(x) K(x,y) \psi_j(y) dy dx \\
&= \int_0^1 \psi_i(x) \left(\int_0^1 y^2 \psi_j(y) dy \right) dx \\
&= \left(\int_0^1 \psi_i(x) dx \right) \left(\int_0^1 y^2 \psi_j(y) dy \right)
\end{aligned}
$$

$$
\mathbf{K} = \begin{bmatrix} \frac{\pi+1}{64\pi^2} \\ \frac{1}{96} \\ -\frac{\pi+1}{64\pi^2} \\ \frac{3\pi+1}{64\pi^2} \\ \frac{7}{96} \\ -\frac{3\pi+1}{64\pi^2} \\ \frac{5\pi+1}{64\pi^2} \\ \frac{19}{96} \\ -\frac{5\pi+1}{64\pi^2} \\ \frac{7\pi+1}{64\pi^2} \\ \frac{37}{96} \\ -\frac{7\pi+1}{96\pi^2} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}
$$

$$\mathbf{K} = \begin{bmatrix} 0 & \frac{\pi+1}{128\pi^2} & 0 & 0 & \frac{\pi+1}{128\pi^2} & 0 & 0 & \frac{\pi+1}{128\pi^2} & 0 & 0 & \frac{\pi+1}{128\pi^2} & 0 \\ 0 & \frac{1}{192} & 0 & 0 & \frac{1}{192} & 0 & 0 & \frac{1}{192} & 0 & 0 & \frac{1}{192} & 0 \\ 0 & \frac{-\pi+1}{128\pi^2} & 0 & 0 & \frac{-\pi+1}{128\pi^2} & 0 & 0 & \frac{-\pi+1}{128\pi^2} & 0 & 0 & \frac{-\pi+1}{128\pi^2} & 0 \\ 0 & \frac{3\pi+1}{128\pi^2} & 0 & 0 & \frac{3\pi+1}{128\pi^2} & 0 & 0 & \frac{3\pi+1}{128\pi^2} & 0 & 0 & \frac{3\pi+1}{128\pi^2} & 0 \\ 0 & \frac{7}{192} & 0 & 0 & \frac{7}{192} & 0 & 0 & \frac{7}{192} & 0 & 0 & \frac{7}{192} & 0 \\ 0 & \frac{-3\pi+1}{128\pi^2} & 0 & 0 & \frac{-3\pi+1}{128\pi^2} & 0 & 0 & \frac{-3\pi+1}{128\pi^2} & 0 & 0 & \frac{-3\pi+1}{128\pi^2} & 0 \\ 0 & \frac{5\pi+1}{128\pi^2} & 0 & 0 & \frac{5\pi+1}{128\pi^2} & 0 & 0 & \frac{5\pi+1}{128\pi^2} & 0 & 0 & \frac{5\pi+1}{128\pi^2} & 0 \\ 0 & \frac{19}{192} & 0 & 0 & \frac{19}{192} & 0 & 0 & \frac{19}{192} & 0 & 0 & \dfrac{19}{192} & 0 \\ 0 & \frac{-5\pi+1}{128\pi^2} & 0 & 0 & \frac{-5\pi+1}{128\pi^2} & 0 & 0 & \frac{-5\pi+1}{128\pi^2} & 0 & 0 & \frac{-5\pi+1}{128\pi^2} & 0 \\ 0 & \frac{7\pi+1}{128\pi^2} & 0 & 0 & \frac{7\pi+1}{128\pi^2} & 0 & 0 & \frac{7\pi+1}{128\pi^2} & 0 & 0 & \frac{7\pi+1}{128\pi^2} & 0 \\ 0 & \frac{37}{192} & 0 & 0 & \frac{37}{192} & 0 & 0 & \frac{37}{192} & 0 & 0 & \frac{37}{192} & 0 \\ 0 & \frac{-7\pi+1}{128\pi^2} & 0 & 0 & \frac{-7\pi+1}{128\pi^2} & 0 & 0 & \frac{-7\pi+1}{128\pi^2} & 0 & 0 & \frac{-7\pi+1}{128\pi^2} & 0 \end{bmatrix}$$

$$\begin{aligned} \therefore U &= (I-\mathbf{K})^{-1}F \\ &= [\frac{-1}{4}, 0, \frac{1}{4}, \frac{-1}{4}, 0, \frac{1}{4}, \frac{-1}{4}, 0, \frac{1}{4}, \frac{-1}{4}, 0, \frac{1}{4}]^T \end{aligned} \tag{34}$$

Putting the values of $\Psi(x)$ and U from eq[ns] (33) and (34) in eq[n] (22), we have

$$u(x) = \sin(8\pi x) \tag{35}$$

which is the CAS wavelet solution of the integral equation (31) .

Legendre wavelet solution

Legendre wavelets $\psi^{(L)}_{n,m}(t) = \psi^{(L)}(k,n,m,t)$ having four arguments; $k = 2, 3, ...,$
$2n-1$, $n = 1, 2, 3, ..., 2^{k-1}$, m is the order of the Legendre polynomial and t is the normalised time, are defined by :

$$\psi^{(L)}_{n,m}(t) = \begin{cases} (m+\frac{1}{2})^{\frac{1}{2}} 2^{\frac{k}{2}} P_m(2^k t - 2n + 1), & if \ \frac{n-1}{2^{k-1}} \leqslant t < \frac{n}{2^{k-1}}, \\ 0, & otherwise. \end{cases} \tag{36}$$

where $P_m(t)$ are Legendre ploynomials of order m (Rehman and Khan [7]). The set $\{\psi^{(L)}_{n,m}\}_{n,m\in\mathbb{Z}}$ of Legendre wavelets forms an orthonormal set. A function $f \in L^2[0,1)$ may be expanded into Legendre wavelet series as:

$$f(t) = \sum_{n=1}^{\infty}\sum_{m=0}^{\infty} c_{n,m}\psi^{(L)}_{n,m}(t), \tag{37}$$

where $c_{n,m} =< f, \psi^{(L)}_{n,m} >$ .The series (37) may be truncated as:

$$(f)(t) \approx \sum_{n=1}^{2^{k-1}}\sum_{m=0}^{M-1} c_{n,m}\psi^{(L)}_{n,m}(t) = C^T\Psi^{(L)}(t) \tag{38}$$

where C and $\Psi^{(L)}(t)$ are $2^{k-1}M \times 1$ matrices given by:

$$\begin{aligned} C &= [c_{1,0}, c_{1,1}, ..., c_{1,M-1}, c_{2,0}, ..., c_{2,M-1}, ..., \\ & \quad c_{2^{k-1},0}, ..., c_{2^{k-1},M-1}]^T \end{aligned}$$

and

$$\begin{aligned} \Psi^{(L)}(t) &= [\psi^{(L)}_{1,0}(t), \psi^{(L)}_{1,1}(t), ..., \psi^{(L)}_{1,M-1}(t), \psi^{(L)}_{2,0}(t), ..., \psi^{(L)}_{2,M-1}(t), ..., \\ & \quad \psi^{(L)}_{2^{k-1},0}(t), ..., \psi^{(L)}_{2^{k-1},M-1}(t)]^T \end{aligned}$$

Similarly, a function $K \in L^2[0,1) \times L^2[0,1)$ may be approximated as:

$$K(x,y) \approx (\Psi^{(L)})^T(x)\mathbf{K}^{(L)}\Psi^{(L)}(y),$$

where $\mathbf{K}^{(L)}$ is $2^{k-1}M \times 2^{k-1}M$ matrix, whose entries are given by

$$\mathbf{K}^{(L)}_{i,j} =< \psi^{(L)}_i(x), < K(x,y), \psi^{(L)}_j(y) >> . \tag{39}$$

For Legendre wavelet solution, take $M = 3, k = 3$ in eq^n (38), then twelve basis functions are given by

$$\begin{aligned}\Psi^{(L)}(x) &= [\psi_{1,0}^{(L)}(x), \psi_{1,1}^{(L)}(x), \psi_{1,2}^{(L)}(x), \psi_{2,0}^{(L)}(x), \psi_{2,1}^{(L)}(x), \psi_{2,2}^{(L)}(x),\\ &\quad \psi_{3,0}^{(L)}(x), \psi_{3,1}^{(L)}(x), \psi_{3,2}^{(L)}(x), \psi_{4,0}^{(L)}(x), \psi_{4,1}^{(L)}(x), \psi_{4,2}^{(L)}(x)]^{T}\end{aligned} \tag{40}$$

where

$$\left.\begin{array}{l}\psi_{1,0}^{(L)}(x) = 2\\ \psi_{1,1}^{(L)}(x) = 2\sqrt{3}(8x-1)\\ \psi_{1,2}^{(L)}(x) = \sqrt{5}(3(8x-1)^2-1)\end{array}\right\} \quad 0 \leqslant x < \frac{1}{4}\,,$$

$$\left.\begin{array}{l}\psi_{2,0}^{(L)}(x) = 2\\ \psi_{2,1}^{(L)}(x) = 2\sqrt{3}(8x-3)\\ \psi_{2,2}^{(L)}(x) = \sqrt{5}(3(8x-3)^2-1)\end{array}\right\} \quad \frac{1}{4} \leqslant x < \frac{1}{2}\,,$$

$$\left.\begin{array}{l}\psi_{3,0}^{(L)}(x) = 2\\ \psi_{3,1}^{(L)}(x) = 2\sqrt{3}(8x-5)\\ \psi_{3,2}^{(L)}(x) = \sqrt{5}(3(8x-5)^2-1)\end{array}\right\} \quad \frac{1}{2} \leqslant x < \frac{3}{4}\,,$$

and

$$\left.\begin{array}{l}\psi_{4,0}^{(L)}(x) = 2\\ \psi_{4,1}^{(L)}(x) = 2\sqrt{3}(8x-7)\\ \psi_{4,2}^{(L)}(x) = \sqrt{5}(3(8x-7)^2-1)\end{array}\right\} \quad \frac{3}{4} \leqslant x < 1\,.$$

$$\mathbf{K}^{(L)} = \begin{bmatrix} \frac{1}{192} & 0 & 0 & \frac{1}{192} & 0 & 0 & \frac{1}{192} & 0 & 0 & \frac{1}{192} & 0 & 0 \\ \frac{\sqrt{3}}{384} & 0 & 0 & \frac{\sqrt{3}}{384} & 0 & 0 & \frac{\sqrt{3}}{384} & 0 & 0 & \frac{\sqrt{3}}{384} & 0 & 0 \\ \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 \\ \frac{7}{192} & 0 & 0 & \frac{7}{192} & 0 & 0 & \frac{7}{192} & 0 & 0 & \frac{7}{192} & 0 & 0 \\ \frac{\sqrt{3}}{128} & 0 & 0 & \frac{\sqrt{3}}{128} & 0 & 0 & \frac{\sqrt{3}}{128} & 0 & 0 & \frac{\sqrt{3}}{128} & 0 & 0 \\ \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 \\ \frac{19}{192} & 0 & 0 & \frac{19}{192} & 0 & 0 & \frac{19}{192} & 0 & 0 & \frac{19}{192} & 0 & 0 \\ \frac{5\sqrt{3}}{384} & 0 & 0 & \frac{5\sqrt{3}}{384} & 0 & 0 & \frac{5\sqrt{3}}{384} & 0 & 0 & \frac{5\sqrt{3}}{384} & 0 & 0 \\ \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 \\ \frac{37}{92} & 0 & 0 & \frac{37}{92} & 0 & 0 & \frac{37}{92} & 0 & 0 & \frac{37}{92} & 0 & 0 \\ \frac{7\sqrt{3}}{384} & 0 & 0 & \frac{7\sqrt{3}}{384} & 0 & 0 & \frac{7\sqrt{3}}{384} & 0 & 0 & \frac{7\sqrt{3}}{384} & 0 & 0 \\ \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 & \frac{\sqrt{5}}{1920} & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} F^{(L)} &= [0, \frac{-\sqrt{3}}{2\pi}, 0, 0, \frac{-\sqrt{3}}{2\pi}, 0, 0, \frac{-\sqrt{3}}{2\pi}, 0, 0, \frac{-\sqrt{3}}{2\pi}, 0]^T, \\ U^{(L)} &= (I - \mathbf{K}^{(L)})^{-1} F^{(L)} \\ &= [0, \frac{-\sqrt{3}}{2\pi}, 0, 0, \frac{-\sqrt{3}}{2\pi}, 0, 0, \frac{-\sqrt{3}}{2\pi}, 0, 0, \frac{-\sqrt{3}}{2\pi}, 0]^T. \end{aligned} \tag{41}$$

Putting the values of $\Psi^{(L)}(x)$ and $U^{(L)}$ from eq^{ns} (40) and (41) in eq^{n} (22), we get the Legendre wavelet solution of the integral equation (31) as:

$$u(x) = -\frac{\sqrt{3}}{2\pi}\psi_{1,1}^{(L)}(x) - \frac{\sqrt{3}}{2\pi}\psi_{2,1}^{(L)}(x) - \frac{\sqrt{3}}{2\pi}\psi_{3,1}^{(L)}(x) - \frac{\sqrt{3}}{2\pi}\psi_{4,1}^{(L)}(x) \tag{42}$$

Haar wavelet solution

The Haar wavelet family for $x \in [0,1]$ is defined as follows:

$$h_i(x) = \begin{cases} 1 & if \;\; x \in [\frac{k}{m}, \frac{k+\frac{1}{2}}{m}), \\ -1 & if x \in [\frac{k+\frac{1}{2}}{m}, \frac{k+1}{m}), \\ 0, & otherwise \end{cases} \tag{43}$$

where $m = 2^b,\ b = 0,1,...,J$ is the level of wavelet; $k = 0,1,...,m-1$ is the translation parameter. J is the maximum level of resulution. i is calculated by $i = m+k+1$. The minimum value of i for $m = 1, k = 0$ is 2. The maximum value of i is $i = 2M = 2^{J+1}$ (Arbabi and Darvishi [6]).

For $i = 1,\ h_1(x)$ is taken to be scaling function which is defined as follows:

$$h_1(x) = \begin{cases} 1 & if \;\; x \in [0,1), \\ 0, & otherwise \end{cases}$$

Any function $f(x)$ can be expressed in terms of Haar wavelets as follows:

$$f(x) = \sum_{i=1}^{2M} a_i h_i(x), \tag{44}$$

where the wavelet coefficients $a_i,\ i = 1,2,...,2M$ are to be determined. For Haar wavelet solution take $J = 3$ in eq$^{\text{n}}$ (43), $b = 0,1,2,3$, then $m = 2^b = 1,2,4,8$. By eq$^{\text{ns}}$ (28) the Haar wavelet coefficients $a_i,\ i = 1,2,...,16$ are given by

$$\begin{aligned}&[-0.008071, 0.001459, 0.002497, 0.001447, 0.000485, 0.006380,\\ &0.000488, -0.000476, 1.000010, 1, 1, 0.988178, 1, 1, 0.999039, 1]\end{aligned} \tag{45}$$

Putting these values of a_i in the eq$^{\text{n}}$ (26), we get the solution of integral equation (31) by Haar wavelet method. The Haar wavelet solutions of integral eq$^{\text{n}}$ 31 are shown in the Table (1).

The exact solution and approximate solutions of Fredholm integral equation (31) obtained by CAS wavelet, Legendre wavelet and Haar wavelet method for different values of x are given in the Table (1).

Table (1)

x	Exact sol^n by eq^n 32	CAS wavelet sol^n by eq^n (35)	Legendre wavelet sol^n by eq^n (42)	Haar wavelet sol^n by eq^n (26)
0	-0.059680	0	0.954930	0.996370
0.1	0.528105	0.587785	0.190986	-1.003630
0.2	-1.010736	-0.951056	-0.572958	0.995399
0.3	0.891376	0.951056	0.572958	0.995399
0.4	-0.647465	-0.587785	-0.190986	0.972689
0.5	-0.059680	0	-0.954930	0.992404
0.6	0.528105	0.587785	0.190986	-1.008083
0.7	-1.010736	-0.951056	-0.572958	-1.007595
0.8	0.891376	0.951056	0.572958	0.987585
0.9	-0.647465	-0.587785	-0.190986	0.989498

The graphs of the exact solution and approximate solutions of integral equation (31) obtained by CAS wavelet, Legendre wavelet and Haar wavelet method are shown in the Fig.(1).

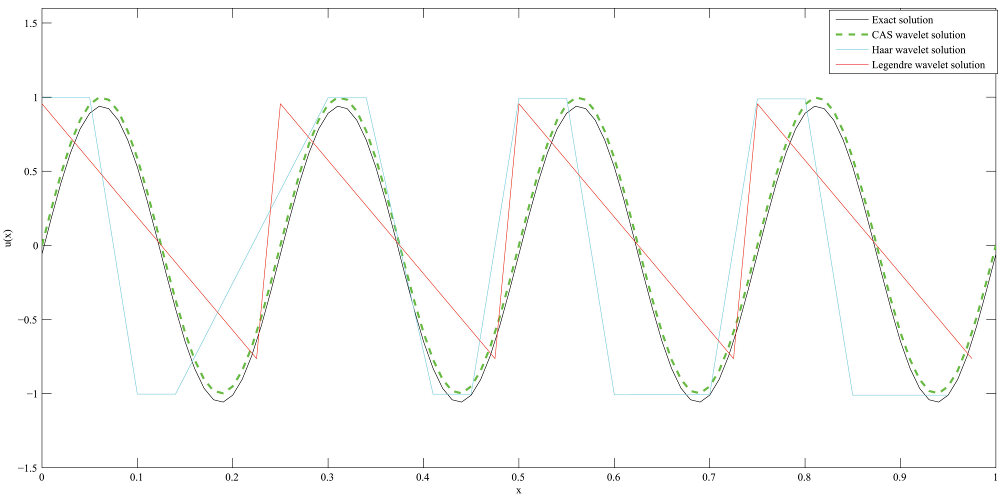


Fig.(1)

By numerical comparison in Table(1) and graphs shown in Fig.(1), it is clear that the solution of Fredholm integral equation (31) by CAS wavelet method is better than solutions obtained by Legendre wavelet and Haar wavelet methods.

Example 2

Consider the Fredholm integral equation:

$$u(x) = \sin(4\pi x) + \int_0^1 xyu(y)dy \; . \tag{46}$$

It is obtained by subtituting $f(x) = \sin(4\pi x)$ and $K(x,y) = xy$, in the Fredholm integral equation (1). The exact solution of Fredholm integral equation (46) is given by

$$u(x) = \sin(4\pi x) - \frac{3x}{8\pi} \tag{47}$$

CAS wavelet solution

For CAS wavelet solution, take $k = 1, M = 1$ in eq^{n} (14), then following the procedure of example (31), we have

$$F^* \;=\; [\frac{-1}{2\sqrt{2}}, 0, \frac{1}{2\sqrt{2}}, \frac{-1}{2\sqrt{2}}, 0, \frac{1}{2\sqrt{2}}]^T,$$

The matrix $\mathbf{K}^*$ is calculated as follows:

$$\begin{aligned} K^*_{i,j} &= \int_0^1 \int_0^1 \psi_i(x) K(x,y) \psi_j(y) dy dx \\ &= \int_0^1 \psi_i(x) \; (\int_0^1 xy\psi_j(y)dy) \; dx \\ &= (\int_0^1 x\psi_i(x)dx) \; (\int_0^1 y\psi_j(y)dy) \end{aligned}$$

$$\mathbf{K} = \begin{bmatrix} \frac{\sqrt{2}}{8\pi} \\ \frac{\sqrt{2}}{8} \\ -\frac{\sqrt{2}}{8\pi} \\ \frac{\sqrt{2}}{8\pi} \\ \frac{\sqrt{2}}{8} \\ -\frac{\sqrt{2}}{8\pi} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{8\pi} & \frac{\sqrt{2}}{8} & -\frac{\sqrt{2}}{8\pi} & \frac{\sqrt{2}}{8\pi} & \frac{\sqrt{2}}{8} & -\frac{\sqrt{2}}{8\pi} \end{bmatrix}$$

$$\mathbf{K}^* = \begin{bmatrix} \frac{1}{32\pi^2} & \frac{1}{32\pi} & \frac{-1}{32\pi^2} & \frac{1}{32\pi^2} & \frac{1}{32\pi} & \frac{-1}{32\pi^2} \\ \frac{1}{32\pi} & \frac{1}{32} & \frac{-1}{32\pi} & \frac{1}{32\pi} & \frac{1}{32} & \frac{-1}{32\pi} \\ \frac{-1}{32\pi^2} & \frac{-1}{32\pi} & \frac{1}{32\pi^2} & \frac{-1}{32\pi^2} & \frac{-1}{32\pi} & \frac{1}{32\pi^2} \\ \frac{1}{32\pi^2} & \frac{1}{32\pi} & \frac{-1}{32\pi^2} & \frac{1}{32\pi^2} & \frac{1}{32\pi} & \frac{-1}{32\pi^2} \\ \frac{3}{32\pi} & \frac{3}{32} & \frac{-3}{32\pi} & \frac{3}{32\pi} & \frac{3}{32} & \frac{-3}{32\pi} \\ \frac{-1}{32\pi^2} & \frac{-1}{32\pi} & \frac{1}{32\pi^2} & \frac{-1}{32\pi^2} & \frac{-1}{32\pi} & \frac{1}{32\pi^2} \end{bmatrix}$$

and

$$U^* = [\frac{-1}{2\sqrt{2}}, 0, \frac{1}{2\sqrt{2}}, \frac{-1}{2\sqrt{2}}, 0, \frac{1}{2\sqrt{2}}]^T.$$

$$u(x) = 1.0188 \sin(4\pi x) - 0.0294 \tag{48}$$

This is the approximate solution of the integral equation (46) by CAS wavelet method.

Legendre wavelet solution

For Legendre wavelet solution, take $M = 3, k = 2$ in eq^n (38), then we have

$$\Psi^{(L)}(x) = [\psi_{1,0}^{(L)}(x), \psi_{1,1}^{(L)}(x), \psi_{1,2}^{(L)}(x), \psi_{2,0}^{(L)}(x), \psi_{2,1}^{(L)}(x), \psi_{2,2}^{(L)}(x)]. \tag{49}$$

Following the procedure of the example (1), we have

$$\begin{aligned} (F^*)^{(L)} &= [0, \frac{-\sqrt{6}}{2\pi}, 0, 0, \frac{-\sqrt{6}}{2\pi}, 0]^T, \\ (U^*)^{(L)} &= [-0.0211, -0.4020, 0, -0.0633, -0.4020, 0]^T \end{aligned} \tag{50}$$

Putting the values of $\Psi^{(L)}(x)$ and $(U^*)^{(L)}$ from eq^{ns} (49) and (50) in eq^n (22), we get the solution of the integral equation (46) by Legendre wavelet method as

$$u(x) = -0.0211\psi_{1,0}^{(L)}(x) - 0.4020\psi_{1,1}^{(L)}(x) - 0.0633\psi_{2,0}^{(L)}(x) - 0.4020\psi_{2,1}^{(L)}(x) \tag{51}$$

$$(\mathbf{K}^*)^{(L)} = \begin{bmatrix} \frac{1}{32} & \frac{\sqrt{3}}{96} & 0 & \frac{3}{32} & \frac{\sqrt{3}}{96} & 0 \\ \frac{\sqrt{3}}{96} & \frac{1}{96} & 0 & \frac{\sqrt{3}}{32} & \frac{1}{96} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{3}{32} & \frac{\sqrt{3}}{32} & 0 & \frac{9}{32} & \frac{\sqrt{3}}{32} & 0 \\ \frac{\sqrt{3}}{96} & \frac{1}{96} & 0 & \frac{\sqrt{3}}{32} & \frac{1}{96} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Haar wavelet solution

For Haar wavelet solution, take $J = 2$ in eq^n (43),$b = 0, 1, 2$ then $m = 2^b = 1, 2, 4$. The Haar wavelet coefficients $a_i, i = 1, 2, ..., 8$ are given by

$$[0.061361, 0.027885, 0.616015, 0.670955, 0.000922, 1.465270, 0.000264, 0.004906]$$

Putting these values of a_i in the eq^n (26), we get the solution of integral equation (46) by Haar wavelet method. The Haar wavelet solutions of integral eq^n 46 are given in the Table (2).

The exact solution and approximate solutions of Fredholm integral equation (46) obtained by CAS wavelet, Legendre wavelet and Haar wavelet method for different values of x are given in the Table (2).

Table (2)

x	Exact sol^n by eq^n (47)	CAS wavelet sol^n by eq^n (48)	Legendre wavelet sol^n by eq^n (51)	Haar wavelet sol^n by eq^n (26)
0	0	-0.0294	0.9549	0.6955
0.1	0.9391	0.9395	0.5610	0.6955
0.2	0.5639	0.5694	0.1671	0.6722
0.3	-0.6236	-0.6282	-0.2268	-0.7701
0.4	-0.9988	-0.9983	-0.6207	-1.7239
0.5	-0.0597	-0.0294	0.8952	0.6194
0.6	0.8794	0.9395	0.5013	0.6194
0.7	0.5042	0.5694	0.1074	-0.3672
0.8	-0.6833	-0.6282	-0.2865	-0.8337
0.9	-1.0585	-0.9983	-0.6803	-0.8532

The graphs of the exact solution and approximate solutions of integral equation (46) obtained by CAS wavelet, Legendre wavelet and Haar wavelet method are shown in the Fig.(2).

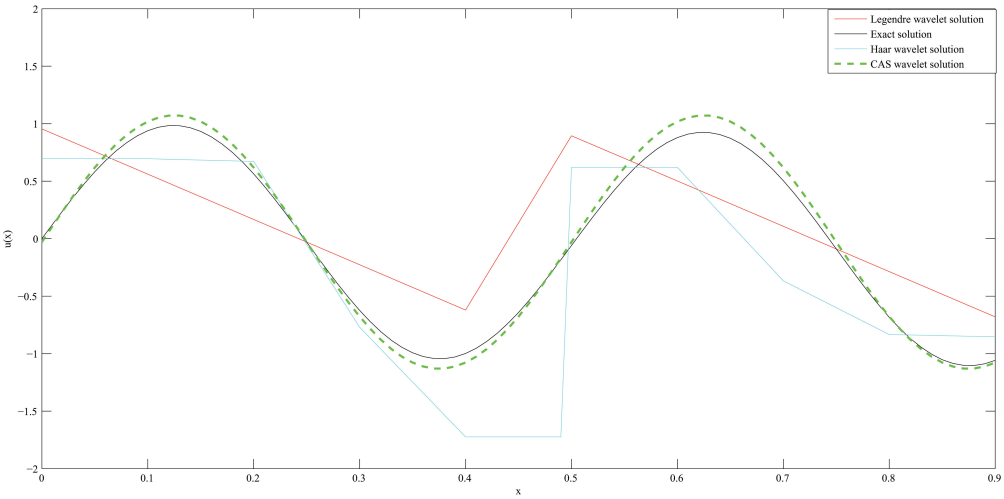


Fig.(2)

By numerical comparison in Table(2) and graphs shown in Fig.(2), it is observed that the solution of Fredholm integral equation (46) by CAS wavelet method is more accurate than solutions obtained by Legendre wavelet and Haar wavelet methods.

Note: The solutions of Fredholm integral equations in examples (1) and (2) by CAS wavelet method propoesd in this research paper and their numerical comparison with Legendre wavelet and Haar wavelet methods show the advantages of CAS wavelet method than Legendre wavelet and Haar wavelet methods.

8 Remarks

1. CAS wavelet approximation of Theorem (5.1) is given by
$E^{(1)}_{2^k,2M+1}(f) = O(\frac{1}{\sqrt{M+1}\ 2^{k(\alpha+1)}})$. $E^{(1)}_{2^k,2M+1}(f) \to 0$ as $M \to \infty,\ k \to \infty$.

CAS wavelet approximation of Theorem (5.2) is given by
$E^{(2)}_{2^k,2M+1}(f) = O(\frac{1}{(M+1)^{\frac{3}{2}}\ 2^{k(\alpha+2)}})$. $E^{(2)}_{2^k,2M+1}(f) \to 0$ as $M \to \infty,\ k \to \infty$.

Therefore, estimators $E^{(1)}_{2^k,2M+1}(f)$ and $E^{(2)}_{2^k,2M+1}(f)$ are best possible in wavelet analysis (Zygmund [5]).

2.
$$\begin{aligned} \because (M+1)^{\frac{3}{2}}2^{k(\alpha+2)} &\geqslant (M+1)^{\frac{1}{2}}2^{k(\alpha+1)},\ M\geqslant 1, k\geqslant 1 \\ \therefore \frac{1}{(M+1)^{\frac{3}{2}}2^{k(\alpha+2)}} &\leqslant \frac{1}{(M+1)^{\frac{1}{2}}2^{k(\alpha+1)}} \\ ie.\ E^{(2)}_{2^k,2M+1}(f) &\leqslant E^{(1)}_{2^k,2M+1}(f). \end{aligned}$$

Hence, estimator $E^{(2)}_{2^k,2M+1}(f)$ is sharper than estimator $E^{(1)}_{2^k,2M+1}(f)$. This shows that the estimator of a function f having $f'' \in H^{\alpha}[0,1)$ is sharper than the estimator of f having $f' \in H^{\alpha}[0,1)$.

3. CAS wavelet method is more effective than Legendre wavelet and Haar wavelet method in finding the solution of Fredholm integral equations (31) and (46).

4. Fredholm integral equation of first kind,

$$\int_0^1 K(x,t)y(t)dt = f(x)$$

can be solved by CAS wavelet method as follows:

$$\int_0^1 \Psi(x)\mathbf{K}\Psi^T(t)\Psi(t)Y = \Psi(x)F$$

Using orthonormality of CAS wavelet, we get $\mathbf{K}Y = F$. By finding the matrix **K** and F as in the case of Fredholm integral of second kind, we can find Y and hence the solution y(x).

9 Acknowledgments

Shyam Lal, one of the authors, is thankful to DST - CIMS for encouragement to this work.

Satish Kumar, one of the authors, is grateful to C.S.I.R. (India) for providing financial assistance in the form of Junior Research Fellowship vide Ref. No. 17/12/2017 (ii) EU-V Dated:13-02-2019 for his research work.

Authors are greatful to the referee for his valuable comments and suggestions, to improve the quality of the research paper.

References

[1] C.K.Chui, Wavelets: A mathematical tool for signal analysis, SIAM, Philadelphia PA,(1997).

[2] Anichini, Conti and Trotta: Some Results for Volterra Integra-differential equations depending on derivative in Unbounded Domains, Ration Mathematica, 37(2019) 55-38.

[3] S. Saha Ray and P.K. Sahu :Numerical Methods for Solving Fredholm Integral Equations of Second Kind. Hindawi Publishing Corporation, 426916(2013).

[4] Xiaoyang Zheng and Zhengyuan Wei : Estimates of Approximation Error by Legendre Wavelet. Applied Mathematics. 694-700(2016).

[5] Zygmund A.: Trigonometric Series, vol.I. Cambridge University Press,Cambridge (1959).

[6] Arbabi,Nazari,Darvishi : A two dimensional Haar wavelets method for solving systems of PDEs. Applied Mathematics and Computation 292 (2017) 33-77.

[7] Rehman and Khan : The Legendre wavelet method for solving fractional differential equations. Commun Nonlinear Sci Numer Simulat 16(2011) 4163-4173

Ratio Mathematica Volume 39, 2020, pp. 213-228

Uniqueness of an entire function sharing a polynomial with its linear differential polynomial

Imrul Kaish*
Nasir Uddin Gazi†

Abstract

In this paper we consider an entire function when it shares a polynomial with its linear differential polynomial. Our result is an improvement of a result of P.Li.

Keywords: Uniqueness; Entire function; Differential Polynomial; Sharing.
2010 AMS subject classifications: 30D35. [1]

*Department of Mathematics and Statistics, Aliah University, Kolkata, West Bengal 700160, India; imrulksh3@gmail.com.
†Department of Mathematics and Statistics, Aliah University, Kolkata, West Bengal 700160, India; nsrgazi@gmail.com
[1]Received on November 2nd, 2020. Accepted on December 17th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.554. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction, Definitions and Results

Let f be a non-constant meromorphic function defined in the open complex plane $\mathbb{C}$ and $a = a(z)$ be a polynomial. We denote by $E(a; f)$ the set of zeros of $f - a$, counted with multiplicities and by $\overline{E}(a; f)$ the set of distinct zeros of $f - a$.

If for two non-constant meromorphic functions f and g, we have $E(a; f) = E(a; g)$, we say that f and g share a CM and if $\overline{E}(a; f) = \overline{E}(a; g)$, we say that f and g share a IM.

We denote by $S(r, f)$ any function satisfying $S(r, f) = o\{T(r, f)\}$, as $r \to \infty$, possibly outside of a set with finite measure.

For an entire function f, we define $deg(f)$ in the following way:

$deg(f) = \infty$, if f is a transcendental entire function and $deg(f)$ is the degree of the polynomial, if f is a polynomial.

The investigation of uniqueness of an entire function sharing two values introduced by L. A. Rubel and C. C. Yang [Rubel and Yang, 1977] in 1977. Following is their result.

Theorem A. *[Rubel and Yang, 1977] Let f be a non-constant entire function. If $E(a; f) = E(a; f^{(1)})$ and $E(b; f) = E(b; f^{(1)})$, for distinct finite complex numbers a and b, then $f \equiv f^{(1)}$.*

In 1979 E. Mues and N. Steinmetz [Mues and Steinmetz, 1979] tried to improve Theorem A by considering IM sharing of values. They proved the following theorem.

Theorem B. *[Mues and Steinmetz, 1979]. Let f be a non-constant entire function and a, b be two distinct finite complex values. If $\overline{E}(a; f) = \overline{E}(a; f^{(1)})$ and $\overline{E}(b; f) = \overline{E}(b; f^{(1)})$, then $f \equiv f^{(1)}$.*

In 1986 G. Jank, E. Mues and L. Volkmann [Jank et al., 1986] considered an entire function sharing a nonzero value with its derivatives and they proved the following result.

Theorem C. *[Jank et al., 1986] Let f be a non-constant entire function and a be a non-zero finite value. If $\overline{E}(a; f) = \overline{E}(a; f^{(1)}) \subset \overline{E}(a; f^{(2)})$, then $f \equiv f^{(1)}$.*

H. Zhong [Zhong, 1995] tried to improve Theorem C by taking higher order derivatives. By the following example he concluded that in Theorem C the second derivative cannot be straight way replaced by any higher order derivatives.

Example 1.1. *[Zhong, 1995] Let $k(\geq 3)$ be a positive integer and $\omega(\neq 1)$ be a $(k-1)$th root of unity. If $f = e^{\omega z} + \omega - 1$, then f, $f^{(1)}$, and $f^{(k)}$ share the value ω CM, but $f \not\equiv f^{(1)}$.*

Considering two consecutive higher order derivatives H. Zhong [Zhong, 1995] improved Theorem C in another direction. The following is the improved result.

Theorem D. *[Zhong, 1995] Let f be a non-constant entire function and a be a non-zero finite value. If $E(a; f) = E(a; f^{(1)})$ and $\overline{E}(a; f) \subset \overline{E}(a; f^{(n)}) \cap \overline{E}(a; f^{(n+1)})$ for $n(\geq 1)$, then $f \equiv f^{(n)}$.*

For further discussion we need the following notation. Let f be a non-constant meromorphic function, $a = a(z)$ be a polynomial and A be a set of complex numbers. We denote by $n_A(t, a; f)$, the number of zeros of $f - a$, counted according to their multiplicities which lie in $A \cap \{z : |z| \leq r\}$. The integrated counting function $N_A(r, a; f)$ of the zeros of $f - a$ which lie in $A \cap \{z : |z| \leq r\}$ is defined as

$$N_A(r, a; f) = \int_0^r \frac{n_A(t, a; f) - n_A(0, a; f)}{t} dt + n_A(0, a; f) \log r,$$

where $n_A(0, a; f)$ denotes the multiplicity of zeros of $f - a$ at origin. $\overline{N}_A(r, a; f)$ be the reduced counting function of zeros of $f - a$ in $A \cap \{z : |z| \leq r\}$. Clearly if $A = \mathbb{C}$ then $N_A(r, a; f) = N(r, a; f)$ and $\overline{N}_A(r, a; f) = \overline{N}(r, a; f)$.

For standard definitions and notations of the value distribution theory we refer the reader to [Hayman, 1964] and [Yang and Yi, 2003].

Recently I. Lahiri and I. Kaish [Lahiri and Kaish, 2017] improved Theorem D by considering a shared polynomial. They proved the following result.

Theorem E. *[Lahiri and Kaish, 2017] Let f be a non-constant entire function and $a = a(z)(\not\equiv 0)$ be a polynomial with $deg(a) \neq deg(f)$. Suppose that $A = \overline{E}(a; f) \Delta \overline{E}(a; f^{(1)})$ and $B = \overline{E}(a, f^{(1)}) \backslash \{\overline{E}(a, f^{(n)}) \cap \overline{E}(a, f^{(n+1)})\}$, where $\triangle$ denotes the symmetric difference of sets and $n(\geq 1)$ is an integer. If*

(i) $N_A(r, a; f) + N_A(r, a; f^{(1)}) = O\{logT(r, f)\}$,

(ii) $N_B(r, a; f^{(1)}) = S(r, f)$ *and*

(iii) each common zero of $f - a$ and $f^{(1)} - a$ has the same multiplicity,

then $f = \lambda e^z$, where $\lambda(\neq 0)$ is a constant.

Throughout the paper we denote by $L = L(f)$ a nonconstant linear differential polynomial generated by f of the form

$$L = L(f) = a_1 f^{(1)} + a_2 f^{(2)} + ... + a_n f^{(n)}, \tag{1}$$

where $a_1, a_2, ..., a_n(\neq 0)$ are constants.

Considering Linear differential polynomial P.Li [Li, 1999] improved Theorem D in the following way.

Theorem F. *[Li, 1999]. Let f be a non-constant entire function and L be defined in (1) and a be a non-zero finite complex number. If $\overline{E}(a; f) = \overline{E}(a; f^{(1)}) \subset \overline{E}(a; L) \cap \overline{E}(a; L^{(1)})$ then $f = f^{(1)} = L$.*

In this paper we extend Theorem D and Theorem F in the following way

Theorem 1.1. *Let f be a non-constant entire function, L be defined in (1) and $a = a(z)(\not\equiv 0)$ be a polynomial with $deg(a) \neq deg(f)$. Suppose that $A = \overline{E}(a; f)\Delta\overline{E}(a; f^{(1)})$ and $B = \overline{E}(a, f^{(1)})\backslash\{\overline{E}(a, L^{(p)}) \cap \overline{E}(a, L^{(q)})\}$ where p, q are integers satisfying $q > p \geq deg(a)$.*

If

(i) $N_A(r, a; f) + N_A(r, a; f^{(1)}) = O\{\log T(r, f)\}$,

(ii) $N_B(r, a; f^{(1)}) = S(r, f)$ *and*

(iii) each common zero of $f - a$ and $f^{(1)} - a$ has the same multiplicity,

then $f = L = \lambda e^z$, where $\lambda(\neq 0)$ is a constant.

Putting $A = B = \emptyset$ we get the following corollary.

Corolary 1.1. *Let f be a non-constant entire function, L be defined in (1) and $a = a(z)(\not\equiv 0)$ be a polynomial with $deg(a) \neq deg(f)$. If $E(a; f) = E(a; f^{(1)})$ and $\overline{E}(a, f^{(1)}) \subset \overline{E}(a, L^{(p)}) \cap \overline{E}(a, L^{(q)})$ where p, q are integers satisfying $q > p \geq deg(a)$, then $f = L = \lambda e^z$, where $\lambda(\neq 0)$ is a constant.*

Remark 1.1. *If in Corollary 1.1, a is a non-zero constant and $p = deg(a) = 0, q = p + 1$ then it is a particular form of Theorem F.*

Remark 1.2. *If in (1), $a_1 = a_2 = ...a_{n-1} = 0$ and $a_n = 1$ then $L = f^{(n)}$ and if in Corollary 1.1, a is a non-zero constant and $p = deg(a), q = p + 1$, then Corollary 1.1 is the Theorem D.*

Remark 1.3. *It is an open problem whether the Theorem 1.1 is valid or not if we omit the condition $p \geq deg(a)$.*

2 Lemmas

In this section we present some necessary lemmas.

Lemma 2.1. *[Lahiri and Kaish, 2017]. Let f be a transcendental entire function of finite order and $a = a(z)(\not\equiv 0)$ be a polynomial and $A = \overline{E}(a; f)\Delta\overline{E}(a; f^{(1)})$.*

If

(i) $N_A(r,a;f) + N_A(r,a;f^{(1)}) = O\{\log T(r,f)\}$,

(ii) each common zero of $f-a$ *and* $f^{(1)}-a$ *has the same multiplicity,*

then $m(r,a;f) = S(r,f)$.

Lemma 2.2. *[Lain, 1993]. Suppose* f *be an entire function,* $a_0, a_1, ...a_n$ *are polynomials and* a_0, a_n *are not identically zero. Then each solution of the linear differential equation* $a_n f^{(n)} + a_{n-1} f^{(n-1)} + ... + a_0 f = 0$ *is of finite order.*

Lemma 2.3. *[Hayman, 1964]. Let* f *be a non-constant meromorphic function and* a_1, a_2, a_3 *be three distinct meromorphic functions satisfying* $T(r,a_\nu) = S(r,f)$ *for* $\nu = 1,2,3$ *then*

$$T(r,f) \le \overline{N}(r,0;f-a_1) + \overline{N}(r,0;f-a_2) + \overline{N}(r,0;f-a_3) + S(r,f).$$

Lemma 2.4. *Let* f *be a transcendental entire function and* $a = a(z)(\not\equiv 0)$ *be a polynomial. Also let* $L(f)$, $L(a)$ *be the linear differential polynomials generated by* f *and* a *respectively. Suppose*

$$h = \frac{(a-a^{(1)})(L^{(p)}(f) - L^{(p)}(a)) - (a - L^{(p)}(a))(f^{(1)} - a^{(1)})}{f-a},$$

$A = \overline{E}(a;f)\backslash\overline{E}(a;f^{(1)})$ *and* $B = \overline{E}(a,f^{(1)})\backslash\{\overline{E}(a,L^{(p)}) \cap \overline{E}(a,L^{(q)})\}$, *where* p,q *are integers satisfying* $0 \le p < q$.

If

(i) $N_A(r,a;f) + N_B(r,a;f^{(1)}) = S(r,f)$,

(ii) each common zero of $f-a$ *and* $f^{(1)}-a$ *has the same multiplicity,*

(iii) h *is a transcendental entire or meromorphic,*

then $m(r,a,f^{(1)}) = S(r,f)$.

Proof. Since $a - a^{(1)} = (f^{(1)} - a^{(1)}) - (f^{(1)} - a)$, if z_0 be a common zero of $f-a$ and $f^{(1)}-a$ with multiplicity $r(\ge 2)$, then z_0 is a zero of $a-a^{(1)}$ with multiplicity $r-1$. So

$$N_{(2}(r,a;f) \le 2N(r,0;a-a^{(1)}) + N_A(r,a;f) = S(r,f), \tag{2}$$

where $N_{(2}(r,a;f)$ be the counting function of multiple zeros of $f-a$.

Using (2) and from the hypothesis we get

$$\begin{aligned} N(r,h) &\le N_A(r,a;f) + N_B(r,a;f^{(1)}) + N_{(2}(r,a;f) + S(r,f) \\ &= S(r,f) \end{aligned}$$

Since $m(r,h)=S(r,f)$, we have $T(r,h)=S(r,f)$
From $h=\dfrac{(a-a^{(1)})(L^{(p)}(f)-L^{(p)}(a))-(a-L^{(p)}(a))(f^{(1)}-a^{(1)})}{f-a}$, we get

$$
\begin{aligned}
f &= a+\frac{1}{h}\{(a-a^{(1)})(L^{(p)}(f)-L^{(p)}(a))-(a-L^{(p)}(a))(f^{(1)}-a^{(1)})\} \\
&= a+\frac{1}{h}\{(a-a^{(1)})(L^{(p)}(f)-a)-(a-L^{(p)}(a))(f^{(1)}-a)\}. \qquad (3)
\end{aligned}
$$

Case 1. Let $p>0$. Differentiating (3) we get

$$
\begin{aligned}
f^{(1)} &= a^{(1)}+(\frac{1}{h})^{(1)}\{(a-a^{(1)})(L^{(p)}(f)-a)-(a-L^{(p)}(a))(f^{(1)}-a)\}+ \\
&\frac{1}{h}\{(a^{(1)}-a^{(2)})(L^{(p)}(f)-a)+(a-a^{(1)})(L^{(p+1)}-a^{(1)})\}- \\
&\frac{1}{h}\{(a^{(1)}-L^{(p+1)}(a))(f^{(1)}-a)+(a-L^{(p)}(a))(f^{(2)}-a^{(1)})\}.
\end{aligned}
$$

This implies
$(f^{(1)}-a)\{1+(\frac{1}{h})^{(1)}(a-L^{(p)}(a))+\frac{1}{h}(a^{(1)}-L^{(p+1)}(a))\}$
$=a^{(1)}-a+(\frac{1}{h})^{(1)}(a-a^{(1)})(L^{(p)}(f)-a)+\frac{1}{h}(a^{(1)}-a^{(2)})(L^{(p)}(f)-a)+\frac{1}{h}(a-a^{(1)})(L^{(p+1)}(f)-a^{(1)})-\frac{1}{h}(a-L^{(p)}(a))(f^{(2)}-a^{(1)})$
$=a^{(1)}-a+(\frac{a-a^{(1)}}{h})^{(1)}(L^{(p)}(f)-L^{(p-1)}(a))+(\frac{a-a^{(1)}}{h})^{(1)}(L^{(p-1)}(a)-a)+\frac{a-a^{(1)}}{h}(L^{(p+1)}(f)-L^{(p)}(a))+\frac{a-a^{(1)}}{h}(L^{(p)}(a)-a^{(1)})-\frac{1}{h}(a-L^{(p)}(a))(f^{(2)}-a^{(1)})$,
or,
$(f^{(1)}-a)\{1+(\frac{a-L^{(p)}(a)}{h}))^{(1)}\}=(a^{(1)}-a)+\{(\frac{a-a^{(1)}}{h})(L^{(p-1)}(a)-a)\}^{(1)}+(\frac{a-a^{(1)}}{h})^{(1)}(L^{(p)}(f)-L^{(p-1)}(a))+\frac{a-a^{(1)}}{h}(L^{(p+1)}(f)-L^{(p)}(a))-\frac{1}{h}(a-L^{(p)}(a))(f^{(2)}-a^{(1)})$,
or

$$
\begin{aligned}
\frac{1}{f^{(1)}-a} &= \frac{h_1}{h_2}-\frac{1}{h_2}(\frac{a-a^{(1)}}{h})^{(1)}(\frac{L^{(p)}(f)-L^{(p-1)}(a)}{f^{(1)}-a}) \\
&+(\frac{a-a^{(1)}}{hh_2})(\frac{L^{(p+1)}(f)-L^{(p)}(a)}{f^{(1)}-a}) \\
&-\frac{1}{hh_2}(a-L^{(p)}(a))(\frac{f^{(2)}-a^{(1)}}{f^{(1)}-a}), \qquad (4)
\end{aligned}
$$

where $h_1=1+(\frac{a-L^{(p)}(a)}{h})^{(1)}$,
$h_2=a^{(1)}-a+\{(\frac{a-a^{(1)}}{h})(L^{(p-1)}(a)-a)\}^{(1)}$.
We now verify that $h_1\not\equiv 0, h_2\not\equiv 0$.

If $h_1 \equiv 0$, then $1 + (\frac{a-L^{(p)}(a)}{h})^{(1)} \equiv 0$. Integrating we get $\frac{1}{h} = \frac{c_1 - z}{a-L^{(p)}(a)}$, where c_1 is a constant. This is a contradiction, because h is transcendental.

If $h_2 \equiv 0$, then $a^{(1)} - a + \{(\frac{a-a^{(1)}}{h})(L^{(p-1)}(a) - a)\}^{(1)} \equiv 0$. Integrating we get $h = \frac{(a-a^{(1)})(L^{(p-1)}(a)-a)}{P(z)}$, where $P(z)$ is a polynomial. This is again a contradiction. Therefore $h_1 \not\equiv 0, h_2 \not\equiv 0$. Again $T(r, h_1) + T(r, h_1) = S(r, f)$, since $T(r, h) = S(r, f)$.

Now from (4) and using Lemma of logarithmic derivative we get $m(r, a; f^{(1)}) = m(r, \frac{1}{f^{(1)}-a}) = S(r, f)$.

Case 2. Let $p = 0$. Then $L^{(p)}(f) = L(f)$.

Suppose $L(f) = a_1 f^{(1)} + a_2 f^{(2)} + ... + a_n f^{(n)}$
and $L(a) = a_1 a^{(1)} + a_2 a^{(2)} + ... + a_n a^{(n)}$, where $a_1, a_2, ..., a_n (\neq 0)$ are constant, $n (\geq 1)$ be an integer.

From the definition of h we get
$f = a + \frac{1}{h}\{(a - a^{(1)})(L(f) - a) - (a - L(a))(f^{(1)} - a)\}$
Differentiating we get

$$\begin{aligned}
f^{(1)} &= a^{(1)} + (\frac{1}{h})^{(1)}\{(a - a^{(1)})(L(f) - a) - (a - L(a))(f^{(1)} - a)\} \\
&\quad + \frac{1}{h}\{(a^{(1)} - a^{(2)})(L(f) - a) + (a - a^{(1)})(L^{(1)}(f) - a^{(1)})\} \\
&\quad - \frac{1}{h}\{(a^{(1)} - L^{(1)}(a))(f^{(1)} - a) - (a - L(a))(f^{(2)} - a^{(1)})\}.
\end{aligned}$$

This implies
$(f^{(1)} - a)\{1 + (\frac{a-L(a)}{h})^{(1)}\} = (a^{(1)} - a) + (\frac{a-a^{(1)}}{h})^{(1)}(L(f) - a) + \frac{a-a^{(1)}}{h}(L^{(1)}(f) - a^{(1)}) - \frac{a-L(a)}{h}(f^{(2)} - a^{(1)}) = (a^{(1)} - a) + (\frac{a-a^{(1)}}{h})^{(1)}(L(f) - L_1(a)) + (\frac{a-a^{(1)}}{h})^{(1)}(L_1(a) - a) + (\frac{a-a^{(1)}}{h})(L^{(1)}(f) - L(a)) + (\frac{a-a^{(1)}}{h})(L(a) - a^{(1)}) - \frac{a-L(a)}{h}(f^{(2)} - a^{(1)}) = (a^{(1)} - a) + \{(\frac{a-a^{(1)}}{h})(L_1(a) - a)\}^{(1)} + (\frac{a-a^{(1)}}{h})^{(1)}(L(f) - L_1(a)) + (\frac{a-a^{(1)}}{h})(L^{(1)}(f) - L(a)) - \frac{a-L(a)}{h}(f^{(2)} - a^{(1)})$

Or,

$$\begin{aligned}
\frac{1}{f^{(1)} - a} &= \frac{h_3}{h_4} - \frac{1}{h_4}(\frac{a - a^{(1)}}{h})^{(1)}(\frac{L(f) - L_1(a)}{f^{(1)} - a}) \\
&\quad + (\frac{a - a^{(1)}}{hh_4})(\frac{L^{(1)}(f) - L(a)}{f^{(1)} - a}) - (\frac{a - L(a)}{hh_4})(\frac{f^{(2)} - a^{(1)}}{f^{(1)} - a}), \quad (5)
\end{aligned}$$

where
$L_1(a) = a_1 a + a_2 a^{(1)} + ... + a_n a^{(n-1)}$,
$h_3 = 1 + (\frac{a-L(a)}{h})^{(1)}$ and
$h_4 = a^{(1)} - a + \{(\frac{a-a^{(1)}}{h})(L_1(a) - a)\}^{(1)}$

Similarly as in Case 1, $h_3 \not\equiv 0$, $h_4 \not\equiv 0$. Also $T(r,h_3)+T(r,h_4)=S(r,f)$. Therefore from (5) and using Lemma of logarithmic derivative we get
$m(r,a;f^{(1)}) = m(r,\frac{1}{f^{(1)}-a}) = S(r,f)$.
This completes the proof of the lemma.

□

Lemma 2.5. *Let f be a transcendental entire function, $a=a(z)(\not\equiv 0)$ be a polynomial and $L=L(f)$ be define in (1). Suppose*

(i) $N_A(r,a;f)+N_A(r,a;f^{(1)})=S(r,f)$, where $A=\overline{E}(a;f)\Delta\overline{E}(a;f^{(1)})$

(ii) $N_B(r,a;f^{(1)}))=S(r,f)$, where $B=\overline{E}(a,f^{(1)})\backslash\{\overline{E}(a,L^{(p)})\cap\overline{E}(a,L^{(q)})\}$ p,q are integers satisfying $q>p\geq deg(a)$,

(iii) each common zero of $f-a$ and $f^{(1)}-a$ has the same multiplicity,

(iv) m(r, a; f) = S(r, f), then $f=L=\lambda e^z$, where $\lambda(\neq 0)$ is a constant.

Proof. Let

$$\alpha = \frac{f^{(1)}-a}{f-a}, \tag{6}$$

From the hypothesis we get,

$$N(r,\alpha)\leq N_A(r,a;f)+S(r,f)=S(r,f)$$

and

$$\begin{aligned} m(r,\alpha) &= m(r,\frac{f^{(1)}-a}{f-a}) \\ &= m(r,\frac{f^{(1)}-a^{(1)}+a^{(1)}-a}{f-a}) \\ &\leq m(r,a;f)+S(r,f) \\ &= S(r,f). \end{aligned}$$

Therefore $T(r,\alpha)=S(r,f)$.

From (6) we get

$$\begin{aligned} f^{(1)} &= \alpha f + a(1-\alpha) \\ &= \alpha_1 f+\beta_1, \end{aligned}$$

where $\alpha_1=\alpha$ and $\beta_1=a(1-\alpha)$

Differentiating we get,

$$f^{(2)}=\alpha_2 f+\beta_2,$$

where $\alpha_2 = \alpha_1^{(1)} + \alpha_1\alpha_1$ and $\beta_2 = \beta_1^{(1)} + \alpha_1\beta_1$.

Similarly,

$$f^{(k)} = \alpha_k f + \beta_k,$$

where $\alpha_{k+1} = \alpha_k^{(1)} + \alpha_1\alpha_k$ and $\beta_{k+1} = \beta_k^{(1)} + \alpha_k\beta_1$.

Clearly $T(r, \alpha_k) + T(r, \beta_k) = S(r, f)$, because $T(r, \alpha) = S(r, f)$.

Now

$$\begin{aligned} L^{(p)} &= \sum_{k=1}^{n} a_k f^{(p+k)} \\ &= (\sum_{k=1}^{n} a_k\alpha_{p+k})f + (\sum_{k=1}^{n} a_k\beta_{p+k}) \\ &= \mu_1 f + \nu_1, \end{aligned} \tag{7}$$

where $\mu_1 = \sum\limits_{k=1}^{n} a_k\alpha_{p+k}$, $\nu_1 = \sum\limits_{k=1}^{n} a_k\beta_{p+k}$

$$\begin{aligned} L^{(q)} &= \sum_{k=1}^{n} a_k f^{(q+k)} \\ &= (\sum_{k=1}^{n} a_k\alpha_{q+k})f + (\sum_{k=1}^{n} a_k\beta_{q+k}) \\ &= \mu_2 f + \nu_2, \end{aligned} \tag{8}$$

where $\mu_2 = \sum\limits_{k=1}^{n} a_k\alpha_{q+k}$, $\nu_2 = \sum\limits_{k=1}^{n} a_k\beta_{q+k}$.

Clearly $T(r, \mu_i) + T(r, \nu_i) = S(r, f)$, $i = 1, 2$.

Let $D = \overline{E}(a; f) \cap \overline{E}(a; f^{(1)}) \cap \overline{E}(a; L^{(p)}) \cap \overline{E}(a; L^{(q)})$.

Note that $D \neq \emptyset$, because otherwise, $N(r, a; f) = S(r, f)$. Then from the hypothesis $T(r, f) = S(r, f)$, a contradiction.

Let $z_1 \in D$ then $f(z_1) = f^{(1)}(z_1) = L^{(p)}(z_1) = L^{(q)}(z_1) = a(z_1)$.

Now from (7) and (8) we get $a(z_1) = \mu_1(z_1)a(z_1) + \nu_1(z_1)$ and $a(z_1) = \mu_2(z_1)a(z_1) + \nu_2(z_1)$

If $\mu_1 a + \nu_1 - a \not\equiv 0$, then

$$\begin{aligned} N(r, a; f) &\leq N_A(r, a; f) + N_B(r, a; f^{(1)}) + N_D(r, a; f) + S(r, f) \\ &\leq N_A(r, 0; \mu_1 a + \nu_1 - a) + S(r, f) \\ &= S(r, f), \end{aligned}$$

a contradiction. Therefore

$$\mu_1 a + \nu_1 - a \equiv 0. \tag{9}$$

Similarly

$$\mu_2 a + \nu_2 - a \equiv 0. \tag{10}$$

From (9) and (10) we get $\mu_1 \equiv \mu_2 \equiv 1$ and $\nu_1 \equiv 0 \equiv \nu_2$.
Then from (7)

$$L^{(p)} \equiv f. \tag{11}$$

Also $\mu_1 \equiv 1$ implies

$$\sum_{k=1}^{n} a_k \alpha_{p+k} \equiv 1. \tag{12}$$

From (12) we see that α has no pole. Because if α has a pole of order $d(\geq 1)$ then the left hand side of (12) has a pole of order $(p+k)d$ but the right hand side is a constant.

Again by simple calculation from (12) we get

$$a_n \alpha^{n+p} + P[\alpha] \equiv 0. \tag{13}$$

where $P[\alpha]$ is a differential polynomial in α with degree not exceeding $(n+p-1)$.

If α is transcendental entire, then by Clunie's Lemma we have $m(r, \alpha) = S(r, \alpha)$, a contradiction.

If α is a nonconstant polynomial then left hand side of (13) is also a nonconstant polynomial, which is again a contradiction.

Therefore α is a constant.

Now from $\frac{f^{(1)}-a}{f-a} = \alpha$, we get $f^{(1)} - \alpha f = a(1-\alpha)$.

Integrating we get

$$\begin{aligned} e^{-\alpha z} f &= (1-\alpha) \int a e^{-\alpha z} dz \\ &= (1-\alpha) P(z) e^{-\alpha z} + \lambda, \end{aligned}$$

where $\lambda(\neq 0)$ is a constant and $P(z)$ is a polynomial of degree atmost $deg(a)$,

or, $f = (1-\alpha)P(z) + \lambda e^{\alpha z}$.

Now $f^{(r+1)} = \lambda \alpha^{r+1} e^{\alpha z}$, if $r = deg(a)$

Therefore

$$\begin{aligned}
L^{(p)} &= \sum_{k=1}^{n} a_k f^{(p+k)} \\
&= (\sum_{k=1}^{n} a_k \alpha^{p+k})\lambda e^{\alpha z} \\
&= \lambda e^{\alpha z} \\
&= \frac{f^{(1)}}{\alpha} - \frac{1-\alpha}{\alpha} p^{(1)}(z), \qquad (14)
\end{aligned}$$

Suppose $\alpha \neq 1$.

Since $D = \overline{E}(a; f) \cap \overline{E}(a; f^{(1)}) \cap \overline{E}(a; L^{(p)}) \cap \overline{E}(a; L^{(q)}) \neq \emptyset$,
we have $f(z_2) = f^{(1)}(z_2) = L^{(p)}(z_2) = L^{(q)}(z_2) = a(z_2)$, for some $z_2 \in D$.
From (14) we get

$$a(z_2) = \frac{a(z_2)}{\alpha} - \frac{1-\alpha}{\alpha} P^{(1)}(z_2)$$

or,

$$a(z_2)(1 - \frac{1}{\alpha}) + \frac{1-\alpha}{\alpha} P^{(1)}(z_2) = 0$$

or,

$$(\alpha - 1)\{a(z_2) - P^{(1)}(z_2)\} = 0$$

or,

$$a(z_2) - P^{(1)}(z_2) = 0.$$

Clearly $a(z) - P^{(1)}(z) \not\equiv 0$, because $deg(P^{(1)}(z))$ is less than $deg(a)$.

$$\begin{aligned}
N(r, a; f) &\leq N_A(r, a; f) + N_B(r, a; f^{(1)}) + N_D(r, a; f) + S(r, f) \\
&\leq N(r, 0; a - P^{(1)}) + S(r, f) \\
&= S(r, f).
\end{aligned}$$

Then from the hypothesis $T(r, f) = S(r, f)$, a contradiction.
Therefore $\alpha = 1$, so $f = \lambda e^z$.
Again

$$\begin{aligned}
L &= \sum_{k=1}^{n} a_k f^{(k)} \\
&= (\sum_{k=1}^{n} a_k \alpha^k)\lambda e^{\alpha z} \\
&= \lambda e^z.
\end{aligned}$$

Therefore $f = L = \lambda e^z$.
This completes the lemma.

$\square$

3 Proof of the Main Theorem

Proof. First we claim that f is a transcendental entire function.
If f is a polynomial, then
$T(r, f) = O(\log r)$ and $N_A(r, a; f) + N_A(r, a; f^{(1)}) = O(\log r)$.
Then from the hypothesis we get $O(\log r) = O(\log T(r, f)) = S(r, f)$, which implies $T(r, f) = S(r, f)$, a contradiction. Therefore $A = \emptyset$.
Similarly $N_B(r, a; f^{(1)}) = S(r, f)$ implies $B = \emptyset$.
Therefore $E(a; f) = E(a; f^{(1)})$ and $\overline{E}(a; f^{(1)}) \subset \overline{E}(a; L^{(p)}) \cap \overline{E}(a; L^{(q)})$.
Let $deg(f) = m$ and $deg(a) = r$. If $m \geq r + 1$ then $deg(f - a) = m$ and $deg(f^{(1)} - a) \leq m - 1$ which contradicts that $E(a, f) = E(a, f^{(1)})$.
If $m \leq r - 1$, then $deg(f - a) = deg(f^{(1)} - a) = r$. Since $E(a, f) = E(a, f^{(1)})$, $(f - a) = t(f^{(1)} - a)$, where $t(\neq 0)$ is a constant.
If $t = 1$, then $f = f^{(1)}$, which is a contradiction because f is a polynomial.
If $t \neq 1$ then $tf^{(1)} - f \equiv (t-1)a$, which is impossible because $deg((t-1)a) = r$ and $deg(tf^{(1)} - f) = m$ and $m < r$. Therefore our claim " f is transcendental entire function " is established. Now we prove the result into two cases.
Case 1. Let $f \equiv L^{(p)}$. Then

$$\begin{aligned} m(r, a; f) &= m(r, \frac{a}{f-a}\frac{1}{a}) \\ &\leq m(r, \frac{a}{f-a}) + S(r, f) \\ &= m(r, \frac{a}{f-a} + 1 - 1) + S(r, f) \\ &\leq m(r, \frac{a}{f-a} + 1) + S(r, f) \\ &\leq m(r, \frac{f}{f-a}) + S(r, f) \\ &= m(r, \frac{L^{(p)}}{f-a}) + S(r, f), \end{aligned} \tag{15}$$

since $p \geq deg(a)$, by Lemma of logarithmic derivative, $m(r, \frac{L^{(p)}}{f-a}) = S(r, f)$. So from (15) $m(r, a; f) = S(r, f)$. Therefore by Lemma 5, $f = L = \lambda e^z, \lambda(\neq 0)$ is a constant.

Case 2. Let $f \not\equiv L^{(p)}$. This case can be divided into two subcases.

Subcase 2.1. Let $f^{(1)} \not\equiv L^{(p)}$.

Since $a-a^{(1)} = (f^{(1)}-a^{(1)})-(f^{(1)}-a)$, a common zero of $f-a$ and $f^{(1)}-a$ of multiplicity $s(\geq 2)$ is a zero of $a-a^{(1)}$ with multiplicity $s-1(\geq 1)$.
Therefore $N_{(2}(r,a;f^{(1)} \mid f=a) \leq 2N(r,0;a-a^{(1)}) = S(r,f)$,

where $N_{(2}(r,a;f^{(1)} \mid f=a)$ denotes the counting function (counted with multiplicities) of those multiple zeros of $f^{(1)}-a$ which are also zeros of $f-a$.

Now

$$\begin{aligned} N_{(2}(r,a;f^{(1)}) &\leq N_A(r,a;f^{(1)}) + N_B(r,a;f^{(1)}) + N_{(2}(r,a;f^{(1)} \mid f=a) + S(r,f) \\ &= S(r,f). \end{aligned} \tag{16}$$

Using (16) and from the hypothesis we get

$$\begin{aligned} N(r,a;f^{(1)}) &\leq N_B(r,a;f^{(1)}) + N(r,\frac{a-L^{(p)}(a)}{a-a^{(1)}};\frac{L^{(p)}(f)-L^{(p)}(a)}{f^{(1)}-a^{(1)}}) + S(r,f) \\ &\leq T(r,\frac{a-L^{(p)}(a)}{a-a^{(1)}};\frac{L^{(p)}(f)-L^{(p)}(a)}{f^{(1)}-a^{(1)}}) + S(r,f) \\ &= N(r,\frac{L^{(p)}(f)-L^{(p)}(a)}{f^{(1)}-a^{(1)}}) + S(r,f) \\ &\leq N(r,a^{(1)};f^{(1)}) + S(r,f). \end{aligned} \tag{17}$$

Again

$$\begin{aligned} m(r,a;f) &= m(r,\frac{f^{(1)}-a^{(1)}}{f-a}\frac{1}{f^{(1)}-a^{(1)}}) \\ &\leq m(r,a^{(1)};f^{(1)}) + S(r,f) \\ &= T(r,f^{(1)}) - N(r,a^{(1)};f^{(1)}) + S(r,f) \\ &= m(r,f^{(1)}) - N(r,a^{(1)};f^{(1)}) + S(r,f) \\ &\leq m(r,f) - N(r,a^{(1)};f^{(1)}) + S(r,f) \\ &= T(r,f) - N(r,a^{(1)};f^{(1)}) + S(r,f), \end{aligned}$$

i.e

$$N(r,a^{(1)};f^{(1)}) \leq N(r,a;f) + S(r,f).$$

So from (17) we get

$$N(r,a;f^{(1)}) \leq N(r,a;f) + S(r,f). \tag{18}$$

Also

$$\begin{aligned} N(r,a;f) &\leq N_A(r,a;f) + N(r,a;f \mid f^{(1)}=a) \\ &\leq N(r,a;f^{(1)}) + S(r,f). \end{aligned} \tag{19}$$

From (18) and (19) we get

$$N(r, a; f^{(1)}) \quad = \quad N(r, a; f) + S(r, f). \tag{20}$$

Let

$h = \dfrac{(a - a^{(1)})(L^{(p)}(f) - L^{(p)}(a)) - (a - L^{(p)}(a))(f^{(1)} - a^{(1)})}{f - a}$, which is defined in Lemma 2.4.

Clearly $T(r, h) = S(r, h)$.

Now

$$\begin{aligned}
T(r, f) &= m(r, f) \\
&= m(r, a + \frac{1}{h}\{(a - a^{(1)})(L^{(p)}(f) - L^{(p)}(a)) - (a - L^{(p)})(f^{(1)} - a^{(1)})\} \\
&\leq m(r, (a - a^{(1)})L^{(p)}(f) - (a - L^{(p)})f^{(1)}) + S(r, f) \\
&\leq m(r, f^{(1)}) + S(r, f) \\
&= T(r, f^{(1)}) + S(r, f) \\
&= m(r, f^{(1)}) + S(r, f) \\
&\leq m(r, f) + S(r, f) \\
&= T(r, f) + S(r, f).
\end{aligned}$$

Therefore

$$T(r, f^{(1)}) \quad = \quad T(r, f) + S(r, f). \tag{21}$$

If h is transcendental, then by Lemma 2.4, $m(r, a; f^{(1)}) = S(r, f)$ and from (20) and (21) $m(r, a; f) = S(r, f)$. So from Lemma 2.5, $f = L = \lambda e^z$, $\lambda(\neq 0)$, is a constant.

If h is rational, then by Lemma 2.2 we see that f is of finite order. So by Lemma 2.1 we get $m(r, a; f) = S(r, f)$.

Therefore from Lemma 2.5, $f = L = \lambda e^z$, $\lambda(\neq 0)$ is a constant.

Subcase 2.2. Let $f^{(1)} \equiv L^{(p)}$. Now

$$\begin{aligned} m(r,a;f) &= m(r, \frac{a^{(1)}}{f-a}\frac{1}{a^{(1)}}) \\ &\leq m(r, \frac{a^{(1)}}{f-a}) + S(r,f) \\ &= m(r, \frac{f^{(1)} - (f^{(1)} - a^{(1)})}{f-a} + S(r,f) \\ &\leq m(r, \frac{f^{(1)}}{f-a}) + S(r,f) \\ &= m(r, \frac{L^{(p)}}{f-a}) + S(r,f). \end{aligned} \tag{22}$$

Since $p \geq deg(a)$, by Lemma of logarithmic derivative, $m(r, \frac{L^{(p)}}{f-a}) = S(r,f)$, so from (22) $m(r,a;f) = S(r,f)$.

Therefore from Lemma 2.5, we get $f = L = \lambda e^z$, $\lambda(\neq 0)$, is a constant.

This completes the proof of the Main Theorem.

□

4 Conclusions

Finally we arrive at the conclusion that a non-constant entire function sharing a polynomial with its linear differential polynomial with some conditions defined in Theorem (1.1) belongs to the class of functions $\mathfrak{F} = \{\lambda e^z : \lambda \in \mathbb{C} \setminus \{0\}\}$.

5 Acknowledgements

Authors are thankful to all the referees for their remarkable suggestions and all the authors of various papers and books which have been consulted to built the work.

References

W. K. Hayman. *Meromorphic Functions*. The Clarendon Press, Oxford, 1964.

G. Jank, E. Mues, and L. Volkmann. Meromorphe funktionen, die mit ihrer ersten und zweiten ableitung einen endlichen wert teilen. *Complex Var. Theory Appl.*, 6:51–71, 1986.

I. Lahiri and I. Kaish. An entire function sharing a polynomial with its derivatives. *Boll. Unione Mat. Ital.*, 10:229–240, 2017.

I. Lain. *Nevanlinna theory and complex differential equations*. Watter de Gruyter, New York, 1993.

P. Li. Entire functions that share one value with their linear differential polynomials. *Kodai Math. j.*, 22:446–457, 1999.

E. Mues and N. Steinmetz. Meromorphe functionen, die mit ihrer ableitung werte teilen. *Manuscripta Math.*, 29:195–206, 1979.

L. A. Rubel and C. C. Yang. Values shared by an entire function and its derivative. *Lecture Notes in Math.(Springer)*, 599:101–103, 1977.

C. C. Yang and H. X. Yi. *Uniqueness Theory of Meromorphic Functions*. Science Press and Kluwer Academic Publishers, New York, 2003.

H. Zhong. Entire functions that share one value with their derivatives. *Kodai Math. J.*, 18:250–259, 1995.

Ratio Mathematica Volume 39, 2020, pp. 229-236

On homomorphism of fuzzy multigroups

J. A. Awolola*
M. A. Ibrahim†

Abstract

In this paper, the homomorphism of fuzzy multigroups is briefly delineated and some related results are shown. In particular, we consider the corresponding isomorphism theorems of fuzzy multigroups.

Keywords: fuzzy multiset, fuzzy multigroup, homomorphism of fuzzy multigroups.
2010 AMS subject classifications: 54A40, 03E72, 20N25, 06D72. [1]

*Department of Mathematics/Statistics/Computer Science, University of Agriculture, P.M.B. 2373, Makurdi, Nigeria; awolola.johnson@uam.edu.ng.
†Department of Mathematics, Ahmadu Bello University, Zaria, Nigeria; amibrahim@abu.edu.ng.
[1]Received on November 2nd, 2020. Accepted on December 17rd, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.553. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

Since the inception of the theory fuzzy multisets introduced by Yager (1986), the subject has become an interesting area for researchers in algebra. The foundation of algebraic structures of fuzzy multisets was laid by Shinoj et al. (2015); Ibrahim and Awolola (2015) discussed further some new results which will bring new openings and development of fuzzy multigroup concept. Some group concepts like subgroups, abelian groups, normal subgroups and direct product of groups have been established (Ejegwa, 2018a,b,d, 2019). The idea of homomorphism of fuzzy multigroups and their alpha-cuts have also been discussed (Ejegwa, 2018c, 2020).

In this paper, more results on homomorphism of fuzzy multigroups are established and the corresponding isomorphism theorems of fuzzy multigroups which analogously exist in group setting are discussed.

2 Preliminaries

We recall here some basic definitions and results used in the sequel. We refer the reader to (Miyamoto, 2001; Shinoj et al., 2015; Ibrahim and Awolola, 2015).

Definition 2.1. (Miyamoto, 2001) Let X be a nonempty set. A *fuzzy multiset* U over X is characterized by count membership function $CM_U : X \to [0,1]$ (giving a multiset of the unit interval $[0,1]$). An expedient notation for a fuzzy multiset U over X is $U = \{(CM_U(a)/a) \mid a \in X\}$ with $CM_U(a) = \{\mu_U^1(a), \mu_U^2(a), ..., \mu_U^m(a), ...\}$, where $\mu_U^1(a), \mu_U^2(a), ..., \mu_U^m(x), ... \in [0,1]$ such that $(\mu_U^1(x) \geq \mu_U^2(a) \geq, ..., \geq \mu_U^m(a), ...)$.

If the fuzzy multiset U is finite, then $CM_U(a) = \{\mu_U^1(a), \mu_U^2(a), ..., \mu_U^m(a)\}$,

where $\mu_U^1(a), \mu_U^2(a), ..., \mu_U^m(a) \in [0,1]$ such that $\mu_U^1(a) \geq \mu_U^2(a) \geq, ..., \geq \mu_U^m(a)$.

The set of all fuzzy multisets over X is denoted by $FMS(X)$. Throughout this paper fuzzy multisets are considered finite.

The usual set operations can be carried over to fuzzy multisets. For instance, let $U, V \in FMS(X)$, then

$$U \subseteq V \Longleftrightarrow CM_U(a) \leq CM_V(a), \forall\, a \in X,$$
$$U \cap V = \{CM_U(a) \wedge CM_V(a)/a \mid a \in X\},$$
$$U \cup V = \{CM_U(a) \vee CM_V(a)/x \mid a \in X\}.$$

Definition 2.2 (Shinoj et al., 2015) Let P and Q be two nonempty sets such that $\varphi : P \to Q$ is a mapping. Consider the fuzzy multisets $U \in FMS(P)$ and $V \in FMS(Q)$. Then,

(i) the image of U under φ is denoted by $\varphi(U)$ has the count membership function

$$CM_{\varphi(U)}(b) = \begin{cases} \bigvee_{\varphi(a)=b} CM_U(a), & \varphi^{-1}(b) \neq \emptyset \\ 0, & \varphi^{-1}(b) = \emptyset \end{cases}$$

(ii) the inverse image of V under φ denoted by $\varphi^{-1}(V)$ has the count membership function $CM_{\varphi^{-1}(V)}(a) = CM_V(\varphi(a))$.

Definition 2.4 (Shinoj et al., 2015) Let X be a group. A fuzzy multiset U over X is called a fuzzy multigroup if

(i) $CM_U(ab) \geq CM_U(a) \bigwedge CM_U(b), \ \forall\, a, b \in X$, and

(ii) $CM_U(a^{-1}) = CM_U(a), \ \forall\, a \in X$.

The immediate consequence is that $CM_U(e) \geq CM_U(a)\, \forall\, a \in X$, where e is the identity element of X. The set all fuzzy multigroups is denoted by $FMG(X)$. The next definition can be found in Shinoj et al. (2015) .

Definition 2.5 Let $U \in FMG(X)$. Then U is called an abelian fuzzy multigroup over X if $CM_U(ab) = CM_U(ba), \ \forall\, a, b \in X$. The set $AFMG(X)$ is the set of all abelian fuzzy multigroups over X.

Definition 2.6 Let $U \in FMS(X)$. Then $U_* = \{x \in X \mid C_U(a) = C_U(e)\}$ **Remark 2.1** For a fuzzy multigroup over a group X, U_* is a group, certainly a subgroup of X Shinoj et al. (2015).

Proposition 2.1 (Ibrahim and Awolola, 2015) Let $U \in FMG(X)$, then $xU = yU \Longleftrightarrow x^{-1}y \in U_*$.

The following propositions are shown in (Ibrahim and Awolola, 2015) .

Proposition 2.2 Let $U \in FMG(X)$. Then the following assertions are equivalent:

(i) $CM_U(ab) = CM_U(ba), \ \forall\, a, b \in X$,

(ii) $CM_U(aba^{-1}) = CM_U(b), \ \forall \, a, b \in X,$

(iii) $CM_U(aba^{-1}) \geq CM_A(b), \ \forall \, a, b \in X,$

(iv) $CM_U(aba^{-1}) \leq CM_U(b), \ \forall \, a, b \in X.$

Proposition 2.3 Let $U \in FMG(X)$. Then $CM_U(ab^{-1}) = CM_U(e)$ implies $CM_U(a) = CM_U(b)$.

As to the converse problem whether $CM_U(a) = CM_U(b)$ implies $CM_U(ab^{-1}) = CM_U(e)$, we give a counter example. Let $X = \{1, s, t, r\}$ be a klein's 4-group and $U = \{(1, 0.7, 0.6, 0.5, 0.5)/1, (0.6, 0.4, 0.2)/s\}$. We see that U is an abelian fuzzy multigroup over X. Then, while $CM_U(t) = CM_U(r) = 0$, we have $CM_U(tr^{-1}) = CM_U(tr) = CM_U(s) = (0.6, 0.4, 0.2) \neq (1, 0.7, 0.6, 0.5, 0.5) = CM_U(1)$. Thus the converse problem above does not hold.

3 Main Results

Proposition 3.1 Let X be a group such that $\varphi : X \to X$ is an automorphism. If $U \in FMG(X)$, then $\varphi(U) = U$ if and only if $\varphi^{-1}(U) = U$.

Proof. Let $a \in X$. Then $\varphi(a) = a$.
Now $CM_{\varphi^{-1}(U)}(a) = CM_U(\varphi(a)) = CM_U(a)$
$\Longrightarrow \varphi^{-1}(U) = U$
Conversely, let $\varphi^{-1}(U) = U$. Since φ is an automorphism, then

$$\begin{aligned} CM_{\varphi(U)}(a) &= \bigvee\{CM_U(a^{'}) \mid a^{'} \in X, \ \varphi(a^{'}) = \varphi(a)\} \\ &= CM_U(\varphi(a)) \\ &= CU_{(\varphi^{-1}(U))}(a) \\ &= CM_U(a) \end{aligned}$$

Hence, the proof.

Proposition 3.2 Let $\varphi : X \to Y$ be a homomorphism of groups such that $U, V \in FMG(Y)$. If U is a constant on $Ker\varphi$, then $\varphi^{-1}(\varphi(U)) = U$.

Proof. Let $\varphi(a) = b$. Then we have
$CM_{\varphi^{-1}(\varphi(U))}(a) = CM_{\varphi(U)}\varphi(a) = CM_{\varphi(U)}(b) = \bigvee\{CM_U(a) \mid a \in X, \ \varphi(a) =$

$b\}$. Since $\varphi(a^{-1}c) = \varphi(a^{-1})\varphi(c) = (\varphi(a))^{-1}\varphi(c) = b^{-1}b = e^{'}, \ \forall\, c \in X$, such that $\varphi(c) = b$, which implies that $a^{-1}c \in Ker\varphi$. Moreover, since U is constant on $Ker\varphi$, then $CM_U(a^{-1}c) = CM_U(e)$. Therefore, $CM_U(a) = CM_U(c)$. This completes the proof.

Proposition 3.3 Let $U \in AFMG(X)$ such that a map $\varphi : X \to X/U$ is defined by $\varphi(a) = aU$. Then φ is a homomorphism with $Ker\varphi = \{a \in X \mid CM_U(a) = CM_U(e)\}$.

Proof. Clearly, φ is a homomorphism. Also,

$$\begin{aligned}
Ker\varphi &= \{a \in X : \varphi(a) = eU\} \\
&= \{a \in X : aU = eU\} \\
&= \{a \in X : CM_U(a^{-1}b) = CM_U(b) \ \forall\, b \in X\} \\
&= \{a \in X : CM_U(a^{-1}) = CM_U(e)\} \\
&= \{a \in X : CM_U(a) = CM_H(e)\} = U_*
\end{aligned}$$

Proposition 3.4 Let $\varphi : X \to Y$ be an epimorphism of groups and $U \in AFMG(X)$, then $X/U_* \cong Y$.

proof. Define $\Psi : X/U_* \to Y$ by $\Psi(xU_*) = \varphi(a) \ \forall\, a \in X$.
Let $aU = bU$ such that $CM_U(a^{-1}b) = CM_U(e)$. This implies that $a^{-1}b \in U_*$. It is easy to show that Ψ is well-defined, homomorphism and epimorphism.

Moreover, $\varphi(a) = \varphi(b)$

$$\begin{aligned}
&\Longrightarrow \varphi(a)^{-1}\varphi(b) = \varphi(e) \\
&\Longrightarrow \varphi(a^{-1})\varphi(b) = \varphi(a^{-1}b) = \varphi(e) \\
&\Longrightarrow a^{-1}b \in U_* \\
&\Longrightarrow CM_U(a^{-1}b) = CM_U(e) \\
&\Longrightarrow aU = bU
\end{aligned}$$

This shows that Ψ is an isomorphism.

Proposition 3.5 If $U, V \in AFMG(X)$ with $CM_U(e) = CM_V(e)$, then $U_*V_*/V \cong U_*/U \cap V$.

Proof. Clearly, for some $x \in U_*V_*$, $a = uv$ such that $u \in U_*$ and $v \in V_*$.
Define $\varphi : U_*V_*/V \to U_*/U \cap V$ by $\varphi(aV) = u(U \cap V)$.
If $aV = bV$ with $b = u_1v_1$, $u_1 \in U_*$ and $v_1 \in V_*$, then

$$\begin{aligned} CM_V(a^{-1}b) &= CM_V((uv)^{-1}u_1v_1) \\ &= CM_V(v^{-1}u^{-1}u_1v_1) \\ &= CM_V(u^{-1}u_1v^{-1}v_1) \\ &= CM_V(e). \end{aligned}$$

Hence, $CM_V(u^{-1}u_1) = CM_V(v^{-1}v_1) = CM_V(e)$. Thus,

$$\begin{aligned} CM_{U\cap V}(u^{-1}u_1) &= CM_U(u^{-1}u_1) \wedge CM_V(u^{-1}u_1) \\ &= CM_U(e) \wedge CM_V(e) \\ &= CM_{U\cap V}(e) \end{aligned}$$

That is, $u(U \cap V) = u_1(U \cap V)$. Therefore, φ is well-defined.

If $aV, bV \in U_*V_*/V$, then $ab = uvu_1v_1$. Since $U \in AFMG(X)$, then $CM_U(vu_1v_1) = CM_U(u_1) \implies vu_1v_1 \in U_*$.
Hence, $\varphi(aVbV) = \varphi(abV) = u(vu_1v_1)(U \cap V) = u(U \cap V)vu_1v_1(U \cap V)$ and

$$\begin{aligned} CM_{U\cap V}(u_1^{-1}(vu_1u_1)) &\geq CM_U(u_1^{-1}vu_1v_1) \wedge CM_V(u_1^{-1}vu_1v_1) \\ &= CM_U(u_1^{-1}(vu_1v_1)) \wedge CM_V(v(u_1^{-1}u_1v_1)) \\ &= CM_U(e) \wedge CM_V(e) \\ &= CM_{U\cap V}(e). \end{aligned}$$

Hence, $vu_1v_1(U \cap V) = u_1(U \cap V)$
That is, $\varphi(aVbV) = u(U \cap V)u_1(U \cap V) = \varphi(aV)\varphi(bV)$, and this shows that φ is a homomorphism. Undeniably, it is also epimorphism.

Furthermore, if $a, b \in U_*V_*$ with $a = uv$ and $b = u_1v_1$, $u, u_1 \in U_*$ and $v, v_1 \in V_*$ and $u(U \cap V) = u_1(U \cap V)$, then $CM_{U\cap V}(u^{-1}u_1) = CM_{U\cap V}(e)$
That is, $CM_U(u^{-1}u_1) \wedge CM_V(u^{-1}u_1) = CM_U(e) \wedge CM_V(e)$.
However, $CM_U(e) = CM_V(e)$ and $CM_U(u^{-1}u_1) = CM_U(e)$
$\implies CM_V(u^{-1}u_1) = CM_V(e)$.

Therefore,

$$\begin{aligned} CM_V(a^{-1}b) &= CM_V((uv)^{-1}u_1v_1) \\ &= CM_V(v^{-1}u^{-1}u_1u_1) = CM_V(u^{-1}u_1v^{-1}v_1) \\ &\geq CM_V(u^{-1}u_1) \wedge CM_V(v^{-1}v_1) \\ &= CM_V(e) \wedge CM_V(e) = CM_V(e) \end{aligned}$$
$$\Longrightarrow CM_V(a^{-1}b) = CM_V(e)$$

Thus, $aV = bV$.
Hence, $U_*V_*/V \cong U_*/U \cap V$.

Proposition 3.6 Let $U, V \in AFMG(X)$ such that $U \subseteq V$ and $CM_U(e) = CM_V(e)$. Then $X/V \cong (X/U)/(V_*/U)$.

proof. Define $\varphi : X/U \to X/V$ by $\varphi(aU) = aV \ \forall\, a \in X$ such that $CM_U(a^{-1}b) = CM_U(e) = CM_V(e) \ \ \forall\ aU = bU$. Since $U \subseteq V$, we have $CM_V(a^{-1}b) \geq CM_U(a^{-1}b) = CM_V(e)$ and thus $CM_V(a^{-1}b) = CM_V(e)$, that is, $aV = bV$, which implies that φ is well-defined. It is homomorphism and epimorphism too.

Moreover,

$$\begin{aligned} Ker\varphi &= \{aU \in X/U : \varphi(aU) = eV\} \\ &= \{aU \in X/U : aV = eV\} \\ &= \{aU \in X/U : CM_V(a) = CM_V(e)\} \\ &= \{aU \in X/U : a \in V_*\} = V_*/U. \end{aligned}$$

Thus, $Ker\varphi = V_*/U$ and so $X/V \cong (X/U)/(V_*/U)$.

References

P.A. Ejegwa. On fuzzy multigroups and fuzzy submultigroups. *Journal of Fuzzy Mathematics*, 26(3):641—-654, 2018a.

P.A. Ejegwa. On abelian fuzzy multigroups. *Journal of Fuzzy Mathematics*, 26 (3):655—-668, 2018b.

P.A. Ejegwa. Homomorphism of fuzzy multigroups. *Applications and Applied Mathematics*, 13(1):114—-129, 2018c.

P.A. Ejegwa. On normal fuzzy submultigroups of a fuzzy multigroup. *Theory and Applications of Mathematics and Computer Science*, 8(1):64—-92, 2018d.

P.A. Ejegwa. Direct product of fuzzy multigroups. *Journal of New Theory*, 28: 62—-73, 2019.

P.A. Ejegwa. On alpha-cuts homomorphism of fuzzy multigroups. *Annals of Fuzzy Mathematics and Informatics*, 19(1):73—-87, 2020.

A.M. Ibrahim and J.A. Awolola. Fuzzification of some results on multigroups. *Journal of the Nigerian Association of Mathematical Physics*, 32:119–124, 2015.

S. Miyamoto. Fuzzy multisets and their generalizations, in: C.s. calude, et al. (eds.). *Multiset Processing, Lecture Notes in Computer Science, Springer, Berlin*, 2235:225—-235, 2001.

T.K. Shinoj, B. Anagha, and J.J. Sunil. On some algebraic structures of fuzzy multisets. *Annals of Fuzzy Mathematics and Informatics*, 9:77–90, 2015.

R.R. Yager. On the theory of bags. *International Journal of General Systems*, 13: 23–37, 1986.

Ratio Mathematica Volume 39, 2020, pp. 237-252

On a class of sets between a-open sets and gδ-open sets

Jagadeesh B.Toranagatti*

Abstract

In this paper, a new class of sets called Da-open sets are introduced and investigated with the help of gδ-open and δ-closed sets. Relationships between this new class and other related classes of sets are established and as an application Da-continuous and almost Da-continuous functions have been defined to study its properties in terms of Da-open sets. Finally, some properties of Da-closed graph and (D,a)-closed graph are investigated.
Keywords:a-open set,δ-open set,gδ-open,Da-open set,Da-closed set.
2010 AMS subject classifications: 54A05, 54A10.[1]

*Department of Mathematics, Karnatak University's Karnatak College, Dharwad - 580001, Karnataka, India; jagadeeshbt2000@gmail.com

[1]Received on August 30th, 2020. Accepted on December 17th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.535. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

The concept of generalized open sets introduced by Levine[Levine, 1970] plays a significant role in General Topology. The study of generalized open sets and its properties found to be useful in computer science and digital topology[Khalimsky et al., 1990, Kovalevsky, 1994, Smyth, 1995]. Since Professor El- Naschie has recently shown in [El Naschie, 1998, 2000, 2005] that the notion of fuzzy topology may be relevant to quantum particle physics in connection with string theory and ϵ^{∞} theory.So,the fuzzy topological version of the notions and results introduced in this paper are very important. Recently, Ekici [Ekici, 2008] introduced the notion of a-open sets as a continuation of research done by Velicko [Velicko, 1968] on the notion of δ-open sets.Dontchev et al., introduced gδ-closed sets and gδ-continuity.In this paper,new generalizations of a-open sets by using $g\delta$-open and δ-closed sets called Da-open sets are presented. Also Da-continuous functions,almost Da-continuous functions,Da-closed graphs and (D,a)-closed graphs have been defined to study its properties in terms of Da-open sets.

2 Prerequisites, Definitions and Theorems

In what follows,spaces always mean topological spaces on which no separation axioms are assumed unless explicitly stated and f:(X,τ) $\rightarrow$ (Y,η) or simply f:X $\rightarrow$Y denotes a function f of a space (X,τ) into a space (Y,η). The δ-closure of a subset A of X is the intersection of all δ-closed sets containing A and is denoted by Cl$_\delta$(A).

Definition 2.1. *In (X,τ),let N $\subset$ X.Then N is called:*
(i)regular closed[Stone, 1937] (resp.,a-closed[Ekici, 2008], δ-preclosed[Raychaudhuri and Mukherjee, 1993], e^*-closed[Ekici, 2009], δ-semiclosed[Park et al., 1997], β-closed[Abd El-Monsef, 1983], semiclosed[Levine, 1963], preclosed[Mashhour, 1982]) if N = Cl(Int(N)) (resp., Cl(Int(Cl$_\delta$(N))) $\subset$ N, Cl(Int$_\delta$(N)) $\subset$ N, Int(Cl(Int$_\delta$(N)) $\subset$ N, Int(Cl$_\delta$(N)) $\subset$ N, Int(Cl(Int(N)) $\subset$ N, Int(Cl(N)) $\subset$ N, Cl(Int(N)) $\subset$ N).
(ii) δ-closed [Velicko, 1968] if N = Cl$_\delta(N)$
where Cl$_\delta(N)$ = $\{p\in X$:Int(Cl(O))$\cap N\neq\phi$,O$\in\tau$ and $p\in O\}$.
(iii)generalized δ-closed (briefly,gδ-closed)[Dontchev et al., 2000] if Cl(N)) $\subset$ G whenever N $\subset$ G and G is δ-open in X.
(iv)generalized closed (briefly,g-closed)[Levine, 1970] if Cl(N)) $\subset$ G whenever N $\subset$ G and G is open in X.
The complements of the above mentioned closed sets are their respective open sets.

The set of all regular open (resp.,δ-open, β-open, δ-preopen, preopen, semiopen, δ-semiopen,e^*-open,gδ-open and a-open) sets of (X,τ) is denoted by RO(X) (resp.δO(X), βO(X), δPO(X), PO(X), SO(X), δSO(X), e^*O(X), GδO(X) and aO(X)).

The a-closure[Ekici, 2008](resp, gδ-closure,δ-closure) of a set N is the intersection of all a-closed(resp, gδ-closed,δ-closed) sets containing N and is denoted by a-Cl(N) (resp., $\mathrm{Cl}_{g\delta}$(N),Cl_{δ}(N)). The a-interior[Ekici, 2008](resp,gδ-interior,δ-interior) of a set N is the union of all a-open(resp, gδ-open,δ-open) sets contained in M and is denoted by a-Int(M)(resp, $\mathrm{Int}_{g\delta}$(M),Int_{δ}(M))

Definition 2.2. *[Ekici, 2005] A topological space (X,τ) is said to be:*
(1) r-T_1 if for each pair of distinct points x and y of X, there exist regular open sets U and V such that $x \in U$, $y \notin U$ and $x \notin V$, $y \in V$.
(2) r-T_2 if for each pair of distinct points x and y of X, there exist regular open sets U and V such that $x \in U$, $y \in V$ and $U \cap V = \phi$.

Theorem 2.1. *Let C and D be subsets of a topological space (X,τ).Then*
(i)If C is gδ-closed,then $Cl_{g\delta}(C) = C$.
(ii) If $C \subset D$,then $Cl_{g\delta}(C) \subset Cl_{g\delta}(D)$.
(iv) $x \in Cl_{g\delta}(C)$ if and only if for each gδ-open set O containing x,$O \cap C \neq \phi$,
(v)$Cl_{g\delta}(C) \cup Cl_{g\delta}(D) \subset Cl_{g\delta}(A \cup D)$.
(vi)$Cl_{g\delta}(C \cap D) \subseteq Cl_{g\delta}(C) \cap Cl_{g\delta}(D)$.

3 Da-Open Sets.

Definition 3.1. *A subset M of a topological space (X,τ) is said to be:*
(1) Da-open if $M \subset Int_{g\delta}(Cl_{\delta}(Int_{g\delta}(M))$,
(2) Da-closed if $Cl_{g\delta}(Int_{\delta}(Cl_{g\delta}(M)) \subset M$.
The collection of all Da-open(resp,Da-closed) sets in (X,τ) is denoted by DaO(X) (resp,DaC(X)).

Theorem 3.1. *Let (X,τ) be a space.Then for any $N \subset X$,*
(i) $N \in \delta O(X)$ implies $N \in aO(X)$[Ekici, 2008].
(ii) $N \in \delta O(X)$ implies $N \in G\delta O(X)$[Dontchev et al., 2000].
(iii)$N \in GO(X)$ implies $N \in G\delta O(X)$[Dontchev et al., 2000].
(iv) $N \in aO(X)$ implies $N \in DaO(X)$.
(v) $N \in G\delta O(X)$ implies $N \in DaO(X)$.
Proof: *(iv) Since $\delta O(X) \subset G\delta O(X)$, $Int_{\delta}(N) \subset Int_{g\delta}(N)$.*
Now,let $N \in aO(X)$, then $N \subset Int(Cl(Int_{\delta}(N))$. Therefore,
$N \subset Int(Cl(Int_{\delta}(N))=Int_{\delta}(Cl(Int_{\delta}(N)) \subset Int_{g\delta}(Cl_{\delta}(Int_{g\delta}(N))$. Hence $N \in DaO(X)$.
(v) Suppose N is gδ-open. Then $Int_{g\delta}(N)=N$.

Therefore, $Int_{g\delta}(N) \subset Cl_\delta(Int_{g\delta}(N)$*.Then*
$N = Int_{g\delta}(N) = Int_{g\delta}(Int_{g\delta}(N)) \subset Int_{g\delta}(Cl_\delta(Int_{g\delta}(N))$*. Hence* $N \in DaO(X)$*.*

Remark 3.1. *The following diagram holds for any subset of a space (X,τ).*

$$\begin{array}{ccccc} \text{open set} & \longleftarrow & \delta\text{-open set} & \longrightarrow & a\text{-open set} \\ & & & & \swarrow \\ \downarrow & & \downarrow & \text{Da-open set} & \\ & & & \nearrow & \\ g\text{-open set} & \longrightarrow & g\delta\text{-open set} & & \end{array}$$

None of these implications is reversible

Example 3.1. *Let* $X=\{p,q,r,s\}$ *and* $\tau=\{X,\phi,\{p\},\{q\},\{p,q\},\{p,r\}.\{p,q,r\}\}$*,then*
$aO(X)=\{X,\phi,\{q\},\{p,r\},\{p,q,r\}\}$
$G\delta O(X)=\{X, \phi,\{p\},\{q\},\{r\},\{p,q\},\{p,r\}\{q,r\},\{p,q,r\}\}$.
$DaO(X)=\{X, \phi,\{p\},\{q\},\{r\},\{p,q\},\{p,r\},\{q,r\}\{p,q,r\}\{p,q,s\},\{q,r,s\}\}$.
Therefore, $\{q,r,s\} \in DaO(X)$ *but* $\{q,r,s\} \notin aO(X)$ *and* $\{q,r,s\} \notin g\delta O(X)$*.*

Lemma 3.1. *If there exists a* $M \in G\delta O(X)$ *such that* $M \subset N \subset Int_{g\delta}(Cl_\delta(M))$*,then N is Da-open.*
Proof: *Since M is gδ-open,* $Int_{\delta g}(M)=M$*. Therefore,*
$Int_{g\delta}(Cl_\delta(Int_{g\delta}(N)) \supset Int_{g\delta}(Cl_\delta(Int_{g\delta}(M)) = Int_{g\delta}(Cl_\delta(M)) \supset N$.
Hence N is Da-open.

Converse of the Lemma 3.1 is not true as shown in Example 3.1.

Example 3.2. *In Example 3.1,* $\{p,q,r\} \in DaO(X)$ *and* $\{p,r\} \in G\delta O(X)$ *but* $\{p,r\} \subseteq \{p,q,r\} \nsubseteq Int_{g\delta}(Cl_\delta(\{p,r\}))=\{p,r\}$ *.*

Lemma 3.2. *For a family* $\{ B_\lambda : \lambda \in \wedge\}$ *of subsets of a space (X,τ),the following hold:*
(1) $Cl_{g\delta}(\bigcap\{B_\lambda:\lambda\in\wedge\}) \subset \bigcap\{Cl_{g\delta}(B_\lambda):\lambda\in\wedge\}$.
(2) $Cl_{g\delta}(\bigcup\{V_\lambda:\lambda\in\wedge\}) \supset \bigcup\{Cl_{g\delta}(B_\lambda):\lambda\in\wedge\}$.
(3) $Cl_\delta(\bigcap\{B_\lambda:\lambda\in\wedge\}) \subset \bigcap \{Cl_\delta(B_\lambda):\lambda\in\wedge\}$.
(4) $Cl_\delta(\bigcup\{B_\lambda:\lambda\in\wedge\}) \supset \bigcup\{Cl_\delta(B_\lambda):\lambda\in\wedge\}$

Theorem 3.2. *If* $\{G_\alpha:\lambda\in\wedge\}$ *is a collection of Da-open sets in a space (X,τ),then* $\bigcup_{\alpha\in\wedge} G_\alpha$ *is a Da-open set in (X,τ) :*
Proof: *Since each* G_α*is Da-open,* $G_\alpha \subset Int_{g\delta}(Cl_\delta(Int_{g\delta}(G_\alpha))$ *for each* $\alpha\in\wedge$ *and hence* $\bigcup_{\alpha\in\wedge} G_\alpha \subset \bigcup_{\alpha\in\wedge} Int_{g\delta}(Cl_\delta(Int_{g\delta}(G_\alpha)) \subset Int_{g\delta}(Cl_\delta(Int_{g\delta}(\bigcup_{\alpha\in\wedge} G_\alpha))$*. Thus* $\bigcup_{\alpha\in\wedge} G_\alpha$ *is Da-open.*

Corolary 3.1. *If* $\{F_\alpha:\alpha\in\wedge\}$ *is a collection of Da-closed sets in a space (X,τ),then* $\bigcap\limits_{\alpha\in\wedge} F_\alpha$ *is a Da-closed set in (X,τ)*

Remark 3.2. *M and N* $\in DaO(X) \not\Rightarrow M\cap N \in DaO(X)$ *as seen from Example 3.1, where both* $M = \{q,r,s\}$ *and* $N = \{p,q,s\} \in DaO(X)$ *but* $M\cap N = \{q,s\} \notin DaO(X)$.

Corolary 3.2. *If* $M\in DaO(X)$ *and* $B\in aO(X)$*,then* $M\cup B\in DaO(X)$.
Proof: *Follows from Theorem 3.1(iv) and Theorem 3.2*

Corolary 3.3. *If* $M\in DaO(X)$ *and* $B\in G\delta O(X)$*,then* $M\cup B\in DaO(X)$.
Proof: *Follows from Theorem 3.1(v) and Theorem 3.2*

Definition 3.2. *In (X,τ),let* $M\subset X$.
(1)The Da-interior of M, denoted by $Int_a^D(M)$ *is defined as*
$Int_a^D(M)=\bigcup\{G:G\subseteq M \text{ and } M\in DaO(X)\}$;
(2)The Da-closure of M, denoted by $Cl_a^D(M)$ *is defined as*
$Cl_a^D(A)=\bigcap\{F:M\subseteq F \text{ and } F\in DaC(X)\}$.

Theorem 3.3. *In (X,τ),let M, N,F* $\subset X$*.Then:*
(1)$M\subset Cl_a^D(M)\subset aCl(M)$, $Cl_a^D(M)\subset Cl_{g\delta}(M)$.
(2) $Cl_a^D(M)$ *is a Da-closed set.*
(3) *If F is a Da-closed set, and* $F\supset M$*,then* $F\supset Cl_a^D(M)$.
i.e.,$Cl_a^D(M)$ *is the smallest Da-closed set containing M.*
(4)*M is Da-closed set if and only if* $Cl_a^D(M)=M$.
(5) $Cl_a^D(Cl_a^D(M)) = Cl_a^D(M)$.
(6)$M\subseteq N$ *implies* $Cl_a^D(M)\subseteq Cl_a^D(N)$.
(7)$p\in Cl_a^D(M)$ *if and only if for each Da-open set V containing p,*$V\cap M\neq\phi$.
(8) $Cl_a^D(M)\cup Cl_a^D(N)\subset Cl_a^D(M\cup N)$.
(9) $Cl_a^D(M\cap N)\subset Cl_a^D(M)\cap Cl_a^D(N)$.
Proof: *(1)It follows from Theorem 3.1(iv) and (v)*
(2)It follows from Definition 3.2 and Corollary 3.1
(3)Let F be a Da-closed set,containing M.$Cl_a^D(M)$ *is the intersection of Da-closed sets containing M, and F is one among these;hence* $F\supset Cl_a^D(M)$.
(4) Let M be Da-closed,then by Definition 3.2(2),$Cl_a^D(M)=M$.
Conversely,let $Cl_a^D(M)=M$*. Then by (2) above,M is Da-closed.*
(5)It follows from (2) and (4).
(6) Obvious.
(7)
$$\begin{aligned} p\notin Cl_a^D(M) &\Leftrightarrow (\exists\, G\in DaC(X))(M\subset G)(p\notin G)\\ &\Leftrightarrow (\exists\, G\in DaC(X))(M\subset G)(p\in G^c)\\ &\Leftrightarrow (\exists\, G^c\in DaO(X))(M\cap G^c=\phi)(p\in G^c)\\ &\Leftrightarrow (\exists\, G^c\in DaO(X,p))(M\cap G^c=\phi)\end{aligned}$$

i.e.,($\exists$ U(=G^c)$\in$ DaO(X,p))(M$\cap$U=ϕ)
(8) and (9) follows from (6).

Remark 3.3. *(1) $Cl_a^D(M) \cup Cl_a^D(N) \neq Cl_a^D(M \cup N)$, in general, as seen from Example 3.1 where $M = \{p\}$, $N = \{r\}$ and $M \cup N = \{p,r\}$.Then $Cl_a^D(M)=\{p\}$, $Cl_a^D(N)=\{r\}$,$Cl_a^D(M)\cup Cl_a^D(N)=\{p,r\}$ and $Cl_a^D(M\cup N)=\{p,r,s\}$;*
(2) $Cl_a^D(M \cap N)\neq Cl_a^D(M) \cap Cl_a^D(N)$, in general,as seen from Example 3.1 where,M = $\{p,q,r\}$, N = $\{s\}$ and $M\cap N = \phi$.Then $Cl_a^D(M) = X$, $Cl_a^D(N) = \{s\}$, $Cl_a^D(M)\cap Cl_a^D(N) = \{s\}$ and $Cl_a^D(M\cap N)=\phi$

Lemma 3.3. *In (X,τ),let M $\subset$ X.Then*
(1) $Cl_a^D(X\backslash M) = X\backslash Int_a^D(M)$,
(2) $Int_a^D(X\backslash M) = X\backslash Cl_a^D(M)$.

Theorem 3.4. *In (X,τ),let M,N,G $\subset$ X,*
(1)$aInt(M) \subseteq Int_a^D(M)\subseteq M$, $Int_{g\delta}(M)\subseteq Int_a^D(M)$.
(2) $Int_a^D(M)$ is a Da-open set.
(3) If G is a Da-open set, and $G \subset M$,then $G \subset Int_a^D(M)$.
i.e.,$Int_a^D(M)$ is the largest Da-open set contained in M.
(4)M is Da-open set if and only if $Int_a^D(M)=M$.
(5) $Int_a^D(Int_a^D(M)) = Int_a^D(M)$.
(6)$M \subseteq N$ implies $Int_a^D(M) \subseteq Int_a^D(N)$.
(7) $p \in Int_a^D(M)$ if and only if there exists Da-open set N containing p such that N $\subseteq$ M.
(8) $Int_a^D(M \cap N)\subseteq Int_a^D(M) \cap Int_a^D(N)$.
(9) $Int_a^D(M) \cup Int_a^D(N) \subseteq Int_a^D(M \cup N)$.
Proof:*Similar to the proof of Theorem 3.3*

Remark 3.4. *(8)$Int_a^D(M \cap N)\neq Int_a^D(M) \cap Int_a^D(N)$, in general, as seen from Example 3.1,where $M = \{p,q,s\}$, $N = \{q,r,s\}$ and $M \cap N = \{q,s\}$.Then $Int_a^D(M) = \{p,q,s\}$, $Int_a^D(N) = \{q,r,s\}$, $Int_a^D(M) \cap Int_a^D(N) = \{q,s\}$ and $Int_a^D(M\cap N) = \{q\}$.*
(9) $Int_a^D(M) \cup Int_a^D(N) \neq Int_a^D(M \cup N)$,in general, as seen from Example 3.1, where $M = \{p,q,r\}$, $N = \{s\}$ and $M \cup N = X$.Then $Int_a^D(M) = \{p,q,r\}$, $Int_a^D(N) = \phi$, $Int_a^D(M) \cup Int_a^D(N) = \{p,q,r\}$ and $Int_a^D(M \cup N) = X$.

Lemma 3.4. *In (X,τ),let M $\subset$ X. Then*
(1)M is Da-open if and only if $M = M \cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M))$.
(2)M is Da-closed if and only if $M = M\cup Cl_{g\delta}(Int_\delta(Cl_{g\delta}(M))$.

Proof:*(1) Let M be an Da-open. Then,*
$M\subseteq Int_{g\delta}(Cl_\delta(Int_{g\delta}(M))$ implies $M\cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M))=M$.
Conversely,let $M = M\cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M))$ implies $M \subset Int_{g\delta}(Cl_\delta(Int_{g\delta}(M))$.
(2)It follows from (1)

Lemma 3.5. *In (X,τ),let M $\subset$ X. Then*
(i)M $\cap$ $Int_{g\delta}$(Cl_δ($Int_{g\delta}$(M)) is Da-open
(ii)M$\cup$ $Cl_{g\delta}$(Int_δ($Cl_{g\delta}$(M)) is Da-closed.
Proof: *(i) $Int_{g\delta}(Cl_\delta(Int_{g\delta}(M \cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M)))))) = Int_{g\delta}(Cl_\delta(Int_{g\delta}(A) \cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M)))))$ = $Int_{g\delta}(Cl_\delta(Int_{g\delta}(M)))$. This implies that*
$M \cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M))) = M \cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M \cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M)))))) \subseteq Int_{g\delta}(Cl_\delta(Int_{g\delta}(M \cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M))))))$. Therefore $M \cap Int_{g\delta}(Cl_\delta(Int_{g\delta}(M)))$ is Da-open.
(ii) From (i) we have $X \backslash (M \cup Cl_{g\delta}(Int_\delta(Cl_{g\delta}(M))))$ = $(X \backslash M) \cap Cl_{g\delta}(Int_\delta(Cl_{g\delta}(X \backslash M)))$ is Da-open so that $M \cup Cl_{g\delta}(Int_\delta(Cl_{g\delta}(M)))$ is Da-closed.

Lemma 3.6. *In (X,τ),let M $\subset$ X. Then*
(i)Int_a^D(M)=M $\cap$ $Int_{g\delta}$(Cl_δ($Int_{g\delta}$(M)).
(ii)Cl_a^D(M)=M$\cup$ $Cl_{g\delta}$(Int_δ($Cl_{g\delta}$(M)).
Proof:*(i)Let N=Int_a^D(M),then N$\subset$M.Since N is Da-open,N$\subset$$Int_{g\delta}$($Cl_\delta$($Int_{g\delta}$(N)) $\subset$$Int_{g\delta}$($Cl_\delta$($Int_{g\delta}$(M)).Then N$\subset$M$\cap$$Int_{g\delta}$($Cl_\delta$($Int_{g\delta}$(M))$\subset$M.Therefore,by Lemma 3.5, it follows that M$\cap$$Int_{g\delta}$($Cl_\delta$($Int_{g\delta}$(M)) is a Da-open set contained in M. But Int_a^D(M) is the largest Da-open set contained in M it follows that*
M$\cap$$Int_{g\delta}$($Cl_\delta$($Int_{g\delta}$(M))$\subset$ Int_a^D(M)=N.Then N=M$\cap$$Int_{g\delta}$($Cl_\delta$($Int_{g\delta}$(M)).
Therefore,Int_a^D(M)=M $\cap$ $Int_{g\delta}$(Cl_δ($Int_{g\delta}$(M)).
(ii)It follows from (i)

4 Da-Continuous functions.

Definition 4.1. *A function f:(X,τ) $\to$ (Y,η) is said be a Da-continuous if for each p$\in$X and each N$\in$O(Y,f(p)), there exists M $\in$ DaO(X,p) such that f(M)$\subset$ N.*

Theorem 4.1. *For a function f:(X,τ) $\to$ (Y,η),the following are equivalent*
(1)f is Da-continuous;
(2)For each N$\in$O(Y),f^{-1}(V)$\in$DaO(X).
Proof:*(1)$\longrightarrow$(2)Let N$\in$O(Y) and p$\in$ f^{-1}(N). Since f(p) $\in$ N,then by(1),there exists M_p $\in$ DaO(X,p) such that f(M_p) $\subset$ N.It follows that*
f^{-1}(N)=$\cup${M_p: p$\in$$f^{-1}$(N)}$\in$DaO(X), by Theorem 3.2 .
(2)$\longrightarrow$(1) Let p $\in$ X and N $\in$O(Y,f(p)).Then,by (2),f^{-1}(N)$\in$DaO(X,p).
Take M = f^{-1}(N), then f(M) $\subset$ N.

Corolary 4.1. *A function f:(X,τ) $\to$ (Y,η) is Da-continuous if and only if f^{-1}(F)$\in$DaC(X) for each F$\in$C(Y).*

Remark 4.1. *The following implications hold for a function f:(X,τ) $\rightarrow$ (Y,η):*

$$\begin{tikzcd}
\textit{continuity} \arrow[dd] & \delta\textit{-continuity} \arrow[l] \arrow[r] \arrow[dd] & a\textit{-continuity} \arrow[d] \\
 & & Da\textit{-continuity} \\
g\textit{-continuity} \arrow[r] & g\delta\textit{-continuity} \arrow[ur] &
\end{tikzcd}$$

Example 4.1. *Consider (X,τ) as in Example 3.1 and η=$\{X,\phi,\{p\},\{q\},\{p,q\},\{p,q,r\}\}$. Define f:(X,$\sigma$)$\rightarrow$(X,$\eta$) by f(p)=s,f(q)=p,f(r)=q and f(s)=r.Then f is Da-continuous but neither a-continuous nor gδ-continuous since $\{p,q,r\}$ is open in (X,η), $f^{-1}(\{p,q,r\})$ = $\{q,r,s\}$$\in$DaO(X) but $\{q,r,s\}$$\notin$aO(X) and $\{q,r,s\}$$\notin$ gδO(X).*
The other Examples are shown in[3,5,21]

Theorem 4.2. *The following conditions are equivalent for a function f:(X,τ) $\rightarrow$ (Y,η):*
(1) f is Da-continuous;
(2) For each subset N of Y, $Cl_{g\delta}(Int_{\delta}(Cl_{g\delta}(f^{-1}(N))) \subset f^{-1}(Cl(N)$;
(3)For each subset N of Y, $f^{-1}(Int(N)) \subset Int_{g\delta}(Cl_{\delta}(Int_{g\delta}(f^{-1}(N))$;
(4)For each subset N of Y,$Cl_a^D(f^{-1}(N)) \subset f^{-1}(Cl(N))$;
(5)For each subset M of X,$f(Cl_a^D(M)) \subset Cl(f(M))$;
(6)For each subset N of Y, $f^{-1}(Int(N)) \subset Int_a^D(f^{-1}(N))$.
Proof: *(1)$\rightarrow$(2) Let N $\subset$ Y.Then by (1),$f^{-1}(Cl(N)) \in DaC(X)$ implies*
$f^{-1}(Cl(N)\supset Cl_{g\delta}(Int_{\delta}(Cl_{g\delta}(f^{-1}(Cl(N)))\supset Cl_{g\delta}(Int_{\delta}(Cl_{g\delta}(f^{-1}(N)))$.
(2)$\rightarrow$(3).Replace N by Y$\setminus$N in (2), we have
$Cl_{g\delta}(Int_{\delta}(Cl_{g\delta}(f^{-1}(Y\setminus N)))\subset f^{-1}(Cl(Y\setminus N)$, and therefore
$f^{-1}(Int(N)) \subset Int_{g\delta}(Cl_{\delta}(Int_{g\delta}(f^{-1}(N))$ for each subset N of Y.
(3)$\rightarrow$(1). Clear
(1)$\rightarrow$(4). Let N $\subset$ Y .Then by (1), $f^{-1}(Cl(N))\in DaC(X)$. Thus
$Cl_a^D(f^{-1}(N)) \subset Cl_a^D(f^{-1}(Cl(N))=f^{-1}(Cl(N)$ by Theorem 3.3(4).
(4)$\rightarrow$(1). Let N $\in$C(Y).Then by (4),
$Cl_a^D(f^{-1}(N)) \subset f^{-1}(Cl(N)=f^{-1}(N)$ implies $Cl_a^D(f^{-1}(N))=f^{-1}(N)$.
Then by Theorem 3.3(4), $f^{-1}(N) \in DaC(X)$.
(4)$\rightarrow$(5).Let M $\subset$ X.Then f(M) $\subset$ Y.By (4), we have
$f^{-1}(Cl(f(M))) \supset Cl_a^D(f^{-1}(f(M))) \supset Cl_a^D(M)$.
Therefore, $f(Cl_a^D(M)) \subset f(f^{-1}(Cl(f(M))) \subset Cl(f(M)$.
(5)$\rightarrow$(4).Let N $\subset$ Y and M=$f^{-1}(N) \subset X$.Then by (5),
$f(Cl_a^D(f^{-1}(N))) \subset Cl(f(f^{-1}(N)) \subset Cl(N)$ implies $Cl_a^D(f^{-1}(N)) \subset f^{-1}(Cl(N))$.
(4)$\rightarrow$(6).Replace N by Y $\setminus$N in (4), we get
$Cl_a^D(f^{-1}(Y\setminus N)) \subset f^{-1}(Cl(Y\setminus N))$ implies $Cl_a^D(X\setminus f^{-1}(N)) \subset f^{-1}(Y\setminus Int(N))$
Therefore,$f^{-1}(int(N)) \subset Int_a^D(f^{-1}(N))$ for each subset N of Y.

(6)→(1).Let $G \subset Y$ be open.Then $f^{-1}(G)=f^{-1}(Int(G)) \subset Int_a^D(f^{-1}(G)$ implies $Int_a^D(f^{-1}(G)=f^{-1}(G)$.So by Theorem 3.4(4),$f^{-1}(G) \in DaO(X)$.

Definition 4.2. *Two non-empty subsets A and B of a topological space (X,τ) are said to be Da-separated if there exist two Da-open sets G and H,such that $A \subset G, B \subset H$, $A \cap H=\phi$ and $B \cap G=\phi$.*

Definition 4.3. *Two non-empty subsets A and B of a topological space (X,τ) are said to be strongly Da-separated if there exist two Da-open sets U and V,such that $A \subset U, B \subset V$ and $U \cap V=\phi$.*

Definition 4.4. *A topological space (X,τ) is said to be*
(1) Da-T_2 if any two distinct points are strongly Da-separated in (X,τ)
(2) Da-T_1 if every pair of distinct points is Da-separated in (X,τ).

Remark 4.2. *The following implications are hold for a topological space (X,τ)*

$$\begin{array}{ccccc} a\text{-}T_2 & \longrightarrow & Da\text{-}T_2 & \longleftarrow & T_2 \\ \downarrow & & \downarrow & & \downarrow \\ a\text{-}T_1 & \longrightarrow & Da\text{-}T_1 & \longleftarrow & T_1 \end{array}$$

Theorem 4.3. *If an injective function $f:(X,\tau) \to (Y,\eta)$ is Da-continuous and (Y,η) is T_1, then (X,τ) is Da-T_1.*
Proof: *Let (Y,σ) be T_1 and $p,q \in X$ with $p \neq q$. Then there exist open subsets G, H in Y such that $f(p) \in G$, $f(q) \notin G$, $f(p) \notin H$ and $f(q) \in H$. Since f is Da-continuous, $f^{-1}(G)$ and $f^{-1}(H) \in DaO(X)$ such that $p \in f^{-1}(G)$, $q \notin f^{-1}(G)$, $p \notin f^{-1}(H)$ and $q \in f^{-1}(H)$. Hence,(X,σ) is Da-T_1 .*

Theorem 4.4. *If an injective function $f: (X,\tau) \to (Y,\eta)$ is Da-continuous and (Y,η) is T_2, then (X,τ) is Da-T_2.*
Proof: *Similar to the proof of Theorem 4.3*

Recall that for a function $f:(X,\tau) \to (Y,\eta)$, the subset $G_f=\{(x,f(x)):x \in X\} \subset X \times Y$ is said to be graph of f.

Definition 4.5. *A graph G_f of a function $f:(X,\tau) \to (Y,\eta)$ is said to be Da-closed if for each $(p,q) \notin G_f$, there exist $U \in DaO(X,p)$ and $V \in O(Y,q)$ such that $(U \times V) \cap G_f = \phi$.*

As a consequence of Definition 4.5 and the fact that for any subsets $C \subset X$ and $D \subset Y$, $(C \times D) \cap G_f=\phi$ if and only if $f(C) \cap D = \phi$,we have the following result.

Lemma 4.1. *For a graph G_f of a function $f:(X,\tau) \to (Y,\eta)$, the following properties are equivalent:*
(1)G_f is Da-closed in $X \times Y$;
(2)For each $(p,q) \notin G_f$, there exist $U \in DaO(X,p)$ and $V \in O(Y,q)$ such that $f(U) \cap V = \phi$.

Theorem 4.5. *If f:(X,τ) $\rightarrow$ (Y,η) is Da-continuous and (Y,η) is T_2 , then G_f is Da-closed in X$\times$Y.*
Proof: *Let (p,q) $\notin G_f$, f(p)$\neq$q. Since Y is T_2, there exist V,W $\in$O(Y) such that f(p)$\in$ V, q$\in$W and V$\cap$W=ϕ. Since f is Da-continuous, f^{-1}(V)$\in$DaC(X,p).Set U =f^{-1}(V), we have f(U)$\subset$ V. Therefore, f(U)$\cap$W=ϕ and G_f is Da-closed in X$\times$Y*

Theorem 4.6. *Let f:(X,τ) $\rightarrow$ (Y,η) have a Da-closed graph G_f. If f is injective, then (X,τ) is Da-T_1.*
Proof:*Let x_1,$x_2$$\in$X with $x_1$$\neq$$x_2$.Then f($x_1$)$\neq$f($x_2$) as f is injective So that (x_1,f(x_2)) $\notin G_f$.Thus there exist U$\in$DaO(X,x_1) and V$\in$O(Y,f(x_2)) such that f(U)$\cap$V = ϕ.Then f(x_2)$\notin$f(U) implies $x_2$$\notin$U and it follows that X is Da-T_1.*

Theorem 4.7. *Let f:(X,τ) $\rightarrow$ (Y,η) have a Da-closed graph G_f. If f is surjective, then (Y,η) is T_1.*
Proof:*Let y_1,$y_2$$\in$Y with $y_1$$\neq$$y_2$.Since f is surjective,f(x)=y_2 for some x$\in$X and (x,y_2)$\notin G_f$.By Lemma 4.1,there exist U$\in$DaO(X,x) and V$\in$O(Y,y_1) such that f(U)$\cap$V = ϕ.It follows that $y_2$$\notin$V.Hence Y is T_1.*

Theorem 4.8. *Let f:(X,τ) $\rightarrow$ (Y,η) have a Da-closed graph G_f. If f is surjective, then (Y,η) is Da-T_1.*
Proof:*Similar to the proof of Theorem 4.7*

Corolary 4.2. *Let f:(X,τ) $\rightarrow$ (Y,η) have a Da-closed graph G_f. If f is bijective, then both (X,τ) and (Y,η) are Da-T_1*
Proof:*Follows from Theorems 4.6 and 4.8*

Definition 4.6. *A graph G_f of a function f:(X,τ) $\rightarrow$ (Y,η) is said to be (D,a)-closed if for each (p,q) $\notin$ G_f, there exist U$\in$DaO(X,p) and V$\in$aO(Y,q) such that (U$\times$aCl(V))$\cap$ G_f = ϕ.*

Lemma 4.2. *For a graph G_f of a function f:(X,τ) $\rightarrow$ (Y,η), the following properties are equivalent:*
(1)G_f is Da-closed in X$\times$Y;
(2)For each (p,q) $\notin G_f$, there exist U$\in$DaO(X,p) and V$\in$aO(Y,q) such that f(U)$\cap$aCl(V)) = ϕ.

Theorem 4.9. *Let M $\subset$ X.Then x$\in$ a-Cl(M) if and only if G $\cap$ M $\neq$ Φ, for every a-open set G containing x.*
Proof:*Similar to the proof of Theorem 3.3(7)*

Theorem 4.10. *Let f:(X,τ) $\rightarrow$ (Y,η) have a (D,a)-closed graph G_f . If f is surjective, then (Y,η) is a-T_2(resp,a-T_1).*
Proof:*Let y_1,$y_2$$\in$Y with $y_1$$\neq$$y_2$.Since f surjective, f($x_1$)=$y_1$ $x_1$$\in$X and hence ($x_1$,$y_2$)$\notin G_f$. By Lemma 4.2,there exist E$\in$DaO(X,x_1) and F$\in$aO(Y,y_2) such that f(E)$\cap$ aCl(F) = ϕ. Now, $x_1$$\in$E implies f($x_1$)=$y_1$$\in$f(E) so that $y_1$$\notin$aCl(F).By Theorem 4.9,there exists D$\in$aO(Y,y_1) such that D$\cap$F=ϕ.Hence Y is a-T_2.*

Theorem 4.11. *Let f:(X,τ) $\rightarrow$ (Y,η) have a (D,a)-closed graph G_f . If f is surjective, then (Y,η) is Da-T_2(resp,Da-T_1).*
Proof:*Similar to the proof of Theorem 4.10*

Theorem 4.12. *Let f:(X,τ) $\rightarrow$ (Y,η) have a (D,a)-closed graph G_f . If f is injective, then (X,τ) is Da-T_1.*
Proof:*Similar to the proof of Theorem 4.6*

Corolary 4.3. *Let f:(X,τ) $\rightarrow$ (Y,η) have a (D,a)-closed graph G_f. If f is bijective, then both (X,τ) and (Y,η) are Da-T_1*
Proof:*Follows from Theorems 4.11 and 4.12*

5 Almost Da-Continuous functions.

Definition 5.1. *A function f:(X,τ) $\rightarrow$ (Y,η) is said to be almost Da-continuous if for each point p $\in$ X and each open subset V of Y containing f(p), there exists U $\in$ DaO(X,p) such that f(U) $\subset$ int(Cl(V)).*

Theorem 5.1. *If f:(X,τ) $\rightarrow$ (Y,η) is Da-continuous function , then f is an almost Da-continuous,but not conversely.*
Proof:*Obvious*

Example 5.1. *Consider (X,τ) and (X,η) as in 4.1. Define f:(X,τ) $\rightarrow$ (X,η) by f(p)=p,f(q)=s,f(r)=q and f(s)=r Then f is almost Da-continuous but not Da-continuous since $\{p,q,r\}$ is open in (X,η), $f^{-1}(\{p,q,r\})=\{p,r,s\}\notin$DaO(X,$\tau$)*

Definition 5.2. *[Noiri and Popa, 1998] A space X is said to be semi-regular if for any open set U of X and each point x $\in$ U there exists a regular open set V of X such that x $\in$ V $\subset$ U.*

Theorem 5.2. *If f:(X,τ) $\rightarrow$ (Y,η) is an almost Da-continuous function and Y is semi-regular, then f is Da-continuous.*
Proof: *Let p $\in$ X and let V $\in$ O(Y,f(p)). By the semi-regularity of Y , there exists G$\in$RO(Y,f(p)) such that G $\subset$ V . Since f is almost Da-continuous, there exists U $\in$ DaO(X, x) such that f(U) $\subset$ Int(Cl(G)) = G $\subset$ V and hence f is Da-continuous.*

Lemma 5.1. *Let (X,τ) be a space and let A be a subset of X. The following statements are true:*
(1) A $\in$ PO(X) if and only if sCl(A) = Int(Cl(A)) [Janković, 1985].
(2) A $\in$ βO(X) if and only if Cl(A) is regular closed [Abd El-Monsef, 1983].

Theorem 5.3. *Let f:(X,τ) $\rightarrow$ (Y,η) be a function. Then the following conditions are equivalent:*
(1) f is almost Da-continuous;
(2) For every N$\in$RO(Y), f^{-1}(N)$\in$DaO(X);
(3) For every M$\in$RC(Y), f^{-1}(M)$\in$DaC(X);
(4) For each subset C of X, f(Cl_a^D(C)) $\subset$ Cl_δ(f(C));
(5)For each subset D of Y, Cl_a^D(f^{-1}(D)) $\subset$ f^{-1}(Cl_δ(D));
(6)For every G$\in\delta$C(Y), f^{-1}(G)$\in$DaC(X);
(7)For every H$\in\delta$O(Y), f^{-1}(H)$\in$DaO(X);
(8) For every N$\in$O(Y), f^{-1}(Int(Cl(N)$\in$DaO(X);
(9) For every M$\in$C(Y), f^{-1}(Cl(Int(M)$\in$DaC(X);
(10) For every N$\in\beta$O(Y), Cl_a^D(f^{-1}(N)) $\subset$ f^{-1}(Cl(N));
(11) For every M$\in\beta$C(Y), f^{-1}(Int(M)) $\subset$ Int_a^D(f^{-1}(M));
(12) For every M$\in$SC(Y), f^{-1}(Int(M)) $\subset$ Int_a^D(f^{-1}(M));
(13) For every N$\in$SO(Y), Cl_a^D(f^{-1}(N)) $\subset$ f^{-1}(Cl(N));
(14) For every M$\in$PO(Y), f^{-1}(M) $\subset$ Int_a^D(f^{-1}(Int(Cl(M));
(15) For each p$\in$ X and each N$\in$O(Y,f(p)), there exists M $\in$ DaO(X,p) such that f(M) $\subset$ sCl(N);
(16) For each p$\in$ X and each N$\in$RO(Y,f(p)), there exists M $\in$ DaO(X,p) such that f(M) $\subset$ N;
(17) For each p$\in$ X and each N$\in\delta$O(Y,f(p)), there exists M $\in$ DaO(X,p) such that f(M) $\subset$ N.
Proof: *(1)$\longrightarrow$(2) Similar to the proof of (1)$\longrightarrow$(2) of Theorem 4.1.*
(2)$\longrightarrow$(3) It follows from the fact that f^{-1}(Y$\setminus$F) = X $\setminus f^{-1}$(F).
(3)$\longrightarrow$(4) Suppose that D$\in$ δC(Y) such that f(C)$\subset$ D. Observe that D = Cl_δ(D) =$\bigcap\{$F:D$\subset$F and F$\in$RC(Y)$\}$ and so f^{-1}(D) = $\bigcap\{f^{-1}$(F):D$\subset$F and F$\in$RC(Y)$\}$. By (3) and Corollary 3.1,we have f^{-1}(D)$\in$DaC(X) and C$\subset$ f^{-1}(D). Hence Cl_a^D(C) $\subset f^{-1}$(D), and it follows that f(Cl_a^D(C)) $\subset$ D. Since this is true for any δ-closed set D containing f(C), we have f(Cl_a^D(C))$\subset$ Cl_δ(f(C)).
(4)$\longrightarrow$(5) Let D $\subset$ Y, then f^{-1}(D) $\subset$ X. By (4), f(Cl_a^D(f^{-1}(D)))$\subset$ Cl_δ(f(f^{-1}(D)))$\subset$$Cl_\delta$(D). So that Cl_a^D(f^{-1}(D)) $\subset$ f^{-1}(Cl_δ(D)).
(5)$\longrightarrow$(6) Let G$\in\delta$C(Y) Then by (5), Cl_a^D(f^{-1}(G)) $\subset$ f^{-1}(Cl_δ(G))=f^{-1}(G). In consequence, Cl_a^D(f^{-1}(G))=f^{-1}(G) and hence by Theorem 3.3(4), f^{-1}(G)$\in$DaC(X).
(6)$\longrightarrow$(7):Clear.
(7)$\longrightarrow$(1): Let p$\in$ X and let O$\in$O(Y,f(p)). Set D = Int(Cl(O)) and C =f^{-1}(D). Since D$\in$ δO(Y), then by (7), C = f^{-1}(D) $\in$ DaO(X). Now, f(p) $\in$ O= Int(O)$\subset$ Int(Cl(O)) = D it follows that p$\in f^{-1}$(D)=C and f(C)=f(f^{-1}(D)$\subset$D=Int(Cl(O).
(2)$\longleftrightarrow$(8): Let N $\in$O(Y). Since Int(Cl(N))$\in$RO(Y),by (2), f^{-1}(Int(Cl(N))$\in$DaO(X). The converse is similar.
(3)$\longleftrightarrow$(9)It is similar to (8)$\longleftrightarrow$(2).

(3)⟶ (10): Let N∈βO(Y).Then by Lemma 5.1(2),Cl(N) ∈ RC(Y).So by(3),f^{-1}(Cl(N)) ∈DaC(X) .Since f^{-1}(N) ⊂ f^{-1}(Cl(N)) and by Theorem 3.3(4),Cl_a^D(f^{-1}(N))⊂ f^{-1}(Cl(N)).
(10)⟶ (11): and (12)⟶ (13):Follows from Lemma 3.3
(11)⟶ (12):It follows from the fact that SC(Y)⊂βC(Y)
(13)⟶ (3):It follows from the fact that RC(Y)⊂SO(Y).
(2)⟷ (14): Let N ∈PO(Y). Since Int(Cl(N)) ∈ RO(Y),then by (2), f^{-1}(Int(Cl(N))) ∈ DaO(X) and hence f^{-1}(N) ⊂ f^{-1}(int(Cl(N))) = Int_a^D(f^{-1}(int(Cl(N)))). Conversely,let N∈RO(Y). Since N ∈ PO(Y), f^{-1}(N) ⊂ Int_a^D(f^{-1}(int(Cl(N)))) =Int_a^D(f^{-1}(N)). In consequence, Int_a^D(f^{-1}(N))=f^{-1}(N) and by Theorem 3.4, f^{-1}(N) ∈ DaO(X).
(1)⟶ (15): Let p∈X and N∈O(Y,f(p)). By (1), there exists M∈ DaO(X,p) such that f(M) ⊂ Int(Cl(N)).Since N∈PO(Y),by Lemma 5.1, f(M) ⊂ sCl(N).
(15)⟶ (16): Let p∈ X and N∈RO(Y,f(p)). Since N∈O(Y,f(p)) and by (15), there exists M∈DaO(X,p) such that f(M)⊂ sCl(N). Since N ∈PO(Y), then by Lemma 5.1, f(M) ⊂Int(Cl(N)) = N.
(16)⟶ (17):Let p∈ X and V∈δO(Y,f(p)). Then, there exists G∈O(Y.f(p))such that G ⊂ Int(Cl(G)) ⊂ N. Since Int(Cl(G))∈RO(Y,f(p)), by (16), there exists M∈ DaO(X,p) such that f(M) ⊂ Int(Cl(G))⊂ N.
(17)⟶(1). Let p∈ X and N∈O(Y,f(p)). Then Int(Cl(N))∈ δO(Y,f(p)). By (17), there exists M∈ DaO(X,p) such that f(M) ⊂ Int(Cl(N)). Therefore,f is almost continuous

Theorem 5.4. *If f:(X,τ) → (Y,η) is an almost Da-continuous injective function and (Y,η) is r-T_1 , then (X,σ) is Da-T_1 .*
Proof: *It is similar to the proof of Theorem 4.3*

Theorem 5.5. *If f:(X,τ) → (Y,σ) is an almost Da-continuous injective function and (Y,σ) is r-T_2 , then (X,τ) is Da-T_2 .*
Proof: *It is similar to the proof of Theorem 4.4*

Lemma 5.2. *[Ayhan and Ozkoç, 2016] Let (X,τ) be a space and let A be a subset of X. Then:*
A ∈e^*O(X) if and only if Cl_δ(A) is regular closed.

Theorem 5.6. *For a function f:(X,τ) → (Y,η),the following are equivalent:*
(a) f is almost Da-continuous;
(b) For every e^*-open set N in Y, f^{-1}(Cl_δ(N)) is Da-closed in X;
(c) For every δ-semiopen subset N of Y, f^{-1}(Cl_δ(N)) is Da-closed set in X;
(d) For every δ-preopen subset N of Y, f^{-1}(Int(Cl_δ(N))) is Da-open set in X;
(e) For every open subset N of Y, f^{-1}(Int(Cl_δ(N))) is Da-open set in X;
(f) For every closed subset N of Y, f^{-1}(Cl(Int_δ(A))) is Da-closed set in X .
Proof: *(a)→(b):Let N∈e^*O(Y) Then by Lemma 5.2,Cl_δ(N)∈RC(Y).*

By (a), $f^{-1}(Cl_\delta(N)) \in DaC(X)$*.*
(b)→(c):Obvious since $\delta SO(Y) \subset e^*O(Y)$*.*
(c)→(d):Let $N \in \delta PO(Y)$*,then* $Int_\delta(Y \backslash N) \in \delta\text{-}SO(Y)$*.By (c),*
$f^{-1}(Cl_\delta(Int_\delta(Y \backslash N)) \in DaC(X)$ *which implies* $f^{-1}(Int(Cl_\delta(N)) \in DaO(X)$*.*
(d)→(e):Obvious since $O(Y) \subset \delta PO(Y)$*.*
(e)→(f):Clear
(f)→(a):Let $N \in RO(Y)$*.Then* $N = Int(Cl_\delta(N))$ *and hence* $Y \backslash N \in C(X)$*. By (f),*
$f^{-1}(Y \backslash N) = X \backslash f^{-1}(Int(Cl_\delta(N))) = f^{-1}(Cl(Int_\delta(Y \backslash N)) \in DaC(X)$*.*
Thus $f^{-1}(N) \in DaO(X)$*.*

Lemma 5.3. *[Ayhan and Ozkoç, 2016] Let* (X,τ) *be a space and let* $A \subset X$*. The following statements are true:*
(a) For each $A \in e^*O(X)$*,* $a\text{-}Cl(A) = Cl_\delta(A)$
(b)For each $A \in \delta SO(X)$*,* $\delta\text{-}pCl(A) = Cl_\delta(A)$*.*
(c)For each $A \in \delta PO(X)$*,*$\delta\text{-}sCl(A) = Int(Cl_\delta(A))$*.*

As a consequence of Theorem 5.6 and Lemma 5.3, we have the following result:

Theorem 5.7. *The following are equivalent for a function* $f:(X,\tau) \to (Y,\eta)$*:*
(a) f is almost Da-continuous;
(b) For every e^**-open subset G of Y,*$f^{-1}(a\text{-}Cl(G))$ *is Da-closed set in X;*
(c) For every δ*-semiopen subset G of Y,*$f^{-1}(\delta\text{-}pCl(G))$ *is Da-closed set in X;*
(d) For every δ*-preopen subset G of Y,*$f^{-1}(\delta\text{-}sCl(G)))$ *is Da-open set in X;*

References

ME Abd El-Monsef. β-open sets and β-continuous mappings. *Bull. Fac. Sci. Assiut Univ.*, 12:77–90, 1983.

Burcu Sünbül Ayhan and Murad Ozkoç. Almost e-continuous functions and their characterizations. *Journal of Nonlinear Sciences & Applications (JNSA)*, 9(12), 2016.

Julian Dontchev, I Arokiarani, and Krishnan Balachandran. On generalized delta-closed sets and almost weakly hausdorff spaces. *Questions and Answers in General Topology*, 18(1):17–30, 2000.

Erdal Ekici. Generalization of perfectly continuous, regular set-connected and clopen functions. *Acta Mathematica Hungarica*, 107(3):193–206, 2005.

Erdal Ekici. On a-open sets a^*-sets and decompositions of continuity and super-continuity. In *Annales Univ. Sci. Budapest*, volume 51, pages 39–51, 2008.

Erdal Ekici. On e*-open sets and (d, s)*-sets. *Mathematica Moravica*, 13(1): 29–36, 2009.

MS El Naschie. On the uncertainty of cantorian geometry and the two-slit experiment. *Chaos, Solitons & Fractals*, 9(3):517–529, 1998.

MS El Naschie. On the unification of heterotic strings, m theory and e^{∞} theory. *Chaos, Solitons & Fractals*, 11(14):2397–2408, 2000.

MS El Naschie. Stability analysis of the two-slit experiment with quantum particles. *Chaos, Solitons & Fractals*, 26(2):291–294, 2005.

DS Janković. A note on mappings of extremally disconnected spaces. *Acta Mathematica Hungarica*, 46(1-2):83–92, 1985.

Efim Khalimsky, Ralph Kopperman, and Paul R Meyer. Computer graphics and connected topologies on finite ordered sets. *Topology and its Applications*, 36 (1):1–17, 1990.

Vladimir Kovalevsky. Some topology-based image processing algorithms. *Ann. New York Acad. of Sci.; Papers on General Topology and Applications*, 728: 174–182, 1994.

Norman Levine. Semi-open sets and semi-continuity in topological spaces. *The American mathematical monthly*, 70(1):36–41, 1963.

Norman Levine. Generalized closed sets in topology. *Rendiconti del Circolo Matematico di Palermo*, 19(1):89–96, 1970.

AS Mashhour. On preconlinuous and weak precontinuous mappings. In *Proc. Math. Phys. Soc. Egypt.*, volume 53, pages 47–53, 1982.

T Noiri and V Popa. On almost β-continuous functions. *Acta Mathematica Hungarica*, 79(4):329–339, 1998.

Jin Han Park, Bu Young Lee, and MJ Son. On δ-semiopen sets in topological space. *J. Indian Acad. Math*, 19(1):59–67, 1997.

S Raychaudhuri and MN Mukherjee. On δ-almost continuity and δ-preopen sets. *Bull. Inst. Math. Acad. Sinica*, 21(4):357–366, 1993.

Michael B Smyth. Semi-metrics, closure spaces and digital topology. *Theoretical computer science*, 151(1):257–276, 1995.

Marshall Harvey Stone. Applications of the theory of boolean rings to general topology. *Transactions of the American Mathematical Society*, 41(3):375–481, 1937.

NV Velicko. H-closed topological spaces. *Amer. Math. Soc. Transl.*, 78:103–118, 1968.

Ratio Mathematica Volume 39, 2020, pp. 253-259

Determine the value *d(M(G))* for non-abelian *p*-groups of order *q* = *pnk* of Nilpotency *c*

Behnam Razzaghmaneshi[*]

Abstract

In this paper we prove that if *n*, *k* and *t* be positive integer numbers such that *t* < *k* < *n* and *G* is a non abelian *p*-group of order *pnk* with derived subgroup of order *pkt* and nilpotency class c, then the minimal number of generators of *G* is at most *p*1 2 ((*nt*+*kt*−2)(2*c*−1)(*nt*−*kt*−1)+*n*. In particular, |*M*(*G*)| _ *p*1 2 (*n*(*k*+1)−2)(*n*(*k*−1)−1)+*n*, and the equality holds in this last bound if and only if *n* = 1 and *G* = *H* ×*Z*, where *H* is extra special *p*-group of order *p*3*n* and exponent *p*, and *Z* is an elementary abelian *p*-group.

Keywords: Schur multiplier, elementary abelian, p-group, extra special

2010 AMS subject classification: 20F45; 20F05, 11Y35.[†]

---

[*] Assistant Professor of mathematics, Algebra, Islamic Azad University, Talesh Branch, Talesh, Iran. B_razzagh@yahoo.com.

[†] Received on October 22nd, 2020. Accepted on December 17th, 2020. Published on December 31th, 2020. doi: 10.23755/rm.v39i0.560. ISSN: 1592-7415. eISSN: 2282-8214.

1. Introduction

Let G be a finite group and $G = FR$ a presentation for G as a factor group of the free group F. Then Schur in [11], show that $M(G) = (F_0 \backslash R)\ [F,R]$.
(1.1) Recall that, for two finite groups A and B, $AB \_= (\ AA_0\ )(\ BB_0\ )$.
Michael R. Jones in years 1973 and 1974 for the finite group G, get some inequalities for $d(M(G))$ and $e(M(G))$, which $d(M(G))$ and $e(M(G))$ the minimal number of generators and exponent of finite group G, respectively. now in current paper we generalized and compute the value $d(M(G))$ and $e(M(G))$ for non-abelian pgroups of order $q = p_{nk}$ and nilpotency c.
Notation: The notation used in this paper is as follows:
(i) If G is a finite group then $E(G)$ denotes exponent of G and $D(G)$ denotes the minimal number of generators of G.
(ii) The the lower central series of a group G is denoted by $G = g_1(G) \_ g_2(G) = G_0 \_ g_3(G) \_ ...$, where for $j \_ 1$, $g_{i+1}(G) = [g_i(G),G]$.
And the upper central series of a group G is denoted by $1 = Z_0(G) \_ Z_1(G) = G_0 \_ Z_2(G) \_ ...$, where for $i \_ 0$, $Z_{i+1}Z_i \_ Z(\ GZ_i(G)\ )$.
The main theorem of this paper as follows.

Main Theorem: Let n, k and t be positive integer numbers such that $t < k < n$ and G is a non abelian p-group of order p_{nk} with derived subgroup of order p_{kt} and nilpotency class c, then the minimal number of generators of G, $(D|M(G)|)$ is $p_{12}\ ((2c-1)n_2-k(k-1)-3n+4$.

2. Some definition, lemma and theorems

The results of this section are several lemma and theorems, where the proofs of their in references [6], [7] and [8], and so we will be omitted.
2.1. Lemma: Let G be a finite group and B a normal subgroup. Set $A = GB$. Let $G = F\ R$ be a presentation for G as a factor group of the free group F and suppose $B = SR$ so that $A = FS$. Then $[F,S]\ [F,R][F,S,F]S_0$ is isomorphic with a factor group of AB.
Proof. See to ([6], Lemma 2.1).
2.2. Corollary. Further to the notation and assumptions of Lemma 2.1, let B 2 be a central subgroup of G. Then $[F,R]\ [F,R]S_0$ is an epimorphic image of AB.
Proof. See to ([6]).
2.3. Definition. Let G be a finite group. We say that G has (special) rank $r(G)$ if every subgroup of G may be generated by $r(G)$ elements and there is at least one subgroup that cannot be generated by fewer than $r(G)$ elements.
Let $G = F\ R$ be a presentation for the finite p-group G as a factor group of a free group F. Let $\Gamma_{i+1} = g_{i+1}(F)$ for all i. Since $G_0 = F_0R\ R$ we have by (1.1), that

$M(G G_0)$ _= $(F_0 \backslash F_0 R)$ $[F,F_0 R]$ = $F_0 [F,F_0 R]$.
With this notation we have:

2.4. Theorem: Let G be a finite p-group of nilpotency class c and $Q_i = G$
$g_i(G)$ for 2 _ i _ c. Then (i) $|G_0||M(G)|$ _ $|M(G G_0)\prod c-1\ i=1\ |Q_{i+1} g_{i+1}(G)|$,
(ii)D(M(G)) _ $D(M(G G_0))+\sum c-1\ i=1\ D(Q_{i+1} g_{i+1}(G))$,
(iii) $E(M(G))$ _ $E(M(G G_0))\prod c-1\ i=1\ E(Q_{i+1} g_{i+1}(G))$.
(i) In the above notation, $|G_0||M(G)| = |\ F_0 [F,R]\ | = |M(G G_0)|\ |[F,F_0 R]$
$[F,R]| = |M(G G_0)|\ |[F,F_{i+2} R]\ [F,R]|\ \prod I\ k=1\ |\ [F,\Gamma_{k+1} R|\ [F,\Gamma_{k+2} R,$ for all i _ 1. Now, 1 =
$g_{c+1}(G) = \Gamma_{c+1} R\ R$ so that Γ_{c+1} _ R and $[F,F_{c+1} R] = [F,R]$.
Next, $g_i(G) = \Gamma_i R\ R$ for all i _ 2. Thus $[F,R](\Gamma_i R)_0[F,\Gamma_i R,F] = [F,R]\Gamma_{i+2} = [F,\Gamma_{i+1} R]$
and (i) follows by Lemma 2.1. (ii) We have,
$r(F_0[F,R])$ _ $r(M(G G_0))+r([F,\Gamma_2 R]\ [F,R]$ so that $D(M(G))$ _ $D(M(G G_0))+\sum c-1$
$i=1\ r([F,\Gamma_{i+1} R]\ [F,\Gamma_{i+2} R])$, and (ii) again follows by Lemma 2.1.
(iii) This follows as for (i) and (ii).

3. The proof of main Theorem

In this section we show that, Let n, k and t be positive integer numbers such that $t < k < n$ and G is a non abelian p-group of order p^{nk} with derived subgroup of order p^{kt} and nilpotency class c, then the minimal number of generators of G, $(D|M(G)|)$ is $p^{\frac{1}{2}((2c-1)n^2-k(k-1)-3n+4}$. For proof of this work we action as follows:
Proof. Let n, k and t be positive integer numbers such that $t < k < n$ and G is a non abelian p-group of order p^{nk} with derived subgroup of order p^{kt} and nilpotency class c. Then by using of Theorem 2.4(ii),we have
$D(M(G))$ _ $D(M(G G_0))+\sum c-1\ i=1\ D(Q_{i+1} g_{i+1}(G))$.
If $D(M(G)) = n$ then the above relation will coming as follows:
$D(M(G))$ _ $\frac{1}{2}((n+k-2)(n-k-1)+1)+n(\sum c-1\ i=1\ g_{i+1}(G))$.
= $\frac{1}{2}((n+k-2)(n-k-1)+1)+n^2(c-1)$. Which the result now follows.
In 1904, Schur [11,12] prove that for every finite groups H and K, then $M(H \times K) = M(H)\times M(K)\times\ H H_0\ \ K K_0$.
In 1957, Green [5] show that if G be a p-group of order p^n, then $|M(G)|$ _ $p^{\frac{1}{2}n(n-1)}$.
In 1967, Gaschatz el al [4] prove that if G be a d-generator p-group of order p^n, G_0 has order p^c and $G\ Z(G)$ is a d- generator group, then $|M(G)|$_ $p^{\frac{1}{2}d(2n-2c-d-1)+2(d-1)c}$.
In 1973, Jones [4-6] show that if G be a p-group of order p^n and $|G_0| = p^k$, then $|M(G)|$ _ $p^{\frac{1}{2}n(n-1)-k}$.
In 1982, Byel and Tappe [2] shown that if G be a Extra especial p-group of order p^{2m+1}, then

(i) If m _ n, than $|M(G)| = p^{2m^2-m-1}$.
(ii) If $m = 1$, then the order of Schur multiplier of D_8, Q_8, E_1 and E_2 are equal 2, 1, p^2 and 1, respectively.
In 1991, Berkovich [1] show that if G be a p-group of order p^n, then $t(G) = 0$ if and only if G _= $Z_p^{(n)}$, and also $t(G) = 1$ if and only if G _= $Z^{(2)}$ or G _= E_1.
In 1994, Zhou [14]prove that if G be a p-group of order p^n, then $t(G) = 2$ if and only if G _= $Z \times Z_{p^2}$ or G _= D_8, G _= $E_1 \times Z_p$.
In 1999, Ellis [3]show that if G be a p-group of order p^n, then $t(G)=3$ if and only if G _= Z_{p^3}, G _= $Z_p^{(2)} \times Z_{p^2}$ or G _= Q_8, G _= E_2, G _= $D_8 \times Z_2$ or G _= $E_1 \times Z_p^{(2)}$.
In 2009, P.Niroomand [10] show that if G be a non-abelian finite p-group of order p^n and $|G'| = p^k$, then $|M(G)|$ is $p^{\frac{1}{2}((n+k-2)(n-k-1)+1}$. In particular, $|M(G)|$ _ $p^{\frac{1}{2}(n-2)(n-1)+1}$, and the equality holds in this last bound if and only if $G = E_1 \times Z$, where Z is an elementary abelian p-group.
The Schur multiplier of abelian groups may be calculated easily by a result [12] which was obtained by Schur. So in this paper, we focus on non-abelianp-groups.
This paper is devoted to the derivation of certain upper bound for the Schur multiplier of non-abelian p-groups of order p^{nk} with derived subgroup of order p^k. We prove that $|M(G)|$ _ $p^{\frac{1}{2}(nk+nt-2)(nk-nt-1)+n}$. In particular, if $|M(G)| = p^{\frac{1}{2}(n(k+1)-2)(n(k-1)-1)+n}$, we characterize the structure of the group G. If G is a p-group of order p^n, Jones [4] proved that $|M(G)||G'|$ _ $p^{\frac{1}{2}n(n-1)}$ which shows that $|M(G)|$ _ $p^{\frac{1}{2}n(n-1)+1}$ when G is a non-abelian p-group of order p^n. So, the general bound given above is better than Joness bound unless $|G| = p^3$, in which case the two bounds are the same.The principal result of this paper is presented in the following theorem.

Main Theorem. Let G be a non-abelian finite p-group of order p^{nk}. If $|G'| = p^{nt}$, then we have $M(G)$ _ $p^{\frac{1}{2}(nk+nt-2)(nk-nt-1)+n}$. In particular $M(G)$ _ $p^{\frac{1}{2}(n(k+1)-2)(n(k-1)-1)+n}$,
and the equality holds in this last bound if and only if $n-1$ and $G = H \times Z$, where H is an extra special p-group of order p^{3n} and exponent p, and Z is an elementary abelian p-group.

Preliminaries and Elementary Theorems.
In this section, we want to several Theorems and Lemmas whose proved in references
[1-14]. At first we list the following theorems, which are used in our proofs.
Our method for the proof is similar to P. Niroomand (2009) and Berkovich, Ya.G.
(1991), which we compute for groups of order p^{nk}.

Theorem 2.1.(See [7,theorem 3.1 and Theorem 4.1].) Let G be a finite p- group and let N be a central subgroup of G. Then $|M(G\ N\ | \_ \ |M(G)||G0 \backslash N| \_ \ |M(G\ N\ ||M(N)||G\ N\ N|$.

Theorem 2.2.(See[9, Theorem 3.3.6].) Let G be an extra special p-group of order $p2m+1$. Then:
(i) If m _ 2, then $M(G) = p2m2-m-1$.
(ii) If m=1, then $M(G)$ _ $p2$, and the equality holds if and only if G is of exponent p.
Theorem 2.3.(See [9, Theorem 2.2.10].) For every finite groups H and K, we Have $M(H \times K \_= M(H)\times M(K)\times\ H\ H0\ \ K\ K0$.
Corollary 2.4. If $G \_= Cm1 \times Cm2 \times ...\times Cmk$, where $mi+1$ divides mi for all i, $1 \_ i \_ k$, then $M(G) \_=Cm2 \times C(2)\ m3 \times ...\times C(k-1)\ mk$.
Proof of the Main Theorem
In this section we want to prove our result. The following technical lemmas shorten the proof of our main Theorem.
Lemma 3.1. Let G be a finite p-group of order pn such that $G\ G0$ is elementary 6 of order $pn-1$, then G is a central product of an extra special p-group H and $Z(G)$ such that $H \backslash Z(G) = G0$.
Proof. Let $H\ G0$ be the complement of $Z(G)\ G0$ in $G\ G0$. Then $G = HZ(G)$, so $G0 = H0$ and $Z(H) = Z(G) \backslash H$. On the other hand, 1 6= $Z(G) \backslash H \_ G0$, and the result follows.
Lemma 3.2. Let G be an abelian p-group of order pn which is elementary abelian. Then $M(G) \_ p1\ 2\ (n-1)(n-2)$.
Proof. the result is obtained obviously if G is cyclic. So, let $G \_=Cpm1\times Cpm2\times ...\times Cpmk$ such that $\sum k\ i=1 mi = n$ and $m1 \_ m2 \_ ... \_ mk$. We know that $m1 \_ 2$, and then, by using *Corollary*2.4, $|M(G)| = pm2+2m3+...+(k-1)mk$
$\_ p(m2+m3+...+mk)+(m3+...+mk)+...+mk \_ p1\ 2\ (n-1)(n-2)$.

Lemma 3.3. Let G be a non- abelian p-group of order pnk with derived subgroup of order p such that $G\ G0$ is not elementary abelian, then $M(G) < p12$
$(nk-1)(nk-2)+1$.
Proof. by using Theorem 2.1 and Lemma 3.2,
$M(G) \_ p-1|M(\ G\ G0)||\ G\ G0\ G0| \_ p-1p1\ 2\ (nk-2)(nk-3)p(nk-1) < p1\ 2\ (nk-1)(nk-2)+1$.
which completes the proof.
Lemma 3.4. let G be a non- abelian p-group of order pnk, such that G
$G0$ is elementary abelian of order $pnk-1$, then $M(G) \_ p1\ 2\ (nk-1)(nk-2)+1$ and the equality holds if and only if $G = H\times Z$, where H is extra special p- group of order $p3n$ and exponent p, and Z is elementary abelian p-group.
Proof. By Lemma 3.1, G is central product of H and $Z(G)$, and Theorem 2.2,
7 we may assume that $|Z(G)| \_ p2$. Let $|H| = p2m+1$, so $|Z(G)| = pn-2m$.

Suppose first that $m \_ 2$. If $Z(G)$ is elementary abelian, let T be a group such that $Z(G) \cong G' \times T$. By using Theorems 2.2 and 2.3, we have
$|M(G)| = |M(H \times T)| = |M(H)||M(T)||\frac{H}{H'} \otimes T| = p^{2m^2-m-1} p^{\frac{1}{2}(n-2m-1)(n-2m-2)} p^{2m(n-2m-1)} = p^{\frac{1}{2}(n^2-3m)} < p^{\frac{1}{2}(n-1)(n-2)+1}$.
Now assume that $Z(G)$ is not elementary abelian. Theorems 2.1 and 2.3 imply That $|M(G)| \leq p|M(H \times Z(G)| = p|M(H)||M(Z(G))||\frac{H}{H'} \otimes Z(G)|$.
Hence by using Theorem 2.2 and Lemma 3.2, we have
$|M(G)| \leq p p^{2m^2-m-1} p^{\frac{1}{2}(n-2m-1)(n-2m-2)} p^{2m(n-2m-1)} < p^{\frac{1}{2}(n-1)(n-2)+1}$.
If H is extra special of order p^{3n} and $Z(G)$ is not elementary abelian, then Theorem 2.1 implies that $|M(G)| \leq p^{-1}|M(\frac{G}{Z(G)})||M(Z(G))||\frac{G}{Z(G)} \otimes Z(G)| \leq p^{\frac{1}{2}nk(nk-3)+1} < p^{\frac{1}{2}(nk-1)(nk-2)+1}$.
By Theorem 2.2, it is easy to see that if $Z(G)$ is elementary abelian, then $|M(G)| = p^{\frac{1}{2}(nk-1)(nk-2)+1}$ if H is extra special of order p^{3n} and exponent p; and in other cases $|M(G)| < p^{\frac{1}{2}(nk-1)(nk-2)+1}$.

Proof of the Main Theorem we prove the theorem by induction on t. if $t = 1$ the result is obtained by Lemma 3.2 and 3.4. Let G be a non-abelian p-group of order p^{nk} with derived subgroup of order p^{nt} $(t \_ 2)$. Choose K in $G' \setminus Z(G)$ of order $p-1$. By using induction hypothesis, we have $|M(\frac{G}{K})| \leq p^{\frac{1}{2}(nk+nt-4)(nk-nt-1)+n}$.
On the other hand, By using Theorem 2.1, implies that $|M(G)| \leq p^{-1}|M(\frac{G}{K})||M(K)||(\frac{G}{G'} \otimes K)| \leq p^{-1} p^{\frac{1}{2}(nk+nt-4)(nk-nt-1)} p^{n-1} p^{(nk-nt)} \leq p^{\frac{1}{2}(nk+nt-4)(nk-nt-1)} p^{n-1} p^{(nk-nt)}\, p^{\frac{1}{2}(nk+nt-2)(nk-nt-1)+n}$.
Now let G be a p-group of order p^{nk} such that $|M(G)| = p^{\frac{1}{2}(nk-1)(nk-2)+n}$. If $|G'| \_ p^{2k}$, then $|M(G)| \leq p^{\frac{1}{2}(n(k-1)-1)(n(k+1)-2)}$, which is a contradiction.
Since $|G'| = p^k$, Lemma 3.3 implies that G/G' is elementary abelian. Hence Lemma 3.4 shows that $G = H \times Z$, where H is an extra special p- group of order p^{3n} and exponent p, and Z is an elementary abelian p-group, so the result follows.

References

[1] Berkovich, Ya.G., *On the order of the commutator subgroups and the Schur multiplier of a finite p-group,*, J. Algebra, 144**.** (1991) 269272.

[2] Beyl, F.R., and J. Tappe, *Group Extensions, Representations and the Schur Multiplicator,*, Vol. 958, SpringerVerlag, Berlin/Heidelberg/New York, 1972.

[3] Ellis, G., *On the Schur multiplier of p-group,*, Comm. Algebra. 27(9), (1999), 4173-4177.

[4] Gaschutz, W., Neubu¨ser, J. and Yen. T., *U¨ ber den Multiplikator von p-Gruppen,* Math. Z. **100** (1967), 93-96.

[5] Green, J.A., *On the number of automorphisms of a finite group,*, Proc. Roy. Soc. A 237 (1956) 574581.

[6] Jones, M.R., *Multiplicators of p-groups,*, Math. Z. 127 (1972) 165166.

[7] Jones, M.R., *Some inequalities for the multiplicator of a finite group,* , Proc. Amer. Math. Soc. 39 (1973) 450456.

[8] Jones, M.R., *Some inequalities for the multiplicator of a finite group II,*, Proc. Amer. Math. Soc. 45 (1974) 167172.

[9] Karpilovsky, G.,The Schur multiplier, London Math. Soc. Monogr. (N.S.) (1987).

[10] Niroomand. P., On the order of Schur multiplier of non-abelian p-groups, J. Algebra 322 (2009), 4479 4482.

[11] Schur, I., *ber die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen,* , J. Reine Angew. Math. 127 (1904) 2050.

[12] Schur, I., *Untersuchungen ber die Darstellung der endlichen Gruppen durchgebrochene lineare Substitutionen,*, J. Reine Angew. Math. 132 (1907) 85137.

[13] Wiegold, J., *The Schur multiplier of p-groups with large derived subgroup groups,*, Arch. Math. 95 (2010), 101-103.

[14] Zhou, X., *On the order of Schur multipliers of finite p-groups*, Comm. Algebra. 22**(1)**, (1994), 1-8.

Ratio Mathematica Volume 39, 2020, pp. 261-268

Thermodynamic behavior of the polytropic gas in cosmology

Prasanta Das*
Kangujam Priyokumar Singh†

Abstract

In this paper, we investigate on the thermodynamic behavior of Polytropic gas as a candidate for dark energy by considering the relation $P = K\rho^{1+\frac{1}{n}}$, where K and n are the Polytropic constant and Polytropic index respectively. Furthermore, P indicates the pressure and ρ is the energy density of the fluid such that $\rho = \frac{U}{V}$ where U and V represent the internal energy and volume, respectively. At first, we find an exact expression for the energy density of the Polytropic gas using thermodynamics and later on, discuss different physical parameters. Finally our study shows that the Polytropic gas may be used to describe the expansion history of the universe from the dust dominated era to the current accelerated era and it is thermodynamically stable.

Keywords: Cosmology; Dark energy; Polytropic gas; Thermodynamics.
2010 AMS subject classification: 83F05, 37D35, 82B30.‡

*Department of Mathematical Sciences, Bodoland University, Kokrajhar, BTR, Assam-783370, India; prasantadasp4@gmail.com.
†Department of Mathematical Sciences, Bodoland University, Kokrajhar, BTR, Assam-783370, India; pkangujam18@gmail.com.
‡Received on October 4th, 2020. Accepted on December 17th, 2020. Published on December 31st, 2020. doi: 10.23755/rm.v39i0.564. ISSN: 1592-7415. eISSN: 2282-8214.

1. Introduction

Cosmologists suggest that our universe expands under an accelerated expansion [1]-[7]. In the standard Friedman Lemaitre Robertson Walker (FLRW) cosmology, a new energy with negative pressure, called dark energy (DE) is responsible for this expansion [8]. The nature of the DE is still unknown and various problems have been proposed by the researchers in this field**.** About 70% of the present energy of the universe is contained in the DE. The cosmological constant with the time independent equation of state is the earliest, simplest and most traditional candidate for the dark energy which can be taken into account as a perfect fluid satisfying the relation $\rho + P = 0$. But it has some problems like fine-tuning and cosmic coincidence puzzles [9], [10]. Besides the cosmological constant, the other dark energy models are quintessence [11], phantom [12], tachyon [13], holographic dark energy [14] [15], K-essence [16] and Chaplygin gas models with various equation of state. Polytropic gas is one of the dynamical dark energy models [17].

In the present study, we want to investigate the thermodynamic behavior of the Polytropic gas. K. Karami et al. investigated the interaction between the Polytropic gas and cold dark matter and found that the Polytropic gas behaves as the phantom dark energy [18]. K. Karami and S. Ghaffari showed that the generalized second law of thermodynamics is always satisfied by a universe filled with a Polytropic gas and a cold dark matter [19]. K. Kleidis and N.K. Spyron used the first law of thermodynamics in the Polytropic gas model and they show that the Polytropic gas behaves as dark energy and this model leads to a suitable fitting with the observational data about the current expanding era [20]. H. Moradpour, A. Abri and H. Ebadi, investigated the thermo dynamical behavior and stability of the Polytropic gas [21]. M. Salti et al. discussed validity of the first and generalized second law of thermodynamics in locally rotationally symmetric Bianchi-type II space time which is dominated by a combination of Polytropic gas and baryonic matter[22]. Moreover, Muzaffer Askin et al. studied the cosmological scenarios of the Polytropic gas dark matter-energy proposal in a Friedmann-Robertson- Walker universe and they found an exact expression for the energy density of the Polytropic gas model according to the thermo dynamical point of views and a relationship between a homogeneous minimally coupled scalar field and the Polytropic gas [23].This paper is organized as follows: in section 2 we construct the basic thermodynamic formalism of the Polytropic gas model and discuss the thermodynamic behavior of this model. Finally in section 3 we provide a brief discussion.

2. Basic Formalism

In this work, we consider the following equation of state which is well known as Polytropic gas equation of state

$$P = K\rho^{1+\frac{1}{n}} \qquad (1)$$

Here $K(> 0)$ and $n(< 0)$ are Polytropic constant and Polytropic index respectively. Moreover, P is the pressure and ρ is the energy density of the fluid such that

$$\rho = \frac{U}{V} \qquad (2)$$

Where U and V are the internal energy and volume filled by the fluid respectively.

First of all, we try to find the internal energy U and energy density ρ of the polytropic gas as a function of its volume V and entropy S.

From the general thermodynamics, we have

$$\left(\frac{\partial U}{\partial V}\right)_S = -P \qquad (3)$$

From the equations (1), (2) and (3), we get

$$\left(\frac{\partial U}{\partial V}\right)_S = -K\left(\frac{U}{V}\right)^{1+\frac{1}{n}} \qquad (4)$$

Integrating the equation (4), we get

$$U = (-1)^{-n}\left(KV^{-\frac{1}{n}} + \xi\right)^{-n} \qquad (5)$$

Where the parameter ξ is the constant of integration which may be a universal constant or a function of entropy S only

The equation (5) also can rewrite in the following form

$$U = (-1)^{-n}K^{-n}V\left(1 + \left(\frac{V}{\varepsilon}\right)^{\frac{1}{n}}\right)^{-n} \qquad (6)$$

Where $$\varepsilon = \left(\frac{K}{\xi}\right)^{n} \qquad (7)$$

And it has a dimension of volume.

Therefore, the energy density ρ of the Polytropic gas is

$$\rho = \frac{U}{V} = (-1)^{-n}K^{-n}\left(1 + \left(\frac{V}{\epsilon}\right)^{\frac{1}{n}}\right)^{-n} \qquad (8)$$

When $n < 0$ then equation (8) gives

$$\rho \sim (-1)^{-n}K^{-n}\frac{\varepsilon}{V} \qquad (9)$$

Now we will use these equations to discuss different physical parameters.

a) Pressure:

Using the equation (8) in the equation (1) we get the pressure of the Polytropic gas as a function of entropy S and volume V in the following form

$$P = (-1)^{n+1} K^{-n} \left(1 + \left(\frac{V}{\epsilon}\right)^{\frac{1}{n}}\right)^{-(n+1)} \qquad (10)$$

We can rewrite the equation (10) in the following form

$$P = -\frac{\rho}{1+\left(\frac{V}{\epsilon}\right)^{\frac{1}{n}}} \qquad (11)$$

When $n < 0$ and ε does not diverge then for small volume i.e. at early stage of universe, $V \ll \varepsilon$ ie $\frac{V}{\varepsilon} \ll 1$, we get $\mathrm{P} \simeq 0$, which represents a dust dominated universe. When $n < 0$ and ε does not diverge then for large volume i.e. at late stage of universe, $V \gg \varepsilon$ ie $\frac{V}{\varepsilon} \gg 1$, we get $\mathrm{P} \simeq -\rho$, which indicates an accelerated expansion of the universe.

b) Caloric equation of state:

Now from the equations (8) and (10) we get the caloric equation of state parameter as

$$\omega = \frac{P}{\rho} = -\frac{1}{1+\left(\frac{V}{\epsilon}\right)^{\frac{1}{n}}} \qquad (12)$$

When $n < 0$ and ε does not diverge then for small volume $V \ll \varepsilon$ ie $\frac{V}{\varepsilon} \ll 1$, we get $\omega \simeq 0$ (Dust dominated)

When $n < 0$ and ε does not diverge then for large volume $V \gg \varepsilon$ ie $\frac{V}{\varepsilon} \gg 1$, we get $\omega \simeq -1$ (Cosmological constant)

Thus the equation of state parameter (ω) of the Polytropic gas with $n < 0$ is decreased from $\omega \simeq 0$ (for small volume) to $\omega \simeq -1$ (for large volume). It indicates that the universe expands from the dust dominated era to the current accelerating era.

c) Deceleration parameter:

We get the deceleration parameter of the Polytropic gas with the help of equation (12)

$$q = \frac{1}{2} + \frac{3}{2}\frac{P}{\rho} = \frac{1}{2} - \frac{3}{2}\frac{1}{1+\left(\frac{V}{\epsilon}\right)^{\frac{1}{n}}} \qquad (13)$$

When $n < 0$ and ε does not diverge then for small volume $V \ll \varepsilon$ ie $\frac{V}{\varepsilon} \ll 1$, we get $q > 0$, which correspond to the deceleration universe.
When $n < 0$ and ε does not diverge then for large volume $V \gg \varepsilon$ ie $\frac{V}{\varepsilon} \gg 1$, we get $q < 0$, which correspond to the accelerated universe.

d) Square velocity of sound:

From the equation (11) we get the velocity of sound (V_s) as

$${V_s}^2 = \left(\frac{\partial P}{\partial \rho}\right)_S = -\frac{1}{1+\left(\frac{V}{\epsilon}\right)^{\frac{1}{n}}} \qquad (14)$$

When $n < 0$ and ε does not diverge then for small volume $V \ll \varepsilon$ ie $\frac{V}{\varepsilon} \ll 1$, we get ${V_s}^2 \simeq 0$ .Since velocity of sound is zero in vacuum. Therefore the Polytropic gas behaves like a pressure less fluid at the early stage of the universe. When $n < 0$ and ε does not diverge then for large volume $V \gg \varepsilon$ ie $\frac{V}{\varepsilon} \gg 1$, we get ${V_s}^2 \simeq -1$, which gives an imaginary speed of sound leading to a perturbation cosmology.

e) Thermodynamic stability:

The conditions of the thermodynamic stability of a fluid are

$$\left(\frac{\partial P}{\partial V}\right)_S < 0 \qquad (15)$$

And $$C_V > 0 \qquad (16)$$

Here C_V is the thermal capacity at constant volume. From the equation (10) we have

$$\left(\frac{\partial P}{\partial V}\right)_S = -\left(1+\frac{1}{n}\right)\frac{P}{V}\frac{1}{1+\left(\frac{V}{\epsilon}\right)^{-\frac{1}{n}}} \qquad (17)$$

If $-1 < n < 0$ and $\varepsilon < 0$ then from (17), we have

$$\left(\frac{\partial P}{\partial V}\right)_S < 0$$

Thus the stability condition (15) of thermodynamics is satisfied.

Now we have to verify the positivity of the thermal capacity at constant volume C_V where $C_V = T\left(\frac{\partial S}{\partial T}\right)_V \qquad (18)$

Now we determine the temperature T of the Polytropic gas as a function of its entropy S and its volume V. The temperature T of the Polytropic gas is determined from the relation

$$T = \left(\frac{\partial U}{\partial S}\right)_V \tag{19}$$

Using (6) in (19) we get

$$T = (-1)^{n+1} V^{1+\frac{1}{n}} \left(K + \xi V^{\frac{1}{n}}\right)^{-(n+1)} \frac{d\xi}{dS} \tag{20}$$

This gives the temperature of the Polytropic gas.
We can rewrite the equation (20) in the following form

$$T = -n \frac{\rho V^{1+\frac{1}{n}}}{1+\left(\frac{V}{\epsilon}\right)^{\frac{1}{n}}} \frac{d\xi}{dS} \tag{21}$$

From (5) we have

$$[\xi]^{-n} = [U] \tag{22}$$

Since $$[U] = [TS] \tag{23}$$

Therefore from the equations (22) & (23) we get

$$\xi = [U]^{-\frac{1}{n}} = [T_* S]^{-\frac{1}{n}} \tag{24}$$

Where T_* (> 0) is a universal constant with temperature dimension.
Differentiating (24) with respect to 'S' we get

$$\frac{d\xi}{dS} = -\frac{1}{n} T_*^{-\frac{1}{n}} S^{-\frac{1}{n}-1} \tag{25}$$

Using (8) & (24) in (25) we get

$$T = (-1)^n V^{1+\frac{1}{n}} \left(T_*^{-\frac{1}{n}} S^{-\frac{1}{n}-1}\right) \left[K + T_*^{-\frac{1}{n}} S^{-\frac{1}{n}} V^{\frac{1}{n}}\right]^{-(n+1)} \tag{26}$$

This leads to the entropy of the Polytropic gas as

$$S = \left[(-1)^{\frac{n}{n+1}} \left(\frac{T_*}{T}\right)^{\frac{1}{n+1}} - 1\right]^n \quad \frac{V}{K^n T_*} \tag{27}$$

We know that entropy (S) of a thermo dynamical system should be positive ie $S > 0$ [24]
Here $S > 0$ if $K^n T_* > 0$
Now the thermal capacity at constant volume is

$$C_V = T \left(\frac{\partial S}{\partial T}\right)_V$$

$$= (-1)^{\frac{2n+1}{n+1}} \left(\frac{n}{n+1}\right) \frac{S}{\left[(-1)^{\frac{n}{n+1}} \left(\frac{T_*}{T}\right)^{\frac{n}{n+1}} - 1\right]} \left(\frac{T_*}{T}\right)^{\frac{1}{n+1}} \tag{28}$$

Therefore, the condition $C_V > 0$ is satisfied if $K^n T_* > 0$. Thus both the conditions of thermo dynamic stability are satisfied. So the Polytropic gas is thermo dynamically stable.

3. Discussion

We have studied the thermo dynamical behavior of the Polytropic gas. Here, we have considered the value of $n < 0$ to study the whole work done in this article. Some important results are given below:

(i) As we have considered $n < 0$, the pressure goes more and more negative as volume increases.

(ii) The equation of state parameter (ω) of the Polytropic gas is $\omega \simeq 0$ at early stage of the universe and $\omega \simeq -1$ at late stage of the universe. This indicates that the universe expands from the dust dominated era to the present accelerated era.

(iii) The deceleration parameter (q) is investigated in the context of thermodynamics as well as Polytropic gas and our analysis shows that universe is decelerated ($q > 0$) at early stage of the universe and accelerated ($q < 0$) at late stage of the universe.

Both the conditions of the thermo dynamical stability of the Polytropic gas are studied for $K^n T_* > 0$ and our analysis shows that the Polytropic gas is thermodynamically stable.

References

[1] A.G. Riess, et al. Type Ia Supernova Discoveries at z > 1 from the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution. Astro phys. J., 607,665-687, 2004.

[2] A.G. Riess, et al. New Hubble space telescope discoveries of type Ia supernovae at z≥ 1: narrowing constraints on the early behavior of dark energy. Astrophys. J., 659, 98–121, 2007.

[3] R. R. Caldwell and M. Doran. Cosmic microwave background and supernova constraints on quintessence: concordance regions and target models. Physical Review D, Vol-69, No.10, Article ID103517, 2004.

[4] T. Koivisto and D. F. Mota. Dark energy anisotropic stress and large scale structure formation. Physical Review D, Vol- 73, No. 8, Article ID 083502, 2006.

[5] S. F. Daniel, R. R. Caldwell, A. Cooray, and A. Melchiorri. Large scale structure as a probe of gravitational slip. Physical Review D, Vol- 77, No. 10, Article ID 103513, 2008.

[6] D. N. Spergel, et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP)*Observations: Determination of Cosmological Parameters. *Ap J S*, 148, 135-159, 2003.

[7] M.Tegmark, et al. Cosmological parameters from SDSS and WMAP. Physical Review D, 69, 103501, 2004.
[8] R. Kalita. Dark Energy. Journal of Modern Physics, 6, 1007-1011, 2015.
[9] E.J. Copland, M. Sami and S. Tsujikawa. Dynamics of Dark energy. Int J Mod. Phys D, 15(11), 1753-1935, 2006.
[10] S. Weinberg. The cosmological constant problem. Reviews of Modern Physics. 61(1),1989.
[11] C. Wetterich. Cosmology and the Fate of Dilatation Symmetry. Nuclear Physics B, 302, 668-696, 1988.
[12] R.R. Caldwell. A Phantom Menace? Cosmological Consequences of Dark Energy Component with Super Negative Equation of State. Physics Letters B, 545, 23-29, 2002.
[13] M.R. Setare, J. Sadeghi and A.R. Amani. Interacting Tachyon Dark Energy in non flat universe. Physics Letters B, 673, 241-246, 2009
[14] M.R. Setare. Holography Chaplygin Gas Model. Physics Letters B, 648, 329-332, 2007.
[15] M.R. Setare. Holographic Chaplygin DGP Cosmologies. International Journal of Modern Physics D, 18, 419-427, 2009.
[16] N. Afshordi, D.J. Chung, and H. Geshnizjani. Casual Field Theory with an Infinite Speed of Sound. Physical Review D, Vol-75, Article ID: 083513, 2007.
[17] U. Mukhopadhyay, S. Ray, S.B. Dutta Choudhury. Dark energy with Polytropic equation of state. Mod. Phys. Lett. A, 23, 3187-3198, 2008.
[18] K. Karami, S. Ghaffari, J. Fehri. Interacting Polytropic gas model of phantom dark energy in non-flat universe. Eur. Phys. J. C., 64, 85-88, 2009.
[19] K. Karami, S. Ghaffari. The generalized second law of thermodynamics for the interacting Polytropic dark energy in non-flat FRW universe enclosed by the apparent horizon. Phys. Letters B, 688, 125-128, 2010.
[20] K. Kleidis and N. K. Spyrou. Polytropic dark matter flows illuminate dark energy and accelerated expansion. Astronomy & Astrophysics, Vol-576, No.A23, 2015.
[21] H. Moradpour, A. Abri and H. Ebadi. Thermodynamic behavior and stability of Polytropic gas. International Journal of Modern Physics D, Vol-25, No.01, 1650014, 2016.
[22] M. Salti, L. Acikgoz and H. Abedim. Polytropic Gas and Gravitational Thermodynamics. Chinese Journal of Physics, 52, 982-991, 2014.
[23] Muzaffer Askin, Mustafa Salti and Oktay Aydogdu. Cosmology via Thermodynamics of Polytropic gas. Modern Physics Letter A, Vol-32, No. 32, 1750177, 2017.
[24] H.B.Callen. Thermo dynamics and thermo statics. Newwork. John Wiley and Sons, 1985.

Ratio Mathematica, 39, 2020

Contents

Published by Accademia Piceno - Aprutina dei Velati in Teramo (A.P.A.V.)