Bulletin of the Iranian Mathematical Society
Vol. 28, No. 1, pp 1-28 (2002)

ON THE MAXIMAL DEGREE OF THE K-STEP OBRECHKOFF'S METHOD

V.R. Ibrahimov

Department of Mathematics of Computation, Baku State University, Faculty of Mechanics and Mathematics, Baku, Z. Khalilov 23, Azerbaijan Republic
ibvag@yahoo.com

Abstract: In this paper the numerical solution of the Cauchy problem for the ordinary differential equations of arbitrary order is considered. In this regard, the k-step Obrechkoff's method is investigated. In a recent work, the maximal value degree for the k-step Obrechkoff's method was found and the natural conditions on its coeffiecients were defined. Taking this result into account, the convergence of the multi-step method depends on its stability. Here we define maximal value of the degree for stable and nonstable k-step Obrechkoff's method of explicit, implicit and forward-jumping types. These results are developments of some results due to G. Dahlquist, Iserles and Norest.

[0] *MSC:* 65L20.
[0] *Keywords and phrases:* Maximal value, Forward-jumping method, K-step Obrechkoff's method, Cauchy problem.
[0] *Received:* 7 October, 2000; *Revised:* 15 December, 2001.

Introduction.

As it is known, in solving many applied problems, usually there appears the necessity of finding the solution of the Cauchy problem for the ordinary differential equations. For this aim either one or multi-step methods or their combinations are used. One of the basic questions for their using is in determination of their accuracy. This question is answered in the work of N.S.Bakhvalov, for the explicit stable k-step method with the constant coefficients for $k \leq 10$ (see [1]) and in Dahlquist's work for the implicit stable k-step method with the constant coefficients and for the stable explicit method, when k is arbitrary. As k-step method is applied to the numerical solution of the first order ODE, but for the numerical solution of ODE of second order usually it is used k-step method with the second derivative. The maximal accuracy of the stable k-step method with the second derivative is determined in [3]. This result was obtained in [4] by different ways. Note that the k-step methode with the second derivative, as the numerical method for the solution of the ODE is investigated by many authors (see, for example [5], [6]).

Works, devoted to the numerical solution of ODE of orders more than 2 are considerably few. So the k-step Obrechkoff's method which can be used, as the numerical method for solving any order ODE is investigated here.

The k-step Obrechkoff's method with the constant coefficients may be written as:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{j=1}^{r} h^j \sum_{i=0}^{k} \beta_i^{(j)} y_{n+i}^{(j)}. \qquad (1)$$

This method for $r = 2$ was investigated in [7] and maximal value of the degree for the A-stable methods is found there. It is evident that the method (1) can be applied for determination of numerical solution of the problem:

$$\begin{cases} y^{(j)} = f(x, y, y', \ldots, y^{(j-1)}), \qquad (j = 1, 2, \ldots, r), \\ y(x_0) = y_0, y^{(v)}(x_0) = y_0^{(v)} \quad (v = 1, \ldots, j-1). \end{cases} \tag{2}$$

It is easy to show, that if $j > 1$, then for determination of the numerical solution of the problem (2), any method of the type (1) can be used and in this case it is necessary to define the solution of the system of the difference equations (hence using some methods for the calculation $y'(x), y''(x), \ldots, y^{(j-1)}(x)$ on the point $x_m (m > 0)$). Convergence and effectiveness of such methods are investigated in [8].

In order to determine the maximal accuracy of the stable method, which is obtained from (1) all over again, one can define the maximal accuracy of the method (1), regardless of its stability.

1. The maximal value of the degree of the k-step Obrechkoff's method.

Usually the concept of the accuracy of a multistep method is concerned with the concept of its order.

Definition 1. The method (1) is said to have the degree p, if for any smooth function $y(x)$,

$$\sum_{i=0}^{k} \left[\alpha_i y(x + ih) - \sum_{j=1}^{r} h^j \beta_i^{(j)} y^{(j)}(x + hi) \right] = O(h^{p+1}), \quad h \to 0. \tag{1.1}$$

It is not difficult to define, that the maximal order of the accuracy for the method (1), coincides with the maximal value of its degree p (see, for exam. [2], [3]). Therefore we shall make busy ourselves with the determination of the maximal value of the degree p, both for stability and for nonstability method, which is received from (1). Consider the next lemma.

Lemma. *Let $y(x)$ be a sufficiently smooth function. Then for implementing relation (1.1), the necessary and sufficient condition is the following:*

$$\begin{cases} \sum_{i=0}^{k} \alpha_i = 0 \ ; \ \sum_{i=0}^{k} \frac{i^v}{v!}\alpha_i = \sum_{j=1}^{v} \sum_{i=0}^{k} \frac{i^{j-1}}{(j-1)!}\beta_i^{(v+1-j)} \ , \ (v = 1, \dots, r), \\ \sum_{i=0}^{k} \frac{i^{r+l}}{(r+l)!}\alpha_i = \sum_{j=1}^{r} \sum_{i=0}^{k} \frac{i^{r+l-j}}{(r+l-j)!}\beta_i^{(j)} \ , \ (l = 1, \dots, p-r). \end{cases} \tag{1.2}$$

For the proof of this lemma, it is sufficient to use the following expansions in (1.1)

$$y(x+ih) = y(x) + \sum_{v=1}^{p} \frac{(ih)^v}{v!} y^{(v)}(x) + O(h^{p+1}),$$

$$y^{(j)}(x+ih) = y^{(j)}(x) + \sum_{v=1}^{p-j} \frac{(ih)^v}{v!} y^{(v+j)}(x) + O(h^{p-j+1}), \ \ (j = 1, 2, \dots, p)$$

and linear independence of the system $1, h, h^2, \dots, h^p$.

The number of equations in (1.2) is equal to $p+1$, but the number of unknowns is equal to $(r+1)(k+1)$. In order for the system (1.2) to have the non-trivial solution, it must be $p+1 < rk+r+k+1$. Hence, $p \le r(k+1)+k-1$. From here it follows that $p_{\max} = r(k+1)+k-1$. But the methods with the maximal orders usually are nonstable.

Definition 2. The formula (1) is called stable, if the modulus of roots of the polynomial

$$\rho(\lambda) \equiv \sum_{i=0}^{k} \alpha_i \lambda^i$$

does not exceed 1, and that the roots of modulus 1 (one) are simple. The method is called stable, if the corresponding formula is stable.

Prior to defining the maximal value of the degree for the stable method, which is received from (1), we shall consider the natural conditions, that we put on coefficients of the formula (1). Suppose that the coefficients of the formula (1) satisfy the following assumptions everywhere.

A. The coefficients $\alpha_i, \beta_i^{(j)}$ $(i = 0, \ldots, k; j = 1, \ldots, r)$ are real and $\alpha_k \neq 0$.

B. The characteristic polynomials $\rho(\lambda)$ and

$$\nu_j(\lambda) \equiv \sum_{i=0}^{k} \beta_i^{(j)} \lambda^i, \quad (j = 1, \ldots, r)$$

have no common factor.

C. The degree of formula (1) satisfies the condition $p \geq r$ and $\nu_r(1) \neq 0$, if $\nu_1(1) = \cdots = \nu_{r-1}(1) = 0$, otherwise $\nu_1(1) \neq 0$ and $p \geq r$.

The necessity of the condition A is evident. Consider, the proof of the necessity of the assumptions B.

Assume that the polynomials $\nu_j(\lambda)$ and $\rho(\lambda)$ have a common factor, degree of which is not less than 1 (one). Then we can write

$$\psi(E)\,(\rho^*(E)y_n - \sum_{j=1}^{r} h^j \nu_j^*(E) y_n^{(j)}) = 0, \tag{1.3}$$

where $\rho(\lambda) \equiv \psi(\lambda)\rho^*(\lambda)$; $\nu_j(\lambda) \equiv \psi(\lambda)\nu_j^*(\lambda)$, $(j = 1, \ldots, r)$, but E is the operator defined by

$$Ey_n = y_{n+1} \quad \text{or} \quad Ey(x) = y(x+h).$$

From (1.3) we have

$$\rho^*(E)y_n = \sum_{j=1}^{r} h^j \nu_j^*(E) y_n^{(j)}. \tag{1.4}$$

It follows from here, that the formula (1) as a difference equation with the order k is equivalent to the difference (1.4) with the order k_1, where $k_1 < k$, i.e. to the difference equation with the lower order. Consequently, for the determination of unique solution of the difference equation (1), it is sufficient to assign the initial values on the first k_1

points. But, as it is known, in this case the solution of difference equation of order $k > k_1$, will be nonunique. The contradiction, which has been obtained, demonstrates the necessity of the condition B.

Now let us prove the necessity of the condition C. Suppose, that the method which is determined by the formula (1) is convergent. Then we can write

$$|y_{n+i} - y(x)| \to 0 \text{ when } h \to 0 \; (x = x_0 + nh), \tag{1.5}$$

here $y(x)$ is exact and y_n is approximate values of the solution of problem (2), calculated by the method, which is determined from the formula (1).

If we substitute (1.5) in (1), then we will have

$$|y(x)| \cdot |\sum_{i=0}^{k} \alpha_i| \le \varepsilon \sum_{i=0}^{k} |\alpha_i| + O(h). \tag{1.6}$$

Taking this into account $y(x) \not\equiv 0$ and going over to the limit in (1.6) when $h \to 0$, we obtain: $\rho(1) = 0$. From the $\rho(1) = 0$ we can write $\rho(\lambda) = (\lambda - 1)\rho'(\lambda)$. Then from (1) we receive

$$\rho'(E)(y_{i+1} - y_i) - h\nu_1(E)y_i' = O(h^2). \tag{1.7}$$

Summarizing (1.7) over i from 0 to n, we have

$$\rho'(E)(y_{n+1} - y_0) = \nu_1(E) \sum_{i=0}^{n} hy_i' + O(h). \tag{1.8}$$

Put $F_n = \sum\limits_{i=0}^{n} hy_i'$ and consider $y_{n+i} \to y(x)$, $y_i \to y(x_0)$.

$$F_{n+i} \to \int_{x_0}^{x} y'(s)ds \quad (i = 0, 1, \ldots, k).$$

Hence

$$\rho'(1)(y(x) - y_0) = \nu_1(1) \int_{x_0}^{x} y'(s)ds. \tag{1.9}$$

Consequently $\rho'(1) = \nu_1(1)$.

For $\psi(\lambda) \equiv 1$ taking into account in the correlation (1.3)

$$\rho(\lambda) = \rho(1) + \rho'(1)(\lambda - 1) + \frac{1}{2}\rho''(1)(\lambda - 1)^2 + O((\lambda - 1)^3),$$
$$\nu_1(\lambda) = \nu_1(1) + \nu_1'(1)(\lambda - 1) + O((\lambda - 1)^2),$$
$$\nu_2(\lambda) = \nu_2(1) + O(\lambda - 1).$$

and $\rho(1) = 0$, $\rho'(1) = \nu_1(1)$, and also $\dfrac{y_{i+1} - y_i}{h} = y_i' + \dfrac{hy_i''}{2} + O(h^2)$ we can write

$$\frac{1}{2}\rho''(1)\left(\frac{y_{i+2} - y_{i+1}}{h} - \frac{y_{i+1} - y_i}{h}\right) - \nu_1'(1)(y_{i+1}' - y_i') - h\left(\nu_2(1) - \frac{h}{2}\rho'(1)\right) y_i'' = O(h^2).$$

Hence

$$\rho''(1)(y_{i+1}' - y_i') - 2\nu_1'(1)(y_{i+1}' - y_i') - h(2\nu_2(1) - \nu_1(1))y_i'' + \frac{h}{2}\rho''(1)(y_{i+1}'' - y_i'') = O(h^2).$$

Summing up the last correlation over i from 0 to n, we obtain

$$(\rho''(1) - 2\nu_1'(1))(y_{i+1}' - y_0') = (2\nu_2(1) - \nu_1(1))\sum_{i=0}^{n} hy_i'' - \frac{h}{2}\rho''(1)(y_{n+1}'' - y_0'') + O(h). \tag{1.10}$$

Going over to the limit in (1.10), when $h \to 0$, we have

$$(\rho''(1) - 2\nu_1'(1))(y'(x) - y_0') = (2\nu_2(1) - \rho'(1))\int_{x_0}^{x} y''(s)ds. \tag{1.11}$$

Hence

$$\rho''(1) = 2\nu_1'(1) + 2\nu_2(1) - \nu_1(1). \tag{1.12}$$

Taking into account that $\rho'(1) = \nu_1(1)$, the correlation (1.12) can be written in the next form:

$$\rho''(1) + \rho'(1) = 2\nu_1'(1) + 2\nu_2(1). \tag{1.13}$$

Using the expansion

$$\begin{aligned}
\rho(\lambda) &= \rho(1) + \rho'(1)(\lambda - 1) + \frac{1}{2}\rho''(1)(\lambda - 1)^2 + \frac{1}{6}\rho'''(1)(\lambda - 1)^3 + O((\lambda - 1)^4), \\
\nu_1(\lambda) &= \nu_1(1) + \nu_1'(1)(\lambda - 1) + \frac{1}{2}\nu_1''(1)(\lambda - 1)^2 + O((\lambda - 1)^3), \\
\nu_2(\lambda) &= \nu_2(1) + \nu_2'(1)(\lambda - 1) + O((\lambda - 1)^2), \\
\nu_3(\lambda) &= \nu_3(1) + O(\lambda - 1), \\
\frac{y_{i+1}^{(j)} - y_i^{(j)}}{h} &= y_i^{(j+1)} + \frac{h}{2}y_i^{(j+2)} + \frac{h^2}{6}y_i^{(j+3)} + O(h^3) \ \ (j = 0, 1)
\end{aligned}$$

in the next expression

$$\rho(E)y_i - h\nu_1(E)y_i' - h^2\nu_2(E)y_i'' - h^3\nu_3(E)y_i''' = O(h^4) \quad (1.14)$$

we receive

$$\begin{aligned}
&\rho'(1)(y_{i+1} - y_i) + \frac{h}{2}\rho''(1)\left(\frac{y_{i+2} - y_{i+1}}{h} - \frac{y_{i+1} - y_i}{h}\right) \\
&+ \frac{h^2}{6}\rho'''(1)\left(\frac{y_{i+3} - y_{i+2}}{h} - 2\frac{y_{i+2} - y_{i+1}}{h} + \frac{y_{i+1} - y_i}{h}\right) \\
=&h\nu_1(1)y_i' + h\nu_1'(1)(y_{i+1}' - y_i') + \frac{h^2}{2}\nu_1''(1)\left(\frac{y_{i+2} - y_{i+1}}{h} - \frac{y_{i+1} - y_i}{h}\right) \\
&+ h^2\nu_2(1)y_i'' + h^2\nu_2'(1)(y_{i+1}'' - y_i'') + h^3\nu_3(1)y_i''' + O(h^4).
\end{aligned}$$

If we use $\rho'(1) = \nu_1(1)$ and correlation (1.13), then we can write

$$\rho'''(1) + 3\rho''(1) + \rho'(1) = 3\nu_1''(1) + 3\nu_1'(1) + 6\nu_2'(1) + 6\nu_3(1). (1.15)$$

Note, that by realizations of the correlations (1.13) and (1.15) we can obtain $p = 2$ and $p = 3$, respectively.

It is obvious, that if we continue this process, before using in (1.14) the expansion of $y_i^{(r)}$, then we will receive correlation similar to (1.13) and (1.15), from which the relation $p = r$ will follow.

Note, that for the receiving of above mentioned correlation we can use

$$\rho(\lambda)(\ln\lambda)^{-1} = \frac{\rho(\lambda)}{\lambda-1}\sum_{i=0}^{\infty} C_i(\lambda-1)^i,$$

$$\nu_1(\lambda) + \sum_{j=1}^{r-1}\nu_j(\lambda)(\ln\lambda)^j = \frac{\rho(\lambda)}{\lambda-1}\sum_{j=0}^{p-1} C_j(\lambda-1)^j + O((\lambda-1)^p), \quad \lambda\to 1,$$

here

$$C_m = \sum_{i=1}^{m}(-1)^{i-1}\frac{C_{m-i}}{i+1}.$$

If the last condition is satisfied then the formula (1) will have the degree equal to p.

The second part of the assumption C, is concerned with the fact that if $\nu_1(\lambda)\not\equiv 0$ and $\nu_1(1)=0$, then as it is obvious from the (1.9), method will be divergent, which contradicts to the assumption.

But if we consider the case $\nu_1(\lambda)\equiv 0, \nu_2(\lambda)\not\equiv 0$, and $\nu_2(1)=0$ then as it is obvious from (1.11) convergence of the considered method will be absent, since values of the function $y'(x)$ are involved into the method. Other cases may be explained by analogy.

Note, that when the second part of the assumptions C is not realized, then the method will be nonstable, what proves validity of the above given reasons.

Frequently there arises the necessity to find relation between k and r, i.e. between order of the k-step method (1) and order of the derivatives of the function $y(x)$, used in (1). For the determination of the relation between k and r, the formula can be written in the next form:

$$\sum_{i=0}^{k}\alpha_i y_{n+i} = \sum_{j=1}^{r}\delta_j h^j \sum_{i=0}^{k}\beta_i^{(j)} y_{n+i}^{(j)}, \tag{1.16}$$

here $\delta_j (j=1,\ldots,r)$ takes values 0 (zero) or 1 (one).

It is clear that if $\delta_1 = 0$, then the method, which is determined by the formula (1.16) cannot be stable. Therefore the notion of l-stability introduced by G.Dahlquist will be used (see [3 p.19]), for $\delta_1=\delta_2=\cdots=\delta_{l-1}=0$ and $\delta_1 = 1$.

Definition 3. Formula (1.16) is said to be l-stable, if the roots of polynomials $\rho(\lambda)$ are located within or on the unit circle and there is not multiple root on the unit circle, except $\lambda = 1$ multiplicity to l.

The method is called l-stable, if the corresponding formula is l-stable. Relation between k and r may be written in the next form:

$$\sum_{j=1}^{r} \delta_j k \geq r \quad \text{or} \quad k \geq \frac{r}{\sum_{j=1}^{r} \delta_j}.$$

If we consider the case $\delta_1 = 0$ and $\delta_2 \neq 0$ then we receive well-known method of Shtermer. In this case $k \geq 2$. Now, we can consider the maximal value of degree for the stable methods, received from the correlation (1).

2. The maximal value of degree for the stable k-step Obrechkoff's method.

For the investigation of the maximal value of degree for the stable k-step Obrechkoff's method, consider in general form, i.e. not taking into account property of explicitly of the considered method, which imposes some limitation on coefficients $\beta_k^{(l)}$ $(l = 1, \ldots, r)$. In general, property of explicitly for formula (1) depends on its application. In particular, if the formula (1) is applied to numerical solution of problem (2), then for $j = r$ formula will be explicit by $\beta_k^{(r)} = 0$ but for $j = 1$ formula will be explicit by $\beta_k^{(l)} = 0$, $(l = 1, \ldots, r)$.

Suppose, that $|\beta_k^{(1)}| + |\beta_k^{(2)}| + \cdots + |\beta_k^{(r)}| \neq 0$ and we shall now prove a theorem, by which relation between p, k, and r can be determined.

Theorem 1. *Suppose, that the formula (1) has the degree p, stable $\alpha_k \neq 0$. Then*

$$p \leq \begin{cases} (k+1)r + 1 & \textit{by even } k \textit{ and odd } r, \\ (k+1)r & \textit{by odd } k \textit{ and even } r. \end{cases}$$

There exists stable formula with the degree $p = (k+1)r+1$ in the case, when k is even and r is odd, but with the degree $p = (k+1)r$ in the other cases, for arbitrary k.

Proof. Taking into account condition of theorem 1, we can write

$$\rho(E)y_n - \sum_{j=1}^{r} h^j \nu_j(E) y_n^{(j)} \sim Ch^{p+1} y_n^{(p+1)} \quad (h \to 0). \tag{2.1}$$

Consider the special case and we put $y(x) = \exp(x)$ (see [2]). Denote by the $\tau = \exp(h)$. Then correlation (2.1) may be written in the next form:

$$\rho(\tau) - \sum_{j=1}^{r} \nu_j(\tau)(\ln \tau)^j \sim C(\tau - 1)^{p+1} \quad (\tau \to 1). \tag{2.2}$$

Replacing by

$$\tau = \frac{(z+1)}{(z-1)} \quad , \quad z = \frac{(\tau+1)}{(\tau-1)}.$$

To use the next notation

$$R(z) = \left(\frac{1}{2}(z-1)\right)^k \rho(\tau) \equiv \sum_{i=0}^{k} a_i z^i,$$

$$S_l(z) = \left(\frac{1}{2}(z-1)\right)^k \nu_1(\tau) \equiv \sum_{i=0}^{k} b_i^{(l)}, z^i \quad (l = 1, \ldots, r),$$

in (2.2) we have, that

$$R(z) - \sum_{j=1}^{r} S_j(z) \left(\ln \frac{z+1}{z-1}\right)^j \sim C \left(\frac{2}{z}\right)^{p-k+1} , \ z \to \infty.$$

From here we can write

$$R(z) \left(\ln \frac{z+1}{z-1}\right)^{-1} - S_1(z) - \sum_{j=2}^{r} S_j(z) \left(\ln \frac{z+1}{z-1}\right)^{j-1} \sim C \left(\frac{2}{z}\right)^{p-k} , \ z \to \infty. \tag{2.3}$$

Considering the following equalities

$$\left(\ln \frac{z+1}{z-1}\right)^{-1} = \frac{z}{2} - \sum_{i=0}^{\infty} \mu_{2i+1} z^{-(2i+1)} \quad (\mu_{2i+1} > 0), \quad \ln \frac{z+1}{z-1} = 2 \sum_{i=0}^{\infty} \frac{z^{-(2i+1)}}{2i+1},$$

in (2.3), then we have

$$R(z)\left(\frac{z}{2}-\sum_{i=0}^{\infty}\mu_{2i+1}z^{-(2i+1)}\right)-S_1(z)-\sum_{v=1}^{r-1}2^v A_v S_{v+1}(z)\ \sim\ C\left(\frac{2}{z}\right)^{p-k}\ ,\ z\to\infty, \tag{2.4}$$

where

$$A_v=\begin{cases}\sum_{s=l-1}^{\infty}C_{2s+1}^{(l)}z^{-(2s+1)} & \text{for}\quad v=2l-1,\\ \sum_{s=l}^{\infty}C_{2s}^{(l)}z^{-2s} & \text{for}\quad v=2l\ (C_m^{(l)}>0,\ m>0).\end{cases}$$

Let the coefficients of the formula (1) satisfy the condition A, B and C, then we can write $a_k=0$, as $\rho(1)=0$.

If the condition of stability from the polynomial $\rho(\lambda)$ carries over to the polynomial $R(\lambda)$, then we have:

1. R(z) has not roots with the positive real parts.

2. R(z) does not have the multiple roots on the imaginary axis The coefficient $a_{k-1}\neq 0$, as $\rho'(1)\neq 0$.

It is clear, that the left-hand side of the relation (2.4) may be written in the next form:

$$R(z)\left(\frac{z}{2}-\sum_{i=0}^{\infty}\mu_{2i+1}z^{-(2i+1)}\right)-S_1(z)-\sum_{v=1}^{r-1}2^v A_v S_{v+1}(z)=\sum_{i=1}^{\infty}C_i z^{-i}. \tag{*}$$

It is easy to determine, that to prove the theorem 1, it will be necessary to investigate consistency of the next system:

$$
\begin{aligned}
& b_k^{(1)} = \frac{1}{2} a_{k-1}, \\
& b_{k-1}^{(1)} + 2C_1^{(1)} b_k^{(2)} = \frac{1}{2} a_{k-2}, \\
& b_{k-2}^{(1)} + 2C_1^{(1)} b_{k-1}^{(2)} + 2^2 C_2^{(1)} b_k^{(3)} = \frac{1}{2} a_{k-3} - \mu_1 a_{k-1}, \\
& b_{k-3}^{(1)} + 2C_1^{(1)} b_{k-2}^{(2)} + 2C_3^{(1)} b_k^{(2)} + 2^2 C_2^{(1)} b_{k-1}^{(3)} + 2^3 C_3^{(2)} b_k^{(4)} = \frac{1}{2} a_{k-4} - \mu_1 a_{k-2}, \qquad (2.5) \\
& \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
& b_0^{(1)} + 2 \cdot \sum_{v=0}^{[\frac{(k+1)}{2}]-1} C_{2v+1}^{(1)} b_{2v+1}^{(2)} + 2^2 \cdot \sum_{v=0}^{[\frac{k}{2}]-1} C_{2v+2}^{(1)} b_{2v+2}^{(3)} \\
& + 2^3 \cdot \sum_{v=1}^{[\frac{(k+1)}{2}]-1} C_{2v+3}^{(2)} b_{2v+3}^{(4)} + \cdots + 2^{r-1} \cdot \sum_{v=[\frac{(r-1)}{2}]}^{\frac{\xi_{k,r}}{2}} C_{2v}^{([\frac{r}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{(k+1)}{2}]-1} a_{2v+1} \mu_{2v+1}
\end{aligned}
$$

with the following system $C_1 = C_2 = \cdots = C_{(r-1)k+r+1} = 0$.

The system (2.5) is received from the $(*)$ by the comparison coefficients of the linearly independent system $z^j \ (j = 0, \ldots, k)$.

Note, that the system (2.5) may be consistent, since the number of the equations and unknowns are identical. (It is not difficult to prove, that the system (2.5) is consistent). Therefore we will investigate the system $C_1 = C_2 = \cdots = C_{(r-1)k+r+1} = 0$, which can be written in the next form:

$$2\sum_{v=0}^{[\frac{k}{2}]} C_{2v+1}^{(1)} b_{2v}^{(2)} + 2^2 \sum_{v=0}^{[\frac{(k+1)}{2}]-1} C_{2v+2}^{(1)} b_{2v+1}^{(3)} + 2^3 \sum_{v=1}^{[\frac{k}{2}]} C_{2v+1}^{(2)} b_{2v}^{(4)} + \dots$$

$$\dots + 2^{r-1} \sum_{v=\frac{(r-2)}{2}}^{\frac{\xi_{k,r}}{2}} C_{2v+1}^{([\frac{k}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{k}{2}]-\xi_k^{(3)}} a_{2v}\mu_{2v+1},$$

$$2\sum_{v=0}^{[\frac{k+1}{2}]-1} C_{2v+3}^{(1)} b_{2v+1}^{(2)} + 2^2 \sum_{v=0}^{[\frac{k}{2}]} C_{2v+2}^{(1)} b_{2v}^{(3)} + 2^3 \sum_{v=0}^{[\frac{(k+1)}{2}]-1} C_{2v+3}^{(2)} b_{2v+1}^{(4)} + \dots$$

$$\dots + 2^{r-1} \sum_{v=\frac{(r-3)}{2}}^{\frac{\xi_{k,r}}{2}} C_{2v+2}^{([\frac{r}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{(k+1)}{2}]-1} a_{2v+1}\mu_{2v+3},$$

....

$$2 \sum_{v=0}^{\frac{(\xi_{k,l}-\xi_l^{(3)})}{2}} C^{(1)}_{2v+l+\xi_l^{(3)}} b^{(2)}_{2v+\xi_l^{(3)}} + 2^2 \sum_{v=0}^{\frac{(\xi_{k,l}-1+\xi_l^{(3)})}{2}} C^{(1)}_{2v+l+1-\xi_l^{(3)}} b^{(3)}_{2v+1-\xi_l^{(3)}} + \dots$$

$$\dots + 2^{r-1} \sum_{v=0}^{\frac{(\xi_{k,l}^{(r)}-\xi_l^{(4)})}{2}} C^{([\frac{r}{2}])}_{2v+l+\xi_l^{(4)}} b^{(r)}_{2v+\xi_l^{(4)}} = - \sum_{v=0}^{[\frac{(k+\xi_l^{(3)})}{2}]-\xi_l^{(3)}-\xi_k^{(3)}} a_{2v+\xi_l^{(3)}} \mu_{2v+l+\xi_l^{(3)}},$$

$$2 \cdot \sum_{v=0}^{\frac{(\xi_{k,l+1}^{(1)}-\xi_{l+1}^{(3)})}{2}} C^{(1)}_{2v+l+1+\xi_{l+1}^{(3)}} b^{(2)}_{2v+\xi_{l+1}^{(3)}} + 2^2 \cdot \sum_{v=0}^{\frac{(\xi_{k,l+1}^{(1)}+1-\xi_{l+1}^{(3)})}{2}} C^{(1)}_{2v+l+2-\xi_{l+1}^{(3)}} b^{(3)}_{2v+1-\xi_{l+1}^{(3)}} + \dots$$

$$\dots + 2^{r-1} \cdot \sum_{v=0}^{\frac{(\xi_{k,l+1}^{(r)}+1-\xi_{l+1}^{(4)})}{2}} C^{([\frac{r}{2}])}_{2v+l+2-\xi_{l+1}^{(3)}} b^{(r)}_{2v+l-\xi_{l+1}^{(4)}} = - \sum_{v=0}^{[\frac{(k+\xi_{l+1}^{(3)})}{2}]-\xi_{l+1}^{(3)}-\xi_k^{(3)}} a_{2v+\xi_{l+1}^{(3)}} \mu_{2v+l+1+\xi_{l+1}^{(3)}}.$$

The system which is received for $k = 2i$ may be divided upon two subsystems. The first of them can be written in the next form:

$$
\begin{gathered}
2\sum_{v=0}^{\frac{k}{2}} C^{(1)}_{2v+1} b^{(2)}_{2v} + 2^2 \cdot \sum_{v=0}^{\frac{(k-2)}{2}} C^{(1)}_{2v+2} b^{(3)}_{2v+1} + \dots \\
\dots + 2^{r-1} \sum_{v=\frac{(r-2)}{2}}^{\frac{\xi_{k,r}}{2}} C^{([\frac{r}{2}])}_{2v+1} b^{(r)}_{2v} = -\sum_{v=0}^{\frac{(k-2)}{2}} a_{2v}\mu_{2v+1},
\end{gathered}
\tag{2.6}
$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$
\begin{gathered}
2\sum_{v=0}^{\frac{k}{2}} C^{(1)}_{2v+l+\xi^{(3)}_r} b^{(2)}_{2v} + 2^2 \cdot \sum_{v=0}^{\frac{(k-2)}{2}} C^{(1)}_{2v+l+1+\xi^{(3)}_r} b^{(3)}_{2v+1} + \dots \\
\dots + 2^{r-1} \sum_{v=\frac{(r-2)}{2}}^{\frac{\xi_{k,r}}{2}} C^{([\frac{r}{2}])}_{2v+l+\xi^{(4)}_r} b^{(r)}_{2v} = -\sum_{v=0}^{\frac{(k-2)}{2}} a_{2v}\mu_{2v+l+2+\xi^{(3)}_r}.
\end{gathered}
$$

The second subsystem can be written in the next form:

$$
\begin{gathered}
2\sum_{v=0}^{\frac{(k-2)}{2}} C^{(1)}_{2v+3} b^{(2)}_{2v+1} + 2^2 \cdot \sum_{v=0}^{\frac{k}{2}} C^{(1)}_{2v+2} b^{(3)}_{2v} + \dots \\
\dots + 2^{r-1} \sum_{v=\frac{(r-3)}{2}}^{\frac{\xi_{k,r}}{2}} C^{([\frac{r}{2}])}_{2v+2} b^{(r)}_{2v} = -\sum_{v=0}^{\frac{(k-2)}{2}} a_{2v+1}\mu_{2v+3},
\end{gathered}
\tag{2.7}
$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$
\begin{gathered}
2\sum_{v=0}^{\frac{(k-2)}{2}} C^{(1)}_{2v+l+2-\xi^{(3)}_r} b^{(2)}_{2v+1} + 2^2 \sum_{v=0}^{\frac{k}{2}} C^{(1)}_{2v+l+1-\xi^{(3)}_r} b^{(3)}_{2v} + \dots \\
\dots + 2^{r-1} \sum_{v=\frac{(r-3)}{2}}^{\frac{\xi_{k,r}}{2}} C^{([\frac{r}{2}])}_{2v+l+1-\xi^{(4)}_r} b^{(3)}_{2v} = -\sum_{v=0}^{\frac{(k-2)}{2}} a_{2v+1}\mu_{2v+l+2-\xi^{(3)}_r},
\end{gathered}
$$

where

$$\xi_{k,v}=\begin{cases}\left[\frac{v+1}{2}\right]-\left[\frac{v+2}{2}\right]+2\left[\frac{k+1}{2}\right], & v=2j,\\ \left[\frac{v+2}{2}\right]-\left[\frac{v+1}{2}\right]+2\left[\frac{k}{2}\right], & v=2j-1,\end{cases}$$

$$\xi_{k,v}^{(1)}=\begin{cases}\left[\frac{v}{2}\right]-\left[\frac{v+1}{2}\right]+2\left[\frac{k+1}{2}\right], & v=2j-1,\\ \left[\frac{v+1}{2}\right]-\left[\frac{v}{2}\right]+2\left[\frac{k}{2}\right], & v=2j,\end{cases}$$

$$l=rk-k+r,\quad \xi_j^{(3)}=\left[\frac{j+2}{2}\right]-\left[\frac{j+1}{2}\right],$$

$$\xi_{l+j}^{(4)}=\begin{cases}\xi_{l+j}^{(3)}, & r=2m\ (0\le j<k),\\ 1-\xi_{l+j}^{(3)}, & r=2m-1,\end{cases}\qquad \xi_{k,v}^{(r)}=\begin{cases}\xi_{k,v}, & v=2n,\\ \xi_{k,v}^{(1)}, & v=2n-1.\end{cases}$$

If we prove, that the system (2.6) or (2.7) is not consistent, then we shall receive that the above mentioned system is not consistent.

Consider the first subsystem. In the system (2.6) number of the equations will be equal to $2ij+j-i+1$, if we assume $r=2j$. It is not difficult to show, that in this case number of the unknowns will be equal to $2ij+j-i$. It is easy to show, that the system for $a_{2n}=0\ \ (n=0,\ldots,k-1)$, will be consistent and in this case it has the trivial solution.

Now consider the second subsystem. In the considered case number of the equations in the system (2.7) is equal to $2ij+j-i$, but number of the unknowns is equal to $2ij+j-i-1$. If we consider, that $a_{k-1}=a_{2i-1}\neq 0$, then we shall receive, that the system (2.7) is not consistent. Really, if we solve the system (2.7), then we have:

$$\gamma_3\sum_{i=0}^{\frac{(k-2)}{2}}a_{2i+1}\mu_{2i+3}+\gamma_5\sum_{i=0}^{\frac{(k-2)}{2}}a_{2i+1}\mu_{2i+5}+\ldots$$
$$\cdots+\gamma_{l+2-\xi_r^{(3)}}\sum_{i=0}^{\frac{(k-2)}{2}}a_{2i+1}\mu_{2i+l+2-\xi_r^{(3)}}=0.$$

It should be noted that all the nonzero coefficients $a_i\ \ (i=0,\ldots,k-1)$ have identical sign.

The received relation we write in the next form:

$$\sum_{i=0}^{\frac{(k-2)}{2}} a_{2i+1}\varphi(i) = 0,$$

here

$$\varphi(i) = \int_{-1}^{1} x^{2i+2}\psi_{l-1-\xi_r^{(3)}}(x)\left(\pi^2 + \ln^2\frac{1+x}{1-x}\right)^{-1}dx,$$
$$\psi_{l-1-\xi_r^{(3)}}(x) = \gamma_3 + \gamma_5 x^2 + \cdots + \gamma_{l+2-\xi_r^{(3)}}x^{l-1-\xi_r^{(3)}}.$$

By the following notation

$$F_k(x) = a_1x^2 + a_3x^4 + \cdots + a_{k-1}x^k,$$

it can be written

$$\int_{-1}^{1} F_k(x)\psi_{l-1-\xi_r^{(3)}}(x)\left(\pi^2\ln^2\frac{1+x}{1-x}\right)^{-1}dx = 0.$$

Hence, using parity of the integrant functions and the mean-value theorem we have

$$\int_0^{\xi} F_k(x)\psi_{l-1-\xi_r^{(3)}}(x)dx = 0 \quad \text{or} \quad F_k(\xi)\int_{\xi_1}^{\xi}\psi_{l-1-\xi_r^{(3)}}(x)dx = 0.$$

If we denote by the

$$\Phi_1'(x) = F_k(x)\psi_{l-1-\xi_r^{(3)}}(x); \quad \Phi_2'(x) = \psi_{l-1-\xi_r^{(3)}}(x)$$

and consider $F_k(\xi) \neq 0$, then we shall have

$$\Phi_1(\xi) = 0; \;\; \Phi_2(\xi) = \Phi_2(\xi_1).$$

Then using Rolle's theorem we obtain, that the polynomial $\psi_{l-1-\xi_r^{(3)}}(x)$ by $\Phi_2(\xi) \neq 0$ has $l+1-\xi_r^{(3)}$ roots, what is impossible. If $\Phi_2(\xi) = 0$ then granting $\Phi_2(0) = 0$ we can write

$$\int_0^{\xi}\psi_{l-1-\xi_r^{(3)}}(x)dx = 0.$$

Then

$$\int_0^\xi (F_k(x) - F_k(\xi)\psi_{l-1-\xi_r^{(3)}}(x))dx = \int_0^\xi F_k'(\eta_k)(x-\xi)\psi_{l-1-\xi_r^{(3)}}(x)dx = 0.$$

Hence, it follows that $\Phi_3(\xi) \neq 0$ $(\Phi_3'(x) = (x-\xi)\psi_{l-1-\xi_r^{(3)}}(x))$ or

$$\int_0^\xi (x-\xi)\psi_{l-1-\xi_r^{(3)}}(x)dx = 0.$$

If $\Phi_3(x) \neq 0$, then we have, that the function $(x-\xi)\psi_{l-1-\xi_r^{(3)}}(x)$ has $l-1-\xi_r^{(3)}$ roots and consequently the system (17) is not consistent. But if $\Phi_3(\xi) = 0$, then using the relation $F_k(x) - F_k(\xi) = F_k'(\xi)(x-\xi) + F_k''(\eta_2)\dfrac{(x-\xi)^2}{2}$ and above described procedure, then we can write

$$\Phi_4(\xi) \neq 0 \ \ (\Phi_4'(x) = (x-\xi)^2\psi_{l-1-\xi_r^{(3)}}(x)) \ \text{ or } \ \int_0^\xi (x-\xi)^2\psi_{l-1-\xi_r^{(3)}}(x)dx = 0.$$

Carrying on by the above-described scheme, we have

$$\int_0^\xi (x-\xi)^v\psi_{l-1-\xi_r^{(3)}}(x)dx = 0 \quad (v = 0, 1, 2, \ldots, k)$$

or $\Phi_k(\xi) \neq 0$. Here $\Phi_k'(x) = (x-\xi)^k\psi_{l-1-\xi_r^{(3)}}(x)$. If $\Phi_k(\xi) \neq 0$ then system (2.7) is not consistent. But if $\Phi_k(\xi) = 0$, then using the last relation we can write

$$\int_0^\xi \varphi(x)\psi_{l-1-\xi_r^{(3)}}(x)dx = 0,$$

where $\varphi(x)$ polynomial of the degree which cannot be more than k.

Obviously, that the received correlation was put on any limitation to coefficients γ_j $(j = 3, 5, \ldots, l+2-\xi_r^{(3)})$ which inadmissible, since they are determined by the solving system (2.7). Particularly, if $r = 2$, then we have $l - 1 - \xi_r^{(3)} = k$. Naturally in this connection we may put $\varphi(x) = \psi_k(x)$. It is clear, that the received relation is not correct. Obviously, that the functions $F_k(x)$ and $\varphi(x)$ has the different properties and therefore they can not coincide. Consequently, the system (2.7) is not consistent. Hence we received, that $p-k \leq rk+r-k$ or $p \leq r(k+1)$.

But if we suppose, that a_{2i+1} changes its sign, then maybe $F_k(\xi) = 0$. Naturally in this connection the system (2.7) may be consistent.

Suppose, that $k = 2i$ and $r = 2j - 1$. In this case number of the equations in the systems (2.6) and (2.7) coincides and equals to $2ij + j - 2i$. It is clear, that the system (2.6) may has the trivial solution. Hence one must they consistent it.

Therefore consider the second system, in which the number of the unknowns is equal to $2ij + j - 2i - 1$. Consequently, the system (2.7) is not consistent, since the number of the equation in that system is equal to $2ij + j - 2i$. Thus we received

$$p - k \leq l + 1 = rk + r - k + 1 \text{ or } p \leq r(k+1) + 1.$$

Note, that in the case $k = 2i$ and $r = 2j$ the last equation in the system (2.6) received as the coefficient $z^{-(l+1)}$, since l is even. Therefore the indicated equation can be written in next form:

$$2C_{l+1}^{(1)} b_0^{(2)} + 2C_{l+3}^{(1)} b_2^{(2)} + \dots .$$

In this case the last equation of the system (2.7) can be written in the following form:

$$2C_{l+1}^{(1)} b_1^{(2)} + 2C_{l+3}^{(1)} b_3^{(2)} + \dots .$$

Now consider the case, when $k = 2i - 1$, that is k is odd. Suppose, that $r = 2j$. Then the number of the equations in the system (2.6) will be equal to $2ij - i + 1$. But number of the unknowns is equal to $2ij - i$. Taking into account, that $a_{2r-2} \neq 0$ can be predicated, then the system (2.6) is not consistent. In this case we may show, that the system (2.7) will be consistent. Consequently,

$$p - k \leq l \quad \text{or} \quad p \leq (k+1)r.$$

Using above mentioned scheme we can prove, that also in the case, when $k = 2i - 1$ and $r = 2j - 1$, the system (2.6) is not consistent, but the system (2.7) is consistent. Consequently,

$$p \leq (k+1)r.$$

After the combination of all the above-mentioned cases, we receive statement of the theorem. ■

Thus, we take for granted the theorem 1. Now consider the case, when $\beta_k^{(j)} = 0 \quad (j = 1(1)r)$, that is investigate the formula, used in the problem (2), which is explicit for all the values of the parameter j. The maximal value for the degree of the stable explicit method may be established by the next theorem.

Theorem 2. *Suppose, that the formula (1) is stable for* $\beta_k^{(j)} = 0 \ (j = 1, \ldots, r)$*, has the degree* p *and* $\alpha_k \neq 0$*. Then* $p \leq rk$*. There exist stable formulas with the degree* $p = rk$ *for the arbitrary* k*.*

Proof. Taking here exactly the same way, as in theorem 1, we receive the systems similar to the systems (2.6) and (2.7). ■

It should be noted, that these systems can not have trivial solution, since in this connection it is received, that the unknowns $\beta_i^{(j)} \quad (i = 0, \ldots, k-1, \ 1 \leq j \leq r)$ for every fixed j can be determined from the system, which consists of the $k+1$ equations. It may be proved, that in this case these systems will not be consistent. Consider the case $k = 2i$. Then the system similar to the systems (2.6) and (2.7) can be written in the next form:

$$2\sum_{v=0}^{\frac{k}{2}} C^{(1)}_{2v+1} b^{(2)}_{2v} + 2^2 \sum_{v=0}^{\frac{(k-2)}{2}} C^{(1)}_{2v+2} b^{(3)}_{2v+1} + \ldots$$

$$\cdots + 2^{r-1} \sum_{v=\frac{(r-2)}{2}}^{\frac{\xi_{k,r}}{2}} C^{([\frac{r}{2}])}_{2v+1} b^{(r)}_{2v} = - \sum_{v=0}^{\frac{(k-2)}{2}} a_{2v}\mu_{2v+1},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \tag{2.8}$$

$$2\sum_{v=0}^{\frac{k}{2}} C^{(1)}_{2v+l+1} b^{(2)}_{2v} + 2^2 \sum_{v=0}^{\frac{(k-2)}{2}} C^{(1)}_{2v+l+2} b^{(3)}_{2v+1} + \ldots$$

$$\cdots + 2^{r-1} \sum_{v=\frac{(r-2)}{2}}^{\frac{(\xi^{(r)}_{k,l+1}-\xi^{(4)}_{l+1})}{2}} C^{([\frac{r}{2}])}_{2v+l+1+\xi^{(4)}_{l+1}} b^{(r)}_{2v+\xi^{(4)}_{l+1}} = - \sum_{v=0}^{\frac{(k-2)}{2}} a_{2v}\mu_{2v+l+1},$$

where $l = (r-1)k$. The second system has the following form:

$$2 \cdot \sum_{v=0}^{\frac{(k-2)}{2}} C^{(1)}_{2v+l+1} b^{(2)}_{2v+1} + 2^2 \cdot \sum_{v=0}^{\frac{k}{2}} C^{(1)}_{2v+2} b^{(3)}_{2v} + \ldots$$

$$\cdots + 2^{r-1} \cdot \sum_{v=\frac{(r-3)}{2}}^{\frac{\xi_{k,r}}{2}} C^{([\frac{r}{2}])}_{2v+2} b^{(r)}_{2v} = - \sum_{v=0}^{\frac{(k-2)}{2}} a_{2v+1}\mu_{2v+1},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \tag{2.9}$$

$$2 \cdot \sum_{v=0}^{\frac{(k-2)}{2}} C^{(1)}_{2v+l+1} b^{(2)}_{2v+1} + 2^2 \cdot \sum_{v=0}^{\frac{k}{2}} C^{(1)}_{2v+1} b^{(3)}_{2v} + \ldots$$

$$\cdots + 2^{r-1} \cdot \sum_{v=\frac{(r-3)}{2}}^{\frac{(\xi^{(r)}_{k,l}-\xi^{(4)}_{l})}{2}} C^{([\frac{r}{2}])}_{2v+l+\xi^{(4)}_{l}} b^{(r)}_{2v+\xi^{(4)}_{l}} = - \sum_{v=0}^{\frac{(k-2)}{2}} a_{2v+1}\mu_{2v+l+1}.$$

In the system (2.8) number of the equations is equal to $(r-1)\ i+1$, but the number of the unknowns is independent of the property of parity of r and equals to $(r-1)i$. Consequently, the system (2.8) is not consistent.

In the system (2.9) number of the equalities is equal to $(r-1)i$, but number of the unknowns equals to $(r-1)i-1$, since one of the unknowns is determined by the coefficient z^0. Hence, it follows that

$$p-k \leq l \quad \text{or} \quad p \leq kr \quad (l = k(r-1)).$$

This theorem for $k = 2i-1$ is proved analogously to the case $k = 2i$.

If we apply the theorem to formula (1.16), then we shall receive the following theorem.

Theorem 3. *Suppose that the formula (1.16) is stable, has the degree p and $\alpha_k \neq 0$. Then*

$$p \leq (k+1)\sum_{j=1}^{r} \delta_j + 1, \quad (p_{\max} = (k+1)\sum_{j=1}^{r} \delta_j + 1).$$

There exists stable formula with the degree $p = p_{\max}$ for $k = 2i$ and $r = 2v-1$, but in other cases there exists stable formula with the degree $p = p_{\max} - 1$ and does not exist stable formula with the degree $p > p_{\max} - 1$.

It is not difficult to determine, that if there exists stable formula with the degree $p > (k+1)r+1$, then it must be in the class of the forward-jumping formulas.

Really, if we consider forward-jumping formula in the next form:

$$\sum_{i=0}^{k-m} \alpha_i y_{n+i} = \sum_{j=1}^{r} h^j \sum_{i=0}^{k} \beta_i^{(j)} y_{n+i}^{(j)}, \tag{2.10}$$

then the theorem 1 can be formulated in the next form:

Theorem 4. *Suppose, that the formula (2.10) is stable, has the degree p and $\alpha_{k-m} \neq 0$. Then*

$$p \leq (k+1)r + m.$$

There exist stable forward-jumping formulas with the degree $p = (k+1)r + m - 1$ for $k = 2i \geq 3m$, $r = 2j$ and $k - m = 2v - 1$.

In other cases there exist stable forward-jumping formulas with the degree $p = (k+1)r + m$ for $k \geq 3m$ if the property parity of k and m is identical and for $k \geq 3m+1$ if the property parity of k and m is not identical.

Proof. Behaving here in exactly the same way, as in theorem 1 and multiplying the polynomial $\rho(\tau)$ and $\nu_l(\tau)$ to $\left(\frac{1}{2}(z-1)\right)^{k-m}$, we receive the next system, consistency of which is questionable

$$\sum_{v=1}^{m} d_v^{(1)} \beta_{k-m+v}^{(1)} + 2 \sum_{v=0}^{[\frac{(k-m)}{2}]} C_{2v+1}^{(1)} b_{2v}^{(2)} + 2^2 \sum_{v=0}^{[\frac{(k-m+1)}{2}]-1} C_{2v+2}^{(1)} b_{2v+1}^{(3)} + \ldots$$

$$\cdots + 2^{r-1} \sum_{v=\frac{(r-2)}{2}}^{\frac{\xi_{k-m,r}}{2}} C_{2v+1}^{([\frac{r}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{(k-m)}{2}]} a_{2v} \mu_{2v+1},$$

$$\sum_{v=1}^{m} d_v^{(2)} \beta_{k-m+v}^{(1)} + 2 \sum_{v=-1}^{[\frac{(k-m+1)}{2}]-1} C_{2v+3}^{(1)} b_{2v+1}^{(2)} + 2^2 \sum_{v=0}^{[\frac{(k-m)}{2}]} C_{2v+2}^{(1)} b_{2v}^{(3)} + \ldots$$

$$\cdots + 2^{r-1} \sum_{v=\frac{(r-4)}{2}}^{\frac{\xi_{k-m,r}}{2}} C_{2v+2}^{([\frac{r}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{(k-m+1)}{2}]-1} a_{2v+1} \mu_{2v+3},$$

....

$$\sum_{v=1}^{m} d_v^{(l)} \beta_{k-m+v}^{(1)} + 2 \sum_{v=-\frac{(l-1)}{2}}^{\frac{\xi_{k-m,l}}{2}} C_{2v+l}^{(1)} b_{2v}^{(2)} + 2^2 \sum_{v=-\frac{(l-2)}{2}}^{\frac{\xi_{k-m,l}^{(1)}}{2}} C_{2v+l}^{(1)} b_{2v}^{(3)} + \ldots$$

$$\cdots + 2^{r-1} \sum_{v=\frac{-(l-r+1)}{2}}^{\frac{\xi_{k-m,l}^{(2)}}{2}} C_{2v+l}^{([\frac{r}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{(k-m+\xi_l^{(3)})}{2}]-\xi_l^{(3)}-\xi_{k-m}^{(3)}} a_{2v+\xi_l^{(3)}} \mu_{2v+l+\xi_l^{(3)}},$$

$$\sum_{v=1}^{m} d_v^{(l+1)} \beta_{k-m+v}^{(1)} + 2 \sum_{v=-\frac{l}{2}}^{\frac{\xi_{k-m,l+1}}{2}} C_{2v+l+1}^{(1)} b_{2v}^{(2)} + 2^2 \sum_{v=-\frac{(l-1)}{2}}^{\frac{\xi_{k-m,l+1}^{(1)}}{2}} C_{2v+l+1}^{(1)} b_{2v}^{(3)} + \ldots$$

$$\cdots + 2^{r-1} \sum_{v=-\frac{(l-r)}{2}}^{\frac{\xi_{k-m,l+1}^{(2)}}{2}} C_{2v+l+1}^{([\frac{r}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{(k-m+\xi_{l+1}^{(3)})}{2}]-\xi_{l+1}^{(3)}-\xi_{k-m}^{(3)}} a_{2v+\xi_{l+1}^{(3)}} \mu_{2v+l+1+\xi_{l+1}^{(3)}},$$

where $l = 2m+kr-k+r-1$. This system is divided into two subsystems. The first of them for $k = 2i$ and $r = 2j$ can be written in the next form:

$$\sum_{v=1}^{m} d_v^{(1)}\beta_{k-m+v}^{(1)} + \sum_{v=0}^{[\frac{(k-m)}{2}]} C_{2v+1}^{(1)} b_{2v}^{(2)} + 2^2 \sum_{v=0}^{[\frac{(k-m+1)}{2}]-1} C_{2v+2}^{(1)} b_{2v+1}^{(3)} + \dots$$
$$\dots + 2^{r-1} \sum_{v=\frac{r}{2}}^{[\frac{(k-m)}{2}]} C_{2v+1}^{([\frac{r}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{(k-m)}{2}]} a_{2v}\mu_{2v+1},$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \qquad (2.11)$$
$$\sum_{v=1}^{m} d_v^{(1)}\beta_{k-m+v}^{(1)} + 2 \sum_{v=-\frac{(l-1)}{2}}^{[\frac{(k-m)}{2}]} C_{2v+l}^{(1)} b_{2v}^{(2)} + 2^2 \sum_{v=-\frac{(l-1)}{2}}^{[\frac{(k-m+1)}{2}]-1} C_{2v+l+1}^{(1)} b_{2v+1}^{(3)} + \dots$$
$$\dots + 2^{r-1} \sum_{v=-\frac{l-r+1}{2}}^{[\frac{(k-m)}{2}]} C_{2v+l}^{([\frac{r}{2}])} b_{2v}^{(r)} = - \sum_{v=0}^{[\frac{(k-m)}{2}]} a_{2v}\mu_{2v+l}.$$

Number of the equations in the system (2.11) is equal to $2ij-i+j+m$. But number of the unknowns depends on parity $k-m$. If $k-m=2n$, then number of the unknowns is equal to $2ij-i+j+m$.

Consequently, the system (2.11) may be consistent. But if $k-m = 2n-1$, then quantity of the unknowns in the system (2.11) is equal to $2ij-i+j+m-1$. It is obvious, that the system (2.11) can not have the trivial solution, since in this connection $\beta_{k-m+v}^{(1)} = 0$ $(v = 1, \dots, m)$. It may be proved, that the mentioned system has not trivial solution (which identically different from zero). Consequently, the system (2.11) is not consistent. Then we have

$$p-k+m \le 2m+rk-k+r-1 \quad \text{or} \quad p \le (k+1)r+m-1.$$

If we consider the case $k = 2i$ and $r = 2j-1$, then the last equation in the system (2.11) can be written in the next form.

$$\sum_{v=1}^{m} d_v^{(l+1)}\beta_{k-m+v}^{(1)} + 2C_1^{(1)} b_{-(l-1)}^{(2)} + \dots .$$

In this case number of unknowns in the system (2.11) for $k-m=2n$ or for $k-m=2n-1$ is equal to $2ij+j-2i+m-1$, but the number of equation is equal to $2ij+j-2i+m+1$.

As it was proved above, here we can prove, that the system (2.11) is not consistent. Consequently,

$$p-k+m \leq l+1 \quad \text{or} \quad p \leq (k+1)r+m.$$

Now consider the case $k=2i$ and $r=2j$, when the system (2.11) is consistent. In this case the second subsystem is written in the next form:

$$\sum_{v=1}^{m} d_v^{(2)} \beta_{k-m+v}^{(1)} + \sum_{v=-1}^{[\frac{(k-m+1)}{2}]-1} C_{2v+3}^{(1)} b_{2v+1}^{(2)} + 2^2 \sum_{v=0}^{[\frac{(k-m)}{2}]} C_{2v+2}^{(1)} b_{2v}^{(3)} + \dots$$
$$\dots + 2^{r-1} \sum_{v=\frac{(r-4)}{2}}^{[\frac{(k-m+1)}{2}]-1} C_{2v+3}^{([\frac{r}{2}])} b_{2v+1}^{(r)} = - \sum_{v=0}^{[\frac{(k-m+1)}{2}]-1} a_{2v+1}\mu_{2v+3},$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \tag{2.12}$$
$$\sum_{v=1}^{m} d_v^{(l+1)} \beta_{k-m+v}^{(1)} + 2 \sum_{v=-\frac{(l-1)}{2}}^{[\frac{(k-m+1)}{2}]-1} C_{2v+l+1}^{(1)} b_{2v+1}^{(2)} + 2^2 \sum_{v=-\frac{(l-1)}{2}}^{[\frac{(k-m)}{2}]} C_{2v+l+1}^{(1)} b_{2v}^{(3)} + \dots$$
$$\dots + 2^{r-1} \sum_{v=-\frac{l-r+1}{2}}^{[\frac{(k-m+1)}{2}]-1} C_{2v+l}^{([\frac{r}{2}])} b_{2v+1}^{(r)} = - \sum_{v=0}^{[\frac{(k-m+1)}{2}]-1} a_{2v+1}\mu_{2v+l+2}.$$

Number of the equations m the system (2.12) for $k-m=2n$ is equal to $2ij+j-i+m$, but number of the unknowns is equal to $2ij+j-i+m-1$. We can prove, that the system (2.12) in this case is not consistent. Hence, it follows that

$$p \leq (k+1)r+m.$$

Now consider the case, when the systems (2.11) and (2.12) are consistent. In order for the consistency of the systems (2.11) and (2.12) to be followed by consistency of the initial system, the unknowns $\beta_{k-m+v}^{(1)}(v=$

$1, \ldots, m)$, found from these systems, must be equal to each other, in general this question depends on parity k and m. If k and m are even or odd simultaneously, then degree of the stable forward-jumping formula has the maximal value for $k \geq 3m$ otherwise for $k \geq 3m + 1$. ■

Consider the case, when in the system (1.16) $\delta_1 = 0$. It is clear, that the formula (1.16) can not be stable. Because in this case we use the notation of 2-stability. If to consider the case $\delta_1 = \delta_2 = \cdots = \delta_{l-1} = 0$ and $\delta_l \neq 0$, then we use the notation of l-stability. It is not difficult to prove, that there exists l-stable method determined by the formula (1.16). For this aim consider the next theorem.

Theorem 5. *Let the formula (1.16) has the degree p, is l-stable, $a_k \neq 0$ and $\delta_1 = \cdots = \delta_{l-1} = 0, \delta_1 = 0$. Then there exists l-stable formula with the degree $p = (\delta_l + \cdots + \delta_r)(k + 1) + l$ in the case, when $k = 2i$, $r = 2j$, $l = 2v$ or $k = 2i$, $r = 2j-1$, $v = 2i-1$. In the other cases there exist l-stable formulas with the degree $p = (\delta_l+\cdots+\delta_r)(k+1)+l-1$.*

Under solving some problems, it is useful to determine beforehand the sign of the coefficients $\beta_k^{(j)}(j = 1, \ldots, r)$, and also the relation between them. For example, in using of two sided methods, just as in construction of the new methods having Obrechkoff's type there arises the question on determination of the sign of some coefficients. For this aim consider the next theorem.

Theorem 6. *Suppose, that the formula (1) is stable, has the degree p, which got a maximal value and $\alpha_k > 0$. Then*

$$\beta_k^{(j)} = (-1)^{j-1} l_j (l_j > 0),\ |\beta_k^{(m)}| > |\beta_k^{(m+1)}|\ (m = 1, \ldots, r-1, j = 1, \ldots, r).$$

But if $\beta_k^{(v)} \neq 0, \beta_k^{(v+1)} = \cdots = \beta_k^{(v+s)} = 0$, $\beta_k^{(v+s+1)} \neq 0$, then $\beta_k^{(v)} \beta_k^{(v+s+1)} < 0$ and $|\beta_k^{(v)}| > |\beta_k^{(v+s+1)}|$.

Let $\beta_k^{(1)} = \beta_k^{(2)} = \cdots = \beta_k^{(s-1)} = 0$ and $\beta_k^{(s)} \neq 0$. Then $\beta_k^{(s)} > 0$.

As it is obvious from the formulation of the theorem 6, here the maximal value of the degree for formula (1) is taken, in every considered cases. For example, $p_{\max} = 5$ in the case $r = 2, k = 2$ and $\beta_k^{(1)} = 0$.

Note. Below we reduced some concrete methods constructed by author several times

$$y_{n+1} = \frac{12y_n - h(f_{n+2} - 8f_{n+1} - 5f_n)}{12} \quad (r = 1, k = 2, p = 3)$$

(local trun. err.$h^4 y_n^{(4)}/24 + O(h^5)$),

$$y_{n+2} = \frac{8y_{n+1} + 11y_n}{19} - \frac{h(f_{n+3} - 24f_{n+2} - 57f_{n+1} - 10f_n}{57} \quad (r = 1, k = 3, p = 5)$$

(local trun. err. $-11h^6 y_n^{(6)}/3420 + O(h^7)$),

$$y_{n+2} = \frac{(416y_{n+1} - 103y_n)}{313} + \frac{h(157f_{n+3} + 11232f_{n+2} + 8451f_{n+1} - 2830f_n)}{25353} - \frac{h^2(11g_{n+3} + 630g_{n+2} - 1557g_{n+1} + 92g_n)}{8451} \quad (r = 2, k = 3, p = 9)$$

(local trun. err.$103h^{10} y_n^{(10)}/212965200 + O(h^{11})$),
here $g(x, y) = f'_x(x, y) + f'_y(x, y)f(x, y), y' = f(x, y)$.

It is noted, that there are concrete methods for which theorem 6 is correct in the case, when the value of the degree of the stable methods is less than maximal.

Obrechkoff's method that is the formula (1.16), in more general form was investigated for $r = 2$ and arbitrary k, in [9].

References

[1] N. S. Bakhvalov, Some remarks to question on the numerical integration of the differential equations, by the method of finite-difference, *RAS USSR (Report of Academy of sciences)*, No. 6, p. 805-809 (in Russian).

[2] G. Dahlquist, Convergence and stability in the numerical integration of ordinary differential equations, *Math. Scand.*, 4 (1956) 33-53.

[3] G. Dahlquist, Stability and error bounds in the numerical integration of ordinary differential equations, *Almqvist & Wiksells boktr.* , 130 (1959) 5-92.

[4] V.R. Ibrahimov, Relation between order and degree for the stable difference formula, *The transactions ASU (Azerbaijan State University)*, series phys-math. sciences, 3 (1978) 40-49 (in Russian).

[5] I. Kobza, Second derivative linear multi-step formula and its stability on the imaginary axis, *Acta UPO. F.R.M.*, 22 (1983) 131-142.

[6] W.H. Enright, Second derivative multi-step methods for stiff ordinary dif. equations, *SIAM T. Numer. Anal.*, 2 (1974) 321-332.

[7] A. Iserles and S.P. Norset, Two-step method and Bi-orthogonality, *Math. of Comput.*, 180 (1987) 543-552.

[8] V.R. Ibrahimov, On the numerical method for the solving differential equations of high orders, *Fourth. Confr. on dif. equat. and applict*, Rousse'89 Bulgaria, abstract of invested and Contrib. papers. Rousse, 1989, p. 384.

[9] V.R. Ibrahimov, About connection between order and degree for the stable forward-jumping formulas, *J. Comput. math. and math. phys.*, 7 (1990) 1045-1056 (in Russian).

Bulletin of the Iranian Mathematical Society
Vol. 28, No. 1, pp 29-42 (2002)

α-VALUATIONS OF SPECIAL CLASSES OF QUADRATIC GRAPHS

"To the memory of Jaromir Abrham (1931-1997)"

Kourosh Eshghi

Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran
eshghi@sharif.edu

Abstract: It is shown that the quadratic graph Q(5,4k) (consisting of 5 cycles of length 4k) has an α-valuation (a stronger form of the graceful valuation) for every positive integer k. Furthermore, additional results are obtained from the main theorem of this paper.

1. BASIC DEFINITIONS

Let $G = (V, E)$ be a graph with $m = |V|$ vertices and $n = |E|$ edges. By the term graph, we mean an undirected finite graph without loops or multiple edges. All parameters in this paper are positive integers. A graceful valuation (or β-valuation) of a graph $G = (V, E)$ is a one-to-one mapping Ψ of the vertex set $V(G)$ into the set $\{0, 1, 2, \ldots, n\}$ with

[0] *MSC (2000):* 05C78
[0] *Keywords:* Graphing labeling, Graceful valuations, α-valuation
[0] *Received:* 26 January, 2001; *Revised:* 6 November, 2001
[0] The author is partially supported by Sharif University of Technology Grant No. 717

this property: If we define, for any edge $e = \{u, v\} \in E(G)$, the value $\Psi^{\bullet}(e) = |\Psi(u) - \Psi(v)|$ then $\Psi^{\bullet}$ is a one-to-one mapping of the set $E(G)$ onto the set $\{1, 2, \ldots, n\}$.

A graph is called graceful if it has a graceful valuation. An α-valuation (or α-labeling) of a graph $G = (V, E)$ is a graceful valuation of G which satisfies the following additional condition: There exists a number $\gamma(0 \leq \gamma \leq |E(G)|)$ such that, for any edge $e \in E(G)$ with end vertices $u, v \in V(G), \min[\Psi(u), \Psi(v)] \leq \gamma < \max[\Psi(u), \Psi(v)]$.

The concept of a graceful valuation and of an α-valuation were introduced by Rosa [7]. Rosa proved that, if G is graceful and if all vertices of G are of even degree, then $|E(G)| \equiv 0$ or 3 (mod 4). This implies that if G has an α-valuation and if all vertices of G are of even degrees, then $|E(G)| \equiv 0$ (mod 4) (G is bipartite). In [7] it is also shown that these conditions are also sufficient if G is a cycle. The symbol Cm will denote a cycle on m vertices. Abrham and Kotzig [2] proved that Rosa's condition is also sufficient for 2-regular graphs with two components.

A snake is a tree with exactly two vertices of degree 1. In [7], it was proved that every snake has an α-valuation. A snake with n edges will be denoted by P_n.

A detailed history of the graph labeling problem and related results appears in Gallian [4,5]. One of the results of Abrham and Kotzig should be mentioned here: If G is a 2-regular graph on n vertices and n edges which has a graceful valuation Ψ then there exists exactly one number $x(0 < x < n)$ such that $\Psi(v) \neq x$ for all $v \in V(G)$; this number x is referred to as the missing value of the graceful valuation[2].
A quadratic graph $Q(r, s)$ is a graph with r components, each of which is an s-cycle.

Here are some of the results published in the references:

1. A $Q(1, s)$-graph (i.e. an s-cycle) is graceful if and only if $s \equiv 0$ or 3 (mod 4). It has an α-valuation if and only if $s \equiv 0$ (mod 4) [7].

2. A $Q(2,s)$-graph has an α-valuation if and only if s is even and $s > 2$ [6].

3. A $Q(3,4k)$-graph has an α-valuation for each $k > 1$. The $Q(3,4)$-graph does not have an α-valuation but it is graceful [6].

4. A $Q(r,3)$-graph is graceful if and only if $r = 1$. A $Q(r,5)$-graph is not graceful for any r [6].

5. A $Q(r,4)$-graph has an α-valuation [6].

2. TRANSFORMATIONS OF LABELING OF A GRAPH

The transformations presented below are used extensively in this paper.

2.1 Transformation Type 1

Lemma 1: (Abrham & Kotzig [1]) *Let r be a non-negative integer and let s be an odd integer, $s = 2k+1$. Then P_s has an α-valuation Ψ with endpoints labeled w and z that satisfies the conditions $z - w = k+1$ and $w = r$.(w.l.o.g., we assume that $w < z$.)*

Suppose that we have two series of vertex labels as follows where $0 \leq n, n+k \leq 2r$ and $2r+1 \leq m, m+k < 4r+1$ and $|4r+1-(m+n+k)| \leq 1$:

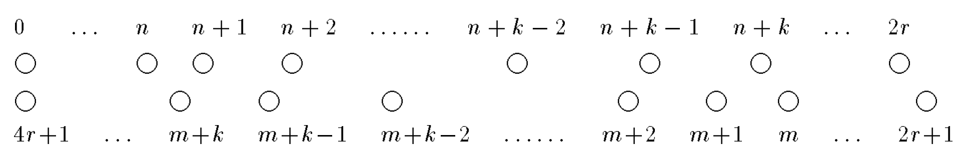


Figure 1: Arrangement of vertex labels in transformation type 1

We apply the transformation type 1 to the vertex labels $(n, n+1, n+2, \ldots, n+k-2, n+k-1, n+k)$ and $(m, m+1, m+2, \ldots, m+k-2, m+k-1, m+k)$ by choosing the vertices w' and z' as end points in the following steps:
Step1: First we modify the vertex labels as follows:

1) From each label $(n, n+1, n+2, \dots, n+k-2, n+k-1, n+k)$ subtract n;

2) From each label $(m, m+1, m+2, \dots, m+k-2, m+k-1, m+k)$ subtract $m-(k+1)$

Step2: According to Lemma 1, we construct an α-valuation for the snake P_{2k+1} on the new labels, with end vertices having labels w and z such that $0 \le w \le k$ and $k+1 \le z \le 2k+1$ and $z-w = k+1$.

Step3: Now if we modify again the new values to the original values in the following way:

1) Add n to each new label $(0, 1, 2, \dots, k-2, k-1, k)$;

2) Add $m-(k+1)$ to each new value $(k+1, k+2, \dots, 2k-1, 2k, 2k+1)$;

Then the end vertices of P_{2k+1} will be labeled $w' = w+n$ and $z' = z+m-(k+1)$ and the edge values will be $m-n-k, m-n-k+1, \dots, m-n+k$.

2.2 Transformation Type 2

Lemma 2: (Abrham & Kotzig [1]) *Let r be a non-negative integer and let s be an even integer, $s = 2k$. Then P_s has an α-valuation Ψ with endpoints labeled w and z that satisfies the conditions $z + w = k$ and $w = r$.(w.l.o.g., we assume that $w < z$.)*

Suppose that we have two series of vertex labels as follows where $0 \le n, n+k \le 2r$ and $2r+1 \le m, m+k-1 < 4r$ and $|4r+1-(m+n+k)| \le 1$:

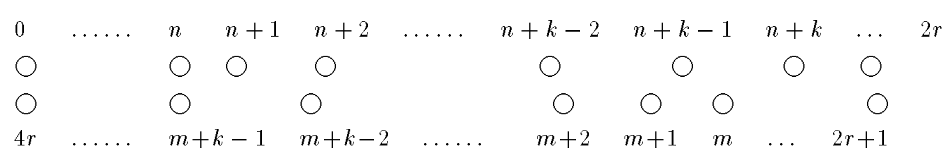


Figure 2: Arrangement of vertex labels in transformation type 2

We apply the transformation type 2 to the vertex labels $(n, n+1, n+2, \ldots, n+k-2, n+k-1, n+k)$ and $(m, m+1, m+2, ..., m+k-2, m+k-1)$ by choosing the vertices w' and z' as end points in the following steps:
Step1: First we modify the vertex labels as follows:

1) From each label $(n, n+1, n+2, \ldots, n+k-2, n+k-1, n+k)$ subtract n;

2) From each label $(m, m+1, m+2, \ldots, m+k-2, m+k-1)$ subtract $m-(k+1)$

Step2: According to Lemma 2, we construct an α-valuation for the snake P_{2k} on the new labels, with end vertices having labels w and z such that $0 \leq w < (k/2) < z \leq k$ and $z + w = k$.

Step3: Now if we transform again the new values to the original values in the following way:

1) Add n to each new label $(0, 1, 2, \ldots, k-2, k-1, k)$;

2) Add $m-(k+1)$ to each new value $(k+1, k+2, \ldots, 2k-1, 2k)$;

Then the end vertices of P_{2k} will be labeled $w' = w + n$ and $z' = z$ and the edge values will be $m-n-k, m-n-k+1, \ldots, m-n+k-1$.

3. BASIC THEOREM

Theorem 1: *The quadratic graph Q(5,4k) has an α-valuation for all* $k \geq 1$.

Proof: The missing value of the α-valuation of this graph is 5k. Now let us assume $k \geq 5$. The vertices of the first C_{4k} will be $[8k, 12k, 8k+1, 12k-1, 8k+2, 12k-2, \ldots, 9k-1, 11k+1, 9k+1, 11k, 9k+2, 11k-1, \ldots, 10k-1, 10k+2, 10k, 10k+1]$; this yields the edge values $4k, 4k-1, 4k-2, 4k-3, \ldots, 2k+2, 2k+1, 2k, ..., 3, 2, 1$. Next we will describe the labeling of the second C_{4k}. The successive vertices will be labeled

as follows:
$[14k{+}1, 6k, 14k, 6k{+}1, \ldots, 13k{+}3, 7k{-}2, 13k{+}2, 7k{-}1, 13k, 7k, \ldots, 12k+2, 8k-2, 12k+1, 8k-1]$. The edge labels of this C_{4k} will be then $8k+1, 8k, 8k-1, \ldots, 6k+5, 6k+4, 6k+3, 6k+2, 6k+1, 6k, \ldots, 4k+4, 4k+3, 4k+2$.
The third C_{4k} will be labeled in three stages as follows:

I. Form the snake $(6k-1, 14k+2, 6k-2, 14k+3, 6k-3, \ldots, 5k+2, 15k-1, 5k+1, 15k, 5k-1, 15k+1, 5k-2, 15k+2, \ldots, 16k-3, 4k+2, 16k-2, 4k+1)$. The values of the edges are then $8k+3, 8k+4, 8k+5, 8k+6, \ldots, 10k-3, 10k-2, 10k-1, 10k+1, \ldots, 12k-5, 12k-4, 12k-3$.

II. Join the vertex labeled $6k-1$ to the vertex labeled $16k-1$ to generate the edge labeled $10k$.

III. Form another snake in such a way that its vertices are labeled as follows: $(16k{-}1, 4k{-}1, 16k{+}2, 4k{-}2, 16k, 4k{+}1)$. The resulting values of the edges of this snake are then $12k-1, 12k, 12k+2, 12k+3, 12k+4$.

Now we construct the fourth cycle C_{4k} according to the following stages:

a) The edge labels $12k+1$ and $12k+5$ are generated by joining the following pairs of vertices respectively: $4k$ and $16k+1; 4k-4$ and $16k+1$.

b) Apply transformation type 2 to the vertex labels $(3k-2, 3k-1, 3k, \ldots, 4k-5, 4k-4, 4k-3)$ and $(16k+3, 16k+4, 16k+5, \ldots, 17k, 17k+1)$ by choosing the two vertices $3k-1$ and $4k-4$ as end vertices. The corresponding edge values of this transformation will be $12k{+}6, 12k{+}7, 12k{+}8, \ldots, 14k{+}2, 14k+3$.

c) The edges labeled $14k+4$ and $14k+5$ are obtained by joining the following pairs of vertices respectively: $3k-1$ and $17k+3$; $4k$ and $18k+5$.

d) Construct the snake $(17k+3, 3k-3, 17k+4, 3k-4, 17k+5, 3k-5, \ldots, 18k, 2k, 18k+1, 2k-1, 18k+5)$. The corresponding edge labels will then be $14k+6, 14k+7, 14k+8, \ldots, 16k, 16k+1, 16k+2, 16k+6$.

Finally the last cycle C_{4k} will be constructed as follows when $k \geq 14$ and $k \neq 22$:

1. The edges labeled $12k-2, 4k+1, 8k+2$ and $16k+3$ are obtained by connecting the following pairs of vertices respectively: $k+3$ and $13k+1$; $9k$ and $13k+1$; $9k$ and $17k+2$; $k-1$ and $17k+2$.

2. Construct the snake $(18k+2, 2k-2, 18k+3, 2k-4, 18k+4, 2k-5, 18k+6, 2k-6, 18k+7, 2k-3)$. This yields the edge labels $16k+4, 16k+5, 16k+7, ..., 16k+13$.

3. The edge labeled $16k+14$ is obtained by joining the two vertex labels $2k-3$ and $18k+11$ together.

4. Apply transformation type 1 to the vertex labels $(k+4, k+5, \ldots, 2k-8, 2k-7)$ and $(18k+8, 18k+9, 18k+10, 18k+11, \ldots, 19k-4, 19k-3)$ by using the two vertices $k+7$ and $18k+11$ as end points. This transformation generates the edge labels $16k+15, 16k+16, ..., 18k-8, 18k-7$.

5. Connect the following pairs of vertices to each other to obtain the edges labeled $18k-6$ and $18k-5$ respectively: 8 and $18k+2$; $k+7$ and $19k+2$.

6. Construct the snake $(19k-2, k+2, 19k-1, k+1, 19k, k, 19k+1, k-1)$. The resulting values of the edges are then $18k-4, 18k-3, \ldots, 18k+2$.

7. The edge labeled $18k+3$ is obtained by joining the two vertices $k+3$ and $19k+6$.

8. Construct the snake $(19k+2, k-2, 19k+3, k-3, 19k+4, k-4, 19k+5, k-5, 19k+6)$. The edge labels $18k+4, 18k+5, 18k+6, \ldots, 18k+11$ are generated by this snake.

9. Connect the two vertices $k-14$ and $19k-2$ to each other to generate the edge label $18k+12$.

10. Finally apply transformation type 2 to the vertex labels $(0, 1, 2, \ldots, 8, \ldots, k-14, \ldots, k-6)$ and $(19k+7, 19k+8, \ldots, 20k-1, 20k)$ by considering the two vertices 8 and $k-14$ as end vertices. The rest of the edge values will be generated by this transformation and the last C_{4k} will be completed.

The construction of an α-valuation of the last C_{4k} when $5 \leq k \leq 13$ or $k = 22$ has been given in the Table 1 as follows:

k	The construction of the fifth cycle C_{4k} in the graph Q(5,4k)
5	[45, 66, 8, 92, 7, 94, 6, 96, 5, 97, 3, 98, 2, 99, 1, 100, 0, 93, 4, 87]
6	[54, 79, 9, 120, 0, 119, 1, 118, 2, 117, 3, 116, 4, 110, 10, 111, 8, 112, 7,114,6, 115, 5, 104]
7	[63, 92, 10, 129, 12, 128, 8, 130, 5, 135, 4, 136, 3, 137, 2, 138, 1, 139, 0, 140, 11, 132, 9, 133, 7, 134, 6, 121]
8	[72, 105, 11, 160, 0, 159, 1, 158, 2, 157, 3, 156, 4, 155, 5, 151, 9, 152, 8, 153, 6, 154, 13, 150, 10, 148, 12, 147, 14, 146, 7, 138]
9	[81, 118, 12, 176, 4, 175, 5, 178, 2, 177, 3, 180, 0, 179, 1, 170, 10, 171, 9, 172, 7, 173, 6, 174, 15, 169, 11, 168, 13, 166, 14, 165, 16, 164, 8, 155]
10	[90, 131, 13, 197, 3, 196, 4, 199, 2, 198, 0, 200, 1, 192, 17, 187, 14, 186, 15, 184, 16, 183, 18, 182, 8, 193, 7, 194, 6, 195, 5, 188, 12, 189, 11, 190, 10, 191, 9, 172]
11	[99, 144, 14, 218, 2, 219, 1, 220, 0, 215, 5, 216, 4, 217, 3, 212, 7, 213, 6, 214, 15, 206, 13, 207, 12, 208, 11, 209, 19, 205, 16, 204, 17, 202, 18, 201, 20, 200, 8, 211, 9, 210, 10, 189]
12	[108, 157, 15, 230, 6, 234, 7, 233, 8, 218, 22, 219, 20, 220, 19, 222, 18, 223, 21, 227, 13, 226, 14, 225, 16, 224, 17, 237,3, 236, 4, 235, 5, 240, 0, 239, 1, 238, 2, 231, 10, 232, 9, 228, 12, 229, 11, 206]
13	[117, 170, 16, 256, 4, 255, 5, 258, 2, 257, 3, 260, 0, 259, 1, 250, 19, 242, 18, 243, 17, 244, 15, 245, 23, 241, 20, 240, 21, 238, 22, 237, 24, 236, 8, 252, 7, 253, 6, 254, 11, 249, 10, 251, 9, 246, 14, 247, 13, 248, 12, 223]
22	[198, 287, 25, 430, 10, 429, 11, 432, 8, 431, 9, 426, 14, 427, 13, 428, 12, 423, 17, 424, 16, 425, 15, 440, 0, 439, 1, 438, 2, 437, 3, 436, 4, 435, 5, 434, 6, 433, 7, 398, 42, 399, 40, 400, 39, 402, 38, 403, 41, 407, 34, 406, 35, 405, 36, 404, 37, 411, 30, 410, 31, 409, 32, 408, 33, 415, 26, 414, 27, 413, 28, 412, 29, 419, 22, 418, 23, 417, 24, 416, 18, 422, 19, 421, 20, 420, 21, 376]

Table 1: The construction of the fifth cycle C_{4k} in the graph Q(5,4k), $5 \leq k \leq 13, k = 22$

For $1 \leq k \leq 4$, the successive vertices of each cycle of Q(5,4k) will be labeled according to the following table:

k	The construction of an α-valuation of Q(5,4k)
1	[0, 18, 1, 20] , [2, 16, 4, 17] , [3, 14, 9, 19] , [6, 13, 7, 15] , [8, 11, 10, 12]
2	[0, 37, 4, 33, 8, 39, 1, 40] , [2, 36, 18, 27, 5, 35, 3, 38] , [6, 34, 7, 31, 11, 30, 9, 32] , [12, 29, 15, 25, 14, 26, 13, 28] , [16, 24, 17, 23, 19, 22, 20, 21]
3	[0, 60, 4, 53, 27, 40, 6, 54, 7, 58, 1, 59] , [2, 57, 12, 49, 8, 52, 9, 51, 5, 55, 3, 56] , [10, 50, 11, 47, 17, 44, 16, 45, 14, 46, 13, 48] , [18, 43, 23, 37, 22, 38, 21, 39, 20, 41, 19, 42] , [24, 36, 25, 35, 26, 34, 28, 33, 29, 32, 30, 31]
4	[0, 80, 5, 74, 4, 70, 36, 53, 7, 75, 3, 76, 2, 78, 1, 79] , [6, 77, 16, 65, 12, 68, 13, 67, 10, 69, 11, 71, 9, 72, 8, 73] , [14, 66, 15, 63, 23, 58, 22, 59, 21, 60, 19, 61, 18, 62, 17, 64] , [24, 57, 31, 49, 30, 50, 29, 51, 28, 52, 27, 54, 26, 55, 25, 56] , [32, 48, 33, 47, 34, 46, 35, 45, 37, 44, 38, 43, 39, 42, 40, 41]

Table 2: The construction of an α-valuation of Q(5,4k) for $1 \leq k \leq 4$

4. THE STANDARD VALUATIONS OF C_{4k}

Definition 1: The standard α-valuation of C_{4k} are given by any of the following sequence of values of the consecutive vertices of C_{4k}:

a) $[4k, 0, 4k-1, 1, 4k-2, 2, \ldots, k-2, 3k+1, k-1, 3k, k+1, 3k-1, k+2, 3k-2, \ldots, 2k+2, 2k-1, 2k+1, 2k]$ with missing value $x = k$.

b) $[0, 4k, 1, 4k-1, 2, 4k-2, \ldots, k-2, 3k+2, k-1, 3k-1, k+1, 3k, k+2, 3k-1, \ldots, 2k-2, 2k+2, 2k, 2k+1]$ with missing value $x = k$.

c) $[4k, 0, 4k-1, 1, 4k-2, 2, \ldots, k-2, 3k+1, k-1, 3k-1, k, 3k-2, \ldots, 2k+1, 2k-2, 2k, 2k-1]$ with missing value $x = 3k$.

d) $[0, 4k, 1, 4k-1, 2, 4k-2, \ldots, k-2, 3k+2, k-1, 3k+1, k, 3k-1, k+1, \ldots, 2k-2, 2k+1, 2k-1, 2k]$ with missing value $x = 3k$.

In Figure 3 one of the standard α-valuations of C_{12} has been shown:

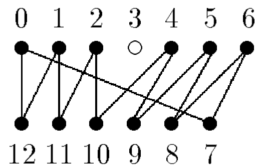


Figure 3: A standard α-valuation of C_{12}

If we suppose that $k_1 + k_2 + \cdots + k_n = k$ and there is an α-valuation for the graph $C_{4k_1} \cup C_{4k_2} \cup \cdots \cup C_{4k_n}$ then in a standard α-valuation of C_{4k}, we can replace C_{4k} by $C_{4k_1} \cup C_{4k_2} \cup \cdots \cup C_{4k_n}$ with its α-valuation and the resulting graph will again have an α-valuation. For example an α-valuation of C_{12} in Figure 3 is replaced by an α-valuation of $2C_6$ in Figure 4:

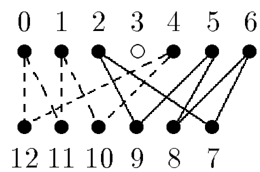


Figure 4: An α-valuation of $2C_6$

Definition 2: The graph C_{4k} has a standard valuation (or standard labeling) if the values of the vertices of C_{4k} can be generated from a standard α-valuation of C_{4k} differ by a constant factor.

For example C_{12} in the α-valuation of $C_{12} \cup C_{20}$ shown in Figure 5 has a standard valuation because it can be generated from a standard α-valuation of C_{12} that differs by a constant factor 10:

Figure 5: An α-valuation of $C_{12} \cup C_{20}$

If a graph C_{4k} has a standard valuation it can be replaced by any α-valuation of $C_{k_1} \cup C_{k_2} \cup \cdots \cup C_{k_n}$ where $k_1 + k_2 + \cdots + k_n = k$ by considering the constant factor. For instance the standard valuation of C_{12} in Figure 5 can be replaced by an α-valuation of $2C_6$ to form an α-valuation of $2C_6 \cup C_{20}$ if we increase the values of the α-valuation $2C_6$ in Figure 4 by constant factor i.e. 10:

Figure 6: An α-valuation of $C_{12} \cup C_{20}$

Theorem 2: *The following graphs have α-valuations:*

a) $\bigcup_{i=1}^{n} C_{4k_i} \bigcup Q(4,4k)$ if $k = \sum_{i=1}^{n} k_i$ and $k_n + k_{n-1} + \cdots + k_{i+2} + k_{i+1} \leq k_i$ for $i = 1,2,3,\ldots,n-1$.

b) $\bigcup_{i=1}^{n} 2C_{4k_i} \cup Q(3,4k)$ if $k = \sum_{i=1}^{n} k_i$ and $k_n + k_{n-1} + \cdots + k_{i+2} + k_{i+1} \leq k_i$ for $i = 1,2,3,\ldots,n-1$.

c) $\bigcup_{i=1}^{n} C_{4k_i} \cup \bigcup_{j=1}^{t} C_{4p_j} \cup Q(3,4k)$ if $k = \sum_{i=1}^{n} k_i = \sum_{j=1}^{t} p_j$ and $k_n + k_{n-1} + \cdots + k_{i+2} + k_{i+1} \leq k_i$ and $p_t + p_{t-1} + \cdots + p_{j+2} + p_{j+1} \leq p_j$ for $i = 1,2,3,\ldots,n-1$ and $j = 1,2,\ldots,t-1$.

d) $\bigcup_{i=1}^{n} (C_{4k_i} \cup C_{4p_j}) \cup C_{4k_n} \cup Q(4,4k)$ if $k = k_n + \sum_{i=1}^{n} (k_i + p_i)$ and $k_i = 2k_{i+1} + p_{i+1}$ for $i = 1,2,3,\ldots,n-1$.

Proof: We know that in construction of α-valuation of Q(5,4k); at least two cycles C_{4k} have standard α-valuation. In order to obtain the different parts of the theorem 2, we replace these two standard valuations with other graphs as follows:

a) Consider one of the standard valuation of C_{4k}. First we replace it by $C_{4k_1} \cup C_{4l_1}$; $l_1 \leq k_1; k = k_1 + l_1$. Then since C_{4l_1} still has a standard

valuation [1], we are able to replace it again by $C_{4k_2} \cup C_{4l_2}$; $l_2 \leq k_2$; $l_1 = k_2 + l_2$. In next stages we continue to replace each C_{4l_i} by $C_{4k_{i+1}} \cup C_{4l_{i+1}}$; $l_{i+1} \leq k_{i+1}$; $l_i = k_{i+1} + l_{i+1}$ for $i = 2, 3, \ldots, n-2$; $k_n = l_{n-1}$.
b) We apply the replacement procedure of part (a) for both C_{4k} which have standard valuations in α-valuation of Q(5,4k).
c) The proof of this part is similar to part (b) except that each standard valuation C_{4k} has been replaced by different disjoint unions of graphs in such a way that their components are not necessarily isomorphic.
d) Consider one of the standard valuation of C_{4k}. First we replace it by $2C_{4k_1} \cup C_{4p_1}$; $k = p_1 + 2k_1$; we know at least one of C_{4k_1} has a standard valuation [3]. Thus we replace C_{4k_1} in the next step by $2C_{4k_2} \cup C_{4p_2}, k_1 = p_2 + 2k_2$. In next stages, we repeat the replacement C_{4k_i} by $2C_{4k_{i+1}} \cup C_{4p_{i+1}}, k_i = p_{i+1} + 2k_{i+1}$; $i = 1, 2, 3, \ldots, n-1$.

References

[1] J. Abrham and A. Kotzig, Graceful valuations of 2-regular graphs with two components, *Discrete Math.*, 150 (1996) 3-15.

[2] J. Abrham and A. Kotzig, On the missing value in graceful numbering of a 2-regular graph, *Cong. Numer.*, 65 (1988) 261-266.

[3] K. Eshghi, *Existence and construction of α-labeling of 2-regular graphs with three components*, Ph.D. Thesis, University of Toronto, 1997.

[4] J. A. Gallian, A survey: recent results, conjectures and open problems on labeling graphs, *J. Graph Theory*, 13 (1989) 491-504.

[5] J. A. Gallian, A guide to the graph labeling zoo, *Discrete Appl. Math.*, 49 (1994) 213-229.

[6] A. Kotzig, β-valuations of quadratic graphs with isomorphic components, *Utilitas Math.*, 7 (1975) 263-279.

[7] A. Rosa, *On certain valuations of the vertices of a graph, Theory of Graphs* (International Symposium, Rome, July 1966), Gordon and Breach, New York and Dunod Paris (1967) 349-355.

Bulletin of the Iranian Mathematical Society
Vol. 28, No. 1, pp 43-57 (2002)

EXTENDING FUNCTIONS IN THE MODEL SUBSPACES OF $H^2(\mathbb{R})$ to $\mathbb{C}$

Javad Mashreghi

Département de mathématiques et de statistique Université Laval Québec, QC Canada G1K 7P4.
Javad.Mashreghi@mat.ulaval.ca

Abstract: It is shown that each function f in a model subspace K_Θ of $H^2(\mathbb{R})$ can be extended to $\mathbb{C}$. The extension to the upper half plane is in $H^2(\mathbb{C}_+)$ and the extension to the lower half plane is in $\Theta\ H^2(\mathbb{C}_-)$. We also show that f is analytic at each point of the real line where Θ is analytic. Finally, we completely characterize K_Θ for $\Theta(x) = e^{i\sigma x}$ and for Θ being a meromorphic Blaschke product.

1. Introduction

Let f be an analytic function in the upper half plane $\mathbb{C}_+ = \{z \in \mathbb{C} : \Im z > 0\}$. Let $\| f_y \|_p = \left(\int_{-\infty}^{\infty} |f(x+iy)|^p \, dx \right)^{\frac{1}{p}}$ and $\| f \|_p = \sup_{y>0} \| f_y \|_p$ for $0 < p < \infty$. The Hardy space $H^p(\mathbb{C}_+)$ consists of all f's with

[0] *MSC(2000):* Primary: 30D50, Secondary: 30B40

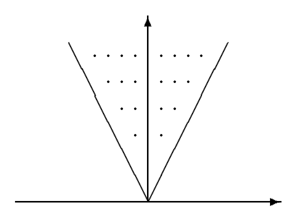
[0] *Keywords:* argument, harmonic conjugate, Hilbert transform, Lipschitz function.

[0] *Received:* 2 May, 2001; *Revised:* 10 January,2002

[0]This research was supported by NSERC (Canada) and FCAR (Québec).

$\| f \|_p < \infty$. The Hardy space $H^\infty(\mathbb{C}_+)$ consists of all bounded analytic functions in the upper half plane. In this case $\| f \|_\infty = \sup_{z \in \mathbb{C}_+} |f(z)|$. For $0 < p < 1$, $H^p(\mathbb{C}_+)$ with the distance $\| f - g \|_p^p$ is a complete metric space. For $1 \leq p < \infty$, $H^p(\mathbb{C}_+)$, $\| \cdot \|_p$ is a Banach space. In particular, $H^2(\mathbb{C}_+)$, $\| \cdot \|_2$ is a Hilbert space. Finally, $H^\infty(\mathbb{C}_+)$ is a Banach algebra [9, pages 70-74].

For each $f \in H^p(\mathbb{C}_+)$, and for almost all $x \in \mathbb{R}$, $\lim_{z \angle\!\!\to x} f(z)$ exists. Denoting this limit by $f(x)$, we have $f \in L^p(\mathbb{R})$, and furthermore, $\| f \|_p = \| f \|_{L^p(\mathbb{R})}$. In the preceding limit, z is required to tend to x from within sectors of opening $< 180°$ having vertex at x, and symmetric about the vertical line passing through x. We frequently say that $f(z) \to f(x)$ as z tends to x *non-tangentially* [4, page 6].



Therefore, there is a canonical correspondence between $H^p(\mathbb{C}_+)$ and a subspace of $L^p(\mathbb{R})$, denoted by $H^p(\mathbb{R})$. The space $H^p(\mathbb{R})$ can also be independently defined as the set of all $f \in L^p(\mathbb{R})$, with (as a distribution) $\hat{f}(x) = 0$ for $x < 0$. The two definitions are equivalent [7, page 172]. The Hardy spaces $H^p(\mathbb{C}_-)$ are defined similarly. The functions in $H^p(\mathbb{C}_-)$ live in the lower half plane and the family of their boundary values, as functions on $\mathbb{R}$, is precisely the space $\overline{H^p(\mathbb{R})}$. See also Chapter 11 of [4].

The function $\Theta \in H^\infty(\mathbb{R})$ is said to be inner if $|\,\Theta(x)\,| = 1$ for almost all $x \in \mathbb{R}$. For each inner function Θ, the set $\Theta\, H^2(\mathbb{R})$ is a closed subspace of the Hilbert space $H^2(\mathbb{R})$ [9, pages 79-80]. Now we are able to introduce our hero.

Definition *The model space K_Θ is the orthogonal complement of*

$\Theta\, H^2(\mathbb{R})$ *in* $H^2(\mathbb{R})$.

In this paper we study the model space K_Θ corresponding to the inner function Θ. In Section 2 we briefly discuss the role of model subspaces in operator theory. In Section 3 an analytic description of K_Θ is given. This new formulation, which is obtained using Hilbert space characteristic of $H^2(\mathbb{R})$, can be exploited to define K_Θ in the Hardy space $H^p(\mathbb{R})$. This representation also enables us to extend each function in K_Θ to the whole complex plane. In Section 4 we extend an $f \in K_\Theta$ to $\mathbb{C}$. This extension has three fundamental properties. First, $\lim_{z \not\angle\to x} f(z) = f(x)$ for almost all $x \in \mathbb{R}$. In these limit, z is allowed to tend to x non-tangentially from *either* half plane. Second, f as a function defined in the upper half plane is in $H^2(\mathbb{C}_+)$. Third, f as a function defined in the lower half plane is in $\Theta\, H^2(\mathbb{C}_-)$. Therefore, f is at least analytic in the upper half plane and is meromorphic in the lower half plane. Furthermore, in Section 5 we show that f is already analytic wherever Θ is on the real line. Finally in Sections 6 and 7 we completely characterize K_Θ corresponding to $\Theta(x) = e^{i\sigma x}$ and for Θ being a meromorphic Blaschke product. In these two cases, and only for them, each $f \in K_\Theta$ is analytic on the whole real line.

2. Link to operator theory

In this section we explain the origin of model subspaces of $H^2(\mathbb{R})$. Let $f \in H^2(\mathbb{R})$. By the Fourier-Plancherel theorem, if we write

$$\hat{f}_N(\lambda) = \int_{-N}^{N} e^{-i\lambda t} f(t)\, dt,$$

then, as $N \to \infty$, the $\hat{f}_N(\lambda)$ tend in $L^2(\mathbb{R})$ to a function $\hat{f}(\lambda)$, called the *Fourier-Plancherel transform* of f. We can characterize an $f \in H^2(\mathbb{R})$ in terms of its Fourier-Plancherel transform. A function $f \in L^2(\mathbb{R})$ is in $H^2(\mathbb{R})$ if and only if $\hat{f}(\lambda) = 0$ for almost every $\lambda < 0$ [9, page 131]. Therefore, there is a canonical isomorphism between $H^2(\mathbb{R})$ and

$L^2((0,\infty))$. Based on the preceding observation, a function $f \in L^2(\mathbb{R})$ is in $\overline{H^2(\mathbb{R})}$ if and only if $\hat{f}(\lambda) = 0$ for almost every $\lambda > 0$.

Let $\delta > 0$. Then the map T_δ

$$\begin{array}{ccc} H^2(\mathbb{R}) & \mapsto & H^2(\mathbb{R}) \\ f(t) & \mapsto & \exp(i\delta\, t)\, f(t), \end{array}$$

is called a *forward shift operator* on $H^2(\mathbb{R})$. Since for each $f \in H^2(\mathbb{R})$

$$\widehat{T_\delta(f)}\,(\lambda) = \hat{f}\,(\lambda - \delta), \qquad \lambda \in \mathbb{R},$$

T_δ shifts the spectrum of f forward by δ units. Beurling in his classical paper [2] characterized the invariant subspaces of $H^2(\mathbb{R})$ for the forward shift operators.

Beurling's theorem: *A closed subspace of $H^2(\mathbb{R})$ is invariant under T_δ, for each $\delta > 0$, if and only if it is of the form $\Theta\, H^2(\mathbb{R})$ for some inner function Θ.*

The adjoint of a forward shift operator, T_δ^*, is called a *backward shift operator*. By direct verification, one verifies that T_δ^* is defined by

$$\widehat{T_\delta^*(f)}(\lambda) = \begin{cases} f\,(\lambda + \delta), & if\ \lambda > 0, \\ 0, & if\ \lambda < 0 \end{cases}$$

for $f \in H^2(\mathbb{R})$ [6]. Therefore, T_δ^* shifts the spectrum of f backward by δ units, and then chops off the negative part of what is thus obtained. In a Hilbert space, a closed subspace M is invariant under a bounded operator T if and only if $M^\perp$ is invariant under T^* [5, page 40]. Therefore according to the Beurling's theorem, A closed subspace of $H^2(\mathbb{R})$ is invariant under T_δ^* for each $\delta > 0$ if and only if it is the orthogonal complement of $\Theta\, H^2(\mathbb{R})$ for some inner function Θ. Therefore, the subspaces K_Θ are precisely those which are invariant under T_δ^* for each $\delta > 0$. That is why some authors call the K_Θ a *coinvariant subspace* of $H^2(\mathbb{R})$.

3. Analytic description of K_Θ

Let Θ be an inner function for the upper half plane. Then $\Theta\, H^2(\mathbb{R})$ is a closed subspace of the Hilbert space $H^2(\mathbb{R})$. According to the notation introduced before, the orthogonal complement of $\Theta H^2(\mathbb{R})$ in $H^2(\mathbb{R})$ is denoted by K_Θ. The following lemma gives an analytic description of K_Θ which can be used as the definition of it in all Hardy spaces $H^p(\mathbb{R})$, $0 < p \le \infty$.

Theorem 3.1. *For each inner function*Θ

$$K_\Theta = H^2(\mathbb{R}) \cap \Theta\, \overline{H^2(\mathbb{R})}.$$

Proof. Uses the properties $\Theta \in H^\infty$ and $\Theta\,\overline{\Theta} = 1$. By definition, $f \in K_\Theta$ if and only if $f \in H^2(\mathbb{R})$ and

$$\int_{-\infty}^{\infty} f(x)\, \overline{\Theta(x)\, g(x)}\, dx = 0$$

for each $g \in H^2(\mathbb{R})$. Thus, $f \in K_\Theta$ if and only if $f \in H^2(\mathbb{R})$

$$\int_{-\infty}^{\infty} \frac{f(x)}{\Theta(x)}\, \overline{g(x)}\, dx = 0$$

for each $g \in H^2(\mathbb{R})$. This condition is equivalent to $\dfrac{f}{\Theta} \in \overline{H^2(\mathbb{R})}$. Therefore $f \in K_\Theta$ if and only if $f \in H^2(\mathbb{R})$ and also $f \in \Theta\, \overline{H^2(\mathbb{R})}$. ∎

4. Extension to upper and lower half planes

Let $h \in L^2(\mathbb{R})$. Then the Poisson integral formula

$$P_h(z) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|\Im z|}{|z-t|^2}\, h(t)\, dt, \qquad z \in \mathbb{C} \setminus \mathbb{R},$$

gives an extension of h to the upper and to the lower half planes. It can be shown that $h \in H^2(\mathbb{R})$ if and only if P_h, as a function defined in the upper half plane, is in $H^2(\mathbb{C}_+)$. Similarly, $h \in \overline{H^2(\mathbb{R})}$ if and only if

P_h, as a function defined in the lower half plane, is in $H^2(\mathbb{C}_-)$ [12]. An $f \in K_\Theta$ belongs in particular to $H^2(\mathbb{R})$. Therefore it has an extension $f(z)$ to the upper half plane, belonging to $H^2(\mathbb{C}_+)$ and given there by the formula

$$f(z) = P_f(z) \;\; for \;\; z \in \mathbb{C}_+.$$

An inner function Θ can be (formally) extended to the lower half plane by putting

$$\Theta(z) = \frac{1}{\overline{\Theta(\bar{z})}}$$

for $z \in \mathbb{C}_-$. The extension of an $f \in K_\Theta$ to the lower half plane is indirect (depending on Θ). For such an f we have $\overline{\Theta}\, f \in \overline{H^2(\mathbb{R})}$ by Theorem 3.1, so, by the preceding observation, $\overline{\Theta}\, f$ has an analytic extension to the lower half plane, equal there to

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{|\Im z|}{|z-t|^2}\, \overline{\Theta(t)}\, f(t)\, dt = P_{\overline{\Theta} f}(z), \qquad z \in \mathbb{C}_-.$$

We then *define* the extension of $f \in K_\Theta$ to $\mathbb{C}_-$ by putting

$$f(z) = \Theta(z)\, P_{\overline{\Theta} f}(z) \;\; for \;\; z \in \mathbb{C}_-,$$

with $\Theta(z)$ defined as above in $\mathbb{C}_-$. This extension is at least meromorphic in the lower half plane.

Remark: We have $\lim_{z \not\angle\to x} \Theta(z) = \Theta(x)$ and $\lim_{z \not\angle\to x} f(z) = f(x)$ for almost all $x \in \mathbb{R}$. In these limits, z is allowed to tend to x non-tangentially from *either* half plane.

With above definitions, Theorem 3.1 yields the following characterization of K_Θ.

Theorem 4.1. *The space K_Θ consists precisely of the functions $f \in L^2(\mathbb{R})$ with extension to the upper half plane belonging to $H^2(\mathbb{C}_+)$ and whose extension to the lower half plane makes $\frac{f}{\Theta} \in H^2(\mathbb{C}_-)$.*

For further applications of this result see [8].

5. Analytic continuation along $\mathbb{R}$

A function $f \in K_\Theta$ can be continued analytically across intervals of $\mathbb{R}$ on which Θ is analytic. This result has important consequences in characterizing elements of K_B when B is a meromorphic Blaschke product.

Theorem 5.1. *If Θ is analytic in a neighborhood of the interval $(a,b) \subset \mathbb{R}$ then any $f \in K_\Theta$ is also analytic there.*

Proof. By Theorem 4.1, f and $\dfrac{f}{\Theta}$ are respectively holomorphic in the upper and lower half planes. Without loss of generality, suppose Θ is holomorphic inside the rectangle $\{\, z\,;\, a < \Re z < b,\ -2 < \Im z < 2\,\}$. Thus $f = \Theta \cdot \dfrac{f}{\Theta}$ is also analytic inside that rectangle except possibly on (α, β), and for almost all $x \in \mathbb{R}$, $\lim\limits_{|y|\to 0} f(x+iy)$ exists. Choose α, $\beta \in (a,b)$ such that this is true for $x = \alpha$ and for $x = \beta$.

With $\varepsilon > 0$, let us take the paths

$$\begin{aligned} \Gamma &= [\alpha+i, \alpha-i] \cup [\alpha-i, \beta-i] \cup [\beta-i, \beta+i] \cup [\beta+i, \alpha+i], \\ \Gamma_\varepsilon &= [\alpha+i\varepsilon, \alpha-i\varepsilon] \cup [\alpha-i\varepsilon, \beta-i\varepsilon] \cup [\beta-i\varepsilon, \beta+i\varepsilon] \cup [\beta+i\varepsilon, \alpha+i\varepsilon], \end{aligned}$$

each oriented counterclockwise. For each point z inside Γ, let $g(z) = \dfrac{1}{2\pi i}\displaystyle\int_\Gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta$. Then g is holomorphic inside Γ. By the Cauchy integral formula,

$$g(z) = f(z) + \frac{1}{2\pi i}\int_{\Gamma_\varepsilon} \frac{f(\zeta)}{\zeta - z}\, d\zeta$$

for $\varepsilon < |\Im z| < 2$. Since f is bounded on the vertical segments through

α and β,

$$\lim_{\varepsilon\to 0}\int_{[\alpha+i\varepsilon,\alpha-i\varepsilon]}\frac{f(\zeta)}{\zeta-z}\,d\zeta=0 \ \ and \ \ \lim_{\varepsilon\to 0}\int_{[\beta-i\varepsilon,\beta+i\varepsilon]}\frac{f(\zeta)}{\zeta-z}\,d\zeta=0.$$

On the horizontal segments

$$\int_{[\alpha-i\varepsilon,\beta-i\varepsilon]}\frac{f(\zeta)}{\zeta-z}\,d\zeta+\int_{[\beta+i\varepsilon,\alpha+i\varepsilon]}\frac{f(\zeta)}{\zeta-z}\,d\zeta=\int_{\alpha}^{\beta}\left(\frac{f(t-i\varepsilon)}{t-i\varepsilon-z}-\frac{f(t+i\varepsilon)}{t+i\varepsilon-z}\right)dt.$$

Both $\dfrac{f(t-i\varepsilon)}{t-i\varepsilon-z}$ and $\dfrac{f(t+i\varepsilon)}{t+i\varepsilon-z}$ converge in $L^2(dt)$ norm to $\dfrac{f(t)}{t-z}$, as $\varepsilon\to 0$. Thus

$$\lim_{\varepsilon\to 0}\int_{[\alpha-i\varepsilon,\beta-i\varepsilon]}\frac{f(\zeta)}{\zeta-z}\,d\zeta+\int_{[\beta+i\varepsilon,\alpha+i\varepsilon]}\frac{f(\zeta)}{\zeta-z}\,d\zeta=0.$$

Hence $g\equiv f$ in the lower and in the upper part of the interior of Γ. Therefore, f is holomorphic on (α,β). Since α and β can be taken as close to a and to b as we want, f is holomorphic on (a,b). ■

6. Paley-Wiener spaces as model subspaces

Let $\sigma>0$. Then, $\Theta(x)=\exp(i\sigma x)$ is an entire inner function. In this case, the functions $f(x)\in K_\Theta$ differ by the factor $e^{i\sigma x/2}$ from those in a Paley-Wiener space.

Theorem 6.1. *Let $\sigma>0$. Then $f\in K_{e^{i\sigma x}}$ if and only if f is an entire function of exponential type, square integrable on the real line, with*

$$-\sigma\le\limsup_{y\to+\infty}\frac{\log|f(iy)|}{y}\le 0 \quad and \quad 0\le\limsup_{y\to-\infty}\frac{\log|f(iy)|}{|y|}\le\sigma.$$

Proof. Since $\Theta(x)=\exp(i\sigma x)$ is analytic across $\mathbb{R}$, each $f\in K_{e^{i\sigma x}}$ is also analytic there. Furthermore, $f\in H^2(\mathbb{C}_+)$ and $\dfrac{f}{\Theta}\in H^2(\mathbb{C}_-)$ imply that f is analytic on $\mathbb{C}_+$ and also on $\mathbb{C}_-$, that $f\in L^2(\mathbb{R})$, and besides that the support of the Fourier-Plancherel transform of f is a subset of $[0,\sigma]$. Thus $\hat f\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$, and for each $z=x\in\mathbb{R}$,

$f(z) = \int_0^\sigma \hat{f}(t)\, e^{izt}\, dt$. By the uniqueness theorem for analytic functions, equality holds everywhere. Therefore f is an entire function of exponential type with the indicated growth conditions on the imaginary axis. The if part is an easy consequence of the celebrated Paley-Wiener theorem. ■

The following Corollary is an immediate consequence of the Paley-Wiener representation of entire functions of exponential type and the preceding theorem. The indicated representation, by itself, shows that each $f \in K_{e^{i\sigma x}}$ is an entire function of exponential type which is square integrable on the real line. The representation, moreover, restricts the rate of growth of f along the imaginary axis.

Corollary 6.2. *Each $f \in K_{e^{i\sigma x}}$ has the representation*

$$f(z) = \int_0^\sigma \hat{f}(t)\, e^{izt}\, dt$$

, where $\hat{f} \in L^2(0,\sigma)$.

7. The model space K_B

Let $\{z_k\}_{k\geq 1}$ be a sequence of complex numbers in the upper half plane $\mathbb{C}_+$. Let $b_k(z) = e^{i\alpha_k} \cdot \dfrac{z - z_k}{z - \bar{z}_k}$, where α_k is so chosen that $e^{i\alpha_k} \cdot \dfrac{i - z_k}{i - \bar{z}_k} \geq 0$. The rational function $B_K = \prod_{k=1}^{K} b_k$ is called a *finite Blaschke product* for the upper half plane; B_K is analytic at each point of the real line and $|\, B_K(x)\,| = 1$ for $x \in \mathbb{R}$. The relation $\sum_{k=1}^{\infty} \dfrac{\Im z_k}{|\, z_k + i\,|^2} < \infty$ is a necessary and sufficient condition for the uniform convergence of B_K on compact sets, disjoint from the closure of $\{\, \bar{z}_k \,;\, k \geq 1\,\}$, to a non-zero analytic function $B = \prod_{k=1}^{\infty} b_k = \lim_{K\to\infty} B_K$, and we call B an *infinite Blaschke product* for the upper half plane [9, page 120]. Furthermore, $|\, B(z)\,| < 1 \quad$ for $z \in \mathbb{C}_+$. Therefore, by Fatou's theorem [9, page 57], for almost all $x \in \mathbb{R}$, $\lim_{z \not\to x} B(z)$ exists. Denoting that limit by $B(x)$ (wherever it exists), one has $|\, B(x)\,| = 1$ almost everywhere [9, page 66]. A Blaschke sequence in the upper half plane, $\{z_k\}$, has no accumulation point on the real line if and only if $\lim_{k\to\infty} |z_k| = \infty$. Here, since the z_k stay away from zero, a necessary and sufficient condition for the uniform convergence of B_K to B on compact sets disjoint from $\{\bar{z}_k\,;\, k \geq 1\}$ is that $\sum_{k=1}^{\infty} \frac{\Im z_k}{|z_k|^2} < \infty$. In this case, B is a meromorphic function with poles at the $\bar{z}_k$. For this reason, it is called a *meromorphic Blaschke product.* The function B is analytic at each point of $\mathbb{R}$, and $|\, B(x)\,| = 1$ for $x \in \mathbb{R}$. Let us multiply B by a constant of modulus one to get $B(0) = 1$. Then for each z different from all the $\bar{z}_k$,

$$B(z) = \prod_{k=1}^{\infty} \left(\frac{\bar{z}_k}{z_k} \cdot \frac{z - z_k}{z - \bar{z}_k} \right).$$

To emphasize legitimacy of repetition, let $\{z_k\}_{k\geq 1}$ be a distinct sequence in the upper half plane with $z_k \to \infty$ and let $\{m_k\}_{k\geq 1}$ be a sequence of positive integers. Suppose that $\sum_{k=1}^{\infty} \frac{m_k \,\Im z_k}{|z_k|^2} < \infty$. Then $B(z) = \prod_{k=1}^{\infty} \left(\frac{\bar{z}_k}{z_k} \cdot \frac{z - z_k}{z - \bar{z}_k} \right)^{m_k}$ is a meromorphic Blaschke product.

Theorem 7.1. *The space K_B consists precisely of the meromorphic functions f with poles of order at most m_k at the $\bar{z}_k$, such that $f \in H^2(\mathbb{C}_+)$ and also $\frac{f}{B} \in H^2(\mathbb{C}_-)$.*

Proof. Let $f \in K_B$. Then by Theorem 4.1, f and $\frac{f}{B}$ are respectively analytic in the upper and lower half planes. Hence $f = B \cdot \frac{f}{B}$ is a meromorphic function in the lower half plane, with poles of order at most m_k at the $\bar{z}_k$. Finally, by Theorem 5.1, f is analytic at each point of the real line. If, on the other hand, $f \in H^2(\mathbb{C}_+)$ and $\frac{f}{B} \in H^2(\mathbb{C}_-)$, then at least $f \in L^2(\mathbb{R})$. Thus $f \in K_B$ by Theorem 4.1. ■

The following result is an easy consequence of Theorem 7.1. It can also be shown that K_B is actually the closed subspace of $H^2(\mathbb{R})$ generated by the elements $\dfrac{1}{(x-\bar{z}_k)^{\ell_k}}$ with $1 \leq \ell_k \leq m_k$ and $k \geq 1$.

Corollary 7.2. *For each ℓ_k, $1 \leq \ell_k \leq m_k$ and $k \geq 1$, we have* $\dfrac{1}{(z-\bar{z}_k)^{\ell_k}} \in K_B$.

The following result gives a complete description of K_B when B is a finite Blaschke product.

Corollary 7.3. *Let B be the finite Blaschke product*

$$B(z) = \prod_{k=1}^{K} \left(\frac{z-z_k}{z-\bar{z}_k} \right)^{m_k}.$$

Then K_B consists precisely of the linear combinations of the simple fractions $\dfrac{1}{(z-\bar{z}_k)^{\ell_k}}$, where $1 \leq k \leq K$ and $1 \leq \ell_k \leq m_k$. Thus $f \in K_B$ if and only if

$$f(z) = \frac{P(z)}{\prod_{k=1}^{K}(\, z - \bar{z}_k \,)^{m_k}},$$

where P is a polynomial of degree $m_1 + \cdots + m_K - 1$.

Every meromorphic Blaschke product can be represented as

$$B(z) = \frac{\overline{E(\bar{z})}}{E(z)} \quad \textit{for} \quad z \in \mathbb{C},$$

where E is an entire function with zeros at the $\bar{z}_k$ [13]. The order of $\bar{z}_k$ as a zero of E is the same as its order as a pole of B. In the general case, E is not necessarily of exponential type. In the following we write $E^*(z)$ for $\overline{E(\bar{z})}$. This observation enables us to give another characterization of K_B.

Theorem 7.4 *The space K_B consists precisely of functions of the form $\dfrac{f}{E}$, where f is an entire function with both $\dfrac{f}{E} \in H^2(\mathbb{C}_+)$ and $\dfrac{f}{E^*} \in H^2(\mathbb{C}_-)$.*

Proof. Let $g \in K_B$. Then by Theorem 7.1, g is a meromorphic function with poles of order at most m_k at the $\bar{z}_k$. Hence $g\,E$ is an entire function, where E is the entire function furnished by before. Put $f = g\,E$. Then $\frac{f}{E} = g \in H^2(\mathbb{C}_+)$, and $\frac{f}{E^*} = \frac{g}{B} \in H^2(\mathbb{C}_-)$. On the other hand, if f satisfies these conditions, then $\frac{f}{E} \in K_B$ by Theorem 7.1. ■

The preceding result enables us to characterize the *minimal* majorant for K_B when B is a meromorphic Blaschke product with zeros in a Stoltz domain [8].

References

[1] R. Boas, *Entire Functions*, academic press, 1954.

[2] A. Beurling, *On two problems concerning linear transformations in Hilbert space*, Acta Math., 81, pp. 239-255, 1949.

[3] John B. Conway, *Functions of one complex variable*, Second Edition, Springer Verlag, 1978.

[4] Peter L. Duren, *Theory of H^p spaces*, Academic Press, 1970.

[5] P. R. Halmos, *Introduction to Hilbert Space and the theory of spectral multiplicity*, Second Edition, Chelsea publishing company, 1957.

[6] P. R. Halmos, *A Hilbert Space Problem Book*, Second Edition, Springer Verlag, 1982.

[7] V. P. Havin, B. Jörike, *The uncertainty principle in harmonic analysis*, Springer-Verlag, 1994.

[8] Victor Havin, Javad Mashreghi, *dmisible majorants for model subspaces of H^2*, preprint.

[9] P. Koosis, *Introduction to H^p spaces*, Cambridge Tracts in Mathematics 115, Second Edition, 1998.

[10] P. Koosis, *The Logarithmic Integral I*, Cambridge studies in advanced mathematics 12, 1988.

[11] P. Koosis, *The Logarithmic Integral II*, Cambridge Studies in Advanced Mathematics 21, 1992.

[12] Javad Mashreghi, *Hilbert transform of* $\log|f|$, Proc. Amer. Math. Soc., 130 (2002), pp. 683-688.

[13] Javad Mashreghi, *Admisible majorants for model subspaces of* H^2, Ph.D. Thesis (2001), McGill University.

[14] A. Zygmund, *Trigonometric Series*, Cambridge, 1968.

Bulletin of the Iranian Mathematical Society
Vol. 28, No. 1, pp 57-65 (2002)

ESTIMATION OF THE MULTIVARIATE NORMAL MEAN UNDER THE EXTENDED REFLECTED NORMAL LOSS FUNCTION

M.Towhidi and J.Behboodian

Department of Statistics, Shiraz University, Shiraz 71454, Iran.

Abstract: This paper considers simultaneous estimation of multivariate normal mean vector using the extended reflected normal loss function (Spiring [9]). It is shown that the sample mean $\overline{X} = (\bar{X}_1, \dots, \bar{X}_p)'$ is admissible when $p \leq 2$, but for $p \geq 3$, we obtain a class of estimators similar to James-Stein estimators which dominate the sample mean in terms of risks.

1. Introduction

Let $X = (X_1, \dots, X_p)'$ be a normal vector with mean vector $\theta = (\theta_1, \dots, \theta_p)'$ and covariance matrix $\sigma^2 I$, where σ^2 is known. We use the notation
$X \sim N_p(\theta, \sigma^2 I)$, in this article. We consider the simultaneous estimation of $\theta = (\theta_1, \dots, \theta_p)'$ by using a random sample $X_1, \dots, X_N$ from $N_p(\theta, \sigma^2 I)$ under the extended reflected normal loss function, given by

$$L(\delta, \theta) = K\left[1 - exp\{-(\delta - \theta)'\Gamma^{-1}(\delta - \theta)\}\right] \tag{1.1}$$

[0] *MSC:* 62A15, 62C15, 62C20, 62H12.

[0] *Keywords:* Admissibility, inadmissibility, James-Stein estimator, reflected normal loss function.

[0] *Received:* 15 September, 2001; *Revised:* 5 January, 2002.

where $K > 0$, Γ is a constant positive definite matrix. In practice the maximum loss can be a function of many things (e.g., Production resources, cost of identification, rework and liabilities) but generally it is finite. As a result the quadratic loss function, with its infinite maximum loss, is often inadequate in describing the loss function associated with a product and has been criticized by some researchers (e.g., Tribus and Szonyi [13], Leon and Wu [8]). The bounded loss function (1.1) was introduced by Spiring [9] for the first time. This loss is a bounded and increasing function of the quadratic loss.

To estimate θ with $N = 1$ and $\sigma = 1$, Stein [10] showed that X is inadmissible when $p \geq 3$ under squared error loss. James and Stein [7] showed that the following estimator, known as J-S estimator,

$$\delta(X) = \left(1 - \frac{p-2}{\sum_{i=1}^{p} X_i^2}\right) X$$

has uniformly smaller risk than X, for all θ. Strawderman [12], Efron and Morris [6], and Casella and Hwang [4] studied the problem of estimating multivariate normal mean vector under quadratic loss function. Brandwein and Strawderman [3] provided minimax estimators for the mean of a spherically symmetric distribution with concave loss. Chung and Kim [5] investigated the admissibility of the sample mean $\bar{X}$ under balanced loss function. (see Zellner [14])

In section 2 of this paper, using the limiting Bayes method, we show that $\bar{X}$ is admissible when $p \leq 2$ under the loss (1.1). In section 3, we obtain an estimator similar to J-S estimator under the loss (1.1) when $p \geq 3$, in the following form

$$\delta^*(\bar{X}) = \left(1 - \frac{c^*}{\bar{X}'\Gamma^{-1}\bar{X}}\right) \bar{X}$$

and we show that δ^* dominates the usual estimator $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i = (\bar{X}_1, \ldots, \bar{X}_p)'$, where $X_i = (X_{i1}, \ldots, X_{ip})'$ and $\bar{X}_j = \frac{1}{N}\sum_{i=1}^{N} X_{ij}; j = 1, \ldots, p$.

2. Admissibility of $\bar{X}$ when $p \leq 2$

In this section, we consider the admissibility of $\bar{X}$ when $p = 1$ and 2. We show that $\bar{X}$ is admissibile, using the standard Blyth's technique [2].

Let $X_1, \ldots, X_N$ be a random sample from $N_p(\theta, I)$ with the prior normal distribution $\pi_a(\theta)$, where θ has the mean vector zero and covariance matrix $\frac{1}{a}I$. It is easy to show that the Bayes estimator of θ w.r.t. $\pi_a(\theta)$ under the extended reflected normal loss function is

$$\delta_a(\bar{X}) = \frac{N\bar{X}}{N+a}$$

with the risk function,

$$\begin{aligned} R(\theta, \delta_a) &= K\left[1 - E\left[exp\left\{-(\frac{N\bar{X}}{N+a} - \theta)'\Gamma^{-1}(\frac{N\bar{X}}{N+a} - \theta)\right\}\right]\right] \\ &= K[1 - (\frac{2\pi}{N})^{-\frac{p}{2}} \int exp\{-(\frac{Nx}{N+a} - \theta)'\Gamma^{-1}(\frac{Nx}{N+a} - \theta) \\ &- \frac{N}{2}(x-\theta)'(x-\theta)\}dx] \end{aligned}$$

Now using the fact that for any matrices C_1 and C_2 of appropriate dimensions,

$$(C_1 + C_2)^{-1} = C_1^{-1} - C_1^{-1}(C_1^{-1} + C_2^{-1})^{-1}C_1^{-1} \qquad (2.1)$$

it follows that the risk function of the estimator δ_a is equal to

$$K[1 - (\frac{2\pi}{N})^{-\frac{p}{2}}(\frac{N+a}{N})^p \int exp[-\frac{1}{2}\{(y-\eta)'(2\Gamma^{-1} + \frac{(N+a)^2}{N}I)(y-\eta)$$
$$+(\frac{a}{N+a})^2\theta'(\frac{1}{2}\Gamma + \frac{N}{(N+a)^2}I)^{-1}\theta\}]dy]$$

or

$$K\left[1 - \frac{(N+a)^p}{N^{p/2}}|2\Gamma^{-1} + \frac{(N+a)^2}{N}I|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\frac{a}{N+a})^2\theta'(\frac{1}{2}\Gamma + \frac{N}{(N+a)^2}I)^{-1}\theta\right\}\right] \qquad (2.2)$$

where η is a function of θ.

Theorem 2.1: *$\bar{X} = (\bar{X}_1, \dots, \bar{X}_p)'$ is admissible under the loss (1.1) when $p = 1, 2$, where $\bar{X}_j = \frac{1}{N}\sum_{i=1}^{N} X_{ij}, j = 1, \dots, p$.*

Proof: Suppose $\bar{X}$ is dominated by some estimator $\delta(\bar{X})$ of θ. Using the continuity of the risk function in θ for an estimator $\delta(\bar{X})$, it follows that there exists some $\theta_0, \epsilon > 0$ and $\xi > 0$ such that

all $R(\theta, \delta) < R(\theta, \bar{X}) - \epsilon$ for $\theta_0 - \xi 1 < \theta < \theta_0 + \xi 1$

where $1 = (1, 1, \dots, 1)'$.

Let r_a, r_a^*, r_a^{**} be defined as follows:

r_a = Bayes risk of the Bayes solution δ_a w.r.t. π_a.

r_a^* = Bayes risk of $\bar{X}$ w.r.t. π_a.

r_a^{**} = Bayes risk of δ w.r.t. π_a.

Then the difference of Bayes risks of $\bar{X}$ and δ is

$$
\begin{aligned}
r_a^* - r_a^{**} &\geq \int_{\theta_0 - \xi 1}^{\theta_0 + \xi 1} \left[R(\theta, \bar{X}) - R(\theta, \delta)\right] \pi_a(\theta) d\theta \\
&\geq \int_{\theta_0 - \xi 1}^{\theta_0 + \xi 1} \epsilon (2\pi)^{-\frac{p}{2}} |\frac{1}{a} I|^{-\frac{1}{2}} exp(-\frac{a}{2}\theta'\theta) d\theta \\
&\geq c a^{\frac{p}{2}}
\end{aligned}
$$

The last inequality holds for all $a < 1$, where c is a positive constant not depending on a.

Also, using (2.2), the difference of Bayes risks of $\bar{X}$ and δ_a is

$$
\begin{aligned}
r_a^* - r_a &= K\{\frac{(N+a)^p}{N^{p/2}}[|2\Gamma^{-1} + \frac{(N+a)^2}{N} I| |\frac{a}{(N+a)^2}(\frac{1}{2}\Gamma + \frac{N}{(N+a)^2} I)^{-1} + I|]^{-\frac{1}{2}} \\
&- N^{\frac{p}{2}} |NI + 2\Gamma^{-1}|^{-\frac{1}{2}}\} \\
&= K\{\frac{(N+a)^p}{N^{p/2}} |(\frac{a}{N} + 1)(2\Gamma^{-1} + \frac{(N+a)^2}{N} I) - \frac{a(N+a)^2}{N^2} I|^{-\frac{1}{2}} \\
&- N^{\frac{p}{2}} |NI + 2\Gamma^{-1}|^{-\frac{1}{2}}\} \\
&= K\left\{(N+a)^{p/2} |2\Gamma^{-1} + (N+a) I|^{-\frac{1}{2}} - N^{p/2} |NI + 2\Gamma^{-1}|^{-\frac{1}{2}}\right\}
\end{aligned}
$$

The second equality is carried out by using the relation (2.1). It can easily be verified that for $p = 1$, the ratio $\frac{r_a^* - r_a^{**}}{r_a^* - r_a}$ tends to infinity as

$a \to 0$ and for $p = 2$, this ratio tends to a positive constant as $a \to 0$. Hence, there exists an $a > 0$ such that $r_a^{**} < r_a$ which contradicts the fact that δ_a is a Bayes solution with respect to π_a. Therefore $\bar{X}$ is for $p = 1, 2$. ■

3. Inadmissibility of $\bar{X}$ for $p \geq 3$

In this section, we consider estimation of $\theta = (\theta_1, \dots, \theta_p)'$ from the model of section 1 under the loss (1.1) and find a class of estimators which have uniformly smaller risk than $\bar{X}$ for $p \geq 3$.

Lemma 3.1: *Let $X = (X_1, \dots, X_p)'$ be distributed as $N_p(\theta, I)$. If $h : \Re^p \to \Re$ is an almost differentiable function with $E\|\nabla h(X)\| < \infty$, then*

$$E[\nabla h(X)] = E[(X - \theta)h(X)]$$

, where $\nabla h(x) = \left(\frac{\partial h(x)}{\partial x_1}, \dots, \frac{\partial h(x)}{\partial x_p}\right)'$.

Proof: See Stein [11]. ■

Theorem 3.1: *Let the positive values $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p$ be the eigenvalues of the matrix Γ. If the estimator δ^c is defined as*

$$\delta^c(\bar{X}) = \left(1 - \frac{c}{\bar{X}'\Gamma^{-1}\bar{X}}\right)\bar{X}$$

where $0 < c < c^*$,$c^* = 2\left[\sum_{i=2}^{p} \frac{1}{2+N\lambda_i} - \frac{1}{2+N\lambda_1}\right]$,then $\delta^c(\bar{X})$ dominates $\bar{X}$ in terms of risks under the extended reflected normal loss function (1.1) for $p \geq 3$, when $c^* > 0$.

bf Proof: For any estimator $\delta(\bar{X})$, we define a function g as

$$g(\theta, \delta) = E\left[exp\left\{-(\delta(\bar{X}) - \theta)'\Gamma^{-1}(\delta(\bar{X}) - \theta)\right\}\right]$$

and show that for all $\theta, g(\theta,\delta^c) \geq g(\theta,\bar{X})$. We observe that

$$\begin{aligned} g(\theta,\delta^c) &= E\left[e^{-(\bar{X}-\theta)'\Gamma^{-1}(\bar{X}-\theta)}e^{-\frac{c^2}{\bar{X}'\Gamma^{-1}\bar{X}}+2c(\bar{X}-\theta)'\frac{\Gamma^{-1}\bar{X}}{\bar{X}'\Gamma^{-1}\bar{X}}}\right] \\ &\geq E\left[e^{-(\bar{X}-\theta)'\Gamma^{-1}(\bar{X}-\theta)}\left\{1-\frac{c^2}{\bar{X}'\Gamma^{-1}\bar{X}}+2c(\bar{X}-\theta)'\frac{\Gamma^{-1}\bar{X}}{\bar{X}'\Gamma^{-1}\bar{X}}\right\}\right] \end{aligned} \tag{3.1}$$

This inequality follows using the fact that

$e^{-x} \geq 1-x \qquad \forall x \in \Re$

Now by defining $\Sigma^{-1} = 2\Gamma^{-1} + NI, A = [a_{ij}]_{p\times p} = \Sigma^{1/2}\Gamma^{-1}\Sigma^{1/2}, Y = (Y_1,\ldots,Y_p)' = \Sigma^{-\frac{1}{2}}\bar{X}$ and $\beta = \Sigma^{-\frac{1}{2}}\theta$, the inequality (3.1) reduces to

$$g(\theta,\delta^c) \geq g(\theta,\bar{X}) - N^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}\left\{E\left[\frac{c^2}{Y'AY}\right] - 2cE\left[(Y-\beta)'\frac{AY}{Y'AY}\right]\right\} \tag{3.2}$$

where Y is distributed as $N_p(\beta, I)$.

Note that by using lemma 3.1, it follows that

$$\begin{aligned} E\left[(Y-\beta)'\frac{AY}{Y'AY}\right] &= E\left[\sum_{i=1}^{p}\frac{\partial}{\partial Y_i}\frac{\sum_{j=1}^{p}a_{ij}Y_j}{\sum_i\sum_j a_{ij}Y_iY_j}\right] \\ &= E\left[\frac{(\sum_i a_{ii})(\sum_i\sum_j a_{ij}Y_iY_j) - 2\sum_i(\sum_j a_{ij}Y_j)^2}{(\sum_i\sum_j a_{ij}Y_iY_j)^2}\right] \\ &= E\left[\frac{tr(A)}{Y'AY} - \frac{2Y'A^2Y}{(Y'AY)^2}\right] \end{aligned}$$

and

$$\begin{aligned} &-E\left[\tfrac{c^2}{Y'AY}\right] + 2cE\left[(Y-\beta)'\tfrac{AY}{Y'AY}\right] \\ &\qquad = E\left\{\frac{Y'[(-c^2+2ctr(A))A - 4cA^2]Y}{(Y'AY)^2}\right\} \end{aligned} \tag{3.3}$$

We know that A is a positive definite matrix and is diagonable as $U'AU = T = diag\{t_1,\ldots,t_p\}$, where the positive values $t_1,\ldots,t_p$ are the eigenvalues of $\mathbf{A}$. Now, we have $U'A^2U = T^2 = diag\{t_1^2,\ldots,t_p^2\}$ and therefore (3.3) reduces to

$-E\left[\frac{c^2}{Y'AY}\right]+2cE\left[(Y-\beta)'\frac{AY}{Y'AY}\right]$

$$= E\left\{\frac{Y'U[(-c^2+2ctr(A))T-4cT^2]U'Y}{(Y'AY)^2}\right\}$$

According to (3.2), we complete the proof by showing that the matrix

$$(-c^2+2ctr(A))T-4cT^2 = diag\{ct_1(-c+2tr(A)-4t_1),\ldots,ct_p(-c+2tr(A)-4t_p)\} \quad (3.4)$$

is positive definite when $0<c<c^*$.

It can be verified that $t_i=\frac{1}{N\lambda_i+2}; i=1,\ldots,p$, where the values $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_p$ are the eigenvalues of Γ. Hence, the diagonal elements of the diagonal matrix (3.4) is positive when

$$0<c<2tr(A)-\frac{4}{N\lambda_i+2} \quad i=1,\ldots,p$$

This condition is equivalent to $0<c<c^*$ with $c^*=2\left[\sum_{i=2}^{p}\frac{1}{2+N\lambda_i}-\frac{1}{2+N\lambda_1}\right]$.

■

Corollary 3.1: *Let the estimator* $\delta^*(\bar{X})$ *be given as*

$$\delta^*(\bar{X})=\left(1-\frac{p-2}{(N+2)\bar{X}'\bar{X}}\right)\bar{X}$$

Now, $\delta^*(\bar{X})$ *dominates* $\bar{X}$ *under the loss function (1.1) with* $\Gamma=I$*, for* $p>2$*. This estimator is similar to J-S estimator.*

Conclusions 3.1: Let the estimator $\delta^*(\bar{X})$ be given as

$$\delta^*(\bar{X})=\left(1-\frac{p-2}{(N+2)\bar{X}'\bar{X}}\right)\bar{X}$$

Now, $\delta^*(\bar{X})$ dominates $\bar{X}$ under the loss function (1.1) with $\Gamma=I$, for $p>2$. This estimator is similar to J-S estimator.

Acknowledgements: The authors would like to thank the referees for their constructive suggestions and also the Research Council of Shiraz University for their support.

References

[1] A. Bekker and J.J.J. Roux, Bayesian multivariate normal analysis with Wishart prior. *Commun. Statist. Theory Meth.*, **Vol. 24**, 10 (1995) 2485-2497.

[2] C. R. Blyth, On minimax statistical decision procedures and their admissibility, *The Ann. Math. Statist.*, **Vol. 22**, (1951) 22-42.

[3] A. C. Brandwein and W. E. Strawderman, Minimax estimation of location parameters for spherically symmetric distributions with concave loss, *The Ann. Statist.*, **Vol. 8**, 2 (1980) 279-284.

[4] G. Casella and J. T. Hwang, Limit expression for the risk of James-Stein estimators, *The Canadian J. Statist.*, **Vol. 10**, (1982) 305-309.

[5] Y. Chung and C. Kim, Simultaneous estimation of the multivariate normal mean under balanced loss function, *Commun. Statist. Theory Meth.*, **Vol. 26**, 7 (1997) 1599-1611.

[6] B. Efron and M. C. Morris, Stein's estimation rule and its competitor-an empirical Bayes approach, *J. American Statist. Assoc.,* **Vol. 68**, (1973) 117-130.

[7] W. James and C. Stein, Estimation with quadratic loss, *Proc. Fourth Berkeley Symp. Math. Statist. and Prob.*, (1961) 361-379.

[8] R. V. Leon and C. F. G. Wu, A theory of performance measures in parameter design, *Statist. Sinica*, **Vol. 2**, 2 (1992) 335-357.

[9] F. A. Spiring, The reflected normal loss function, *The Canadian J. Statist.*, **Vol. 21**, 3 (1993) 321-330.

[10] C. Stein, Inadmissibility of the usual estimator of a multivariate normal distribution, *Proc. Third Berkeley Symp. Statist. and Prob.*, **Vol. 1**, (1955) 197-206.

[11] C. Stein, Estimation of a multivariate normal distribution, *The Ann. Statist.*, **Vol. 9**, (1981) 1135-1151.

[12] W. E. Strawderman, Proper Bayes minimax estimators of the multivariate normal distribution, *The Ann. Math. Statist.*, **Vol. 42**, (1971) 385-388.

[13] M. Tribus G. Szonyi, An alternate view of the Taguchi approach, *Quality Progress, May*, (1989) 46-52.

[14] A. Zellner, Bayesian and Non-Bayesian estimation using balanced loss functions, *Statistical Decision Theory and Related Topics V*, NewYork, Springer-Verlag, (1994) 377-390.

Bulletin of the Iranian Mathematical Society
Vol. 28, No. 1, pp 67-86 (2002)

CATEGORY OF POLYGROUP OBJECTS

S.N. Hosseini, S.Sh. Mousavi and M.M. Zahedi

Dept. of Math., Shahid Bahonar University of Kerman, Kerman, IRAN.
zahedi@arg3.uk.ac.ir

Abstract: In this manuscript we generalize of the notion of polygroup in an arbitrary category $\mathcal{E}$, which is also a generalization of the notion of group object. We then define the category, $PGrp(\mathcal{E})$, of polygroup objects in $\mathcal{E}$, and we investigate some of its properties such as having limits and colimits. We also show that $PGrp(\mathcal{E})$ is a concrete category over the category $Mon(\mathcal{E})$ of monoid objects in $\mathcal{E}$, and that it has free objects and is geometric and essentially algebraic as such. Finally the preservation and reflection of epimorphism and monomorphism by the forgetful functor from $PGrp(\mathcal{E})$ to $Mon(\mathcal{E})$ is investigated.

1. Introduction

The hyperstructure theory was introduced by F. Marty in 1934 [6] at the 8th Congress of Scandinavian Mathematicians.

---

[0] *MSC (2000):* 20N20, 18A30

[0] *Keywords:* Polygroup, Polygroup object, Essentially algebraic

[0] *Received:* 28 December, 2001; *Revised:* 14 March, 2002.

[0] Supported by a Grant from Ministry of Science, Research and Technology.

Bonansinga and Corsini used this structure and introduced the notion of quasicanonical hypergroup [2], called polygroup by Comer [3], which is a generalization of the notion of a group.

The theory has found applications in many branches of mathematics such as analysis, algebra, geometry, automata and fuzzy set. In this paper we give a generalization of the above notion by categorical methods, having two goals in mind. One is to embed the category of polygroups in a complete category, as limits have not been computed in the category of polygroups yet. Another is to give the hyperstructure theory a categorical organization. Having established these goals, we have come up with several other results, such as those mentioned in the abstract. First we give some notions that are needed in the sequel.

Definition 1.1 Let $\mathcal{E}$ be a category with finite products. We call the triple $(P, *, \overline{E})$ a monoid object in $\mathcal{E}$ if:
(a) $* : P \times P \to P$ is a morphism in $\mathcal{E}$ such that the following diagram commutes.

$$\begin{array}{ccc} P^3 & \stackrel{id_P \times *}{\longrightarrow} & P^2 \\ * \times id_P \downarrow & & \downarrow * \\ P^2 & \stackrel{*}{\longrightarrow} & P \end{array}$$

(b) $\overline{E} : 1 \to P$ is a morphism in $\mathcal{E}$ such that the following diagram commutes.

$$\begin{array}{ccc} P & \stackrel{\langle id_P, \overline{E}!_P \rangle}{\longrightarrow} & P^2 \\ \langle \overline{E}!_P, id_P \rangle \downarrow & \searrow id_P & \downarrow * \\ P^2 & \stackrel{*}{\longrightarrow} & P \end{array}$$

i.e., $id_P * \overline{E}!_P = \overline{E}!_P * id_P = id_P$, where $!_P : P \to 1$ is the unique morphism from P to the terminal object 1. We call $\overline{E}$ the identity.

If $(P, *, \overline{E})$ and $(P', *', \overline{E}')$ are two monoid objects in $\mathcal{E}$, a morphism of monoid objects $f : (P, *, \overline{E}) \longrightarrow (P', *', \overline{E}')$ is a morphism $f : P \to P'$ in $\mathcal{E}$ such that $f* = *' f^2$, and $f\overline{E} = \overline{E}'$.

Proposition 1.2 *Let* $(P, *, \overline{E})$ *and* $(P', *', \overline{E}')$ *be monoid objects in*

$\mathcal{E}$ and $f : (P, *, \overline{E}) \to (P', *', \overline{E}')$ be a morphism. If $E = \overline{E}!_P$ and $E' = \overline{E}'!_{P'}$, then $fE = E'f$.

Proof: Straightforward. ■

Theorem 1.3 *The collection of monoid objects in $\mathcal{E}$ together with morphisms forms a category, which is denoted by $Mon(\mathcal{E})$.*

Proof: The proof is obvious. ■

Notation In the special case, where $\mathcal{E} = Set$ is the category of sets and
functions, we denote the category $Mon(Set)$ by Mon.

Theorem 1.4 *Let $\{(P_\alpha, *_\alpha, \overline{E}_\alpha)\}_{\alpha\in I}$ be a collection of objects in $Mon(\mathcal{E})$. If $\mathcal{E}$ has products then $(\Pi P_\alpha, \Pi *_\alpha, \Pi \overline{E}_\alpha)$ is a product of $\{(P_\alpha, *_\alpha, \overline{E}_\alpha)\}$ in $Mon(\mathcal{E})$. In particular $Mon(\mathcal{E})$ has finite products.*

Proof: Straightforward. ■

Definition 1.5 (see [7], Page 98) Let $\mathcal{E}$ be a category with finite products. We call the quadruple $(H, *, e, i)$ a group object in $\mathcal{E}$ if:
(a) $* : H \times H \to H$ is a morphism in $\mathcal{E}$ such that the following diagram commutes.

$$\begin{array}{cccc} & H^3 & \xrightarrow{1\times *} & H^2 & \\ * \times 1 & \downarrow & & \downarrow & * \\ & H^2 & \xrightarrow{*} & H & \end{array}$$

(b) $e : 1 \to H$ is a morphism in $\mathcal{E}$ such that the follwoing diagram commutes.

$$\begin{array}{ccccc} & H & \xrightarrow{\langle 1, e!_H\rangle} & H^2 & \\ \langle e!_H, 1\rangle & \downarrow & \overset{1}{\searrow} & \downarrow & * \\ & H^2 & \xrightarrow{*} & H & \end{array}$$

(c) $i : H \to H$ is a morphism in $\mathcal{E}$ such that the following diagram commutes.

$$\begin{array}{ccccc} & H & \xrightarrow{\langle 1,i\rangle} & H^2 & \\ \langle i,1\rangle & \downarrow & \overset{e!_H}{\searrow} & \downarrow & * \\ & H^2 & \xrightarrow{*} & H & \end{array}$$

We denote the quadruple $(H, *, e, i)$ by $\hat{H}$.

If $\hat{H}$ and $\hat{H}'$ are two group objects in $\mathcal{E}$. A morphism of group objects $\hat{f} : \hat{H} \longrightarrow \hat{H}'$ is defined by a morphism $f_H : H \rightarrow H'$ in $\mathcal{E}$ such that $f_H * = *' f_H$.

Theorem 1.6 *The collection of group objects in $\mathcal{E}$ together with morphisms forms a category, which is denoted by $Grp(\mathcal{E})$.*

Proof: Straightforward. ■

Notation In the special case, where $\mathcal{E} = Set$, we denote the category $Grp(Set)$ by Grp.

Theorem 1.7 (see [5], pp 237-238) *If G is a group object in $\mathcal{E}$, then each hom-set $Hom_{\mathcal{E}}(X, G)$ has the structure of a group, natural in X. Conversely, a group structure on $Hom_{\mathcal{E}}(X, G)$ for each object X of $\mathcal{E}$, natural in X, gives G the structure of an internal group. Equivalently G is a group object in $\mathcal{E}$ if and only if $Hom_{\mathcal{E}}(-, G)$ is a group object in $Set^{\mathcal{E}^{op}}$.*

Definition 1.8 A hyperstructure is a nonempty set H together with a map

$$* : H \times H \longrightarrow P^*(H)$$

which is called hyperoperation, where $P^*(H)$ denotes the set of all non-empty subsets of H.

Remark 1.9 A hyperoperation $* : H \times H \longrightarrow P^*(H)$ yields an operation $\otimes : P^*(H) \times P^*(H) \longrightarrow P^*(H)$, defined by $A \otimes B = \bigcup_{a \in A, b \in B} a * b$. Conversely an operation on $P^*(H)$ yields a hyperoperation on H, defined by $x * y = \{x\} \otimes \{y\}$.

Definition 1.10 A hyperstructure $(H, *)$ is called a polygroup if it satisfies the following conditions
(a) $x * (y * z) = (x * y) * z$, for all x, y and z in H (associative law)
(b) there exists $e \in H$, such that $e * x = x * e = \{x\}$, for all $x \in H$ (identity element)
(c) for all $x \in H$, there exists a unique element x' of H such that $e \in x * x' \cap x' * x$ (inverse element)
(d) for all x, y and z in H we have

$$z \in x * y \Rightarrow x \in z * y' \Rightarrow y \in x' * z \qquad \text{(reversibility property)}$$

A morphism from $(H, *)$ into $(H', *)$ is defined by a map $f : H \to H'$ such that $f(x * y) = f(x) * f(y)$.

Remark 1.11 In Definition 1.10 (c) the uniqueness of x' is not necessary, in fact we can obtain this property from the other conditions, provided that we replace condition (d) by $z \in x * y \Rightarrow \forall y',\ x \in z * y' \Rightarrow \forall x',\ y \in x' * z$.

Theorem 1.12 *The collection of polygroups together with morphisms forms a category, which is denoted by PG.*

Proof: Straightforward. ■

Definition 1.13 (see [4], Page 16) Let $\mathcal{E}$ be a category with finite products, and $r : A \to B^2$ be a monomorphism in $\mathcal{E}$ (that is r is a relation on B). Let $\alpha, \beta : X \to B$ be morphisms in $\mathcal{E}$. We say that $\alpha \leq_r \beta$ if there exists a morphism $h : X \to A$ in $\mathcal{E}$ such that $rh = \langle \alpha, \beta \rangle$.

Definition 1.14 (see [4], Page 16) Let $r : A \to B^2$ be a monomorphism in $\mathcal{E}$. Then we say that r is
(a) reflexive if for every morphism $\alpha : X \to B$ in $\mathcal{E}$ we have $\alpha \leq_r \alpha$.
(b) transitive if for every morphisms α, β and $\gamma : X \to B$ in $\mathcal{E}$, $\alpha \leq_r \beta$ and $\beta \leq_r \gamma$, implies that $\alpha \leq_r \gamma$.
(c) a preorder if it is reflexive and transitive.

The background category in the definition of a polygroup is the category *Set*. In order to generalize the notion of a polygroup we need to replace the category *Set* by an arbitrary category $\mathcal{E}$. To achieve this we need to free the definition of a polygroup from element taking, so we make the following observations.

Given a polygroup $(H, *)$, by Remark 1.9, we have an operation on $P^*(H)$. It can be easily verified that this operation is associative if and only if the given hyperoperation on H is associative.

The element $e \in H$ yields the function $E : P^*(H) \to P^*(H)$ taking each set A to the singleton $\{e\}$. Observe that this function factors through the terminal object. Also we have a function $e : H \to H$ taking each element x to e. If $s : H \to P^*(H)$ is the singleton function, that is the function that takes x to $\{x\}$, then it can be easily seen that $Es = se$. Let us also notice here that the singleton function satisfies the condition: $sx \subseteq sy \Rightarrow x = y$.

The existence of a unique inverse yields a function $i : H \to H$. To interpret $e \in x * x'$ in an arbitrary category, we replace it by $\{e\} \subseteq x * x'$, and since $\{e\}$ and $x * x'$ are elements of $P^*(H)$, we observe that we have a relation "$\subseteq$" on $P^*(H)$ and that $\{e\}$ is related to $x * x'$. In other words we have the relation $R = \{(A, B) : A \subseteq B\}$, the inclusion $r : R \to P^{*^2}(H)$, and that $(\{e\}, x * x') \in R$.

So we have a multiple $(H, P^*(H), R, s, r, *, E, i)$ satisfying conditions (a)-(d) of Definition 1.10, rewritten appropriately.

Using these observations we arrive at our definition of a polygroup object in an arbitrary category that is given in the next section.

2. Category of polygroup objects

Definition 2.1 Let $\mathcal{E}$ be a category with finite products. A polygroup object in $\mathcal{E}$ is a multiple $(H, P, R, s, r, *, E, i)$ where H, P and R are objects in $\mathcal{E}$ and $s, r, *, E$ and i are morphisms in $\mathcal{E}$ such that, $s : H \to P$ is a monomorphism and $r : R \to P^2$ is a preorder on P. Moreover for

all morphisms $\alpha, \beta : B \to H$ in $\mathcal{E}$ if $s\alpha \leq_r s\beta$, then $\alpha = \beta$ and:
(a) $* : P^2 \to P$ makes the following diagram commutative.

$$\begin{array}{ccc} P^3 & \xrightarrow{1\times *} & P^2 \\ *\times 1 \downarrow & & \downarrow * \\ P^2 & \xrightarrow{*} & P \end{array}$$

and hence we say that $*$ is associative.
(b) $E : P \to P$ makes the following diagram commutative.

$$\begin{array}{ccc} P & \xrightarrow{\langle 1,E\rangle} & P^2 \\ \langle E,1\rangle \downarrow & \overset{1}{\searrow} & \downarrow * \\ P^2 & \xrightarrow{*} & P \end{array}$$

That is, $1 * E = E * 1 = 1$, and also Es factors through s, i.e., there exists a morphism $e : H \to H$ in $\mathcal{E}$ such that $Es = se$. Moreover there exist morphisms $\bar{E} : 1 \to P$ and $\bar{e} : 1 \to H$ in $\mathcal{E}$ such that $E = \bar{E}!_P$ and $e = \bar{e}!_H$. We call E the identity.
(c) $i : H \to H$ satisfies $Es \leq_r 1 * i$ and $Es \leq_r i * 1$ where $1 * i = *s^2\langle 1, i\rangle$ and $i * 1 = *s^2\langle i, 1\rangle$. We call i an inverse.
(d) For all morphisms α, β and $\gamma : B \to H$ in $\mathcal{E}$ we have the following implications:

$$s\alpha \leq_r s\beta * s\gamma \Rightarrow s\beta \leq_r s\alpha * si\gamma \Rightarrow s\gamma \leq_r si\beta * s\alpha$$

In this definition we denote the multiple $(H, P, R, s, r, *, E, i)$ by $\hat{H}$.

Theorem 2.2 *Let* $\hat{H} = (H, P, R, s, r, *, E, i)$ *be a polygroup object in* $\mathcal{E}$ *and* $\star : H^2 \to H$ *be a morphism in* $\mathcal{E}$ *such that* $s\star = *s^2$*. Then* $(H, \star, e_H, i)$ *is a group object in* $\mathcal{E}$.

Proof: It is enough to show that $1 \star i = e$. From Definition 2.1(c) it follows that $Es \leq_r s(1 \star i)$ and so $se \leq_r s(1 \star i)$. Thus $1 \star i = e$. Similarly $i \star 1 = e$. ■

Theorem 2.3 *Let* $(H, *, e, i)$ *be a group object in* $\mathcal{E}$*. Set* $H = P = R$ *and let* $r = \Delta = \langle 1, 1\rangle$ *be the diagonal morphism from* H

into H^2, *and* $s = id_H$ *be the identity morphism on* H *in* $\mathcal{E}$. *Then* $(H, H, H, id_H, \Delta, *, e, i)$ *is a polygroup object in* $\mathcal{E}$.

Proof: Straightforward. ■

Proposition 2.4 *Let* $\hat{H}$ *be a polygroup object in* $\mathcal{E}$, *then the following statements hold:*
(i) $Es = se = \overline{E}!_H$
(ii) $s\overline{e} = \overline{E}$,
(iii) $1 * \overline{E}!_P = 1$,
(iv) $i^2 = 1$,
(v) $i\overline{e} = \overline{e}$,
(vi) $ie = ei = e$,
(vii) $E^2 = E$, *and*
(viii) $E * E = E$.

Proof: Straightforward. ■

Notation Let $\mathcal{E}$ be a category and A be an object in $\mathcal{E}$. We denote the functor $Hom(-, A) : \mathcal{E}^{op} \longrightarrow Set$ by $\overline{A}$ and if $f : A \rightarrow B$ is a morphism in $\mathcal{E}$ we denote the natural transformation $Hom(-, f) : \overline{A} \stackrel{\cdot}{\longrightarrow} \overline{B}$, by $\overline{f}$.

Lemma 2.5 *Let* $r : R \rightarrow P^2$ *be a monomorphism in* $\mathcal{E}$. *For all objects* F *and morphisms* $f, g : F \stackrel{\cdot}{\longrightarrow} \overline{P}$ *in* $Set^{\mathcal{E}^{op}}$ *we have*

$$f \leq_{\overline{r}} g \Longleftrightarrow \forall A \in \mathcal{E}, \forall x \in F(A), f_A(x) \leq_r g_A(x).$$

Proof: Let $f \leq_{\overline{r}} g$, so there exists a morphism $h : F \stackrel{\cdot}{\longrightarrow} \overline{R}$ in $Set^{\mathcal{E}^{op}}$ such that $\overline{r}h = \langle f, g\rangle$. Thus we get $h_A(x) : A \longrightarrow R$, so $f_A(x) \leq_r g_A(x)$, for all $A \in \mathcal{E}$ and $x \in F(A)$. Thus $f_A(x) \leq_r g_A(x)$.

Now suppose for all $A \in \mathcal{E}$ and $x \in F(A)$ we have $f_A(x) \leq_r g_A(x)$. So there exists a morphism $h_{A,x} : A \rightarrow R$ such that $rh_{A,x} = \langle f_A(x), g_A(x)\rangle$ $\forall A \in \mathcal{E}$ and $x \in F(A)$. Now define $h_A : F(A) \longrightarrow Hom(A, R)$ by $h_A(x) = h_{A,x}$. It easily follows that $h : F \rightarrow \overline{R}$ is a natural transformation and $(\overline{r}h)_A(x) = \langle f, g\rangle_A(x)$, so $f \leq_{\overline{r}} g$. ■

Lemma 2.6 *Let $r : R \to P^2$ be a monomorphism in $\mathcal{E}$. Then for all morphisms $\alpha, \beta : B \to P$ in $\mathcal{E}$ we have: $\alpha \leq_r \beta \iff \overline{\alpha} \leq_{\overline{r}} \overline{\beta}$.*

Proof: First suppose that $\alpha \leq_r \beta$. So there exists a morphism $h : B \to R$ such that $rh = \langle \alpha, \beta \rangle$. If $A \in \mathcal{E}$ and $x : A \to B$ is a morphism in $\mathcal{E}$, so we get $(\overline{r}\overline{h})_A(x) = \overline{r}_A(\overline{h}_A(x))$. Thus $\overline{\alpha} \leq_{\overline{r}} \overline{\beta}$.

Now suppose $\overline{\alpha} \leq_{\overline{r}} \overline{\beta}$. Then there exists a morphism $\overline{h} : \overline{B} \longrightarrow \overline{R}$ in $Set^{\mathcal{E}^{op}}$ such that $\overline{r}\overline{h} = \langle \overline{\alpha}, \overline{\beta} \rangle$. By Yoneda lemma we have $\overline{h} = Hom(-, h)$, where $h = \overline{h}_B(1)$. So we have $(\overline{r}\overline{h})_B(1) = \langle \alpha, \beta \rangle$, thus $rh = \langle \alpha, \beta \rangle$. Therefore $\alpha \leq_r \beta$. ∎

Remark 2.7 Let $f' : \overline{A} \to \overline{B}$ be a morphism in $Set^{\mathcal{E}^{op}}$. By Yoneda Lemma we get $f' = \overline{f}$,where $f = f'_A(1_A)$.

Theorem 2.8 *Let $H, P, R, s, r, *, E, i$ be as in the statement of Definition if $\hat{\overline{H}} = (\overline{H}, \overline{P}, \overline{R}, \overline{s}, \overline{r}, \overline{*}, \overline{e}, \overline{i})$ is a polygroup 2.1. Then $\hat{H} = (H, P, R, s, r, *, E, i)$ is a polygroup object in $\mathcal{E}$ if and only object in $Set^{\mathcal{E}^{op}}$.*

Proof: Follows from Lemmas 2.5 and 2.6. ∎

Remark 2.9 Theorem 2.8 is the generalized version of Theorem 1.7.

Definition 2.10 Let $\hat{H}$ and $\hat{H}'$ be polygroup objects in $\mathcal{E}$. A morphism $\hat{f} : \hat{H} \to \hat{H}'$ is a triple (f_H, f_P, f_R) where $f_H : H \to H'$, $f_P : P \to P'$ and $f_R : R \to R'$ are morphisms in $\mathcal{E}$ such that
(a) $f_P s = s' f_H$,
(b) $f_P^2 r = r' f_R$,
(c) $f_P * = *' f_P^2$,
(d) $f_P E = E' f_P$.

Theorem 2.11 *If $\hat{f} = (f_H, f_P, f_R) : \hat{H} \to \hat{H}'$ is a morphism, then:*
(i) $f_H e = e' f_H$,
(ii) $f_H i = i' f_H$.

Proof: (i) Straightforward.

(ii) We know that $Es \leq_r 1 * i$, thus:

$$\begin{aligned} & \overline{E}!_H \leq_r 1 * i \text{ ; by Proposition 2.4(i)} \\ \Rightarrow\ & f_P s\overline{e}!_H \leq_r f_P * s^2\langle 1, i\rangle \text{ ; by Proposition 2.4(ii)} \\ \Rightarrow\ & s' f_H \overline{e}!_H \leq_r f_P * s^2\langle 1, i\rangle \text{ ; by Definition 2.10(a) and Proposition 2.4(i)} \\ \Rightarrow\ & E's' f_H \leq_r f_P * s^2\langle 1, i\rangle \text{ ; by Theorem 2.11(i) and Proposition 2.4(i)} \\ \Rightarrow\ & E's' f_H \leq_r *'(s' f_H)^2\langle 1, i\rangle \text{ ; by Definition 2.10(a)and(c)} \\ \Rightarrow\ & s'\overline{e}'!_H \leq_r s' f_H * s' f_H i \text{ ; by Proposition 2.4(i)-(ii)} \\ \Rightarrow\ & s' f_H i \leq_r s' i' f_H * s'\overline{e}'!_H . \text{ ; by Proposition 2.4(ii)-(iii)} \end{aligned}$$

Hence $f_H i = i' f_H$. ■

Lemma 2.12 *Let $\hat{H}$ and $\hat{G}$ be two polygroup objects in $\mathcal{E}$ and $\hat{f} : \hat{H} \to \hat{G}$ be a morphism. If $!_H : H \to 1$ is an epimorphism then $f_H \overline{e} = \overline{e}'$.*

Proof: Straightforward. ■

Theorem 2.13 *The collection of all polygroup objects in $\mathcal{E}$ together with polygroup morphisms forms a category, which is denoted by $PGrp(\mathcal{E})$.*

Sketch of proof: For $\hat{f} : \hat{H} \to \hat{H}'$ and $\hat{g} : \hat{H}' \to \hat{H}''$, the composition $\hat{f} \circ \hat{g}$ is defined by $(f_H \circ g_{H'}, f_P \circ g_{P'}, f_R \circ g_{R'})$. The identity morphism $\hat{id} : \hat{H} \to \hat{H}$ is defined to be the triple (id_H, id_P, id_R), where id_H, id_P, and id_R are the identities on H, P and R respectively in $\mathcal{E}$. ■

Notation In the special case, where $\mathcal{E} = Set$, we denote $PGrp(Set)$ by $PGrp$.

Remark 2.14 By Theorems 2.2 and 2.3, the category $Grp(\mathcal{E})$ can be embedded in the category $PGrp(\mathcal{E})$, i.e. $Grp(\mathcal{E}) \subseteq PGrp(\mathcal{E})$. It then follows that

$$Mon \subseteq Grp \subseteq PG \subseteq PGrp$$

Theorem 2.15 *Let $\{\hat{H}_\alpha\}_{\alpha \in I}$ be a collection of objects in $PGrp(\mathcal{E})$, where*

$$\hat{H}_\alpha = (H_\alpha, P_\alpha, R_\alpha, s_\alpha, r_\alpha, *_\alpha, E_\alpha, i_\alpha)$$

If $\mathcal{E}$ has products then

$$\Pi\hat{H}_\alpha = (\Pi H_\alpha, \Pi P_\alpha, \Pi R_\alpha, \Pi s_\alpha, \Pi r_\alpha, \Pi *_\alpha, \Pi E_\alpha, \Pi i_\alpha)$$

is a product of $\{\hat{H}_\alpha\}$ in $PGrp(\mathcal{E})$. In particular $GPGrp(\mathcal{E})$ has finite products.

Proof: Straightforward. ∎

Theorem 2.16 *Let $\mathcal{E}$ be a category such that for all objects H in $\mathcal{E}$, $!_H : H \longrightarrow 1$ is an epimorphism. If $\mathcal{E}$ has equalizers, then so does $PGrp(\mathcal{E})$.*

Proof: Let $\hat{H} = (H, P, R, s, r, *, E, i)$ and $\hat{G} = (G, P', R', s', r', *', E', i')$ be two polygroup objects. Given a pair of morphisms $\hat{f}, \hat{g} : \hat{H} \to \hat{G}$ in $PGrp(\mathcal{E})$, let $\epsilon_K : K \to H$, $\epsilon_{P''} : P'' \to P$, and $\epsilon_{R''} : R'' \to R$ be equalizers of (f_H, g_H), (f_P, g_P), and (f_R, g_R), respectively. Using the fact that ϵ_K, $\epsilon_{P''}$, and $\epsilon^2_{P''}$ are equalizers, we obtain the morphisms $s'' : K \to P''$, $r'' : R'' \to P''^2$, $*'' : P''^2 \to P''$, $E'' : P'' \to P''$, and $i'' : K \to K$. It is straightforward though tedious to show that $\hat{K} = (K, P'', R'', s'', r'', *'', E'', i'')$ is an equalizer of $\hat{f}, \hat{g}$. ∎

Corollary 2.17 *Let $\mathcal{E}$ be a category in which for all objects H in $\mathcal{E}$, $!_H : H \longrightarrow 1$ is an epimorphism. If $\mathcal{E}$ has limits then so does $PGrp(\mathcal{E})$.*

Proof: Since $\mathcal{E}$ has products and equalizers, by Theorems 2.15 and 2.16, so does $PGrp(\mathcal{E})$. Therefore $PGrp(\mathcal{E})$ has limits. ∎

3. Free Polygroup Objects

Proposition 3.1 *There exists a faithful functor U_1 from $PGrp(\mathcal{E})$ into $Mon(\mathcal{E})$, where $U_1(\hat{H} \xrightarrow{\hat{f}} \hat{H}') = ((P, *, \overline{E}) \xrightarrow{f_P} (p', *', \overline{E}'))$, that is $PGrp(\mathcal{E})$ is concrete over $Mon(\mathcal{E})$.*

Proof: Let $\hat{H} = (H, P, R, s, r, *, E, i)$ and $\hat{H}' = (H', P', R', s', r', *', E', i')$ be two arbitrary objects in $PGrp(\mathcal{E})$ and $\hat{f} = (f_H, f_P, f_R)$ be a morphism from $\hat{H}$ into $\hat{H}'$ in $PGrp(\mathcal{E})$. Then it is easy to see that $U_1 : PGrp(\mathcal{E}) \to$

$Mon(\mathcal{E})$ is a functor. Now we show that U_1 is faithful. Let $\hat{f}, \hat{g} : \hat{H} \to \hat{H}'$ be morphisms in $PGrp(\mathcal{E})$, such that $U_1(\hat{f}) = U_1(\hat{g})$. Thus $f_P = g_P$. From Definition 2.10(a) we know that $f_P s = s' f_H$ and $g_P s = s' g_H$. So $s' f_H = s' g_H$ and since s' is a monomorphism we get that $f_H = g_H$. Also we have $f_P^2 r = r' f_R$ and $g_P^2 r = r' g_R$, by Definition 2.10 (b), so $r' f_R = r' g_R$. Since r' is a monomorphism, thus $f_R = g_R$. Therefore $\hat{f} = \hat{g}$, that is U_1 is faithful. ■

Definition 3.2 Let $(P, *, \overline{E})$ be an object in $Mon(\mathcal{E})$ and $\Delta : P \to P^2$ be the diagonal morphism. Then $(1, P, P, \overline{E}, \Delta, *, \overline{E}!_P, id_1)$ is an object in $PGrp(\mathcal{E})$. We denote this object by $\hat{P}^1$.

Theorem 3.3 *The concrete category* $(PGrp(\mathcal{E}), U_1)$ *has free objects.*

Proof: Let $(P, *, \overline{E})$ be an object in $Mon(\mathcal{E})$. We claim that $\hat{P}^1$ is a free object over $(P, *, \overline{E})$. For this reason we show that the identity morphism

$$id_{(P,*,\overline{E})} : (P, *, \overline{E}) \longrightarrow U_1(\hat{P}^1) = (P, *, \overline{E})$$

is a universal arrow over $(P, *, \overline{E})$. Let $\hat{H}' = (H', P', R', s', r', *', E', i')$ be another object in $PGrp(\mathcal{E})$ and $g : (P, *, \overline{E}) \longrightarrow U_1(\hat{H}') = (P', *', \overline{E})$ be a morphism in $Mon(\mathcal{E})$. Thus by Definition 1.1 we have $g* = *' g^2$ and $g\overline{E}!_P = (\overline{E}'!_{P'})g$. Since $\hat{H}'$ is an object in $PGrp(\mathcal{E})$ thus by Definition 2.1(b) we get $\overline{e}'!_{H'} = e'$ and we know that $g \leq_\Delta g$, thus there exists a morphism $g' : P \to R'$ such that $r'g' = \langle g, g \rangle = g^2\Delta$. It is easy to check that $\hat{g} := (\overline{e}', g, g')$ is a morphism from $\hat{P}$ into $\hat{H}'$ in $PGrp(\mathcal{E})$, and $U_1(\hat{g}) = g$, which implies that $U_1(\hat{g}) \circ id_{(P,*,\overline{E})} = g$. Now suppose that $\hat{f} = (f_1, f_2, f_3)$ be an arbitrary morphism from $\hat{P}$ into $\hat{H}$ such that $U_1(\hat{f}) \circ id_{(P,*,\overline{E})} = g$. Therefore we get $U_1(\hat{f}) = g = U_1(\hat{g})$. Since U_1 is faithful, thus $\hat{f} = \hat{g}$. So $\hat{g}$ is a unique morphism such that $U_1(\hat{g}) \circ id_{(P,*,\overline{E})} = g$. Therefore $\hat{P}^1$ is a free object over$(P, *, \overline{E})$. ■

Proposition 3.4 *The mapping* F_1 *defined by*

$$F_1((P, *, \overline{E}) \xrightarrow{f} (P', *', \overline{E}')) = (\hat{P}^1 \overset{(id_1, f, f)}{\longrightarrow} \hat{P'}^1)$$

is a functor from $Mon(\mathcal{E})$ into $PGrp(\mathcal{E})$.

Proof: Straightforward. ■

Lemma 3.5

(i) $\eta_{1_{(P,*,\overline{E})}} = id_{(P,*,\overline{E})} : (P,*,\overline{E}) \longrightarrow U_1F_1(P,*,\overline{E}) = (P,*,\overline{E})$ *defines* $\eta_1 : Id_{Mon(\mathcal{E})} \longrightarrow U_1F_1$ *as a natural transformation.*

(ii) For every object $\hat{H} = (H,P,R,s,r,*,E,i)$ *in* $PGrp(\mathcal{E})$*, the mapping* $\epsilon_{1_{\hat{H}}} = (e', id_P, h) : F_1U_1(\hat{H}) = \hat{P}^1 \to \hat{H}$ *defines* $\epsilon_1 : F_1U_1 \to Id_{PGrp(\mathcal{E})}$ *as a natural transformation, where* $h : P \to R$ *is a morphism in* $\mathcal{E}$ *such that* $rh = \langle id_P, id_P\rangle$.

Proof: Straightforward. ■

Theorem 3.6 *Suppose that* U_1*,* F_1*,* η_1 *and* ϵ_1 *as in Propositions 3.1, 3.4 and Lemma 3.5 respectively. Then we have* $(\eta_1, \epsilon_1) : F_1 \vdash U_1 : (PGrp(\mathcal{E}), Mon(\mathcal{E}))$ *is an adjoint situation.*

Proof: By Lemma 3.5 it is enough to show that $U_1\epsilon_1 \circ \eta_1 U_1 = id_{U_1}$ and $\epsilon_1 F_1 \circ F_1\eta_1 = id_{F_1}$. We only show the former, the latter is similar. Let
$\hat{H} = (H,P,R,s,r,*,E,i)$ be an object in $PGrp(\mathcal{E})$. So we have

$$\begin{aligned}(U_1\epsilon_1 \circ \eta_1 U_1)(\hat{H}) &= U_1\epsilon_1(\hat{H}) \circ (\eta_1 U_1)(\hat{H}) \\ &= U_1\epsilon_{1_{\hat{H}}} \circ \eta_{1_{U_1(\hat{H})}} \\ &= id_{(P,*,\overline{E})} \circ id_{(P,*,\overline{E})} \text{ ; by Proposition 3.1} \\ &= U_1(\overline{e}, id_P, h) \circ id_{(P,*,\overline{E})} \text{ ; by Lemma 3.5 (i)} \\ &= id_{U_1(\hat{H})}.\end{aligned}$$

So $U_1\epsilon_1 \circ \eta_1 U_1 = id_{U_1}$. ■

Theorem 3.7 *Let* $T_1 = (T_1, \eta_1, \mu_1)$ *be the monad associated with the adjoint situation given in Theorem 3.6, where* $T_1 = U_1F_1$*. Then* T_1 *is the trivial monad.*

Proof: Let x be an object in the Eilenberg-Moore category $(Mon(\mathcal{E})^{T_1}, U_1^{T_1})$ [see 1]. So we have $x : T_1(P,*,\overline{E}) \to (P,*,\overline{E})$, for some object $(P,*,\overline{E})$

in $Mon(\mathcal{E})$, that satisfies
(a) $x \circ \eta_{1_{(P,*,\overline{E})}} = id_{(P,*,\overline{E})}$, and
(b) $x \circ T_1 x = x \circ \mu_{1_{(P,*,\overline{E})}}$.

By (a) and Lemma 3.5 (i) we have $x = id_{(P,*,\overline{E})}$. Since $T_1 = U_1 F_1$, so we get $T_1 = Id_{Mon(\mathcal{E})}$, and by condition (b) we conclude that $\mu_1 = id_{T_1}$. Thus we have $\mu_1 = \eta_1 = id_{T_1}$. Then the monad T_1 is trivial monad $(Id_{Mon(\mathcal{E})}, id_{T_1}, id_{T_1})$. ■

Corollary 3.8 *The Eilenberg-Moore category* $(Mon(\mathcal{E})^{T_1}, U_1^{T_1})$ *is concretely isomorphic to* $(Mon(\mathcal{E}), Id_{Mon(\mathcal{E})})$.

Proof: Straightforward. ■

Theorem 3.9 *Let* $(\eta_2, \epsilon_2) : F_2 \to U_2 : Mon(\mathcal{E}) \to \mathcal{E}$ *be an adjoint situation and* $T_2 = (T_2, \eta_2, \mu_2)$ *be its associated monad. Then the Eilenberg-Moore category* $(\mathcal{E}^{T_2}, U_2^{T_2})$ *is concretely isomorphic to the category* $Mon(\mathcal{E})$.

Sketch of Proof: If $x : T_2 P \to P$ is a T_2-algebra, multiplication in P is defined by $* = x \circ U_2 *_2 \circ l_P \times l_P$ where $l_P : P \to U_2 F_2(P)$ is an universal arrow and $F_2(P) = (F_2 P, *_2, \overline{E}_2)$ is a free object over P in $Mon(\mathcal{E})$. And $\overline{E}$ defined by $\overline{E} = x \circ U_2 \overline{E}_2$. Then $(P, *, \overline{E})$ is an object in $Mon(\mathcal{E})$. ■

Let $(\eta_1, \epsilon_1) : F_1 \vdash U_1 : (PGrp(\mathcal{E}), Mon(\mathcal{E}))$ be adjoint situation as in Theorem 3.6. Let $U = U_2 U_1$, $F = F_1 F_2$, $\eta = U_2 \eta_1 F_2 \circ \eta_2$, $\epsilon = \epsilon_1 \circ F_1 \epsilon_2 U_1$ and $\mu = U \epsilon F$. Since composition of adjoint situations is an adjoint situations (see Proposition 19.13[1]), so we have $(\eta, \epsilon) : F \vdash U : (PGrp(\mathcal{E}), \mathcal{E})$ is an adjoint situation. Suppose that $T = (T, \eta, \mu)$ be its associated monad, then we have the following theorem:

Theorem 3.10 *The Eilenberg-Moore category* $(\mathcal{E}^T, U^T)$ *is concretely isomorphic to the* $Mon(\mathcal{E})$.

Proof: The proof is obvious by Theorems 3.8 and 3.9. ■

Proposition 3.11 *If the category $\mathcal{E}$ has finite products, then the functor*
$F_1 : Mon(\mathcal{E}) \to PGrp(\mathcal{E})$ given in Proposition 3.4, preserves finite products.

Proof: Let $\{(P_\alpha, *_\alpha, \overline{E}_\alpha)\}_{\alpha \in I}$ be a finite family of objects in $Mon(\mathcal{E})$. Then by Theorem 1.4 the object $(\Pi P_\alpha, \Pi *_\alpha, \Pi \overline{E}_\alpha)$ is a product of the above family in $Mon(\mathcal{E})$. Let $\Delta : \Pi P_\alpha \to (\Pi P_\alpha)^2$ be the diagonal morphism, then $Pr_\alpha^2 \Delta = \Delta_\alpha Pr_\alpha$, for all $\alpha \in I$, where $Pr_\alpha : \Pi P_\alpha \to P_\alpha$ is the canonical projection morphism. By Theorem 2.15 $F(\Pi P_\alpha, \Pi *_\alpha, \Pi \overline{E}_\alpha) = \widehat{\Pi P_\alpha}^1 = (1, \Pi P_\alpha, \Pi P_\alpha, \Pi \overline{E}_\alpha, \Delta, \Pi *_\alpha, \Pi \overline{E}_\alpha !_{\Pi P_\alpha}, id_1)$ together with the canonical projection $\widehat{P_{r_\alpha}} = (id_1, Pr_\alpha, Pr_\alpha)$ is a product of the family $\{F(P_\alpha, *_\alpha, \overline{E}_\alpha) = \widehat{P_\alpha}^1 = (1, P_\alpha, P_\alpha, \overline{E}_\alpha, \Delta_\alpha, *_\alpha, \overline{E}_\alpha !_{P_\alpha}, id_1)\}$ in $PGrp(\mathcal{E})$. ■

Proposition 3.12 *Let $\mathcal{E}$ have equalizers. If the following diagram*

$$(P, *, \overline{E}) \xrightarrow{f} (P', *', \overline{E}') \underset{h}{\overset{g}{\rightrightarrows}} (P'', *'', \overline{E}'') \ (1)$$

is an equalizer in $Mon(\mathcal{E})$, then

$$F_1(P, *, \overline{E}) = \hat{P}^1 \xrightarrow{\hat{f} = (id_1, f, f)} F_1(P', *', \overline{E}') = \hat{P'}^1 \underset{\hat{h} = (id_1, h, h)}{\overset{\hat{g} = (id_1, g, g)}{\rightrightarrows}} F_1(P'', *'', \overline{E}'') = \hat{P''}^1$$

is an equalizer in $PGrp(\mathcal{E})$.

Proof: Since $gf = hf$, so $\hat{g}\hat{f} = \hat{h}\hat{f}$. Let $\hat{H}_1 = (H_1, P_1, R_1, s_1, r_1, *_1, E_1, i_1)$ be an object in $PGrp(\mathcal{E})$ and $\hat{k}_1 = (k_{H_1}, k_{P_1}, k_{R_1}) : \hat{H}_1 \to \hat{P'}^1$ be a morphism in $PGrp(\mathcal{E})$, such that $\hat{g}\hat{k}_1 = \hat{h}\hat{k}_1$. By Theorem 2.13 we have $gk_{P_1} = hk_{P_1}$ and $gk_{R_1} = hk_{R_1}$.

Since $(P_1, *_1, \overline{E}_1)$ is an object in $Mon(\mathcal{E})$ and the diagram (1) is an equalizer in $Mon(\mathcal{E})$, so we get the unique morphism $t_{P_1} : (P_1, *_1, \overline{E}_1) \to (P, *, \overline{E})$ in $Mon(\mathcal{E})$, such that $ft_{P_1} = k_{P_1}$. Let $(F_2(R_1), *_2, \overline{E}_2)$ be the free monoid over R_1 in $Mon(\mathcal{E})$, so we have the unique morphism $h' : (F_2(R_1), *_2, \overline{E}_2) \to (P', *', \overline{E}')$ in $Mon(\mathcal{E})$ such that $U_2(h') \circ l = k_{R_1}$, where $l : R_1 \to U_2F_2(R_1)$ is a U_2-universal morphism in $\mathcal{E}$. Thus we have $h' \circ l = k_{R_1}$. Since $g : (P', *', \overline{E}') \to (P'', *'', \overline{E}'')$ is a morphism in $Mon(\mathcal{E})$, then we get $gh' = hh'$. So by diagram (I) we get a unique morphism $t' : (F_2(R_1), *_2, \overline{E}_2) \to (P, *, \overline{E})$ in $Mon(\mathcal{E})$ such that $ft' = h'$. Now, define $t_{R_1} := U_2(t') \circ l$, that is, $t_{R_1} = t' \circ l$. Since the concrete category $(Mon(\mathcal{E}), U_2)$ over $\mathcal{E}$, has free objects, and f is a monomorphism in $Mon(\mathcal{E})$, then by Theorem 8.38 [1], we have $U_2(f) = f$ is a monomorphism in $\mathcal{E}$. Now it easily follows that $\hat{t}_1 = (!_{H_1}, t_{P_1}, t_{R_1}) : \hat{H}_1 \to \hat{P}$ is a morphism in $PGrp(\mathcal{E})$. Also $\hat{f}\hat{t}_1 = \hat{k}_1$, and $\hat{t}_1$ is a unique morphism such that $\hat{f}\hat{t}_1 = \hat{k}_1$. ∎

Theorem 3.13 *If $\mathcal{E}$ has finite limits, then the pair (F_1, U_1) where $F_1 \vdash U_1$ is a geometric morphism.*

Proof: The proof is obvious by Propositions 3.11 and 3.12, see also page 26 of [4]. ∎

Remark 3.14 By Theorem 3.3, the concrete category $(PGrp(\mathcal{E}), U_1)$ over

$Mon(\mathcal{E})$ has free objects. So by Theorem 8.38 [1], we get that U_1 preserves and reflects monomorphisms, and by Proposition 7.44 [1], U_1 reflects epimorphisms.

Example 3.15 We know that $(Z, 0, \overline{E})$ and $(Q, 0, \overline{E}')$ are objects in Mon. Let f be the inclusion homomorphism from $(Z, 0, \overline{E})$ into $(Q, 0, \overline{E}')$, i.e., $f(n) = n$ for all $n \in Z$. We have that f is an epimorphism, by some manipulation (see Example 7.40 (5) of [1]). But $f : Z \hookrightarrow Q$ as a function in Set is not an epimorphism. We know that $\hat{Z}^1$ and $\hat{Q}^1$ are objects in $PGrp$, and $\hat{f} = (id_1, f, f) : \hat{Z}^1 \to \hat{Q}^1$ is a morphism in $PGrp$. It is easy to see that $\hat{f}$ is an epimorphism in $PGrp$. But $U(\hat{f}) = f : Z \hookrightarrow Q$, as a morphism in Set, is not an epimorphism.

Notation The full subcategory of $PGrp$ whose objects are $(1, P, R, s, r, *, E, id_1)$ is denoted by P_1Grp.

Theorem 3.16 *If $\hat{f} = (f_1, f_2, f_3) : \hat{P} \to \hat{P}'$ is an epimorphism in P_1Grp, then $U_1(\hat{f}) = f_2$ is an epimorphism in Mon.*

Proof: Let $(P,*,\overline{E}) \xrightarrow{f_2} (P',*',\overline{E}') \overset{g_2}{\underset{h_2}{\rightrightarrows}} (P'',*'',\overline{E}'')$ be a diagram in Mon, such that $g_2f_2 = h_2f_2$. We know that $r' : R' \to P'^2$ and $r'(R')$ is the set defined by $\{(p'_1,p'_2) | \exists x \in R'$, such that $r'(x) = (p'_1,p'_2)\}$. Now let R'' be the smallest preorder relation on the set

$$\{(g_2(p'_1), g_2(p'_2)) | (p'_1,p'_2) \in r'(R')\} \cup \{(h_2(p'_1), h_2(p'_2)) | (p'_1,p'_2) \in r'(R')\}$$

and $r'' : R'' \hookrightarrow P''^2$ be the inclusion map. Define $s'' : 1 \to P''$ by $s'' = g_2s'$. Thus s'' is a monomorphism in Set. Since for any morphisms $\alpha,\beta : B \to 1$ in Set, we have $\alpha = \beta$, therefore if $s''\alpha \leq_{r''} s''\beta$, then $\alpha = \beta$. Now define $\hat{P}'' = (1, P'', R'', s'', r'', *'', \overline{E}''!_{P''}, id_1)$. It is easy to check that $\hat{P}''$ is an object in $P_1Grp(Set)$. It can easily be checked that $\hat{g} = (id_1, g_2, g_2^2r')$ and $\hat{h} = (id_1, h_2, h_2^2r')$ are morphisms in P_1Grp from $\hat{P}'$ to $\hat{P}''$. Therefore $\hat{h}\hat{f} = \hat{g}\hat{f}$, and since $\hat{f}$ is an epimorphism in P_1Grp thus $\hat{h} = \hat{g}$. That implies $g_2 = h_2$, so $f_2 : (P,*,\overline{E}) \to (P',*',\overline{E}')$ is an epimorphism in Mon. ■

Proposition 3.17 *Let $\hat{f} = (f_H, f_P, f_R) : \hat{H} \to \hat{H}'$ be a morphism in $PGrp(\mathcal{E})$. $\hat{f}$ is an isomorphism in $PGrp(\mathcal{E})$ if and only if f_H, f_P and f_R are isomorphisms in $\mathcal{E}$.*

Proof: Straightforward. ■

Remark 3.18 In Example 3.15, we showed that $\hat{f} = (id_1, f, f) : \hat{Z}^1 \to \hat{Q}^1$ is an epimorphism in $PGrp$. Since the concrete category $(PGrp, U)$ over Set has free objects, so by Theorem 8.38 [2] we get $\hat{f}$ is also a monomorphism in $PGrp$. But $\hat{f}$ is not an isomorphism in $PGrp$, because if $\hat{f}$ is an isomorphism in $PGrp$, then $U(\hat{f}) = f : Z \hookrightarrow Q$ is an isomorphism in Set, which is a contradiction. So the category $PGrp$ is not balanced, and so it is not a topos.

4. Essentially Algebraic Property

We start this section by assuming that the forgetful functor $U_2 : Mon(\mathcal{E}) \to \mathcal{E}$ is (generating, Mono-source)-factorizable. For example, we know that the forgetful functor $U_2 : Mon \to Set$ is an adjoint, and so by Proposition 18.3 [1], it is (generating,-)-factorizable, hence it is (generating, Mono-source)-factorizable.

Throughout this Section we use $U = U_2U_1 : PGrp(\mathcal{E}) \to \mathcal{E}$, where U_1 and U_2 are the functors introduced in section 3.

Theorem 4.1: *The functor $U : PGrp(\mathcal{E}) \to \mathcal{E}$ is (generating, Mono-source)-factorizable.*

Proof: Let $\{\hat{H}_j = (H_j, P_j, R_j, s_j, r_j, *_j, E_j, i_j)\}_{j\in I}$ be a collection of objects in $PGrp(\mathcal{E})$, and $(X \xrightarrow{f_j} U(\hat{H}_j) = P_j)_{j\in I}$ be a U-structure source. Since the functor $U_2 : Mon(\mathcal{E}) \to \mathcal{E}$ is (generating, Mono-source)-factorizable, thus there exists a U_2-generating morphism $e : X \to U_2(G, *, \overline{E})$ and a mono-source $((G, *, \overline{E}) \xrightarrow{m_j} (P_j, *_j, \overline{E}_j))_{j\in I}$ in $Mon(\mathcal{E})$, such that $f_j = U_2(m_j)\circ e$; for all $j \in I$. On the other hand, $U_2(G, *, \overline{E}) = U_2U_1F_1(G, *, \overline{E}) = U(\hat{G}^1)$. Therefore we have the morphism $e : X \to U(\hat{G}^1)$. We claim that e is U-generating. To show this, let $\hat{G}^1 \underset{\hat{h}}{\overset{\hat{g}}{\rightrightarrows}} \hat{K}$ be morphisms in $PGrp(\mathcal{E})$, such that $U(\hat{g})\circ e = U(\hat{h})\circ e$. So $U_2(U_1(\hat{g}))\circ e = U_2(U_1(\hat{h})) \circ e$, since e is U_2-generating thus $U_1(\hat{g}) = U_1(\hat{h})$. Therefore $\hat{g} = \hat{h}$, because U_1 is faithful. Now we want to get a mono-source $(\hat{G}^1 \xrightarrow{\hat{m}_j} \hat{H}_j)_{j\in I}$ in $PGrp(\mathcal{E})$. Since $r_j : R_j \to P_j^2$ is reflexive, for all $j \in I$, thus $m_j \leq_{r_j} m_j$. Hence for all $j \in I$, there exists a morphism $h_j : G \to R_j$ in $\mathcal{E}$, such that $r_jh_j = m_j^2\Delta$. Now define $\hat{m}_j : \hat{G}^1 \to \hat{H}_j$ by $\hat{m}_j = (\hat{e}_j, m_j, h_j)$. By Proposition 2.4, it is easy to check that $\hat{m}_j$ is a morphism in $PGrp(\mathcal{E})$. Now we show that $(\hat{G}^1 \xrightarrow{\hat{m}_j} \hat{H}_j)_{j\in I}$ is a mono-source in $PGrp(\mathcal{E})$. Let $\hat{k} = (k_1, k_2, k_3)$ and $\hat{g} = (g_1, g_2, g_3)$ be morphisms in $PGrp(\mathcal{E})$ from $\hat{k}$ to $\hat{G}^1$ such that $\hat{m}_j \circ \hat{k} = \hat{m}_j \circ \hat{g}$, for all $j \in I$. So we have $U(\hat{m}_j \circ \hat{k}) = U(\hat{m}_j \circ \hat{g})$, for all $j \in I$. Therefore $m_j \circ k_2 = m_j \circ g_2$, for all $j \in I$, and since $(m_j)_{j\in I}$ is a mono source, we have $k_2 = g_2$. By Proposition 3.1, we get $\hat{k} = \hat{g}$.
Thus $(\hat{G}^1 \xrightarrow{\hat{m}_j} \hat{H}_j)_{j\in I}$ is a mono-source in $PGrp(\mathcal{E})$, and we have:

$$U(\hat{m}_j) \circ e = U_2(U_1(\hat{m}_j)) \circ e = U_2(m_j) \circ e = f_j$$

We conclude that every U-structure has a (generating, Mono-source)-factorization. ■

Theorem 4.2 *The forgetful functor $U : PGrp(\mathcal{E}) \longrightarrow \mathcal{E}$, creates isomorphisms.*

Proof: Let $\hat{H} = (H, P, R, s, r, *, E, i)$ be an object in $PGrp(\mathcal{E})$, and $f : X \to U(\hat{H}) = P$ be an $\mathcal{E}$-isomorphism. Define $\hat{X}_H = (H, X, R, s', r', *', E', i)$, where $s' = f^{-1}s$, $r' = f^{-2}r$, $*' = f^{-1} * f^2$, $\overline{E}' = f^{-1}\overline{E}$ and $E' = f^{-1}Ef$. It is easy to check that $\hat{X}_H$ is an object in $PGrp(\mathcal{E})$ and $\hat{f} = (id_H, f, id_R) : \hat{X}_H \to \hat{H}$ is an isomorphism in $PGrp(\mathcal{E})$. Also we have $U(\hat{f}) = f$, and since U is faithful, thus $\hat{f}$ is unique morphism in $PGrp(\mathcal{E})$ such that $U(\hat{f}) = f$. ■

Corollary 4.3 *The concrete category $(PGrp(\mathcal{E}), U)$ over $\mathcal{E}$ is essentially algebraic.*

Proof: By Theorem 4.1 and 4.2, we have U is essentially algebraic, so the concrete category $(PGrp(\mathcal{E}), U)$ over $\mathcal{E}$, is essentially algebraic. ■

Corollary 4.4

(i) The concrete category $(PGrp(\mathcal{E}), U)$ has equalizers.
(ii) The functor U detects colimits.
(iii) The functor U preserves and creates limits.
(iv) If $\mathcal{E}$ is complete, then $PGrp(\mathcal{E})$ is complete.
(v) If $\mathcal{E}$ has coproducts, then $PGrp(\mathcal{E})$ is cocomplete.
(vi) If $\mathcal{E}$ is wellpowered, then $PGrp(\mathcal{E})$ is wellpowered.

Proof: The proof is concluded by Corollary 23.10, Theorem 23.11, and Proposition 23.12 of [1]. ■

Remark 4.5 By Corollaries 4.3 and 4.4 we get that $PGrp$ is complete, cocomplete and wellpowered.

References

[1] J. Adamek, H. Herrlich and G.E. Strecker, *Abstract and Concrete Categories*, John Wiely and Sons, Inc, 1990.

[2] P. Bonansinga and P. Corsini, Sugli omomorphismi di semi-ipergruppi e di ipergruppi, *Bollettino U.M.I.*, Vol(6), 1-B (1982) 717-727.

[3] S.D. Comer, Polygroups derived from cogroups, *Journal of Algebra*, Vol.89, 2 (1984) 397-405.

[4] P.T. Johnstone, *Topos Theory*, Academic Press, 1977.

[5] S. Maclane and I. Moerdijk, *Sheaves in Geometry and Logic. A First Introduction to Topos Theory*, Springer-Verlag, 1992.

[6] F. Marty, Sur une generalization de la notion de groupe. *Actes du huitieme congres des mathematiciens scandinaves*, Stockolm, 1934, pp. 45-49.

[7] H. Schubert, *Categories*, Springer-Verlag. Berlin Heidelberg New-York 1972.

عنوان و چکیدهٔ فارسی مقالات

برچسب‌گذاری نوع α برای رده‌های خاصی از گرافهای درجهٔ دوّم

در این مقاله ابتدا نشان داده شده است که گراف درجه دوم $Q(S,4K)$ که متشکل از پنج دور به طول $4K$ است، به ازای هر عدد صحیح K دارای برچسب‌گذاری نوع α می‌باشد. سپس با استفاده از آن وجود برچسب‌گذاری α در رده‌های دیگری از گرافهای درجه دوم نیز نشان داده شده است.

توسیع توابع در زیرفضاهای مدل $H^2(\mathbb{R})$ به $\mathbb{C}$

در این مقاله زیرفضای مدل K_Θ از زیرفضای هاردی $H^2(\mathbb{R})$ مطالعه شده است. در اینجا نشان می‌دهیم که هر عضو $f \in K_\Theta$ که روی $\mathbb{R}$ تعریف شده است را می‌توان به تمام صفحهٔ مختلط تعمیم داد. این تعمیم به گونه‌ای است که در تمام نقاط خط حقیقی، به جز مجموعه‌ای با انگازه لبگ صفر، f دارای حد مماسی از هر دو طرف می‌باشد.
همچنین اگر Θ روی بازه‌ای از خط حقیقی تابعی تحلیلی باشد، f نیز روی این بازه تحلیلی خواهد بود. کاربرد این نتیجه در دو مورد حائز اهمیت می‌باشد. اوّلاً موقعی که تابع Θ تابع‌نمایی، $e^{i\sigma x}, \sigma$ باشد، این مورد در مقالهٔ مهم Buezling-Malliavinمطالعه شده است. ثانیاً موقعی که Θ یک تابع Blaschke مروموروف باشد، در این حالت f نیز تابعی مروموروف با همان قطبهای Θ می‌باشد. نتایج بیشتر در این مورد در [8] ذکر شده است.

برآوردیابی بردار میانگین توزیع نرمال چند بعدی تحت تابع زیان نرمال برگردانده شده‌ی تعمیم یافته

در این مقاله با در نظر گرفتن تابع زیان نرمال برگردانده شده‌ی تعمیم یافته، برآوردگرهای میانگین بردار نرمال p بعدی مورد بررسی و مقایسه قرارگرفته‌اند. روا بودن بردار $\overline{X} = (\overline{X}_1, ..., \overline{X}_p)'$ برای $p \leq 2$ اثبات گردیده و ناروا بودن این بردار برای $p \geq 3$ نشان داده شده است. همچنین کلاسی از برآوردگرها معرفی شده که از نظر مقدار تابع ریسک بر بردار $\overline{X}$ برتری دارند.

رسته اشیاء پلی گروهی

در این مقاله، مفهوم پلی گروه را در یک رسته دلخواه $\mathcal{E}$ تعمیم می‌دهیم، که همچنین تعمیمی از مفهوم شیء گروهی نیز می‌باشد. سپس رسته $PGrp(\mathcal{E})$ از اشیاء پلی گروهی در $\mathcal{E}$ را تعریف می‌کنیم و بعضی از خواص آن، همچون حد و هم-حد، را در آن تحقیق می‌کنیم. همچنین نشان می‌دهیم که $PGrp(\mathcal{E})$ یک رسته ملموس روی رسته $Mon(\mathcal{E})$ از اشیاء تکواره‌ای روی $\mathcal{E}$ می‌باشد و دارای اشیاء آزاد، هندسی و اساساً جبری است. سرانجام درباره حفظ و منعکس نمودن برونریختی و تکریختی توسط فانکتور فراموشی از $PGrp(\mathcal{E})$ به $Mon(\mathcal{E})$ بحث می‌کنیم.