ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ UNIVERSITY OF IOANNINA

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ DEPARTMENT OF MATHEMATICS

Number 9 **Volume June 2003**

A HYPOTHESIS TESTING

by K. X. Karakostas
Department of Mathematics, University of Ioannina, Creece

1. - Introduction

Let $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_m$ be two independent and identically distributed (i.i.d) random samples (r.s) from normal populations with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2, respectively. It is well known (see e.g Casella, Berger (1990)) that when the variances of the two populations are unknown, but equal, the test statistic for testing the hypothesis $H_o:\mu_X=\mu_Y$, is the following

$$t_{Old} = \frac{\bar{X}-\bar{Y}}{\sqrt{(n-1)S_X^2+(m-1)S_Y^2/(n+m-2)}\sqrt{(n+m)/(nm)}}, \tag{1}$$

where $\bar{X}$, $\bar{Y}$, S_X^2 and S_Y^2 are the sample means and variances of the two samples, respectively. Under the null hypothesis H_o the distribution of (1) is the central t with n+m-2 degrees of freedom. This test is both likelihood and a significance test. When the population variances are unknown and unequal the test statistic, for testing the same hypothesis, is,

$$t_{Welch} = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_x^2}{n}+\frac{S_y^2}{m}}} \tag{2}$$

The distribution of the above statistic, under the null hypothesis, is approximated by a central t distribution whose degrees of freedom are obtained approximately from various formulas. (See Cochran and Cox (1950), Satterthwaite (1946)).

Let's now think of a slightly different problem. More precisely we consider the previous hypothesis H_o, but this time we suppose that the variance σ_X^2 is known and σ_Y^2 is unknown. Problems of such a type can arise when we test the effectiveness of two methods and the precision of one of them is known, e.g. from previous experience.

2. -Testing the hypothesis H_o:$\mu_X=\mu_Y$ with σ^2_X known and σ^2_Y unknown

In the sequel we will present three methods for testing the previously mentioned hypothesis. The first is based on the significance test; the second one is the well-known likelihood ratio technique, while the third is rather different. Without loss of generality, in what follows, we can assume that $\sigma^2_X=1$.

2.1. -Significance test approach

It is well known that $\bar{X}\sim N(\mu_X,\ 1/n)$ and $\bar{Y}\sim N(\mu_Y,\ \sigma^2_Y/m)$. Hence, under the null hypothesis H_o:$\mu_X=\mu_Y$ $\bar{X}-\bar{Y}\sim N\left(0,\ \frac{1}{n}+\frac{\sigma^2_Y}{m}\right)$. Assuming that $1/n\approx 0$ and using the fact that $\frac{1}{\sigma^2_Y}\sum_{j=1}^{N}(Y_j-\bar{Y})^2\sim \aleph^2_{N-1}$ and $\bar{X}$, $\bar{Y}$ and S^2_Y are independent we get that the statistic $t_{New}=\frac{\sqrt{m}(\bar{X}-\bar{Y})}{S_Y}$ follows the central t-distribution with m-1 degrees of freedom. That is

$$t_{New}=\frac{\sqrt{m}(\bar{X}-\bar{Y})}{S_Y}\sim t_{m-1} \qquad (3)$$

The above test is an approximate test of significance, and as it will be come clear in the next section it cannot be derived using the likelihood approach.

2.2. - Likelihood ratio approach

Under H_o:$\mu_X=\mu_Y(=\mu)$ the likelihood function of the two samples is given by

$$L_o=\frac{1}{\left(\sqrt{\sigma^2_y}\right)^m\left(\sqrt{2\pi}\right)^{n+m}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}\exp\left\{-\frac{1}{2\sigma^2_y}\sum_{j=1}^{m}(y_j-\mu)^2\right\}. \qquad (4)$$

To find the maximum likelihood estimates (m.l.e), $\hat{\mu}$ and $\hat{\sigma}^2_{y,o}$, of μ and σ^2_Y, respectively, we take the natural logarithm of both sides and we get

$$\ln L_o=\ln\left(\frac{1}{\sqrt{2\pi}}\right)^{n+m}-\frac{m}{2}\ln\sigma^2_y-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2-\frac{1}{2\sigma^2_y}\sum_{j=1}^{m}(y_j-\mu)^2.$$

Hence

$$\frac{\partial \ln L_o}{\partial\mu}=\sum_{i=1}^{n}(x_i-\mu)+\frac{1}{\sigma^2_y}\sum_{j=1}^{m}(y_j-\mu) \quad\text{and}\quad \frac{\partial \ln L_o}{\partial\sigma^2_y}=-\frac{m}{2\sigma^2_y}+\frac{1}{2\sigma^4_y}\sum_{j=1}^{m}(y_j-\mu)^2.$$

By putting each of the above equations to zero we get the next system of equations

$$\left(m+n\sigma_y^2\right)\mu = m\overline{Y} + n\sigma_y^2\overline{X} \quad \text{και} \quad m\sigma_y^2 = \sum_{j=1}^{n}(y_j-\mu)^2 . \qquad (5)$$

If we solve this system (For the solution see Appendix), with respect of μ and σ_Y^2, we get the m.l. e., $\hat{\mu}$ and $\hat{\sigma}_{Y,o}^2$ of μ and σ_Y^2, respectively..

Under H_α:$\mu_X \neq \mu_Y$ the likelihood function of the two samples is given by

$$L_\alpha = \frac{1}{\left(\sqrt{2\pi}\right)^{n+m}\left(\sqrt{\sigma_Y^2}\right)^m} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_x)^2\right\} \exp\left\{-\frac{1}{2\sigma_y^2}\sum_{j=1}^{m}(y_j-\mu_y)^2\right\}. \qquad (6)$$

Working as above we obtain that the maximum likelihood estimates of μ_X, μ_Y and σ_Y^2 are given by the formulae

$$\hat{\mu}_X = \overline{X},\ \hat{\mu}_Y = \overline{Y} \ \text{και}\ \hat{\sigma}_{Y,\alpha}^2 = \frac{1}{m}\sum_{j=1}^{m}\left(Y_j - \overline{Y}\right)^2 . \qquad (7)$$

From (4), (6), (5) και (7) we have the maximum likelihood ratio test, λ, is given by

$$\lambda = \frac{\left(\sqrt{\hat{\sigma}_{y,o}^2}\right)^{-m} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2\right\} \exp\left\{-\frac{1}{2\hat{\sigma}_{y,o}^2}\sum_{j=1}^{m}(y_j-\hat{\mu})^2\right\}}{\left(\sqrt{\hat{\sigma}_{y,\alpha}^2}\right)^{-m} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\overline{X})^2\right\} \exp\left\{-\frac{m}{2}\right\}} \le k \ \text{ or}$$

$$\lambda = \left(\frac{\hat{\sigma}_{y,\alpha}^2}{\hat{\sigma}_{y,o}^2}\right)^{m/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2\right\} \exp\left\{-\frac{1}{2\hat{\sigma}_{y,o}^2}\sum_{j=1}^{m}(y_j-\hat{\mu})^2\right\} \exp\left\{\frac{1}{2}\sum_{i=1}^{n}(x_i-\overline{X})^2\right\} \le k^*$$

The term $-\frac{1}{2\hat{\sigma}_{y,o}^2}\sum_{j=1}^{m}(y_j-\hat{\mu})^2$ can be written as

$$-\frac{1}{2\hat{\sigma}_{y,o}^2}\sum_{j=1}^{m}(y_j-\hat{\mu})^2 + \frac{1}{2\hat{\sigma}_{y,o}^2}m\hat{\sigma}_{y,o}^2 - \frac{1}{2\hat{\sigma}_{y,o}^2}m\hat{\sigma}_{y,o}^2 = -\frac{1}{2\hat{\sigma}_{y,o}^2}\left\{\sum_{j=1}^{m}(y_j-\hat{\mu})^2 - m\hat{\sigma}_{y,o}^2\right\} - \frac{m}{2}.$$

From the second of the equations (5) we have that $\sum_{j=1}^{m}(y_j-\hat{\mu})^2 - m\hat{\sigma}_{y,o}^2 = 0$, and hence the maximum likelihood ratio takes the form

$$\lambda = \left(\frac{\hat{\sigma}_{y,\alpha}^2}{\hat{\sigma}_{y,o}^2}\right)^{m/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left[(x_i-\hat{\mu})^2 - (x_i-\overline{X})^2\right]\right\} \le k^{**}.$$

But $\sum_{i=1}^{n}\left[(x_i-\hat{\mu})^2 - (x_i-\overline{X})^2\right] = \sum_{i=1}^{n}\left(x_i-\hat{\mu}+x_i-\overline{X}\right)\left(x_i-\hat{\mu}-x_i+\overline{X}\right)$

$$= \sum_{i=1}^{n}\left(2x_i-\overline{X}-\hat{\mu}\right)\left(\overline{X}-\hat{\mu}\right)$$

$$=\left(\bar{X}-\hat{\mu}\right)\sum_{i=1}^{n}\left(2x_i-\bar{X}-\hat{\mu}\right)=\left(\bar{X}-\hat{\mu}\right)\left(2n\bar{X}-n\bar{X}-n\hat{\mu}\right)$$

$$=\left(\bar{X}-\hat{\mu}\right)\left(n\bar{X}-n\hat{\mu}\right)=n\left(\bar{X}-\hat{\mu}\right)^2.$$

Therefore $\lambda=\left(\frac{\hat{\sigma}^2_{y,\alpha}}{\hat{\sigma}^2_{y,o}}\right)^{m/2}\exp\left\{-\frac{n}{2}\left(\bar{X}-\hat{\mu}\right)^2\right\}\le k$.

Using the first of the equations (5) we obtain

$$\bar{X}-\hat{\mu}=\bar{X}-\frac{m\bar{Y}+n\bar{X}\hat{\sigma}^2_{y,o}}{m+n\hat{\sigma}^2_{y,o}}=\frac{m\bar{X}-m\bar{Y}}{m+n\hat{\sigma}^2_{y,o}}=\frac{m\left(\bar{X}-\bar{Y}\right)}{m+n\hat{\sigma}^2_{y,o}}$$

and finally

$$\lambda=\left(\frac{\hat{\sigma}^2_{y,\alpha}}{\hat{\sigma}^2_{y,o}}\right)^{m/2}\exp\left\{-\frac{n}{2}\frac{m^2\left(\bar{X}-\bar{Y}\right)^2}{\left(m+n\hat{\sigma}^2_{y,o}\right)^2}\right\}\le k^{***}. \qquad (8)$$

2.2.1- Case m/n≈0

The exact distribution of the quantity in (8) cannot be derived analytically. To simplify things we use the approximation m/n≈0. This can be the case e.g. when for the old method a wealth of observations is available whereas this is not the case for the new method, (e.g. because of cost).

If, in equations (5), we put m/n≈0, we get

$$\hat{\mu}=\bar{X}\ \text{και}\ \hat{\sigma}^2_{y,o}=\frac{1}{m}\sum_{j=1}^{m}(y_j-\bar{X})^2.$$

Also, in this case, (8) can be written as $\lambda=\left(\frac{\hat{\sigma}^2_{y,\alpha}}{\hat{\sigma}^2_{y,o}}\right)^{m/2}\le k^{***}$. Substituting $\hat{\sigma}^2_{y,\alpha}$ from (7) and $\hat{\sigma}^2_{y,o}$ from the previous relationship, we get

$$\lambda=\frac{\sum_{j=1}^{m}(y_j-\bar{Y})^2}{\sum_{j=1}^{m}(y_j-\bar{X})^2}\le k^{***}. \qquad (9)$$

But

$$\sum_{j=1}^{m}(y_j-\bar{Y})^2=\sum_{j=1}^{m}\left[(y_j-\bar{X})+(\bar{X}-\bar{Y})\right]^2=\sum_{j=1}^{m}\left[(y_j-\bar{X})+2(y_j-\bar{X})(\bar{X}-\bar{Y})+(\bar{X}-\bar{Y})^2\right]$$

$$=\sum_{j=1}^{m}(y_j-\bar{X})^2+2(\bar{X}-\bar{Y})\sum_{j=1}^{m}(y_j-\bar{X})+m(\bar{X}-\bar{Y})^2$$

$$= \sum_{j=1}^{m}(y_j - \overline{X})^2 - 2m(\overline{X} - \overline{Y})^2 + m(\overline{X} - \overline{Y})^2$$

$$= \sum_{j=1}^{m}(y_j - \overline{X})^2 - m(\overline{X} - \overline{Y})^2 .$$

Hence (9) can be written as

$$\lambda = \frac{\sum_{j=1}^{m}(y_j - \overline{Y})^2}{\sum_{j=1}^{m}(y_j - \overline{Y})^2 + m(\overline{X} - \overline{Y})^2} \leq k^{***} \quad \text{or} \quad \lambda = \frac{1}{1 + \frac{m(\overline{X} - \overline{Y})^2}{\sum_{j=1}^{m}(y_j - \overline{Y})^2}} \leq k^{***} \quad \text{or}$$

$$\lambda = \frac{1}{1 + \frac{t_{New}^2}{m-1}} \leq k^{**},$$

where

$$t_{New} = \frac{\sqrt{m(m-1)}(\overline{X} - \overline{Y})}{\sqrt{\sum_{j=1}^{m}(y_j - \overline{Y})^2}} = \frac{\sqrt{m}(\overline{X} - \overline{Y})}{S_Y} . \qquad (10)$$

Under the null hypothesis H_o:μ_X=μ_Y, the distribution of the above statistic, (10), is the central t-distribution with m-1 degrees of freedom.

Remark The assumption n/m≈0 it is not so helpful. This is so because, in this case, the likelihood statistic in (8) can be written in a form, which follows the chi-square distribution with one (1) degree of freedom.

2.3 - A different approach

In this section we will present a different approach for testing the hypothesis H_o:μ_X=μ_Y . This approach it is based on the "synthesis" of two other hypotheses and resembles, in some way, the union-intersection method of Roy (1957).

More precisely suppose that the hypothesis for testing is the H_o:μ_X=μ_Y while the alternative is the H_α:$\mu_X \neq \mu_Y$. The null hypothesis H_o can be separated into two partial hypotheses

$$H_o^{(1)} : \mu_X = \mu_o \text{ και } H_o^{(2)} : \mu_Y = \mu_o ,$$

where μ_o is a specified constant. Suppose now that for that value of μ_o, at least one of the hypothesis $H_o^{(1)} : \mu_X = \mu_o$ and $H_o^{(2)} : \mu_Y = \mu_o$ is being rejected. Then H_o:μ_X=μ_Y has to be rejected too. Alternatively, if both of the hypothesis $H_o^{(1)}: \mu_X = \mu_o$ and $H_o^{(2)}: \mu_Y = \mu_o$ can not be rejected, then we can not reject H_o:μ_X=μ_Y too. The value of μ_o is the estimated common value of μ_X and μ_Y, under H_o:μ_X=μ_Y, as it is given by the solution of the system (5). Denote by γ_z and γ_t the powers of testing the hypothesis $H_o^{(1)}: \mu_X = \mu_o$ and $H_o^{(2)}: \mu_Y = \mu_o$, respectively. Then the power of the different approach procedure, $\gamma_{DiffAppr}$ is given by $\gamma_{DiffAppr}=\gamma_z(1-\gamma_t)+(1-\gamma_z)\gamma_t+\gamma_z\gamma_t$, or $\gamma_{DiffAppr}=\gamma_z+\gamma_t-\gamma_z\gamma_t$. Obviously $\gamma_{DiffAppr}\geq\gamma_z$ and $\gamma_{DiffAppr}\geq\gamma_t$.

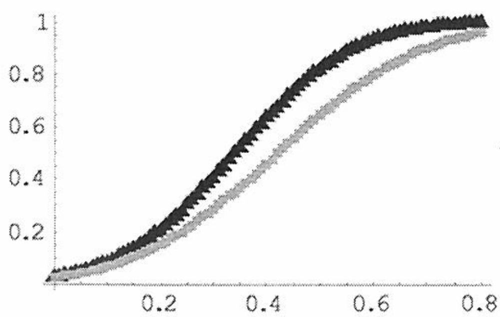
3. – Power comparisons

As it is well known, (see e.g. Cassella, Berger (1990)), the power, γ, of a test is defined as γ=P(reject H_o $|H_\alpha$ is true).More ever, if the test statistic follows, under the null hypothesis, a t distribution with v degrees of freedom then, under the alternative hypothesis H_α, the distribution of the test statistic follows, in general, a non-central t-distribution with v degrees of freedom and non-centrality parameter δ. The analytical form of such a distribution is

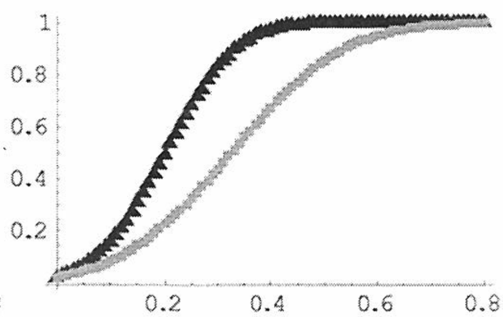
$$f(t;v,\delta) = \frac{1}{2^{(v+1)/2}\Gamma(\frac{v}{2})\sqrt{\pi v}} \int_0^\infty x^{(v-1)/2} \exp\left\{-\tfrac{1}{2}\left[x + \left(t\sqrt{\tfrac{x}{v}} - \delta\right)^2\right]\right\} dx, \; t\in R \qquad (11)$$

Significance test approach.

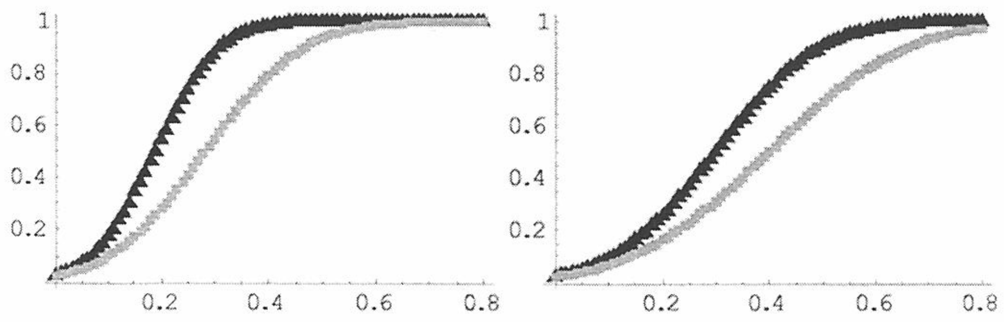
A) Variances are equal. Suppose that the variances σ_X^2 and σ_Y^2 are equal, say to σ^2. In this case we have to compare the powers of the test statistics in (1) and (3). The power, γ_{Old}, of this test statistic in (1), is γ_{Old}= P ($|t_{Old}|>t_{\alpha/2;n+m-2}|H_\alpha$ is true), while the power, γ_{New} , for the test statistic in (3), is γ_{New}= P($|t_{New}|>t_{\alpha/2;m-1}|H_\alpha$ is true). The non-centrality parameters are, $\delta_{Old}=|\mu_X-\mu_Y|\frac{1}{\sigma}\left(\frac{nm}{n+m}\right)^{1/2}$ and $\delta_{New}=|\mu_X-\mu_Y|(m/\sigma)^{1/2}$, respectively. No theoretical results can be obtained to see if γ_{Old} is greater or smaller to γ_{New}. However numerical computations have shown that γ_{New} is greater to γ_{Old}. (See graph below)



a



b



c d

Graph1. Power comparisons for the test statistic t_{old} with t_{New}, when the variances are equal for α=0.05. a) m=35, n=60; b) m=97, n=60; c) m=110, n=87; d) m=45, n=55. (The black curve is the power of the new test while the gray is for the old one).

If the sample sizes are equal, i.e m=n (=N), then we can show, theoretically, that γ_{New} is greater to γ_{Old}. To see that note that in this case the power, γ_{Old}, is $\gamma_{Old}= P(|t_{Old}|>t_{\alpha/2;2(N-1)}|H_\alpha$ is true), while the power, γ_{New} is $\gamma_{New}= P(|t_{New}|>t_{\alpha/2;N-1}|H_\alpha$ is true). Also, the non centrality parameters are, $\delta_{Old}=|\mu_X-\mu_Y|(N/2\sigma)^{1/2}$ and $\delta_{New}=|\mu_X-\mu_Y|(N/\sigma)^{1/2}$, respectively.

Proposition In the case n=m (=N), if the variances σ_X^2 and σ_Y^2 are equal, say to σ^2, and 1/N can be taken, approximately, to be equal to zero, then

$$\gamma_{New}\geq\gamma_{Old}.$$

Proof. Obviously

$\gamma_{Old}= P(|t_{Old}|>t_{\alpha/2;2(N-1)}|H_\alpha$ is true)=P($t_{Old}> t_{\alpha/2;2(N-1)}|H_\alpha$ is true)+ P($t_{Old}<-t_{\alpha/2;2(N-1)}|H_\alpha$ is true).

Hence, using (11), we can write

$$\gamma_{Old}=\frac{1}{2^N\Gamma(N-1)\sqrt{2(N-1)\pi}}\left[\int_0^\infty x^{N-2}\left(\int_{t_{\alpha/2;2(N-1)}}^\infty \exp\left\{-\frac{1}{2}\left[x+\left(t\sqrt{\frac{x}{2(N-1)}}-\delta_{Old}\right)^2\right]\right\}dt\right)dx\right.$$

$$\left.+\int_0^\infty x^{N-2}\left(\int_{-\infty}^{-t_{\alpha/2;2(N-1)}} \exp\left\{-\frac{1}{2}\left[x+\left(t\sqrt{\frac{x}{2(N-1)}}-\delta_{Old}\right)^2\right]\right\}dt\right)dx\right]$$

and similarly

$$\gamma_{New}=\frac{1}{2^{N/2}\Gamma((N-1)/2)\sqrt{(N-1)\pi}}\left[\int_0^\infty x^{(N-2)/2}\left(\int_{t_{\alpha/2;N-1}}^\infty \exp\left\{-\frac{1}{2}\left[x+\left(t\sqrt{\frac{x}{N-1}}-\delta_{New}\right)^2\right]\right\}dt\right)dx\right.$$

$$\left.+\int_0^\infty x^{(N-2)/2}\left(\int_{-\infty}^{-t_{\alpha/2;N-1}} \exp\left\{-\frac{1}{2}\left[x+\left(t\sqrt{\frac{x}{N-1}}-\delta_{New}\right)^2\right]\right\}dt\right)dx\right]$$

Now, since $t_{\alpha/2;N-1}\approx t_{\alpha/2;2(N-1)}$, especially in the case where $1/N\approx0$, we get that $\gamma_{New}\geq\gamma_{Old}$ if

$$\frac{1}{2^{N/2}\Gamma((N-1)/2)\sqrt{(N-1)\pi}}\exp\left\{-\tfrac{1}{2}\left[x+\left(t\sqrt{\tfrac{x}{N-1}}-\delta_{New}\right)^2\right]\right\}\geq$$

$$\frac{1}{2^{N}\Gamma(N-1)\sqrt{2(N-1)\pi}}\exp\left\{-\tfrac{1}{2}\left[x+\left(t\sqrt{\tfrac{x}{2(N-1)}}-\delta_{Old}\right)^2\right]\right\}$$

or equivalently

$$-\frac{1}{2}\left[x+\left(t\sqrt{\frac{x}{N-1}}-\delta_{New}\right)^2\right]\geq-\frac{1}{2}\left[x+\left(t\sqrt{\frac{x}{2(N-1)}}-\delta_{Old}\right)^2\right]+c$$

where $c=\ln\left(\Gamma(\tfrac{N-1}{2})/2^{\frac{5}{2}}\Gamma(N-1)\right)$. Note that c is a negative number. By putting, in the last inequality, $\delta_{Old}=\delta_{New}\frac{\sqrt{2}}{2}$ and after some algebraic manipulation we arrive at

$$\delta_{New}^2-2t\sqrt{\frac{x}{N-1}}\delta_{New}+t^2\frac{x}{N-1}-4c\geq 0.$$

But this inequality is always true, and hence we have completed the proof of the proposition.

B) Variances are unequal. If the variances are unequal, that is $\sigma_X^2\neq\sigma_Y^2$, then the power of the test statistic in (3) must be compared with that of test statistic in (2). In this case, see Brentari, Carpita and Dancelli (2001), $\gamma_{Welch}=P(|t_{Welch}|>t_{\alpha/2;m-1}|H_\alpha\text{ is true})$ and the non-centrality parameter is $\delta_{Welch}=|\mu_X-\mu_Y|(m/\sigma_Y)^{1/2}$. Hence, in this case, the powers of the two test statistics coincide.

Likelihood ratio approach

Case $m/n\approx 0$. If the variances are equal, then $\gamma_{New}=P(|t_{New}|>t_{\alpha/2;m-1}|H_\alpha$ is true), where now t_{New} is the statistic in (10) and $\gamma_{Old}=P(|t_{Old}|>t_{\alpha/2};n|H_\alpha$ is true), where t_{Old} is the statistic in (1), written appropriately. Then, under the alternative hypothesis, each of these statistics follow the non-central t distribution with non-centrality parameters $\delta_{New}=|\mu_X-\mu_Y|/\sigma m^{1/2}$ and $\delta_{Old}=|\mu_X-\mu_Y|/\sigma m^{1/2}$, respectively. In this case $\delta_{New}=\delta_{Old}$ and hence $\gamma_{New}\leq\gamma_{Old}$. The last inequality comes from the Pearson-Hartley (1951) table for the power of the $F_{1,v}$ distribution.

If the variances are unequal, then $\gamma_{Welch}=P(|t_{Welch}|>t_{\alpha/2;m-1}|H_\alpha\text{ is true})$, where t_{Welch} is the statistic in (2). Following the work of Brentari, Carpita and Dancelli (2001) we can show that the two tests coincide. Hence $\gamma_{New}=\gamma_{Welch}$.

Different approach. In this case, if the variances are equal, ($\sigma_X=\sigma_Y=1$) then $\gamma_{DiffAppr}$ is slightly les than γ_{Old}. For example if n=46, m=37, α=0.05 $\mu_X=5$ and $\mu_Y=5.6$ we get that

$\gamma_{DiffAppr}$=0.72 and γ_{Old}=0.76. In the case of unequal variances $\gamma_{DiffAppr}$ is much better than γ_{Welch}. For example for n=46, m=37, α=0.05, μ_X=4, μ_Y=8, σ_X=1 and σ_Y=11 we obtain $\gamma_{DiffAppr}$=0.77 and γ_{Welch}=0.11. In both cases (variances equal or unequal) the value of μ_o is obtained from the formula (12)

4. – Numerical comparisons

In following three examples the mean of one population is double the mean of the other, so we expect the test to reject the hypothesis of equality of the two means.

Example 1. (*Case m=n*)

Using a computer we created two independent random samples from the normal distributions N(4, 49) and N(7, 100), each of size 45. The two samples gave $\overline{X} = 4{,}034$, $S_x = 6{,}75$, $\overline{Y} = 7{,}19$ and $S_Y = 8{,}93$. In order our first sample to be from a normal distribution with variance equal to one (1) we divide each observation, for both samples, by seven (7) to get N(0.57, 1) and N(1, 2.04). The Levene's test does not reject the hypothesis of equal variances, so we can make use of the statistic in (1). This statistic cannot reject the null hypothesis, (p=0,062). However, the t test given by relation (3) rejects the null hypothesis (p=0,01).

Example 2 (*Case m/n≈0*)

We created two independent random samples. The first, of size n=270, from the normal N(4, 1) and the second, of size m=20, from the normal distribution N(8, 121). The two samples gave $\overline{X} = 3{,}98$, $S_x = 1{,}03$, $\overline{Y} = 8{,}59$ και $S_Y = 9{,}56$. The test, for the equality of variances (Leven's test), rejects the hypothesis that the variances are equal, (p=<0,001). For that reason we use the Welch's (1938) test, (see equation (2)) for testing the hypothesis $H_o:\mu_X=\mu_Y$. The Welch's statistic gave p=0,0439. The corresponding value for the test statistic in (10) is p=0,0152.

Example 3 (*Case § 2.3*)

We created two independent random samples. The first, of size n=35, from the normal N(4, 1) and the second, of size m=42, from the normal distribution N(8, 225). The two samples gave $\overline{X} = 4{,}23$, $S_x = 0{,}97$, $\overline{Y} = 7{,}72$ και $S_Y = 14{,}26$. The solution of the system (5) gave $\hat{\mu}$=4,62. From this we conclude that for α=0,05 the null hypothesis is rejected since p_z=0.0214 and p_t=0.1661. The test statistic t_{Welch} in (2) cannot reject the null hypothesis, (p=0,1212).

5. Conclusions

From the previous analysis we can state the following:

a) The significance test t_{New}, in (3) is more powerful than the t_{Old} in (1), in the case of equal variances, and coincides with the t_{Welch} in (2), in the case of unequal variances.

b) The likelihood ratio test in (10), in the case of equal variances, is less powerful than the test in (1), while in the case of unequal variances, coincides with the test in (2).

c) The power for the different approach, presented in § 2.3, is slightly less powerful than the test in (1), in the case of equal variances, but it is more powerful than the test in (2), if the variances are unequal.

Appendix - Solution of the system (5)

From the first of equations (5) we get

$$\hat{\mu}=\frac{m\bar{Y}+n\bar{X}\hat{\sigma}^2_{y,o}}{m+n\hat{\sigma}^2_{y,o}}. \tag{12}$$

Making use of this relationship we can write

$$\sum_{j=1}^{m}(y_j-\hat{\mu})^2=\sum_{j=1}^{m}\left(y_j-\frac{m\bar{Y}+n\bar{X}\hat{\sigma}^2_{y,o}}{m+n\hat{\sigma}^2_{y,o}}\right)^2=\sum_{j=1}^{m}\left(\frac{my_j+n\hat{\sigma}^2_{y,o}y_j-m\bar{Y}-n\bar{X}\hat{\sigma}^2_{y,o}}{m+n\hat{\sigma}^2_{y,o}}\right)^2$$

$$=\frac{1}{\left(m+n\hat{\sigma}^2_{y,o}\right)^2}\sum_{j=1}^{m}\left\{m(y_j-\bar{Y})+n(y_j-\bar{X})\hat{\sigma}^2_{y,o}\right\}^2$$

$$=\frac{1}{\left(m+n\hat{\sigma}^2_{y,o}\right)^2}\sum_{j=1}^{m}\left\{m^2(y_j-\bar{Y})^2+n^2(y_j-\bar{X})^2\left(\hat{\sigma}^2_{y,o}\right)^2+2mn\hat{\sigma}^2_{y,o}(y_j-\bar{Y})(y_j-\bar{X})\right\}$$

$$=\frac{1}{\left(m+n\hat{\sigma}^2_{y,o}\right)^2}\left\{m^2\sum_{j=1}^{m}(y_j-\bar{Y})^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2\sum_{j=1}^{m}(y_j-\bar{X})^2+2mn\hat{\sigma}^2_{y,o}\sum_{j=1}^{m}(y_j-\bar{Y})(y_j-\bar{X})\right\}$$

$$=\frac{1}{\left(m+n\hat{\sigma}^2_{y,o}\right)^2}\left\{m^2\sum_{j=1}^{m}(y_j-\bar{Y})^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2\sum_{j=1}^{m}(y_j-\bar{X})^2+2mn\hat{\sigma}^2_{y,o}\sum_{j=1}^{m}(y_j-\bar{Y})y_j\right\}$$

$$=\frac{1}{\left(m+n\hat{\sigma}^2_{y,o}\right)^2}\left\{m^2\sum_{j=1}^{m}(y_j-\bar{Y})^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2\sum_{j=1}^{m}(y_j-\bar{X})^2+2mn\hat{\sigma}^2_{y,o}\sum_{j=1}^{m}(y_j-\bar{Y})^2\right\}$$

From the second of the equations (5) we get $m\hat{\sigma}^2_{y,o}=\sum_{j=1}^{m}(y_j-\hat{\mu})^2$ or, by replacing this term on the right hand side of the equal sign with the last result we get

$$m\hat{\sigma}^2_{y,o}\left(m+n\hat{\sigma}^2_{y,o}\right)^2=$$

$$=\left\{m^2\sum_{j=1}^{m}(y_j-\bar{Y})^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2\sum_{j=1}^{m}(y_j-\bar{X})^2+2mn\hat{\sigma}^2_{y,o}\sum_{j=1}^{m}(y_j-\bar{Y})^2\right\}$$

or $m\hat{\sigma}^2_{y,o}\left[m^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2+2mn\hat{\sigma}^2_{y,o}\right]=$

$$=\left\{m^2\sum_{j=1}^{m}(y_j-\overline{Y})^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2\sum_{j=1}^{m}(y_j-\overline{X})^2+2mn\hat{\sigma}^2_{y,o}\sum_{j=1}^{m}(y_j-\overline{Y})^2\right\}$$

or $m^3\hat{\sigma}^2_{y,o}+mn^2\left(\hat{\sigma}^2_{y,o}\right)^3+2m^2n\left(\hat{\sigma}^2_{y,o}\right)^2=$

$$=\left\{m^2\sum_{j=1}^{m}(y_j-\overline{Y})^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2\sum_{j=1}^{m}(y_j-\overline{X})^2+2mn\hat{\sigma}^2_{y,o}\sum_{j=1}^{m}(y_j-\overline{Y})^2\right\}$$

and finally

$$mn^2\left(\hat{\sigma}^2_{y,o}\right)^3+\left\{2m^2n-n^2\sum_{j=1}^{m}(y_j-\overline{X})^2\right\}\left(\hat{\sigma}^2_{y,o}\right)^2+\left\{m^3-2mn\sum_{j=1}^{m}(y_j-\overline{Y})^2\right\}\hat{\sigma}^2_{y,o}-$$

$$-m^2\sum_{j=1}^{m}(y_j-\overline{Y})^2=0.$$

Putting $A=mn^2$, $B=2m^2n-n^2\sum_{j=1}^{m}(y_j-\overline{X})^2$, $\Gamma=m^3-2mn\sum_{j=1}^{m}(y_j-\overline{Y})^2$ and

$\Delta=-m^2\sum_{j=1}^{m}(y_j-\overline{Y})^2$ we obtain

$$A(\hat{\sigma}^2_{y,o})^3+B(\hat{\sigma}^2_{y,o})^2+\Gamma\hat{\sigma}^2_{y,o}+\Delta=0 \qquad (13)$$

References

Brentari, E., Carpita, M. and Dancelli, L (2001) *An Improvement in the Welch Approximate t Test*, Metron LIX, 99-112.

Casella, G, Berger, R. L. (1990) *Statistical Inference*, Duxbury Press, Belmont, California

Cochran W. G, Cox, G. M (1950) *Experimental Designs*, New York, John Wiley & Sons, Inc.

Pearson, E. S., Hartley, H. O. (1951) *Charts of the Power Function of the Analysis of Variance Tests, Derived from the non-central F Distribution*, Biometrika 38, 112-130

Roy, S. N.(1957) *Some Aspects of Multivariate Analysis*, John Wiley

Satterthwaite, F. E. (1946) *An Approximate Distribution of Estimates of Variance Components*, Biometrics Bulletin (now called Biometrics) 12, 110-114

Welch, B. L (1938) *The Significance of the Difference Between Two Means when the Population Variances are Unequal*, Biometrika 29, 350-362.

Stochastic ordering properties of the exponential, upper truncated and lower truncated families of distributions

S. Papachristos, K. Skouri

Dept. of Mathematics, Statistics and O.R. Section,

University of Ioannina,

45110 Ioannina, Greece

Abstract

Using Bayes approach we obtain the predictive density functions for the exponential, range of type A and range of type B families of distributions, conditioned on a sufficient statistics. We then examine stochastic ordering properties useful to inventory control theory.

Key words: Sufficient statistic, predictive densities, stochastic ordering, Bayesian inventory control.

*Corresponding author. Email: spapachr@cc.uoi.gr

1. INTRODUCTION

The concept of stochastic ordering on probability densities is rather old. During the years, various types of stochastic ordering have been proposed. Each one of them has served different needs and scientific settings. The definition given here is based on the cumulative distribution function.

Definition. *Let φ and ψ be two density functions and $\Phi(x)$, $\Psi(x)$ their corresponding cumulative distribution functions (c.d.f). If $\Phi(x) \geq \Psi(x)$ for all $x \in R$, we say that the density φ is stochastically smaller than ψ and we write it as $\varphi \subset \psi$.*

The first application of this type of stochastic ordering seems to appear in the paper of Mann and Withney (1947), who tried to characterize the alternatives when testing for equality of two distributions.

Stochastic ordering has proved an important research tool in many scientific areas: Queuing theory (Stoyan, 1983), reliability theory (Boland et al. 1992), stochastic scheduling (Ross, 1983), economics and operations research (Savage, 1972, Masse, 1962). A most extensive treatment of this concept and its applications is given in the book by Shaked and Shanthikumar (1994).

In the area of inventory control the concept of stochastic ordering enters through the demand function for the product in question. Although not absolutely clear, it seems that it has appeared for the first time, in the pioneering paper by Scarf (1959). Scarf, using adaptive dynamic programming, managed to compare optimal ordering levels under different statistical information settings. This work was followed by the work of Karlin (1960). Karlin used the stochastic ordering in a more systematic way. In the first part of his paper he produced some very interesting and fairly general results. In the second part he specialized on the densities describing the system's demand by considering two classes of probability densities, namely the exponential and range family of type A. For each of these families he studied the stochastic ordering properties of their predictive densities and then he applied the obtained results to produce some very interesting ordering relations for the optimal ordering levels of the inventory system under study. Following the work of Scarf and Karlin, Iglehart (1964) extended their results, and produced new ones. His contribution was important in two directions: (i) Establishing stochastic ordering properties for the predictive density functions and (ii) transferring this ordering to monotonicity properties on the optimal ordering levels. Closing this important paper, Iglehart proposed the study of the range family of type B, another family of truncated densities, which could be used to describe the demand fluctuations for some inventory systems. Iglehart suspected that this family might also have nice stochastic ordering properties. Work in the field continued with Brown and Rogers (1972), Papachristos (1977a, 1977b), Azoury (1985).

The common characteristic of the research mentioned in inventory control theory, was that the densities used to describe demand fluctuations accepted a sufficient statistic for the unknown parameter ω. This statistics was based on all past demand records supposed available. However there are cases, where not all past demand records are completely available. Such are the cases with censored, or missing data. In such cases we need to consider families, which can accept, sufficient statistics based on censored or missing data. In this direction we mention the most important contributions, which are the work of Braden and Freimer (1991) and Nahmias (1994).

In this paper we study the stochastic ordering properties of the predictive densities of three families of distributions i.e. the exponential, the range family of type A, and the range of type B. Section 2 is devoted to the exponential family. We review the existing results and give a new one. We then proceed with section 3 where we deal with the range family of type A. We also review the existing results, and we present a new one. Section 4 is devoted to the study the range family of type B, proposed by Iglehart (1964). For this family we establish stochastic ordering properties, for its predictive density function. This ordering is related to the information available for the parameter ω. Finally in section 5 we give some concluding remarks and we are making proposals, for possible applications of the produced results and further research in the area of inventory control.

2. THE EXPONENTIAL FAMILY

Let ξ be a continuous random variable with density belonging to the exponential family

$$\varphi(\xi|\,\omega)=\beta(\omega)\exp(-\omega\xi)r(\xi) \quad \xi\geq 0 \tag{1}$$

where

$$(\beta(\omega))^{-1}=\int_0^\infty \exp(-\omega\xi)r(\xi)d\xi \tag{2}$$

ω is unknown parameter taking values at the interval $I_1=[0,\infty)$ and $r(\xi)$ is a strictly positive known function such that the integral in (2) exists for every $\omega \in I_1$. It 's obvious that $\beta(\omega)$ is a strictly increasing function of ω. From this family we can obtain the distributions, Gamma in the continuous case, Poisson and Negative Binomial in the discrete case.

Let $\xi_1, \xi_2,\ldots, \xi_n$ be n independent observations of the r.v. ξ and $Q=\sum_{i=1}^{n}\xi_i$ and $s=\frac{Q}{n}$. Then

$$\varphi(\xi_1, \xi_2,\ldots, \xi_n \mid \omega)=\{\beta(\omega)\}^n \exp(-\omega ns)\prod_{i=1}^{n} r(\xi_i) \tag{3}$$

From Fisher-Neyman factorization theorem (Raiffa and Schlaifer, 1961) we have that the pair (Q, n) or equivalently (s, n) is a sufficient statistic for the parameter ω. If $f(\omega)$ is the prior probability density (p.d.) of ω then using the Bayes theorem we find the posterior density of ω given Q, n as

$$f(\omega| Q, n)=\frac{\{\beta(\omega)\}^n \exp(-\omega Q)f(\omega)d\omega}{\int_0^\infty\{\beta(\omega)\}^n \exp(-\omega Q)f(\omega)d\omega} \tag{4}$$

The predictive probability density function (p.p.d.f.) of ξ is now given by

$$\varphi(\xi| Q, n)=\int_0^\infty \varphi(\xi|\omega)f(\omega|Q,n)d\omega = r(\xi)\frac{\int_0^\infty\{\beta(\omega)\}^{n+1} \exp(-\omega\xi)\exp(-\omega Q)f(\omega)d\omega}{\int_0^\infty\{\beta(\omega)\}^n \exp(-\omega Q)f(\omega)d\omega}$$

$$=r(\xi)\frac{A(n+1, Q+\xi)}{A(n, Q)} \tag{5}$$

where $A(n, y)=\int_0^\infty\{\beta(\omega)\}^n \exp(-\omega y)f(\omega)d\omega$

If we replace Q by ns on the right hand side of (4) and (5) we obtain a similar expression for $f(\omega| s,n)$ and $\varphi(\xi| s, n)$.

The p.d $\varphi(\xi| Q, n)$ possesses some very interesting stochastic order properties which are examined in theorem 2.1, 2.2 and 2.3. Karlin (1960), taking (Q, n) as a sufficient statistic for the parameter ω, stated and proved theorem 2.1a and 2.2. His proofs are based in a crucial way on the concept of "variation diminishing transformations". Here we give alternative proofs using the property of monotone likelihood ratio of the exponential family. Iglehart (1964), taking (s, n) as a sufficient statistic for the parameter ω, stated and proved theorem 2.1b. Here we give alternative proof to this result. Iglehart also stated a theorem parallel to 2.2 taking the pair (s, n) as a sufficient statistic for ω. His result is not valid, as it has shown in (Papachristos, 1977a). So taking (Q, n) as a sufficient statistic for ω we prove theorem 2.2. Theorem 2.3 is new and gives a (stochastic) upper bound for $\varphi(\xi| Q, n)$.

Theorem 2.1. *For any $n\geq 1$ and Q, Q', s, s' with $Q<Q'$ and $s<s'$ we have*

a) $\varphi(\xi/Q, n)\subset\varphi(\xi/Q', n)$

b) $\varphi(\xi/s, n) \subset\varphi(\xi/s', n)$

PROOF. If we set

$$\Delta(\xi, Q, Q')= \varphi(\xi| Q, n)- \varphi(\xi| Q', n)$$

then we must prove that

$$\int_0^x \Delta(\xi, Q, Q')d\xi \geq 0 \quad \forall x \in R.$$

Since $\varphi(\xi| Q, n)$ and $\varphi(\xi| Q', n)$ are p.d. we have

$$\int_0^\infty \Delta(\xi, Q, Q')d\xi = 0$$

This proves that the function $\Delta(\xi, Q, Q')$ changes sing at least once (from positive to negative or reversely) as ξ traverses the interval $[0, \infty)$. The function $\Delta(\xi, Q, Q')$ can be written

$$\Delta(\xi,Q, Q')= \int_0^\infty \varphi(\xi|\omega)\frac{\{\beta(\omega)\}^n \exp(-\omega Q')}{A(n,Q)}t_1(\omega)f(\omega)d\omega \tag{6}$$

with $t_1(\omega)=\exp\{-\omega(Q-Q')\}-\dfrac{A(n, Q)}{A(n, Q')}$

Since $Q<Q'$ it follows that $A(n, Q)>A(n, Q')$ and it is easy to see that if ω traverses the interval $[0, \infty)$ the function $t_1(\omega)$ changes sign exactly once from negative to positive values.

Now based on the two lemmas, given in the appendix we conclude that $\Delta(\xi, Q, Q')$ changes sign at most once from positive to negative values as ξ traverses the interval $[0, \infty)$. So $\Delta(\xi, Q, Q')$

changes sing exactly once from positive to negative values and this proves the theorem. Replacing (Q, Q΄) with (ns, ns΄) the part b of theorem follows

Theorem 2.2. *For any n≥1 and Q we have,*

$$\varphi(\xi/Q,\ n+1)\subset\varphi(\xi/Q,\ n)$$

PROOF. The proof is similar to that of the previous theorem, we only note the following

If

$$\Delta(\xi,\ n)=\ \varphi(\xi|\ Q,\ n+1)-\ \varphi(\xi|\ Q,\ n)$$

Then

$$\Delta(\xi,\ n)=\int_0^\infty \varphi(\xi|\omega)\frac{(\beta(\omega))^n\exp(-\omega Q)}{A(n+1,Q)}t_2(\omega)f(\omega)d\omega\,, \tag{7}$$

with $t_2(\omega)=\beta(\omega)-\dfrac{A(n+1,Q)}{A(n,Q)}$. Since $\beta(\omega)$ is a strictly increasing function we have

$$\beta(0)A(n,\ Q)<A(n+1,\ Q).$$

But $\beta(\omega)\to\infty$ as $\omega\to\infty$, and so the function $t_2(\omega)$ changes sign once from negative to positive values as ω traverses the interval $[0,\ \infty)$.

Theorem 2.3. *For any n≥1 and Q, s we have*

a) $\varphi(\xi/Q,\ n)\subset\varphi(\xi/0)$

b) $\varphi(\xi/s,\ n)\subset\varphi(\xi/0)$

where $\varphi(\xi|\ 0)$ is given by (1) for $\omega=0$

PROOF. Let us set

$$\Delta(\xi)=\ \varphi(\xi|\ Q,\ n\ )-\varphi(\xi|\ 0)=r(\xi)\int_0^\infty(\beta(\omega)\exp(-\omega\xi)-\beta(0))f(\omega|Q,n)d\omega$$

Then we must prove that

$$\int_0^x\Delta(\xi)d\xi\ \geq 0\quad\forall\ x\in R.$$

But $\Delta(0)=r(0)\int_0^{\infty}(\beta(\omega)-\beta(0))f(\omega|Q,n)d\omega>0$ because $r(0)>0$ and $\beta(\omega)>\beta(0)$

Since

$$\int_0^{\infty}\Delta(\xi)d\xi=0$$

It follows that the function $\Delta(\xi)$ changes sign at least once as ξ traverses the interval $[0, \infty)$

Moreover the function $\beta(\omega)\exp(-\omega\xi)-\beta(0)$ is strictly decreasing with respect to ξ and goes to $-\beta(0)$ as $\xi\to\infty$. So $\Delta(\xi)$ changes sign exactly once from positive to negative values and this proves the theorem.

3. THE RANGE FAMILY OF TYPE A

Let ξ be a continuous random variable with density

$$g(\xi|\omega)=\beta(\omega)q(\xi)\psi(\omega, \xi),\ \xi\geq 0 \tag{8}$$

where

$$\psi(\omega, \xi)=\begin{cases}1 \text{ if } \xi\leq\omega \\ 0 \text{ if } \omega<\xi\end{cases},$$

$$(\beta(\omega))^{-1}=\int_0^{\omega}q(\xi)d\xi, \tag{9}$$

ω is an unknown parameter taking values in some interval $I_2=[0, \infty)$ and $q(\xi)$ is a strictly positive known function, such that the integral in (9) exists for every $\omega\in I_2$. We call this family the "the range family of type A". Defining suitably the function $q(\xi)$ we can obtain from this family some known upper truncated distributions.

Let $\xi_1, \xi_2,\ldots, \xi_n$ be the values of n independent observations of the random variable ξ whose density is $g(\xi|\omega)$. The likelihood function of $\xi_1, \xi_2,\ldots, \xi_n$ is

$$g(\xi_1, \xi_2,\ldots, \xi_n|\omega)=(\beta(\omega))^n\prod_{i=1}^{n}q(\xi_i)\psi(\omega, \upsilon)$$

where

$$\upsilon = \max_{1 \le i \le n} \xi_i$$

This expression shows that the pair (υ, n) is a sufficient statistic for the parameter ω and also that υ is the maximum likelihood estimator (m.l.e.) based on $\xi_1, \xi_2, \ldots, \xi_n$ (Raiffa and Schlaifer, 1961).

If $f(\omega)$ is the prior distribution for ω then using the Bayes theorem we find, the posterior density of ω

$$f(\omega \mid \upsilon, n) = \frac{\{\beta(\omega)\}^n f(\omega)\psi(\omega, \upsilon)}{\int_0^\infty \{\beta(\omega)\}^n f(\omega)\psi(\omega, \upsilon)d\omega}$$

$$= \begin{cases} \dfrac{\{\beta(\omega)\}^n f(\omega)}{B(\upsilon, n)} & \text{if } \omega \ge \upsilon \\ 0 & \text{if } \omega < \upsilon \end{cases}$$

where

$$B(\upsilon, n) = \int_\upsilon^\infty \{\beta(\omega)\}^n f(\omega)d\omega$$

The p.p.d.f. of ξ is now given by

$$g(\xi \mid \upsilon, n) = q(\xi)\frac{B(\max(\xi, \upsilon), n+1)}{B(\upsilon, n)} \quad \upsilon < \infty$$

This family was first studied by Karlin (1960) and later by Iglehart (1964) who stated and proved theorems 3.1 and 3.2 which follows. The proofs given by these authors are based on the concept of "variation diminishing transformation". Here we give alternatives proofs, which do not require this concept. Theorem 3.3 is new and gives an (stochastic) upper bound of p.p.d.f of the family (8).

Theorem 3.1. *For any $n \ge 1$ and υ, υ' with $\upsilon < \upsilon' < \infty$ we have*

$$g(\xi / \upsilon, n) \subset g(\xi / \upsilon', n).$$

PROOF. We suppose that $0 < \upsilon < \infty$, and let us set

$$\Delta(\xi, \upsilon, \upsilon') = g(\xi \mid \upsilon, n) - g(\xi \mid \upsilon', n).$$

Then we must prove that

$$\int_0^x \Delta(\xi, \upsilon, \upsilon')d\xi \geq 0 \ \forall\ x .$$

Obviously

$$\int_0^\infty \Delta(\xi, \upsilon, \upsilon')d\xi = 0$$

This last relation shows that $\Delta(\xi, \upsilon, \upsilon')$ change sing at least once as ξ traverses the interval $[0, \infty)$. The result will follow if we prove that the function $\Delta(\xi, \upsilon, \upsilon')$ changes sign exactly once from positive to negative values as ξ traverses the interval $[0, \infty)$.

The function $\Delta(\xi, \upsilon, \upsilon')$ can be written as

$$\Delta(\xi, \upsilon, \upsilon')=q(\xi)\ S(\xi, \upsilon, \upsilon')$$

where

$$S(\xi, \upsilon, \upsilon')=\frac{B(\max(\xi, \upsilon), n+1)}{B(\upsilon, n)}-\frac{B(\max(\xi, \upsilon'), n+1)}{B(\upsilon', n)}$$

It is obvious that $B(\upsilon, n)>B(\upsilon', n)$.

If $\xi<\upsilon$ then we have

$$S(\xi, \upsilon, \upsilon')=\frac{B(\upsilon, n+1)}{B(\upsilon, n)}-\frac{B(\upsilon', n+1)}{B(\upsilon', n)}$$

We shall prove that this is positive

If

$$D=B(\upsilon, n+1)B(\upsilon', n)-B(\upsilon, n)B(\upsilon', n+1)$$

then

$$D= B(\upsilon', n)\left(B(\upsilon', n+1)+ \int_\upsilon^{\upsilon'}(\beta(\omega))^{n+1} f(\omega)d\omega\right) - \left(B(\upsilon', n)+ \int_\upsilon^{\upsilon'}(\beta(\omega))^{n} f(\omega)d\omega \right) B(\upsilon', n+1)$$

$$= B(\upsilon', n) \int_\upsilon^{\upsilon'}(\beta(\omega))^{n+1} f(\omega)d\omega - B(\upsilon', n+1) \int_\upsilon^{\upsilon'}(\beta(\omega))^{n} f(\omega)d\omega$$

$$>\beta(\upsilon')\, B(\upsilon', n) \int_\upsilon^{\upsilon'}(\beta(\omega))^{n} f(\omega)d\omega - \beta(\upsilon')\, B(\upsilon', n) \int_\upsilon^{\upsilon'}(\beta(\omega))^{n} f(\omega)d\omega = 0$$

since $\beta(\omega)$ is strictly decreasing. So $S(\xi, \upsilon, \upsilon')>0$.

If $\upsilon\leq\xi\leq\upsilon'$ then

$$S(\xi, \upsilon, \upsilon')=\frac{B(\xi, n+1)}{B(\upsilon, n)}-\frac{B(\upsilon', n+1)}{B(\upsilon', n)}$$

is monotonically decreasing.

If $\upsilon'\leq\xi$ then

$$S(\xi, \upsilon, \upsilon')=\frac{B(\xi, n+1)}{B(\upsilon, n)}-\frac{B(\xi, n+1)}{B(\upsilon', n)}<0$$

So the function $\Delta(\xi, \upsilon, \upsilon')$ changes sign once from positive to negative values as ξ traverses the interval $[0, \infty)$

The case $\upsilon=0$ can be treated in exactly the same way

Theorem 3.2. *For any $n\geq 1$*

a) $g(\xi/\upsilon, n+1)\subset g(\xi/\upsilon, n)$, if $\upsilon<\infty$

b) $g(\xi/\upsilon, n+1)= g(\xi/\upsilon, n)$, if $\upsilon=\infty$

PROOF. Suppose first that $0<\upsilon<\infty$. We follow the same line of proof as in theorem 3.1. So if

$$\Delta(\xi, \upsilon, n)=g(\xi \mid \upsilon, n+1)-g(\xi \mid \upsilon, n)$$

we are required to prove that

$$\int_0^x \Delta(\xi,\upsilon, n)d\xi \geq 0 \; \forall \; x.$$

The result will follow if we prove that the function $\Delta(\xi, \upsilon, n)$ changes sign exactly once from positive to negative values as ξ traverses the interval $[0, \infty)$. Again $\int_0^x \Delta(\xi,\upsilon, n)d\xi=0$. The function $\Delta(\xi, \upsilon, n)$ can be written as

$$\Delta(\xi, \upsilon, n)=q(\xi)\, S(\xi, \upsilon, n)$$

where

$$S(\xi, \upsilon, n)=\frac{B(\max(\xi, \upsilon), n+2)}{B(\upsilon, n+1)}-\frac{B(\max(\xi, \upsilon), n+1)}{B(\upsilon, n)}$$

For $\xi \leq \upsilon$ we have

$$S(\xi, \upsilon, n)=\frac{B(\upsilon, n+2)}{B(\upsilon, n+1)}-\frac{B(\upsilon, n+1)}{B(\upsilon, n)}$$

This is positive because

$D= B(\upsilon, n+2)B(\upsilon, n)-[B(\upsilon, n+1)]^2$

$$=\int_{\upsilon}^{\infty}[(\beta(\omega))^{\frac{n+2}{2}}(f(\omega))^{\frac{1}{2}}]^2 d\omega \int_{\upsilon}^{\infty}[(\beta(\omega))^{\frac{n}{2}}(f(\omega))^{\frac{1}{2}}]^2 d\omega - [B(\upsilon, n+1)]^2$$

$$>\left[\int_{\upsilon}^{\infty}(\beta(\omega))^{n+1} f(\omega)d\omega\right]^2 -[B(\upsilon, n+1)]^2=0$$

where the last inequality was obtained using Schwartz inequality for integrals.

For $\xi \geq \upsilon$ we have

$$S(\xi, \upsilon, n)=\frac{B(\xi, n+2)}{B(\upsilon, n+1)}-\frac{B(\xi, n+1)}{B(\upsilon, n)}$$

and the function $S(\xi, \upsilon, n)$ is strictly decreasing. Since $\beta(\omega)$ is a decreasing function it follows that

$$S(\xi, \upsilon, n)< B(\xi, n+1)[\frac{\beta(\xi)}{B(\upsilon, n+1)}-\frac{1}{B(\upsilon, n)}]$$

and taking ξ large enough we can make $S(\xi, \upsilon, n)$ negative because $\beta(\infty)B(\upsilon, n)<B(\upsilon, n+1)$

Similar arguments can give the proof for the case $\upsilon=0$

If $\upsilon=\infty$ then since $\upsilon \leq \omega$ it follows that $\omega=\infty$ and so

$$g(\xi \mid \upsilon, n)=\beta(\infty)q(\xi)$$

independently of n which proves the relation.

Theorem 3.3. *For any $n \geq 1$ and υ,*

$$g(\xi/\upsilon, n) \subset g(\xi/\infty), \text{ where } g(\xi/\infty) \text{ is given by (8) for } \omega=\infty$$

PROOF. From (8) we have

$$g(\xi| \infty)=\beta(\infty)q(\xi),\ \xi \geq 0$$

if

$$\Delta(\xi)= g(\xi| \upsilon, n)- g(\xi| \infty)$$

Then

$$\Delta(\xi)=q(\xi)\,[\frac{B(\max(\xi,\upsilon),n+1)}{B(\upsilon,n)}-\beta(\infty)]$$

Suppose first that $0<\upsilon<\infty$. Since $\beta(\omega)$ is a decreasing function we shall have

$$B(\upsilon, n+1)>\beta(\infty)B(\upsilon, n)$$

So if $\xi<\upsilon$ then $\Delta(\xi)>0$. If $\xi>\upsilon$ then the function

$$A(\xi)=[\frac{B(\xi,n+1)}{B(\upsilon,n)}-\beta(\infty)]$$

is strictly decreasing in ξ and $\lim_{\xi\to\infty} A(\xi)=-\beta(\infty)$. So $\Delta(\xi)$ changes sign once from positive to negative values as ξ traverses the interval $[0, \infty)$.

If $\upsilon=0$ then

$$\Delta(\xi)=q(\xi)\,[\frac{B(\xi,n+1)}{B(0,n)}-\beta(\infty)]$$

It is easy to see that for sufficiently small values of ξ $\Delta(\xi)$ is positive. This and the fact that $B(\xi, n+1)$ decreases monotonically to zero as $\xi\to\infty$, while $\beta(\infty)>0$ establish the result.

4. THE RANGE FAMILY OF TYPE B

Let ξ be a continuous random variable with density belonging to the family

$$\sigma(\xi|\,\omega)=\gamma(\omega)q(\xi)\psi(\omega,\xi),\ \xi\geq 0 \tag{10}$$

where

$$\psi(\omega,\xi)=\begin{cases}1 \text{ if } \omega\leq\xi<\infty\\ 0 \text{ if } \xi<\omega\end{cases},$$

$$(\gamma(\omega))^{-1}=\int_{\omega}^{\infty} q(\xi)d\xi, \tag{11}$$

ω is an unknown parameter taking values in the interval $I_3=[0,\infty)$ and $q(\xi)$ is a strictly positive and bounded function on $[0,\infty)$, such that the integral in (11) exists for every $\omega \in I_3$. We call this family the "the range family of type B". Defining suitably the function $q(\xi)$ we can obtain from this family some known below truncated distributions.

Let $\xi_1, \xi_2, \ldots, \xi_n$ be the values of n independent observations of the random variable ξ whose p.d. is $\sigma(\xi|\omega)$, and $\tau = \min_{1\le i\le n} \xi_i$. The likelihood function of $\xi_1, \xi_2, \ldots, \xi_n$ is

$$\sigma(\xi_1, \xi_2, \ldots, \xi_n | \omega) = (\gamma(\omega))^n \prod_{i=1}^{n} q(\xi_i)\psi(\omega, \tau)$$

This expression shows that the pair (τ, n) is a sufficient statistic for the parameter ω and also that τ is the m.l.e of ω based on $\xi_1, \xi_2, \ldots, \xi_n$.

If $f(\omega)$ is the prior p.d. of ω then using the Bayes theorem we find, for $\tau>0$, the posterior density of ω

$$f(\omega | \tau, n) = \frac{(\gamma(\omega))^n f(\omega)\psi(\omega, \tau)}{\int_0^\infty (\gamma(\omega))^n f(\omega)\psi(\omega, \tau)d\omega}$$

$$= \begin{cases} \dfrac{(\gamma(\omega))^n f(\omega)}{A(\tau, n)} & \text{if } \omega \le \tau \\ 0 \ \text{ if } \omega > \tau \end{cases}$$

where

$$A(\tau, n) = \int_0^\tau (\gamma(\omega))^n f(\omega)\, d\omega$$

The p.p.d.f. of ξ is now given by

$$\sigma(\xi | \tau, n) = q(\xi)\frac{A(\min(\xi, \tau), n+1)}{A(\tau, n)}, \ \xi\ge 0, \ \tau>0.$$

We give an example of a distribution belonging to this family and the respective p.p.d.f.

Example. If we define the density function $\sigma(\xi|\omega)=\exp(\omega-\xi)$ $\omega<\xi<\infty$ with a prior for parameter ω $f(\omega)=\exp(-\omega)$ then the p.p.d.f is given by

$$\sigma(\xi | \tau, n) = \begin{cases} \dfrac{n-1}{n}\dfrac{\exp[(n-1)\xi]-\exp(-\xi)}{\exp[(n-1)\tau]-1} & \xi < \tau \\[2ex] \dfrac{n-1}{n}\exp(-\xi)\dfrac{\exp(n\tau)-1}{\exp[(n-1)\tau]-1} & \xi > \tau \end{cases}$$

The study of this family was suggested by Iglehart (1964) in his pioneering article. In the next three theorems we establish results parallel to those obtained by Iglehart for the range family of type A.

Theorem 4.1. *For any $n \geq 1$ and τ, τ' with $\tau < \tau'$ we have,*

$$\sigma(\xi/\tau, n) \subset \sigma(\xi/\tau', n)$$

PROOF. If we set

$$\Delta(\xi, \tau, \tau') = \sigma(\xi|\tau, n) - \sigma(\xi|\tau', n)$$

Then we must prove that

$$\int_0^x \{\sigma(\xi|\tau, n) - \sigma(\xi|\tau', n)\} d\xi \geq 0 \ \forall\ x$$

The theorem will be valid if we prove that $\Delta(\xi, \tau, \tau')$ changes sign exactly once from positive to negative values as ξ traverse the interval $[0, \infty)$ (Again it is easily seen that $\int_0^\infty \Delta(\xi, \tau, \tau') d\xi = 0$)

Consider first the case $\tau > 0$. The difference $\Delta(\xi, \tau, \tau')$ can be written as

$$\Delta(\xi, \tau, \tau') = q(\xi)\, S(\xi, \tau, \tau')$$

where

$$S(\xi, \tau, \tau') = \frac{A(\min(\xi, \tau), n+1)}{A(\tau, n)} - \frac{A(\min(\xi, \tau'), n+1)}{A(\tau', n)}$$

If $0 \leq \xi \leq \tau$ then

$$S(\xi, \tau, \tau') = \frac{A(\xi, n+1)}{A(\tau, n)} - \frac{A(\xi, n+1)}{A(\tau', n)} > 0$$

since $A(\tau, n) < A(\tau', n)$

If $\tau \leq \xi \leq \tau'$ then

$$S(\xi, \tau, \tau') = \frac{A(\tau, n+1)}{A(\tau, n)} - \frac{A(\xi, n+1)}{A(\tau', n)}$$

and is strictly decreasing in ξ.

If $\tau' \leq \xi$ then

$$S(\xi, \tau, \tau')=\frac{A(\tau, n+1)}{A(\tau, n)}-\frac{A(\tau', n+1)}{A(\tau', n)}<0.$$

This last inequality can be established as follows.

If

$$D=A(\tau, n+1)A(\tau', n)-A(\tau, n)A(\tau', n+1)$$

then

$$D= A(\tau, n+1)\{ A(\tau, n)+ \int_{\tau}^{\tau'}(\gamma(\omega))^{n} f(\omega)d\omega) - \{ A(\tau, n+1)+ \int_{\tau}^{\tau'}(\gamma(\omega))^{n+1} f(\omega)d\omega) A(\tau, n)$$

$$= A(\tau, n+1) \int_{\tau}^{\tau'}(\gamma(\omega))^{n} f(\omega)d\omega - A(\tau, n) \int_{\tau}^{\tau'}(\gamma(\omega))^{n+1} f(\omega)d\omega$$

$$<\gamma(\tau) A(\tau, n) \int_{\tau}^{\tau'}(\gamma(\omega))^{n} f(\omega)d\omega - \gamma(\tau) A(\tau, n) \int_{\tau}^{\tau'}(\gamma(\omega))^{n} f(\omega)d\omega =0$$

since $\gamma(\omega)$ is strictly increasing. So in the interval $[\tau, \tau']$ there will exist a unique point ξ_0 such that $S(\xi_0, \tau, \tau')=0$ and so $\Delta(\xi, \tau, \tau')\geq(\leq)0$ for $\xi\geq\xi_0$ ($\xi\leq\xi_0$)

If $\tau=0$, then since $0\leq\omega\leq\tau$ we have $\omega=0$ and

$$\sigma(\xi| 0, n)=\gamma(0)q(\xi), \quad \xi\geq0$$

So

$$\Delta(\xi, \tau, \tau')=q(\xi)[\gamma(0)- \frac{A(\min(\xi, \tau'), n+1)}{A(\tau', n)}]$$

The quantity inside the brackets is a decreasing function of ξ, because $A(\tau, n)$ is an increasing function of τ and $\min(\xi, \tau')$ is increasing in τ'. Moreover for $\xi\geq\tau'$ it is negative, while for ξ close to zero it is positive. So it changes sign exactly once as ξ traverses the interval $[0, \infty)$ and this prove the theorem

Theorem 4.2. *For any $n\geq1$ and τ we have*

a) $\sigma(\xi/\tau, n)\subset\sigma(\xi/\tau, n+1)$, if $\tau>0$

b) $\sigma(\xi/\tau, n)= \sigma(\xi/\tau, n+1)$, if $\tau=0$

PROOF. Let us set

$$\Delta(\xi, \tau, n)= \sigma(\xi| \tau, n)- \sigma(\xi| \tau, n+1)=q(\xi)S(\xi, \tau, n), \text{and } \tau>0$$

Where

$$S(\xi, \tau, n)= \frac{A(\min(\xi, \tau), n+1)}{A(\tau, n)} - \frac{A(\min(\xi, \tau), n+2)}{A(\tau, n+1)}$$

Similar with theorem 2.1 we are required to prove that

$$\int_0^x \Delta(\xi,\tau, n)d\xi > 0 \ \forall \ x.$$

If $\xi \le \tau$ then

$$S(\xi, \tau, n)= \int_0^\xi (\gamma(\omega))^{n+1} f(\omega)[\frac{1}{A(\tau, n)} - \gamma(\omega)\frac{1}{A(\tau, n+1)}]d\omega$$

Since $\gamma(\omega)$ is strictly increasing $A(\tau, n+1) > \gamma(0)A(\tau, n)$ and so for $\xi \in [0, \varepsilon]$, where $\varepsilon > 0$ is a suitably small number, we have $S(\xi, \tau, n) > 0$.

If $\tau \le \xi$ then

$$S(\xi, \tau, n)= \frac{A(\tau, n+1)}{A(\tau, n)} - \frac{A(\tau, n+2)}{A(\tau, n+1)} < 0$$

This follows because

$$A(\tau, n)A(\tau, n+2)= \int_0^\tau ((\gamma(\omega))^{n/2} (f(\omega))^{1/2})^2 d\omega \int_0^\tau ((\gamma(\omega))^{(n+2)/2} (f(\omega))^{1/2})^2 d\omega$$

$$> [\int_0^\tau (\gamma(\omega))^{n+1} f(\omega)d\omega]^2 = [A(\tau, n+1)]^2$$

This last inequality is deduced by applying Schwartz 's inequality for integrals. So $S(\xi, \tau, n)$ changes sign once from positive to negative values in the interval $[0, \tau]$ and this prove the first part

If $\tau=0$ then $\sigma(\xi| 0, n)=\gamma(0)q(\xi)$ for all n

Theorem 4.3. *For any $n \ge 1$ and τ we have $\sigma(\xi/\tau, n) \subset \sigma(\xi/ \xi_1)$ where ξ_1 is the first given observation and $\sigma(\xi/ \xi_1)$ is given by (10) for $\omega= \xi_1$*

PROOF. Suppose first that $\tau > 0$.

If $\xi < \xi_1$ then $\sigma(\xi| \xi_1)=0$ and so

$$\sigma(\xi| \tau, n) - \sigma(\xi| \xi_1) > 0$$

If $\xi \geq \xi_1$ then

$$\sigma(\xi| \tau, n) - \sigma(\xi| \xi_1) = q(\xi)\left(\frac{A(\tau, n+1)}{A(\tau, n)} - \gamma(\xi_1)\right) < 0$$

This is negative because since $\tau \leq \xi_1$ and $\gamma(\omega)$ is strictly increasing

$$A(\tau, n+1) < \gamma(\tau) A(\tau, n) < \gamma(\xi_1) A(\tau, n)$$

So there is a jump from positive to negative values at the point $\xi = \xi_1$ and this prove the theorem for $\tau > 0$

If $\tau = 0$ then

$$\sigma(\xi| 0, n) - \sigma(\xi| \xi_1) = q(\xi)\{\gamma(0) - \gamma(\xi_1)\}\psi(\xi_1, \xi)$$

and the same is true.

5. CONCLUDING REMARKS

In this article we have studied stochastic ordering for p.p.d.f. for a random variable ξ, with p.d. belonging to, exponential, range of type A and range of type B families of distributions. Members of these families are designated by an unknown parameter ω, which has a known prior density. Moreover all families accept a sufficient statistic for the parameter ω, which is based on full past history of observations on ξ.

These families of densities can be used to describe the demand fluctuations for various inventory systems. The fact that they accept a sufficient statistics, say w, gives the flexibility to use a Bayesian approach to study the inventory systems of interest. Scarf (1959 and 1960), Karlin (1960), Iglehart (1964), Papachristos (1977a) have proved that the stochastic ordering on the predictive densities, related to the values of w, is transferred into inequalities on the optimal ordering levels for some classes of inventory systems. This high level qualitative result, describes the behaviour of optimal levels in relation to the statistical information for the parameter ω contained in the sufficient statistics. The range family of type B studied here can very well serve to describe the demand for product in certain inventory systems. This is a quite suitable model in cases where some lower bound for the demand can be supposed. The results produced here for the range family of type B, can be used to extend the results of Scarf, Karlin and Iglehart for systems with demand described by the range family of type B. Iglehart (1964) also produced an asymptotic expansion for the optimal inventory level in the case of systems with infinite horizon. We do believe that a similar expansion can be established for the optimal inventory level in the case of the range family of type B. The results contained in theorems 2.3, 3.3, 4.3 can be used to establish in a more rigorous mathematical way the adaptive dynamic programming equation, which is used in the study of inventory systems. They also could be used in finding bounds on, (i) the optimal inventory levels, (ii) the rate of variation (derivatives) of cost functions and (iii) on the cost functions themselves of the models under study, which would be independent from the available statistical information for the parameter ω contained in the sufficient statistics.

Additionally to the above we believe that the reduction of the dimension of the state space, proved by Scarf (1960), Azoury (1985), Lariviere and Porteus (1999) could also be established for the range family of type B.

The case of sufficient statistics based on censored data is also very interested. This arises in inventory systems operating with lost sales status.

APPENDIX

Lemma 1. *The exponential family as defined by 1 has an increasing likelihood ratio*

The proof of this lemma is trivial and omitted

Lemma 2. *If $p(x/\omega)$ has an increasing likelihood ratio and $t(\omega)$ changes sign at most once in its domain, then, for any distribution function $F(\omega)$, the function*

$$g(x)=\int p(x/\omega)\ t(\omega)d\,F(\omega)$$

changes sign at most once. Moreover if $t(\omega)$ changes sign in some direction as ω traverses its domain, $g(x)$ changes sign in the opposite direction.

This lemma is a slight modification of that given by Karlin and Rubin, 1956 (p. 276) and its proof can be derived using the same reasoning.

REFERENCES

Azoury, K. S. (1985), Bayes solution to dynamic inventory models under unknown demand distribution. *Manag. Science*, 31, 1150-1160.

Boland, P.J. , El – Neweihi, E., and Proschan, F. (1992), Stochastic order for redundancy allocations in series and parallel systems, *Advances in applied probability* 24, 161-171.

Braden, D. J. and Freimer M. (1991) Informational dynamics of censored observations. *Manag. Science*, 17, 1390-1404.

Brown G. F. and W. F. Rogers (1972) A Bayesian approach to demand estimation and inventory provisioning *Nav. Res. Log. Quart.* 19, 607-624.

Dunsmorre, I. R. and Aitchison, J. (1975) *Statistical Prediction Analysis* Cambridge University Press.

Iglehart, D. L. (1964) The dynamic inventory problem with unknown demand distributions *Manag. Science*, 10, 429-440.

Karlin, S. (1957) Polya type distributions II. *Ann. Math. Stat.*, 28, 281-308.

Karlin, S. (1960) Dynamic inventory policy with varying stochastic demands *Manag Science*, 6, 231-258.

Karlin S., Rubin, H. (1956) The theory of decision procedures for distributions with a monotone likelihood ratio *Ann. Math Stat* 27, 272-299.

Lariviere, M. A. and Porteus E. L. (1999), Stalking information: Bayesian inventory management with unobserved lost sales *Manag. Science* 45, 346-363.

Mann, H. B. and Withney D. R. (1947) On a test of whether one of two random variables is stochastically larger than the other *Ann. Math Stat* 18, 50-60.

Masse, P. (1962) *Optimal investment decisions* Prentice Hall, Englewood Cliffs, NJ.

Nahmias, S. (1994) Demand estimation in lost sales inventory systems. *Nav. Res. Log.* 41, 739-757.

Papachristos, S. (1977a) *Adaptive dynamic programming and inventory control*. Ph.d theses Faculty of Economics, University of Manchester England.

Papachristos, S. (1977b) A note on the dynamic inventory problem with unknown demand distribution. *Manag. Science.* 23, 1248-1251.

Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory.* Harvard Business School.

Ross, S. (1983) *Introduction to stochastic dynamic programming*. Academic New York.

Savage, L. J. (1972) *The foundations of statistics* 2nd ed. Dover. New York

Scarf, H. (1959) Bayes solutions of the statistical inventory problem *Ann. Math. Stat.* 30, 490-508.

Scarf, H. (1960) Some remarks on Bayes solutions to the inventory problem. *Nav. Res. Log. Quart.* 7, 591-596.

Shaked, M. and J. G. Shanthikumar (1994) *Stochastic orders and their applications* Academic New York.

Stoyan, D. (1983) *Comparison methods for queues and other stochastic models*, D. J. Dalley ed Willey New York.

PERTURBED FREDHOLM BOUNDARY VALUE PROBLEMS FOR DELAY DIFFERENTIAL SYSTEMS

A.A.Boichuk and M.K.Grammatikopoulos

Abstract.

Boundary value problems for systems of ordinary differential equations with a small parameter ε and with a finite number of measurable delays of the argument are considered. Under the assumption that the number m of boundary conditions not exceeds the dimension n of the differential system, it is proved that the point $\varepsilon = 0$ generates $\rho = (n-m)$ - parametric families of solutions of the initial problem. Bifurcation conditions of such solutions are established. Also, it is shown that the index of the operator, which is determined by the initial boundary value problem, is equal to ρ and coincides with the index of the unperturbed problem. Finally, an algorithm for construction of solutions (in the form of Laurent series with a finite number terms of negative power of ε) of the boundary value problem under consideration is suggested.

Keywords: delay equations, Fredholm operator, pseudo-inverse matrix, orthoprojector, boundary value problems, index, families of solutions, functionals.

2000 Mathematical Subject Classification: 34K10, 34K06, 34K18.

1. INTRODUCTION.

We consider in Banach spaces the problem of existence and construction of solutions $z : [a,b] \to R^n$ of systems of ordinary differential equations with a small parameter ε and with a finite number of measurable delays of argument of the form

$$\dot{z}(t) = \sum_{i=1}^{k} A_i(t)z(h_i(t)) + \varepsilon \sum_{i=1}^{k} B_i(t)z(h_i(t)) + g(t), \quad t \in [a,b], \quad h_i(t) \le t, \qquad (1)$$

with the initial conditions

$$z(s) = \psi(s), \quad \text{if} \quad s < a < b,$$

and subject to the boundary conditions

$$lz = \alpha, \qquad \alpha \in R^m. \qquad (2)$$

In this connection, we suppose that the unperturbed problem ($\varepsilon = 0$) does not have solutions for arbitrary nonhomogeneities $g(t)$ belonging to the space considered below and $\alpha \in R^m$ and for arbitrary initial function $\psi : R^1 \setminus [a, b] \to R^n$. Moreover, we suppose that the number m of boundary conditions (2) not exceeds the dimension n of the differential system (1). Further, we establish conditions for the perturbed coefficients $B_i(t)$ and for the delays $h_i(t)$, under which the boundary value problem (1)-(2) admits a family of solutions or a single solution. Finally, we suggest an algorithm for the construction of such solutions.

In the case, where there is no delay effect ($h_i(t) = t, i = 1, ..., k$), and $m = n$, the problem (1)-(2) has been studied in [2, page 252]. Also, in the case, where there is no delay effect ($h_i(t) = t, i = 1, ..., k$), and $A_i(t) = 0$, the periodic ($lz := z(a) - z(b) = 0$) boundary value problem (1)-(2) has been considered in [6].

2. INITIAL VALUE PROBLEMS.

Consider the linear equation with concentrated delay

$$\dot{z}(t) - \sum_{i=1}^{k} A_i(t) z(h_i(t)) = g(t), \quad t \in [a, b], \quad z(s) = \psi(s), \quad \text{if} \quad s < a, \tag{3}$$

where $A_i(t)$ are $(n \times n)$-matrices, while the functions $h_i(t) \leq t$ are measurable for $t \in [a, b]$.

Usually (see [3, 8]) a solution of the delay differential equation (3) is constructing in the space of continuously differentiable functions as a continuous extension of the defined on an initial set function $\psi(s)$ to the interval $[a, b]$. A such definition requires the initial function $\psi(s)$ and the solution $z(s)$ to be "continuously joined" at the point $s = a$, i.e. $\psi(a) = z(a)$. This leads to the notion of infinite-dimensional fundamental matrix (introduced for the investigation of the initial problem (3)), whose dimension coincides with the dimension of the basis of the space of initial functions.

Following [1], we will present here basic notions concerning the initial problem (3) for delay differential systems with a finite-dimensional fundamental matrix.

Let $h_i : [a, b] \to R^1$ and $\psi : R^1 \setminus [a, b] \to R^n$ be given functions. Define

$$(S_{h_i} z)(t) = \begin{cases} z(h_i(t)), & \text{if } h_i(t) \in [a, b], \\ 0, & \text{if } h_i(t) \notin [a, b], \end{cases} \tag{4}$$

where S_{h_i} denotes (see [1, page 10]) the operator of so-called inner composition, and put

$$\psi^{h_i}(t) = \begin{cases} 0, & \text{if } h_i(t) \in [a, b], \\ \psi(h_i(t)), & \text{if } h_i(t) \notin [a, b]. \end{cases} \tag{5}$$

Now, in view of (4) and (5), the equation (3) can be rewritten in the form

$$(Lz)(t) := \dot{z}(t) - \sum_{i=1}^{k} A_i(t)(S_{h_i} z)(t) = \varphi(t), \tag{6}$$

where

$$\varphi(t) = g(t) + \sum_{i=1}^{k} A_i(t)\psi^{h_i}(t). \tag{7}$$

The transformations (4) and (5) allow to join the initial function $\psi(s)$, $s < a$ of (3) to the absolute term and to apply to the equation (6) the well-developed methods of linear functional analysis. We will investigate the equation (6) under the assumption that the bounded on $[a, b]$ operator L maps the Banach space $D_p^n[a, b]$ of absolutely continuous functions $z : [a, b] \to R^n$ with the norm

$$||z||_{D_p^n} = ||\dot{z}||_{L_p^n} + ||z(a)||_{R^n}$$

into the Banach space $L_p^n[a, b]$ $(1 < p < \infty)$ of integrable vector functions $\varphi : [a, b] \to R^n$ with the standard in these spaces norm.

According to [1, page 13], the absolutely continuous on $[a, b]$ vector function $z(t) \in D_p^n[a, b]$, for which $\dot{z}(t) \in L_p^n[a, b]$, is called a *solution of the delay differential system* (6), if $z(t)$ satisfies the system (6) almost everywhere on $[a, b]$.

In sequel, we will consider the equation (6) rewritten in the form

$$\dot{z}(t) = A(t)(S_h z)(t) + \varphi(t).$$

Here $A(t) = (A_1(t), ..., A_k(t))$ is an $(n \times N)$ - matrix $(N = nk)$ consisting of $(n \times n)$ -matrices $A_i(t)$, $(S_h z)(t) = \operatorname{col}[(S_{h_1} z)(t), ..., (S_{h_k} z)(t)]$ is an N-dimensional column vector and $\varphi(t)$ is an n-dimensional column vector given by (7). The operator of inner composition S_h maps the space D_p^n into the space $L_p^N = \underbrace{L_p^n \times \cdots \times L_p^n}_{k\,\text{time}}$; i.e. $S_h : D_p^n \to L_p^N$. For the operator $S_{h_i} : D_p^n \to L_p^n$ we have the following representation

$$(S_{h_i} z)(t) = \int_a^b \chi_{h_i}(t, s)\dot{z}(s)ds + \chi_{h_i}(t, a)z(a), \tag{8}$$

where $\chi_{h_i}(t, s)$ is the characteristic function of the set

$$\Omega = \{(t, s) \in [a, b] \times [a, b] : a \le s \le h_i(t) \le b\},$$

and it means (see [1, page 17] or [4]) that

$$\chi_{h_i}(t, s) = \begin{cases} 1, & \text{if } (t, s) \in \Omega, \\ 0, & \text{if } (t, s) \notin \Omega. \end{cases}$$

It is well-known, that the nonhomogeneous delay operator equation (6) is solvable for any right-hand side $\varphi(t) \in L_p^n[a, b]$ and admits an n-parametric family of solutions in the form

$$z(t) = X(t)c + \int_a^b K(t, \tau)\varphi(\tau)\, d\tau, \tag{9}$$

where the $(n \times n)$-matrix $K(t, \tau)$ is called *Cauchy matrix*. For any fixed τ this matrix is a solution to the following matrix Cauchy problem:

$$\frac{\partial K(t, \tau)}{\partial t} = A(t)(S_h K(\cdot, \tau))(t), \qquad K(\tau, \tau) = I.$$

In what follows, we assume that the matrix $K(t,\tau)$ is defined in the square $[a,b]\times[a,b]$, where $K(t,\tau)\equiv 0$ for $a\le t<\tau\le b$. The finite dimensional fundamental $(n\times n)$-matrix of the homogeneous ($\varphi(t)\equiv 0$) delay equation corresponding to (6) is of the form $X(t)=K(t,a)$. By $(S_hK(\cdot,\tau))(t)$ we denote the $(N\times n)$-matrix, whose columns are obtained by applying the operator of inner composition S_h to the corresponding columns of $(n\times n)$-matrix $K(t,\tau)$.

3. FREDHOLM BOUNDARY VALUE PROBLEMS.

Consider the following linear nonhomogeneous boundary value problem

$$(Lz)(t):=\dot{z}(t)-A(t)(S_hz)(t)=\varphi(t),\qquad t\in[a,b], \tag{10}$$

$$lz=\alpha. \tag{11}$$

Here $L\,:\,D_p^n[a,b]\to L_p^n[a,b]$ is the bounded linear delay differential operator, $l=\mathrm{col}\,[l_1,...,l_m]$ is an m-dimensional bounded vector functional, the number m of components of which, in generally, is not equal to the dimension n of the differential system. Functionals l_i map the space $D_p^n[a,b]$ into the space R , while $l\ :D_p^n[a,b]\to R^m$; $\alpha\in R^m$. Moreover, the rows of the matrices $A_i(t)$ and the column vector $\varphi(t)$ belong to the space $L_p^n[a,b]$, i.e. $A_i(t),\varphi(t)\in L_p^n[a,b]$. It is well-known (see [1, page 33] or [7, page 86]), that this boundary value problem defines a Fredholm operator, which maps the space $D_p^n[a,b]$ into the space $L_p^n[a,b]\times R^m$.

Here we are interested in necessary and sufficient conditions for solvability of the above problem as well as in finding a representation of its solution $z(t)\in D_p^n[a,b]$.

The general solution of the equation (10) is of the form (9). So, substituting (9) into the boundary conditions (11), we obtain the algebraic (with respect to $c\in R^n$) system

$$Qc=\alpha-l\int_a^b K(\cdot,\tau)\varphi(\tau)\,d\tau \tag{12}$$

with the $(m\times n)$-dimensional constant matrix $Q=lX(\cdot)$ and with $rank\ Q=n_1$. From the system (12) we can find the constant $c\in R^n$ for which the solution (9) of the system (10) is also a solution of the boundary value problem (10)-(11).

Using the theory of pseudo-inverse matrices and ortoprojectors (see, for example [9], or [2, Theorem 3.9, page 92]), we receive necessary and sufficient conditions for solvability of the algebraic system (12) and for the existence of solutions for the boundary value problem (10)-(11).

Let $P_Q\,:\,R^n\to N(Q)=kerQ$ and $P_{Q^*}\,:\,R^m\to N(Q^*)=kerQ^*=cokerQ$ denote, respectively, the $(n\times n)$ and $(m\times m)$-dimensional matrices-orthoprojectors on the kernel and the cokernel of the matrix Q with the properties: $P_Q^2=P_Q=P_Q^*$;

$P_{Q^*}^2 = P_{Q^*} = P_{Q^*}^*$, where the symbol $*$ means the operation of transposition. Also let Q^+ denote the $(n \times m)$-matrix, which is a Moore-Penrose pseudo-inverse to Q.

The algebraic system (12) is solvable if and only if its right-hand side belongs to the orthogonal complement $N^{\perp}(Q^*) = R(Q)$ of the subspace $N(Q^*)$. This means that the equality

$$P_{Q^*}\Big\{\alpha - l\int_a^b K(\cdot,\tau)\varphi(\tau)\,d\tau\Big\} = 0$$

holds. Since $rank\ P_Q = n - rank\ Q = n - n_1 = r$ and $rank\, P_{Q^*} = m - rank\, Q^* = m - n_1 = d$, we use the symbol $P_{Q_d^*}$ to denote the $(d \times m)$-matrix, whose rows represent a complete set of d linearly independent rows of the $(m \times m)$-matrix P_{Q^*}. Let P_{Q_r} be an $n \times r$ matrix, whose columns represent a complete set of r linearly independent columns of the $(n \times n)$ matrix P_Q. Then the last condition is expressed by the equality

$$P_{Q_d^*}\Big\{\alpha - l\int_a^b K(\cdot,\tau)\varphi(\tau)\,d\tau\Big\} = 0. \tag{13}$$

If (13) holds, then

$$c = Q^+\Big(\alpha - l\int_a^b K(\cdot,\tau)\varphi(\tau)d\tau\Big) + P_{Q_r}c_r, \quad P_{Q_r}c_r \in N(Q), \forall c_r \in R^r$$

is a solution of the algebraic system (12). Substituting the obtained value of c into (9), we receive the general solution of the boundary value problem (10)-(11)

$$z(t,c) = X(t)P_{Q_r}c_r + X(t)Q^+\alpha + \int_a^b K(t,\tau)\varphi(\tau)d\tau - X(t)Q^+l\int_a^b K(\cdot,\tau)\varphi(\tau)d\tau.$$

This solution can be rewritten in the form

$$z(t,c_r) = X_r(t)c_r + (G\varphi)(t) + X(t)Q^+\alpha, \tag{14}$$

where $X_r(t) = X(t)P_{Q_r}$ is the fundamental matrix of the homogeneous boundary value problem

$$\dot z(t) = A(t)(S_h z)(t), \qquad lz = 0. \tag{15}$$

The operator $(G\varphi)(t)$ is defined as

$$(G\varphi)(t) = \int_a^b K(t,\tau)\varphi(\tau)d\tau - X(t)Q^+l\int_a^b K(\cdot,\tau)\varphi(\tau)d\tau$$

and is called *generalized Green operator* for the boundary value problem (10)-(11) (see [2, page 134]).

From the above observation follows

Theorem 1. *Consider the boundary value problem (10)-(11). Then:*

1) the operator $\Lambda_0 : D_p^n[a,b] \to L_p^n[a,b] \times R^m$ defined by the formula

$$(\Lambda_0 z)(t) \stackrel{def}{=} col\ [\ \dot z(t) - A(t)(S_h z)(t), \quad lz\ ] \tag{16}$$

is a Fredholm one with

$$ind\ \Lambda_0 = dimker\ \Lambda_0 - dimker\ \Lambda_0^* = \rho = r - d = n - m,$$

where the operator Λ_0^* *is the adjoint one to* Λ_0;

2) the homogeneous boundary value problem (15) has r *and only* r *linearly independent solutions* $X_r(t)c_r,\ \forall c_r \in R^r$ ($dimker\ \Lambda_0 = r = n - rank\ Q = n - n_1$);

3) the nonhomogeneous boundary value problem (10)- (11) is solvable for those and only those $\varphi(t) \in L_p^n[a,b]$ *and* $\alpha \in R^m$ *which satisfy (13)* ($dimker\ \Lambda_0^* = d = m - rank\ Q^* = m - n_1$) *and its solutions form the* r*-parametric family (14).*

These results will essentially be applied for obtaining new existence conditions for the solutions of perturbed linear and nonlinear boundary value problems for delay equations.

Remark 1. If the vector functional l satisfies the relation

$$l \int_a^b K(\cdot,\tau)\varphi(\tau)d\tau = \int_a^b lK(\cdot,\tau)\varphi(\tau)d\tau, \tag{17}$$

then the generalized Green operator $(G\varphi)(t)$ obtains the form

$$(G\varphi)(t) = \int_a^b G(t,\tau)\varphi(\tau)d\tau.$$

The $(n \times n)$ - matrix $G(t,\tau)$ is the kernel of the integral representation of the operator $(G\varphi)(t)$ and has the form

$$G(t,\tau) = K(t,\tau) - X(t)Q^+lK(\cdot,\tau) \tag{18}$$

and is called *generalized Green matrix*. Without loss of generality, we will assume below that the condition (17) is fulfilled.

For example, the relation (17) holds for periodic $lz := z(a) - z(b) = 0$ and for multipoint $lz = \sum_{i=1}^{k} M_i z(t_i)$ boundary conditions as well as for the conditions of the form of Riemann-Stieltjes integral

$$lz = \int_a^b d\Phi(t)z(t),$$

where $\Phi(t)$ is an $m \times n-$ matrix, whose components are functions with bounded variation on $[a,b]$. In the last case

$$lK(\cdot,\tau) = \int_\tau^b d\Phi(t)K(t,\tau),$$

because $K(t,\tau) \equiv 0$ for $t < \tau$.

Remark 2. The solvability condition (13) for the problem (10)-(11) holds, provided that the initial function ψ is appropriately chosen. In fact, using (5), we can represent the condition (13) in the form

$$P_{Q_d^*}\Big\{\alpha - l\int_a^b K(\cdot,\tau)\Big[g(\tau)+\sum_{i=1}^{k} A_i(\tau)\psi^{h_i}(\tau)\Big]\,d\tau\Big\} = 0.$$

This allows us to get the solvability of the problem (10)-(11) by varying the function ψ. But, if nonhomogeneities $g(t)\in L_p^n[a,b]$ and $\alpha\in R^m$ and the initial vector-function $\psi : R^1\setminus[a,b]\to R^n$ are arbitrary, then the solvability condition (13) for the problem (10)-(11) does not hold. So, it is necessary to suggest a method for regularization of a boundary value problem which is not everywhere solvable.

4. PERTURBED BOUNDARY VALUE PROBLEMS.

Consider the perturbed nonhomogeneous linear boundary value problem (1)-(2), which, in view of (4) and (5), can be rewritten in the form

$$\dot z(t) = A(t)(S_h z)(t) + \varepsilon B(t)(S_h z)(t) + \varphi(t),\quad lz=\alpha,\quad t\in[a,b]. \tag{19}$$

As before, we will assume that $A(t)=(A_1(t),...,A_k(t))$ and $B(t)=(B_1(t),...,B_k(t))$ are $(n\times N)$ - matrices $(N=nk)$ consisting respectively, of $(n\times n)$ -matrices $A_i(t)\in L_p^n[a,b]$ and $B_i(t)\in L_p^n[a,b]$. Assume that the generating boundary value problem

$$\dot z(t) = A(t)(S_h z)(t) + \varphi(t),\quad lz=\alpha, \tag{20}$$

which follows from (19) for $\varepsilon=0$, has no solution for arbitrary nonhomogeneities $\varphi(t)\in L_p^n[a,b]$ and $\alpha\in R^m$. Then Theorem 1 shows that the solvability criterion (13) does not hold for the problem (20), because the nonhomogeneities are arbitrary. Thus we arrive at the following question.

Is it possible to make the problem (20) to become solvable by means of linear perturbations and, if it is possible, then what kind should be the perturbations $B_i(t)$ and the delays $h_i(t)$, in order to become the boundary value problem (19) everywhere solvable?

We can answer this question with the help of the $(d\times r)$-matrix

$$B_0=\int_a^b H(\tau)B(\tau)(S_hX_r)(\tau)\,d\tau,\quad H(\tau)=P_{Q_d^*}lK(\cdot,\tau),$$

the construction of which involves the coefficients of the problem (19).Using the method of [10] we can find conditions, when solutions of the boundary value problem (19) appear in the form of Laurent series (in powers of a small parameter ε) with finite number terms of negative power of ε.

Below we will prove a statement, which enables us to solve the above problem. In order to state this result, we remind that by P_{B_0} we denote an $(r\times r)$-matrix - orthoprojector, projecting R^r onto the null-space $N(B)$ of the $(d\times r)$-matrix B_0, and by $P_{B_0^*}$ denote a $(d \times d)$-matrix -orthoprojector, projecting R^d onto the null-space $N(B_0^*)$ of the $(r \times d)$-matrix $B_0^* = B_0^t$. Now we can formulate the following

Lemma. *Consider the boundary value problem (19) and assume that for arbitrary nonhomogeneities $\varphi(t) \in L_p^n[a,b]$ and $\alpha \in R^m$ the generating boundary value problem (20) has no solutions.*

If the following equivalent relations

$$P_{B_0^*} = 0 \quad \Longleftrightarrow \quad rank\ B_0 = d \tag{21}$$

hold, then for arbitrary $\varphi(t) \in L_p^n[a,b]$ and $\alpha \in R^m$ the boundary value problem (19) has at least one solution in the form of the series

$$z(t,\varepsilon) = \sum_{i=-1}^{\infty} \varepsilon^i z_i(t), \tag{22}$$

converging for $\varepsilon \in (0, \varepsilon_*]$, where ε_* is an appropriate constant characterizing the domain of the convergence of the series (22).

Proof. Substitute (22) into (19) and equate the coefficients at equal powers of ε. For ε^{-1}, we obtain the homogeneous boundary value problem

$$\dot{z}_{-1} = A(t)(S_h z_{-1})(t), \qquad l z_{-1} = 0, \tag{23}$$

which determines $z_{-1}(t)$.

By the hypothesis of Theorem 1 , the homogeneous boundary value problem (23) has an r-parametric $(r = n - n_1)$ family of solutions $z_{-1}(t, c_{-1}) = X_r(t)c_{-1}$, where the r-dimensional column vector $c_{-1} \in R^r$ can be determined from the solvability condition of the problem for $z_0(t)$.

For ε^0, we get the boundary value problem

$$\dot{z}_0 = A(t)z_0 + B(t)(S_h z_{-1})(t) + \varphi(t), \qquad l z_0 = \alpha, \tag{24}$$

which determines $z_0(t)$.

It is an implication of Theorem 1 that the solvability criterion for the problem (24) has the form

$$P_{Q_d^*}\alpha - \int_a^b H(\tau)\{\varphi(\tau) + B(\tau)(S_h X_r)(\tau)c_{-1}\}d\tau = 0,$$

from which we receive with respect to $c_{-1} \in R^r$ the algebraic system

$$B_0 c_{-1} = P_{Q_d^*}\alpha - \int_a^b H(\tau)\varphi(\tau)d\tau, \tag{25}$$

where

$$B_0 = \int_a^b H(\tau)B(\tau)(S_h X_r)(\tau)\,d\tau, \quad H(\tau) = P_{Q_d^*} l K(\cdot, \tau).$$

The last system is solvable for arbitrary $\varphi(t) \in L_p^n[a,b]$ and $\alpha \in R^m$ if and only if the condition $P_{B_0^*} = 0$ is satisfied. The system (25) becomes resolvable with respect to $c_{-1} \in R^r$ up to an arbitrary constant vector $P_{B_0}c$ $(\forall c \in R^r)$ from the null-space of matrix B_0 with

$$c_{-1} = -B_0^+\Big\{P_{Q_d^*}\alpha - \int_a^b H(\tau)\varphi(\tau)\,d\tau\Big\} + P_{B_0}c.$$

This solution can be rewritten in the form

$$c_{-1} = \bar{c}_{-1} + P_{B_\rho}c_\rho, \quad \forall\, c_\rho \in R^\rho, \tag{26}$$

where

$$\bar{c}_{-1} = -B_0^+\Big\{P_{Q_d^*}\alpha - \int_a^b H(\tau)\varphi(\tau)\,d\tau\Big\} \tag{27}$$

and P_{B_ρ} is an $(r \times \rho)$ - dimensional matrix, whose columns are complete set of ρ linearly independent columns of $(r \times r)$ - dimensional matrix P_{B_0}, with

$$\rho = rank\ P_{B_0} = r - rank\ B_0 = r - d = n - m.$$

So, for the solutions of the problem (23) we have the following expression

$$z_{-1}(t, c_\rho) = \bar{z}_{-1}(t, \bar{c}_{-1}) + X_r(t)P_{B_\rho}c_\rho, \quad \forall\, c_\rho \in R^\rho; \quad \bar{z}_{-1}(t, \bar{c}_{-1}) = X_r(t)\bar{c}_{-1}.$$

Assuming that (21) holds, the boundary value problem (24) has the r -parametric family of solutions

$$z_0(t, c_0) = X_r(t)c_0 + X(t)Q^+\alpha + \tag{28}$$

$$+\int_a^b G(t,\tau)[\varphi(\tau) + B(\tau)S_h(\bar{z}_{-1}(\cdot, \bar{c}_{-1}) + X_r(\cdot)P_{B_\rho}c_\rho)(\tau)]d\tau.$$

Here c_0 is an r-dimensional constant vector, which is determined at the next step from the solvability condition of the boundary value problem for $z_1(t)$.

For ε^1, we get the boundary value problem

$$\dot{z}_1 = A(t)z_1 + B(t)(S_h z_0)(t), \qquad l z_1 = 0, \tag{29}$$

which determines $z_1(t)$. The solvability criterion for the problem (29) has the form

$$\int_a^b H(\tau)B(\tau)S_h\{X_r(\star)c_0 + X(\star)Q^+\alpha +$$

$$+\int_a^b G(\star, s)[\varphi(s) + B(s)S_h(\bar{z}_{-1}(\cdot, \bar{c}_{-1}) + X_r(\cdot)P_{B_\rho}c_\rho)(s)]ds\}(\tau)d\tau = 0$$

or equivalently the form

$$B_0c_0 = \int_a^b H(\tau)B(\tau)S_h\{X(\star)Q^+\alpha + \tag{30}$$

$$+\int_a^b G(\star, s)[\varphi(s) + B(s)S_h(\bar{z}_{-1}(\cdot, \bar{c}_{-1}) + X_r(\cdot)P_{B_\rho}c_\rho)(s)]ds\}(\tau)d\tau.$$

The algebraic system (30) has the following family of solutions

$$c_0 = B_0^+ \int_a^b H(\tau)B(\tau)S_h\{X(\star)Q^+\alpha+ \tag{31}$$

$$+\int_a^b G(\star,s)[\varphi(s)+B(s)(S_h\bar{z}_{-1}(\cdot,\bar{c}_{-1}))(s)]ds\}(\tau)d\tau+$$

$$+[I_r + B_0^+ \int_a^b H(\tau)B(\tau)S_h\{\int_a^b G(\star,s)B(s)(S_hX_r(\cdot))(s)ds\}(\tau)d\tau]P_{B_\rho}c_\rho =$$

$$= \bar{c}_0 + [\cdot,\cdot,\cdot]P_{B_\rho}c_\rho,$$

where

$$\bar{c}_0 = B_0^+ \int_a^b H(\tau)B(\tau)S_h\{X(\star)Q^+\alpha+ \tag{32}$$

$$+\int_a^b G(\star,s)[\varphi(s)+B(s)(S_h\bar{z}_{-1}(\cdot,\bar{c}_{-1}))(s)]ds\}(\tau)d\tau,$$

and

$$[\cdot,\cdot,\cdot] = I_r + B_0^+ \int_a^b H(\tau)B(\tau)S_h\{\int_a^b G(\star,s)B(s)(S_hX_r(\cdot))(s)ds\}(\tau)d\tau.$$

So, for the ρ -parametric family of solutions of the problem (23) we have the following expression

$$z_0(t,c_\rho) = \bar{z}_0(t,\bar{c}_0) + \bar{X}_0(t)P_{B_\rho}c_\rho, \quad \forall\ c_\rho \in R^\rho,$$

where

$$\bar{z}_0(t,\bar{c}_0) = X_r(t)\bar{c}_0 + X(t)Q^+\alpha+ \tag{33}$$

$$+\int_a^b G(t,\tau)[\varphi(\tau)+B(\tau)(S_h\bar{z}_{-1}(\cdot,\bar{c}_{-1}))(\tau)]d\tau$$

and

$$\bar{X}_0(t,\bar{c}_0) = X_r(t)[I_r + B_0^+ \int_a^b H(\tau)B(\tau)S_h\{\int_a^b G(\star,s)B(s)(S_hX_r(\cdot))(s)ds\}(\tau)d\tau].$$

Again, assuming that (21) holds, the boundary value problem (29) has the r -parametric family of solutions

$$z_1(t,c_1) = X_r(t)c_1+ \tag{34}$$

$$+\int_a^b G(t,\tau)B(\tau)S_h(\bar{z}_0(\cdot,\bar{c}_0)+\bar{X}_0(\cdot)P_{B_\rho}c_\rho)(\tau)d\tau.$$

Here c_1 is an r-dimensional constant vector, which is determined at the next step from the solvability condition of the boundary value problem for $z_2(t)$

$$\dot{z}_2 = A(t)z_2 + B(t)(S_hz_1)(t), \qquad lz_2 = 0. \tag{35}$$

The solvability criterion for the problem (35) has the form

$$\int_a^b H(\tau)B(\tau)S_h\{X_r(\star)c_1 + \int_a^b G(\star, s)B(s)S_h(\bar{z}_0(\cdot, \bar{c}_0) + \bar{X}_0(\cdot)P_{B_\rho}c_\rho)(s)ds\}(\tau)d\tau = 0,$$

or the form

$$B_0c_1 = -\int_a^b H(\tau)B(\tau)(S_h\{\int_a^b G(\star, s)B(s)S_h(\bar{z}_0(\cdot, \bar{c}_0) + \bar{X}_0(\cdot)P_{B_\rho}c_\rho)(s)ds\})(\tau)d\tau.$$

Under the condition (21), the last equation has the $\rho-$ parametric family of solutions

$$c_1 = \bar{c}_1 + \{.,.,.\},$$

where

$$\bar{c}_1 = -B^+\int_a^b H(\tau)B(\tau)(S_h\{\int_a^b G(\star, s)B(s)(S_h\bar{z}_0(\cdot, \bar{c}_0))(s)ds\})(\tau)d\tau$$

and

$$\{.,.,.\} = \{I_r - B^+\int_a^b H(\tau)B(\tau)(S_h\{\int_a^b G(\star, s)B(s)(S_h\bar{X}_0(\cdot))(s)ds\})(\tau)d\tau\}P_{B_\rho}c_\rho.$$

So, for the coefficient $z_1(t, c_1) = z_1(t, \rho)$ we have the following expression

$$z_1(t, c_\rho) = \bar{z}_1(t, \bar{c}_1) + \bar{X}_1(t)P_{B_\rho}c_\rho, \quad \forall\ c_\rho \in R^\rho,$$

where

$$\bar{z}_1(t, \bar{c}_0) = X_r(t)\bar{c}_1 + \int_a^b G(t, \tau)B(\tau)(S_h\bar{z}_0(\cdot, \bar{c}_0))(\tau)d\tau; \tag{36}$$

and

$$\bar{X}_1(t) = X_r(t)[I_r + B_0^+\int_a^b H(\tau)B(\tau)S_h\{\int_a^b G(\star, s)B(s)(S_h\bar{X}_0(\cdot))(s)ds\}(\tau)d\tau].$$

Continuing this process, by assuming that (21) holds, it follows by induction that the coefficients $z_i(t, c_i) = z_i(t, c_\rho)$ of the series (22) can be determined from the relevant boundary value problems as follows

$$z_i(t, c_\rho) = \bar{z}_i(t, \bar{c}_i) + \bar{X}_i(t)P_{B_\rho}c_\rho, \qquad \forall\ c_\rho \in R^\rho, \tag{37}$$

where

$$\bar{z}_i(t, \bar{c}_i) = X_r(t)\bar{c}_1 + \int_a^b G(t, \tau)B(\tau)S_h\bar{z}_{i-1}(\cdot, \bar{c}_{i-1})(\tau)d\tau;$$

$$\bar{c}_i = -B^+\int_a^b H(\tau)B(\tau)(S_h\{\int_a^b G(\star, s)B(s)S_h\bar{z}_{i-1}(\cdot, \bar{c}_{i-1})(s)ds\})(\tau)d\tau, \ (i = 1, 2, ...);$$

$$\bar{X}_i(t) = X_r(t)[I_r + B_0^+\int_a^b H(\tau)B(\tau)S_h\{\int_a^b G(\star, s)B(s)(S_h\bar{X}_{i-1}(\cdot))(s)ds\}(\tau)d\tau],$$

$$i = 0, 1, 2, ..., \quad \text{and} \quad \bar{X}_{-1}(t) = X_r(t).$$

Since the convergence of the series (22) can be proved by traditional methods of majorization, the proof of the lemma is complete.

From the above Lemma we have the following conclusions.

The boundary value problem (19) determines the bounded operator

$$(\Lambda_\varepsilon z)(t) \stackrel{def}{=} col[\dot{z}(t) - A(t)(S_h z)(t) - \varepsilon B(t)(S_h z)(t), \quad lz] \tag{38}$$

which is acting from the space $D_p^n[a,b]$ to the space $L_p^n[a,b] \times R^m$, $1 < p < \infty$. Under the assumption (21), the problem (19) is always solvable in the Banach spaces under consideration. This means that the image of the operator Λ_ε coincides with the whole space $L_p^n[a,b] \times R^m$, i.e. Im $\Lambda_\varepsilon = L_p^n[a,b] \times R^m$. Therefore, Λ_ε is a normally-solvable operator (see [5, 7]), while the boundary value problem adjoint to the homogeneous one

$$\dot{z}(t) = A(t)(S_h z)(t) + \varepsilon B(t)(S_h z)(t), \qquad lz = 0 \in R^m, \tag{39}$$

has only trivial solutions, i.e. $dimker\ \Lambda_\varepsilon^* = 0$, $\varepsilon \neq 0$, where the operator Λ_ε^* is the adjoint one to Λ_ε in the corresponding spaces. Note that our problem do not need the construction of the adjoint problem. Such a construction for the unperturbed boundary value problem (10)-(11) is given in [1, p.36].

As it shown in the proof of Lemma, $dimker\ \Lambda_\varepsilon = \rho = r - d$. This, together with the mentioned above property $dimker\ \Lambda_\varepsilon^* = 0$, means that the normally-solvable operator Λ_ε is a Fredholm one. Now, it is not difficult to see that for the differential operator (38) with delayed arguments, the well-known fact from the theory of operators (see [7, page 86] or [5]), concerning the maintenance under small perturbations of the index of the Fredholm operator Λ_0 (16), is satisfied. Indeed, since by Theorem 1,

$$dimker\ \Lambda_0 = r, \ dimker\ \Lambda_0^* = d,$$

and, by Lemma,

$$dimker\ \Lambda_\varepsilon = r - d, \quad dimker\ \Lambda_\varepsilon^* = 0, \quad \varepsilon \neq 0,$$

it follows that

$$ind\ \Lambda_0 = ind\ \Lambda_\varepsilon.$$

From the previous discussion, we have

Theorem 2. *Consider the boundary value problem*

$$\dot{z}(t) = A(t)(S_h z)(t) + \varepsilon B(t)(S_h z)(t) + \varphi(t), \quad lz = \alpha \in R^m, \tag{40}$$

and assume that for arbitrary nonhomogeneities $\varphi(t) \in L_p^n[a,b]$ and $\alpha \in R^m$ the generating boundary value problem (20) has no solutions.

If the condition

$$rank\ [B_0 = \int_a^b H(\tau)B(\tau)(S_h X_r)(\tau)\, d\tau] = d = m - n_1 \quad (\ rank\ Q = n_1) \tag{41}$$

holds, then:

1) the operator $\Lambda_\varepsilon : D_p^n[a,b] \to L_p^n[a,b] \times R^m$ $(1 < p < \infty)$ *defined by the formula(38) is a Fredholm one with*

$$ind\Lambda_\varepsilon = dimker\Lambda_\varepsilon - dimker\Lambda_\varepsilon^* = \rho = r - d = n - m,$$

$$ind\Lambda_0 = dimker\Lambda_0 - dimker\Lambda_0^* = \rho = r - d = n - m;$$

where the operator Λ_ε^* *is the adjoint one to* Λ_ε, $(\ dimker\ \Lambda_0 = r,\ dimker\ \Lambda_0^* = d\ )$;

2) the homogeneous boundary value problem (39) has a $\rho-$ *parametric family of solutions*

$$z_0(t,\varepsilon,c_\rho) = \sum_{i=-1}^{\infty} \varepsilon^i \bar{X}_i(t) P_{B_\rho} c_\rho, \quad \forall\ c_\rho \in R^\rho, \quad (\rho = dimker\Lambda_\varepsilon) \tag{42}$$

with the properties

$$z_0(\cdot,\varepsilon,c_\rho) \in D_p^n[a,b], \quad \dot{z}_0(\cdot,\varepsilon,c_\rho) \in L_p^n[a,b], \quad z_0(t,\cdot,c_\rho) \in C(0,\varepsilon_*];$$

3) the boundary value problem adjoint to (39) *has only trivial solutions* $(dimker\ \Lambda_\varepsilon^* = 0,\ \ \varepsilon \neq 0\ )$;

4) for arbitrary $\varphi(t) \in L_p^n[a,b]$ *and* $\alpha \in R^m$ *the boundary value problem (40) has a* ρ*-parametric set of solutions* $z(t,\varepsilon) = z(t,\varepsilon,c_\rho)$ *with the properties*

$$z(\cdot,\varepsilon,c_\rho) \in D_p^n[a,b], \quad \dot{z}(\cdot,\varepsilon,c_\rho) \in L_p^n[a,b], \quad z(t,\cdot,c_\rho) \in C(0,\varepsilon_*],$$

in the form of the series

$$z(t,\varepsilon,c_\rho) = \sum_{i=-1}^{\infty} \varepsilon^i [\bar{z}_i(t,\bar{c}_i) + \bar{X}_i(t) P_{B_\rho} c_\rho] \quad \forall\ c_\rho \in R^\rho \tag{43}$$

converging for $\varepsilon \in (0,\varepsilon_*]$, *where* ε_* *is as in Lemma and the coefficients* $\bar{z}_i(t,\bar{c}_i)$, $\bar{c}_i$ *and* $\bar{X}_i(t)$ *can be determined from (37).*

In the case, when the number m of boundary conditions is equal to the dimension n of the differential system (40), from the condition (41) we have

$$rank\ [B_0 = \int_a^b H(\tau)B(\tau)(S_h X_\tau)(\tau)\, d\tau] = r = d$$

and from Theorem 2 we receive the following

Corollary. *Consider the boundary value problem*

$$\dot{z}(t) = A(t)(S_h z)(t) + \varepsilon B(t)(S_h z)(t) + \varphi(t), \quad lz = \alpha \in R^n, \tag{44}$$

and assume that for arbitrary nonhomogeneities $\varphi(t) \in L_p^n[a,b]$ *and* $\alpha \in R^n$ *the generating boundary value problem (20) has no solutions. If the condition*

$$det\ B_0 \neq 0 \tag{45}$$

holds, then:
1) the operator $\Lambda_\varepsilon : D_p^n[a,b] \to L_p^n[a,b] \times R^n$, *defined by the formula*

$$(\Lambda_\varepsilon z)(t) \stackrel{def}{=} col[\dot z(t) - A(t)(S_h z)(t) - \varepsilon B(t)(S_h z)(t), \quad lz] \tag{46}$$

is a Fredholm index zero operator with

$$ind\ \Lambda_\varepsilon = dimker\ \Lambda_\varepsilon - dimker\ \Lambda_\varepsilon^* = 0,$$

$$ind\ \Lambda_0 = dimker\ \Lambda_0 - dimker\ \Lambda_0^* = 0, \qquad (\ dimker\ \Lambda_0 = dimker\ \Lambda_0^* = r = d\ );$$

2) the homogeneous boundary value problem

$$\dot z(t) = A(t)(S_h z)(t) + \varepsilon B(t)(S_h z)(t), \quad lz = 0 \in R^n \tag{47}$$

has only trivial solutions $(dimker\ \Lambda_\varepsilon = 0, \ \ \varepsilon \neq 0\ )$;

3) the boundary value problem adjoint to (47) has only trivial solutions $(dimker\ \Lambda_\varepsilon^* = 0, \ \ \varepsilon \neq 0\ )$;

4) for arbitrary $\varphi(t) \in L_p^n[a,b]$ *and* $\alpha \in R^n$ *the boundary value problem (44) has the unique solution* $z(t,\varepsilon)$ *with the properties*

$$z(\cdot,\varepsilon) \in D_p^n[a,b], \ \dot z(\cdot,\varepsilon) \in L_p^n[a,b], \ z(t,\cdot) \in C(0,\varepsilon_*],$$

in the form of the series

$$z(t,\varepsilon) = \sum_{i=-1}^{\infty} \varepsilon^i \bar z_i(t,\bar c_i), \tag{48}$$

converging for $\varepsilon \in (0,\varepsilon_*]$, *where* ε_* *is as in Lemma and the coefficients* $\bar z_i(t,\bar c_i)$, $\bar c_i$ *can be determined from (37).*

Remark 3. If (41) does not hold, then in order to obtain sufficient conditions for existence of solutions of the boundary value problem (40) for arbitrary nonhomogeneities $\varphi(t) \in L_p^n[a,b]$ and $\alpha \in R^m$, the solution $z(t,\varepsilon)$ of the problem (40) is constructed in the form of series (22) with $i \leq -2$.

Remark 4. If

$$rank\ [B_0 = \int_a^b H(\tau)B(\tau)(S_h X_r)(\tau)\,d\tau] = d,$$

then the nonlinear boundary value problem with the measurable delays $h_i(t)$

$$\dot{z}(t) = \sum_{i=1}^{k} A_i(t)z(h_i(t)) + g(t) + \varepsilon \sum_{i=1}^{k} B_i(t)z(h_i(t)) + \varepsilon \sum_{i=1}^{k} R_i(z(h_i(t)), t, \varepsilon),$$

$$z(s) = \psi(s), \text{ if } s < a, \quad lz = \alpha \in R^m, \qquad t \in [a, b]$$

has at least one solution $z(t, \varepsilon)$ with the properties

$$z(\cdot, \varepsilon) \in D_p^n[a, b], \quad \dot{z}(\cdot, \varepsilon) \in L_p^n[a, b],$$

where

$$A_i(t), \; B_i(t), \; g(t), \; R_i(z, \cdot, \varepsilon) \in L_p^n[a, b], \; R_i(z, t, \varepsilon) = o(z^2).$$

5. APPLICATIONS.

Example 1. Consider the linear boundary value problem for the delay differential equation

$$\dot{z}(t) = \varepsilon \sum_{i=1}^{k} B_i(t)z(h_i(t)) + g(t), \quad t \in [0, T], \quad z(s) = \psi(s), \text{ if } s < 0, \tag{49}$$

$$z(0) = z(T),$$

where $B_i(t)$ are $(n \times n)$ - matrices, $B_i(t), g(t) \in L_p^n[0, T]$; $\psi : R^1 \setminus [a, b] \to R^n$; $h_i(t)$ are measurable functions . Using the symbols S_{h_i} and ψ^{h_i} (see (4),(5)), we arrive at the following operator system

$$\dot{z}(t) = \varepsilon B(t)(S_h z)(t) + \varphi(t), \quad lz = z(0) - z(T) = 0,$$

where $B(t) = (B_1(t), ..., B_k(t))$ is an $(n \times N)$ - matrix $(N = nk)$, and

$$\varphi(t) = g(t) + \sum_{i=1}^{k} B_i(t)\psi^{h_i}(t) \in L_p^n[0, T].$$

It is easily verify that

$$X(t) = E, \quad \dot{z}(t) = 0; \quad lX(\cdot) = Q = 0; P_Q = P_{Q^*} = E; (r = n, d = m = n);$$

$$K(t, \tau) = \begin{cases} E, 0 \le \tau \le t \le T; \\ 0, \qquad\qquad \tau > t; \end{cases} \quad lK(\cdot, \tau) = K(0, \tau) - K(T, \tau) = -E;$$

$$H(\tau) = P_{Q^*} lK(\cdot, \tau) = -E.$$

According to the representation (8), we have the following expressions

$$(S_{h_i}E)(t) = \chi_{h_i}(t, 0)E = E \begin{cases} 1, \text{ if } \; 0 \le h_i(t) \le T, \\ 0, \text{ if } \qquad h_i(t) < 0; \end{cases}$$

and

$$B_0 = -\int_0^T H(\tau)B(\tau)(S_h E)(\tau)d\tau =$$
$$= -\int_0^T \sum_{i=1}^k B_i(\tau)(S_{h_i}E)(\tau)d\tau = -\sum_{i=1}^k \int_0^T B_i(\tau)\chi_{h_i}(\tau,0)d\tau,$$

where B_0 is an $(n \times n)$ matrix.

If $det\ B_0 \neq 0$, then the problem (49) has the unique solution $z(t,\varepsilon)$ with the properties

$$z(\cdot,\varepsilon) \in D_p^n[0,T], \quad \dot z(\cdot,\varepsilon) \in L_p^n[0,T], \quad z(t,\cdot) \in C(0,\varepsilon_*]$$

for arbitrary $g(t) \in L_p^n[0,T]$, $\psi : R^1 \setminus [a,b] \to R^n$, and for measurable delays $h_i(t)$.

If, for example, $h_i(t) = t - \Delta_i$, where $0 < \Delta_i = const < T$, $(i = 1,...,k)$, then

$$\chi_{h_i}(t,0) = \begin{cases} 1, \text{ if } 0 \le h_i(t) = t - \Delta_i \le T, \\ 0, \text{ if } \quad h_i(t) = t - \Delta_i < 0, \end{cases} = \begin{cases} 1, \text{ if } \Delta_i \le t \le T + \Delta_i, \\ 0, \text{ if } \quad t < \Delta_i. \end{cases}$$

For this reason the $(n \times n)-$ matrix B_0 can be rewritten in the form

$$B_0 = -\int_0^T H(\tau)\sum_{i=1}^k B_i(\tau)\chi_{h_i}(\tau,0)d\tau =$$
$$= -\sum_{i=1}^k \int_0^T B_i(\tau)\chi_{h_i}(\tau,0)d\tau = -\sum_{i=1}^k \int_{\Delta_i}^T B_i(\tau)d\tau,$$

while the solvability condition of the boundary value problem (49) has the form

$$det\ [B_0 = -\sum_{i=1}^k \int_{\Delta_i}^T B_i(\tau)d\tau] \neq 0.$$

In the case, where there is no delay effect ($\Delta_i = 0$, $i = 1,...,k$), the last solvability condition coincide with such one of [2, 6].

Example 2. Consider the linear boundary value problem for the delay differential equation

$$\dot z(t) = \varepsilon \sum_{i=1}^k B_i(t)z(h_i(t)) + g(t), \quad t \in [0,T], \quad z(s) = \psi(s), \text{ if } s < 0, \tag{50}$$

$$lz := [1, \underbrace{0 \cdots 0}_{(n-1)\,\text{time}}]z(0) - [1, \underbrace{0 \cdots 0}_{(n-1)\,\text{time}}]x(T) = \alpha \in R, (m = 1).$$

Using the symbols S_{h_i} and ψ^{h_i}, we arrive at the following boundary value problem for operator system

$$\dot z(t) = \varepsilon B(t)(S_h z)(t) + \varphi(t),$$

$$lz := [1, \underbrace{0 \cdots 0}_{(n-1)\,\text{time}}]z(0) - [1, \underbrace{0 \cdots 0}_{(n-1)\,\text{time}}]z(T) = \alpha \in R, (m = 1),$$

where $B(t) = (B_1(t), ..., B_k(t))$ is an $(n \times N)$ - matrix $(N = nk)$, and

$$\varphi(t) = g(t) + \sum_{i=1}^{k} B_i(t)\psi^{h_i}(t) \in L_p^n[0, T].$$

It is easy to see that

$$X(t) = E, \quad \dot{z}(t) = 0; \quad lX(\cdot) = Q = [\underbrace{0 \cdots 0}_{n\,\text{time}}]; P_Q = E; P_{Q^*} = 1;$$

$$(rankQ = n_1 = 0, r = n, d = m - n_1 = 1);$$

$$K(t, \tau) = \begin{cases} E, 0 \leq \tau \leq t \leq T, \\ 0, \qquad \tau > t; \end{cases}$$

$$lK(\cdot, \tau) = [1, \underbrace{0 \cdots 0}_{(n-1)\,\text{time}}]K(0, \tau) - [1, \underbrace{0 \cdots 0}_{(n-1)\,\text{time}}]K(T, \tau) = -[1, \underbrace{0 \cdots 0}_{(n-1)\,\text{time}}];$$

$$H(\tau) = P_{Q^*}lK(\cdot, \tau) = -[1, \underbrace{0 \cdots 0}_{(n-1)\,\text{time}}].$$

According to the representation (8), we have the following expression

$$(S_{h_i}E)(t) = \chi_{h_i}(t, 0)E = E \begin{cases} 1, \text{ if } 0 \leq h_i(t) \leq T, \\ 0, \text{ if } \quad h_i(t) < 0. \end{cases}$$

In order to obtain the solvability conditions for the problem (50), it suffices to consider only the first row of the matrices

$$B_i(t) = \begin{pmatrix} b_{11}^{(i)}(t) & b_{12}^{(i)}(t) & * & * & * & b_{1n}^{(i)}(t) \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix}, \qquad (i = 1, ..., k).$$

Indeed, the $(1 \times n)$ matrix has the form

$$B_0 = -\int_0^T H(\tau)B(\tau)(S_hE)(\tau)d\tau = -\int_0^T H(\tau)\sum_{i=1}^{k} B_i(\tau)(S_{h_i}E)(\tau)d\tau =$$

$$= -\int_0^T H(\tau)\sum_{i=1}^{k} B_i(\tau)\chi_{h_i}(\tau, 0)d\tau = -[\sum_{i=1}^{k}\int_0^T b_{11}^{(i)}(\tau)\chi_{h_i}(\tau, 0)d\tau,$$
$$\sum_{i=1}^{k}\int_0^T b_{12}^{(i)}(\tau)\chi_{h_i}(\tau, 0)d\tau, \cdots, \sum_{i=1}^{k}\int_0^T b_{1n}^{(i)}(\tau)\chi_{h_i}(\tau, 0)d\tau].$$

If one of the inequalities

$$\sum_{i=1}^{k}\int_0^T b_{1j}^{(i)}(\tau)\chi_{h_i}(\tau, 0)d\tau \neq 0 \quad (j = 1, ..., n)$$

is true, then rank $B_0 = d = 1$, and for arbitrary $\varphi(t) \in L_p^n[a,b]$, $\alpha \in R$, and for measurable delays $h_i(t)$ the boundary value problem (50) has a $\rho = (n-1)$–parametric set of solutions $z(t, c_\rho, \varepsilon)$ with the properties

$$z(\cdot, c_\rho, \varepsilon) \in D_p^n[0,T], \quad \dot{z}(\cdot, c_\rho, \varepsilon) \in L_p^n[0,T], \quad z(t, c_\rho, \cdot) \in C(0, \varepsilon_*]$$

in the form of the series (43).

If, for example, $h_i(t) = t - \Delta_i$, where $0 < \Delta_i = const < T$, $(i = 1, ..., k)$, then

$$\chi_{h_i}(t,0) = \begin{cases} 1, \text{ if } & \Delta_i \le t \le T + \Delta_i, \\ 0, \text{ if } & t < \Delta_i. \end{cases}$$

For this reason the $(1 \times n)$ - dimensional matrix B_0 can be rewritten in the form

$$\begin{aligned} B_0 &= -\int_0^T H(\tau) \sum_{i=1}^k B_i(\tau) \chi_{h_i}(\tau, 0) d\tau = \\ &= -[\sum_{i=1}^k \int_{\Delta_i}^T b_{11}^{(i)}(\tau) d\tau, \ \sum_{i=1}^k \int_{\Delta_i}^T b_{12}^{(i)}(\tau) d\tau, \cdots, \sum_{i=1}^k \int_{\Delta_i}^T b_{1n}^{(i)}(\tau) d\tau \ ], \end{aligned}$$

and one of the solvability conditions of the problem (50) is of the form

$$\sum_{i=1}^k \int_{\Delta_i}^T b_{1j}^{(i)}(\tau) d\tau \neq 0 \quad (j = 1, ..., n).$$

Acknowledgments. The essential part of the present work was finished, while the first author was visiting the University of Ioannina during the winter of 2003, who would like to thank the Ministry of Economy and Finance of Hellenic Republik for providing the NATO Science Fellowship (Ref. No: DOO 850) and University of Ioannina for its hospitality.

References

[1] N.V.Azbelev, V.P.Maksimov and L.F.Rakhmatullina, *Introduction to the theory of functional differential equations*, Nauka, Moscow, 1991 (Russian).

[2] A.A.Boichuk, V.F.Zhuravlev and A.M.Samoilenko, *Generalized inverse operators and Noether boundary-value problems*, Inst. Math. Nat. Acad. Sci. Ukraine, Kyiv, 1995 (Russian).

[3] J.K.Hale, *Theory of Functional Differential Equations*, Springer-Verlag, New York, 1975.

[4] P.R.Halmoc, *Measure Theory*, D.Van Nostrand, Princeton, 1950.

[5] T.Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1966.

[6] I.T.Kiguradze, *Boundary value problems for the systems of ordinary differential equations*, Itogi Nauki i Tekhniki, **30**, (1987), 3-103 (Russian).

[7] S.G.Krein, *Linear Equations in Banach Space*, Nauka, Moscow, 1971 (Russian).

[8] A.D.Myshkis, *Linear Differential Equations with Retarded Argument*, Nauka, Moscow, 1975 (Russian).

[9] C.R.Rao and S.K.Mitra, *Generalized Inverse of Matrices and Its Applications*, Wiley, New York, 1971.

[10] M.I.Vishik and L.A.Lyusternik, *Solvability of some problems concerning perturbations in case of matrices and selfadjoint and non-selfadjoint differential equations*, Uspekhi Mat,Nauk. **15**, (1960),no. 3, 3-80 (Russian).

Alexander A. Boichuk:

The National Academy of Science of Ukraine

Institute of Mathematics, 3 Tereshchenkivs'ka Str., 016 01 Kyiv, Ukraine

and

Department of Applied Mathematics, Faculty of Science

University of Zilina, J.M.Hurbana 15, Slovak Republic

E-mail: **boichuk@imath.kiev.ua**

Myron K. Grammatikopoulos:

Department of Mathematics, University of Ioannina,
451 10 Ioannina, Hellas

E-mail: **mgrammat@cc.uoi.gr**

Symmetries and similarity solutions: An application to fluid mechanics

Vassilios K. Kalpakides and Konstantinos G. Balassas

Department of Mathematics, University of Ioannina, Ioannina, GR-45110, Greece

Abstract

The free convective boundary–layer problem due to the motion of an elastic surface into an electrically conducting fluid is studied with group–theoretical methods. The symmetry groups admitted by the corresponding boundary value problem are obtained. Particular attention is paid on the group of scaling which provides the similarity solution of the problem. Also, the admissible form of the data, in order to be conformed to the obtained symmetries, is provided. Finally, with the use of the entailed similarity solution the problem is transformed to a boundary value problem of ODEs and is solved numerically.

Keywords: Symmetry groups; Similarity solutions; Infinitesimal generator; Laminar boundary-layer flow

1 Introduction

In this paper, we apply the so–called symmetry methods for a particular problem of fluid mechanics. The main advantage of such methods is that they can successfully be applied to non–linear differential equations. The symmetries of a differential equation are those continuous groups of transformations under which the differential equation remains invariant, that is, a symmetry group maps any solution to another solution. The interesting point is that, having obtained the symmetries of a specific problem, one can proceed further to find out the group–invariant solutions, which, in the case of the scaling group of transformations, are nothing but the well–known similarity solutions. The similarity solutions are quite popular because they result in the reduction of the independent variables of the problem. In our case, the problem under investigation is two–dimensional. Hence, any similarity solution will transform the system of PDEs to a system of ODEs.

To obtain a symmetry of a differential equation is equivalent to the determination of the infinitesimal generator of the transformation group associated with this symmetry. In [1, 2, 3, 4], one can find the general theory of Lie groups as well as the implied methods for determining the infinitesimal generator components. An alternative way being based on exterior calculus for determining of the infinitesimal generator can be found in the book of Edelen [5]. It is worth to note that there is an extensive literature where the methods arising from exterior calculus are used to attack symmetry problems of continuum mechanics [6, 7, 8, 9, 10, 11, 12].

Most of the researchers in the field of fluid mechanics usually try to obtain the similarity solutions by introducing a general similarity transformation with unknown parameters into the differential equation obtaining in this way an algebraic system. Then, the solution of this system, if exists, determines the values of the unknown parameters. In our opinion, it is better to attack any problem of similarity solutions from the outset, i.e, to find out the full list of the symmetries of the problem and then to study which of them are appropriate to provide group–invariant (or more specifically similarity solutions).

We apply this procedure to a boundary layer problem which arises from the motion of an elastic surface into an electrically conducting, incompressible viscous fluid. Particular variants of this problem have been studied by a numerous of researchers since 1961. We mention here the work of [13, 14, 15, 16, 17, 18, 19, 20]. It is remarkable that all of them have used the above described heuristic method to obtain the similarity transformations and the associated similarity solutions of the problem. That is, assuming particular boundary conditions and considering a particular form of the magnetic field[2], they try to fit a similarity solution in these data.

We set the problem on another base. First, we do not guess any kind of probable symmetry. The question of any possible symmetry for the system of PDEs is examined generally. In the same spirit, we do not make any assumption about the data of the problem. We consider the most general form for the boundary conditions and the magnetic field function involved in the system. Both the form of the functions on the boundaries and the form of magneting field arise as a consequence of the requirement to respect the obtained symmetries.

Next, having established the admissible symmetries of the boundary value problem, we proceed to the determination of the similarity solutions which, in turn, are used to tranform the system to a two point boundary value problem of ODEs. Finally, the reduced problem is solved numerically and its solutions are depicted for different values of the physical parameters.

2 Preliminaries

In this section, we give some preliminary notions necessary for the following analysis. Our analysis is based on the application of Lie group theory to differential equations. Any Lie group of transformation is related with the so–called infinitesimal generator, which can reproduce the finite group.

Definition 1. The infinitesimal generator of the one-parameter Lie group of transformation

$$\mathbf{x}^* = \mathbf{X}(\mathbf{x}; \varepsilon) \tag{2.1}$$

is the vector field

$$\mathbf{V} = \mathbf{V}(\mathbf{x}) = \xi(\mathbf{x}) \cdot \nabla = \sum_{i=1}^{n} \xi_i(\mathbf{x}) \frac{\partial}{\partial x_i}. \tag{2.2}$$

[2]For instance, in [17, 19] constant magnetic field is assumed, while in [20] a magnetic field depended on x is considered.

The components ξ_i provide the infinitesimal part of the transformation group, that is eq. (2.1) can be written as

$$\mathbf{x}^* = \mathbf{x} + \varepsilon\xi(\mathbf{x}) + o(\varepsilon) \tag{2.3}$$

Let $F(\mathbf{x}) = F(x_1, x_2, \ldots, x_n)$ be any differentiable function, and let the vector field $\mathbf{V}$ represents a transformation group. The differential operator $\mathbf{V}$ acts on the function F as follows

$$\mathbf{V}F(\mathbf{x}) = \xi(\mathbf{x}) \cdot \nabla F(\mathbf{x}) = \sum_{i=1}^{n} \xi_i(\mathbf{x}) \frac{\partial F(\mathbf{x})}{\partial x_i}. \tag{2.4}$$

If $\mathbf{V}F \equiv 0$, we will say that F is invariant under the action of the associated transformation group.
We are mainly interested in the notion of invariance of a PDE under the action of some group of transformation. Thus, the space over which the transformation group will act is made up by the dependent variables $\mathbf{u}$ and independent ones $\mathbf{x}$ involved in the differential equation. In that case, a typical transformation group has the form

$$\mathbf{x}^* = \mathbf{X}(\mathbf{x}, \mathbf{u}; \varepsilon), \tag{2.5}$$
$$\mathbf{u}^* = \mathbf{U}(\mathbf{x}, \mathbf{u}; \varepsilon). \tag{2.6}$$

The corresponding infinitesimal generator is denoted by

$$\mathbf{V} = \xi_i(x_j, u^\mu)\frac{\partial}{\partial x_i} + \eta^l(x_j, u^\mu)\frac{\partial}{\partial u^l}, \quad i, j = 1, \ldots, n, \quad l, \mu = 1, \ldots, m, \tag{2.7}$$

where n and m are the number of independent and dependent variables, respectively.
Consider now a *kth* order partial differential equation defined on a domain $\Omega_{\mathbf{x}}$ in $\mathbf{x}$-space

$$\Delta(\mathbf{x}, \mathbf{u}, \mathbf{u}^{[1]}, ..., \mathbf{u}^{[k]}) = 0, \tag{2.8}$$

where $\mathbf{u}^{[l]}$ denotes any l-order partial derivative of $\mathbf{u}$ with respect to $\mathbf{x}$. Noting that the differential operator Δ admits an enlarged argument containing in addition the partial derivatives of $\mathbf{u}$, we must also enlarge the infinitesimal generator in the following way [1, 2]

$$\begin{aligned} \mathbf{V} = \xi_i(\mathbf{x}, \mathbf{u})\frac{\partial}{\partial x_i} + \eta^l(\mathbf{x}, \mathbf{u})\frac{\partial}{\partial u^l} + \eta_i^{\mu(1)}(\mathbf{x}, \mathbf{u}, \mathbf{u}^{[1]})\frac{\partial}{\partial u_i^\mu} \\ + \cdots + \eta_{i_1 i_2 \ldots i_k}^{\mu(k)}(\mathbf{x}, \mathbf{u}, \mathbf{u}^{[1]}, \ldots, \mathbf{u}^{[k]})\frac{\partial}{\partial u_{i_1 i_2 \ldots i_k}^\mu}, \end{aligned} \tag{2.9}$$

where by $u_{i_1 i_2 \ldots i_k}^\mu$ is denoted the mixed partial derivative of u^μ with respect to $x_{i_1}, \ldots, x_{i_k}$ and $\eta_{i_1 i_2 \ldots i_k}^{\mu(k)}$ denotes the additional components of the infinitesimal generator given by the recursion formula

$$\begin{gathered} \eta_{i_1 \ldots i_l}^{\mu(l)} = D_{i_l}\eta_{i_1 \ldots i_{l-1}}^{\mu(l-1)} - (D_{i_l}\xi_j)u_{i_1 \ldots i_{l-1} j}^\mu, \\ 1 \le l \le k, \; 1 \le i_l \le n, \; 1 \le \mu \le m. \end{gathered} \tag{2.10}$$

Let us now associate with the PDE (2.8) the boundary conditions

$$B_\alpha(\mathbf{x}, \mathbf{u}, \mathbf{u}^{[1]}, ..., \mathbf{u}^{[k-1]}) = 0 \tag{2.11}$$

prescribed on the boundary surfaces

$$\omega_\alpha(\mathbf{x}) = 0, \quad a = 1, 2, ..., s. \tag{2.12}$$

We are ready to give the main definition concerning the invariance of a boundary value problem under the action of a transformation group.

Definition 2. The one-parameter Lie group of transformation given by eqs. (2.5)–(2.6) constitutes a symmetry group for the boundary value problem described by eqs. (2.8) and (2.11)–(2.12) if and only if

$$(i)\ \mathbf{V}^{(k)}\Delta(\mathbf{x}, \mathbf{u}, \mathbf{u}^{[1]}, ..., \mathbf{u}^{[k]}) = 0, \quad \text{when} \quad \Delta(\mathbf{x}, \mathbf{u}, \mathbf{u}^{[1]}, ..., \mathbf{u}^{[k]}) = 0, \tag{2.13}$$

$$(ii)\ \mathbf{V}\omega_\alpha(\mathbf{x}) = 0 \quad \text{when} \quad \omega_\alpha(\mathbf{x}) = 0, \tag{2.14}$$

$$(iii)\ \mathbf{V}^{(k-1)}B_\alpha(\mathbf{x}, \mathbf{u}, \mathbf{u}^{[1]}, ..., \mathbf{u}^{[k-1]}) = 0,$$
$$\text{when} \quad B_\alpha(\mathbf{x}, \mathbf{u}, \mathbf{u}^{[1]}, ..., \mathbf{u}^{[k-1]}) = 0 \quad \text{on} \quad \omega_\alpha(\mathbf{x}) = 0. \tag{2.15}$$

3 The mathematical description of the problem

We consider a free convective, laminar boundary-layer flow of an electrically conducting incompressible viscous fluid over a vertical porous and elastic surface. The surface is stretched vertically upward along the positive x-axis, with a prescribed velocity

$$u(x, y = 0) = u_0(x), \tag{3.1}$$

while the origin $(x, y) = (0, 0)$ is kept fixed. The y–axis is vertical to the surface, as it is depicted[3] in Fig. 1. Also, due to the fact that the elastic surface is porous, there is a component of the velocity of the fluid which has vertical direction to the surface given by

$$v(x, y = 0) = v_0(x). \tag{3.2}$$

The motion of the surface within the fluid creates a boundary layer, which is extended along the x-axis. All the system is under the influence of a magnetic field $B = B(x, y)$, which applies to the y–direction. We consider that the temperature of the surface changes along the x-axis and its distribution is described by a given function $T_0(x)$.

Under the assumption that the viscous dissipation term in the energy equation and the induced magnetic field can be neglected, the basic boundary layer equations of

[3] Although the problem is set in the two dimensional space of x and y, the expression "the surface $y = 0$" is used instead of the more accurate "the curve $y = 0$".

the mass, momentum and energy for the steady flow of Boussinesq type are respectively as follows:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0, \tag{3.3}$$

$$u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=\nu\frac{\partial^2 u}{\partial y^2}-\frac{\sigma u B^2}{\rho}+g\beta(T-T_\infty), \tag{3.4}$$

$$u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}=a\frac{\partial^2 T}{\partial y^2}, \tag{3.5}$$

where σ is the electric conductivity, β is the volumetric coefficient of thermal expansion, ν is the kinematic viscosity, ρ is the mass density and a is the thermal diffusivity, which are assumed to be constants. Also, g is the gravity field assumed to be parallel to the $x-$axis, $T=T(x,y)$ is the temperature field and T_∞ is the temperature at infinity. According to the above description, the boundary conditions of the problem should be of the form

$$y=0: \quad \begin{array}{l} u(x,0)=u_0(x), \\ v(x,0)=v_0(x), \quad x\succ 0 \\ \theta(x,0)=\theta_0(x). \end{array} \tag{3.6}$$

$$y\to\infty: \quad u\to 0, \quad \theta\to 0, \quad x\succ 0, \tag{3.7}$$

where $\theta=T-T_\infty$. Also, $\theta_0=T_0-T_\infty$ is a prescribed function along the boundary surface $y=0$.

We introduce now the stream function ψ, which is related to the components of the velocity field by the equations

$$u=\frac{\partial\psi}{\partial y}, \quad v=-\frac{\partial\psi}{\partial x}. \tag{3.8}$$

Inserting eqs. (3.8) into the field equations (3.3)–(3.5), we obtain

$$\psi_y\psi_{xy}-\psi_x\psi_{yy}-\nu\psi_{yyy}+\kappa B^2\psi_y-\phi\theta=0, \tag{3.9}$$

$$\psi_y\theta_x-\psi_x\theta_y-a\theta_{yy}=0, \tag{3.10}$$

where $\kappa=\sigma/\rho$ and $\phi=g\beta$.

The associated boundary conditions can be written as

$$y=0: \quad \begin{array}{l} \psi_y(x,0)=u_0(x)\equiv\Psi_1(x), \\ -\psi_x(x,0)=v_0(x)\equiv\Psi_2(x), \quad x\succ 0 \\ \theta(x,0)=\theta_0(x). \end{array} \tag{3.11}$$

$$y\to\infty: \quad \psi_y\to 0, \quad \theta\to 0, \quad x\succ 0. \tag{3.12}$$

4 Determination of the symmetry groups

In this section, we will look for any possible symmetry group of the boundary value problem described by PDEs (3.9)–(3.10) and boundary conditions (3.11)–(3.12).

The infinitesimal generator associated with the possible symmetries for the system under study has the general form[1, 2]:

$$\mathbf{V} = \xi_1(x,y,\psi,\theta)\frac{\partial}{\partial x} + \xi_2(x,y,\psi,\theta)\frac{\partial}{\partial y} + \eta^1(x,y,\psi,\theta)\frac{\partial}{\partial \psi} + \eta^2(x,y,\psi,\theta)\frac{\partial}{\partial \theta}. \qquad (4.1)$$

Due to the order of the PDEs of our system, we need the third extension of (4.1). This is given by

$$\begin{aligned}\mathbf{V}^{(3)} = \mathbf{V} &+ \eta_1^{1(1)}\frac{\partial}{\partial \psi_x} + \eta_2^{1(1)}\frac{\partial}{\partial \psi_y} + \eta_1^{2(1)}\frac{\partial}{\partial \theta_x} + \eta_2^{2(1)}\frac{\partial}{\partial \theta_y} + \eta_{11}^{1(2)}\frac{\partial}{\partial \psi_{xx}} \\ &+\eta_{12}^{1(2)}\frac{\partial}{\partial \psi_{xy}} + \eta_{22}^{1(2)}\frac{\partial}{\partial \psi_{yy}} + \eta_{11}^{2(2)}\frac{\partial}{\partial \theta_{xx}} + \eta_{12}^{2(2)}\frac{\partial}{\partial \theta_{xy}} + \eta_{22}^{2(2)}\frac{\partial}{\partial \theta_{yy}} \\ &+\eta_{111}^{1(3)}\frac{\partial}{\partial \psi_{xxx}} + \eta_{112}^{1(3)}\frac{\partial}{\partial \psi_{xxy}} + \eta_{122}^{1(3)}\frac{\partial}{\partial \psi_{xyy}} + \eta_{222}^{1(3)}\frac{\partial}{\partial \psi_{yyy}} + \\ &\eta_{111}^{2(3)}\frac{\partial}{\partial \theta_{xxx}} + \eta_{112}^{2(3)}\frac{\partial}{\partial \theta_{xxy}} + \eta_{122}^{2(3)}\frac{\partial}{\partial \theta_{xyy}} + \eta_{222}^{2(3)}\frac{\partial}{\partial \theta_{yyy}}, \end{aligned} \qquad (4.2)$$

where the components $\eta^{\gamma(1)}$, $\eta^{\gamma(2)}$ and $\eta^{\gamma(3)}$ $(\gamma = 1,2)$ depend on the quantities $(x,y,\psi,\theta,\psi^{[1]},\theta^{[1]})$, $(x,y,\psi,\theta,\ \psi^{[1]},\psi^{[2]},\theta^{[1]},\theta^{[2]})$ and $(x,y,\psi,\theta,\ \psi^{[1]},\psi^{[2]},\psi^{[3]},$ $\theta^{[1]},\theta^{[2]},\theta^{[3]})$, respectively. The general form of any component of the extended infinitesimal generator is given by the following relation [2]:

$$\eta^{\mu(l)}_{i_1\dots i_l} = D_{i_l}\eta^{\mu(l-1)}_{i_1\dots i_{l-1}} - (D_{i_l}\xi_j)u^{\mu}_{i_1\dots i_{l-1}j}, \quad 1 \le l \le 3,\ 1 \le i_l, \mu \le 2. \qquad (4.3)$$

In the last formula, we used the denotation $u^1 := \psi$ and $u^2 := \theta$. Also, we must note that $u^{[l]}$ denotes any l-order partial derivative of u with respect to x and y. It is important to note that it is not necessary to compute all the components which appear in eq. (4.2). Having in mind the form of the PDEs of our system, we easily conclude that the components which will survive after the application of the infinitesimal generator on the system will be:

$$\eta_1^{1(1)},\ \eta_2^{1(1)},\ \eta_1^{2(1)},\ \eta_2^{2(1)},\ \eta_{12}^{1(2)},\ \eta_{22}^{2(2)},\ \eta_{22}^{1(2)},\ \eta_{222}^{1(3)}.$$

Hence, in order to develop the prescribed method, we must first calculate these components. One can easily carry out such calculation using the formula (4.3).

Let us proceed further by applying the invariance requirement on the system given by eqs. (3.9) –(3.10). The condition (2.13) will take the form:

$$\mathbf{V}^{(3)}(\psi_y\psi_{xy} - \psi_x\psi_{yy} - \nu\psi_{yyy} + kB^2\psi_y - \phi\theta) = 0, \qquad (4.4)$$

when eqs. (3.9)–(3.10) hold

and

$$\mathbf{V}^{(3)}(\psi_y\theta_x - \psi_x\theta_y - a\theta_{yy}) = 0, \qquad (4.5)$$

when eqs. (3.9)–(3.10) hold.

Developing eqs. (4.4)–(4.5), one can straightforwardly take

$$2\xi_1\kappa BB_x\psi_y + 2\xi_2\kappa BB_y\psi_y - \phi\eta^2 - \eta_1^{1(1)}\psi_{yy} + \eta_2^{1(1)}\psi_{xy} + \kappa B^2\eta_2^{1(1)} + \\ +\eta_{12}^{1(2)}\psi_y - \eta_{22}^{1(2)}\psi_x - \nu\eta_{222}^{1(3)} = 0, \quad (4.6)$$

when the equations (3.9), (3.10)) hold

and

$$-\eta_1^{1(1)}\theta_y + \eta_2^{1(1)}\theta_x + \eta_1^{2(1)}\psi_y - \eta_2^{2(1)}\psi_x - a\eta_{22}^{2(2)} = 0, \quad (4.7)$$

when the equations (3.9), (3.10) hold.

Let us now explore the conditions given by eqs. (4.6) and (4.7). Inserting eqs. (3.9) and (3.10) into both eqs. (4.6) and (4.7) as well as the expressions of all η's given by eq. (4.3), we will obtain two identically zero polynomials in terms of various combinations of the derivatives of ψ and θ. Thus, their coefficients, which will contain ξ_1, ξ_2, η^1 and η^2 and their derivatives, should be zero. This procedure provides the so–called determining equations, i.e., PDEs in ξ_1, ξ_2, η^1 and η^2, the solutions of which determine the components of the infinitesimal generator. In our problem, the determining equations associated with the conditions (4.6) and (4.7) are given by

$$\xi_{1,\theta} = 0, \quad \xi_{1,y} = 0, \quad \xi_{1,\psi} = 0, \quad \xi_{2,\theta} = 0, \quad \xi_{2,\psi} = 0, \quad (4.8)$$

$$\xi_{2,yy} = 0, \quad \eta_x^1 = 0, \quad \eta_y^1 = 0, \quad -\phi\eta^2 + \eta_\psi^1\phi\theta - 3\phi\theta\xi_{2,y} = 0, \quad (4.9)$$

$$\eta_\theta^1 = 0, \quad \eta_{\psi\psi}^1 = 0, \quad \xi_{2,xy} = 0, \quad \eta_\psi^1 - \xi_{1,x} + \xi_{2,y} = 0, \quad (4.10)$$

$$2\kappa B(x,y)B_x(x,y)\xi_1 + 2\kappa B(x,y)B_y(x,y)\xi_2 + 2\kappa B^2(x,y)\xi_{2,y} = 0, \quad (4.11)$$

$$\eta_y^2 = 0, \quad \eta_x^2 = 0, \quad \eta_\psi^2 = 0, \quad \eta_{\theta\theta}^2 = 0. \quad (4.12)$$

One can confirm[4] that the solutions of the system (4.8)–(4.12) are given as follows:

$$\begin{aligned} \xi_1 &= \kappa_1 x + \kappa_1', \\ \xi_2 &= A(x) + sy + \kappa_2', \\ \eta^1 &= (\kappa_1 - s)\psi + \kappa_3', \\ \eta^2 &= (\kappa_1 - 4s)\theta, \end{aligned} \quad (4.13)$$

where $\kappa_1, s, \kappa_1', \kappa_3'$ are arbitrary constants and A is an arbitrary function.
Furthermore, we have to apply the invariance condition to the boundary conditions [1, 2]; that is, on the relation which should be held along the boundary surfaces and on the boundary surfaces themselves. We start with the latter requirement; We can write the two boundary surfaces of the problem under study as follows

$$\omega_1(x,y) = y, \qquad \omega_2(x,y) = y - K, \quad K \in I\!R^+.$$

Thus, the invariance requirement for the former takes the form

$$\mathbf{V}[\omega_1] = 0 \;\; \text{when} \;\; \omega_1 = 0 \Rightarrow \xi_2(x,0) = 0. \quad (4.14)$$

[4]We remark that the system (4.8)–(4.12) is overdetermined, so one can obtain its solution by simple integrations.

The latter fulfils the invariance requirement identically, thus, it does not impose any constraint on the generator.
Inserting eq. (4.14) into eqs. (4.13), we obtain the final form for the components of the infinitesimal generator

$$\begin{aligned} \xi_1 &= \kappa_1 x + \kappa_1^{'}, \\ \xi_2 &= sy, \\ \eta^1 &= (\kappa_1 - s)\psi + \kappa_3^{'}, \\ \eta^2 &= (\kappa_1 - 4s)\theta. \end{aligned} \tag{4.15}$$

What it remains is to require invariance of the data which must be held on the boundary surfaces. This requirement means

$$\begin{aligned} &\mathbf{V}^{(1)}[\psi_y - \Psi_1(x)] = 0 \quad \text{when} \quad \psi_y(x,0) = \Psi_1(x), \\ &\mathbf{V}^{(1)}[-\psi_x - \Psi_2(x)] = 0 \quad \text{when} \quad \psi_x(x,0) = -\Psi_2(x), \\ &\mathbf{V}^{(1)}[\theta - \theta_0(x)] = 0 \quad \text{when} \quad \theta(x,0) = \theta_0(x). \end{aligned} \tag{4.16}$$

Examining the above conditions, we obtain the following differential equations

$$\begin{aligned} &(\kappa_1 x + \kappa_1^{'})\Psi_1^{'} - (\kappa_1 - 2s)\Psi_1 = 0, \\ &(\kappa_1 x + \kappa_1^{'})\Psi_2^{'} + s\Psi_2 = 0, \\ &(\kappa_1 x + \kappa_1^{'})\theta_0^{'} - (\kappa_1 - 4s)\theta_0 = 0, \end{aligned}$$

which directly give the admissible form for the functions Ψ_1, Ψ_2 and θ_0

$$\begin{aligned} \Psi_1(x) &= c_1|\kappa_1 x + \kappa_1^{'}|^{\frac{\kappa_1 - 2s}{\kappa_1}} \\ \Psi_2(x) &= c_2|\kappa_1 x + \kappa_1^{'}|^{-\frac{s}{\kappa_1}}, \\ \theta_0(x) &= c_3|\kappa_1 x + \kappa_1^{'}|^{\frac{\kappa_1 - 4s}{\kappa_1}}. \end{aligned} \tag{4.17}$$

Consequently, a set of boundary conditions conformed to the symmetries (4.15) should be of the form

$$\begin{aligned} y = 0: \qquad &\psi_y(x,0) = c_1|\kappa_1 x + \kappa_1^{'}|^{\frac{\kappa_1 - 2s}{\kappa_1}}, \\ &\psi_x(x,0) = -c_2|\kappa_1 x + \kappa_1^{'}|^{-\frac{s}{\kappa_1}}, \quad x \succ 0 \\ &\theta(x,0) = c_3|\kappa_1 x + \kappa_1^{'}|^{\frac{\kappa_1 - 4s}{\kappa_1}}. \end{aligned} \tag{4.18}$$

and

$$y \to \infty: \qquad \psi_y \to 0, \quad \theta \to 0, \quad x \succ 0. \tag{4.19}$$

It is interesting to note that the system of determining equations provides a constraint for the magnetic field function, as well. Indeed, eq. (4.11) takes the form

$$(\kappa_1 x + \kappa_1^{'})B_x + syB_y + sB = 0.$$

Hence, B is constrained to be of the form

$$B(x,y) = \frac{1}{y} G(\frac{|\kappa_1 x + \kappa_1'|^s}{y^{\kappa_1}}), \tag{4.20}$$

where G is an arbitrary function.

Concluding, the following statement has been proved.

Proposition. *The boundary value problem described by eqs.* (3.9)–(3.10) *and the data* (3.11)–(3.12), *admits the following multi–parameter group of symmetries*

$$\begin{aligned}
x^* &= x + \varepsilon(\kappa_1 x + \kappa_1'), \\
y^* &= y + \varepsilon s y, \\
\psi^* &= \psi + \varepsilon((\kappa_1 - s)\psi + \kappa_3'), \\
\theta^* &= \theta + \varepsilon((\kappa_1 - 4s)\theta).
\end{aligned} \tag{4.21}$$

Moreover, the admissible form of the data[5] *on the boundaries should be of the form given by eq.* (4.17) *and the magnetic field function is constrained to be of the form given by eq.* (4.20).

Looking at the transformation equations (4.21), one can recognize two kind of symmetries. Vanishing the parameters κ_1' and κ_3', the scaling group parameterized by κ_1 and s arises. On the other hand, vanishing κ_1 and s, the group of translations with respect to x and ψ is obtained.

5 Group–invariant solutions

The next question is whether the symmetry group we have obtained in the last section gives any of the so–called group–invariant solutions. A group–invariant solution is nothing else but a solution of the BVP (3.9)–(3.12), which is also invariant under the group (4.21) [1, 2]. Suppose (ψ, θ) is a solution of the problem (3.9)–(3.12). In order this solution to be invariant under the transformation group (4.21), the following system of partial differential equations must hold

$$\xi_1 \frac{\partial \psi}{\partial x} + \xi_2 \frac{\partial \psi}{\partial y} = \eta^1, \quad \xi_1 \frac{\partial \theta}{\partial x} + \xi_2 \frac{\partial \theta}{\partial y} = \eta^2$$

or

$$-(\kappa_1 x + \kappa_1')\frac{\partial \psi}{\partial x} - sy\frac{\partial \psi}{\partial y} + (\kappa_1 - s)\psi = -\kappa_3', \tag{5.1}$$

$$-(\kappa_1 x + \kappa_1')\frac{\partial \theta}{\partial x} - sy\frac{\partial \theta}{\partial y} + (\kappa_1 - 4s)\theta = 0. \tag{5.2}$$

[5] In the sense that they respect the obtained symmetries and so they do not destroy the symmetry of the boundary value problem

Using the method of characteristics, we can solve the system (5.1)–(5.2)

$$\psi(x,y) = |\kappa_1 x + \kappa_1^{'}|^{\frac{\kappa_1 - s}{\kappa_1}} F(X_1) - \frac{k_3^{'}}{k_3}, \tag{5.3}$$

$$\theta(x,y) = |\kappa_1 x + \kappa_1^{'}|^{\frac{\kappa_1 - 4s}{\kappa_1}} H(X_1), \tag{5.4}$$

where F, H are arbitrary functions and X_1 is the similarity variable given by the relation

$$X_1 = \frac{|\kappa_1 x + \kappa_1^{'}|^s}{y^{\kappa_1}}. \tag{5.5}$$

Equations (5.3)–(5.4) give the general form for any group–invariant solution of our problem. The interesting point here is that such a solution has the property to reduce the number of the independent variables of the problem. Thus, inserting the solution (5.3)–(5.4) and the admissible form of the magnetic field (4.20) into the field equations (3.9)–(3.10), we obtain

$$\begin{aligned}
-\kappa_1^2(\kappa_1 - 2s) X_1^{\frac{2\kappa_1+2}{\kappa_1}} F^{'2} - \kappa_1(\kappa_1 - s)(\kappa_1+1) X_1^{\frac{\kappa_1+2}{\kappa_1}} F F^{'} + \\
\nu \left[\kappa_1(\kappa_1+1)(\kappa_1+2) X_1^{\frac{\kappa_1+3}{\kappa_1}} F^{'} + \kappa_1^2 (3\kappa_1 + 3) X_1^{\frac{2\kappa_1+3}{\kappa_1}} F^{''} + \kappa_1^3 X_1^{\frac{3\kappa_1+3}{\kappa_1}} F^{'''} \right] - \\
\kappa_1 \kappa G^2 X_1^{\frac{\kappa_1+3}{\kappa_1}} F^{'} + \kappa_1^2(\kappa_1 - s) X_1^{\frac{2\kappa_1+2}{\kappa_1}} F F^{''} - \phi H = 0,
\end{aligned} \tag{5.6}$$

$$\begin{aligned}
-(\kappa_1 - 4s) X_1^{\frac{\kappa_1+1}{\kappa_1}} F^{'} H - \kappa_1(\kappa_1 - s) X_1^{\frac{\kappa_1+1}{\kappa_1}} F H^{'} - \alpha \Big[\kappa_1(\kappa_1+1) X_1^{\frac{\kappa_1+2}{\kappa_1}} H^{'} + \\
\kappa_1^2 X_1^{\frac{2\kappa_1+2}{\kappa_1}} H^{''} \Big] = 0,
\end{aligned} \tag{5.7}$$

Also, inserting eqs. (5.3)–(5.4) into the boundary conditions (4.18)–(4.19), we take the following boundary conditions

$$F(0) = -\frac{c_2}{\kappa_1 - s}, \quad F^{'}(0) = -\frac{c_1}{\kappa_1} \lim_{X_1 \to 0+} \left[X_1^{-\frac{\kappa_1+1}{\kappa_1}} \right], \quad H(0) = c_3, \tag{5.8}$$

$$\lim_{X_1 \to \infty} F^{'}(X_1) = 0, \quad \lim_{X_1 \to \infty} H(X_1) = 0, \tag{5.9}$$

where $\kappa_1 \in [-1, 0)$ and $\kappa_1 \neq s$.

Eqs. (5.6)–(5.9) describe the new form of our problem. Thus, the initial boundary value problem of PDEs has been transformed to a boundary value problem of ODEs which is generally easier to be solved by some numerical method.

6 Numerical results for the scaling symmetry

Proceeding further to numerical results, we are confined to the case of scaling symmetry, consequently we choose $\kappa_1^{'} = \kappa_3^{'} = 0$. Furthermore, we examine two distinct cases. The first one for the case $\kappa_1 = -1$, corresponding to the problem already studied by [20] and the second one for the case $\kappa_1 = -1/2$, which corresponds to a magnetic field depended on both variables x and y.

Case 1 $\kappa_1 = -1$
Moreover, in order to produce results comparable to [20], we set $s = \frac{m-1}{2}$, where m is an arbitrary parameter. Thus, we examine the case in which the transformation equations are of the following form:

$$\begin{aligned} x^* &= e^{-\varepsilon} x, \\ y^* &= e^{\varepsilon s} y, \\ \psi^* &= e^{-(1+s)\varepsilon} \psi, \\ \theta^* &= e^{-(1+4s)\varepsilon} \theta. \end{aligned} \tag{6.1}$$

Also, under this choice of the parameters, the similarity solutions (5.3)–(5.4) take the form

$$\psi(x,y) = x^{\frac{m+1}{2}} F(X_1), \tag{6.2}$$

$$\theta(x,y) = x^{2m-1} H(X_1), \tag{6.3}$$

while the similarity variable (5.5) reduces to

$$X_1 = y x^{\frac{m-1}{2}}. \tag{6.4}$$

As far as the magnetic field is concerned, we choose the function G to be of the form

$$G(X_1) = C X_1,$$

where C an arbitrary non–vanishing constant. The above equation with the aid of eq. (4.20) leads to a magnetic field depended only on x

$$B(x,y) = B_0 x^{\frac{m-1}{2}}, \tag{6.5}$$

where B_0 is a constant.
Also, the boundary conditions associated with this choice of parameters become

$$\begin{aligned} y = 0: \qquad & \psi_y(x,0) = c_1 x^m, \\ & \psi_x(x,0) = -c_2 x^{\frac{m-1}{2}}, \quad x \succ 0 \\ & \theta(x,0) = c_3 x^{2m-1} \end{aligned} \tag{6.6}$$

$$y \to \infty: \qquad \psi_y \to 0, \quad \theta \to 0, \quad x \succ 0. \tag{6.7}$$

Finally, the reduced system of ODEs corresponding to the above problem is given by the equations:

$$\nu F''' + \frac{m+1}{2} F F'' - m F'^2 - M F' + \phi H = 0, \tag{6.8}$$

$$\frac{1}{Pr} H'' + \left(\frac{m+1}{2\nu}\right) F H' - \left(\frac{2m-1}{\nu}\right) F' H = 0, \tag{6.9}$$

where $M = \kappa B_0^2 = \sigma B_0^2/\rho$.
Also, the associated boundary conditions take the form

$$F(0) = -\frac{2c_2}{m+1} = F_m, \quad F^{'}(0) = c_1, \quad H(0) = c_3, \tag{6.10}$$

$$\lim_{X_1 \to \infty} F^{'}(X_1) = 0, \quad \lim_{X_1 \to \infty} H(X_1) = 0. \tag{6.11}$$

Eqs. (6.8)–(6.11) arise either from the initial problem (3.9)–(3.12) by the insertion of eqs. (6.2)–(6.5) accompanied by boundary conditions (6.6)–(6.7) or directly from the transformed problem given by eqs. (5.6)–(5.9), substituting the chosen parameters.

In order to face numerically the problem (6.8)–(6.11), we have used a numerical solver of MATLAB package which solves any two–point boundary value problem for ODEs by collocation. We have obtained results for the fields $F'(X_1)$ and $H(X_1)$ corresponding to the velocity and temperature fields, respectively. In Fig. 2, it is shown the behaviour of velocity field versus X_1 for different values of the parameter m. We remark that the velocity profile decreases as the value of m increases. In Fig. 3, it is examined the behaviour of the velocity for different choices of the boundary data at $X_1 = 0$. The influence of the kinematic viscosity on the velocity is shown in Fig. 4. As one expects, the velocity increases as the viscosity increases. In Fig. 5, it is presented the behaviour of velocity field for various values of M, which is related with the magnetic field intensity. One can see that as the magnetic field intensity grows up the velocity field decreases. The temperature field follows a similar to the velocity field behaviour (see Fig 6, 7, 8), except the case of correlation with the magnetic field intensity. In Fig. 9, one can notice that the temperature field, unlike the velocity field, increases as the parameter M increases.

Case 2 $\kappa_1 = -\frac{1}{2}$
In this case, as the function G is concerned, we keep the same form as in the previous case, but we choose $\kappa_1 = -\frac{1}{2}$ and $s = -\frac{1}{4}$. This choice allows the function B to depend on both x and y variables. Indeed, in virtue of eq. (4.20), we easily conclude that

$$B(x) = 2^{\frac{1}{4}} B_0 x^{-\frac{1}{4}} y^{-\frac{1}{2}}.$$

The above chosen parameters provide the following group of transformations

$$\begin{aligned} x^* &= e^{-\frac{\varepsilon}{2}} x, \\ y^* &= e^{-\frac{\varepsilon}{4}} y, \\ \psi^* &= e^{-\frac{\varepsilon}{4}} \psi, \\ \theta^* &= e^{\frac{\varepsilon}{2}} \theta. \end{aligned} \tag{6.12}$$

Eqs. (5.3)–(5.4) provide the similarity solution for the case under study as follows

$$\psi(x, y) = 2^{-\frac{1}{2}} x^{\frac{1}{2}} F(X_1), \tag{6.13}$$

$$\theta(x, y) = 2x^{-1} H(X_1), \tag{6.14}$$

with similarity variable

$$X_1 = 2^{\frac{1}{4}} x^{-\frac{1}{4}} y^{\frac{1}{2}}. \tag{6.15}$$

Also, the acceptable boundary conditions for these values of parameters, become

$$\begin{aligned} y = 0: \qquad & \psi_y(x,0) = c_1 \\ & \psi_x(x,0) = -2^{\frac{1}{2}} c_2 x^{-\frac{1}{2}}, \quad x \succ 0 \qquad (6.16) \\ & \theta(x,0) = 2c_3 x^{-1} \\ y \to \infty: \qquad & \psi_y \to 0, \quad \theta \to 0, \quad x \succ 0. \qquad (6.17) \end{aligned}$$

Finally, for the similarity solution given by eqs. (6.13)–(6.14), the system of partial differential equations (3.9)–(3.10) reduce to the following system of ODEs

$$FF^{'} + X_1FF^{''} + 6\nu X_1^{-2}F^{'} - 6\nu X_1^{-1}F^{''} + 2\nu F^{'''} - 8MF^{'} + \phi H = 0, \qquad (6.18)$$

$$4F^{'}H + FH^{'} - 2\alpha X_1^{-2}H^{'} + 2\alpha X_1^{-1}H^{''} = 0 \qquad (6.19)$$

and the boundary conditions take the form

$$F(0) = 4c_2, \quad \lim_{X_1 \to 0+} \frac{F^{'}(X_1)}{X_1} = 2c_1, \quad H(0) = c_3, \qquad (6.20)$$

$$\lim_{X_1 \to \infty} F^{'}(X_1) = 0, \quad \lim_{X_1 \to \infty} H(X_1) = 0. \qquad (6.21)$$

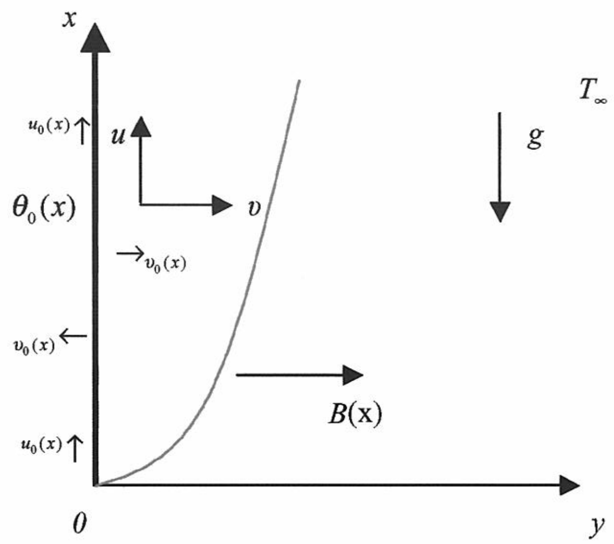
The boundary value problem given by eqs. (6.18)–(6.21) is solved numerically on the interval $[0.01798, 8]$, for $c_1 = 1, c_2 = \frac{5}{4}$ and $c_3 = 1$. In Figures 10 and 11, the behaviour of the fields $F^{'}/X_1$ and H versus X_1 is depicted.

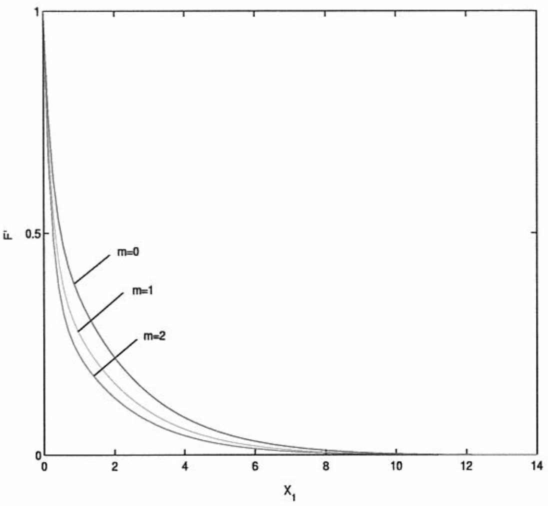
7 Concluding remarks

Using group–theoretical methods, we have obtained the transformation groups for the problem under study. We have found that apart from the scaling group the system admits a group of translations, as well. Concerning the group of scaling and the associated similarity solutions, our results are in full agreement with the work of [20]. Moreover, due to the generality of our procedure and the lack of unnecessary assumptions, we have obtained the general form of the functions involved in the boundary conditions (see eqs. 4.18) and the admissible form of the magnetic field function (eq. 4.20). This enlarges the range of particular problems, probably of practical interest, which can be solved with the similarity methods. Exploiting this fact we have provided a particular example where the magnetic field function, unlike the already existing results, depends on both space variables.

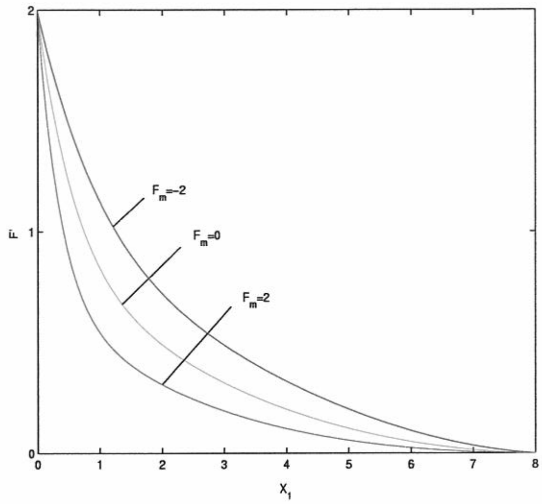
Acknowledgement–The authors thank Prof. A. Raptis for the valuable discussions about this work they had with him. Also, the financial support by the Research Committee of the University of Ioannina, under the "Sotiris Dakaris" project (No 1220), is gratefully acknowledged.

Fig. 1 Boundary-layer around a stretching surface

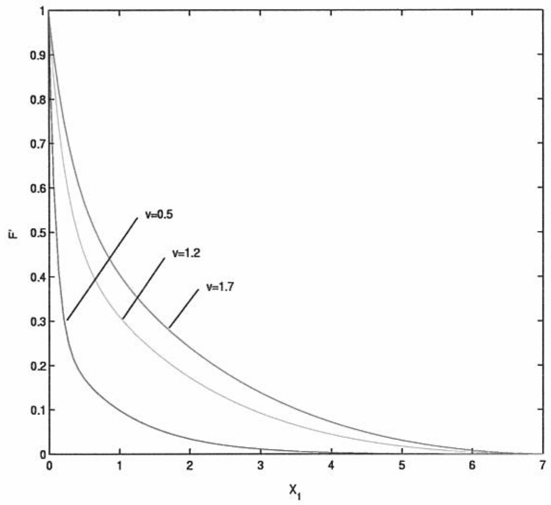




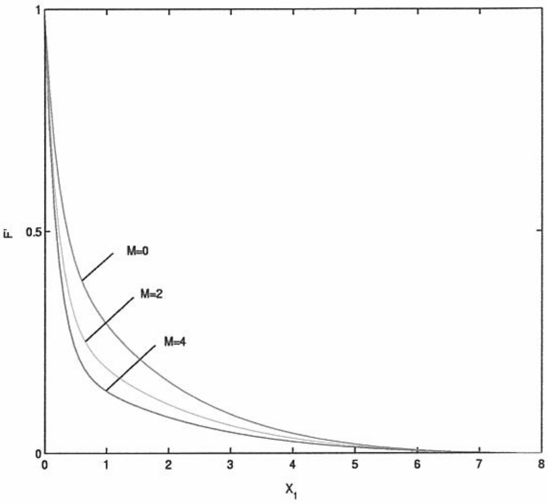
(a) Fig. 2 Behaviour of F' for different m



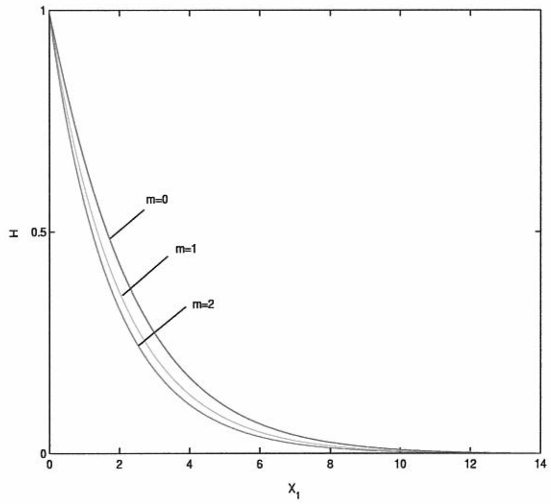
(b) Fig. 3 Behaviour of F' for different F_m



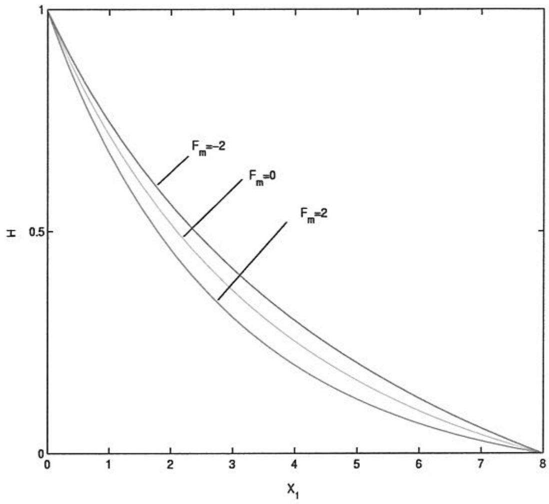
(c) Fig. 4 Behaviour of F' for different v



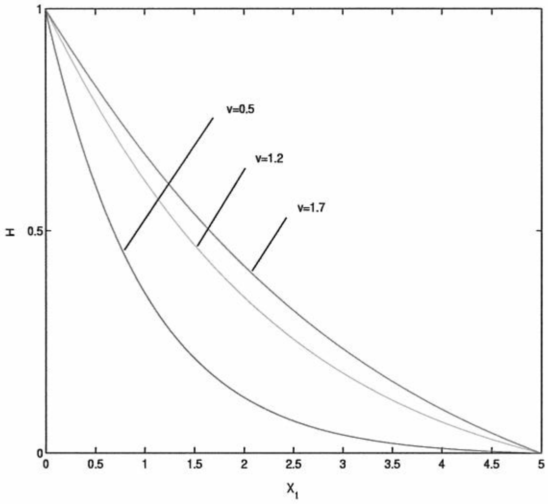
(d) Fig. 5 Behaviour of F' for different M



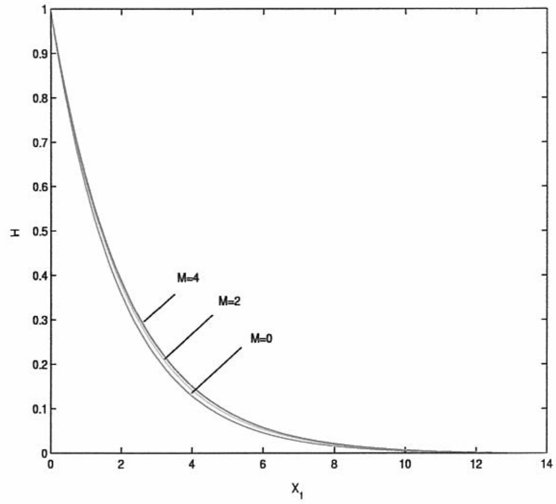
(e) Fig. 6 Behaviour of H for different m



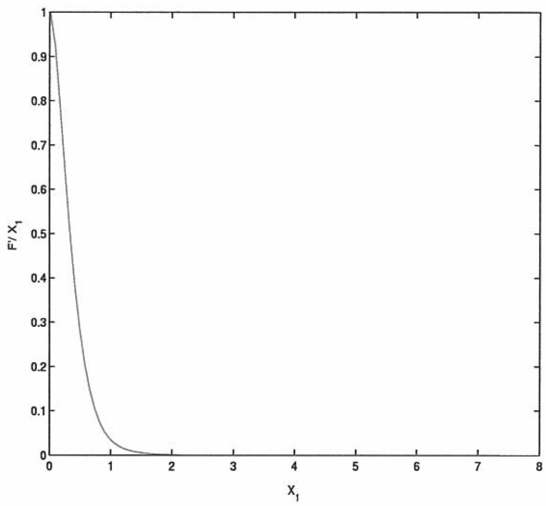
(f) Fig. 7 Behaviour of H for different F_m



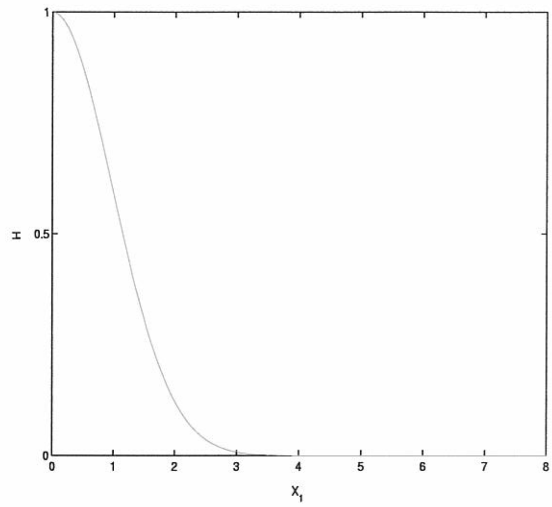
(g) Fig. 8 Behaviour of H for different v



(h) Fig. 9 Behaviour of H for different M



(i) Fig. 10 Behaviour of F'/X_1 when $B = B(x, y)$



(j) Fig. 11 Behaviour of H when B is (x, y)-depended

References

[1] P. J. Olver, *Application of Lie Groups to Differential Equations,* 2nd edition, Springer, New York, (1993).

[2] J. W. Bluman and S. Kumei, *Symmetries and Differential Equations,* Springer, New York, (1989).

[3] N. H. Ibragimov, *Transformation Groups Applied to Mathematical Physics,* D. Reidel, Dordrecht, (1985).

[4] N. H. Ibragimov, *Elementary Lie Group Analysis and Ordinary Differential Equations,* John Wiley and Sons, Chichester, (1999).

[5] Dominic G. B. Edelen, *Applied Exterior Calculus,* John Wiley and Sons, New York, (1985).

[6] E. S. Suhubi, Isovector fields and similarity solutions for general balance equations. *Int. J. Engng. Sci.* **29,** 133 (1991).

[7] E. S. Suhubi, Symmetry groups for one-dimensional longitudinal motions of thermoelastic solids. *ARI* **50,** 57 (1997).

[8] M. Pakdemirli and E. S. Suhubi, Similarity solutions of boundary–layer equations for second order fluids. *Int. J. Engng. Sci.* **30,** 611 (1992).

[9] V. K. Kalpakides, Isovector fields and similarity solutions of nonlinear thermoelasticity. *Int. J. Engng. Sci.* **36,** 1103 (1998).

[10] V. K. Kalpakides, On the symmetries and similarity solutions one–dimensional nonlinear thermoelasticity. *Int. J. Engng. Sci.* **39,** 1863 (2001).

[11] Th. Koureas, A. Charalambopoulos and V. K. Kalpakides, On isovector fields and similarity solutions of generalized dynamic thermoelasticity. *Int. J. Engng. Sci.* **39,** 2071 (2001).

[12] Th. Koureas, A. Charalambopoulos and V. K. Kalpakides, Symmetry groups and Group-invariant solutions of 3D non-linear thermoelasticity. (submitted for publication).

[13] B. C. Sakiadis, Boundary-layer behaviour on continuous solid surfaces. *Amer. Int. Chem. Eng.* **7,** 26 (1961).

[14] L. E. Erickson and L. T. Fan and V. G. Fox , Heat and mass transfer on a moving continuous flat plate with suction or injection. *Ind. Eng. Chem. fund.* **5,** 19 (1966).

[15] F. K. Tsou and E. M. Sparrow and R. J. Goldstein, Flow and heat transfer in the boundary layer on a continuous moving surface. *Int. J. Heat Mass Transfer* **10,** 219 (1967).

[16] P. S. Gupta and A. S. Gupta, Heat and mass transfer on a stretching sheet with suction or blowing. *Can. J. Chem. Eng.* **55,** 744 (1977).

[17] A. Chakrabarti and A. S. Gupta, Hydromagnetic flow and heat transfer over a stretching sheet. *Quart. Appl. Math.* **37,** 73 (1979).

[18] C. K. Chen and M. Char , Heat transfer over a continuous flat stretching surface with suction and blowing. *J. Math. Anal. Appl.* **135,** 568 (1988).

[19] K. Vajravelu and J. Nayfeh, Convective heat transfer at a stretching sheet. *Acta Mechanica* **96,** 47 (1993).

[20] J. Fan, J. Shi and X. Xu, Similarity solutions of free convective boundary–layer behavior at a stretching surface. *Heat and Mass Transfer* **39,** 191 (1999).

Polynomial algebras of parabolic invariants as modules over the Dickson algebra.*

Nondas E. Kechagias

15 May, 2003

Abstract

A free module basis for the ring of upper triangular invariants over the Dickson algebra has been given firstly by Campbell and Hughes [1] and later by the author [3] using different methods. The analogue case for the ring of Borel invariants as a module over parabolic invariants has also been studied [4]. We extend the method employed in [3] to provide free bases of parabolic rings of invariants over the Dickson algebra for particular families of groups and a basis for any group consisting of two blocks. The multiplicative transfer which applies in mod-p cohomology is also studied between appropriate rings of invariants.

1 Introduction

Let $G = GL_n$, B_n, or U_n be the general linear group, the Borel subgroup, and the upper triangular subgroup with 1's on the diagonal, respectively. G acts as usual on V, an n-dimensional $\mathbb{F}_p$-vector space. We write $\mathbb{F}_p[V]$ for the ring of polynomial functions on $\overline{\mathbb{F}_p} \otimes_{\mathbb{F}_p} V$, where $\overline{\mathbb{F}_p}$ is an algebraic closure of $\mathbb{F}_p$.

$$\mathbb{F}_p[V] = \mathbb{F}_p[y_1, \cdots, y_n]$$

**1991 Mathematics Subject Classification.* Primary 13F20, 13A50, 55S10; Secondary 20J05, 18G10.

Key words and phrases. Parabolic invariants, Dickson algebra, Cyclotomic polynomials, Transfer.

Part of this work was presented at the 924th A.M.S. meeting in Montreal Canada.

Here, $V^* = \langle y_1, \cdots y_n\rangle$ and $\mathbb{F}_p[V]$ is a graded polynomial algebra with $|y_i| = 2$ (for topological reasons) or 1, if $p = 2$. The group G acts on $\mathbb{F}_p[V]$ via $(gf)(u) = f(g^{-1}u)$.

The Dickson algebra $D_n :=\mathbb{F}_p[V]^{GL(n,\mathbb{F}_p)}$ plays a fundamental role in modular invariant theory of finite groups and serves as a computational tool in algebraic topology. If G is a finite group acting on a finite dimensional vector space V over the field $\mathbb{F}_p$, the ring $\mathbb{F}_p[V]^G$ is a finite extension of D_n.

It is well known that $\mathbb{F}_p[V]^{U_n}$, $\mathbb{F}_p[V]^{P(n_1,n-n_1)}$ and D_n are all polynomial algebras. Moreover, D_n serves as a homogeneous system of parameters and in fact both $\mathbb{F}_p[V]^{U_n}$ and $\mathbb{F}_p[V]^{P(n_1,n-n_1)}$ are free D_n-modules.

A free basis has been given for $\mathbb{F}_p[V]^{U_n}$ as a module over D_n ([1] and [3]) and $\mathbb{F}_p[V]^{P(n_1,n-n_1)}$ [4]. Here $P(n_1, n - n_1)$ is a parabolic subgroup of GL_n. In this work, we study the case $\mathbb{F}_p[V]^{P(n_1,n-n_1)}$ as a free module over D_n for some families of natural numbers $(n_1, n-n_1)$. The problem is focussed on the numbers: n and n_1. Only for certain families of pairs (n, n_1), the situation is completely analogous to the ones above, Proposition 27, but in general it is different than the previous ones and more complicated depending on the divisibility between natural numbers and certain relations between them. For the later case, we provide appropriate formulas (in section 2) which serve as a central tool to construct a free basis in Proposition 29. For the general case a module basis for $\mathbb{F}_p[V]^{P(n_1,n-n_1)}$ over D_n is given in Theorem 31:

Theorem 31 *The set* $B' = \left\{ \prod_{i=0}^{n_1-1} d_{n_1,i}^{m_i} \,|\, 0 \le m_i \le A_i \right\}$ *is a module basis for* $\mathbb{F}_p(n_1, n_2)$ *over* D_n.

Here the bounds A_i shall be defined in definition 17 and depend on n and n_1.

The interested reader can extend our method to obtain a free basis for particular choices using our formulas provided in section two and the preliminary subsection of section three.

Besides our invariant theoretic interest, we are motivated by topological properties of the transfer between cohomology of certain subgroups of the symmetric group. Namely, we are interested in multiplicative properties of the transfer studied by Kuhn and Priddy, because of their application in stable homotopy theory [6]. The transfer between parabolic subgroups is studied in the last section. Theorem 41 asserts that the natural map between $\mathbb{F}_p[V]^{B_n}$ and $\mathbb{F}_p(N, n)$ equals the transfer map and the analogue of Kuhn and Priddy result is expressed in Theorem 44:

Theorem 44 *Let* $\tau^* : \mathbb{F}_p[V]^G \to D_n$ *be the transfer map, where* $G = U_n$ *or* B_n. *Let* $x, y \in \mathbb{F}_p[V]^{U_n}$ *or* $\mathbb{F}_p[V]^{B_n}$. *Then* $\tau^*(xy) = \tau^*(x)\tau^*(y)$ *for all* y *iff* $x \in D_n$.

Acknowledgment: We thank Professor Gregor Kemper for helping us out with example 30.

We finish this introduction by recalling some well known Theorems relevant to the Dickson algebra.

Let $\sum_{p^n}$ be the symmetric group acting on V by permutations and ${\mathbb{F}_p}^n$ its subgroup consisting of all translations. Then $Aut({\mathbb{F}_p}^n) \cong GL_n$ and its Weyl subgroup, $W_{\sum_{p^n}}({\mathbb{F}_p}^n) \cong GL_n$, acts on V^* as follows:

$$(a_{i,j})y_k := \sum_i a_{i,k}y_i$$

We repeat some classical results from the literature. First the Dickson algebra, D_n, is described. Let

$$h_i = \prod_{a \in \langle y_1, \cdots, y_{i-1} \rangle} (y_i - a), \quad L_n = \prod h_i, \quad d_{n,i} = \frac{\begin{vmatrix} y_1 & \cdots & y_n \\ \vdots & \cdots & \vdots \\ y_1^{p^n} & \cdots & y_n^{p^n} \end{vmatrix}}{L_n}$$

In the last determinant the p^i -th power is missing. The degrees of the previous elements are p^{i-1}, $\frac{p^n-1}{p-1}$, and $p^n - p^i$ respectively. We note that there are other descriptions for the polynomials above [8].

Theorem 1 (Dickson) $D_n := \mathbb{F}_p[V]^{GL_n} = \mathbb{F}_p[d_{n,0}, \cdots, d_{n,n-1}]$.

Theorem 2 (Mùi) *i)* $H_n := \mathbb{F}_p[V]^{U_n} = \mathbb{F}_p[h_n, \cdots, h_1]$.
ii) $\mathbb{F}_p[V]^{B_n} = \mathbb{F}_p[(h_n)^{p-1}, \cdots, (h_1)^{p-1}]$.

Relations between the generators of rings of invariants are given as follows:

Proposition 3 *[3]* $d_{n,n-i} = \sum_{1 \le j_1 < \cdots < j_i \le n} \prod_{s=1}^{i} \left(h_{j_s}^{p-1}\right)^{p^{n-i+s+j_s}}$.

It is known that any subgroup between B_n and GL_n is conjugate to a parabolic subgroup. Let $N = (n_1, \cdots, n_\ell)$ be a sequence of non-negative integers such that $\sum n_i = n$. Let $P(N,n)$ be the so called parabolic subgroup of GL_n:

$$\begin{pmatrix} GL_{n_1} & * & * \\ 0 & \ddots & * \\ 0 & 0 & GL_{n_\ell} \end{pmatrix}$$

Theorem 4 *(Kuhn and Stong).*

$$\mathbb{F}_p(N,n) := \mathbb{F}_p[V]^{P(N,n)} = \mathbb{F}_p[d_{\nu_i,\nu_i-k_i} \mid 1 \le i \le \ell,\ 1 \le k_i \le n_i,\ \nu_i = \sum_{t=1}^{i} n_t].$$

2 Relations between parabolic and Dickson generators

Since D_n is a subalgebra of $\mathbb{F}_p[V]^{P(N,n)}$, any Dickson generator can be decomposed in terms of generators of the later algebra. We shall describe these relations in this section. For simplicity only the case $N = (n_1, n_2 = n - n_1)$ will be considered. The interested reader can extend those formulas to any number of blocks.

A Dickson generator $d_{n,n-i}$ consists of the sum of all possible combinations of i elements from $\{h_1^{p-1}, \cdots, h_n^{p-1}\}$ in certain p-th exponents. Let $d'_{n,n-i}$ be the polynomial which is given as $d_{n,n-i}$ but elements are from $\{h_{n_1+1}^{p-1}, \cdots, h_n^{p-1}\}$ on the same exponents. Here $n - i \ge n_1$. It is obvious that the new polynomial is a summand of the old one and it will be expressed in terms of old generators.

Example 5 *Let $n = 5$ and $n_1 = 2$.*

i) $d'_{5,4} = h_3^{(p-1)p^2} + h_4^{(p-1)p} + h_5^{(p-1)} = d_{5,4} - d_{2,1}^{p^3}$

ii) $d'_{5,3} = h_5^{(p-1)} h_4^{(p-1)} + h_5^{(p-1)} h_3^{(p-1)p} + h_4^{(p-1)p} h_3^{(p-1)p} = d_{5,3} - d_{2,1}^{p^2} d'_{5,4} - d_{2,0}^{p^3} = d_{5,3} - d_{2,1}^{p^2} d_{5,4} + d_{2,1}^{p^2+p^3} - d_{2,0}^{p^3}$

iii) $d'_{5,2} = h_5^{(p-1)} h_4^{(p-1)} h_3^{(p-1)} = d_{5,2} - d_{2,1}^{p} d'_{5,3} - d_{2,0}^{p^2} d'_{5,4} = d_{5,2} - d_{2,1}^{p} d_{5,3} + d_{2,1}^{p^2+p} d_{5,4} - d_{2,1}^{p+p^2+p^3} + d_{2,0}^{p^3} d_{2,1}^{p} - d_{2,0}^{p^2} d_{5,4} + d_{2,0}^{p^2} d_{2,1}^{p^3}$

Proposition 6 *Let $n = n_1 + n_2$ and $n_2 \geq i \geq 1$. Then*

$$d'_{n,n-i} = d_{n,n-i} + \sum_{t=1}^{i}\sum_{\ell=1}^{t} d_{n,n-i+t} \sum_{\substack{j_1+\ldots+j_l=t\\ 0<j_s\leq n_1}} (-1)^{\ell} d_{n_1,n_1-j_1}^{p^{n_2-i+t}} d_{n_1,n_1-j_2}^{p^{n_2-i+t-j_1}} \ldots d_{n_1,n_1-j_l}^{p^{n_2-i+t-j_1-\ldots-j_{l-1}}}$$

Here $d_{k,k} = 1$.

Proof. First we decompose $d'_{n,n-i}$ in terms of $d'_{n,n-t}$ for $t < i$.

$$d'_{n,n-i} = d_{n,n-i} - \sum_{t=1}^{i} d'_{n,n-i+t} d_{n_1,n_1-t}^{p^{n_2-i+t}}$$

Proof by induction on $1 \leq i \leq n$ and use of Proposition 3.
For the proof of the statement induction is used on $1 \leq i \leq n$. For the general step, each $d'_{n,n-i+t}$ is decomposed (by induction hypothesis) and terms are collected with respect to $d_{n,n-i+c}$'s for $1 \leq c \leq i$.

$$\left[\sum_{t=1}^{c} d_{n_1,n_1-t}^{p^{n_2+t-i}} \sum_{\ell=0}^{c-t}(-1)^{\ell+1} \sum_{j_1+\ldots+j_l=i-t} d_{n_1,n_1-j_1}^{p^{n_2-i+c}} d_{n_1,n_1-j_2}^{p^{n_2-i+c-j_1}} \ldots d_{n_1,n_1-j_l}^{p^{n_2-i+c-j_1-\ldots-j_{l-1}}}\right] d_{n,n-i+c}$$

We have to show that the expression above coincides with the required expression:

$$\sum_{\ell=1}^{c} d_{n,n-i+c} \sum_{j_1+\ldots+j_l=i} (-1)^{\ell} d_{n_1,n_1-j_1}^{p^{n_2-i+t}} d_{n_1,n_1-j_2}^{p^{n_2-i+t-j_1}} \ldots d_{n_1,n_1-j_l}^{p^{n_2-i+t-j_1-\ldots-j_{l-1}}}$$

It is obvious that the last expression contains the previous one. The other direction is shown by considering an individual member of the one above. ■

Now the next proposition is obvious.

Proposition 7 $\mathbb{F}_p(n_1, n_2) :=\mathbb{F}_p[V]^{P(n_1,n_2)} =\mathbb{F}_p[d_{n_1,j},\, d'_{n,n-i} \,|\, 0 \leq j \leq n_1 - 1$ *and* $1 \leq i \leq n_2]$.

Our last task in this section is to provide relations between parabolic and Dickson invariants.

Example 8 *We continue our last example.* $d_{5,0} = (h_1 \cdots h_5)^{p-1} = d_{2,0}d'_{5,2} = d_{2,0}d_{5,2} - d_{2,0}d^{p}_{2,1}d_{5,3} + d_{2,0}d^{p+p^2}_{2,1}d_{5,4} - d_{2,0}d^{p+p^2+p^3}_{2,1} + d^{1+p^3}_{2,0}d^{p}_{2,1} - d^{1+p^2}_{2,0}d_{5,4} + d^{1+p^2}_{2,0}d^{p^3}_{2,1}$.

$d_{5,1} = (h_2 \cdots h_5)^{p-1} + (h_1 \cdots h_4)^{p(p-1)} + \cdots + (h_1 \cdots h_3)^{p^2(p-1)}h_5^{p-1} = d_{2,1}d'_{5,2} + d^{p}_{2,0}d'_{5,3} = d_{2,1}d_{5,2} - d^{1+p}_{2,1}d_{5,3} + d^{1+p+p^2}_{2,1}d_{5,4} - d^{1+p+p^2+p^3}_{2,1} - d^{p^3}_{2,0}d^{1+p}_{2,1} - d^{p^2}_{2,0}d_{2,1}d_{5,4} + d^{p^2}_{2,0}d^{1+p^3}_{2,1} + d^{p}_{2,0}d_{5,3} - d^{p}_{2,0}d^{p^2}_{2,1}d_{5,4} + d^{p}_{2,0}d^{p^2+p^3}_{2,1} - d^{p+p^3}_{2,0}$.

Proposition 9 *Let* $0 \le i \le n_1 - 1$. *Then*

$$d_{n,i} = \sum_{t \ge \max(0, i-n_2)}^{i} d^{p^{i-t}}_{n_1,t} d'_{n,n_1+i-t}$$

Proof. We use formula 3 and consider each product like being divided in to two parts according to indices of the h_i^{p-1}'s. The first part consists of those h_i^{p-1}'s such that their indices are less or equal than n_1. Then we group monomials according to the second part. ■

3 $\mathbb{F}_p(n_1, n_2)$ as a free module over D_n

3.1 Preliminary

Let p be a prime number and k, s and r_t be natural numbers such that $1 \le t \le s$ and $1 \le r_t < r_{t+1}$.

Definition 10 *Given two sequences consisting of* s *natural numbers* $(k + r_s, ..., k + r_1)$ *and* $(r_s, ..., r_1)$, r_i *as above, the symbol* $\left[{k+r_s,...,k+r_1 \atop r_s,...,r_1}\right]$ *stands for the fraction* $\frac{\prod_{i=1}^{s}(k+r_i)}{\prod_{i=1}^{s} r_i}$. $\left[{k+r_s,...,k+r_1 \atop r_s,...,r_1}\right]$ *is called admissible, if the number of times a particular natural number is a divisor of the* r_t*'s is less or equal than the number of times that particular number is also a divisor of the* $k + r_j$*'s.*

Lemma 11 *If* $\left[{k+r_s,...,k+r_1 \atop r_s,...,r_1}\right]$ *is admissible, then* $\left[{k+r_s,...,k+r_1 \atop r_s,...,r_1}\right]$ *is integral.*

Example 12 *a)* $\begin{bmatrix} 13, & 12, & 11, & 10, & 9, & 8 \\ 6, & 5, & 4, & 3, & 2, & 1 \end{bmatrix}$ *is admissible:* $\{6, 3, 2\}$, $\{5\}$, $\{4, 2\}$, $\{3\}$, $\{2\}$ *and* $\{13\}$, $\{12, 6, 4, 3, 2\}$, $\{11\}$, $\{10, 5, 2\}$, $\{8, 4, 2\}$, $\{9, 3\}$.

b) $\begin{bmatrix} 30, & 28, & 23, & 22, & 20 \\ 12, & 10, & 5, & 4, & 2 \end{bmatrix}$ *is integral but it is not admissible:*
i) $20x22x23x28x30/2x4x5x10x12 = 1771$;
ii) $\{\underline{12}, 6, 4, 3, 2\}$, $\{10, 5, 2\}$, $\{5\}$, $\{2, 4\}$, $\{2\}$ *and* $\{30, 15, 10, 6, 5, 3, 2\}$, $\{28, 14, 7, 4, 2\}$, $\{23\}$, $\{22, 11, 2\}$, $\{20, 10, 5, 4, 2\}$.

Let us reserve the symbol $\left[{k+r \atop r}\right]$ for the sequence $r_t = t$.

Proposition 13 $\left[{k+r \atop r}\right]$ *is admissible for* k *and* $r \geq 1$.

Proof. Using double induction and the well known formula $\binom{k+r}{r} = \binom{k+r-1}{r} + \binom{k+r-1}{r-1}$, the integrality of $\binom{k+r}{r}$ is proved. We assume that $\left[{k+r-1 \atop r-1}\right]$ is admissible. If r is a prime number or divides $(k+r)$, then the statement follows. Otherwise, let $\frac{r}{q}$ be a divisor of r. Then $\frac{r}{q}$ is also a divisor of $m\frac{r}{q}$ for $1 \leq m \leq q$. We need to show that there exist at least q natural numbers $k + r'$ between $k+1$ and $k+r$ which are divisible by $\frac{r}{q}$. Let $k = x\frac{r}{q} + l$ with $0 \leq l \leq \frac{r}{q} - 1$. Then $k + r = (x+q)\frac{r}{q} + l$ and the statement follows. ■

Lemma 14 *If* $\left[{k+r_s, \ldots, k+r_1 \atop r_s, \ldots, r_1}\right]$ *is admissible, then* $\left[{m+k+r_s, \ldots, m+k+r_1 \atop r_s, \ldots, r_1}\right]$ *is also admissible, where* m *is a multiple of* $\operatorname{lcm}(r_s, \ldots, r_1)$.

We shall define a peculiar division in $\left[{k+r \atop r}\right]$ as follows:

Definition 15 *For each element* r_t *of* $(r_s, \ldots, r_1)$ *we let* r_t *divide* $k + r_l$, $r_t "\backslash" k + r_l$, *if* $\frac{k+r_l}{r_t}$ *is integral,* $l \leq t$ *and* l *is maximal with this property.*

Let $I_{exact} = \{r_t \mid r_t | k + r_t\} \subset \{1, \ldots, r\}$. Now we define a partition of $\{1, \ldots, r\}$ according to the given division as follows: For each $r_{t(0)} \in I_{exact}$ let $I_{r_{t(0)}}$ contain all natural numbers l between r and $r_{t(0)}$ such that there is a subsequence $(l = r_{t_l}, r_{t_l - 1}, \ldots, r_{t_l - s_l} = r_{t(0)})$ with $r_{t_l - s} "\backslash" k + r_{t_l - s - 1}$. Now $\{1, \ldots, r\} = \bigsqcup_{r_{t(0)}} I_{r_{t(0)}}$.

Example 16 *a)* $\left[{23=12+11 \atop 11}\right]$, $I_{exact} = \{6, 4, 3, 2, 1\}$, $I_6 = \{9, 6\}$, $I_4 = \{11, 10, 8, 4\}$, $I_3 = \{5, 3\}$, $I_2 = \{7, 2\}$ *and* $I_1 = \{1\}$. $\begin{bmatrix} 21, & 18 \\ 9, & 6 \end{bmatrix}$, $\begin{bmatrix} 23, & 22, & 20, & 16 \\ 11, & 10, & 8, & 4 \end{bmatrix}$, $\begin{bmatrix} 17, & 15 \\ 5, & 3 \end{bmatrix}$, $\begin{bmatrix} 19, & 14 \\ 7, & 2 \end{bmatrix}$ *are all admissible.*

b) $\left[{29=18+11 \atop 11}\right]$, $I_{exact} = \{9,6,3,2,1\}$, $I_9 = \{9\}$, $I_6 = \{8,6\}$, $I_3 = \{7,3\}$, $I_2 = \{11,4,2\} \cup \{10,2\} \cup \{5,2\}$ *and* $I_1 = \{1\}$. $\begin{bmatrix} 27 \\ 9 \end{bmatrix}, \begin{bmatrix} 26, & 24 \\ 8, & 6 \end{bmatrix}, \begin{bmatrix} 25, & 21 \\ 7, & 3 \end{bmatrix}$ *are admissible sequences but* $\begin{bmatrix} 29, & 28, & 23, & 22, & 20 \\ 11, & 10, & 5, & 4, & 2 \end{bmatrix}$ *is not even integral.*

Remark 1 *We note that only sequences satisfying certain properties will be considered in this work due to great technicalities.*

Definition 17 *According to division defined above between sequences and for any prime number p, we define* A_{r_t} *to be the natural number*

$$A_{r_t} := \begin{cases} p^{r_t - r_{t-i(t)}} \frac{p^{k+r_{t-i(t)}}-1}{p^{r_t}-1}, r_t \neq r_{t-i(t)} \\ \frac{p^{k+r_t}-p^{r_t}}{p^{r_t}-1}, r_t = r_{t-i(t)} \end{cases}$$

And

$$X\begin{bmatrix} k+r_s, \dots, k+r_1 \\ r_s, \dots, r_1 \end{bmatrix} := \prod_{t=1}^{s} \frac{p^{k+r_t}-1}{p^{r_t}-1}$$

$$X\begin{bmatrix} k+r \\ r \end{bmatrix} := \prod_{t=1}^{r} \frac{p^{k+t}-1}{p^t-1}$$

Remark 2 *Note that* A_i *is the same in* $X\left[{k+r \atop r}\right]$ *and* $X\left[{k+r-1 \atop r-1}\right]$ *for* $i = 1, \dots, r-1$.

Proposition 18 $X\left[{k+r_s,\dots,k+r_1 \atop r_s,\dots,r_1}\right]$ *is integral if and only if* $\left[{k+r_s,\dots,k+r_1 \atop r_s,\dots,r_1}\right]$ *is admissible.*

Proof. It is known that $x^k - 1 = \prod_{d|k} C_d$, where C_d stands for the d-th cyclotomic polynomial. Thus $\prod_{t=1}^{s} \frac{x^{k+r_t}-1}{x^{r_t}-1} = \prod_{t=1}^{s} \frac{\prod_{d|(k+r_t)} C_d}{\prod_{d'|r_t} C_{d'}}$ and the statement follows because of the definition of admissibility. ■

Lemma 19 *Let* $\left[{k+r_s,\dots,k+r_1 \atop r_s,\dots,r_1}\right]$ *such that* $r_t | k + r_{t-1}$, $r_1 | k$ *and for each divisor* d *of* $\gcd(r_1, r_2)$
$d | k + r_s$ *or* $\exists t : d \nmid r_t$ *and* $d | k + r_{t-1}$. *Then* $X\left[{k+r_s,\dots,k+r_1 \atop r_s,\dots,r_1}\right]$ *is integral.*

Remark 3 *k in the last lemma must satisfy the following conditions: $k = l' \operatorname{lcm}(r_1, ...r_s) + l \operatorname{lcm}(r_1, r_2) - r_1$ such that l' is a non-negative integer and l is a positive integer and $l \operatorname{lcm}(r_1, r_2) - r_1 + r_{i-1} \equiv 0 \bmod r_i$. If $r_s = s$, then $k = l \operatorname{lcm}(2, ..., s) + 1$.*

Example 20 $\begin{bmatrix} 19,18,16 \\ 9,8,4 \end{bmatrix}$ *is integral although* $X \begin{bmatrix} 19,18,16 \\ 9,8,4 \end{bmatrix}$ *is not integral.*

Proposition 21 *Let $\begin{bmatrix} k+r_s,...,k+r_1 \\ r_s,...,r_1 \end{bmatrix}$ such that the r_t's are as in the last lemma. Then $X \begin{bmatrix} k+r_s,...,k+r_1 \\ r_s,...,r_1 \end{bmatrix} = \sum_{i=0}^{s} \prod_{t=0}^{i} A_{r_t}$. Here $A_{r_i} = p^{r_i - r_{i-1}} \frac{p^{k+r_{i-1}}-1}{p^{r_i}-1}$, $A_{r_i} = p^{r_1} \frac{p^k-1}{p^{r_1}-1}$ and $A_0 = 1$.*

Proof. We use induction on s. $1 + A_{r_1} = \frac{p^{k+r_1}-1}{p^{r_1}-1}$. For the general step: $X \begin{bmatrix} k+r_{s-1},...,k+r_1 \\ r_{s-1},...,r_1 \end{bmatrix} + \prod_{t=0}^{s} A_{r_t} = \prod_{t=1}^{s} \frac{p^{k+r_t}-1}{p^{r_t}-1} \Leftrightarrow \prod_{t=0}^{s} A_{r_t} = X \begin{bmatrix} k+r_{s-1},...,k+r_1 \\ r_{s-1},...,r_1 \end{bmatrix} p^{r_s} \frac{p^k-1}{p^{r_s}-1}$ $\Leftrightarrow p^{r_s} \frac{p^k-1}{p^{r_1}-1} \prod_{t=2}^{s-1} \frac{p^{k+r_{t-1}}-1}{p^{r_t}-1} = \prod_{t=1}^{s} \frac{p^{k+r_t}-1}{p^{r_t}-1}$. ■

The next corollary serves as a collection of relations between generators of D_{n_1}.

Corollary 22 *a) Let $0 \leq i \leq n_1 - 1$ and $(n_1 - i)$ divides $n - i$. Then $d_{n,i}$ contains $d_{n_1,i}^{A_i+1}$ as a summand. Moreover d'_{n,n_1} contains $d_{n_1,i}^{A_{n_1-i}}$ as a summand.*

b) Let $0 \leq i \leq n_1 - 1$ and $(n_1 - i) \in I_{r_{t(0)}}$. Then $d_{n,i}$ contains $d_{n_1,i} d_{n_1,n_1-r_{t(0)}}^{A_{r_{t(0)}}}$ as a summand.

c) Let i and v_i such that $0 \leq i \leq n_1 - 1$, $n - i - v_i = l_i(n_1 - i)$ and $0 \leq v_i \leq n_1 - i$ where v_i is the smallest with this property. Then $d_{n,i+v_i} d_{n_1,i} - d_{n,i} d_{n_1,i+v_i}$ contains $d_{n_1,i}^{1+A_{n_1-i}}$ as a summand.

d) Let i and t such that $0 \leq i < t \leq n_1 - 1$, $(n_1 - i)" \nmid "(n - t)$ and $0 \leq t' < t$ where $(n_1 - t')"|"(n - t)$ and t' is maximal with this property. Then $d_{n,t} d_{n_1,i} - d_{n,i} d_{n_1,t}$ contains $d_{n_1,i} d_{n_1,t'}^{A_{n_1-t'}}$ as a summand.

Proof. The proof depends on Propositions 9 and 6.
a) Since $(n_1 - i)$ divides $(n - i) = (n - n_1 + n_1 - i)$ and $(n - n_1)$, $(p^{n_1} - p^i)$ divides $(p^n - p^i)$ and this is how A_{n_1-i} has been defined in 17. Moreover, $d_{n,i}$ contains $d_{n_1,i} d'_{n,n_1}$ and d'_{n,n_1} contains $d_{n_1,i}^{A_{n_1-i}}$ (Proposition 6).
b) By definition $r_{t(0)}$ divides $n_2 + r_{t(0)}$, hence $r_{t(0)}$ divides n_2. This implies, as in a), that d'_{n,n_1} contains $d_{n_1,n_1-r_{t(0)}}^{A_{r_{t(0)}}}$.

c) We use last Proposition for $i = i + v_1$ and $t = i$ and Proposition 6 for the decomposition of d'_{n,n_1+v_i}. Then $d'_{n,i+v_i}$ contains $d_{n_1,i}^{\left(\sum_{t=1}^{l_i-1} p^{(v_i+(n_1-i)t)}\right)}$ as a summand. Let us recall definition 17: $A_{n_1-i} = v_i \sum_{t=0}^{l_i-1} p^{(n_1-i)t}$.

d) Let $s = t-t'$ in $d_{n,t}d_{n_1,i} - d_{n,i}d_{n_1,t} = \sum_{s=1}^{t} d'_{n,n_1+s}d_{n_1,t-s}d_{n_1,i} - \sum_{s=1}^{i} d'_{n,n_1+s}d_{n_1,i-s}d_{n_1,t}$, then $d'_{n,n_1+t-t'}$ contains $d_{n_1,t'}^{A_{n_1-t'}-1}$. ■

We shall close this subsection by considering some maps which will let us connect $\mathbb{F}_p(n_1, n_2)$ with $\mathbb{F}_p(n_1 - 1, n_2)$. Those maps have been used by Campbell and Hughes in [1].

Let $\pi_{y_1} : \mathbb{F}_p[y_1, \cdots, y_n] \to \mathbb{F}_p[y_1, \cdots, y_n]$ be the map induced by

$$\pi_{y_1}(y_i) = \begin{cases} 0, & i = 1 \\ y_i, & i > 1 \end{cases}$$

and $sh_{(-1)} : \mathbb{F}_p[y_1, \cdots, y_n] \twoheadrightarrow \mathbb{F}_p[y_1, \cdots, y_{n-1}]$ the one induced by

$$sh_{(-1)}(y_i) = \begin{cases} 0, & i = 1 \\ y_{i-1}, & i > 1 \end{cases}$$

Now let us consider the induced maps on $\mathbb{F}_p(n_1, n_2)$: $sh_{(-1)} \cdot \pi_{y_1}(d_{n_1,i}) = d^p_{n_1-1,i-1}$. It follows that $\mathrm{Im}(sh_{(-1)} \cdot \pi_{y_1}) = \mathbb{F}_p[d^p_{n_1-1,i-1}, d^p_{n-1,j-1} \mid i = 0, ..., n_1 - 2, j = n_1 - 1, ..., n-2]$. We need one more map $\varepsilon_{(1)} : \mathrm{Im}(sh_{(-1)} \cdot \pi_{y_1}) \to \mathbb{F}_p(n_1, n_2)$ given by

$$\begin{aligned} \varepsilon_{(1)}(d^p_{n_1-1,i-1}) &= d_{n_1,i} \\ \varepsilon_{(1)}(d^p_{n-1,j-1}) &= d_{n,j} \end{aligned}$$

Lemma 23 *Let $f \in \mathbb{F}_p(n_1, n_2)$, then $\varepsilon_{(1)} \cdot sh_{(-1)} \cdot \pi_{y_1}(f) - f \in (d_{n_1,0})$.*

Here $(d_{n_1,0})$ is the ideal in $\mathbb{F}_p(n_1, n_2)$ generated by the top Dickson generator $d_{n_1,0}$.

Proof. Let $d \in \mathbb{F}_p(n_1, n_2)$ be a non-trivial monomial. We consider two cases. Let $d \notin (d_{n_1,0})$, then d is not divisible by y_1. Otherwise, $d \in (d_{n_1,0})$, since $d \in \mathbb{F}_p(n_1, n_2)$. Thus $d \in (d_{n_1,1}, ..., d_{n,n-1})$ and $d = \prod_{i>0, j\geq n_1} d^{m_i}_{n_1,i} d^{m_j}_{n,j}$. Now, $sh_{(-1)} \cdot \pi_{y_1}(d) = \prod_{i>0, j\geq n_1} d^{pm_i}_{n_1-1,i} d^{pm_j}_{n-1,j}$ and $\varepsilon_{(1)} \cdot sh_{(-1)} \cdot \pi_{y_1}(d) = d$.

Let $d \in (d_{n_1,0}) \Leftrightarrow \pi_{y_1}(d) = 0$. Thus, if $f \in \mathbb{F}_p(n_1, n_2)$ and $g \in \mathrm{Im}(sh_{(-1)} \cdot \pi_{y_1})$ such that $g = sh_{(-1)} \cdot \pi_{y_1}(f)$, then $\varepsilon_{(1)}(g) - f \in (d_{n_1,0})$. ■

3.2 Main Results

Since U_n is a p-Sylow subgroup of GL_n and H_n is a polynomial algebra, $\mathbb{F}_p(n_1, n_2)$ is Cohen-Macaulay [[8], Proposition 8.3.1]. Hence, $\mathbb{F}_p(n_1, n_2)$ is a free module over D_n and we shall provide a free basis for particular choices of n_1 and n_2.

In the sequel $B_A(A')$ stands for a free module basis of the algebra A' over the algebra A and $B'_A(A')$ for a module basis respectively.

Bases are known for H_n and $\mathbb{F}_p[V]^{B_n}$ over D_n.

Theorem 24 *[1], [3]. The set* $B_{D_n}(H_n) = \{h_1^{r_1} \cdots h_n^{r_n} \mid 0 \leq r_i < p^{n-i+1} - 1\}$ *is a free module basis for* H_n *over* D_n.

We extended the last theorem for the parabolic subgroups following the method used in [3].

Theorem 25 *[4]i) The set* $B_{\mathbb{F}_p[V]^{B_n}}(H_n) = \{h_1^{r_1} \cdots h_n^{r_n} \mid 0 \leq r_i < p - 1\}$ *is a free module basis for* H_n *over* $\mathbb{F}_p[V]^{B_n}$.

ii) The set

$$B_{\mathbb{F}_p(N,n)}\left(\mathbb{F}_p[V]^{B_n}\right) = \left\{h_1^{(p-1)r_1} \cdots h_n^{(p-1)r_n} \mid 0 \leq r_i < \tfrac{p^{\nu_s - i+1} - 1}{p-1},\ \nu_{s-1} < i \leq \nu_s\right\}$$

is a free module basis for $\mathbb{F}_p[V]^{B_n}$ *over* $\mathbb{F}_p(N, n)$. *Here,* $\nu_i = \sum_{t=1}^{i} n_t$.

We shall make a few remarks concerning the key-points of the proof of theorems above.

i) Let G be one of the groups under consideration. There is a close relation between degrees of generators and the order of the group: $\mathbb{F}_p[V]^G = \mathbb{F}_p[\alpha_1, \cdots, \alpha_n]$, $\prod_1^n |\alpha_i| = 2^n|G|$. The rank of H_n over $\mathbb{F}_p[V]^G$ is $[G : U_n]$.

ii) Let $P(G, t)$ denote the Poincaré series of $\mathbb{F}_p[V]^G$. Note that $|h_i|$ divides $|\alpha_i|$ and hence $P\left(\mathbb{F}_p[V]^G, t\right) / P\left(H_n, t\right) = \prod_i \left(1 + t^{|h_i|} + t^{2|h_i|} + \cdots + t^{\left(\frac{|\alpha_i|}{|h_i|} - 1\right)|h_i|}\right)$.

The last statement along with proposition 3 provides a strong hint for a set of free basis generators, namely those expressed in the last theorems.

This is not the case for $\mathbb{F}_p(n_1, n_2)$ over D_n as the next example suggests.

Example 26 *Let* $n = 5$ *and* $n_1 = 2$. $|d_{5,0}| = p^5 - 1$, $|d_{2,0}| = p^2 - 1$ *and* $\frac{p^5-1}{p^2-1}$ *is not integral. Hence the method described above does not apply. But* $X\left[{5 \atop 2}\right]$

is integral and according to Proposition 21

$$X\begin{bmatrix}5\\2\end{bmatrix} = 1 + A_1 + A_1A_2 = 1 + p^3 + p^2 + p + (p^3 + p^2 + p)(p^3 + p)$$

Proposition 27 *Let $n_2 = l\,\mathrm{lcm}(2, ..., n_1)$. Then*

$$B_{D_n}(\mathbb{F}_p(n_1, n_2)) = \left\{ \prod_{i=0}^{n_1-1} d_{n_1,i}^{m_i} \,|\, 0 \le m_i \le A_{n_1-i} \right\}$$

is a free module basis for $\mathbb{F}_p(n_1, n_2)$ over D_n.

Proof. Let us recall that $A_{n_1-i} + 1 = \sum_{t=0}^{n_2/(n_1-i)} p^{(n_1-i)t}$. The proof depends on the following facts:
i) The hypothesis about n_2 guarantees that $\frac{p^{n-i}-1}{p^{n_1-i}-1}$ is integral for each i.
ii) $d_{n_1,i}^{A_{n_1-i}+1}$ is a summand in the decomposition of $d_{n,i}$ (22-a)).
iii) Mimic the proof of last theorem. ■
We continue our last example.

Example 28 $d_{5,1} = d_{2,1}d_{5,2} - d_{2,1}^{1+p}d_{5,3} + d_{2,1}^{1+p+p^2}d_{5,4} - d_{2,1}^{1+p+p^2+p^3} - d_{2,0}^{p^3}d_{2,1}^{1+p} - d_{2,0}^{p^2}d_{2,1}d_{5,4} + d_{2,0}^{p^2}d_{2,1}^{1+p^3} + d_{2,0}^{p}d_{5,3} - d_{2,0}^{p}d_{2,1}^{p^2}d_{5,4} + d_{2,0}^{p}d_{2,1}^{p^2+p^3} - d_{2,0}^{p+p^3}$. *This equation provides a bound for the top degree of $d_{2,1}$. We shall also find bounds for the monomial $d_{2,0}^{i}d_{2,1}^{j}$.*
$d_{5,0} = d_{2,0}d_{5,2} - d_{2,0}d_{2,1}^{p}d_{5,3} - d_{2,0}d_{2,1}^{p+p^2}d_{5,4} - d_{2,0}d_{2,1}^{p+p^2+p^3} + d_{2,0}^{1+p^3}d_{2,1}^{p} - d_{2,0}^{1+p^2}d_{5,4} + d_{2,0}^{1+p^2}d_{2,1}^{p^3}$. *We use the last equation for the monomial $d_{2,0}d_{2,1}^{p+p^2+p^3}$.* $d_{5,1}d_{2,0} - d_{5,0}d_{2,1} = -2\,d_{2,0}^{1+p^3}d_{2,1}^{p+1} + d_{2,0}^{1+p}d_{5,3} - d_{2,0}^{1+p}d_{2,1}^{p^2}d_{5,4} + d_{2,0}^{1+p}d_{2,1}^{p^3+p^2} - d_{2,0}^{1+p+p^3}$. *The following set provides a free basis $\{ d_{2,0}^{j_0}d_{2,1}^{j_1} \,|\, 0 \le j_0 \le p^3 + p$, if $0 \le j_1 < p^3 + p^2 + p$; $0 \le j_1 \le p^3 + p^2 + p$ if $j_0 = 0\}$. The cardinality of the last set is $X\begin{bmatrix}5\\2\end{bmatrix}$.*

Proposition 29 *Let $n_2 = l\,\mathrm{lcm}(2, ..., n_1) + 1$. Then*

$$B_{D_n}(\mathbb{F}_p(n_1, n_2)) = \{1\} \cup \left\{ \prod_{t=i}^{n_1-1} d_{n_1,t}^{m_t} \,|\, 1 \le m_i \le A_{n_1-i} \text{ and } 0 \le m_t < A_{n_1-t}, t > i \right\}_{i=0}^{n_1-1}$$

is a free module basis for $\mathbb{F}_p(n_1, n_2)$ over D_n.

Proof. Let us recall that the given hypothesis has been studied in Lemma 21 and the Proposition after as well as the number $\left[{n_2+n_1 \atop n_1}\right]$. Because the cardinality of $B_{D_n}(\mathbb{F}_p(n_1,n_2))$ is the right one, we only have to show that this set is actually a generating set. Let us explain how the claimed set has been deduced. Combining the facts that in this case $t|n_2+t-1$ and Corollary 22 a) and d) we have two families of relations: i) $d_{n_1,i}d_{n_1,n_1-1}^{A_1}$ is a summand in the decomposition of $d_{n,i}$. ii) $d_{n_1,i}d_{n_1,i+t}^{A_{n_1-i-t}}$ is contained in $d_{n,i+t+1}d_{n_1,i}-d_{n,i}d_{n_1,i+t+1}$ for $t\geq 1$. We shall also show that $d_{n,i+t+1}d_{n_1,i}$ and $d_{n,i}d_{n_1,i+t+1}$ have only one monomial in common. So let $d_{n_1,i+t+1}d^m$ be such a monomial in $d_{n,i+t+1}$. Then $k=p^n-p^{n_1}=p^{n_1}(p^{n_2}-1)$. Thus a factor of the degree of possible elements which is not a p-th power must divide $(p^{n_2}-1)$. Because $n_2=l\,\mathrm{lcm}(2,...,n_1)+1$ and that divisor is a product of $(p^{n_1-t}-1)$, we conclude that there is only such element, namely $d_{n_1,n_1-1}^{A_{n_1-1}}$. To prove that the claimed set satisfies the required property we use double induction on, the total degree, $\sum m_i(p^{n_1}-p^i)$, of a typical monomial $\prod_{i=0}^{n_1-1} d_{n_1,i}^{m_i}$ and n_1. For $n_1=1$, this case is an application of the last Proposition. Since

$$d_{n,0}=d_{1,0}d'_{n,1}=d_{1,0}(d_{n,1}+\sum_{t=1}^{n-1}(-1)^t d_{n,1+t}d_{1,0}^{p+...+p^t})$$

$d_{1,0}^{1+p+...+p^t}$ decomposes with respect to the given basis. Let $f=\prod_{i=0}^{n_1-1} d_{n_1,i}^{m_i}$. Then $\varepsilon_{(1)}\cdot sh_{(-1)}\cdot\pi_{y_1}(f)-f\in(d_{n_1,0})$ (please see lemma 23). Let $g^p=sh_{(-1)}\cdot\pi_{y_1}(f)$ and we decompose g in $\mathbb{F}_p(n_1-1,n_2)$ by induction and the remark 2. Let us apply $\varepsilon_{(1)}$ on the p-th power of the last decomposition and call that element $g(n)$. Then $g(n)$ fulfils the requirements of our basis. Thus $f-g(n)=d_{n_1,0}h$ and by induction h can be decomposed as h'. Finally, $d_{n_1,0}h'$ decomposes according to our relations regarding $d_{n_1,0}^{A_{n_1}+1}$ and $d_{n_1,0}d_{n_1,i}^{A_{n_1-i}}$. ■

We combine the last two Propositions by applying our method in the case $n=11$ and $n_1=5$.

Example 30 *Let $n=11$ and $n_1=5$. We compute A_i for $i=5,..,1$. $I_1=\{1\}$, $I_2=\{5,4,2\}$ and $I_3=\{3\}$. $A_5=p(p^5+1)$, $A_4=p^2(p^4+1)$, $A_3=p^6+p^3$, $A_2=p^6+p^4+p^2$, and $A_1=p^6+...+p$. Now a free basis $B_{D_{11}}(\mathbb{F}_p(5,6))$ follows:*
$\{d_{5,2}^{m_2}d_{5,4}^{m_4}|0\leq m_2\leq A_3, 0\leq m_4\leq A_1\}\cup\{d_{5,2}^{m_2}d_{5,3}^{m_3}d_{5,4}^{m_4}|0\leq m_2\leq A_3, 0\leq m_3\leq A_2, 0\leq m_4\leq A_1\}\cup\{d_{5,1}^{m_1}d_{5,2}^{m_2}d_{5,3}^{m_3}d_{5,4}^{m_4}|1\leq m_1\leq A_4, 0\leq m_2\leq A_3, 0\leq m_3\leq$

$A_2, 0 \leq m_4 \leq A_1\} \cup \{d_{5,0}^{m_0} d_{5,1}^{m_1} d_{5,2}^{m_2} d_{5,3}^{m_3} d_{5,4}^{m_4} | 1 \leq m_0 \leq A_5, 0 \leq m_1 \leq A_4, 0 \leq m_2 \leq A_3, 0 \leq m_3 \leq A_2, 0 \leq m_4 \leq A_1\}$.

Theorem 31 *The set* $B' = \left\{ \prod_{i=0}^{n_1-1} d_{n_1,i}^{m_i} \,|\, 0 \leq m_i \leq A_i \right\}$ *is a module basis for* $\mathbb{F}_p(n_1, n_2)$ *over* D_n.

Here A_i is as in definition 17.

Proof. Firstly, we recall that $d_{n_1,i}^{A_{n_1-i+1}}$ decomposes with respect to the claimed basis because of Corollary 22. Let $f = \prod_{i=0}^{n_1-1} d_{n_1,i}^{m_i}$ be a typical element. We use double induction on the total degree of f and n_1. We follow the proof of Proposition 29.

Let $m_0 = 0$, then $sh_{(-1)} \cdot \pi_{y_1}(f) = \prod_{i=0}^{n_1-2} d_{n_1-1,i-1}^{pm_i} = \left(\prod_{i=0}^{n_1-2} d_{n_1-1,i-1}^{m_i} \right)^p = \left(\sum_I f_I \right)^p$ with respect to $B'_{D_{n-1}}(\mathbb{F}_p(n_1 - 1, n_2))$. Now, $sh_{(-1)} \cdot \pi_{y_1}(f) = \sum_I f'_I$ with respect to B'. Thus $f - \sum_I f'_I = d_{n_1,0}h$ and induction hypothesis can be applied.

If $m_0 > 0$, two cases should be considered: i) $A_{n_1} + 1 \geq m_0$ and ii) $A_{n_1} \leq m_0$. For i) we use induction on $\prod_{i=1}^{n_1-1} d_{n_1,i}^{m_i}$ and for ii) the relation $d_{n_1,0}^{A_{n_1}+1} = d_{n_1,0} d_{n,i(0)} - d_{n_1,i(0)} d_{n,0} -$ "*others*". ■

4 The transfer between parabolic subgroups

The main results in this section are Theorems 41 and 44. Let us recall part of material which has been appeared in [4].

Let H be a subgroup of a finite group G, then the inclusion $i : H \hookrightarrow G$ induces the transfer map $tr^* : H^*(H,\mathbb{F}_p) \longrightarrow H^*(G,\mathbb{F}_p)$ given by $tr^*(u) = |G : H|^{-1} \sum_{g \in G/H} gu$. If $W_G(H)$ is the Weyl subgroup, then the inclusion above induces $i^* : H^*(G,\mathbb{F}_p) \longrightarrow H^*(H,\mathbb{F}_p)^{W_G(H)}$.

$\sum_{p^n}$ acts on V and if we regard $(\mathbb{F}_p)_i$ as the subgroup of translations in the i-th component, then $\sum_{p^n,p} := (\mathbb{F}_p)_1 \int \cdots \int (\mathbb{F}_p)_n$ is a p-Sylow subgroup of $\sum_{p^n}$. We can regard ${\mathbb{F}_p}^n$ as the subgroup of all translations of V in $\sum_{p^n}$. The Weyl subgroups of ${\mathbb{F}_p}^n$ in $\sum_{p^n;p}$, $\sum_{p^{n_1}} \int \sum_{p^{n_2}}$ and $\sum_{p^n}$ respectively are the upper triangular group U_n, $P(n_1, n_2)$ and the general linear group $GL(n,p)$. The induced inclusion $W_{\sum_{p^n,2p}}(\mathbb{F}_p^n) \to W_{\sum_{p^{n_1}} \int \sum_{p^{n_2}}}(\mathbb{F}_p{}^n) \to W_{\sum_{p^n}}(\mathbb{F}_p^n)$ induces $H^*(\mathbb{F}_p^n, \mathbb{F}_p)^{U_n} \xrightarrow{\tau^*} H^*(\mathbb{F}_p^n, \mathbb{F}_p)^{P(n_1,n_2)} \xrightarrow{\tau^*} H^*(\mathbb{F}_p^n, \mathbb{F}_p)^{GL(n,p)}$ given by $\tau^*(f) = \sum_{g \in G/H} gf$. Here f is a U_n or $P(n_1, n_2)$-invariant polynomial in $\mathbb{F}_p[V]$.

Here V is a $\mathbb{F}_pG$-module. In our case the transfer is surjective and $\mathbb{F}_p[V]^{GL(n,p)}$ is a direct summand.

The following diagram is commutative [5].

$$\begin{array}{ccccc} H^*\left(\sum_{p^n,p}\right) & \xrightarrow{\tau^*} & H^*\left(\sum_{p^{n_1}} \int \sum_{p^{n_2}}\right) & \xrightarrow{\tau^*} & H^*\left(\sum_{p^n}\right) \\ \uparrow & & \uparrow & & \uparrow \\ H^*(\mathbb{F}_p^n, \mathbb{F}_p)^{U_n} & \xrightarrow{\tau^*} & H^*(\mathbb{F}_p^n, \mathbb{F}_p)^{P(n_1,n_2)} & \xrightarrow{\tau^*} & H^*(\mathbb{F}_p^n, \mathbb{F}_p)^{GL(n,p)} \end{array} \qquad (1)$$

Campbell and Hughes ([1]) have studied the transfer for the case:

$$\tau^* : H_n \to D_n$$

We recall from [1] a set of coset representatives for GL_n over U_n.

$\mathbb{F}_p^n$ is the $(p^n - 1)$-st cyclotomic field over $\mathbb{F}_p$ and let σ_n be a primitive $(p^n - 1)$-st root of unity over $\mathbb{F}_p$ of an irreducible factor of the $(p^n - 1)$-st cyclotomic polynomial $x^{p^n-1} - 1$. Then $\mathbb{F}_p^{n*} =< \sigma_n >$. Let $\Phi : \mathbb{F}_p^{n*} \cong V - \{\bar{0}\} \to< \sigma_n >$ be the natural correspondence. σ_n acts linearly on V as follows: $\sigma_n \bar{0} = \bar{0}$ and $\sigma_n u = \sigma_n^{k+1}$ where $\Phi(u) = \sigma_n^k$. Note that the fixed point set of σ_n^m is the zero vector set.

Let $W^{n-m+1} = \langle y_{n-m+1}, \cdots, y_n \rangle$, then $(W^{n-m+1})^* = \langle \sigma_m \rangle$ as a multiplicative group and we denote the corresponding linear transformation again by the same symbol as above.

Theorem 32 *[1]The set* $\{\sigma_n^{i_n} \cdots \sigma_1^{i_1} \mid 0 \leq i_m < p^m - 1\}$ *is a set of left coset representatives for* GL_n *over* U_n.

We generalize the previous theorem for parabolic subgroups. Let $\Im_n(p-1)$ be the set of coset representatives of GL_n over B_n:

$$\{< \sigma_n > / < \sigma_n^{l_{p-1}} > \text{ such that } | < \sigma_n^{l_{p-1}} > | = p - 1\}$$

Note that this set is induced by $\mathbb{F}_p^{n*} \cong< y_1, ..., y_n > -\{\bar{0}\}$. $|\Im_n(p-1)| = \frac{p^n-1}{p-1}$.

Theorem 33 *i) The set* $\{\sigma_n^{i_n}\cdots\sigma_1^{i_1} \mid 0\le i_m < p-1\}$ *is a set of left coset representatives for* B_n *over* U_n. *Here* $<\sigma_i>\cong<y_i>-\{\bar{0}\}$.

ii) The set $\{\sigma_n^{i_n}\cdots\sigma_1^{i_1} \mid 0\le i_m < \frac{p^m-1}{p-1}\}$ *is a set of left coset representatives for* GL_n *over* B_n. *Here* $\sigma_m^{i_m}\in\Im_m(p-1)$ *induced by* $<y_{n-m+1},\cdots,y_n> -\{\bar{0}\}$.

iii) Let $N=\{n_1,\cdots,n_\ell\}$ *and* $\nu_t=n_1+\cdots+n_t$. *The set* $\{\prod_{t=1}^{\ell}\sigma_{1,t}^{i_{1,t}}\cdots\sigma_{n_t,t}^{i_{n_t,t}} \mid 1\le s\le n_t,\ 0\le i_{s,t}<\frac{p^s-1}{p-1}\}$ *is a set of left coset representatives for* $P(N,n)$ *over* B_n. *Here* $\sigma_{k_t,t}^{i_{k_t,t}}\in\Im_{k_t,t}(p-1)$ *induced by* $<y_{\nu_t-k_t+1},\cdots,y_{\nu_t}>-\{\bar{0}\}$.

Proof. We use the following embedding:

$$\begin{pmatrix}1 & 0\\ 0 & GL_k\end{pmatrix}\hookrightarrow\begin{pmatrix}GL_{k+1}\end{pmatrix}\text{ and }\begin{pmatrix}1&0&0&0\\0&\ddots&&\vdots\\ \vdots&&1&0\\0&\cdots&0&GL_{n_\ell}\end{pmatrix}\hookrightarrow\begin{pmatrix}GL_{n_{\ell-1}}&0\\0&GL_{n_\ell}\end{pmatrix}$$

We recall that $|GL_n|=\prod_{i=1}^{n}p^n-p^{i-1}=(p^n-1)p^{n-1}|GL_{n-1}|$. The idea is to construct a set of representatives for the last block and to extend it step by step until we cover the whole group.

i) We use induction. Let $C(B_{n-1}/U_{n-1})$ be a set of representatives for B_{n-1} over U_{n-1}. Let σ_n be a primitive element in $\langle y_1\rangle^*$. This element has order $p-1$, acts linearly and moves every element of $\langle y_1\rangle$. Since U_n fixes y_1, $\{\sigma_n^{i_n}\mid 0\le i_n<p-1\}C(B_{n-1}/U_{n-1})=C(B_n/U_n)$.

ii) Let $C(GL_{n-1}/B_{n-1})$ be a set of representatives for GL_{n-1} over B_{n-1}. Let σ_n^i be an element in $\Im_n(p-1)$. This element has order p^n-1, acts linearly and moves every element of $\langle y_1,\cdots,y_n\rangle$. Moreover, $\sigma_n^i y_1^{p-1}\ne y_1^{p-1}$. Since B_n fixes (y_1^{p-1}), $\Im_n(p-1)C(GL_{n-1}/B_{n-1})=C(GL_n/B_n)$.

iii) We only have to show that

$$C(\mathbb{F}_p(n_1,n_2)/B_{n_1+n_2})=C(GL_{n_1}/B_{n_1})C(GL_{n_2}/B_{n_2})$$

and then the argument follows by induction on the number of blocks. To prove the claim we use induction on n_1. For $n_1=2$, $\Im_{1,1}(p-1)=\{e\}$ and $\Im_{2,1}(p-1)=\{\sigma_{2,1}^i\mid 0\le i<p+1\}$. Now the claim follows. ∎

The method used above to determine coset representatives for $\mathbb{F}_p(n_1, n_2)$ over B_n can not be applied for GL_n over $\mathbb{F}_p(n_1, n_2)$ for the same reasons as in a free module basis construction of the appropriate rings of invariants.

Proposition 34 *Let* $GL_n = \bigsqcup_{i=1}^{|GL_n:B_n|} \sigma_i B_n$ *and* $\mathbb{F}_p(n_1, n_2) = \bigsqcup_{j=1}^{|\mathbb{F}_p(n_1,n_2):B_n|} \tau_j B_n$ *as constructed before. Then* $GL_n = \bigsqcup_{k=1}^{|GL_n:\mathbb{F}_p(n_1,n_2)|} u_k B_n$ *where* $u_k = \sigma_{i(k)}$ *such that* $\sigma_{i(k)} \notin \mathbb{F}_p(n_1, n_2)$ *and* $\sigma_{i(k)} \notin \sigma_{i(j)}\mathbb{F}_p(n_1, n_2)$ *for* $k > j$.

Definition 35 *Let the trace over the matrix* σ_m *be defined by* $Tr_m(-) = (p^m - 1)^{-1} \sum_{i=1}^{p^m-1} \sigma_m^i(-)$ *and* $Tr_{m,t}^{(p-1)}(-) = (\frac{p^m-1}{p-1})^{-1} \sum_{i=0}^{\frac{p^m-1}{p-1}} \sigma_{m,t}^{(p-1)i}(-)$.

Then $\tau^*(f) = \prod_m Tr_m(f)$ for H_n over G and $\tau^*(f) = \prod_t \prod_m Tr_{m,t}^{(p-1)}(f)$ for B_n over G. Here G is $\mathbb{F}_p(N, n)$ or GL_n.

Because of the properties of the matrices σ_m and the generators for D_n, $\mathbb{F}_p(N, n)$, B_n, and H_n we conclude:

$$\tau^*(h_1^{i_1} \cdots h_m^{i_m} \cdots h_n^{i_n}) = Tr_n(h_1^{i_1} \cdots Tr_{n-m+1}(h_m^{i_m} \cdots Tr_1(h_n^{i_n}) \cdots) \cdots)$$

Definition 36 *Let* $b_i = h_i^{p-1}$ *and* $TR_{i,t} = Tr_{i,t}^{(p-1)}$.

$$\tau^*(b_1^{i_1} \cdots b_{n_1}^{i_{n_1}} b_{n_1+1}^{i_{n_1+1}} \cdots b_{\nu_{\ell-1}+1}^{i_{\nu_{\ell-1}+1}} \cdots b_{\nu_\ell}^{i_{\nu_\ell}}) = \prod_t TR_{n_t,t}(b_{\nu_{t-1}+1}^{i_{\nu_{t-1}+1}} \cdots TR_{1,t}\left(b_{\nu_t}^{i_{\nu_t}}\right) \cdots).$$

Let $\xi : H_n \longrightarrow D_n$ be the natural $\mathbb{F}_p$-epimorphism. Campbell and Hughes have shown that the map ξ is actually the induced transfer map between U_n and GL_n-invariants. We showed that the same is true for $\xi : H_n \longrightarrow \mathbb{F}_p(n_1, n_2)$ in [3].

Next we consider the map above for parabolic subgroups: $\xi : \mathbb{F}_p(n_1, n_2) \longrightarrow D_n$.

The next lemma plays a key role in the proof of the main theorem of this section.

Lemma 37 $\sum_{u \in V} u^r \begin{cases} = 0 & if\ r \neq 0 \mod(p^n - 1) \\ \neq 0 & if\ r = 0 \mod(p^n - 1) \end{cases}$ *in* $\mathbb{F}_p[V]$.

Proof. Let $u \in V^*$. Since $|V^*| = (p-1)(p^{n-1}+\cdots+1)$, we write its elements as follows:

$\{a_i(y_i+v_i) \mid a_i \in \mathbb{F}_p{}^*,\ v_i \in < y_{i-1}, \cdots, y_1 >\}$. We use induction on n. For $n=1$, $\sum a^{(p-1)k} y_1^{(p-1)k} = (p-1)y_1^{(p-1)k}$. If $r \neq (p-1)k$, then $\sum a^r = 0$. Let $r = 0 \mod(p^n-1)$, then $a_i^r = 1$. Hence, $\sum_{a_n} a_n^r(y_n+v_n)^r = (p-1)(y_n+v_n)^r$. $\sum_{v_n}(y_n+v_n)^{(p^n-1)} = \sum_{v_n}\sum y_n^{p^n-1-t}v_n^t$. The last sum contains the term $\sum_{v_n} y_n^{p^n-p^{n-1}} v_n^{p^{n-1}-1}$ which is non-zero by induction. The second claim follows. For the first claim, if $r \not\equiv 0 \mod(p-1)$, then $\sum_{a_i} a_i^r(y_i+v_i)^r = 0$. Otherwise, $\sum_{a_i} a_i^r(y_i+v_i)^r = (p-1)(y_i+v_i)^r$. Now, $\sum_{u\in V} u^r = (p-1)\sum_{i,v_i}(y_i+v_i)^r = \sum_{v_n}(p-1)(y_n+v_n)^r + \nu_n^r$. Let $y_n = 0$, then our sum is zero and hence divisible by y_n. On the other hand, this sum is a GL_n-invariant. Thus, this sum is divisible by $d_{n,0}$. ■

Remark 4 *Let $r = (p-1)k \not\equiv 0 \mod(p^n-1)$. By lemma above, $\sum_{u\in V} u^r = (p-1)\sum_{i,v_i}(y_i+v_i)^r = 0$. Hence $\sum_{i,v_i}(y_i+v_i)^{(p-1)k} = 0$.*

Lemma 38 *$TR_{n_t-m+1,t}(b^{i_m}_{\nu_{t-1}+m}) = 0$ for $b^{i_m}_{\nu_{t-1}+m} = h^{(p-1)i_m}_{\nu_{t-1}+m}$ in $B_{\mathbb{F}_p(N,n)}\left(\mathbb{F}_p[V]^{B_n}\right)$.*

Proof. By definition, $TR_{n_t-m+1,t}(b^{i_m}_{\nu_{t-1}+m}) = \sum_{s=0} \sigma^s_{n_t-m+1,t}(b^{i_m}_{\nu_{t-1}+m})$. Here $0 \leq i_m < \frac{p^m-1}{p-1}$, $1 < m \leq n_t$. For the rest of the proof, let us simplify the coefficients $\sigma^s_{n_t-m+1} \to \sigma^s_m$ and $b^{i_m}_{\nu_{t-1}+m} \to b^{i_m}_m$. We recall that σ^s_m is a primitive element in $< y_{\nu_{t-1}+m}, \cdots, y_{\nu_t} >^*$ and fixes y_t for $t < \nu_{t-1}+m$. Recall that $h_m(y_m) = y_m^{p^{m-1}} + \sum(-1)^i d_{m-1,i} y_m^{p^i}$, h_m is linear in y_m. Now, we evaluate the sum in the argument. $\sum_{s=0}\sigma^s_m(h_m^{(p-1)i_m}) = \sum_{s=0}\left(h_m^{(p-1)i_m}(\sigma^s_m y_m)\right) = \sum_{i,u_i}\left(h_m^{(p-1)i_m}(y_i+u_i)\right)$. Here $i = \nu_{t-1}+m, ..., \nu_t$ and $u_i \in < y_{\nu_{t-1}+m}, \cdots, y_i >$. The last sum is zero, because of remark 4. ■

A direct application of the last lemma and remark 35 is the next theorem.

Theorem 39 *i) Let $h \neq 1$ be an element of the set $B_{\mathbb{F}_p[V]^{B_n}}(H_n)$. Then $\tau^*(h) = 0$.*

ii) Let $h(p-1) \neq 1$ be an element of the set $B_{\mathbb{F}_p(N,n)}\left(\mathbb{F}_p[V]^{B_n}\right)$. Then $\tau^*(h(p-1)) = 0$.

The next theorem follows from last.

Theorem 40 *i) Let* $\xi : H_n \longrightarrow \mathbb{F}_p[V]^{B_n}$ *be the natural* $\mathbb{F}_p$*-epimorphism. Then* $\xi = \tau^*$.

ii) Let $\xi : \mathbb{F}_p[V]^{B_n} \longrightarrow \mathbb{F}_p(N, n)$ *be the natural* $\mathbb{F}_p$*-epimorphism. Then* $\xi = \tau^*$.

The obvious step at this point is to extend the statements above to $\tau^* : \mathbb{F}_p(n_1, n_2) \to D_n$. Let B denote the set $B_{D_n}(\mathbb{F}_p(n_1, n_2))$.

Theorem 41 *Let* $\xi : \mathbb{F}_p(n_1, n_2) \longrightarrow D_n$ *be the natural* $\mathbb{F}_p$*-epimorphism with respect to basis* B*. Then* $\xi = \tau^*$.

Proof. The statement follows from the following commutative diagram

$$\begin{array}{ccc} \mathbb{F}_p[V]^{B_n} & \overset{\tau=\xi}{\rightarrow} & \mathbb{F}_p(n_1, n_2) \\ & \searrow^{\tau=\xi} & \downarrow \\ & & D_n \end{array}$$

Here τ and ξ refer to the right map. ■

Kuhn and Priddy have examined a multiplicative property of the transfer between certain subgroups of the symmetric group in [5]. We examine an analogy between certain rings of invariants.

Lemma 42 *Let* h^I *be an element of the set* $B_{D_n}(H_n)$*. Then* $\exists h^J$*, an other basis element, such that* $\tau^*(h^{I+J}) \neq 0$.

Proof. We provide two different types of elements.
i) Let $I = (i_1, ..., i_n)$, we define $J = (j_1, ..., j_n)$ such that

$$j_t = \begin{cases} p^{n-t+1} - 1 - i_t, i_t \neq 0 \\ 0, i_t = 0 \end{cases}$$

Claim: $\tau^*(h^{I+J}) = \prod_{t=0}^{n-1} d_{n,t}^{\varepsilon_t}$ where $\varepsilon_t = 1$ if $i_t \neq 0$ and 0 otherwise. We proceed by induction on the non-zero indexes of I. It is known that $h_t^{p^{n-t+1}-1} = d_{n,t-1}$+"$others$" ([3]). Let us recall that this is the way the basis $B_{D_n}(\mathbb{F}_p[V]^{B_n})$ was constructed. Moreover, $\tau^*(f) = 0$ for all monomials f involved in

"*others*". Thus $h^{I+J} = \prod_{t=0}^{n-1} d_{n,t}^{\varepsilon_t} +$ "*others*" and $\tau^*(f) = 0$ for all monomials f involved in "*others*".
ii) There also exists an other element $h^{J'}$ of degree less or equal than the degree of h^J defined above. Let $l = \max\{t | i_t \neq 0\}$ and $\{i_{s_1}, ..., i_{s_r} = l\}$ be the non-zero elements of I. Define

$$j'_t = \begin{cases} p^{n-l+1} - 1 - i_l \\ p^{n-i_{s_t}+1} - 1 - i_{s_t} \text{ for } t < r, i_{s_t} > p^{i_{s_{t+1}} - i_{s_t}}(p^{n-i_{s_{t+1}}+1} - 1) \\ p^{i_{s_{t+1}} - i_{s_t}}(p^{n-i_{s_{t+1}}+1} - 1) - i_{s_t} \text{ for } t < r, 0 < i_{s_t} \leq p^{i_{s_{t+1}} - i_{s_t}}(p^{n-i_{s_{t+1}}+1} - 1) \\ 0 \text{ for } i_{s_t} = 0 \text{ or } t > l \end{cases}$$

Then $\tau^*(h^{I+J}) \neq 0$. This is because
$h_{i_{s_{t+1}}}^{p^{n-i_{s_{t+1}}+1}-1} = d_{n,i_{s_{t+1}}-1} - h_{i_{s_t}}^{p^{i_{s_{t+1}}-i_{s_t}}(p^{n-i_{s_{t+1}}+1}-1)} +$ "*others*". ■

Proposition 43 *Let* $h = \sum_{h^I \in B_{D_n}(H_n)} h^I$, *then* $\tau^*(hf) \neq 0$ *for any non-zero element of* H_n.

Proof. We apply last lemma for f a monomial of the form $h'd$ where h' is a basis element and $d \in D_n$. Then we extend linearly for any element of H_n. ■

Remark 5 *Lemma and Proposition above extend easily to the case* $B_{D_n}\left(\mathbb{F}_p[V]^{B_n}\right)$.

Theorem 44 *Let* $\tau^* : \mathbb{F}_p[V]^G \to D_n$ *be the transfer map, where* $G = U_n$ *or* B_n. *Let* $x, y \in H_n$ *or* $\mathbb{F}_p[V]^{B_n}$. *Then* $\tau^*(xy) = \tau^*(x)\tau^*(y)$ *for all* y *iff* $x \in D_n$.

Proof. Let $\tau^*(x) = z$ and $\tau^*(y) = w$. Now decompose x and y with respect to the given basis, then $x = z + h_x$ and $y = w + h_y$. Hence $\tau^*(xy) = zw + \tau^*(h_x h_y)$. The last implies that $\tau^*(h_x h_y) = 0$ and $\tau^*(h_x) = 0$. We shall show that $h_x = 0$. Assume that $h_x \neq 0$. Choose $h_y = \sum_{h^I \in B_{D_n}(H_n)} h^I$ then $h_x h_y = d_{xy} + h_{xy}$ and $\tau^*(h_{xy}) = 0$. Thus $\tau^*(h_x h_y) = d_{xy}$. ■

References

[1] H. E. A. Campbell and I. P. Hughes, The ring of upper triangular invariants as a module over the Dickson algebra. Mathematishes Annalen. Volume 306 (1996), 429-443.

[2] L. E. Dickson, A fundamendal system of invariants of the general modular linear group with a solution of the form problem. Trans. A. M. S. 12 (1911), 75-98.

[3] Nondas E. Kechagias, The transfer between rings of modular invariants of subgroups of GL(n,p). The Fields Institute Communications of the A. M. S.. Stable-Unstable Homotopy. S. Kochman and R. Selick, editors. Volume **19** (1998), pp. 165-180.

[4] Nondas E. Kechagias, Modular invariant theory, homology operations, and the transfer between rings of invariants of parabolic subgroups. Bulletin of the Greek Mathematical Society. Volume 38 (1996), pp. 79-88.

[5] N. J. Kuhn, Chevalley group theory and the transfer in the homology of symmetric groups. Topology Vol 24, No. 3. pp 247-264. 1985.

[6] N. J. Kuhn and S. B. Priddy, The transfer and Whitehead's conjecture. Math. Proc. Camb. Phil. Soc. (1985) 98, 459-480.

[7] Huyhn Mui, Modular invariant theory and the cohomology algebras of the symmetric groups, J. Fac. Sci. Univ. Tokyo, IA (1975), 319-369.

[8] L. Smith, Polynomial invariants of finite groups. Research Notes in Mathematics, vol. 6. A. K. Peters.

[9] C. W. Wilkerson, A primer on Dickson invariants, Cont. Math. 19 (1983), 421-434.

E-MAIL: NKECHAG@CC.UOI.GR

Oscillation Properties of First Order Neutral Differential Equations Near the Critical States

Y. Domshlak, N. Partsvania, and I. P. Stavroulakis

Department of Mathematics, Ben-Gurion University of the Negev, 84105 Beer-Sheva, Israel,
E-mail: domshlak@indigo.bgu.ac.il

A. Razmadze Mathematical Institute of the Georgian Academy of Sciences,
1 M. Aleksidze St., 380093 Tbilisi, Georgia,
E-mail: ninopa@rmi.acnet.ge

Department of Mathematics, University of Ioannina, 451 10 Ioannina, Greece,
E-mail: ipstav@cc.uoi.gr

1 Introduction

In the present paper we consider the first order NDE with the almost constant coefficient

$$\ell[y] := y'(t) - py'(t-1) + Q(t)y(t-\sigma) = 0, \quad t \geq t_0, \tag{1}$$

where $\sigma \geq 0$, $p = \text{const}$, and

$$\lim_{t\to\infty} Q(t) = q. \tag{2}$$

Its "limiting" equation is

$$x'(t) - px'(t-1) + qx(t-\sigma) = 0, \quad t \geq 0. \tag{3}$$

Eq.(1) some times (in the critical situations) does not inherit the oscillation properties of Eq.(3). For example, in the case $\{p = 0;\ q = \frac{1}{\sigma e}\}$ these properties depend on the character of the convergence (2) (see [5, Corollary 4.5] and [7]; compare also with the results in [2] and [3]).

The situation $\{p = 1;\ q = 0\}$ is the same. It was considered in Theorem 3.2.4 and Corollary 3.2.2 in [8] and in [1] as well. However, the case $\{p = +1;\ q = 0\}$ is special: the solutions' behavior for Eq.(1) in this situation is rather like the second order ordinary differential equations (see [1]). Other critical situations for Eq.(1) have not been studied at all.

It is to be emphasized that this is the first paper dealing with this problem in the general form.

2 The main result and some preliminary remarks

Let

$$\begin{cases} p > \dfrac{1-\sigma}{e(2-\sigma)} & \text{for } 0 \leq \sigma < 2 \\ p \text{ be arbitrary} & \text{for } \sigma \geq 2 \end{cases} . \tag{4}$$

Then the equality

$$p = e^{-s}\,\frac{1-s\sigma}{1-s(\sigma-1)} \tag{5}$$

defines the unique real number s,

$$\begin{cases} -\dfrac{1}{1-\sigma} < s < 1 & \text{for } 0 \le \sigma < 1 \\ -\infty < s < 1 & \text{for } 1 \le \sigma \le 2 \\ -\infty < s < \dfrac{1}{\sigma-1} & \text{for } \sigma > 2 \end{cases}. \tag{6}$$

Denote

$$h[\sigma;s] := \sigma(\sigma-1)s^2 + (1-2\sigma)s + 2,$$

$$\mathcal{K}[\sigma;s] := \frac{e^{-s\sigma}h[\sigma;s]}{8[1-s(\sigma-1)]} \quad \text{and} \quad q := \frac{e^{-s\sigma}s^2}{1-s(\sigma-1)}. \tag{7}$$

It is easy to check that $h[\sigma;s] > 0$ in domains (6).

The following statement can be called as "Kneser-like Theorem" for the first order neutral differential equation.

Theorem 1 *Let p and $s < 1$ be defined by* (4) *and* (5)–(6). *Assume*

$$\liminf_{t\to\infty} t^2\big[Q(t) - q\big] = C > \mathcal{K}[\sigma;s]. \tag{8}$$

Then:

1^0) *all solutions of Eq.*(1) *are oscillatory;*

2^0) *any solution of Eq.*(1) *has at least one zero on each interval* $(T-1; T\exp\frac{\pi}{\nu})$ *for sufficiently large T and*

$$4\nu^2 < \frac{C}{\mathcal{K}} - 1. \tag{9}$$

Remark 1 This result is sharp in the sence that the strict inequality in (8) cannot be replaced by the non-strict one. Indeed, define

$$y_0(t) := \sqrt{t}\cdot e^{-st}, \quad p = e^{-s}\,\frac{1-s\sigma}{1-s(\sigma-1)}, \quad Q(t) := \frac{py_0'(t-1) - y_0'(t)}{y_0(t-\sigma)}.$$

One can easily check that

$$\lim_{t\to\infty} t^2\Big[Q(t) - \frac{e^{-s\sigma}\cdot s^2}{1-s(\sigma-1)}\Big] = \mathcal{K}[\sigma;s]. \tag{10}$$

On the other hand, Eq.(1) has a non-oscillatory solution $y_0(t)$.

Remark 2 Consider two important particular cases:

1. $\{p = 1;\ \sigma \ge 0\}$. Then $s = 0$, $q = 0$, $\mathcal{K}[\sigma;0] = \frac{1}{4}$, and the first statement of Theorem 1 turns to the known result from [1] for the equation

$$x'(t) - x'(t-1) + Q(t)x(t-\sigma) = 0.$$

The second statement of Theorem 1 is new even for this particular case;

2. $\{p = 0;\ \sigma > 0\}$. Then $s = \frac{1}{\sigma}$, $q = \frac{1}{e\sigma}$, $\mathcal{K}[\sigma;\frac{1}{\sigma}] = \frac{\sigma}{8e}$, and Theorem 1 turns to the known result from [7] for the equation

$$x'(t) + Q(t)x(t-\sigma) = 0.$$

3 Critical states of the autonomous first order NDE

In this section we present for the first time the complete description of the critical states of Eq.(3) with respect to its oscillation properties.

Definition 1 We say that NDE (3) is in the *critical state* with respect to its oscillation properties if there exists at least one eventually positive solution of Eq.(3), while the equation

$$z'(t) - pz'(t-1) + (q+\varepsilon)z(t-\sigma) = 0, \quad t \geq t_0$$

has oscillatory solutions only $\forall \varepsilon > 0$. The pair $\{p; q\}$ will be called a *critical pair.*

It is well-known (see, for example, [9]) that all solutions of Eq.(3) are oscillatory if and only if its characteristic equation

$$F(s) := F(\{p;q\}, s) = -s + spe^s + qe^{s\sigma} = 0 \tag{11}$$

has no real roots.

This fact gives us the possibility to discern all critical pairs for Eq.(3). Indeed, in case the pair $\{p; q\}$ is critical,

$$\exists\, \overline{s} \in (-\infty, \infty) : \;\; F(\overline{s}) = 0, \;\; F'(\overline{s}) = 0. \tag{12}$$

Remark 3 The inverse statement is not true. The pair $\{p; q\}$ will be critical if $F(\{p; q+\varepsilon\}; s) > 0$ $\forall s \in (-\infty, \infty)$ and $\forall \varepsilon > 0$ only. Therefore, among a few pairs $\{p; q_i\}$, $i = 1, 2, \ldots, m$ given rise to the solvable system (12), the pair $[\{p; \overline{q}\} : \; \overline{q} = \max_i q_i]$ only will be critical.

Remark 4 Note that the pair $\{p; q\}$, $q < 0$ can not be critical because $F(0) = q < 0$ and $\lim_{s \to -\infty} F(s) = +\infty$. Therefore, we suppose $q \geq 0$ in (2)–(3) from the beginning.

Let the system (12) be solvable with the solution s (maybe, not unique!). Then

$$\begin{cases} F(s) = -s + spe^s + qe^{s\sigma} = 0 \\ F'(s) = -1 + p(1+s)e^s + q\sigma e^{s\sigma} = 0 \end{cases} \Longleftrightarrow \begin{cases} p = e^{-s}\dfrac{1 - s\sigma}{1 - s(\sigma - 1)} \\ q = e^{-s\sigma}\dfrac{s^2}{1 - s(\sigma - 1)} \end{cases}. \tag{13}$$

The system (13) can be considered as the parametric representation of all pairs $\{p; q\}$ so that the system (12) is solvable. In view of Remark 4 we must restrict the interval of the parameters in (13) by

$$1 - s(\sigma - 1) > 0 \Longleftrightarrow \begin{cases} -\dfrac{1}{1-\sigma} < s & \text{for } 0 \leq \sigma < 1 \\ -\infty < s < \infty & \text{for } \sigma = 1 \\ s < \dfrac{1}{\sigma - 1} & \text{for } \sigma > 1 \end{cases}. \tag{14}$$

Therefore the possible cases of all critical pairs $\{p; q\}$ of Eq.(3) look as follows:

a) Case $\sigma = 0$. There exists a critical pair (13) for $p > 0$. No critical pairs for $p \leq 0$. All points of the curve (13) present a critical pair;

b) Case $0 < \sigma < 1$. There exists a critical pair (13) for $p \geq p(s_0) = \frac{-4\sigma^2}{[1+\sqrt{1+4\sigma(1-\sigma)}]^2} e^{-s_0}$, where

$$s_0 := \frac{4}{2\sigma - 1 + \sqrt{1 + 4\sigma(1-\sigma)}}$$

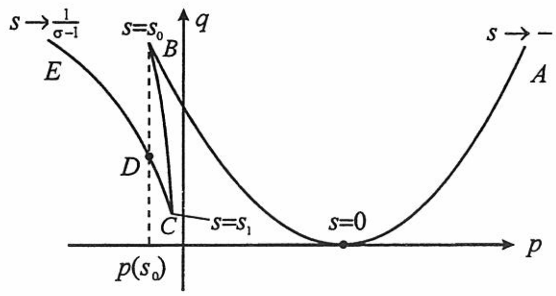
is the smallest root of the equation

$$h[\sigma; s] = 0. \tag{15}$$

No critical pairs for $p < p(s_0)$. The piece of the curve (13) coresponding to $s > s_0$ does not present any critical pair. The point $\{p(s_0), q(s_0)\}$ is an inflection point of the curve (13);

c) Case $\sigma = 1$. The analogous situation with $s_0 = 2$ and $p \geq -\frac{1}{e^2}$.

d) Case $1 < \sigma < \frac{1}{2}(1+\sqrt{2})$. The set $(AB) \cup (CD)$ of the critical pairs $\{p; q\}$ is not connected! There exists a critical pair for any p. The number s_1 is a second root of Eq.(15). The points B and C are the inflection points of the curve (13).



e) The case $\sigma \geq \frac{1}{2}(1+\sqrt{2})$. There exists a critical pair for any p. The set of all critical pairs is connected again.

4 Auxiliary results from Sturmian Comparison method for NDE

Further investigations are based on our approach elaborated in [4] and [6]. Below we state two basic results from [6] in an easier and convenient form (see Lemma 1 together with Theorem 2 from [6] for Theorem A and Lemma 2 together with Theorem 2 from [6] for Theorem B).

Theorem A (for the case $\sigma \neq 1$) *Let $\varphi(t) > 0$ and $k(t)$ be continuous functions such that for sufficiently large t,*

$$\int^{\infty} \varphi(\xi)\, d\xi = \infty, \quad \int_t^{t+\rho} \varphi(\xi)\, d\xi < \frac{\pi}{2}, \quad \rho := \max\{1; \sigma - 1\}, \tag{16}$$

$$\left.\begin{array}{ll} -\varphi(t)\operatorname{ctg} \displaystyle\int_{t-1+\sigma}^{t} \varphi(\xi)\, d\xi < k(t) < \varphi(t)\operatorname{ctg} \displaystyle\int_t^{t+1} \varphi(\xi)\, d\xi & \text{for } 0 \leq \sigma < 1 \\ k(t) < \varphi(t)\operatorname{ctg} \displaystyle\int_t^{t+\rho} \varphi(\xi)\, d\xi & \text{for } \sigma > 1 \end{array}\right\}. \tag{17}$$

Define $\widetilde{Q}(t)$ by

$$\widetilde{Q}(t+\sigma-1) := \operatorname{cosec} \int_t^{t+\sigma-1} \varphi(\xi)\, d\xi \cdot \exp\Big(-\int_{t-1}^{t-1+\sigma} k(\xi)\, d\xi\Big) \times$$

$$\times\left[\varphi(t-1)\cos\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi - k(t-1)\sin\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi - p\varphi(t)\exp\int\limits_{t-1}^{t}k(\xi)\,d\xi\right] \tag{18}$$

and define $S(t)$ *by*

$$S(t-\sigma+1) := \sec\int\limits_{t}^{t+\sigma-1}\varphi(\xi)\,d\xi\cdot\exp\Big(-\int\limits_{t-1}^{t+\sigma-1}k(\xi)\,d\xi\Big)\times$$

$$\times\left[\varphi(t-1)\sin\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi + k(t-1)\cos\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi - pk(t)\exp\int\limits_{t-1}^{t}k(\xi)\,d\xi\right]. \tag{19}$$

Assume

$$\widetilde{Q}(t)\ge 0 \tag{20}$$

and

$$\widetilde{Q}(t)\ge S(t)\ \ \textit{for}\ \ t>T. \tag{21}$$

Then if

$$Q(t)\ge\widetilde{Q}(t),\ \ t>T, \tag{22}$$

all solutions of Eq.(1) *are oscillatory and, what is more, any solution has at least one zero on each interval* $(a-\rho, b+1)$ *for* $a>T$*, where* T *is a sufficiently large number and* $\int_a^b\varphi(\xi)\,d\xi=\pi$.

Theorem B (for the case $\sigma=1$) *Let* $\varphi(t)>0$ *be defined as in Theorem* A, $k(t)$ *be a continuous solution of the equation*

$$L(t) := \varphi(t)\cos\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi - k(t-1)\sin\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi - p\varphi(t)\exp\int\limits_{t-1}^{t}k(\xi)\,d\xi = 0,\ \ t>T \tag{23}$$

such that the condition

$$k(t)<\varphi(t)\operatorname{ctg}\int\limits_{t}^{t+1}\varphi(\xi)\,d\xi \tag{24}$$

holds. Define $\widetilde{Q}(t)$ *by the equality*

$$\widetilde{Q}(t) := -pk(t)+\exp\Big(-\int\limits_{t-1}^{t}k(\xi)\,d\xi\Big)\left[\varphi(t-1)\sin\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi + k(t-1)\cos\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi\right]. \tag{25}$$

If

$$Q(t)\ge\widetilde{Q}(t)\ge 0,\ \ t>T, \tag{26}$$

then the statement of Theorem A *holds.*

5 Proof of the main result

The proof of Theorem 1 in case $\sigma \neq 1$ is based on Theorem A where we define

$$\varphi(t) := \frac{\nu}{t}, \quad k(t) := s + \frac{1}{2t} + \frac{z}{t^2}. \tag{27}$$

The number $s < 1$ is defined by (4), (5), (6), the numbers $\nu > 0$ and z will be chosen later.

Notation $f(t) \cong g(t)$ means $f(t) = g(t) + o(t^{-2})$ as $t \to \infty$.

We write down some needed asymptotics:

$$\left.\begin{array}{c}
\frac{t}{\nu}\varphi(t-1) \cong 1 + \frac{1}{t} + \frac{1}{t^2}; \quad \cos\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi \cong 1 - \frac{\nu^2}{2t^2};\\
\cos\int\limits_{t}^{t-1+\sigma}\varphi(\xi)\,d\xi \cong 1 - \frac{\nu^2(\sigma-1)^2}{2t^2};\\
\frac{t}{\nu}\sin\int\limits_{t-1}^{t}\varphi(\xi)\,d\xi \cong 1 + \frac{1}{2t} + \frac{2-\nu^2}{\sigma t^2};\\
\frac{t}{\nu(\sigma-1)}\sin\int\limits_{t}^{t-1+\sigma}\varphi(\xi)\,d\xi \cong 1 + \frac{\sigma-1}{2t} - \frac{(\sigma-1)^2(2-\nu^2)}{\sigma t^2};\\
\cos\int\limits_{t-1}^{t-1+\sigma}\varphi(\xi)\,d\xi \cong 1 - \frac{\nu^2\sigma^2}{2t^2};\\
\frac{t}{\nu\sigma}\sin\int\limits_{t-1}^{t-1+\sigma}\varphi(\xi)\,d\xi \cong 1 - \frac{\sigma-2}{2t} + \frac{2\sigma^2-6\sigma+6-\sigma^2\nu^2}{6t^2};\\
\sec\int\limits_{t}^{t+\sigma-1}\varphi(\xi)\,d\xi \cong 1 + \frac{\nu^2(\sigma-1)^2}{2t^2};\\
\frac{\nu(\sigma-1)}{t}\operatorname{cosec}\int\limits_{t}^{t+\sigma-1}\varphi(\xi)\,d\xi \cong 1 + \frac{\sigma-1}{2t} + \frac{(\sigma-1)^2(2\nu^2-1)}{12t^2};\\
k(t-1) \cong s + \frac{1}{2t} + \frac{1+2z}{2t^2};\\
\int\limits_{t-1}^{t-1+\sigma}k(\xi)\,d\xi = s\sigma + \frac{\sigma}{t} + \frac{2\sigma-\sigma^2+4\sigma z}{4t^2};\\
\exp\int\limits_{t-1}^{t-1+\sigma}k(\xi)\,d\xi \cong e^{s\sigma}\left(1 + \frac{\sigma}{2t} + \frac{4\sigma-\sigma^2+8\sigma z}{8t^2}\right);\\
\exp\left(-\int\limits_{t-1}^{t-1+\sigma}k(\xi)\,d\xi\right) = e^{-s\sigma}\left(1 - \frac{\sigma}{2t} + \frac{-4\sigma+3\sigma^2-8\sigma z}{8t^2}\right).
\end{array}\right\} \tag{28}$$

In view of (6) it is easy to see that the conditions (16)–(17) hold for sufficiently small $\nu > 0$ and for sufficiently large t.

We omit all intermediate calculations and state the final asymptotics for $\widetilde{Q}(t)$ defined by (18) and for $S(t)$ defined by (19):

$$\widetilde{Q}(t+\sigma-1) \cong q + \frac{1}{t^2}\left\{8\mathcal{K}[\sigma;s]\cdot\frac{z}{\sigma-1} + B(s;\sigma)\cdot(1+4\nu^2)\right\}, \tag{29}$$

$$S(t+\sigma-1) \cong q + \frac{1}{t^2}\mathcal{K}[\sigma;s]\cdot(1+4\nu^2), \tag{30}$$

where q is defined in (7).

In view of $C > \mathcal{K} > 0$ one can define $\nu > 0$ and z such that $C > \mathcal{K}(1+4\nu^2)$ and

$$\mathcal{K}(1+4\nu^2) < 8\mathcal{K}\cdot\frac{z}{\sigma-1} + B(s;\sigma)\cdot(1+4\nu^2) < C. \tag{31}$$

(The exact form of the expression $B(s;\sigma)$ is not essential).

Due to (31) we obtain

$$Q(t) > \widetilde{Q}(t) > S(t) > 0 \tag{32}$$

and therefore Theorem 1 is proved based on Theorem A.

The proof of Theorem 1 in case $\sigma = 1$ is based on Theorem B and on the following

Lemma 1 *Consider on the half-axis (t_0,∞) the following non-linear integral equation*

$$u(t) = B(t)\left[1 - \exp\int_t^{t+1} u(\xi)\,d\xi\right] + A(t) := (\Phi u)(t), \tag{33}$$

where

$$|B(t)| \le C, \quad \lim_{t\to\infty} B(t) = b > 0 \tag{34}$$

and

$$\lim_{t\to\infty} t^m A(t) = 0 \ \textit{for some}\ m \ge 0. \tag{35}$$

Then Eq.(33) has at least one solution $u_0(t)$ such that

$$\lim_{t\to\infty} t^m u_0(t) = 0. \tag{36}$$

Proof We define the operator Φ on the Banach space $\mathbb{C} = \mathbb{C}(t_0,\infty)$ with the norm $\|u\| = \sup_{t>t_0}|u(t)|$.

Consider $\mathcal{M} := \{u \in \mathbb{C} : \ -d_1 \le u(t) \le d_2\}$, $d_1, d_2 > 0$. The bounded set $\mathcal{M}$ is convex because

$$u_1,\, u_2 \in \mathcal{M} \Longrightarrow \vartheta u_1 + (1-\vartheta)u_2 \in \mathcal{M},\ \ 0 < \vartheta < 1.$$

Choose d_1, d_2 such that

$$\Phi\mathcal{M} \subset \mathcal{M}. \tag{37}$$

Suppose $u \in \mathcal{M}$ and $v = \Phi u$. Then

$$v(t) = A(t) + B(t)\left[1 - \exp\int_t^{t+1} u(\xi)\,d\xi\right] \le A(t) + B(t)(1-e^{-d_1}),$$

$$v(t) \ge A(t) + B(t)(1-e^{d_2}).$$

For $v \in \mathcal{M}$ it should be

$$\begin{cases} A(t) + B(t)[1 - e^{-d_1}] \le d_2 \\ A(t) + B(t)[1 - e^{d_2}] \ge -d_1 \end{cases} \iff$$

$$\iff A(t) + B(t)[1 - e^{-d_1}] \le d_2 \le \ln\Big[1 + \frac{A(t)}{B(t)} + \frac{d_1}{B(t)}\Big]. \tag{38}$$

We choose d_1 and d_2 as follows: suppose $d_1 > 0$ is sufficiently large such that $b < \ln(1 + \frac{d_1}{b})$ and $b(1 - e^{-d_1}) < \ln(1 + \frac{d_1}{b})$. Then we can define d_2 by ,

$$b(1 - e^{-d_1}) < d_2 < \ln\Big(1 + \frac{d_1}{b}\Big). \tag{39}$$

Thus, in view of (34) and (35), the inequality (38) holds for sufficiently large t_0. Therefore, $\Phi\mathcal{M} \subset \mathcal{M}$.

Now we define

$$\mathcal{M}_0 := \Big\{\forall u \in \mathcal{M} : \lim_{t\to\infty} t^m u(t) = 0 \text{ uniformly}\Big\}.$$

That is $\forall \varepsilon > 0\ \exists T_\varepsilon > t_0$ such that for all $u \in \mathcal{M}_0$, $|u(t)| < \varepsilon$ for $t > T_\varepsilon$.

Evidently, $\mathcal{M}_0$ is closed in $\mathbb{C}$. Indeed, all previous reasons are valid. On the other hand, if $u \in \mathcal{M}_0$, then $\lim\limits_{t\to\infty} t^m(\Phi u)(t) = 0$,

$$\lim_{t\to\infty} t^m\Big[\exp\int_t^{t+1} u(\xi)\, d\xi - 1\Big] = 0,$$

and so, $\Phi\mathcal{M}_0 \subset \mathcal{M}_0$. Show, in addition, that the set $\{\Phi\mathcal{M}_0\}$ is compact in $\mathbb{C}(t_0, \infty)$. We have to check whether there exists a finite partition of the axis $(t_0, \infty) = \bigcup\limits_1^N I_j$ such that

$$\forall\varepsilon > 0 \text{ and } \forall t_1,\, t_2 \in I_j : |v(t_1) - v(t_2)| < \varepsilon \text{ for all } v \in \Phi\mathcal{M}_0.$$

According to the definition of $\mathcal{M}_0$, let T_ε be such that $|v(t)| < \frac{\varepsilon}{8C}$ for $t > T_\varepsilon$ and for any $v \in \Phi\mathcal{M}_0$. Suppose $I_j := [t_0 + (j-1)\delta, t_0 + j\delta]$, $j = \overline{1, N-1}$, $\delta := \frac{T_\varepsilon - t_0}{N-1}$, $d = \max\{d_1, d_2\}$, and $I_N := (T_\varepsilon, \infty)$.

In view of the equality $e^{z_1} - e^{z_2} = e^{\theta z_1 + (1-\theta)z_2}(z_1 - z_2)$, $0 < \theta < 1$ and the notation $z(t) := \int\limits_t^{t+1} u(\xi)\, d\xi$, we have

$$|e^{z(t_1)} - e^{z(t_2)}| \le e^{2d}|z(t_1) - z(t_2)| = e^{2d}\Big|\int_{t_1}^{t_1+1} u(\xi)\, d\xi - \int_{t_2}^{t_2+1} u(\xi)\, d\xi\Big| =$$

$$= e^{2d}\Big|\int_{t_1}^{t_2} u(\xi)\, d\xi - \int_{t_1+1}^{t_2+1} u(\xi)\, d\xi\Big| \le e^{2d}\Big|\int_{t_1}^{t_2} |u(\xi)|\, d\xi + \int_{t_1+1}^{t_2+1} |u(\xi)|\, d\xi\Big| \le$$

$$\le 2de^{2d}C|t_1 - t_2| < 2Cde^{2d}\delta$$

$\forall t_1,\, t_2 \in I_j$, $j = \overline{1, N-1}$.

Suppose N is sufficiently large, such that $2dCe^{2d}\delta < \frac{\varepsilon}{3}$, and $\forall t_1,\, t_2 \in (t_0, T_\varepsilon)$, $|t_1 - t_2| < \delta$, the following inequalities hold:

$$|A(t_1) - A(t_2)| < \frac{\varepsilon}{3}, \quad |B(t_1) - B(t_2)| < \frac{\varepsilon}{3(1+e^d)}.$$

Then $\forall t_1,\, t_2 \in I_j$, $j = \overline{1, N-1}$, we obtain

$$|v(t_1) - v(t_2)| \leq |A(t_1) - A(t_2)| + |B(t_1) - B(t_2)| + \Bigg| B(t_1) \exp \int_{t_1}^{t_1+1} u(\xi)\, d\xi -$$

$$-B(t_2) \exp \int_{t_2}^{t_2+1} u(\xi)\, d\xi \Bigg| \leq \frac{\varepsilon}{3} + |B(t_1) - B(t_2)| + |B(t_2) - B(t_1)| \exp \int_{t_1}^{t_1+1} u(\xi)\, d\xi +$$

$$+|B(t_2)| \Bigg| \exp \int_{t_2}^{t_2+1} u(\xi)\, d\xi - \exp \int_{t_1}^{t_1+1} u(\xi)\, d\xi \Bigg| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3(1+e^d)} \cdot (1+e^d) + C2de^{2d}\delta < \varepsilon.$$

For $t_1,\, t_2 \in I_N$ $(t_1,\, t_2 > T_\varepsilon)$ we obtain

$$\big|v(t_1) - v(t_2)\big| \leq 4C \Bigg[\int_{t_1}^{t_1+1} |u(\xi)|\, d\xi + \int_{t_2}^{t_2+1} |u(\xi)|\, d\xi \Bigg] < 8C \frac{\varepsilon}{8C} = \varepsilon.$$

And so, the set $\Phi \mathcal{M}_0 \subset \mathbb{C}_0$ is compact.

Thus the continuous operator Φ transforms the convex bounded set $\mathcal{M}_0 \subset \mathbb{C}$ into its compact subset. Therefore, according to Schauder's Fixed Point Theorem, Eq.(33) has a solution $u_0(t)$ such that (36) holds. Lemma 1 is proved. □

Proof of Theorem 1 for case $\sigma = 1$. We consider now the NDE

$$y'(t) - py'(t-1) + Q(t)y(t-1) = 0, \quad t \geq t_0, \tag{40}$$

where $p > 0$ and

$$\lim_{t \to \infty} Q(t) = q. \tag{41}$$

Let $s < 1$ be the (unique) root of the equation $e^{-s}(1-s) = p$ and the pair $\{p; q\}$ be critical (that is $q = s^2 \cdot e^{-s}$). Define $\varphi(t) := \frac{\nu}{t}$ in Theorem B and consider Eq.(23), in which we introduce the new variable $u(t)$ by

$$k(t) := s + \frac{1}{2t} + \frac{z}{t^2} + u(t) := k_0(t) + u(t) \tag{42}$$

and z will be chosen later.

Using the suitable asymptotics from (28) we find (see (23))

$$L_0(t) := L[k_0(t)] = -\frac{\nu}{t^2} \left[(2-s)z + \frac{(3-s)(1+4\nu^2)}{24} \right] + o(t^{-3}).$$

Choose $z := -\frac{(3-s)(1+4\nu^2)}{24(2-s)}$. Then we get $L_0(t) = o(t^{-3})$. Substituting (42) into Eq.(23) we obtain the equation relative to the variable $u(t)$:

$$\sin \int_{t-1}^{t} \varphi(\xi)\, d\xi \cdot u(t-1) = L_0(t) - p\varphi(t) \exp \int_{t-1}^{t} k_0(\xi)\, d\xi \left[\exp \int_{t-1}^{t} u(\xi)\, d\xi - 1 \right].$$

The equation obtained is Eq.(33) with

$$A(t) := \operatorname{cosec} \int_{t}^{t+1} \varphi(\xi)\, d\xi \cdot L_0(t+1) = o(t^{-2}),$$

$$B(t) := p\varphi(t+1) \cdot \operatorname{cosec} \int_{t}^{t+1} \varphi(\xi)\, d\xi \exp \int_{t}^{t+1} k_0(\xi)\, d\xi, \quad \lim_{t\to\infty} B(t) = 1 - s > 0.$$

Then, according to Lemma 1, Eq.(23) has a (exact) solution of the form (42), where $u(t) = o(t^{-2})$ and, in view of $s < 1$, (24) holds. Substituting (42) into (25) and omitting all intermediate calculation, we find

$$\widetilde{Q}(t) = e^{-s} \cdot s^2 + \frac{e^{-s}(2-s) \cdot (1+4\nu^2)}{8t^2} + o(t^{-2}) =$$
$$= q + \mathcal{K}[1; s] \cdot (1+4\nu^2) \cdot \frac{1}{t^2} + o(t^{-2}).$$

So, it is possible to choose $\nu > 0$ such that $C > \mathcal{K}[1; s] \cdot (1+4\nu^2)$, and therefore Theorem 1 for the case $\sigma = 1$ is proved.

References

[1] S. S. Cheng and J. U. Zhang, On a functional differential equation related to the Emden–Fowler equation. *Funct. Differ. Eqs.* **4**(1997), No. 1–2, 215–221.

[2] J. Diblik, Positive and oscillating solutions of differential equations with delay in critical case. *J. Comput. Appl. Math.* **88**(1998), 185–202.

[3] J. Diblik and N. Koksch, Positive solutions of the equation $x'(t) = -c(t)x(t-\tau)$ in the critical case. *J. Math. Anal. Appl.* **250**(2000), 635–659.

[4] Y. Domshlak, On oscillation and location of the solutions' zeros for neutral differential equations. (Russian) *Izv. Akad. Nauk Azerb.* **2**(1984), 24–30.

[5] Y. Domshlak, On oscillation properties of delay differential equations with oscillating coefficients. *Funct. Differ. Eqs.* **2**(1994), 59–68.

[6] Y. Domshlak, G. Kvinikadze, and I. P. Stavroulakis, Sturmian comparison method: the version for the first order neutral differential equations. *Math. Inequal. Appl.* **5**(2002), 247–256.

[7] Y. Domshlak and I. P. Stavroulakis, Oscillation of first order delay differential equations in a critical state. *Appl. Anal.* **61**(1996), 359–371.

[8] L. Erbe, Q. Kong, and B. G. Zhang, Oscillation theory for functional differential equations. *Marcel Dekker*, 1994.

[9] Y. G. Sficas and I. P. Stavroulakis, Necessary and sufficient conditions for oscillations of neutral differential equations. *J. Math. Anal. Appl.* **123**(1987), No. 2, 494–507.

OSCILLATION CRITERIA FOR SECOND-ORDER DELAY DIFFERENTIAL EQUATIONS

J. DŽURINA[1] AND I.P. STAVROULAKIS[2]

ABSTRACT. The aim of this paper is to establish some new oscillation criteria for the second order retarded differential equation

$$\left(r(t)|u'(t)|^{\alpha-1}u'(t)\right)' + p(t)|u[\tau(t)]|^{\alpha-1}u[\tau(t)] = 0.$$

The results obtained essentially improve known results in the literature.

INTRODUCTION

In this paper we study oscillatory properties of the retarded functional differential equation

$$\left(r(t)|u'(t)|^{\alpha-1}u'(t)\right)' + p(t)|u[\tau(t)]|^{\alpha-1}u[\tau(t)] = 0 \qquad (E_1)$$

under the following hypothesis (H):

(H1) α is a positive number;
(H2) $r(t) \in C^1(t_0, \infty)$, $r(t) > 0$; $R(t) := \int_{t_0}^{t} r^{-\frac{1}{\alpha}}(s)\, ds \to \infty$ as $t \to \infty$;
(H3) $p(t) \in C(t_0, \infty)$, $p(t) > 0$;
(H4) $\tau(t) \in C^1(t_0, \infty)$, $\tau(t) \le t$, $lim_{t\to\infty}\tau(t) = \infty$.

By a solution of (E_1) we mean a function $u \in C^1[T_u, \infty)$, $T_u \ge t_0$, which has the property $r(t)|u'(t)|^{\alpha-1}u'(t) \in C^1[T_u, \infty)$ and satisfies (E_1) on T_u, ∞. We consider only those solutions $u(t)$ of (E_1) which satisfy $\sup\{|u(t) : t \le T|\} > 0$ for all $T \le T_u$. We assume that (E_1) possesses such a solution. A nontrivial solution of (E_1) is said to be oscillatory it if has arbitrarily large zeros; otherwise it is called nonoscillatory. Equation (E_1) is oscillatory if all of its solutions are oscillatory.

Recently, Mirzov[8,9,10], Elbert [3,4], Kusano et al [5,6,7], Chern et al [2], Agarwal et al[1] have observed some similar properties between equation $(E_1)$and the corresponding linear equation

$$\left(r(t)y'\right)' + q(t)y\left(\tau(t)\right) = 0.$$

In this paper we shall continue in this direction the study of oscillatory properties of (E_1). The purpose of this paper is to improve the above-mentioned results. We shall

---

1991 *Mathematics Subject Classification.* Primary 34C10.
Key words and phrases. neutral equation, delayed argument.
Research supported by S.G.A., Grant No. 1/74466/00

establish some new oscillatory criteria for (E_1) and for the following partial case of (E_1)

$$\left(|u'(t)|^{\alpha-1}u'(t)\right)' + p(t)|u(\tau(t))|^{\alpha-1}u[\tau(t)] = 0. \tag{E_2}$$

As is customary, all functional inequalities are assumed to hold eventually, that is for all large t.

Main Results

First consider the case where $\alpha \geq 1$.

Theorem 1. *Let $\alpha \geq 1$. Assume that for some $k \in (0,1)$*

$$\int^{\infty} \left(R^{\alpha}[\tau(t)]p(t) - \frac{\alpha\tau'(t)}{4kR[\tau(t)]r^{1/\alpha}[\tau(t)]}\right)dt = \infty \tag{1}$$

Then Eq.(E_1) is oscillatory.

Proof. Assume the converse. Let $u(t)$ be a nonoscillatory solution of (E_1). Without loss of generality we may assume that $u(t) > 0$. This implies

$$\left(r(t)|u'(t)|^{\alpha-1}u'(t)\right)' = -p(t)(u[\tau(t)])^{\alpha} < 0$$

Hence, the function $r(t)|u'(t)|^{\alpha-1}u'(t)$ is decreasing and therefore we shall consider the following two cases

(i) $u'(t) > 0$,
(ii) $u'(t) < 0$.

But the condition (ii) implies that for some positive constant M

$$r(t)|u'(t)|^{\alpha-1}u'(t) \leq -M < 0.$$

That is

$$-u'(t) \geq \left(\frac{M}{r(t)}\right)^{1/\alpha}.$$

Integrating the above inequality from t_1 to t, we obtain

$$u(t) \leq u(t_1) - M^{1/\alpha}(R(t) - R(t_1)).$$

Letting $t \to \infty$, in the above inequality and using (H2), we get $u(t) \to -\infty$. This contradiction proves that (i) holds. On the other hand, using the fact that $[r(t)(u'(t))^{\alpha}]^{1/\alpha}$ is nonincreasing, we see that for any $k_1 \in (0,1)$ and all large t

$$\begin{aligned} u[\tau(t)] &\geq \int_{t_1}^{\tau(t)} u'(s)\,ds = \int_{t_1}^{\tau(t)} \frac{1}{r^{1/\alpha}(s)}\left(r^{1/\alpha}(s)u'(s)\right)ds \\ &\geq r^{1/\alpha}[\tau(t)]u'[\tau(t)]\left(R[\tau(t)] - R(t_1)\right) > k_1 R[\tau(t)]r^{1/\alpha}[\tau(t)]u'[\tau(t)]. \end{aligned} \tag{2}$$

Define

$$w(t) = R^{\alpha}[\tau(t)]\frac{r(t)(u'(t))^{\alpha}}{(u[\tau(t)])^{\alpha}}. \tag{3}$$

Then $w(t) > 0$ and

$$w'(t) = \frac{\alpha \cdot \tau'(t) R^{\alpha-1}[\tau(t)]}{r^{1/\alpha}[\tau(t)]} \cdot \frac{r(t)\big(u'(t)\big)^{\alpha}}{\big(u[\tau(t)]\big)^{\alpha}} - R^{\alpha}[\tau(t)]p(t)$$

$$- \alpha R^{\alpha}[\tau(t)] \frac{r(t)\big(u'(t)\big)^{\alpha}}{\big(u[\tau(t)]\big)^{\alpha+1}} u'[\tau(t)]\tau'(t). \tag{4}$$

Taking into account (2) and the monotonicity of $r(t)(u'(t))^{\alpha}$, we conclude that

$$\frac{u'[\tau(t)]}{u[\tau(t)]} = \frac{1}{r[\tau(t)]} \cdot \frac{r[\tau(t)]\big(u'[\tau(t)]\big)^{\alpha}}{\big(u[\tau(t)]\big)^{\alpha}} \cdot \left(\frac{u[\tau(t)]}{u'[\tau(t)]}\right)^{\alpha-1}$$

$$\geq \frac{r(t)\big(u'(t)\big)^{\alpha}}{u[\tau(t)]^{\alpha}} \cdot \frac{kR^{\alpha-1}[\tau(t)]}{r^{1/\alpha}[\tau(t)]} ,$$

where $k = k_1^{\alpha-1} \in (0,1)$. Combining the previous inequality with (4) we get

$$w'(t) \leq \frac{\alpha\tau'(t)}{R[\tau(t)]r^{1/\alpha}[\tau(t)]} w(t) - \frac{k\alpha\tau'(t)}{R[\tau(t)]r^{1/\alpha}[\tau(t)]} w^2(t) - R^{\alpha}[\tau(t)]p(t).$$

Since the polynomial $P(w) = w - kw^2 \leq 1/(4k)$, the previous inequality implies

$$w'(t) \leq \frac{\alpha\tau'(t)}{4R[\tau(t)]r^{1/\alpha}[\tau(t)]} - R^{\alpha}[\tau(t)]p(t).$$

Integrating this estimate from t_1 to t, we have

$$w(t) \leq w(t_1) - \int_{t_1}^{t} \left[R^{\alpha}[\tau(t)]p(s) - \frac{\alpha\tau'(s)}{4R[\tau(s)]r^{1/\alpha}[\tau(s)]}\right] ds .$$

Letting $t \to \infty$ we get in view of (1) that $w(t) \to -\infty$. This contradiction completes the proof.

Corollary 1. *Let $\alpha \geq 1$. Assume that*

$$\liminf_{t\to\infty} \frac{R^{\alpha+1}[\tau(t)]r^{1/\alpha}[\tau(t)]p(t)}{\tau'(t)} > \frac{\alpha}{4}. \tag{5}$$

Then Eq.(E_1) is oscillatory.

Proof. It is not hard to verify that (5) yields the existence of $k \in (0,1)$ and $\epsilon > 0$ such that for all large t

$$\frac{R^{\alpha+1}[\tau(t)]r^{1/\alpha}[\tau(t)]p(t)}{\tau'(t)} > \frac{\alpha}{4k} + \epsilon .$$

This means that

$$R^{\alpha}[\tau(t)]p(t) - \frac{\alpha\tau'(t)}{4kR[\tau(t)]r^{1/\alpha}[\tau(t)]} > \epsilon \frac{\tau'(t)}{R[\tau(t)]r^{1/\alpha}[\tau(t)]} . \tag{6}$$

Now, it is obvious that (6) implies (1) and the assertion of this corollary follows from Theorem 1.

Imposing stronger condition on the function $r(t)$ we are able to prove another oscillation criterion for Eq.(E_1).

Theorem 2. *Let $r'(t) > 0$ and $\alpha \geq 1$. Assume that for some $k \in (0,1)$*

$$\int^{\infty} \left(R^{\alpha}[\tau(t)]p(t) - \frac{\alpha\tau'(t)R^{\alpha-2}[\tau(t)]r^{1-\frac{2}{\alpha}}[\tau(t)]}{4k\tau^{\alpha-1}(t)} \right) dt = \infty. \tag{8}$$

Then Eq.(E_1) is oscillatory.

Proof. Assume that $u(t)$ is an eventually positive solution of (E_1). From the proof of Theorem 1 we know that $u'(t) > 0$ and that $r(t)\big(u'(t)\big)^{\alpha}$ is decreasing. Moreover, since

$$0 > \left(r(t)\big(u'(t)\big)^{\alpha}\right)' = r'(t)\big(u'(t)\big)^{\alpha} + \alpha r(t)\big(u'(t)\big)^{\alpha-1}u''(t) ,$$

we see that $u''(t) < 0$. It is easy to verify that for any $k_1 \in (0,1)$ and all large t

$$u[\tau(t)] \geq \int_{t_1}^{\tau(t)} u'(s)ds \geq u'[\tau(t)]\big(\tau(t) - t_1\big) \geq k_1\tau(t)u'[\tau(t)] . \tag{9}$$

Let $w(t)$ be defined as in (3). Then $w(t) > 0$ and (4) holds. Using (9) and the monotonicity of $r(t)(u'(t))^{\alpha}$ we conclude that

$$\begin{aligned}\frac{u'[\tau(t)]}{u[\tau(t)]} &= \frac{1}{r[\tau(t)]} \cdot \frac{r[\tau(t)]\big(u'[\tau(t)]\big)^{\alpha}}{\big(u[\tau(t)]\big)^{\alpha}} \cdot \left(\frac{u[\tau(t)]}{u'[\tau(t)]}\right)^{\alpha-1} \\ &\geq \frac{1}{r[\tau(t)]} \cdot \frac{r(t)\big(u'(t)\big)^{\alpha}}{\big(u[\tau(t)]\big)^{\alpha}} \big(k_1\tau(t)\big)^{\alpha-1}.\end{aligned}$$

Using this estimate in (3), we have

$$w'(t) \leq \frac{\alpha\tau'(t)}{R[\tau(t)]r^{1/\alpha}[\tau(t)]} w(t) - \frac{\alpha k\tau'(t)\tau^{\alpha-1}(t)}{R^{\alpha}[\tau(t)]r[\tau(t)]} w^2(t) - R^{\alpha}[\tau(t)]p(t) ,$$

where $k = k_1^{\alpha-1} \in (0,1)$. It is easy to check that

$$\begin{aligned}w'(t) \leq &\frac{\alpha\tau'(t)R^{\alpha-2}[\tau(t)]r^{1-\frac{2}{\alpha}}[\tau(t)]}{4k\tau^{\alpha-1}(t)} - R^{\alpha}[\tau(t)]p(t) \\ &- \frac{\alpha k\tau'(t)\tau^{\alpha-1}(t)}{R^{\alpha}[\tau(t)]r[\tau(t)]} \left[w(t) - \frac{R^{\alpha-1}[\tau(t)]r^{1-\frac{1}{\alpha}}[\tau(t)]}{2k\tau^{\alpha-1}(t)} \right]^2\end{aligned}$$

Consequently,

$$w'(t) \leq \frac{\alpha\tau'(t)R^{\alpha-2}[\tau(t)]r^{1-\frac{2}{\alpha}}[\tau(t)]}{4k\tau^{\alpha-1}(t)} - R^{\alpha}[\tau(t)]p(t) .$$

Now, we can proceed exactly as in the proof of Theorem 1 to obtain desirable contradiction with the positivity of $w(t)$. So, this part of the proof can be omitted.

Remark 1. Conditions (1) in Theorem 1 and (8) of Theorem 2 are equivalent for $\alpha = 1$. Moreover, it can be easily checked that for $\alpha = 1$ we can let $k = 1$ in (1) and (8), respectively.

Remark 2. Theorems 1 and 2 improve Theorem 1 in [2] and Theorem 1 improves Theorem 2.3 in [1].

The conclusions of Theorems 1 and 2 lead to the following results for Eq.(E_2).

Theorem 3. *Let $\alpha \geq 1$. Assume that for some $k \in (0,1)$*

$$\int^{\infty} \left(\tau^{\alpha}(t)p(t) - \frac{\alpha\tau'(t)}{4k\tau(t)} \right) dt = \infty .$$

Then Eq.(E_2) is oscillatory.

Proof. Assuming the converse, we admit that $x(t)$ is an eventually positive solution of (E_2). Following the steps of the proof of Theorem 1 (or Theorem 2) and setting

$$w(t) = \left(\frac{\tau(t)u'(t)}{u[\tau(t)]} \right)^{\alpha}$$

we get a desirable contradiction.

Corrolary 1 leads to the following:

Corollary 2. *Let $\alpha \geq 1$. Assume that*

$$\liminf_{t\to\infty} \frac{\tau^{\alpha+1}(t)p(t)}{\tau'(t)} > \frac{\alpha}{4} .$$

Then Eq.(E_2) is oscillatory.

Now consider the case where $0 < \alpha < 1$.

Theorem 4. *Let $0 < \alpha < 1$. Assume that*

$$\int^{\infty} \left(R^{\alpha}[\tau(t)]p(t) - \frac{\alpha\tau'(t)}{4R^{2-\alpha}[\tau(t)]r^{\frac{2}{\alpha}-1}[\tau(t)]\tilde{P}(t)} \right) dt = \infty, \tag{10}$$

where

$$\tilde{P}(t) = \left(\frac{1}{r[\tau(t)]} \int_t^{\infty} p(s)ds \right)^{\frac{1-\alpha}{\alpha}} .$$

Then Eq.(E_1) is oscillatory.

Proof. It can be assumed that Eq.(E_1) has an eventually positive solution $u(t)$. Then, exactly as in the proof of Theorem 1 we conclude that $u'(t) > 0$ and moreover $r(t)(u'(t))^{\alpha}$ is decreasing. Using these facts and integrating (E_1) from t to ∞ we have

$$r(\tau(t))\left(u'(\tau(t))\right)^{\alpha} \geq r(t)(u'(t)) \geq \int_{t_1}^{\infty} p(s)u^{\alpha}[\tau(s)]ds \geq u^{\alpha}[\tau(t)] \int_t^{\infty} p(s)ds.$$

Thus

$$\left(\frac{u'(\tau(t))}{u(\tau(t))} \right)^{1-\alpha} \geq \tilde{P}(t) .$$

Defining $w(t)$ as in (3), we see that (4) holds. I can easily be checked that

$$\frac{u'[\tau(t)]}{u[\tau(t)]} = \frac{1}{r[\tau(t)]} \cdot \frac{r[\tau(t)](u'[\tau(t)])^{\alpha}}{(u[\tau(t)])^{\alpha}} \cdot \left(\frac{u'[\tau(t)]}{u[\tau(t)]} \right)^{1-\alpha} \geq \frac{\tilde{P}(t)}{r[\tau(t)]} \cdot \frac{r(t)(u'(t))^{\alpha}}{(u[\tau(t)])^{\alpha}}$$

Combining this with (4) we have

$$w'(t) \leq \frac{\alpha\tau'(t)}{R[\tau(t)]r^{1/\alpha}[\tau(t)]}w(t) - \frac{\alpha\tau'(t)\tilde{P}(t)}{R^{\alpha}[\tau(t)]r[\tau(t)]}\,w^2(t) - R^{\alpha}[\tau(t)]p(t)\ .$$

Direct computation shows that

$$\begin{aligned} w'(t) \leq\ & \frac{\alpha\tau'(t)R^{\alpha-2}[\tau(t)]r^{1-\frac{2}{\alpha}}[\tau(t)]}{4\tilde{P}(t)} - R^{\alpha}[\tau(t)]p(t) \\ & - \frac{\alpha\tau'(t)\tilde{P}(t)}{R^{\alpha}[\tau(t)]r[\tau(t)]}\left[w - \frac{R^{\alpha-1}[\tau(t)]r^{1-\frac{1}{\alpha}}[\tau(t)]}{2\tilde{P}(t)}\right]^2 \end{aligned}$$

Therefore,

$$w'(t) \leq \frac{\alpha\tau'(t)}{4R^{2-\alpha}[\tau(t)]r^{\frac{2}{\alpha}-1}[\tau(t)]\tilde{P}(t)} - R^{\alpha}[\tau(t)]p(t)\ .$$

Integrating this estimate from t to ∞ we obtain in view of (10) that $\lim_{t\to\infty} w(t) = -\infty$. This contradiction proves the theorem.

For the partial case of (E_1) we immediately have the following

Theorem 5. *Let $0 < \alpha < 1$. Denote $\tilde{P}_1(t) = \left(\int_t^{\infty} p(s)ds\right)^{\frac{1-\alpha}{\alpha}}$. If*

$$\int^{\infty}\left(\tau^{\alpha}(t)p(t) - \frac{\alpha\tau'(t)}{4\tau^{2-\alpha}(t)\tilde{P}_1(t)}\right)dt = \infty, \qquad (11)$$

then Eq.(E_2) is oscillatory.

The proof of this theorem is similar to the proof of Theorem 2 and therefore is omitted.

Examples

Example 1. Consider

$$\left[|u'(t)|^{\alpha-1}u'(t)\right]' + \frac{a}{t^{\alpha+1}}\,|u[\lambda t]|^{\alpha-1}\,u[\lambda t] = 0, \quad 0<\lambda<1, \quad \alpha\geq 1, \quad a>0 \qquad (12)$$

If

$$\frac{\lambda^{\alpha}}{\alpha}a > \frac{1}{4}$$

then, from Corrolary 2, it follows that Eq.(12) is oscillatory. Observe that our condition essentially improves the condition

$$\frac{\lambda^{\alpha}}{\alpha}a > 1$$

given in Theorem 2.3 in [1].

On the other hand, the corresponding Theorems 1 and 5 in [2] fail for Eq.(12).

Example 2. For the retarded differential equation

$$\left(|u'(t)|u'(t)\right)' + \frac{a}{t^2}u^2[\sqrt{t}]\,\mathrm{sgn}\,u[\sqrt{t}] = 0, \quad a > 0 \tag{13}$$

Corollary 2 implies oscillation of (13) if

$$a > 1/4.$$

Observe however that Theorem 2.3 in [1] requires the stronger condition

$$a > 2.$$

Also Theorem 1 and 5 in [2] cannot be applied to (13).

Example 3. Consider Eq.(12) with $0 < \alpha < 1$. If

$$\frac{\lambda^\alpha}{\alpha}a > \frac{1}{4^\alpha}$$

then, by Theorem 5, all solutions of this equation oscillate. Observe that this condition essentially improves the condition

$$\frac{\lambda^\alpha}{\alpha}a > 1$$

given in Theorem 2.3 in [1]. Moreover Theorems 1 and 5 in [2] fail for this equation.

References

1. R.P. Agarwal, S.L. Shieh and C.C. Yeh, *Oscillation criteria for second-order retarded differential equations*, Mathl. Comput. Modelling **26** (1997), 1–11.
2. J.L. Chern, W.Ch. Lian and C.C. Yeh, *Oscillation criteria for second order half-linear differential equations with functional arguments*, Publ. Math. Debrecen **48** (1996), 209–216.
3. Á. Elbert, *A half-linear second order differential equation*, Colloquia Math. Soc. János Bolyai 30: Qualitative Theory of Differential Equations (1979), 153–180, Szeged.
4. Á. Elbert, *Oscillation and nonoscillation theorems for some nonlinear ordinary differential equations*, Lecture Notes in Mathematics, Ordinary and Partial Differential Equations **964** (1982), 187–212.
5. T. Kusano and Y. Naito, *Oscillation and nonoscillation criteria for second order quasilinear differential equations*, (preprint).
6. T. Kusano, Y. Naito and A. Ogata, *Strong oscillation and nonoscillation of quasilinear differential equations of second order*, Differential Equations and Dynamical Systems **2** (1994), 1–10.
7. T. Kusano and N. Yoshida, *Nonoscillation theorems for a class of quasilinear differential equations of second order*, J. Math. Anal. Appl. **189** (1995), 115–127.
8. D.D. Mirzov, *On the oscillation of a system of nonlinear differential equations*, Diferencia?nye Uravnenija 9 (1973), 581–583.
9. D.D. Mirzov, *On some analogs of Sturm's and Kneser's theorems for nonlinear systems*, J. Math. Anal. Appl. **53** (1976), 418–425.
10. D.D. Mirzov, *On the oscillation of solutions of a system of differential equations*, Mat. Zametki **23** (1978), 401–404.

[1]Jozef Džurina, Department of Mathematical Analysis, Faculty of Sciences, Šafárik University, Jesenná 5, 041 54 Košice, Slovakia, e-mail: **dzurina@kosice.upjs.sk**

[2]I. P. Stavroulakis, Department of Mathematics, University of Ioannina, 451 10 Ioannina, Greece, e-mail: ipstav@cc.uoi.gr