ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
2004

ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
MATHEMATICA 43 (2004)

MATHEMATICA 43

PALACKÝ UNIVERSITY OLOMOUC
OLOMOUC 2004

The autors are responsible for the translation of the papers.

ISBN 80-244-0965-8
ISSN 0231-9721

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004)

CONTENTS

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 7–20

Further Ultimate Boundedness of Solutions of some System of Third Order Nonlinear Ordinary Differential Equations

A. U. AFUWAPE[1], M. O. OMEIKE[2]

[1]*Department of Mathematics, Obafemi Awolowo University, Ile-Ife, Nigeria*
e-mail: aafuwape@oauife.edu.ng

[2]*Department of Mathematical Sciences, University of Agriculture Abeokuta, Nigeria*
e-mail: moomeike@yahoo.com

(Received September 3, 2003)

Abstract

In this paper, we shall give sufficient conditions for the ultimate boundedness of solutions for some system of third order non-linear ordinary differential equations of the form

$$\dddot{X} + F(\ddot{X}) + G(\dot{X}) + H(X) = P(t, X, \dot{X}, \ddot{X})$$

where $X, F(\ddot{X})$, $G(\dot{X})$, $H(X)$, $P(t, X, \dot{X}, \ddot{X})$ are real n-vectors with F, G, $H : \mathbb{R}^n \to \mathbb{R}^n$ and $P : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ continuous in their respective arguments. We do not necessarily require that $F(\ddot{X}), G(\dot{X})$ and $H(X)$ are differentiable. Using the basic tools of a complete Lyapunov Function, earlier results are generalized.

Key words: Ultimate boundedness, complete Lyapunov functions, nonlinear third order system.

2000 Mathematics Subject Classification: 34D40, 34D20, 34C25

1 Introduction

In a sequence of results, Afuwape [1, 2, 3], Ezeilo [5], Ezeilo and Tejumola [8, 9], Meng [10] and Tiryaki [12] studied particular cases of the third-order nonlinear system of differential equations of the form

$$\dddot{X} + F(\ddot{X}) + G(\dot{X}) + H(X) = P(t, X, \dot{X}, \ddot{X}) \tag{1.1}$$

where $X, F(\ddot{X}), G(\dot{X}), H(X), P(t, X, \dot{X}, \ddot{X})$ are real n-vectors with $F, G, H : \mathbb{R}^n \to \mathbb{R}^n$ and $P : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ continuous in the respective arguments.

Boundedness and Periodicity results were discussed by imposing differentiability conditions in [5, 8, 9, 12] on the nonlinear functions in the particular cases of (1.1), while not necessarily differentiable conditions were imposed in [1, 3, 10] for the study of ultimate boundedness of particular cases of (1.1). Furthermore, the Lyapunov second method was used with the aid of a suitable differentiable Lyapunov function.

For $n = 1$ and $f(\ddot{x}) = a\ddot{x}, g(\dot{x}) = b\dot{x}$ this reduces to

$$\dddot{x} + a\ddot{x} + b\dot{x} + h(x) = p(t, x, \dot{x}, \ddot{x}) \tag{1.2}$$

which was studied by Ezeilo [6,7]. In [7], Ezeilo studied the ultimate boundedness and convergence of solutions of (1.2) by assuming

$$\frac{h(\xi + \eta) - h(\eta)}{\xi} \in I_0 \tag{1.3}$$

for some designated $\xi, \eta (\neq 0)$ with $I_0 \equiv [\delta, kab]$ where $\delta > 0$ is an arbitrary constant and $0 < k < 1$. I_0 is a subset of the generalized Routh–Hurwitz interval $(0, ab)$.

When $\eta = 0$, $\xi \neq 0$ in (1.3) we have

$$H_0 = H_0(\xi) \equiv \frac{\{h(\xi) - h(0)\}}{\xi} \tag{1.4}$$

and

$$H_0 = \frac{h(\xi)}{\xi} \quad \text{if } h(0) = 0. \tag{1.5}$$

On the other hand if $F(\ddot{X}) = A\ddot{X}, G(\dot{X}) = B\dot{X}$ in (1.1) we have

$$\dddot{X} + A\ddot{X} + B\dot{X} + H(X) = P(t, X, \dot{X}, \ddot{X}) \tag{1.6}$$

where A, B are real symmetric $n \times n$ matrices.

Afuwape [1] and Meng [10] studied (1.6) for the ultimate boundedness and periodicity of solutions for which H is of class $C(\mathbb{R}^n)$ by satisfying

$$H(X) = H(Y) + A(X, Y)(X - Y) \tag{1.7}$$

where $A(X,Y)$ is a real $n \times n$ operator for any X,Y in $\mathbb{R}^n$, and having real eigenvalues $\lambda_i(A(X,Y))$ $(i = 1,2,\ldots,n)$.

It was assumed that these eigeuvalues satisfy

$$0 < \delta_h \leq \lambda_i(A(X,X)) \leq \Delta_h \tag{1.8}$$

with δ_h, Δ_h as fixed constants.

Moreover, the matrices A, B have real positive eigenvalues $\lambda_i(A)$ and $\lambda_i(B)$ respectively with $\delta_a = \min \lambda_i(A), \delta_b = \min \lambda_i(B), \Delta_a = \max \lambda_i(A)$, $\Delta_b = \max \lambda_i(B), i = 1,2,\ldots,n$ and that for some constant $k(< 1)$ the "generalized" Routh–Hurwitz condition,

$$\Delta_h \leq k\delta_a\delta_g \tag{1.9}$$

was satisfied. Furthermore, when $F(\ddot{X}) = A\ddot{X}$ in (1.1) we have

$$\dddot{X} + A\ddot{X} + G(\dot{X}) + H(X) = P(t, X, \dot{X}, \ddot{X}) \tag{1.10}$$

where A is a real symmetric $n \times n$ matrix.

In [3], Afuwape studied (1.10) for the ultimate boundedness of solutions for which G, H are of class $C(\mathbb{R}^n)$ by satisfying

$$G(Y_1) = G(Y_2) + B_g(Y_1,Y_2)(Y_1 - Y_2) \tag{1.11a}$$

$$H(X_1) = H(X_2) + C_h(X_1,X_2)(X_1 - X_2) \tag{1.11b}$$

where $B_g(Y_1,Y_2)$, $C_h(X_1,X_2)$ are $n \times n$ real continuous operators, having real eigenvalues $\lambda_i(B_g(Y_1,Y_2))$, $\lambda_i(C_h(X_1,X_2))$, $(i = 1,2,\ldots,n)$ respectively and which satisfy

$$0 < \delta_g \leq \lambda_i(B_g(Y_1,Y_2)) \leq \Delta_g \tag{1.12a}$$

$$0 < \delta_h \leq \lambda_i(C_h(X_1,X_2)) \leq \Delta_h \tag{1.12b}$$

with $\delta_g, \delta_h, \Delta_g, \Delta_h$ as fixed constants.

Also, the matrix A has real positive eigenvalues $\lambda_i(A)$ with $\delta_a = \min \lambda_i(A)$, $\Delta_a = \max \lambda_i(A)$, $i = 1,2,\ldots,n$ and that for some constant $k(< 1)$ the "generalized" Routh Hurwitz condition (1.9) was satisfied.

In this paper, we shall extend earlier results of [1, 3, 5, 8, 9, 10, 12] to systems of the form (1.1) and for which generalized Routh–Hurwitz condition (1.9) is satisfied. A new differentiable Lyapunov function which is a modification of the one used in [10] is used to prove ultimate boundedness of solutions of (1.1). In addition to (1.11a) and (1.11b) we assume that F is of class $C(\mathbb{R}^n)$ and satisfies

$$F(Z_1) = F(Z_2) + A_f(Z_1,Z_2)(Z_1 - Z_2) \tag{1.11c}$$

where $A_f(Z_1,Z_2)$ is $n \times n$ real continuous operator having real eigenvalues $\lambda_i(A_f(Z_1,Z_2))$ $(i = 1,2,\ldots,n)$. These real eigenvalues satisfy

$$0 < \delta_f \leq \lambda_i(A_f(Z_1,Z_2)) \leq \Delta_f \tag{1.12c}$$

with δ_f, Δ_f as fixed constants.

Furthermore, these eigenvalues satisfy, for some constant $k(k<1$, defined later) the "generalized" Routh–Hurtwitz condition (1.9).

Finally, we shall assume that $P(t,X,Y,Z)$ satisfies

$$\|P(t,X,Y,Z)\| \leq p_1(t) + p_2(t)\left\{\|X(t)\|^2 + \|Y(t)\|^2 + \|Z(t)\|^2\right\}^{\rho/2} + p_3(t)\left\{\|X(t)\|^2 + \|Y(t)\|^2 + \|Z(t)\|^2\right\}^{1/2} \quad (1.13)$$

for any X,Y,Z in $\mathbb{R}^n$, where $p_1(t), p_2(t), p_3(t)$ are continuous functions in t and $0 \leq \rho \leq 1$.

Remark 1 The estimate (1.13) reduces to [8, 1.3 (3)] if $p_3(t) = \delta_0$. When specialized to the case $n = 1$, the estimate (1.13) reduces to estimate (4.96) of [11, p. 339] if $p_3(t) = q$.

2 Notations

We shall use the notations as given in [1]. Throughout this paper, δ's and Δ's with or without suffices will denote positive constants whose magnitudes depend on vector functions F, G, H and P. The δ's and Δ's with numerical or alphabetical suffices shall retain fixed magnitudes, while those without suffices are not necessarily the same at each occurrences.

Finally, we shall denote the scalar product $\langle X, Y\rangle$ of any vectors X, Y in $\mathbb{R}^n$, with respective components $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ by $\sum_{i=1}^n x_i y_i$. In particular, $\langle X, X\rangle = \|X\|^2$.

3 Statement of the results

Our first main result in this paper is the following:

Theorem 1 *Suppose $F(0) = G(0) = H(0) = 0$, and that*

(i) there exist $n \times n$ real continuous operators

$$A_f(Z_1, Z_2), \quad B_g(Y_1, Y_2), \quad C_h(X_1, X_2)$$

for any vectors $X_1, X_2, Y_1, Y_2, Z_1, Z_2$ in $\mathbb{R}^n$, such that the functions F, G, H are of class $C(\mathbb{R}^n)$, satisfy (1.11a,b,c), with the eigenvalues, $\lambda_i(A_f(Z_1, Z_2))$, $\lambda_i(B_g(Y_1, Y_2))$, $\lambda_i(C_h(X_1, X_2))$ $(i = 1, 2, \ldots, n)$ satisfying (1.12a,b,c);

(ii) the operators A_f, B_g and C_h are associative and commute pairwise, and

(iii) the vector function P satisfies inequality (1.13) for all X, Y, Z in $\mathbb{R}^n$, where $p_1(t)$, $p_2(t)$ and $p_3(t)$ are continuous functions of t, with $0 \leq \rho < 1$.

Then, there exist constants $\rho_3, \Delta_1, \Delta_2, \Delta_3$ such that if $|p_3(t)| \leq \rho_3$, for all t in $\mathbb{R}$, with ρ_3 chosen small enough, then every solution $X(t)$ of (1.1) with $X(t_0) =$

X_0, $\dot{X}(t_0) = Y_0$, $\ddot{X}(t_0) = Z_0$, and for any constant r, whatever in the range $\frac{1}{2} \le r \le 1$, satisfies

$$\{\|X(t)\|^2 + \|\dot{X}(t)\|^2 + \|\ddot{X}(t)^2\|\}^r \le \Delta_1 \exp\{-\Delta_2(t-t_0)\}$$
$$+ \Delta_3 \int_{t_0}^{t} \left\{p_1^{2r}(\tau) + p_2^{2r/(1-\rho)}(\tau)\right\} \exp\{-\Delta_2(t-\tau)\}\, d\tau; \tag{3.1}$$

for all $t \ge t_0 \ge 0$, where $\Delta_1 \equiv \Delta_1(X_0, Y_0, Z_0)$.

Remark 2 (1) When specialized to the case $n = 1$ with P dependent only on t the above estimate (3.1) reduces to the estimate (4.86) of [11, Theorem (4.24) p. 335].

(2) In fact this result generalizes Theorem 1 of [3] if $\rho_3 = \delta_0$: A number of quite important results can be deduced from the above. For example, we have

Corollary 1 *If $P \equiv 0$ and all the conditions of Theorem 1 hold, then every solution $X(t)$ of (1.1) satisfies*

$$\{\|X(t)\|^2 + \|\dot{X}(t)\|^2 + \|\ddot{X}(t)^2\|\} \longrightarrow 0 \tag{3.2}$$

as $t \to \infty$, provided that ρ_3 is small enough.

Indeed by setting $\rho_1(t) = 0 = \rho_2(t)$ in (1.13), we have that

$$\{\|X(t)\|^2 + \|\dot{X}(t)\|^2 + \|\ddot{X}(t)^2\|\}^r \le \Delta_1 \exp\{-\Delta_2(t-t_0)\}, \quad t \ge t_0$$

from which (3.2) follows on letting $t \to \infty$.

Remark 3 When specialized to the case $n = 1$ with $p_1(t) = p_2(t) = 0$ i.e. satisfying condition (C'') of [11, Theorem 4.25] then the above estimate (3.2) reduces to the estimate (4.97) of [11, Theorem 4.25].

Further, if $P \ne 0$, but such that

$$\int_{t}^{t+\mu} \left\{p_1^{\nu}(\tau) + p_2^{\nu/(1-\rho)}(\tau)\right\} d\tau \longrightarrow 0 \tag{3.3}$$

as $t \to \infty$, then we have

Corollary 2 *Suppose that there are some fixed constants ν ($1 \le \nu \le 2$), and $\mu > 0$, such that (3.3) is true, and all the conditions of Theorem 1 hold. Then, every solution $X(t)$ of (1.1) satisfies (3.2) as $t \to \infty$.*

Remark 4 This result is a direct generalization of [6, Theorem 2] when specialized to the case $n = 1$. Its proof can be obtained from (3.1) by using an obvious modification of the arguments in [6, §3.2].

The next result is on the ultimate boundedness of solutions of (1.1).

Theorem 2 *Suppose that $F(0) = G(0) = H(0) = 0$ and that all the conditions of Theorem 1 hold. Suppose further that $|p_3(t)| \leq \rho_3$ for all t in $\mathbb{R}$ with ρ_3 sufficiently small and that the functions $p_1(t), p_2(t)$ satisfy*

$$|p_1(t)| \leq \delta_0 \quad \text{and} \quad |p_2(t)| \leq \delta_1$$

for all t in $\mathbb{R}$.

Then, there exists a constant Δ_4 such that every solution $X(t)$ of (1.1) ultimately satisfies.

$$\{\|X(t)\|^2 + \|\dot{X}(t)\|^2 + \|\ddot{X}(t)\|^2\} \leq \Delta_4 \tag{3.4}$$

Remark 5 (1) If $|p_1(t)| \leq \delta_0$, $|p_2(t)| \leq \delta_1$ and $|p_3(t)| \leq \rho_3$, with ρ_3 sufficiently small, then Theorem 2 reduces to Corollary 3 of [8] for which equation (1.6) was considered.

(2) If $\rho = 0$ in (1.13) we have the estimates (3.6) of [1, Theorem 1] which improves on estimates (3.4) of [1, Theorem 1] and (1.8) of [10, Theorem 1]. Thus, Theorem 2 reduces to Theorem 1 of [1,10] for which (1.6) was considered. Moreover, the estimate (1.13) is a generalization of all the bounds on $P(t, X, Y, Z)$ mentioned earlier.

4 Some preliminary results

We shall state, for completeness, some standard results needed in the proofs of our results.

Lemma 1 (1,§4) *Let Q, D be real symmetric commuting $n \times n$ matrices. Then,*

(i) for any X in $\mathbb{R}^n$,

$$\delta_d\|X\|^2 \leq \langle DX, X\rangle \leq \Delta_d\|X\|^2 \tag{4.1}$$

where δ_d, Δ_d are respectively, the least and greatest eigenvalues, of matrix D;

(ii) the eigenvalues $\lambda_i(QD)$, $(i = 1, 2, \ldots, n)$ of the product matrix QD are all real and satisfy

$$\min_{1\leq j,k\leq n} \lambda_j(Q)\lambda_k(D) \leq \lambda_i(QD) \leq \max_{1\leq j,k\leq n} \lambda_j(Q)\lambda_k(D) \tag{4.2}$$

(iii) the eigenvalues $\lambda_i(Q + D)$, $(i = 1, 2, \ldots, n)$ of the sum of Q and D are all real and satisfy

$$\begin{aligned} \left\{\min_{1\leq j\leq n} \lambda_j(Q) + \min_{1\leq k\leq n} \lambda_k(D)\right\} &\leq \lambda_i(Q + D) \\ &\leq \left\{\max_{1\leq k\leq n} \lambda_j(Q) + \max_{1\leq k\leq n} \lambda_k(D)\right\} \end{aligned} \tag{4.3}$$

where $\lambda_j(Q)$ and $\lambda_k(D)$ are respectively the eigenvalues of Q and D.

5 The function V

Our main tool in the proof of the results is the continuous function $V = V(X, Y, Z)$ defined for any X, Y, Z in $\mathbb{R}^n$ by

$$2V = \beta(1-\beta)\delta_g^2\|X\|^2 + \beta\delta_g\|Y\|^2 + \alpha\delta_g\delta_f^{-1}\|Y\|^2 + \alpha\delta_f^{-1}\|Z\|^2 + \|Z + \delta_f Y + (1-\beta)\delta_g X\|^2. \quad (5.1)$$

where $0 < \beta < 1$ and $\alpha > 0$

The following result is immediate from (5.1):

Lemma 2 *Assume that all the hypothesis on vectors $F(Z), G(Y)$ and $H(X)$ in Theorem 1 are satisfied. Then, there exist positive constants δ_2 and δ_3 such that*

$$\delta_2(\|X\|^2 + \|Y\|^2 + \|Z\|^2) \leq 2V \leq \delta_3(\|X\|^2 + \|Y\|^2 + \|Z\|^2) \quad (5.2)$$

Proof The proof follows if we use Lemma 1 repeatedly and then choose

$$\delta_2 = \min\left\{\beta(1-\beta)\delta_g^2; \delta_g(\beta + \alpha\delta_f^{-1}); \alpha\delta_f^{-1}\right\}$$

and

$$\delta_3 = \max\left\{\delta_g(1-\beta)(1+\delta_g+\delta_f); \delta_g(\beta+\alpha\delta_f^{-1}) + \delta_f[1+\delta_g(1-\beta)+\delta_f]; 1 + \alpha\delta_f^{-1} + \delta_f + \delta_g(1-\beta)\right\}$$

□

6 Proof of Theorem 1

Let us replace system of differential equations of form (1.1) in the equivalent system form

$$\dot{X} = Y, \quad \dot{Y} = Z, \quad \dot{Z} = -F(Z) - G(Y) - H(X) + P(t, X, Y, Z) \quad (6.1)$$

for which a typical solution will be $(X(t), Y(t), Z(t))$.

To prove Theorem 1, it suffices to show that the function V (defined in (5.1)) satisfies for any solution $(X(t), Y(t), Z(t))$ of (6.1) and for any r in the range $\frac{1}{2} \leq r \leq 1$.

$$\dot{V} \leq -\delta_4\psi^2 + \delta_5\left\{p_1^{2r}(t) + p_2^{\frac{2r}{(1-\rho)}}(t)\right\}\psi^{2(1-r)} \quad (6.2)$$

for some constants δ_4, δ_5 where $\psi^2 = \{\|X(t)\|^2 + \|Y(t)\|^2 + \|Z(t)\|^2\}$. We note that from Lemma 2, (6.2) becomes

$$\dot{V} \leq -\delta_6 V + \delta_7\left\{p_1^{2r}(t) + p_2^{\frac{2r}{(1-\rho)}}(t)\right\}V^{(1-r)} \quad (6.3)$$

with $\delta_6 = \delta_2\delta_4$ and $\delta_7 = \delta_3\delta_5$. If we choose $U = V^r$, this reduces to

$$\dot{U} \leq -r\delta_6 U + r\delta_7\left\{p_1^{2r}(t) + p_2^{\frac{2r}{(1-\rho)}}(t)\right\}. \quad (6.3)$$

which can be solved for U to obtain

$$\begin{aligned} U(t) &\le U(t_0)\exp\{-r\delta_6(t-t_0)\} \\ &+ \Delta_5\int_{t_0}^{t}\left\{p_1^{2r}(\tau)+p^{\frac{2r}{(1-\rho)}}(\tau)\right\}\exp\{-r\delta_6(t-\tau)\}\,d\tau \end{aligned} \tag{6.4}$$

for all $t \ge t_0$.

Rewriting this with $V^r = U$ and applying Lemma 2, we shall obtain (3.1) with

$$\begin{aligned} \Delta_1 &= \delta\{\|X(t_0)\|^2+\|Y(t_0)\|^2+\|Z(t_0)\|^2\}^r; \\ \Delta_2 &= r\delta_6 \quad\text{and}\quad \Delta_3=\delta\Delta_5 \end{aligned}$$

Thus the proof of Theorem 1 is complete as soon as inequality (6.2) is proved.

7 The derivative of V and the proof of (6.2)

Let $(X(t), Y(t), Z(t))$ be any solution of (6.1). The total derivative of V, with respect to t along the solution path after simplification is

$$\dot{V} = -W_1 - W_2 - W_3 - W_4 - W_5 - W_6 - W_7 + W_8 \tag{7.1}$$

where

$$\begin{aligned}
W_1 &= \{\gamma_1\delta_g(1-\beta)\langle X, H(X)\rangle + \eta_1\delta_f\langle Y, G(Y)-\delta_g(1-\beta)Y\rangle \\
&\quad + \xi_1\alpha\delta_f^{-1}\langle Z, F(Z)\rangle + \langle Z, F(Z)-\delta_f Z\rangle\Big\} \\
W_2 &= \Big\{\gamma_2\delta_g(1-\beta)\langle X, H(X)\rangle + \xi_2\alpha\delta_f^{-1}\langle Z, F(Z)\rangle + (1+\alpha\delta_f^{-1})\langle Z, H(X)\rangle\Big\} \\
W_3 &= \{\gamma_3\delta_g(1-\beta)\langle X, H(X)\rangle + \eta_2\delta_f\langle Y, G(Y)-\delta_g(1-\beta)Y\rangle + \delta_f\langle Y, H(X)\rangle\} \\
W_4 &= \Big\{\gamma_4\delta_g(1-\beta)\langle X, H(X)\rangle + \xi_3\alpha\delta_f^{-1}\langle Z, F(Z)\rangle \\
&\quad + \delta_g(1-\beta)\langle X, F(Z)-\delta_f Z\rangle\} \\
W_5 &= \{\gamma_5\delta_g(1-\beta)\langle X, H(X)\rangle + \eta_3\delta_f\langle Y, G(Y)-\delta_g(1-\beta)Y\rangle \\
&\quad + \delta_g(1-\beta)\langle X, G(Y)-\delta_g Y\rangle\} \\
W_6 &= \Big\{\xi_4\alpha\delta_f^{-1}\langle Z, F(Z)\rangle + \eta_4\delta_f\langle Y, G(Y)-\delta_g(1-\beta)Y\rangle \\
&\quad + (1+\alpha\delta_f^{-1})\langle Z, G(Y)-\delta_g Y\rangle\Big\} \\
W_7 &= \Big\{\xi_5\alpha\delta_f^{-1}\langle Z, F(Z)\rangle + \eta_5\delta_f\langle Y, G(Y)-\delta_g(1-\beta)Y\rangle + \delta_f\langle Y, F(Z)-\delta_f Z\rangle\} \\
W_8 &= \Big\{\langle(1-\beta)\delta_g X + \delta_f Y + (1+\alpha\delta_f^{-1})Z,\ P(t,X,Y,Z)\rangle\Big\}
\end{aligned}$$

with $\xi_i, \eta_i, \gamma_i;\ (i = 1, 2, 3, 4, 5)$ are strictly positive constants such that

$$\sum_{i=1}^{5}\xi_i = 1; \quad \sum_{i=1}^{5}\eta_i = 1 \quad \text{and} \quad \sum_{i=1}^{5}\gamma_i = 1.$$

To arrive at (6.2), we first prove the following:

Lemma 3 *Subject to a conveniently chosen value of k in (1.9), we have for all X, Y, Z in $\mathbb{R}^n$*

$$W_j \geq 0, \quad (j = 2, 3, 4, 5, 6, 7).$$

Proof For strictly positive constants k_1, k_2, conveniently chosen later, we have

$$\begin{aligned} &\langle (1+\alpha\delta_f^{-1})Z, H(X)\rangle = \\ &= \|k_1(1+\alpha\delta_f^{-1})^{1/2}Z + 2^{-1}k_1^{-1}(1+\alpha\delta_f^{-1})^{1/2}H(X)\|^2 \\ &- \langle k_1^2(1+\alpha\delta_f^{-1})Z, Z\rangle - \langle 4^{-1}k_1^{-2}(1+\alpha\delta_f^{-1})H(X),\ H(X)\rangle \end{aligned} \tag{7.2a}$$

and

$$\begin{aligned} \langle \delta_f Y, H(X)\rangle &= \|k_2\delta_f^{1/2}Y + 2^{-1}k_2^{-1}\delta^{1/2}H(X)\|^2 \\ &- \langle k_2^2\delta_f Y, Y\rangle - \langle 4^{-1}k_2^{-2}\delta_f H(X), H(X)\rangle. \end{aligned} \tag{7.2b}$$

Now, using (1.11) and the assumptions that $F(0) = G(0) = H(0) = 0$, we have

$$\begin{aligned} W_2 &= \|k_1(1+\alpha\delta_f^{-1})^{1/2}Z + 2^{-1}k_1^{-1}(1+\alpha\delta_f^{-1})^{1/2}H(X)\|^2 \\ &+ \langle Z, \xi_2\alpha\delta_f^{-1}F(Z) - k_1^2(1+\alpha\delta_f^{-1})Z\rangle \\ &+ \langle H(X), \gamma_2\delta_g(1-\beta)X - 4^{-1}k_1^{-2}(1+\alpha\delta_f^{-1})H(X)\rangle \end{aligned} \tag{7.3a}$$

and

$$\begin{aligned} W_3 &= \|k_2\delta_f^{1/2}Y + 2^{-1}k_2^{-1}\delta^{1/2}H(X)\|^2 \\ &+ \langle Y, \eta_2\delta_f[G(Y) - \delta_g(1-\beta)Y] - k_2^2\delta_f Y\rangle \\ &+ \langle H(X), \gamma_3\delta_g(1-\beta)X - 4^{-1}k_2^{-2}\delta_f H(X)\rangle. \end{aligned} \tag{7.3b}$$

Furthermore, by using Lemma 1 repeatedly, we obtain for all X, Z in $\mathbb{R}^n$,

$$W_2 \geq 0 \tag{7.4a}$$

if $k_1^2 \leq \frac{\xi_2\alpha\delta_f}{\alpha+\delta_f}$ with

$$\Delta_h \leq \frac{4\gamma_2\xi_2\alpha(1-\beta)\delta_f^2\delta_g}{(\alpha+\delta_f)^2} \tag{7.5a}$$

and for all X, Y in $\mathbb{R}^n$,

$$W_3 \geq 0. \tag{7.4b}$$

If $k_2^2 \leq \eta_2\beta\delta_g$ with

$$\Delta_h \leq 4\gamma_3\eta_2\beta(1-\beta)\delta_g^2/\delta_f. \tag{7.5b}$$

Combining all the inequalities in (7.3) and (7.4), we have for all X, Y, Z in $\mathbb{R}^n$, $W_2 \geq 0$ and $W_3 \geq 0$, if $\Delta_h \leq k\delta_f\delta_g$ with

$$k = \min\left\{\frac{4\gamma_2\xi_2\alpha(1-\beta)\delta_f}{(\alpha+\delta_f)^2}; \frac{4\eta_2\gamma_3\beta(1-\beta)\delta_g}{\delta_f^2}\right\} < 1. \tag{7.6}$$

To complete the proof of Lemma 3, we need to show that for all X, Y, Z in $\mathbb{R}^n$

$$W_i \geq 0 \quad (i = 4, 5, 6, 7).$$

By hypothesis (1.11) the assumptions that $F(0) = G(0) = H(0) = 0$, and for strictly positive constants k_3, k_4, k_5, k_6 conveniently chosen later, we have

$$\begin{aligned}\langle \delta_g(1-\beta)X, F(Z) - \delta_f Z\rangle &= \langle \delta_g(1-\beta)X, [A_f(Z,O) - \delta_f I]Z\rangle \\ &= \|2^{-1}k_3^{-1}\delta_g^{1/2}(1-\beta)^{1/2}[A_f(Z,O) - \delta_f I]^{1/2}X \\ &\quad + k_3\delta_g^{1/2}(1-\beta)^{1/2}[A_f(Z,O) - \delta_f I]^{1/2}Z\|^2 \\ &\quad - \langle 4^{-1}k_3^{-2}\delta_g(1-\beta)[A_f(Z,O) - \delta_f I]X, X\rangle \\ &\quad - \langle k_3^2\delta_g(1-\beta)[A_f(Z,O) - \delta_f I]Z, Z\rangle \end{aligned} \tag{7.7a}$$

$$\begin{aligned}\delta_g(1-\beta)\langle X, G(Y) - \delta_g Y\rangle &= \langle \delta_g(1-\beta)X, [B_g(Y,O) - \delta_g I]Y\rangle \\ &= \|2^{-1}k_4^{-1}\delta_g^{1/2}(1-\beta)^{1/2}[B_g(Y,O) - \delta_g I]^{1/2}X \\ &\quad + k_4\delta_g^{1/2}(1-\beta)^{1/2}[B_g(Y,O) - \delta_g I]^{1/2}Y\|^2 \\ &\quad - \langle 4^{-1}k_4^{-2}\delta_g(1-\beta)[B_g(Y,O) - \delta_g I]X, X\rangle \\ &\quad - \langle k_4^2\delta_g(1-\beta)[B_g(Y,O) - \delta_g I]Y, Y\rangle \end{aligned} \tag{7.7b}$$

$$\begin{aligned}(1+\alpha\delta_f^{-1})\langle Z, G(Y) - \delta_g Y\rangle &= \langle (1+\alpha\delta_f^{-1})Z, [B_g(Y,O) - \delta_g I]Y\rangle \\ &= \|2^{-1}k_5^{-1}(1+\alpha\delta_f^{-1})^{1/2}[B_g(Y,O) - \delta_g I]^{1/2}Z \\ &\quad + k_5(1+\alpha\delta_f^{-1})^{1/2}[B_g(Y,O) - \delta_g I]^{1/2}Y\|^2 \\ &\quad - \langle 4^{-1}k_5^{-2}(1+\alpha\delta_f^{-1})[B_g(Y,O) - \delta_g I]Z, Z\rangle \\ &\quad - \langle k_5^2(1+\alpha\delta_f^{-1})[B_g(Y,O) - \delta_g I]Y, Y\rangle \end{aligned} \tag{7.7c}$$

$$\begin{aligned}&\delta_f\langle Y, F(Z) - \delta_f Z\rangle = \langle \delta_f Y, [A_f(Z,O) - \delta_f I]Z\rangle \\ &= \|2^{-1}k_6^{-1}\delta_f^{1/2}[A_f(Z,O) - \delta_f I]^{1/2}Y + k_6\delta_f^{1/2}[A_f(Z,O) - \delta_f I]^{1/2}Z\|^2 \\ &\quad - \langle 4^{-1}k_6^{-2}\delta_f[A_f(Z,O) - \delta_f I]Y, Y\rangle \\ &\quad - \langle k_6^2\delta_f[A_f(Z,O) - \delta_f I]Z, Z\rangle. \end{aligned} \tag{7.7d}$$

Thus,

$$\begin{aligned}W_4 &= \|2^{-1}k_3^{-1}\delta_g^{1/2}(1-\beta)^{1/2}[A_f(Z,O) - \delta_f I]^{1/2}X \\ &\quad + k_3\delta_g^{1/2}(1-\beta)^{1/2}[A_f(Z,O) - \delta_f I]^{1/2}Z\|^2 \\ &\quad + \langle X, \{\gamma_4\delta_g(1-\beta)C_h(X,O) - 4^{-1}k_3^{-2}\delta_g(1-\beta)[A_f(Z,O) - \delta_f I]\}X\rangle \\ &\quad + \langle Z, \{\xi_3\alpha\delta_g^{-1}A_f(Z,O) - k_3^2\delta_g(1-\beta)[A_f(Z,O) - \delta_f I]\}Z\rangle \end{aligned} \tag{7.8a}$$

$$\begin{aligned}W_5 &= \|2^{-1}k_4^{-1}\delta_g^{1/2}(1-\beta)^{1/2}[B_g(Y,O)-\delta_g I]^{1/2}X \\ &\quad + k_4\delta_g^{1/2}(1-\beta)^{1/2}[B_g(Y,O)-\delta_g I]^{1/2}Y\|^2 \\ &\quad + \langle X, \{\gamma_5\delta_g(1-\beta)C_h(X,0) - 4^{-1}k_4^{-2}\delta_g(1-\beta)[B_g(Y,O)-\delta_g I]\}X\rangle \\ &\quad + \langle Y, \{\eta_3\delta_f[B_g(Y,O)-\delta_g(1-\beta)I] - k_4^2\delta_g(1-\beta)[B_g(Y,O)-\delta_g I]\}Y\rangle \qquad (7.8b)\end{aligned}$$

$$\begin{aligned}W_6 &= \|2^{-1}k_5^{-1}(1+\alpha\delta_f^{-1})^{1/2}[B_g(Y,O)-\delta_g I]^{1/2}Z \\ &\quad + k_5(1+\alpha\delta_f^{-1})^{1/2}[B_g(Y,O)-\delta_g I]^{1/2}Y\|^2 \\ &\quad + \langle Z, \{\xi_4\alpha\delta_g^{-1}A_f(Z,O) - 4^{-1}k_5^{-2}(1+\alpha\delta_f^{-1})[B_g(Y,O)-\delta_g I]\}Z\rangle \\ &\quad + \langle Y, \{\eta_4\delta_f[B_g(Y,O)-\delta_g(1-\beta)I] \\ &\quad - k_5^2(1+\alpha\delta_f^{-1})[B_g(Y,O)-\delta_g I]\}Y\rangle \qquad (7.8c)\end{aligned}$$

and

$$\begin{aligned}W_7 &= \|2^{-1}k_6^{-1}\delta_f^{1/2}[A_f(Z,O)-\delta_f I]^{1/2}Y + k_6\delta_f^{1/2}[A_f(Z,0)-\delta_f I]^{1/2}Z\|^2 \\ &\quad + \langle Y, \{\eta_5\delta_f[B_g(Y,O)-\delta_g(1-\beta)I] - 4^{-1}k_6^{-2}\delta_f[A_f(Z,O)-\delta_f I]\}Y\rangle \\ &\quad + \langle Z, \{\xi_5\alpha\delta_f^{-1}A_f(Z,O) - k_6^2\delta_f[A_f(Z,O)-\delta_f I]\}Z\rangle. \qquad (7.8d)\end{aligned}$$

Thus, for all X, Z in $\mathbb{R}^n$

$$W_4 \geq 0 \qquad (7.9a)$$

if

$$\frac{\Delta_f - \delta_f}{4\gamma_4\delta_h} \leq k_3^2 \leq \frac{\xi_3\alpha}{(1-\beta)(\delta_g - \delta_f)}. \qquad (7.10a)$$

For all X, Y in $\mathbb{R}^n$

$$W_5 \geq 0 \qquad (7.9b)$$

if

$$\frac{\Delta_g - \delta_g}{4\gamma_5\delta_h} \leq k_4^2 \leq \frac{\eta_3\beta\delta_f}{(1-\beta)(\Delta_g - \delta_g)}\,. \qquad (7.10b)$$

For all Y, Z in $\mathbb{R}^n$

$$W_6 \geq 0 \qquad (7.9c)$$

if

$$\frac{\delta_g(\alpha+\delta_f)(\Delta_g - \delta_g)}{4\xi_4\alpha\delta_f^2} \leq k_5^2 \leq \frac{\beta\eta_4\delta_g\delta_f^2}{(\alpha+\delta_f)(\Delta_g - \delta_g)}\,. \qquad (7.10c)$$

Also, for all Y, Z in $\mathbb{R}^n$

$$W_7 \geq 0 \qquad (7.9d)$$

if

$$\frac{\Delta_f - \delta_f}{4\eta_5\beta\delta_g} \leq k_6^2 \leq \frac{\alpha\xi_5}{\delta_f(\Delta_f - \delta_f)}\,. \qquad (7.10d)$$

This completes the proof of Lemma 3. □

We are now left with the estimates for W_1 and W_8.

From (7.1), we clearly have

$$W_1 \geq \gamma_1\delta_g\delta_h(1-\beta)\|X\|^2 + \eta_1\delta_f\delta_g\beta\|Y\|^2 + \xi_1\alpha\|Z\|^2$$
$$\geq \delta_8(\|X\|^2 + \|Y\|^2 + \|Z\|^2) \tag{7.11}$$

where $\delta_8 = \min\{\gamma_1\delta_g\delta_h; \eta_1\delta_f\delta_g\beta; \xi_1\alpha\}$. For the remaining part of the proof of (6.2); let us for convenience denote $(\|X\|^2 + \|Y\|^2 + \|Z\|^2)$ by ψ^2.

Since $P(t, X, Y, Z)$ satisfies (1.5), Schwarz's inequality gives for W_8.

$$|W_8| \leq \left\{(1-\beta)\delta_g\|X\| + \delta_f\|Y\| + (1+\alpha\delta_1^{-1})\|Z\|\right\}\|P(t,X,Y,Z)\|$$
$$\leq 3^{1/2}\delta_9\left\{p_3(t)\psi^2 + p_2(t)\psi^{(1+\rho)} + p_1(t)\psi\right\}; \tag{7.12}$$

where $\delta_9 = \max\left\{(1-\beta)\delta_g; \delta_f; (1+\alpha\delta_f^{-1})\right\}$.

Combining inequalities (7.3), (7.11) and (7.13) with the assumption that $|p_3(t)| \leq \rho_3$ for all t in $\mathbb{R}$, we obtain from (7.1) that

$$\dot{V} \leq -(\delta_8 - 3^{1/2}\delta_9\rho_3)\psi^2 + 3^{1/2}\delta_9\left\{p_2(t)\psi^{(1+\rho)} + p_1(t)\psi\right\}. \tag{7.14}$$

This we can rewrite as

$$\dot{V} \leq -\delta_{10}\psi^2 + \psi_1 + \psi_2 \tag{7.15}$$

where

$$3\delta_{10} = \delta_8 - 3^{1/2}\delta_9\rho_3, \qquad \psi_1 = \{\delta_{11}p_1(t) - \delta_{10}\psi\}\psi;$$

and

$$\psi_2 = \left\{\delta_{11}p_2(t)\psi^{(1+\rho)} - \delta_{10}\psi^2\right\}.$$

If we choose ρ_3 small enough such that $\delta_{10} > 0$ (following [6, p. 306]), with the necessary modification we obtain

$$\psi_1 \leq \delta_{12}\psi^{2(1-r)}p_1^{2r}(t) \tag{7.16a}$$

and

$$\psi_2 \leq \delta_{13}\psi^{2(1-r)}p_2^{2r/(1-\rho)}(t) \tag{7.16b}$$

for any constant r in the range $\frac{1}{2} \leq r \leq 1$.

Thus, (7.15) reduces to

$$\dot{V} \leq -\delta_{10}\psi^2 + \delta_{14}\left\{p_1^{2r}(t) + p_2^{2r/(1-\rho)}(t)\right\}\psi^{2(1-r)} \tag{7.17}$$

with

$$\delta_{14} = \max\{\delta_{12}; \delta_{13}\}$$

This is (6.2) with $\delta_4 = \delta_{10}$ and $\delta_5 = \delta_{14}$.

8 Proof of Theorem 2

As pointed out in [1], to prove Theorem 2, if suffices to prove that the function V satisfies

(i) $V(X, Y, Z) \to \infty$ as $(\|X\|^2 + \|Y\|^2 + \|Z\|^2) \to \infty$; and

(ii) $\dot{V} \leq -1$

along paths of any solution $(X(t), Y(t), Z(t))$ of (6.1) for which $(\|X(t)\|^2 + \|Y(t)\|^2 + \|Z(t)\|^2)$ is large enough. We only need to concern ourselves with property (ii), since by Lemma 2, inequality (5.3), property (i) has been taken care of.

If all the conditions of Theorem 1 are satisfied, then, for any solution $(X(t), Y(t), Z(t))$ of (6.1), $\dot{V}$ satisfies inequality (7.17). That is

$$\dot{V} \leq -\delta_{10}\psi^2 + \delta_{14}\left\{p_1^{2r}(t) + p_2^{2r/(1-\rho)}(t)\right\}\psi^{2(1-r)}$$

for any r in the range $\frac{1}{2} \leq r \leq 1$.

Now, if $p_1(t)$ and $p_2(t)$ are bounded for all t in $\mathbb{R}$, then there exists some constant $\delta_{15} > 0$ such that

$$\dot{V} \leq -\delta_{10}\psi^2 + \delta_{15}\psi^{2(1-r)} \leq -1$$

if

$$\psi \geq \delta_{16} > (\delta_{10}^{-1}\delta_{15})^{1/2r}.$$

Thus property (ii) is proved for V, and this completes the proof of Theorem 2.

References

[1] Afuwape, A. U.: *Ultimate boundedness results for a certain system of third-order non-linear differential equation.* J. Math. Anal. Appl. **97** (1983), 140–150.

[2] Afuwape, A. U.: *Uniform dissipative solutions for a third-order non-linear differential equation.* In: Differential equations (J. W. Knowles and R. T. Lewis, eds.), Elsevier, North Holland, 1984, 1–6.

[3] Afuwape, A. U.: *Further ultimate boundedness results for a third order non-linear system of differential equations.* Analisi Funzionale e Appl. **6**, 99–100, N. I. (1985), 348–360.

[4] Afuwape, A. U., Ukpera, A. S.: *Existence of solutions of periodic boundary value problems for some vector third order differential equations.* J. of Nig. Math. Soc. **20** (2001), 1–17.

[5] Ezeilo, J. O. C.: *n-dimensioinal extensions of boundedness and stability theorems for some third order differential equations.* J. Math. Anal. Appl. **18** (1967), 395–416.

[6] Ezeilo, J. O. C.: *Stability Results for the Solutions of some third and fourth order differential equations.* Ann. Mat. Pura. Appl. **66**, 4 (1964), 233–250.

[7] Ezeilo, J. O. C.: *New properties of the equation $x''' + ax'' + bx' + h(x) = p(t, x, \dot{x}, x'')$ for certain special values of the incrementary ratio $y^{-1}\{h(x + y) - h(x)\}$.* In: Equations Differentielles et Functionelles Non-lineares (P. Janssens, J. Mawhin and N. Rouche, eds.), Hermann, Paris, 1973, 447–462.

[8] Ezeilo, J. O. C., Tejumola, H. O.: *Boundedness and periodicity of solutions of a certain system of third-order non-linear differential equations.* Ann. Math. Pura e Appl. **74** (1966), 283–316.

[9] Ezeilo, J. O. C., Tejumola, H. O.: *Further results for a system of third order ordinary differential equations.* Atti. Acad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. **58** (1975), 143–151.

[10] Meng, F. W.: *Ultimate boundedness results for a certain system of third order nonlinear differential equations.* J. Math. Anal. Appl. **177** (1993), 496–509.

[11] Reissig, R., Sansone, G., Conti, R.: Non-Linear Differential Equations of Higher Order. *Noordhoff, Groningen*, 1974.

[12] Tiryaki, A.: *Boundedness and periodicity results for a certain system of third order non-linear differential equations.* Indian J. Pure Appl. Math. **30**, 4 (1999). 361–372.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 21–26

On Special Almost Geodesic Mappings of Type π_1 of Spaces with Affine Connection *

Vladimir BEREZOVSKY [1], Josef MIKEŠ [2]

[1] *Department of Mathematics, Agricultural Academy of Uman'*
Institutskaya 2, Uman', Ukraine

[2] *Department of Algebra and Geometry*
Faculty of Science, Palacký University
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: Mikes@inf.upol.cz

(Received February 23, 2004)

Abstract

N. S. Sinyukov [5] introduced the concept of an *almost geodesic mapping* of a space A_n with an affine connection without torsion onto $\overline{A}_n$ and found three types: π_1, π_2 and π_3. The authors of [1] proved completness of that classification for $n > 5$.

By definition, special types of mappings π_1 are characterized by equations

$$P^h_{ij,k} + P^\alpha_{ij} P^h_{\alpha k} = a_{ij}\delta^h_k,$$

where $P^h_{ij} \equiv \overline{\Gamma}^h_{ij} - \Gamma^h_{ij}$ is the deformation tensor of affine connections of the spaces A_n and $\overline{A}_n$.

In this paper geometric objects which preserve these mappings are found and also closed classes of such spaces are described.

Key words: Almost geodesic mappings, affine connection space.

2000 Mathematics Subject Classification: 53B05, 53B99

*Supported by grant No. 201/02/0616 of The Grant Agency of the Czech Republic.

1 Introduction

In this paper the theory of almost geodesic mappings of type π_1^* of spaces with affine connection without torsion is studied. These mappings are a special case of almost geodesic mappings of type π_1 introduced by N. S. Sinyukov [5].

General properties of mappings π_1^* are shown and invariant objects with respect to these mappings are found. Mappings π_1^* of spaces of constant curvature and affine spaces are studied.

First we recall basic formulas and properties of the theory of almost geodesic mappings of spaces A_n with affine connection which are described in [5], [6].

A curve ℓ defined in a space with affine connection A_n is called *almost geodesic* if there exists a two-dimensional parallel distribution along ℓ, to which the tangent vector of this curve belongs at every point.

A diffeomorphism $f\colon A_n \to \overline{A}_n$ is an *almost geodesic mapping* if, as a result of f, every geodesic of the space A_n passes into an almost geodesic curve of the space $\overline{A}_n$.

A mapping f from A_n onto $\overline{A}_n$ is almost geodesic if and only if, in a common coordinate system $x \equiv (x^1, x^2, \dots, x^n)$ with respect to the mapping f, the connection deformation tensor $P_{ij}^h(x) \equiv \overline{\Gamma}_{ij}^h(x) - \Gamma_{ij}^h(x)$ satisfies the relations [5]

$$A^h_{\alpha\beta\gamma}\lambda^\alpha\lambda^\beta\lambda^\gamma \equiv a\, P^h_{\alpha\beta}\lambda^\alpha\lambda^\beta + b\,\lambda^h,$$

where $A^h_{ijk} \equiv P^h_{ij,k} + P^\alpha_{ij}P^h_{\alpha k}$, Γ^h_{ij} $(\overline{\Gamma}^h_{ij})$ are components of the affine connection of spaces A_n $(\overline{A}_n)$, λ^h is any vector, a and b are some functions of variables x^h and λ^h. Hereafter "," denotes the covariant derivative with respect to the connection of the space A_n.

Three types of almost geodesic mappings, π_1, π_2 and π_3, were found in [5]. We proved [1] that for $n > 5$ other types do not exist. Almost geodesic mappings of type π_1 are characterized by the following conditions

$$A^h_{(ijk)} = \delta^h_{(i}\, a_{jk)} + b_{(i}\, P^h_{jk)},$$

where a_{ij} is a symmetric tensor, b_i is a covector, δ^h_i is the Kronecker symbol, (ijk) is the symmetrization of indices.

Many papers are dedicated to study of mappings π_2 and π_3 (see [5], [6], [4]) in contrast to mappings π_1. The reason is that basic equations of these mappings are difficult to study. Therefore in this paper we deal with a special case of mappings π_1. This special case does not imply that geodesic mappings are either π_2 or π_3 mappings.

2 Almost geodesic mappings π_1^*

Let a diffeomorphism from A_n onto $\overline{A}_n$ satisfy

$$P^h_{ij,k} + P^\alpha_{ij}P^h_{\alpha k} = a_{ij}\,\delta^h_k, \tag{1}$$

where a_{ij} is a symmetric tensor.

Diffeomorphisms of this kind are a special case of almost geodesic mappings of type π_1. We denote them by π_1^*.

Let us derive the integrability condition arising from (1). We differentiate (1) covariantly by x^m and then alternate with respect to the indices k and m. Next in the integrability condition of (1) we contract with respect to the indices h and m. After editing we have

$$(n-1)\,a_{ij,k} = P^{\alpha}_{ij}R_{\alpha k} - P^{\beta}_{\alpha(i}R^{\alpha}_{j)\beta k} - (n-1)\,P^{\alpha}_{ij}a_{\alpha k}, \tag{2}$$

where R^h_{ijk} is the Riemannian tensor in A_n, $R_{ij} \equiv R^{\alpha}_{ij\alpha}$ is the Ricci tensor, $(i\,j)$ denotes the symmetrization of indices.

Evidently, equations (1) and (2) represent a system of differential equations of Cauchy type in the space A_n which is solvable with respect to unknown functions $P^h_{ij}(x)$ and $a_{ij}(x)$, which, naturally, satisfy the algebraic conditions

$$P^h_{ij}(x) = P^h_{ji}(x), \qquad a_{ij}(x) = a_{ji}(x). \tag{3}$$

We have

Theorem 1 *The space A_n with affine connection admits an almost geodesic mapping π_1^* onto $\overline{A}_n$ if and only if there exists a solution P^h_{ij} and a_{ij} of system of Cauchy type (1) and (2) satisfying (3).*

Integrability conditions of this system have the form

$$-P^{\alpha}_{ij}R^h_{\alpha km} + P^h_{\alpha(i}R^{\alpha}_{j)km} =$$
$$\frac{1}{n-1}\left[(P^{\alpha}_{ij}R_{\alpha m} - P^{\beta}_{\alpha(i}R^{\alpha}_{j)m\beta})\delta^h_k - (P^{\alpha}_{ij}R_{\alpha k} - P^{\beta}_{\alpha(i}R^{\alpha}_{j)k\beta})\delta^h_m\right],$$

$$(n-1)a_{\alpha(i}R^{\alpha}_{j)km} = P^{\alpha}_{ij}R^h_{\alpha[k,m]} + P^{\beta}_{\alpha(i}R^{\alpha}_{j)mk,\beta}$$
$$+ R_{[mk]}a_{ij} + P^{\beta}_{\gamma[m}R^h_{|i|k]\beta}P^{\gamma}_{\alpha j} + P^{\alpha}_{ij}P^{\beta}_{\alpha\gamma}R^{\gamma}_{[km]\beta} - P^{\alpha}_{ij}P^{\beta}_{\gamma[k}R^{\gamma}_{|\alpha|m]\beta},$$

where $[i\,j]$ denotes the alternation of indices.

3 Invariant object of mappings π_1^*

If P^h_{ij} is the deformation tensor ([5]) then Riemannian tensors R^h_{ijk} and $\overline{R}^h_{ijk}$ of spaces A_n and $\overline{A}_n$ satisfy the following condition

$$\overline{R}^h_{ijk} = R^h_{ijk} + P^h_{i[k,j]} + P^{\alpha}_{i[k}P^h_{j]\alpha}. \tag{4}$$

Using formulas (1) and (4) we obtain

$$\overline{\overset{*}{W}}{}^h_{ijk} = \overset{*}{W}{}^h_{ijk}, \tag{5}$$

where

$$\overset{*}{W}{}^h_{ijk} \equiv R^h_{ijk} - \frac{1}{n-1}R_{i[j}\delta^h_{k]}, \qquad \overline{\overset{*}{W}}{}^h_{ijk} \equiv \overline{R}^h_{ijk} - \frac{1}{n-1}\overline{R}_{i[j}\delta^h_{k]}. \tag{6}$$

Clearly, $\overset{*}{W}{}^h_{ijk}$ and $\overline{\overset{*}{W}}{}^h_{ijk}$ is a tensor of type $\binom{1}{3}$ in the space A_n and $\overline{A}_n$, respectively.

Condition (5) shows that this tensor is invariant with respect to almost geodesic mappings π_1^*.

We contract condition (5) in indices h and i to obtain the equality

$$W_{ij} = \overline{W}_{ij}, \tag{7}$$

where

$$W_{ij} \equiv R_{[ij]}, \qquad \overline{W}_{ij} \equiv \overline{R}_{[ij]}, \tag{8}$$

Subtract (7) from (5) to write

$$W^h_{ijk} = \overline{W}^h_{ijk}, \tag{9}$$

where W^h_{ijk} and $\overline{W}^h_{ijk}$ are Weyl projective curvature tensors of spaces A_n and $\overline{A}_n$, respectivelly. We get

Theorem 2 *The Weyl projective curvature tensor W^h_{ijk} and tensors $\overset{*}{W}{}^h_{ijk}$ and W_{ij}, which are defined by (6) and (8), are invariant with respect to almost geodesic mappings π_1^*.*

4 Mappings π_1^* of affine and projective-euclidean spaces

From Theorem 2 it follows

Theorem 3 *If a projective-euclidean space or equiaffine space admits an almost geodesic mapping π_1^* onto $\overline{A}_n$ then $\overline{A}_n$ is also a projective-euclidean space or an equiaffine space.*

The proof of Theorem 3, evidently, follows from the condition $W^h_{ijk} = 0$ in the projective-euclidean space and from the condition $W_{ij} = 0$ in the equiaffine space.

It means that projective-euclidean spaces and equiaffine spaces make up closed classes with respect to mappings π_1^*.

Clearly, the Riemannian tensor is preserved by mappings π_1^* if and only if the tensor a_{ij} vanishes. In this case basic equations have the form

$$P^h_{ij,k} = -P^\alpha_{ij} P^h_{\alpha k}. \tag{10}$$

Equations (10) are completely integrable in the affine space. Evidently, these equations have a solution for any initial conditions $P^h_{ij}(x_o)$.

If the initial conditions are such that $P^h_{ij}(x_o) \not\equiv \delta^h_{(i}\psi_{j)}(x_o)$ then every solution generates a mapping π_1^* which is not a geodesic mapping of the affine space A_n onto the affine space $\overline{A}_n$. Therefore we can write

Theorem 4 *Mappings π_1^* of an affine space A_n onto itself exist. All lines map into planar curves (not necessary lines).*

Moreover, integrability conditions (1) and (2) in affine space are always true. We obtain

Theorem 5 *Riemannian spaces V_n with non constant curvature admit non geodesic mappings π_1^* which are necessarilly mappings of type π_3 and preserve the quadratic complex of geodesics.*

Proof Let a Riemannian space V_n with non constant curvature K admit a non geodesic mapping π_1^*. Integrability conditions (1) then have the form

$$K(P^h_{k(i}g_{j)l} - P^h_{l(i}g_{j)k}) + \delta^h_l B_{ijk} - \delta^h_k B_{ijl} = 0, \tag{11}$$

where $B_{ijk} \equiv a_{ij,k} + P^\alpha_{ij}(a_{\alpha k} + K\, g_{\alpha k})$, g_{ij} is the metric tensor of the space V_n.

From the last formula it follows

$$P^h_{ij} = P^h g_{ij} \tag{12}$$

where P^h is a vector. Then the mapping is F-planar [4]. Clearly, on the basis of results in [1], such mappings are almost geodesic mappings of type π_3. It is proved in the paper [1] that mappings $\pi_1 \cap \pi_3$ preserve the quadratic complex of geodesics [3].

After substituting (12) in (1) we have

$$P^h_{,k} + P^h P_k = \alpha \delta^h_k,$$

where α is a function, P_k is a covector.

These conditions characterize concircular vector fields P^h, which always exist in spaces with constant curvature. □

5 Examples of almost geodesic mappings π_1^*

We present an example of an almost geodesic mapping of type π_1^* of an affine space A_n onto an affine space $\overline{A}_n$.

Let $x^1, x^2, \dots, x^n$ and $\overline{x}^1, \overline{x}^2, \dots, \overline{x}^n$ be affine coordinate in A_n and $\overline{A}_n$, respectively.

The mapping

$$\overline{x}^h = \frac{1}{2} C^h_\alpha (x^\alpha - C^\alpha)^2 + x^h_o, \tag{13}$$

where C^h_i, C^h, x^h_o are some constants, $x^h \neq C^h$, and the determinant $\det\left|C^h_i\right| \not\equiv 0$, defines an almost geodesic mapping π_1^* of the space A_n onto $\overline{A}_n$.

We can prove directly that the deformation tensor P^h_{ij} in the coordinate system $x^1, x^2, \dots, x^n$ has the form

$$P^i_{ii} = \frac{1}{x^i - C^i}, \qquad i = \overline{1, n},$$

and the other components are equal to zero.

Evidently, the tensor P_{ij}^h corresponds to equations (10). This mapping is not of type π_2 or π_3.

Lines in the space A_n which are defined by equations $x^h = a^h + b^h\, t$ where t is the parameter, map into parabolas (or lines) of the space $\overline{A}_n$, which are defined by equations

$$\overline{x}^h = D^h + E^h\, t + F^h\, t^2$$

where

$$D^h = \frac{1}{2} C_\alpha^h (a^\alpha - C^\alpha)^2, \quad E^h = C_\alpha^h (a^\alpha - C^\alpha) b^\alpha, \quad F^h = \frac{1}{2} C_\alpha^h (b^\alpha)^2$$

in this mapping.

The image is a line if vectors E^h and F^h are collinear.

Finally we remark that formula (13) generates a system of almost geodesic mappings of type π_1 of planar spaces if the coefficients C_i^h, C^h and x_o^h are continuous.

References

[1] Berezovsky, V. E., Mikeš, J.: *On the classification of almost geodesic mappings of affine-connected spaces.* In: Proc. Conf., Dubrovnik (Yugoslavia) 1988, 41–48 (1989).

[2] Berezovsky, V. E., Mikeš, J.: *On almost geodesic mappings of the type* π_1 *of Riemannian spaces preserving a system* n*-orthogonal hypersurfaces.* Suppl. Rend. Circ. Mat. Palermo, II. Ser. 59, 103–108 (1999).

[3] Chernyshenko, V. M.: *Räume mit einem speziellen Komplex von geodätischen Linien.* Tr. Semin. Vektor. Tenzor. Anal. **11** (1961), 253–268 (in Russian).

[4] Mikeš, J.: *Holomorphically projective mappings and their generalizations,* J. Math. Sci., New York **89**, 3 (1998), 1334–1353.

[5] Sinyukov, N. S.: On geodesic mappings of Riemannian spaces. *Nauka, Moscow*, 1979 (in Russian).

[6] Sinyukov, N. S.: *Almost geodesic mappings of affine connected and Riemannian spaces.* Itogi Nauki Tekh., Ser. Probl. Geom. **13** (1982), 3–26 (in Russian); J. Sov. Math. **25** (1984), 1235–1249.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 27–32

Deductive Systems of BCK-Algebras

Sergio A. CELANI

CONICET and Facultad de Ciencias Exactas
Universidad Nacional del Centro
Pinto 399, 7000 Tandil, Argentina
e-mail: scelani@exa.unicen.edu.ar

(Received October 16, 2003)

Abstract

In this paper we shall give some results on irreducible deductive systems in BCK-algebras and we shall prove that the set of all deductive systems of a BCK-algebra is a Heyting algebra. As a consequence of this result we shall show that the annihilator F^* of a deductive system F is the the pseudocomplement of F. These results are more general than that the similar results given by M. Kondo in [7].

Key words: BCK-algebras, deductive system, irreducible deductive system, Heyting algebras, annihilators.

2000 Mathematics Subject Classification: 03F35, 03G25

1 Introduction and preliminaries

In [7] it was shown that the set of all ideals (or deductive systems, in our terminology) of a BCK-algebra $\mathbf{A}$ is a pseudocomplement distributive lattice and that the annihilator F^* of a deductive system F of $\mathbf{A}$ is the pseudocomplement of F. Related results on annihilators in Hilbert algebras and Tarski algebras (or also called commutative Hilbert algebras [6] or Abbot's implication algebras) are given in [2] and [3]. On the other hand, it was shown in [9] that the set of deductive systems $Ds(\mathbf{A})$ of a BCK-algebra $\mathbf{A}$ is an infinitely distributive lattice, and thus it is a Heyting algebra. In this note we will give a description of this fact and we shall prove that the annihilator F^* of the deductive system F can be obtained as $F^* = F \Rightarrow \{1\}$, where $\Rightarrow$ is the Heyting implication defined in the lattice $Ds(\mathbf{A})$.

In the remaining part of this section we shall review some results on BCK-algebras. In section 2 we shall study the notion of irreducible deductive system. In particular, we shall give a generalization of a result given in [8] for BCK-algebras with supremum. In Section 3 we shall prove that the lattice of deductive system of a BCK-algebra is a Heyting algebra.

Definition 1 An algebra $\mathbf{A} = \langle A, \rightarrow, 1 \rangle$ of type $(2, 0)$ is a *BCK-algebra* if for all $a, b, c \in A$ the following conditions hold:

1. $a \rightarrow a = 1$,
2. $(a \rightarrow b) \rightarrow ((b \rightarrow c) \rightarrow (a \rightarrow c)) = 1$,
3. $a \rightarrow (b \rightarrow c) = b \rightarrow (a \rightarrow c)$,
4. $a \rightarrow (b \rightarrow a) = 1$
5. $a \rightarrow b = 1$ and $b \rightarrow a = 1$, implies $a = b$.

If $\mathbf{A}$ is a BCK-algebra and we define the binary relation $\leq$ on $\mathbf{A}$ by $a \leq b$ if and only if $a \rightarrow b = 1$, then $\leq$ is a partial order in $\mathbf{A}$.

Let us recall that in all BCK-algebras $\mathbf{A}$ the following properties are satisfied:

P1 $1 \rightarrow a = a$,

P2 $a \rightarrow ((a \rightarrow b) \rightarrow b) = 1$

P3 $a \rightarrow b \leq (c \rightarrow b) \rightarrow (c \rightarrow a)$,

P4 $a \rightarrow b = ((a \rightarrow b) \rightarrow b) \rightarrow b$,

P5 if $a \leq b$, then $c \rightarrow a \leq c \rightarrow b$ and $b \rightarrow c \leq a \rightarrow c$.

A BCK-algebra *with supremum*, or BCK$^{\vee}$-algebra is an algebra

$$\mathbf{A} = \langle A, \rightarrow, \vee, 1 \rangle$$

where $\langle A, \rightarrow, 1 \rangle$ is a BCK-algebra, $\langle A, \vee, 1 \rangle$ is a join-semilattice, and $a \rightarrow b = 1$ if and only if $a \vee b = b$. For $a, b \in A$ we define inductively $a \rightarrow_n b$ as $a \rightarrow_0 b = b$ and $a \rightarrow_{n+1} b = a \rightarrow ((a \rightarrow_n b))$.

Let $\mathbf{A}$ be a BCK-algebra. A *deductive system* or *filter* of $\mathbf{A}$ is a nonempty subset F of A such that $1 \in F$, and for every $a, b \in A$, if a, $a \rightarrow b \in F$, then $b \in F$. It is clear that if F is a deductive system, $a \leq b$ and $a \in F$, then $b \in F$. The set of all deductive system of a BCK-algebra $\mathbf{A}$ is denoted by $Ds(\mathbf{A})$. The deductive system generated by a set $X \subseteq A$ is denoted by $\langle X \rangle$. Let us recall that

$$\langle X \rangle = \{a \in A : x_1 \rightarrow (\ldots (x_n \rightarrow a) \ldots) = 1 \text{ for some } x_1, \ldots, x_n \in X\}.$$

In particular, $\langle x \rangle = \{a \in A : x \rightarrow (\ldots (x \rightarrow a) \ldots) = x \rightarrow_n a = 1\}$.

Let $\mathbf{A}$ be a BCK-algebra. In [9] (see also [10]) it was proved that the structure $\langle Ds(\mathbf{A}), \vee, \wedge, \{1\}, A \rangle$ is a bounded (infinitely) distributive lattice where the operations $\wedge$ and $\vee$ are defined by:

$$\begin{aligned} F_1 \wedge F_2 &= F_1 \cap F_2 \\ F_1 \vee F_2 &= \{a \in A : \exists (x, y) \in F_1 \times F_2;\ x \rightarrow (y \rightarrow a) = 1\}. \end{aligned}$$

We note that

$$F \vee \langle a \rangle = \{c \in A : a \to_n c \in F \text{ for some } n \geq 0\}$$

for $F \in Ds(\mathbf{A})$ and $a \in A$. Indeed, let $c \in F \vee \langle a \rangle$. Then there exist $x \in F$ and $n \geq 0$ such that $x \to (y \to c) = 1$ and $a \to_n y = 1$. Since $x \to (y \to c) = 1 \in F$, $y \to c \in F$. So, $y \to c \leq (a \to_n y) \to (a \to_n c) = 1 \to (a \to_n c) = a \to_n c \in F$.

2 Irreducible deductive systems

In [8] the separation theorem for BCK$^{\vee}$-algebras was proved. In this section following the paper [1], we prove a separation theorem for any BCK-algebra.

Let $\mathbf{A}$ be a BCK-algebra. A deductive system F is *irreducible* if and only if for any $F_1, F_2 \in Ds(\mathbf{A})$ such that $F = F_1 \cap F_2$, we have $F = F_1$ or $F = F_2$. We denote by $X(\mathbf{A})$ the set of all irreducible deductive systems of a BCK-algebra $\mathbf{A}$.

Lemma 2 *Let* $\mathbf{A}$ *be a BCK-algebra. Let* $F \in Ds(\mathbf{A})$. *Then* F *is irreducible if and only if for every* $a, b \notin F$ *there exist* $c \notin F$ *and* $n \geq 0$ *such that* $a \to_n c$, $b \to_n c \in F$.

Proof $\Rightarrow$) Let $a, b \notin F$. Let us consider the deductive systems $F_a = \langle F \cup \{a\} \rangle = F \vee \langle a \rangle$ and $F_b = \langle F \cup \{b\} \rangle = F \vee \langle b \rangle$. Since $F \neq F_a$ and $F \neq F_b$, then by irreducibility of F we have $F \subset F_a \cap F_b$. It follows that there exists $c \in (F_a \cap F_b) - F$. Then $a \to_n c \in F$ and $b \to_m c \in F$ for some $n, m \geq 0$. If we assume that $n \geq m$, then by property P4 we have that $b \to_m c \leq b \to_n c$. So, $a \to_n c \in F$ and $b \to_n c \in F$.

$\Leftarrow$). Let $F_1, F_2 \in Ds(\mathbf{A})$ such that $F = F_1 \cap F_2$. Suppose that $F \neq F_1$ and $F \neq F_2$. Then there exist $a \in F_1 - F$ and $b \in F_2 - F$. So, by the assumption, there exists $c \notin F$ and $n \geq 0$ such that $a \to_n c \in F$ and $b \to_n c \in F$. As, a, $a \to_n c \in F_1$ and $F_1 \in Ds(\mathbf{A})$, then $c \in F_1$. Similarly, $c \in F_2$. Thus, $c \in F_1 \cap F_2 = F$, which is a contradiction. □

Let $\mathbf{A}$ be a BCK-algebra. A subset I of A is called an *ideal* of $\mathbf{A}$ if:

1. If $b \in I$ and $a \leq b$, then $a \in I$.
2. If $a, b \in I$ there exists $c \in I$ such that $a \leq c$ and $b \leq c$.

The set of all ideals of $\mathbf{A}$ will be denoted by $Id(\mathbf{A})$.

Theorem 3 *Let* $\mathbf{A}$ *be a BCK-algebra. Let* $F \in Ds(\mathbf{A})$ *and* $I \in Id(\mathbf{A})$ *such that* $F \cap I = \emptyset$. *Then there exists* $P \in X(\mathbf{A})$ *such that* $F \subseteq P$ *and* $P \cap I = \emptyset$.

Proof Let us consider the following subset of $Ds(\mathbf{A})$:

$$\mathcal{F} = \{H \in Ds(\mathbf{A}) : F \subseteq H \text{ and } H \cap I = \emptyset\}.$$

Since $F \in \mathcal{F}$, then $\mathcal{F} \neq \emptyset$. It is clear that the union of a chain of elements of $\mathcal{F}$ is also in $\mathcal{F}$. So, by Zorn's lemma, there exists a maximal element P of $\mathcal{F}$. We prove that $P \in X(\mathbf{A})$. Let $a, b \notin P$ and let us consider the deductive systems $P_a = \langle P \cup \{a\} \rangle$ and $P_b = \langle P \cup \{b\} \rangle$. Clearly, $P \subset P_a \cap P_b$. Then, $P_a, P_b \notin \mathcal{F}$. Thus, $P_a \cap I \neq \emptyset$ and $P_a \cap I \neq \emptyset$. It follows that there exist $x, y \in I$ such that $a \to_n x \in P$ and $b \to_m y \in P$ for some $n, m \geq 0$. Suppose that $m \leq n$. Then $b \to_m y \leq b \to_n y \in P$. Since I is an ideal, there exists $c \in I$ such that $x \leq c$ and $y \leq c$. So, $a \to_n x \leq a \to_n c \in P$ and $b \to_n y \leq b \to_n c \in P$. Therefore, by Lemma 2, we conclude that $P \in X(\mathbf{A})$. □

Corollary 4 *Let* $\mathbf{A}$ *be a BCK-algebra. Let* $F \in Ds(\mathbf{A})$.

1. *For each* $a \notin F$ *there exists* $P \in X(\mathbf{A})$ *such that* $a \notin P$ *and* $F \subseteq P$.
2. $F = \bigcap \{P \in X(\mathbf{A}) : F \subseteq P\}$.

3 Annihilators

Let us recall that a *Heyting algebra* is an algebra $\langle A, \vee, \wedge, \Rightarrow, 0, 1 \rangle$ of type $(2, 2, 2, 0, 0)$ such that $\langle A, \vee, \wedge, 0, 1 \rangle$ is a bounded distributive lattice and the operation $\Rightarrow$ satisfies the condition: $a \wedge b \leq c$ if and only if $a \leq b \Rightarrow c$, for all $a, b, c \in A$. The *pseudocomplement* of an element $x \in A$ is the element $x^* = x \Rightarrow 0$.

Let $\mathbf{A}$ be a BCK-algebra. Let $a \in A$. Define the set $[a) = \{x \in A : a \leq x\}$. We note that in general the set $[a) \notin Ds(\mathbf{A})$.

For each pair $F, H \in Ds(\mathbf{A})$ let us define the subset $F \Rightarrow H$ of A as follows:

$$F \Rightarrow H = \{a \in A : [a) \cap F \subseteq H\}.$$

Theorem 5 *Let* $\mathbf{A}$ *be a BCK-algebra. Let* $F, H \in Ds(\mathbf{A})$. *Then*

1. $F \Rightarrow H \in Ds(\mathbf{A})$.
2. $F \Rightarrow H = \{x \in A : (x \to f) \to f \in H$ *for each* $f \in F\}$.
3. $\langle Fi(A), \vee, \wedge, \Rightarrow, \{1\}, A \rangle$ *is a Heyting algebra.*

Proof 1. Since, $[1) \cap F = \{1\} \subseteq H$, then $1 \in F \Rightarrow H$.

Let x, $x \to y \in F \Rightarrow H$. Then, $[x) \cap F \subseteq H$ and $[x \to y) \cap F \subseteq H$. Let $z \in [y) \cap F$. As, $y \leq z$, then by the property P5, $x \to y \leq x \to z$. By property P4., we have $x \to z \in F$. Thus,

$$x \to z \in [x \to y) \cap F.$$

On the other hand, as $x \leq (x \to z) \to z$ and $z \leq (x \to z) \to z$, we get $(x \to z) \to z \in [x) \cap F$. Therefore,

$$x \to z,\ (x \to z) \to z \in H,$$

and consequently $z \in H$. So, $F \Rightarrow H \in Ds(\mathbf{A})$.

2. We prove that

$$F \Rightarrow H \subseteq G = \{x \in A : (x \to f) \to f \in H \text{ for each } f \in F\}.$$

Let $x \in A$ such that $[x) \cap F \subseteq H$. Let $f \in F$. Since, $x \leq (x \to f) \to f$ and $f \leq (x \to f) \to f$, then $(x \to f) \to f \in [x) \cap F \subseteq H$. Thus, $x \in G$.

Let $x \in G$. Let $y \in A$ such that $x \leq y$ and $y \in F$. Since $(x \to y) \to y \in H$ and $x \to y = 1$, then $1 \to y = y \in H$. Thus, $x \in F \Rightarrow H$.

3. Let $F, H, K \in Ds(\mathbf{A})$. Then it is easy to check that

$$F \cap H \subseteq K \text{ if and only if } F \subseteq H \Rightarrow K.$$

Thus, $\langle Ds(\mathbf{A}), \vee, \wedge, \Rightarrow, \{1\}, A\rangle$ is a Heyting algebra. □

As a corollary we have the following result, first given by M. Kondo in [7].

Corollary 6 *Let* $\mathbf{A}$ *be a BCK-algebra. The annihilator of a deductive system* F *is the deductive system*

$$F^* = F \Rightarrow \{1\} = \{x \in A : [x) \cap F = \{1\}\}.$$

Proof It is immediate by the above theorem. □

For BCK$^{\vee}$-algebras we can give the following result which generalize a similar result given by M. Kondo in [7] for commutative BCK-algebras.

Proposition 7 *Let* $\mathbf{A}$ *be a BCK*$^{\vee}$*-algebra. Then for every* $F \in Ds(\mathbf{A})$

$$F^* = \{x \in A : x \vee f = 1 \text{ for each } f \in F\}.$$

Proof Let $x \in A$ such that $x \vee f = 1$ for each $f \in F$. We prove that $[x) \cap F = \{1\}$. Let $a \in A$ such that $x \leq a$ and $a \in F$. Then $a = x \vee a = 1$. Thus, $x \in F^*$.

Let $x \in F^*$. Then $[x) \cap F = \{1\}$. Since $x \leq x \vee f$, $f \leq x \vee f$, for each $f \in F$, and as F is increasing, then $x \vee f \in [x) \cap F$. Thus, $x \vee f = 1$, for each $f \in F$. □

Now we prove that the annihilator of a subset X is the annihilator of the deductive system generated by X. This result was proved for Tarski algebras in [2].

Theorem 8 *Let* $\mathbf{A}$ *be a BCK*$^{\vee}$*-algebra. Then for every subset* X *of* A*, we have* $X^* = \langle X\rangle^*$.

Proof Since $X \subseteq \langle X\rangle$, then $\langle X\rangle^* \subseteq X^*$. Let $x \in X^*$. We prove that for every $a \in \langle X\rangle$, $x \vee a = 1$. Suppose that there exists $a \in \langle X\rangle$ such that $a \vee x \neq 1$. Then there exist $x_1, \ldots, x_k \in X$ such that

$$x_1 \to (x_2 \to \ldots (x_k \to a) \ldots) = 1.$$

As $x \in X^*$, $x \vee x_i = 1$ for every $x_i \in \{x_1, \ldots, x_k\}$. Since, $a \vee x \neq 1$, by Theorem 3 there exists an irreducible deductive system P such that $x \notin P$, $a \notin P$ and taking into account that $x \vee x_i = 1$, then $x_i \in P$ for every $x_i \in \{x_1, \ldots, x_k\}$. But since, $x_1 \to (x_2 \to \ldots (x_k \to a) \ldots) = 1 \in P$, then $a \in P$, which is a contradiction. Thus, $a \vee x = 1$ for every $a \in \langle X\rangle$ and consequently $x \in \langle X\rangle^*$. □

References

[1] Celani, S.: *A note on Homomorphisms of Hilbert Algebras.* International Journal of Mathematics and Mathematical Science **29**, 1 (2002), 55–61.

[2] Chajda, I., Halaš, R., Zedník, J.: *Filters and Annihilators in Implication Algebras.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **37** (1998), 41–45.

[3] Chajda, I., Halaš, R., Jun, J. B. *Annihilators and deductive systems in commutative Hilbert algebras.* Comment. Math. Univ. Carolinae **43**, 3 (2002) 407–417.

[4] Jun, Y. B., Roh, E. H., Meng, J.: *Annihilators in BCI-algebras.* Math. Japonica **43**, 3 (1996), 559–562.

[5] Halaš, R.: *Annihilators of BCK-algebras.* Czech. Math. Journal **53**, 128 (2003), 1001–1007.

[6] Halaš, R.: *Remarks on commutative Hilbert algebras.* Math. Bohemica **127**, 4 (2002), 525–529.

[7] Kondo, M.: *Annihilators in BCK-algebras.* Math. Japonica **49**, 3 (1999), 407–410.

[8] Pałasiński, M.: *Ideals in BCK-algebras which are lower semilattices.* Math. Japonica **26**, 2 (1981), 245–250.

[9] Pałasiński, M.: *On ideal and congruence lattices of BCK-algebras.* Math. Japonica **26**, 5 (1981), 543–544.

[10] Wei, S. M., Jun, Y. B.: *Ideal lattices of BCI-algebras.* Math. Japonica **44**, 2 (1996), 303–305.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 33–53

Impulsive Periodic Boundary Value Problem

Jan DRAESSLER

University of Hradec Králové
e-mail: jan.draessler@uhk.cz

(Received February 10, 2004)

Abstract

In the paper we consider the impulsive periodic boundary value problem with a general linear left hand side. The results are based on the topological degree theorems for the corresponding operator equation $(I - F)u = 0$ on a certain set Ω that is established using properties of strict lower and upper functions of the boundary value problem.

Key words: Boundary value problem, topological degree, upper and lower functions, impulsive problem, periodic solution, differential equation.

2000 Mathematics Subject Classification: 34B37, 34C25

1 Introduction

In this paper we will study the boundary value problem

$$x'' + a(t)x' + b(t)x = f(t, x, x') \tag{1.1}$$

$$x(t_1+) = J(x(t_1)), \quad x'(t_1+) = M(x'(t_1-)), \tag{1.2}$$

$$x(0) = x(2\pi), \quad x'(0) = x'(2\pi). \tag{1.3}$$

We suppose, that a, b are Lebesgue integrable functions on $[0, 2\pi]$ and f fulfils the Carathéodory conditions on $[0, 2\pi] \times \mathrm{R}^2$. Furthermore, we suppose that $t_1 \in (0, 2\pi)$ and

$$\begin{gathered} J, M \text{ are continuous mappings } R \to R \text{ and,} \\ J \text{ is increasing on } R \text{ and } M \text{ is nondecreasing on } R. \end{gathered} \tag{1.4}$$

Our main assertions (Theorem 3.2 and Theorem 3.6) are based on the properties of the Leray–Schauder topological degree. We search an operator problem $u = Fu$ which corresponds to (1.1)–(1.3) and such that operator $I - F$ has nonzero topological degree on a certain set Ω. For establishing Ω the existence of strict lower and upper functions of the problem is assumed.

We consider two cases of ordering of strict lower and upper functions σ_1 and σ_2:

i) The functions are well ordered i.e. $\sigma_1(t) < \sigma_2(t)$ for all $t \in [0, 2\pi]$. In this case, we get the existence of a solution u which lies between the strict lower and upper functions i.e. $\sigma_1(t) < u(t) < \sigma_2(t)$ on $[0, 2\pi]$ (Corollary 3.3).

ii) The functions are in the opposite order i.e. $\sigma_2(t) < \sigma_1(t)$ for all $t \in [0, 2\pi]$. In this case, we get the existence of a solution u, at least one point of which lies between the strict functions i.e. $\sigma_2(t_u) < u(t_u) < \sigma_1(t_u)$ for some $t_u \in [0, 2\pi]$ (Corollary 3.7).

This work generalizes the results published in [1],[2] where the equation $x'' = f(t, x, x')$, which is a special case of the equation (1.1), has been studied.

1.1 Definitions

$L[0, 2\pi]$ is the Banach space of the Lebesgue integrable functions on $[0, 2\pi]$ with the norm $\|x\|_1 = \int_0^{2\pi} |x(t)| dt$.

$L_\infty[0, 2\pi]$ denotes the Banach space of essentially bounded functions on $[0, 2\pi]$ with the norm $\|x\|_\infty = \operatorname{ess\,sup}\{|x(t)|; t \in [0, 2\pi]\}$.

$C[0, 2\pi]$ and $C^1[0, 2\pi]$ are the spaces of functions continuous on $[0, 2\pi]$ and of functions with continuous first derivatives on $[0, 2\pi]$, respectively.

Similarly, $AC[0, 2\pi]$ and $AC^1[0, 2\pi]$ denote spaces of functions absolutely continuous on $[0, 2\pi]$ and of functions with absolutely continuous first derivatives on $[0, 2\pi]$, respectively.

Let $t_1 \in (0, 2\pi)$. Then $\widetilde{C}^1[0, 2\pi]$ means the set of functions

$$u(t) = \begin{cases} u_1(t) \text{ for } 0 \le t \le t_1 \\ u_2(t) \text{ for } t_1 < t \le 2\pi \end{cases},$$

where $u_1 \in C^1[0, t_1]$ and $u_2 \in C^1[t_1, 2\pi]$. $\widetilde{AC}^1[0, 2\pi]$ specifies the set of functions $u \in \widetilde{C}^1[0, 2\pi]$ with absolutely continuous first derivatives on $(0, t_1)$ and on $(t_1, 2\pi)$. For $u \in \widetilde{C}^1[0, 2\pi]$ we establish

$$u'(0) = \lim_{\tau \to 0+} u'(\tau), \quad u'(2\pi) = \lim_{\tau \to 2\pi-} u'(\tau),$$

$$u'(t_1) = \lim_{\tau \to t_1-} u'(\tau),$$

$$\|u\|_{\widetilde{C}^1} = \|u\|_\infty + \|u'\|_\infty.$$

Moreover, for $u \in \widetilde{C}^1[0, 2\pi]$ and $t \in (0, 2\pi)$ we will use notations

$$\Delta u(t) = u(t+) - u(t), \quad \Delta u'(t) = u'(t+) - u'(t). \tag{1.5}$$

$\widetilde{C}^1[0, 2\pi]$ with the norm $\| \cdot \|_{\widetilde{C}^1}$ is the Banach space.

Definition 1.1 By a solution of the impulsive problem (1.1)–(1.3) we call $u \in \widetilde{AC}^1[0, 2\pi]$ which fulfils the equation (1.1) for a.e. $t \in [0, 2\pi]$ and satisfies conditions (1.2) and (1.3).

By a solution of the problem (1.1), (1.3) (without impulses) we call $u \in AC^1[0, 2\pi]$ which fulfils the equation (1.1) for a.e. $t \in [0, 2\pi]$ and satisfies conditions (1.3).

Definition 1.2 A function $\sigma_1 \in \widetilde{AC^1}[0, 2\pi]$ is a lower function of (1.1)–(1.3) if

$$\sigma_1'' + a(t)\sigma_1' + b(t)\sigma_1 \geq f(t, \sigma_1, \sigma_1') \ \textit{for a.e.}\ t \in [0, 2\pi], \tag{1.6}$$

$$\sigma_1(t_1+) = J(\sigma_1(t_1)),\ \sigma_1'(t_1+) \geq M(\sigma_1'(t_1)), \tag{1.7}$$

$$\sigma_1(0) = \sigma_1(2\pi),\ \sigma_1'(0) \geq \sigma_1'(2\pi). \tag{1.8}$$

Definition 1.3 A function $\sigma_2 \in \widetilde{AC^1}[0, 2\pi]$ is an upper function of (1.1)–(1.3) if

$$\sigma_2'' + a(t)\sigma_2' + b(t)\sigma_2 \leq f(t, \sigma_2, \sigma_2') \ \textit{for a.e.}\ t \in [0, 2\pi], \tag{1.9}$$

$$\sigma_2(t_1+) = J(\sigma_2(t_1)),\ \sigma_2'(t_1+) \leq M(\sigma_2'(t_1)), \tag{1.10}$$

$$\sigma_2(0) = \sigma_2(2\pi),\ \sigma_2'(0) \leq \sigma_2'(2\pi). \tag{1.11}$$

Definition 1.4 A lower function σ_1 of (1.1)–(1.3) is a strict lower function of (1.1)–(1.3) if it is not a solution of (1.1)–(1.3) and there exists $\varepsilon > 0$ such that

$$\sigma_1'' + a(t)y + b(t)x \geq f(t, x, y) \text{ for a.e. } t \in [0, 2\pi] \tag{1.12}$$

and each

$$x \in [\sigma_1(t), \sigma_1(t) + \varepsilon],\ y \in [\sigma_1'(t) - \varepsilon, \sigma_1'(t) + \varepsilon].$$

Similarly, an upper function σ_2 of (1.1)–(1.3) is a strict upper function of (1.1)–(1.3) if it is not a solution of (1.1)–(1.3) and there exists $\varepsilon > 0$ such that

$$\sigma_2'' + a(t)y + b(t)x \leq f(t, x, y) \text{ for a.e. } t \in [0, 2\pi] \tag{1.13}$$

and each

$$x \in [\sigma_2(t) - \varepsilon, \sigma_2(t)],\ y \in [\sigma_2'(t) - \varepsilon, \sigma_2'(t) + \varepsilon].$$

2 Auxiliary problem

In this chapter we will study the auxiliary Dirichlet problem and present assertions which consist of the relation of the strict lower and upper functions to a solution of the auxiliary problem. The assertions will be used in next chapters.

Consider the boundary value problem

$$x'' + a(t)x' + b(t)x = h(t), \tag{2.1}$$

$$x(0) = x(2\pi) = c, \tag{2.2}$$

where $h \in L[0, 2\pi]$ and $c \in$R and the corresponding homogeneous problem

$$x'' + a(t)x' + b(t)x = 0, \tag{2.3}$$

$$x(0) = x(2\pi) = 0. \tag{2.4}$$

We study two cases of the problem:

i) The problem (2.3), (2.4) has only the trivial solution. In this case there is the Green function of (2.3), (2.4) and we can prove that there exists an operator F corresponding to (2.1), (2.2) such that every solution u of $x = Fx$ fulfils

$$u(t_1+) = u(t_1) + d, \;\; u'(t_1+) = u'(t_1) + e, \quad d, e \in \mathrm{R}. \tag{2.5}$$

ii) The problem (2.3), (2.4) has the nontrivial solution. In this case we transform the problem to an equivalent form to be able to use the way in i).

Lemma 2.1 *Let the homogeneous boundary value problem (2.3), (2.4) has only the trivial solution. Then there exists a unique solution* $u \in \widetilde{AC}^1[0, 2\pi]$ *of the impulsive problem (2.1), (2.2), (2.5).*

The solution can be written in the form

$$u = c + \widetilde{g}(t, t_1)d + g(t, t_1)e + \int_0^{2\pi} g(t, s)[h(s) - cb(s)]ds, \tag{2.6}$$

where $g(t, s)$ is the Green function of (2.3),(2.4) and $\widetilde{g}(t, s)$ is a function which fulfills (2.3) for a.e. $t \in [0, s) \cup (s, 2\pi]$ and each fixed $s \in [0, 2\pi]$ and satisfies conditions (2.4) and

$$\widetilde{g}(s+, s) = \widetilde{g}(s, s) + 1, \qquad \left.\frac{\partial \widetilde{g}(t, s)}{\partial t}\right|_{t \to s+} = \left.\frac{\partial \widetilde{g}(t, s)}{\partial t}\right|_{t \to s-} \tag{2.7}$$

for each $s \in (0, 2\pi)$. At first, we need to prove that such function $\widetilde{g}(t, s)$ exists.

Lemma 2.2 *Let the homogeneous boundary value problem (2.3), (2.4) has only the trivial solution. Then for each fixed* $s \in [0, 2\pi]$ *there exists a function* $\widetilde{g}$ *which fulfills (2.3) for a.e* $t \in [0, 2\pi]$ *and satisfies (2.4), (2.7).*

Proof Consider fixed $s \in [0, 2\pi]$ and a problem (2.4),

$$x''(t) + a(t)x'(t) + b(t)x(t) = \widetilde{h}_s(t), \tag{2.8}$$

where

$$\widetilde{h}_s(t) = h_s^*(t) + \frac{1}{2\pi}a(t) - \Big(1 - \frac{1}{2\pi}t\Big)b(t), \qquad h_s^*(t) = \begin{cases} b(t) & \text{for } t \le s \\ 0 & \text{for } s < t \end{cases}.$$

Since the corresponding homogeneous problem has only the trivial solution and $\widetilde{h}_s \in L[0, 2\pi]$ then there exists a solution $w_s \in AC^1[0, 2\pi]$

$$w_s(t) = \int_0^{2\pi} g(t, \tau)\widetilde{h}_s(\tau)d\tau$$

satisfying for a.e. $t \in [0, 2\pi]$

$$w_s''(t) + a(t)w_s'(t) + b(t)w_s(t) = h_s^*(t) + \frac{1}{2\pi}a(t) - \Big(1 - \frac{1}{2\pi}t\Big)b(t),$$

$$w_s(0) = 0, \quad w_s(2\pi) = 0.$$

Denote

$$u_s(t) = \begin{cases} w_s(t) - \frac{1}{2\pi}t & \text{for } t \le s \\ w_s(t) + 1 - \frac{1}{2\pi}t & \text{for } t > s \end{cases}.$$

Then $u_s \in AC^1([0, 2\pi]\backslash\{s\})$ and

$$u_s''(t) + a(t)u_s'(t) + b(t)u_s(t) = w_s''(t) + a(t)\left[w_s'(t) - \frac{1}{2\pi}\right]$$

$$+ b(t)\left[w_s(t) - \tfrac{1}{2\pi}t\right] = h_s^*(t) - b(t) = 0$$

for a.e. $t \in (0, s)$ and

$$u_s''(t) + a(t)u_s'(t) + b(t)u_s(t) = w_s''(t) + a(t)\left[w_s'(t) - \frac{1}{2\pi}\right]$$

$$+ b(t)\left[w_s(t) + 1 - \tfrac{1}{2\pi}t\right] = h_s^*(t) = 0$$

for a.e. $t \in (s, 2\pi)$. Moreover

$$\Delta u_s(s) = w_s(s) + 1 - \frac{1}{2\pi}s - \left[w_s(s) - \frac{1}{2\pi}s\right] = 1,$$

$$\Delta u_s'(s) = w_s'(s) - \frac{1}{2\pi} - \left[w_s'(s) - \frac{1}{2\pi}\right] = 0,$$

$$u_s(0) = 0, \quad u_s(2\pi) = 0.$$

Hence we can define $\widetilde{g}(t, s) = u_s(t)$ for each fixed $s \in [0, 2\pi]$. □

Proof of Lemma 2.1 Now, we prove that u given by (2.6) is a solution of (2.1), (2.2), (2.5). For fixed $t_1 \in [0, 2\pi]$ we denote

$$\phi(t) = g(t, t_1), \quad \widetilde{\phi}(t) = \widetilde{g}(t, t_1),$$

$$u_1(t) = \widetilde{\phi}(t)d + \phi(t)e, \quad u_2(t) = \int_0^{2\pi} g(t, s)[h(s) - cb(s)]\, ds.$$

In view to properties of functions $g, \widetilde{g}$ we have $\phi, \widetilde{\phi} \in \widetilde{AC}^1[0, 2\pi]$ and

$$\Delta\phi(t_1) = 0, \quad \Delta\phi'(t_1) = 1,$$

$$\Delta\widetilde{\phi}(t_1) = 1, \quad \Delta\widetilde{\phi}'(t_1) = 0.$$

Then $u_1 \in \widetilde{AC}^1[0, 2\pi]$ is a solution of (2.3)–(2.5). Moreover $u_2 \in AC^1[0, 2\pi]$ is a solution of (2.4),

$$x'' + a(t)x' + b(t)x = h(t) - cb(t)$$

i.e. $u_2 + c$ is a solution of the problem (2.1), (2.2) without impulses. Thus $u = c + u_1 + u_2 \in \widetilde{AC}^1[0, 2\pi]$ is a solution of the impulsive problem (2.1), (2.2), (2.5). □

Lemma 2.3 *Let*

$$b(t) \le 0 \text{ for a.e. } t \in [0, 2\pi] \text{ and } \int_0^{2\pi} b(t)\, dt \ne 0. \tag{2.9}$$

Then the homogeneous boundary value problem (2.3), (2.4) has only the trivial solution.

Proof On the contrary, suppose that there exists a nontrivial solution u of (2.3), (2.4). Since $-u$ is a solution of (2.3), (2.4), as well, without loss of generality we can suppose that there exists a maximum point

$$\max_{t \in J} u(t) = u(t_M) > 0, \quad u'(t_M) = 0, \quad t_M \in (0, 2\pi).$$

Then, with respect to (2.4), there exists $t_0 \in (t_M, 2\pi)$ such that $u(t) > 0$ for all $t \in (t_M, t_0)$ and $u'(t_0) < 0$. On the contrary

$$u'(t_0) = -e^{-A(t_0)} \int_{t_M}^{t_0} e^{A(s)} b(s) u(s)\, ds \ge 0,$$

where $A(t) = \int_{t_M}^t a(s)\, ds$, a contradiction. □

Lemma 2.4 *Let (2.9) be fulfilled, let $\sigma_2(t)$ be a strict upper function of the problem (1.1)–(1.3) and let $\widetilde{f}(t,x,y)$ satisfy Carathéodory conditions on* $[0,2\pi]\times \mathrm{R}^2$ *and*

$$\widetilde{f}(t,x,y) > f(t,\sigma_2,y) \quad \textit{for a.e. } t\in[0,2\pi],\ x>\sigma_2,\ y\in \mathrm{R}. \tag{2.10}$$

Then

$$u(t)\le \sigma_2(t) \tag{2.11}$$

is valid for $t\in[0,2\pi]$ for every solution u of (1.2),

$$x''+a(t)x'+b(t)x=\widetilde{f}(t,x,x'), \tag{2.12}$$

which fulfils

$$u(0)=u(2\pi)\le\sigma_2(0). \tag{2.13}$$

Proof Denote $v(t)=u(t)-\sigma_2(t)$ for $t\in[0,2\pi]$.

(i) Let there exist $t_0\in(0,t_1)\cup(t_1,2\pi)$ such that $v(t_0)=\max\{v(t): t\in(0,t_1)\cup(t_1,2\pi)\}>0$, $v'(t_0)=0$. Then there exists $\delta>0$ such that $v(t)>0$, $|v'(t)|<\varepsilon$ for all $t\in(t_0,t_0+\delta)$, where ε is from (1.13) and so for a.e. $t\in(t_0,t_0+\delta)$

$$\begin{aligned} v''(t)&=u''(t)-\sigma_2''(t)=\widetilde{f}(t,u(t),u'(t))-a(t)u'(t)-b(t)u(t)-\sigma_2''(t)\\ &> f(t,\sigma_2(t),u'(t))-a(t)u'(t)-b(t)u(t)-\sigma_2''(t)\ge -b(t)(u(t)-\sigma_2(t))\ge 0. \end{aligned}$$

Hence, $v'(t)>0$ and $v(t)>v(t_0)$ for each $t\in(t_0,t_0+\delta)$, a contradiction.

(ii) Now, we suppose that $v(t_1)>v(t)$ for all $t\in(0,t_1)$ and $v(t_1)>0$. Then $u'(t_1)-\sigma_2'(t_1)=v'(t_1)\ge 0$ and $u(t_1)>\sigma_2(t_1)$. From the properties of J and M we get

$$v(t_1+)=J(u(t_1))-J(\sigma_2(t_1))>0,\quad v'(t_1+)=M(u'(t_1))-M(\sigma_2'(t_1))\ge 0.$$

Let $v'(t_1+)>0$. Then in view to (2.13) there is a maximum point $t_0\in(t_1,2\pi)$ and $v(t_0)>0$ which contradicts to (i). Then $v'(t_1+)=0$ and there exists $\beta\in(t_1,2\pi)$ such that $v'(\beta)<0,\ v(t)>0,\ |v'(t)|<\varepsilon$ for all $t\in(t_1,\beta)$, where ε is from (1.13) and then

$$v''(t)=u''(t)-\sigma_2''(t)=\widetilde{f}(t,u(t),u'(t))-a(t)u'(t)-b(t)u(t)-\sigma_2''(t)\ge 0,$$

for a.e. $t\in(t_1,\beta)$ and hence $v'(\beta)\ge 0$, a contradiction.

(iii) Suppose that $v(t_1+)>v(t)$ for all $t\in(t_1,2\pi]$ and $v(t_1+)>0$. Then $u'(t_1+)-\sigma_2'(t_1+)=v'(t_1+)\le 0$ and $u(t_1+)>\sigma_2(t_1+)$. If $v'(t_1+)=0$ then we get a contradiction as in (i). Hence $v'(t_1+)<0$. From the properties of functions J,M we get

$$v(t_1)=u(t_1)-\sigma_2(t_1)>0,\quad v'(t_1)=u'(t_1)-\sigma_2'(t_1)<0.$$

In view to (2.13) there exists a maximum point $t_0\in(0,t_1)$ such that $v(t_0)>0$, a contradiction with (i). □

Lemma 2.5 *Let (2.9) be fulfilled and σ_2 be a strict upper function of the problem (1.1)–(1.3). Then*

$$u(t) < \sigma_2(t) \ on \ [0, 2\pi] \tag{2.14}$$

is valid for every solution u of (1.1)–(1.3) which satisfies (2.11).

Proof Denote $v(t) = u(t) - \sigma_2(t)$ for $t \in [0, 2\pi]$.

(i) Let there exist $t_0 \in (0, t_1)$, such that $v(t_0) = \max\{v(t) : t \in (0, t_1)\} = 0$. Then $v'(t_0) = 0$ and there exist α, β such that $0 \le \alpha < t_0 < \beta \le t_1$ and $-\varepsilon < v(t) \le 0$, $|v'(t)| < \varepsilon$ for each $t \in (\alpha, \beta)$. From the property of the strict upper function, we get for a.e. $t \in (\alpha, \beta)$

$$v''(t) = u''(t) - \sigma_2''(t) = f(t, u(t), u'(t)) - a(t)u'(t) - b(t)u(t) - \sigma_2''(t) \ge 0$$

and so $v'(t) \ge 0,\ v(t) \ge 0$ for $t \in (t_0, \beta)$ and $v'(t) \le 0,\ v(t) \ge 0$ for $t \in (\alpha, t_0)$. With respect to (2.11) it is possible only if $v(t) = v'(t) = 0$ for $t \in (\alpha, \beta)$ where $\alpha = 0, \beta = t_1$. From (1.3), (1.11) we get $v(2\pi) = v(0) = 0$, $v'(2\pi) \le v'(0) = 0$ i.e. $v(2\pi) = v'(2\pi) = 0$ and hence, we obtain $v(t) = v'(t) = 0$ for $t \in (t_1, 2\pi]$, as well. Then $u(t) = \sigma_2(t)$ for $t \in [0, 2\pi]$, which contradicts to the definition of the strict upper function. In the case $t_0 \in (t_1, 2\pi)$, we can use the same arguments to get a contradiction.

(ii) Let $v(t) < v(t_1) = 0$ for $t \in [0, t_1)$. Then $v'(t_1) \ge 0$ and

$$v(t_1+) = J(u(t_1)) - J(\sigma(t_1)) = 0, \quad v'(t_1+) = M(u'(t_1)) - M(\sigma'(t_1)) \ge 0.$$

If $v'(t_1+) > 0$ then there exists $\gamma_1 \in (t_1, 2\pi)$ such that $v(\gamma_1) > 0$, a contradiction. Thus $v'(t_1+) = v(t_1+) = 0$ and so $v'(t) = v(t) = 0$ for $t \in (t_1, 2\pi]$. Using boundary value conditions we get $v(t) = 0$ for $t \in [0, t_1)$, as well. Then $u(t) = \sigma_2(t)$ for $t \in [0, 2\pi]$, a contradiction.

(iii) Now, let $v(t) < v(t_1+) = 0$ for $t \in (t_1, 2\pi]$. Then $v'(t_1+) \le 0$. Suppose $v'(t_1+) = 0$. Then there is $\beta \in (t_1, 2\pi]$ such that $0 > v(t) > -\varepsilon$ and $|v'(t)| < \varepsilon$ for $t \in (t_1, \beta)$ where $\varepsilon > 0$ is the constant from Definition 1.4. Thus, we get $v'(t) = 0$ for all $t \in (t_1, \beta)$ with $\beta = 2\pi$ and the same result we get on $[0, t_1)$, a contradiction. Then $v'(t_1+) < 0$ and from the properties of functions J and M we obtain $v(t_1) = 0$, $v'(t_1) < 0$. Hence there exists $\gamma_2 \in (0, t_1)$ such that $v(\gamma_2) > 0$, a contradiction.

(iv) Let $v(0) = v(2\pi) = 0$. From (1.2), (1.11) we get $v'(0) = v'(2\pi) = 0$. We get a contradiction as in (i). □

Lemma 2.6 *Let (2.9) be fulfilled, let $\sigma_1(t)$ be a strict lower function of the problem (1.1)–(1.3) and let $\widetilde{f}(t, x, y)$ satisfy Carathéodory conditions on $[0, 2\pi] \times \mathrm{R}^2$ and*

$$\widetilde{f}(t, x, y) < f(t, \sigma_1, y) \ for \ a.e. \ t \in [0, 2\pi],\ x < \sigma_1,\ y \in \mathrm{R}. \tag{2.15}$$

Then

$$u(t) \ge \sigma_1(t) \tag{2.16}$$

is valid for $t \in [0, 2\pi]$ for every solution u of (2.12), (1.2) which fulfils

$$u(0) = u(2\pi) \geq \sigma_1(0). \tag{2.17}$$

Proof We can use similar arguments as in the proof of Lemma 2.4. □

Lemma 2.7 *Let (2.9) be fulfilled, let $\sigma_1(t)$ be a strict lower function of the problem (1.1)–(1.3). Then*

$$u(t) > \sigma_1(t) \text{ on } [0, 2\pi] \tag{2.18}$$

is valid for every solution u of (1.1)–(1.3) which satisfies (2.16).

Proof We can use similar arguments as in the proof of Lemma 2.5. □

We can rewrite the periodic conditions (1.3) to the equivalent form of Dirichlet type boundary conditions

$$x(0) = x(0) + x'(0) - x'(2\pi),\ x(2\pi) = x(0) + x'(0) - x'(2\pi). \tag{2.19}$$

In view to Lemma 2.1 and (2.19), we can rewrite problem (1.1)–(1.3) to the operator form

$$\begin{aligned}(Fx)(t) &= x(0) + x'(0) - x'(2\pi) \\ &+ \int_0^{2\pi} g(t,s)\big[f(s, x(s), x'(s)) - \big(x(0) + x'(0) - x'(2\pi)\big)b(s)\big]\,ds \\ &+ \widetilde{g}(t, t_1)[J(x(t_1)) - x(t_1)] + g(t, t_1)[M(x'(t_1)) - x'(t_1)], \quad t \in [0, 2\pi].\end{aligned} \tag{2.20}$$

The operator $F : \widetilde{C}^1[0, 2\pi] \to \widetilde{C}^1[0, 2\pi]$ is completely continuous (see [2], Lemma 3.1) and every fixed point $u \in \widetilde{C}^1[0, 2\pi]$ of F is a solution of (1.1)–(1.3).

Now, assume that problem (2.3), (2.4) has a nontrivial solution. Then we choose an arbitrary $\mu \in (-\infty, 0)$ and instead of (1.1) we will use the equation

$$x'' + a(t)x' + \mu x = f_\mu(t, x, x'), \tag{2.21}$$

where

$$f_\mu(t, x, x') = f(t, x, x') + (\mu - b(t))x. \tag{2.22}$$

Then in view to Lemma 2.3 the corresponding homogeneous problem

$$x'' + a(t)x' + \mu x = 0, \tag{2.23}$$

(2.4) has only the trivial solution and hence we can rewrite problem (2.21), (1.2), (1.3) to the operator form

$$\begin{aligned}(F_\mu x)(t) &= x(0) + x'(0) - x'(2\pi) \\ &+ \int_0^{2\pi} g_\mu(t,s)\big[f_\mu(s, x(s), x'(s)) - \big(x(0) + x'(0) - x'(2\pi)\big)\mu\big]\,ds \\ &+ \widetilde{g}_\mu(t, t_1)[J(x(t_1)) - x(t_1)] + g_\mu(t, t_1)[M(x'(t_1)) - x'(t_1)], \quad t \in [0, 2\pi],\end{aligned} \tag{2.24}$$

where g_μ is the Green function of (2.23), (2.2) and $\widetilde{g}_\mu$ is a function which fulfils (2.23) for a.e. $t \in [0,s) \cup (s,2\pi]$ and each fixed $s \in [0,2\pi]$ satisfies (2.4) and

$$(2.25) \qquad \widetilde{g}_\mu(s+,s) = \widetilde{g}_\mu(s,s) + 1, \quad \frac{\partial \widetilde{g}_\mu(t,s)}{\partial t}\bigg|_{t=s+} = \frac{\partial \widetilde{g}_\mu(t,s)}{\partial t}\bigg|_{t=s-}.$$

The operator $F_\mu : \widetilde{C}^1[0,2\pi] \to \widetilde{C}^1[0,2\pi]$ is completely continuous and every its fixed point $u \in \widetilde{C}^1[0,2\pi]$ is a solution of (1.1)–(1.3).

3 Strict lower and upper functions and topological degree

Lemma 3.1 *Suppose that $r_0 \in (0,\infty)$, $p \in L[0,2\pi]$, $q \in L_\infty[0,2\pi]$, p,q are positive a.e. on $[0,2\pi]$. Then there exists $r^* \in (0,\infty)$ such that for each $x \in \widetilde{AC}^1[0,2\pi]$ fulfilling (1.2), (1.3),*

$$(3.1) \qquad \|x\|_\infty < r_0$$

and

$$(3.2) \qquad |x'' + a(t)x' + b(t)x(t)| \le (1+|x'|)(p(t)+q(t)|x'|)$$

for a.e. $t \in [0,2\pi]$, the estimate

$$(3.3) \qquad \|x'\|_\infty < r^*$$

is valid.

Proof Let (3.1), (3.2) be valid. In view to the mean value theorem there exist $\tau_1 \in [0,t_1), \tau_2 \in (t_1,2\pi]$ such that

$$|x'(\tau_1)| \le \frac{2r_0}{t_1}, \quad |x'(\tau_2)| \le \frac{2r_0}{2\pi - t_1}.$$

Denote

$$(3.4) \qquad \begin{aligned} A(t) &= \exp[\textstyle\int_0^t a(\tau)\,d\tau\,], \\ y(t) &= A(t)x'(t), \quad \widetilde{A} = \|A\|_\infty, \\ N &= \|p\|_1 + 2\|q\|_\infty r_0 + \|b\|_1 r_0, \\ \widetilde{r} &> \max\Big\{A(\tau_1)\tfrac{2r_0}{t_1}, A(\tau_2)\tfrac{2r_0}{2\pi - t_1}\Big\}. \end{aligned}$$

(i) At first, suppose $x'(t_{sup}) = \sup\{x'(t) : t \in [0,2\pi]\} > 0$.

Assume $0 \le t_{sup} \le t_1$. Then there are α, β such that $0 \le \alpha < \beta \le t_1$ and $t_{sup} \in [\alpha,\beta]$ and such that $x'(t) > 0$ for each $t \in [\alpha,\beta]$. From (3.2) we get for a.e. $t \in [\alpha,\beta]$

$$|x''(t) + a(t)x'(t)| \le (1+x'(t))[p(t)+q(t)x'(t)] + |b(t)|r_0,$$

$$|[A(t)x'(t)]'| \le A(t)(1+x'(t))(p(t)+q(t)x'(t)) + A(t)|b(t)|r_0,$$

$$\left|\frac{y'(t)}{\widetilde{A}+y(t)}\right| \le \frac{|[A(t)x'(t)]'|}{A(t)(1+x'(t))} \le p(t)+q(t)x'(t)+|b(t)|r_0,$$

$$-p(t)-q(t)x'(t)-|b(t)|r_0 \le \frac{y'(t)}{\widetilde{A}+y(t)} \le p(t)+q(t)x'(t)+|b(t)|r_0. \tag{3.5}$$

Let $\tau_1 \le t_{sup}$. Then we can choose τ_1 such that $x'(\tau_1) \ge 0$ and $x'(t) > 0$ on (τ_1, t_{sup}). Then by integrating of the right hand inequality of (3.5) on (τ_1, t_{sup}), we get

$$\ln\left(\frac{\widetilde{A}+y(t_{sup})}{\widetilde{A}+y(\tau_1)}\right) \le \|p\|_1 + 2\|q\|_\infty r_0 + \|b\|_1 r_0 = N,$$

$$x'(t_{sup}) \le \frac{1}{A(t_{sup})}[(\widetilde{A}+\widetilde{r})e^N - \widetilde{A}]. \tag{3.6}$$

Let $\tau_1 \ge t_{sup}$. Then we can choose τ_1 such that $x'(\tau_1) \ge 0$ and $x'(t) > 0$ on (t_{sup}, τ_1). Then we get (3.6) by integrating of the left hand inequality of (3.5). Similarly we get (3.6) with τ_2 instead of τ_1 for $t_1 < t_{sup} \le 2\pi$.

Assume $x'(t_1+) > x'(t)$ for each $t \in (t_1, 2\pi]$. Then there exists $\beta \in (t_1, 2\pi)$ such that $x'(t) > 0$ on (t_1, β). Thus (3.5) is valid for each $t \in (t_1, \beta)$. We can choose τ_2 such that $x'(\tau_2) \ge 0$ and $x'(t) > 0$ on (t_1, τ_2). By integrating of the left hand inequality of (3.5) on (t_1, τ_2) we get

$$x'(t_1+) \le \frac{1}{A(t_1)}[(\widetilde{A}+\widetilde{r})e^N - \widetilde{A}].$$

(ii) Now, suppose $x'(t_{inf}) = \inf\{x'(t) : t \in [0, 2\pi]\} < 0$.

Assume $0 \le t_{inf} \le t_1$. Then there are α, β such that $0 \le \alpha < \beta \le t_1$ and $t_{inf} \in [\alpha, \beta]$ and such that $x'(t) < 0$ for each $t \in [\alpha, \beta]$. From (3.2) we get for a.e. $t \in [\alpha, \beta]$

$$|x''(t) + a(t)x'(t)| \le (1-x'(t))[p(t) - q(t)x'(t)] + |b(t)|r_0,$$

$$|[A(t)x'(t)]'| \le A(t)(1-x'(t))(p(t)-q(t)x'(t)) + A(t)|b(t)|r_0,$$

$$\left|\frac{y'(t)}{\widetilde{A}-y(t)}\right| \le \frac{|[A(t)x'(t)]'|}{A(t)(1-x'(t))} \le p(t)-q(t)x'(t)+|b(t)|r_0,$$

$$-p(t)+q(t)x'(t)-|b(t)|r_0 \le \frac{y'(t)}{\widetilde{A}-y(t)} \le p(t)-q(t)x'(t)+|b(t)|r_0. \tag{3.7}$$

Let $\tau_1 \le t_{inf}$. Then we can choose τ_1 such that $x'(\tau_1) \le 0$ and $x'(t) < 0$ on (τ_1, t_{inf}). By integrating of the right hand inequality of (3.7) on (τ_1, t_{inf}), we get

$$-\ln\left(\frac{\widetilde{A} - y(\tau_1)}{\widetilde{A} - y(t_{inf})}\right) \le \|p\|_1 + 2\|q\|_\infty r_0 + \|b\|_1 r_0 = N,$$

$$x'(t_{inf}) \ge -\frac{1}{A(t_{inf})}[(\widetilde{A} + \widetilde{r})e^N - \widetilde{A}]. \tag{3.8}$$

Let $\tau_1 \ge t_{inf}$.Then we can choose τ_1 such that $x'(\tau_1) \le 0$ and $x'(t) < 0$ on (t_{inf}, τ_1). By integrating of the left hand inequality of (3.7) on (t_{inf}, τ_1) we get (3.8), as well. Similarly we get (3.8) with τ_2 instead of τ_1 for $t_1 < t_{inf} \le 2\pi$.

Assume $x'(t_1+) < x'(t)$ for each $t \in (t_1, 2\pi]$. Then there exists $\beta \in (t_1, 2\pi)$ such that $x'(t) < 0$ on (t_1, β). Thus (3.7) is valid for each $t \in (t_1, \beta)$. We can choose τ_2 such that $x'(\tau_2) \ge 0$ and $x'(t) > 0$ on (t_1, τ_2). By integrating of the left right inequality of (3.7) on (t_1, τ_2) we get

$$x'(t_1+) \ge -\frac{1}{A(t_1)}[(\widetilde{A} + \widetilde{r})e^N - \widetilde{A}].$$

Hence for

$$r^* > \frac{1}{\min\{A(t) : t \in [0, 2\pi]\}}[(\widetilde{A} + \widetilde{r})e^N - \widetilde{A}]$$

the inequality (3.3) is valid. □

Theorem 3.2 *Let $\sigma_1, \sigma_2 \in \widetilde{AC}^1[0, 2\pi]$ be strict lower and upper functions of (1.1)–(1.3) such that*

$$\sigma_1(t) < \sigma_2(t) \;\; for \;\; t \in [0, 2\pi] \tag{3.9}$$

and let there exist functions $p, q \in L[0, 2\pi]$ positive a.e. on $[0, 2\pi]$ such that

$$|f(t, x, y)| \le (1 + |y|)(p(t) + q(t)|y|) \tag{3.10}$$

for a.e. $t \in [0, 2\pi]$ and all $(x, y) \in [\sigma_1(t), \sigma_2(t)] \times \mathrm{R}$. Then

$$d[I - F_\mu, \Omega] = 1$$

for any $\mu \in (-\infty, 0)$ and F_μ defined by (2.24), where

$$\begin{aligned} \Omega = \{x \in \widetilde{C}^1[0, 2\pi] : \sigma_1(t) < x(t) < \sigma_2(t) \;\; on \;\; [0, 2\pi], \\ \sigma_1(t_1+) < x(t_1+) < \sigma_2(t_1+), \;\; \|x'\|_\infty < C\}, \end{aligned} \tag{3.11}$$

$$C > \left[1 + \left(\|\sigma_1\|_{\widetilde{C}_1} + \|\sigma_2\|_{\widetilde{C}_1}\right) \max\left\{\frac{1}{t_1}; \frac{1}{2\pi - t_1}\right\}\right] e^{\|p\|_1 + 2\|q\|_\infty r_0 + \|b\|_1 r_0},$$

$$r_0 = \max\{\|\sigma_1\|_\infty, \|\sigma_2\|_\infty\}.$$

Proof Let us choose a constant C satisfying (3.11) and define auxiliary functions

$$\alpha(t,x)=\begin{cases}\sigma_1(t) & \text{for } x<\sigma_1(t)\\ x & \text{for } \sigma_1(t)\le x\le\sigma_2(t)\\ \sigma_2(t) & \text{for } \sigma_2(t)<x\end{cases}, \tag{3.12}$$

$$\beta(y)=\begin{cases}-C & \text{for } \ y<-C\\ \ y & \text{for } -C\le y\le C\\ \ C & \text{for } C<y\end{cases}, \tag{3.13}$$

$$\overline{f}_\mu(t,x,y)=f_\mu(t,x,\beta(y))=f(t,x,\beta(y))+(\mu-b(t))x, \tag{3.14}$$

$$\widetilde{f}_\mu(t,x,y)=\begin{cases}\overline{f}_\mu(t,\sigma_1(t),y)-\frac{\sigma_1(t)-x}{1+\sigma_1(t)-x} & \text{for } x<\sigma_1(t)\\ \overline{f}_\mu(t,x,y) & \text{for } \sigma_1(t)\le x\le\sigma_2(t)\\ \overline{f}_\mu(t,\sigma_2(t),y)+\frac{x-\sigma_2(t)}{1+x-\sigma_2(t)} & \text{for } \sigma_2(t)<x\end{cases}, \tag{3.15}$$

and an operator

$$\begin{aligned}(\widetilde{F}_\mu x)(t)&=\alpha(0,x(0)+x'(0)-x'(2\pi))+\widetilde{g}(t,t_1)[J(\alpha(t_1,x(t_1)))\\ &\quad-\alpha(t_1,x(t_1))]+g(t,t_1)[M(\beta(x'(t_1-)))-\beta(x'(t_1-))]\\ &\quad+\int_0^{2\pi} g(t,s)[\widetilde{f}_\mu(s,x(s),x'(s))-\alpha(0,x(0)+x'(0)-x'(2\pi))b(s)]\,ds\end{aligned} \tag{2.24}$$

for $t\in[0,2\pi]$ and $\mu\in(-\infty,0)$. We can see that $\widetilde{f}_\mu$ fulfills the Carathéodory conditions on $[0,2\pi]\times \mathrm{R}^2$ and $\alpha:\mathrm{R}^2\to\mathrm{R}$, β: $\mathrm{R}\to\mathrm{R}$ are continuous mappings. Therefore $\widetilde{F}_\mu:\widetilde{C}^1[0,2\pi]\to\widetilde{C}^1[0,2\pi]$ is completely continuous. Consider the parameter system of operator equations

$$x-\lambda\widetilde{F}_\mu x=0,\quad \lambda\in[0,1]. \tag{3.17}$$

With respect to (3.12)–(3.16), we can show that there is $r\in(0,\infty)$ that

$$\|\widetilde{F}_\mu x\|_{\widetilde{C}^1}\le r \text{ for } x\in\widetilde{C}^1[0,2\pi]. \tag{3.18}$$

Hence, there is $\rho>0$ such that for any $\lambda\in[0,1]$ every solution of (3.17) lies inside the set

$$K(\rho)=\{x\in\widetilde{C}^1[0,2\pi];\|x\|_{\widetilde{C}^1}<\rho\}.$$

Then $G=I-\lambda\widetilde{F}_\mu$ is a homotopic map on $\overline{K(\rho)}\times[0,1]$ and

$$d[I-\widetilde{F}_\mu,K(\rho)]=d[I\ ,K(\rho)]=1.$$

Therefore there exists a solution u of (3.17) with $\lambda = 1$. In view to (3.16), u is a solution of

$$x'' + a(t)x' + \mu x = \widetilde{f}_\mu(t, x, x'), \tag{3.19}$$

$$x(t_1+) = \widetilde{J}(x(t_1)),\ x'(t_1+) = \widetilde{M}(x'(t_1)), \tag{3.20}$$

$$x(0) = x(2\pi) = \alpha(0, x(0) + x'(0) - x'(2\pi)), \tag{3.21}$$

where

$$\begin{aligned} \widetilde{J}(x) &= x + J(\alpha(t, x)) - \alpha(t, x) \text{ for } x \in \mathrm{R}, \\ \widetilde{M}(y) &= y + M(\beta(y)) - \beta(y) \text{ for } y \in \mathrm{R}. \end{aligned} \tag{3.22}$$

With respect to (3.12) we have

$$\sigma_1(0) \le \alpha(0, u(0) + u'(0) - u'(2\pi)) \le \sigma_2(0),$$

$$\widetilde{f}_\mu(t, x, y) > \overline{f}_\mu(t, \sigma_2, y) \text{ for a.e. } t \in [0, 2\pi],\ \text{each } x \in (\sigma_2, \infty), y \in \mathrm{R},$$

$$\widetilde{f}_\mu(t, x, y) < \overline{f}_\mu(t, \sigma_1, y) \text{ for a.e. } t \in [0, 2\pi],\ \text{each } x \in (-\infty, \sigma_1), y \in \mathrm{R}.$$

In view to (3.22), $\widetilde{J}$ and $\widetilde{M}$ satisfy conditions (1.4). Let $\varepsilon > 0$ be from Definition 1.4. Since $\|\sigma_1\|_{\widetilde{C}^1} + \|\sigma_2\|_{\widetilde{C}^1} < C$, then there exists $\varepsilon_1 < \varepsilon$ such that $\|\sigma_1\|_\infty + \varepsilon_1 < C$ and $\|\sigma_2\|_\infty + \varepsilon_1 < C$. Then $\overline{f}_\mu(t, x, y) = f_\mu(t, x, y)$ for $(x, y) \in [\sigma_1(t), \sigma_1(t) + \varepsilon_1] \times [\sigma_1'(t) - \varepsilon_1, \sigma_1'(t) + \varepsilon_1]$ i.e. σ_1 fulfils condition (1.12), (1.7), (1.8) with $\overline{f}_\mu(t, x, y)$ instead of $f(t, x, y)$. Hence σ_1 is a strict lower function of

$$x'' + a(t)x' + \mu x = \overline{f}_\mu(t, x, x'), \tag{3.23}$$

(1.2), (1.3). Similarly $\overline{f}_\mu(t, x, y) = f_\mu(t, x, y)$ for $(x, y) \in [\sigma_2(t) - \varepsilon_1, \sigma_2(t)] \times [\sigma_2'(t) - \varepsilon_1, \sigma_2'(t) + \varepsilon_1]$ i.e. σ_2 fulfils conditions (1.13), (1.10), (1.11) with $\overline{f}_\mu(t, x, y)$ instead of $f(t, x, y)$. Hence σ_2 is a strict upper function of (3.23), (1.2), (1.3). In view to (3.12) and (3.21) we have $\sigma_1(0) \le u(0) = u(2\pi) \le \sigma_2(0)$. Then using lemmas 2.4-2.7 with $\overline{f}_\mu(t, x, y)$ and $\widetilde{f}_\mu(t, x, y)$ instead of $f(t, x, y)$ and $\widetilde{f}(t, x, y)$ we get

$$\sigma_1(t) < u(t) < \sigma_2(t) \text{ on } [0, 2\pi]. \tag{3.24}$$

Furthermore, $\widetilde{f}(t, x, y)$ fulfills (3.10) and thus from Lemma 3.1 we get

$$\|u'\|_\infty < C.$$

Moreover, in view to (3.15) for a.e. $t \in [0, 2\pi]$ we have

$$\widetilde{f}_\mu(t, u(t), u'(t)) = f_\mu(t, u(t), u'(t)).$$

Then we get that u is a solution of the equation (2.21) and satisfies condition (1.2) and $u(0) = u(2\pi)$. Now, we need to prove the second condition in (1.3) i.e. we prove that

$$u'(0) = u'(2\pi).$$

Especially, we prove

$$\sigma_1(0) \le u(0) + u'(0) - u'(2\pi) \le \sigma_2(0). \tag{3.25}$$

On the contrary, suppose that

$$u(0) + u'(0) - u'(2\pi) > \sigma_2(0). \tag{3.26}$$

Then from (3.21) we get

$$u(0) = u(2\pi) = \alpha(0, u(0) + u'(0) - u'(2\pi)) = \sigma_2(0) = \sigma_2(2\pi) \tag{3.27}$$

and using (3.26)

$$u'(0) > u'(2\pi). \tag{3.28}$$

On the other side we proved

$$u(t) \le \sigma_2(t) \quad t \in [0, 2\pi]$$

and with (3.27) and (3.28) this yields

$$\sigma_2'(0) \ge u'(0) > u'(2\pi) \ge \sigma_2'(2\pi)$$

which contradicts to (1.11). Similarly we will prove that

$$\sigma_1(0) \le u(0) + u'(0) - u'(2\pi).$$

With respect to (3.21) and (3.12) we have $u'(0) = u'(2\pi)$. Thus, we have proved that every solution of (3.17) with $\lambda = 1$ is a solution of (2.21), (1.2), (1.3) which satisfies (3.24). Hence $u \in \Omega$. Since $F_\mu = \widetilde{F}_\mu$ on $\overline{\Omega}$ and $x \ne F_\mu x$ for $x \in \partial\Omega$, we use the excision property of the topological degree and get

$$d(I - F_\mu, \Omega) = d(I - \widetilde{F}_\mu, \Omega) = d(I - \widetilde{F}_\mu, K(\rho)) = 1. \qquad \square$$

Corollary 3.3 *Let the assumptions of Theorem 3.2 be satisfied. Then the problem (1.1)–(1.3) has a solution u, which fulfills (3.24).*

Lemma 3.4 *Let σ_1, σ_2 be strict lower and upper functions such that*

$$\sigma_2(t) < \sigma_1(t) \text{ for each } t \in [0, 2\pi]. \tag{3.29}$$

Moreover, let $p, q \in L[0, 2\pi]$ be positive a.e. on $[0, 2\pi]$ such that for a.e. $t \in [0, 2\pi]$ and all $x, y \in R$

$$|f(t, x, y) - b(t)x| < p(t) + q(t)|y|. \tag{3.30}$$

Then for every solution $u \in \widetilde{AC}^1[0,2\pi]$ *of (1.1)–(1.3) which fulfills*

$$\sigma_2(t_u) < u(t_u) < \sigma_1(t_u) \quad \textit{for some } t_u \in [0,2\pi] \tag{3.31}$$

the estimate

$$\|u'\|_{\widetilde{C}} < N_1, \qquad \|u\|_{\widetilde{C}} < N_2, \tag{3.32}$$

where

$$N_1 = (2 + \|\sigma_1'\|_\infty + \|\sigma_2'\|_\infty) e^{\|p\|_1 + \|q\|_1 + \|a\|_1}, \quad N_2 = \|\sigma_1\|_\infty + \|\sigma_2\|_\infty + 2\pi N_1,$$

is valid.

Proof At first, we prove the existence of such $r \in (0,\infty)$ that the condition

$$|u'(s_u)| < r \tag{3.33}$$

is valid for some $s_u \in [0,2\pi]$. Denote $v_i(t) = (-1)^i(\sigma_i(t) - u(t))$, $i = 1,2$

i) Let

$$v_i'(s_u) = 0 \tag{3.34}$$

for some $s_u \in (0,t_1) \cup (t_1,2\pi)$. Then

$$|u'(s_u)| = |\sigma_i'(s_u)| \le \|\sigma_i'\|_\infty.$$

ii) Assume that (3.34) is not valid. Then from (1.3),(1.8) and (1.11) we have $v_i'(t) < 0$ for $t \in (0,t_1)$ and $v_i'(t) > 0$ for $t \in (t_1,2\pi)$ i.e.

$$v_i'(t_1) \le 0 \text{ and } v_i'(t_1+) \ge 0.$$

On the other hand,

$$v_i'(t_1+) \le (-1)^i [M(\sigma_i'(t_1)) - M(u'(t_1))] \le 0$$

and hence $v_i'(t_1+) = 0$. Then $|u'(t_1+)| = |\sigma_i'(t_1+)|$ and there exists $s_u \in (t_1,2\pi)$, that

$$|u'(s_u)| \le \|\sigma_i'\|_\infty + 1.$$

The condition (3.33) is proved for $r = \|\sigma_1'\|_\infty + \|\sigma_2'\|_\infty + 1$. Now, suppose that (3.30) is valid. Then for a.e. $t \in [0,2\pi]$ we get

$$|u''(t) + a(t)u'(t)| = |f(t,u(t),u'(t)) - b(t)u(t)| \le p(t) + q(t)|u'(t)|,$$

$$|u''(t)| \le p(t) + (q(t) + |a(t)|)|u'(t)|,$$

$$-p(t) - q(t) - |a(t)| \le \frac{u''(t)u'(t)}{1 + u'^2(t)} \le p(t) + q(t) + |a(t)|. \tag{3.35}$$

We integrate the left inequality of (3.35) on (t, s_u) for $t \in (t_1, s_u)$ and the right inequality of (3.35) on (s_u, t) for $t \in (s_u, 2\pi]$ and using (1.3) we can extend that for $t \in [0, t_1)$. Thus we have for each $t \in [0, 2\pi]$

$$1 + u'^2(t) \le (1 + u'^2(s_u))e^{2(\|p\|_1+\|q\|_1+\|a\|_1)} \le (1 + |u'(s_u)|)^2 e^{2(\|p\|_1+\|q\|_1+\|a\|_1)},$$

$$|u'(t)| \le (1 + |u'(s_u)|)e^{\|p\|_1+\|q\|_1+\|a\|_1} < N_1.$$

Moreover for each $t \in [0, t_1) \cup (t_1, 2\pi]$

$$|u(t)| \le |u(t_u)| + \left| \int_{t_u}^{t} u'(\tau)d\tau \right| < \|\sigma_1\|_\infty + \|\sigma_2\|_\infty + 2\pi N_1 = N_2,$$

is valid and then we get (3.32). □

Remark 3.5 Let $p, q \in L[0, 2\pi]$ be positive a.e. on $[0, 2\pi]$ such that for a.e. $t \in [0, 2\pi]$ and all $x, y \in R$ (3.30) is satisfied. Then

$$|f_\mu(t, x, y) - \mu x| < p(t) + q(t)|y| \tag{3.36}$$

is valid for a.e. $t \in [0, 2\pi]$ and each $x, y \in R$ where $f_\mu(t, x, y)$ is given by (2.22).

Theorem 3.6 *Let σ_1 and σ_2 be respectively strict lower and upper functions of (1.1)–(1.3) which fulfill (3.29), let $M(0) = 0$ and let there exist $p, q \in L[0, 2\pi]$ positive a.e. on $[0, 2\pi]$ such that (3.30) is satisfied. Then for any $\mu \in (-\infty, 0)$*

$$d[I - F_\mu, \Omega_2] = -1, \tag{3.37}$$

where F_μ is defined by (2.24),

$$\begin{aligned} \Omega_2 = \{x \in \widetilde{C}^1[0, 2\pi];\ \|x\|_\infty < \widetilde{N}_2,\ \|x'\|_\infty < \widetilde{N}_1, \\ \sigma_2(t_x) < x(t_x) < \sigma_1(t_x) \text{ for some } t_x \in [0, 2\pi]\}, \end{aligned} \tag{3.38}$$

$$\widetilde{N}_1 = (1 + \|\sigma_1'\|_\infty + \|\sigma_2'\|_\infty)e^{2(\|a\|_1+\|q\|_1+3\|p\|_1)}$$

and

$$\widetilde{N}_2 = \|\sigma_1\|_\infty + \|\sigma_2\|_\infty + 2\pi\widetilde{N}_1.$$

Proof Let $\varrho > \widetilde{N}_2$. Denote

$$\widetilde{f}(t, x, y) = \begin{cases} f(t, x, y) + p(t) & \text{for } x \ge \varrho \\ f(t, x, y) + \frac{x-\widetilde{N}_2}{\varrho-\widetilde{N}_2}p(t) & \text{for } \widetilde{N}_2 < x < \varrho \\ f(t, x, y) & \text{for } -\widetilde{N}_2 \le x \le \widetilde{N}_2 \\ f(t, x, y) - \frac{x+\widetilde{N}_2}{-\varrho+\widetilde{N}_2}p(t) & \text{for } -\varrho < x < -\widetilde{N}_2 \\ f(t, x, y) - p(t) & \text{for } x \le -\varrho \end{cases}. \tag{3.39}$$

Then for a.e. $t \in [0, 2\pi]$ and each $x, y \in R$

$$|\widetilde{f}(t, x, y) - b(t)x| \le |f(t, x, y) - b(t)x| + p(t) \le 2p(t) + q(t)|y|.$$

In view to (3.30)

$$-p(t) < f(t,x,0) - b(t)x < p(t)$$

is valid for a.e. $t \in [0, 2\pi]$ and each $x \in \mathrm{R}$ and hence

$$\widetilde{f}(t,\varrho,0) - b(t)\varrho = f(t,\varrho,0) - b(t)\varrho + p(t) > 0,$$

$$\widetilde{f}(t,-\varrho,0) + b(t)\varrho = f(t,-\varrho,0) + b(t)\varrho - p(t) < 0.$$

Consider a problem (1.3),

$$x'' = \overline{f}(t,x,x'), \tag{3.40}$$

$$x(t_1+) = \widetilde{J}(x(t_1)), \quad x'(t_1+) = M(x'(t_1)), \tag{3.41}$$

where

$$\overline{f}(t,x,y) = \widetilde{f}(t,x,y) - a(t)y - b(t)x, \tag{3.42}$$

$$\widetilde{J}(x) = \begin{cases} x & \text{for } x \le -\varrho, \\ -\frac{x+\widetilde{N}_2}{-\varrho+\widetilde{N}_2}\varrho + \big[1 - \frac{x+\widetilde{N}_2}{-\varrho+\widetilde{N}_2}\big]J(-\widetilde{N}_2) & \text{for } -\varrho < x < -\widetilde{N}_2, \\ J(x) & \text{for } -\widetilde{N}_2 \le x \le \widetilde{N}_2, \\ \frac{x-\widetilde{N}_2}{\varrho-\widetilde{N}_2}\varrho + \big[1 - \frac{x-\widetilde{N}_2}{\varrho-\widetilde{N}_2}\big]J(\widetilde{N}_2) & \text{for } \widetilde{N}_2 < x < \varrho, \\ x & \text{for } x \ge \varrho. \end{cases} \tag{3.43}$$

We can see that $\widetilde{J}$ is a continuous and increasing on R and

$$\widetilde{J}(\varrho) = \varrho, \quad \widetilde{J}(-\varrho) = -\varrho.$$

Moreover σ_1, σ_2 are strict lower and strict upper functions of (1.3), (3.40), (3.41). For a.e. $t \in J$ and each $x, y \in \mathrm{R}$ define a function

$$\overline{h}(t,x,y) = \begin{cases} \overline{f}(t,-\varrho,y) - \omega_1(t, \frac{-\varrho-x}{1-\varrho-x}) & \text{for } x < -\varrho \\ \overline{f}(t,x,y) & \text{for } -\varrho < x < \varrho\ , \\ \overline{f}(t,\varrho,y) + \omega_2(t, \frac{x-\varrho}{1+x-\varrho}) & \text{for } x > \varrho \end{cases} \tag{3.44}$$

where

$$\omega_i(t,\varepsilon) = \sup_{y \in [-\varepsilon,\varepsilon]} \{|\overline{f}(t,(-1)^i\varrho,y) - \overline{f}(t,(-1)^i\varrho,0)|\}, \quad i = 1,2 \tag{3.45}$$

for $\varepsilon > 0$. ω_i is positive and nondecreasing with the second variable and with respect to (3.42) and (3.30) we have

$$\omega_i(t,\varepsilon) = \sup_{y \in [-\varepsilon,\varepsilon]} \{|\widetilde{f}(t,(-1)^i\varrho,y) - \widetilde{f}(t,(-1)^i\varrho,0) - a(t)y|\},$$

$$\omega_i(t,\varepsilon) \le 4p(t) + (q(t) + |a(t)|)|y| \tag{3.46}$$

for a.e. $t \in [0, 2\pi]$ and each $y \in [-\varepsilon, \varepsilon]$. Now, consider the problem (1.3), (3.41)

$$x'' = \overline{h}(t, x, x'). \tag{3.47}$$

Choose $\eta > 0$ and put $\sigma_3(t) = -\varrho - \eta$, $\sigma_4(t) = \varrho + \eta$ for $t \in [0, 2\pi]$. Then for a.e. $t \in [0, 2\pi]$

$$\overline{h}(t, \varrho + \eta, 0) = \overline{f}(t, \varrho, 0) + \omega_2\Big(t, \frac{\eta}{1+\eta}\Big) = f(t, \varrho, 0) - b(t)\varrho + \omega_2\Big(t, \frac{\eta}{1+\eta}\Big) > 0.$$

For $\varepsilon = \frac{\eta/2}{1+\eta/2}$ and for $x \in [\varrho+\eta-\varepsilon, \varrho+\eta]$, $y \in [-\varepsilon, \varepsilon]$ we obtain $x \in (\varrho+\eta/2, \varrho+\eta]$ and $|y| < \frac{x-\varrho}{1+x-\varrho}$ i.e. $\omega_2(t, |y|) \le \omega_2(t, \frac{x-\varrho}{1+x-\varrho})$.

Hence, in view of (3.44), we get

$$\begin{aligned} \overline{h}(t, x, y) &= \overline{f}(t, \varrho, y) + \omega_2\Big(t, \frac{x-\varrho}{1+x-\varrho}\Big) \\ &\ge \overline{f}(t, \varrho, 0) - |\overline{f}(t, \varrho, y) - \overline{f}(t, \varrho, 0)| + \omega_2\Big(t, \frac{x-\varrho}{1+x-\varrho}\Big) > 0. \end{aligned}$$

Thus σ_4 is a strict upper function of (3.47), (3.41), (1.3). Similarly we can prove, that σ_3 is a strict lower function of (3.47), (3.41), (1.3). Now, we choose an arbitrary $\mu \in (-\infty, 0)$ and rewrite the equation to the form

$$x'' + a(t)x' + \mu x = h_\mu(t, x, x'), \tag{3.48}$$

$$h_\mu(t, x, y) = h(t, x, y) + (\mu - b(t))x, \tag{3.49}$$

$$h(t, x, y) = \overline{h}(t, x, y) + a(t)y + b(t)x. \tag{3.50}$$

Then σ_1, σ_3 are strict lower and σ_2, σ_4 strict upper functions of (3.48), (3.41), (1.2) such that

$$\sigma_3(t) < \sigma_2(t) < \sigma_1(t) < \sigma_4(t) \text{ for all } t \in [0, 2\pi]. \tag{3.51}$$

Denote

$$\begin{aligned} \widetilde{\Omega} &= \{x \in \widetilde{C}^1[0, 2\pi] : \|x\|_\infty < \varrho + \eta, \|x'\|_\infty < \widetilde{N}_1\}, \\ \Delta_1 &= \{x \in \widetilde{\Omega} : x(t) > \sigma_1(t) \text{ for } t \in [0, 2\pi]\}, \\ \Delta_2 &= \{x \in \widetilde{\Omega} : x(t) < \sigma_2(t) \text{ for } t \in [0, 2\pi]\}. \end{aligned}$$

In view to (3.44)–(3.48) there exist functions $\widetilde{p}, \widetilde{q} \in L[0, 2\pi]$ positive a.e. on $[0, 2\pi]$ such that for a.e. $t \in [0, 2\pi]$ and each $(x, y) \in [-\varrho - \eta, \varrho + \eta] \times \mathrm{R}$ the inequality

$$|h_\mu(t, x, y)| \le \widetilde{p}(t) + \widetilde{q}(t)|y|$$

is fulfilled. Then by Theorem 3.2 we obtain

$$d[I - H_\mu, \widetilde{\Omega}] = 1,\ d[I - H_\mu, \Delta_1] = 1 \text{ and } d[I - H_\mu, \Delta_2] = 1,$$

where

$$(H_\mu x)(t) = x(0) + x'(0) - x'(2\pi)$$
$$+ \int_0^{2\pi} g_\mu(t,s)\big[h_\mu(s,x(s),x'(s)) - \big(x(0)+x'(0)-x'(2\pi)\big)\mu\big]\,ds$$
$$+ \widetilde{g}_\mu(t,t_1)[\widetilde{J}(x(t_1)) - x(t_1)] + g_\mu(t,t_1)[M(x'(t_1)) - x'(t_1)].$$

Denote

$$\Delta = \widetilde{\Omega}\backslash\overline{(\Delta_1 \cup \Delta_2)}. \tag{3.52}$$

Then, from the additivity of the Leray–Schauder topological degree, we have

$$d[I - H_\mu, \Delta] = -1.$$

Thus there is a solution u of the problem

$$(I - H_\mu)x = 0 \tag{3.53}$$

which for some $t_u \in [0, 2\pi]$ satisfies (3.31). Moreover from (3.39), (3.42), (3.44), (3.46), (3.50) we can see that u is a solution of the equation

$$x'' + a(t)x' + b(t)x = h(t, x, x'),$$

and we have for a.e. $t \in [0, 2\pi]$

$$|u''| = |\overline{h}(t,u,u')| \le |\overline{f}(t,u,u')| + (|a(t)| + q(t))|u'| + 4p(t)$$

$$\le |\widetilde{f}(t,u,u') - a(t)u' + b(t)u| + (|a(t)| + q(t))|u'| + 4p(t) \le 2(|a(t)| + q(t))|u'| + 6p(t),$$

$$\left|\frac{u''u'}{1+u'^2}\right| \le 2(|a(t)| + q(t)) + 6p(t).$$

Integrating this inequality on (t_u, t) we get for each $t \in [0, 2\pi]$

$$1 + u'^2(t) \le (1 + u'^2(t_u))e^{4(\|a\|_1 + \|q\|_1 + 3\|p\|_1)}$$

$$|u'(t)| \le (1 + |u'(t_u)|)e^{2(\|a\|_1 + \|q\|_1 + 3\|p\|_1)} < \widetilde{N}_1.$$

Then we have $\|u'\|_\infty < \widetilde{N}_1$, $\|u\|_\infty < \widetilde{N}_2$ for every solution $u \in \Delta$ of (3.48) and from the excision property of the degree we have

$$d[I - H_\mu, \Delta] = d[I - H_\mu, \Omega_2] = -1$$

Finally, from (3.42)–(3.44) and (3.52) $H_\mu = F_\mu$ for $x \in \overline{\Omega}_2$ follows and so

$$d[I - H_\mu, \Omega_2] = d[I - F_\mu, \Omega_2] = -1. \qquad \square$$

Corollary 3.7 *Let the assumptions of Theorem 3.6 be satisfied. Then the problem (1.1)–(1.3) has a solution u, which fulfills*

$$\sigma_2(t_u) < u(t_u) < \sigma_1(t_u)$$

for some $t_u \in [0, 2\pi]$.

References

[1] Rachůnková, I., Tvrdý, M.: *Impulsive periodic boundary value problem and topological degree.* Functional Differential Equations **9**, 3-4 (2002), 471–498.

[2] Rachůnková, I., Tvrdý, M.: *Periodic boundary value problem for nonlinear second order differential equations with impulses – part I.* Preprint 148/2002 MÚAV ČR, Prague.

[3] Draessler, J., Rachůnková, I.: *On three solutions of the second order periodic boundary value problem.* Nonlinear Oscillations **4** (2002), 471–486.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 55–60

Kronecker Modules and Reductions of a Pair of Bilinear Forms *

Giovanni FALCONE [1], M. Alessandra VACCARO [2]

Dipartimento di Matematica ed Applicazioni
Università degli Studi di Palermo
Via Archirafi, 34, I-90123 Palermo, Italia
e-mail: [1]*falcone@math.unipa.it*
[2]*vaccaro@math.unipa.it*

(Received December 17, 2003)

Abstract

We give a short overview on the subject of canonical reduction of a pair of bilinear forms, each being symmetric or alternating, making use of the classification of pairs of linear mappings between vector spaces given by J. Dieudonné.

Key words: Kronecker modules, bilinear forms.

2000 Mathematics Subject Classification: 11E04

The problem of a simultaneous reduction of a pair of symmetric bilinear forms over a given field is classic: this problem has been solved, for fields of characteristic zero, in 1868 by K. Weierstrass, under the assumption that both the forms are not degenerate. Two papers, the first of which by L. Kronecker [4], dated 1890, the second by L. E. Dickson [1], dated 1909, give a complete answer for fields of characteristic zero. Later J. Williamson [9] (1935), [10] (1945) showed that similar results were also valid for any field of characteristic $\neq 2$, but the condition that one of the form is not degenerate is needed again. The case where both the forms are degenerate has been solved by W. Waterhouse [7] (1976), as well as the case of a pair of symmetric bilinear forms (even degenerate) over a field of characteristic 2, [8] (1977).

*Research supported by M.U.R.S.T.

Two papers of the 70's by P. Gabriel [3] and R. Scharlau [6] showed that the classification of pairs of linear mappings (or Kronecker modules) by J. Dieudonné [2] (1946), which goes back to the mentioned paper of Kronecker, plays a fundamental role in studying pairs of bilinear forms. More precisely, Scharlau gives a complete answer for a pair of alternating bilinear forms, as pointed out by Waterhouse in [8].

The case where one of the forms is symmetric and the other is alternating has been treated by several authors and can be found in two papers by C. Riehm [5] and Gabriel [3], but the arguments used by Riehm, as well as the ones used by Waterhouse, do not concern any longer the theory of Kronecker modules.

We provide a statement (Theorem 2 and following discussion) which gives an overview on the subject from the point of view of Kronecker modules. This allows us to give an alternative proof of some results which had been given in the mentioned papers.

1. A *Kronecker module over the field* K is a pair

$$\Phi = (\varphi_1 : V' \to V'';\ \varphi_2 : V' \to V'')$$

of linear mappings from a K-vector space V' into a K-vector space V''. We write for short $\Phi = (V', V''; \varphi_1, \varphi_2)$, or simply $\Phi = (V',\ V'')$. An *isomorphism* $\iota : \Phi \to \Psi$ from Φ onto the Kronecker module $\Psi = (W', W''; \psi_1, \psi_2)$ is a pair of bijective linear mappings $\iota = (\iota' : V' \to W'; \iota'' : V'' \to W'')$ such that $\iota''\varphi_h = \psi_h\iota'$ $(h = 1, 2)$.

From the Kronecker module Φ we obtain two further Kronecker modules: the *opposite* of Φ, that is the Kronecker module

$$\Phi^\circ := (V', V''; \varphi_2, \varphi_1),$$

and the *transpose* of Φ, that is the Kronecker module

$${}^t\Phi := (V''^*, V'^*; {}^t\varphi_1, {}^t\varphi_2),$$

where, for a given linear mapping φ, we denote by ${}^t\varphi$ the *transpose* of φ, from the dual V''^* of V'' into the dual V'^* of V', defined by

$${}^t\varphi(x''^*)(x') = x''^*(\varphi(x'))$$

for all $x' \in V'$ and $x''^* \in V''^*$. The Kronecker module Φ is *self-transpose* if there exists an isomorphism $\Phi \to {}^t\Phi$.

Any Kronecker module can be decomposed into the direct sum of indecomposable submodules and, for two such decompositions, the Krull–Remak–Schmidt Theorem applies. This means $\Phi(F) = \Phi_1 \bigoplus \ldots \bigoplus \Phi_t$ for a fixed number t of indecomposable submodules Φ_i, determined up to permutations and isomorphisms. Indecomposable Kronecker modules were classified by Kronecker [4] and Dieudonné [2]: let

$$\begin{aligned}
&\Phi_\varphi = (K^n, K^n; id, \varphi) && n > 0, \quad \varphi \in \mathbf{End}_K K^n, \\
&\Phi_n = (K^n, K^{n+1}; \varphi_1, \varphi_2) && n \geq 0, \text{ where} \\
& && \varphi_1 : (x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, 0), \\
& && \varphi_2 : (x_1, \ldots, x_n) \mapsto (0, x_1, \ldots, x_n),
\end{aligned} \tag{1}$$

then an indecomposable Kronecker module is isomorphic to one of $\Phi_n, {}^t\Phi_n, \Phi_\varphi$, or Φ_φ°, for a suitable endomorphism $\varphi \in \mathbf{End}_K(K^n)$ which makes K^n into an indecomposable $K[\varphi]$-module. Note that Φ_φ is self-transpose, whereas Φ_n is not, hence Φ is self-transpose precisely if $\dim V' = \dim V''$.

The Krull–Remak–Schmidt Theorem has the following useful corollary (the *exchange theorem*): let $\Phi = \Upsilon_1 \bigoplus \Upsilon_2$ and $\Phi = \bar{\Upsilon}_1 \bigoplus \bar{\Upsilon}_2$ be two decompositions of Φ. Assume that no indecomposable component of Υ_1 (resp. $\bar{\Upsilon}_1$) is isomorphic to any indecomposable component of Υ_2 (resp. $\bar{\Upsilon}_2$), then $\Phi = \Upsilon_1 \bigoplus \bar{\Upsilon}_2 = \bar{\Upsilon}_1 \bigoplus \Upsilon_2$.

2. Let $f_h : V \times V \to K$, $h = 1, 2$, be a pair of bilinear forms, each being symmetric or alternating, defined on a K-vector space V. We can associate to the triple $F = (V; f_1, f_2)$ the self-transpose Kronecker module

$$\Phi(F) := (\bar{f}_1 : V \to V^*, \ \bar{f}_2 : V \to V^*),$$

where, for $x \in V$, $\bar{f}_h(x)$ is the mapping $y \mapsto f_h(x, y)$. For a subspace U of V, we can set

$$U^\perp = \{v \in V : f_1(v, x) = f_2(v, x) = 0 \text{ for any } x \in U\}.$$

We say that F is *decomposable* if $V = U + U^\perp$ for some nontrivial subspace U.

Manifestly, any decomposition of V into the direct sum of two subspaces U_1 and U_2, orthogonal with respect to both f_1 and f_2, provides a decomposition of $\Phi(F)$. The converse is generally not true.

The canonical identification $V = V^{**}$ yields the consequent identification $\Phi(F) = {}^t\Phi(F)$. Hence, the number of components of $\Phi(F)$ isomorphic to Φ_n is the same of the ones isomorphic to ${}^t\Phi_n$. This provides a decomposition of $\Phi(F)$ into self-transpose submodules, having no isomorphic components in common, which gives in turn an orthogonal decomposition of V, as the following lemma claims.

Lemma 1 *Let $\Phi(F) = \Upsilon_1 \bigoplus \Upsilon_2$ with self-transpose Υ_h, $h = 1, 2$. Assume that no component of Υ_1 is isomorphic to any component of Υ_2, then F decomposes.*

Proof Let $\Upsilon_h \equiv (U_h, W_h^*)$, then $V = U_1 \bigoplus U_2$ and $V^* = W_1^* \bigoplus W_2^*$. Consequently $V = W_1 \bigoplus W_2$, corresponding to the decomposition $\Phi(F) = {}^t\Phi(F) = {}^t\Upsilon_1 \bigoplus {}^t\Upsilon_2$. By the exchange theorem, we have the further decompositions $\Phi(F) = {}^t\Upsilon_1 \bigoplus \Upsilon_2 = \Upsilon_1 \bigoplus {}^t\Upsilon_2$, hence $V = W_1 \bigoplus U_2 = U_1 \bigoplus W_2$. As we have $\bar{f}_1(x), \bar{f}_2(x) \in W_i^*$ for any $x \in U_i$, then for any $y \in W_j$, $j \neq i$, it follows $f_h(x, y) = 0$, that is, the latter decompositions of V are orthogonal. $\square$

In view of the above lemma, indecomposable F correspond to Kronecker modules $\Phi(F)$ isomorphic to either $(\Phi_\varphi)^r$ or $(\Phi^\circ_\varphi)^r$, or $(\Phi_n)^s \bigoplus ({}^{\mathrm{t}}\Phi_n)^s$. Moreover, direct computations on the bases show that $s = 1$ and

$r = 1$ for an indecomposable pair of symmetric forms,
$r = 1, 2$ for an indecomposable pair, where one is symmetric and the other is alternating,
$r = 2$ for an indecomposable pair of alternating bilinear forms,

according to [7], [8], [5] and [6]. Therefore we have

Theorem 2 *Let F be indecomposable. Then the Kronecker module $\Phi(F)$ is isomorphic to either Φ_φ or ${\Phi_\varphi}^\circ$, or $\Phi_n \bigoplus {}^{\mathrm{t}}\Phi_n$.*

Let $\mathcal{U}$, $\mathcal{W}$ be bases such that $\Phi(F)$ is represented by matrices (S_1, S_2), the entries of which are given in (1), and A be the matrix of the rowed coordinates of the vectors in $\mathcal{W}^*$ with respect to $\mathcal{U}^*$. A simultaneous reduction to canonical form is now reached through the condition that the product S_hA, $(h = 1, 2)$, as a representation of F, is symmetric or alternating.

In particular, if both the forms are degenerate, i.e. $\Phi(F) = \Phi_n \bigoplus {}^{\mathrm{t}}\Phi_n$, by definition of Φ_n the bases

$$\mathcal{U} = \{u'_1, \ldots, u'_n, u''_1, \ldots, u''_{n+1}\} \quad \text{and} \quad \mathcal{W} = \{w'_1, \ldots, w'_{n+1}, w''_1, \ldots, w''_n\}$$

of V are such that,

$$\begin{aligned} f_1(u'_i, w'_j) = f_1(w''_i, u''_j) &= \delta_{i,j}, \\ f_2(u'_i, w'_j) = f_2(w''_i, u''_j) &= \delta_{i,j-1}, \end{aligned} \tag{2}$$

where $1 \le i \le n$ and $1 \le j \le n + 1$.

Let (J_1, J_2) be the matrix representation of $\Phi(F)$ with respect to $\mathcal{U}$ and $\mathcal{W}^*$, the entries of which are given by (2). Then, the matrix A fulfills the equations $J_hA = S_h$ $(h = 1, 2)$. The partitions $\mathcal{U} = \{u'_1, \ldots, u'_n\} \cup \{u''_1, \ldots, u''_{n+1}\}$ and $\mathcal{W} = \{w'_1, \ldots, w'_{n+1}\} \cup \{w''_1, \ldots, w''_n\}$ allow one to write the equations $J_hA = S_h$ in blocks as

$$\begin{pmatrix} \mathbf{J}_h & \mathbf{0} \\ \mathbf{0} & {}^{\mathrm{t}}\mathbf{J}_h \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} S_h^{11} & S_h^{12} \\ \varepsilon\, {}^{\mathrm{t}}S_h^{12} & S_h^{22} \end{pmatrix} \qquad \varepsilon = \pm 1, \tag{3}$$

where we put

$$J_h = \begin{pmatrix} \mathbf{J}_h & \mathbf{0} \\ \mathbf{0} & {}^{\mathrm{t}}\mathbf{J}_h \end{pmatrix},$$

and

$$(\mathbf{J}_1, \mathbf{J}_2) = \left(\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{pmatrix} \right)$$

is the current representation of the Kronecker module Φ_n.

One sees that the products ${}^{\mathrm{t}}\mathbf{J}_h A_{22} = S_h^{22}$ are symmetric or alternating just if $A_{22} = \mathbf{0}$, hence $S_h^{22} = \mathbf{0}$. Furthermore, the square matrix A_{12} cannot be singular, so, up to replacing each of the vectors $u''_1, \ldots, u''_{n+1}$ by a suitable linear combination of themselves, we may assume $\mathcal{U}$ such that $A_{12} = \mathbf{I}_{n+1}$. Therefore, the equations (3) turn into

$$\begin{pmatrix} \mathbf{J}_h & \mathbf{0} \\ \mathbf{0} & Y_h \end{pmatrix} \begin{pmatrix} A_{11} & \mathbf{I}_{n+1} \\ A_{21} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} S_h^{11} & \mathbf{J}_h \\ \varepsilon\,{}^{\mathrm{t}}\mathbf{J}_h & \mathbf{0} \end{pmatrix}, \tag{4}$$

for suitable matrices Y_h which play no role for our purposes. Hence F has a representation

$$\left(\begin{pmatrix} T_1 & \mathbf{J}_1 \\ \varepsilon\,{}^{\mathrm{t}}\mathbf{J}_1 & \mathbf{0} \end{pmatrix}, \begin{pmatrix} T_2 & \mathbf{J}_2 \\ \varepsilon\,{}^{\mathrm{t}}\mathbf{J}_2 & \mathbf{0} \end{pmatrix} \right), \tag{5}$$

where we write T_h instead of S_h^{11}.

Replace now each of the vectors u'_i in $\mathcal{U}$ by $u'_i + \sum_{j=1}^{n+1} c_{ij} u''_j$, then

$$f_1\left(u'_r + \sum_{j=1}^{n+1} c_{rj} u''_j, u'_s + \sum_{j=1}^{n+1} c_{sj} u''_j \right) = f_1(u'_r, u'_s) + c_{rs} + c_{sr},$$

$$f_2\left(u'_r + \sum_{j=1}^{n+1} c_{rj} u''_j,\ u'_s + \sum_{j=1}^{n+1} c_{sj} u''_j \right) = f_2(u'_r, u'_s) + c_{r,\,s+1} + c_{s,\,r+1},$$

for $r, s = 1, \ldots, n$. Let $\mathrm{char} K \neq 2$, then it is possible to find entries c_{ij} which make the above quantities zero. Let $\mathrm{char} K = 2$, then it is still possible to do that, provided $r \neq s$.

The above arguments can be summarized in the following result, which has been proved by Scharlau [6] for pairs of alternating forms, while Waterhouse [7], [8] proved it for pairs of symmetric forms, but there he made use of other tecniques.

Theorem 3 *Let F be an indecomposable pair of degenerate bilinear forms on a K-vector space V, each being symmetric or alternating. Then, V has odd dimension $2n+1$ over K and F has a representation*

$$\left(\begin{pmatrix} D_1 & \mathbf{J}_1 \\ \varepsilon\,{}^{\mathrm{t}}\mathbf{J}_1 & \mathbf{0} \end{pmatrix}, \begin{pmatrix} D_2 & \mathbf{J}_2 \\ \varepsilon\,{}^{\mathrm{t}}\mathbf{J}_2 & \mathbf{0} \end{pmatrix} \right)$$

for suitable diagonal matrices D_1, D_2. Moreover, if the characteristic of K is not 2, there exists a representation with $D_1 = D_2 = \mathbf{0}$.

References

[1] Dickson, L. E.: *Equivalence of pairs of bilinear or quadratic forms under rational transformations.* Trans. AMS **10** (1909), 347–360.

[2] Dieudonné, J.: *Sur la reduction canonique des couples de matrices.* Bull. Soc. math. France **74** (1946), 130–146.

[3] Gabriel, P.: *APPENDIX: Degenerate Bilinear Forms.* J. Algebra **31** (1974), 67–72.

[4] Kronecker, L.: *Algebraische Reduktion der Scharen bilinearer Formen.* Sitzungsber. Akad. Berlin (1890), 763–776.

[5] Riehm, C.: *The equivalence of bilinear forms.* J. Algebra **31** (1974), 45–66.

[6] Scharlau, R.: *Paare alternierender Formen.* Math. Z. **147** (1976), 13–19.

[7] Waterhouse, W.: *Pairs of quadratic forms.* Inventiones Math. **37** (1976), 157–164.

[8] Waterhouse, W.: *Pairs of symmetric bilinear forms in characteristic 2.* Pacific J. Math. **69**, 1 (1977), 275–283.

[9] Williamson, J.: *The equivalence of non-singular pencils of Hermitian matrices in an arbitrary field.* Amer. J. Math. **57** (1935), 475–490.

[10] Williamson, J.: *Note on the equivalence of nonsingular pencils of Hermitian matrices.* Bull. AMS **51** (1945), 894–897.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 61–64

Class Preserving Mappings of Equivalence Systems

Ivan CHAJDA

Department of Algebra and Geometry, Faculty of Science,
Palacký University, Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: chajda@inf.upol.cz

(Received February 6, 2004)

Abstract

By an equivalence system is meant a couple $\mathcal{A} = (A, \theta)$ where A is a non-void set and θ is an equivalence on A. A mapping h of an equivalence system $\mathcal{A}$ into $\mathcal{B}$ is called a class preserving mapping if $h([a]_\theta) = [h(a)]_{\theta'}$ for each $a \in A$. We will characterize class preserving mappings by means of permutability of θ with the equivalence Φ_h induced by h.

Key words: Equivalence relation, equivalence system, relational system, homomorphism, strong homomorphism, permuting equivalences.

2000 Mathematics Subject Classification: 08A02, 08A35, 03E02

For the basic concepts, the reader is referred to [1],[2],[3]. Let R and S be binary relations on a non-void set A. As usually, their *relational product* will be denoted by $R \circ S$, i.e. $R \circ S = \{\langle a, b\rangle \in A^2; \exists c \in A \text{ with } \langle a, c\rangle\rangle \in R$ and $\langle c, b\rangle \in S\}$. We will say that R, S *permute* (or they are *permutable*) if $R \circ S = S \circ R$.

Lemma 1 *Let R, S be symmetric relations on A. Then $R \circ S \subseteq S \circ R$ is equivalent to $R \circ S = S \circ R$.*

Proof If $R \circ S \subseteq S \circ R$ then, due to symmetry,

$$S \circ R = S^{-1} \circ R^{-1} = (R \circ S)^{-1} \subseteq (S \circ R)^{-1} = R^{-1} \circ S^{-1} = R \circ S$$

thus S, R permute. The converse is trivial. □

By a *relational system* is meant a pair $\mathcal{A} = (A, R)$, where $A \neq \emptyset$ is a set and R is a binary relation on A. If R is an equivalence relation, $\mathcal{A} = (A, R)$ will be called an *equivalence system*.

We are going to introduce a quotient relational system as follows.

Definition 1 Let $\mathcal{A} = (A, R)$ be a relational system and Φ be an equivalence on A. Define a binary relation R/Φ on the factor set (i.e. a partition) A/Φ as follows: $\langle [a]_\Phi, [b]_\Phi \rangle \in R/\Phi$ iff there exist $x \in [a]_\Phi, y \in [b]_\Phi$ with $\langle x, y \rangle \in R$. Then $\mathcal{A}/\Phi = (A/\Phi, R/\Phi)$ will be called a *quotient relational system* of $\mathcal{A}$ by Φ.

Remark 1 It is evident that if R is reflexive or symmetric then R/Φ has the corresponding property.

Lemma 2 *Let $\mathcal{A} = (A, R)$ be a relational system and R be transitive. Let Φ be an equivalence on A and $\Phi \circ R \subseteq R \circ \Phi$. Then R/Φ is transitive, too.*

Proof Suppose $\langle [a]_\Phi, [b]_\Phi \rangle \in R/\Phi$ and $\langle [b]_\Phi, [c]_\Phi \rangle \in R/\Phi$. Then there exist $x, y, y', z \in A$ such that $x \in [a]_\Phi$, $y, y' \in [b]_\Phi$, $z \in [c]_\Phi$ and $\langle x, y \rangle \in R$, $\langle y', z \rangle \in R$. Hence $\langle x, z \rangle \in R \circ \Phi \circ R \subseteq R \circ R \circ \Phi \subseteq R \circ \Phi$. Thus there is $w \in A$ with $\langle x, w \rangle \in R$ and $\langle w, z \rangle \in \Phi$, i.e. $w \in [z]_\Phi = [c]_\Phi$. By the Definition, $\langle [a]_\Phi, [c]_\Phi \rangle \in R/\Phi$ proving transitivity of R/Φ. □

Let $\mathcal{A} = (A, R), \mathcal{B} = (B, S)$ be relational systems. A mapping $h : A \to B$ is called a *homomorphism* of $\mathcal{A}$ into $\mathcal{B}$ if $\langle a, b \rangle \in R$ implies $\langle h(a), h(b) \rangle \in S$.

A homomorphism h of $\mathcal{A}$ into $\mathcal{B}$ is called *strong* if for each $\langle x, y \rangle \in S$ there exist $a, b \in A$ such that $\langle a, b \rangle \in R$ and $h(a) = x, h(b) = y$. Let $\mathcal{A} = (A, \theta), \mathcal{B} = (B, \theta')$ be equivalence systems. A mapping $h : A \to B$ is called *class preserving* if $h([a]_\theta) = [h(a)]_{\theta'}$ for each $a \in A$.

Lemma 3 *Let $\mathcal{A} = (A, \theta)$, $\mathcal{B} = (B, \theta')$ be equivalence systems and $h : A \to B$ a surjective class preserving mapping. Then h is a strong homomorphism of $\mathcal{A}$ onto $\mathcal{B}$.*

Proof It is evident that $\langle a, b \rangle \in \theta$ implies $\langle h(a), h(b) \rangle \in \theta'$, i.e. it is a surjective homomorphism of $\mathcal{A}$ onto $\mathcal{B}$. Suppose $\langle c, d \rangle \in \theta'$. Then there is $a \in A$ with $h(a) = c$ and $d \in [c]_{\theta'} = [h(a)]_{\theta'}$. Hence, there exists $x \in [a]_\theta$ such that $h(x) = d$. Since $\langle a, x \rangle \in \theta$, h is a strong homomorphism. □

Example 1 The converse of Lemma 3 does not hold in general. Consider e.g. $\mathcal{A} = (A, \theta)$, $\mathcal{B} = (B, \theta')$ where $A = \{x_1, x_2, y_1, y_2, z_1, z_2\}$, $B = \{a, b, c\}$, $\theta' = B \times B$ and θ is determined by the partition $\{x_1, x_2\}, \{y_1, y_2\}, \{z_1, z_2\}$. Let $h : A \to B$ is defined as follows: $h(x_1) = h(y_1) = a$, $h(x_2) = h(z_1) = b$, $h(y_2) = h(z_2) = c$. Then h is a surjective strong homomorphism of $\mathcal{A}$ onto $\mathcal{B}$ but it is not a class preserving mapping; e.g. for x_1 we have

$$h([x_1]_\theta) = h(\{x_1, x_2\}) = \{a, b\} \neq \{a, b, c\} = [a]_{\theta'} = [h(x_1)]_{\theta'}.$$

Theorem 1 *Let $\mathcal{A} = (A, \theta)$, $\mathcal{B} = (B, \theta')$ be equivalence systems and $h : A \to B$ a surjective mapping. The following are equivalent:*

(a) h is a class preserving mapping;

(b) h is a homomorphism of $\mathcal{A}$ onto $\mathcal{B}$ and for each $x, y \in A$ with $\langle h(x), h(y)\rangle \in \theta'$ there exists $z \in A$ such that $\langle x, z\rangle \in \theta$ and $h(y) = h(z)$.

Proof (a) $\Rightarrow$ (b) by Lemma 3 and its proof. Prove (b) $\Rightarrow$ (a). Since h is a homomorphism, we easily get $h([a]_\theta) \subseteq [h(a)]_{\theta'}$. Suppose $c \in [h(a)]_{\theta'}$. Then $c = h(w)$ for some $w \in A$. By (b) there exists $z \in A$ such that $\langle a, z\rangle \in \theta$ and $h(z) = h(w) = c$. Since $z \in [a]_\theta$, we conclude $h([a]_\theta) = [h(a)]_{\theta'}$. □

Let $h : A \to B$ be a mapping. Denote by Φ_h the so-called *h-induced equivalence* on A, i.e.

$$\langle x, y\rangle \in \Phi_h \quad \text{if and only if} \quad h(x) = h(y).$$

Let Φ be an equivalence on A. Denote by h_Φ the so-called *natural mapping* $h_\Phi : A \to A/\Phi$ defined by $h_\Phi(a) = [a]_\Phi$.

Theorem 2 *Let $\mathcal{A} = (A, \theta)$ be an equivalence system and Φ be an equivalence on A. Suppose that θ, Φ permute. Then the natural mapping h_Φ is a class preserving mapping of $\mathcal{A}$ onto the quotient equivalence system $\mathcal{A}/\Phi = (A/\Phi, \theta/\Phi)$.*

Proof By Lemma 2 and the previous Remark, $\mathcal{A}/\Phi$ is clearly a quotient equivalence system. Of course, h_Φ is a surjective mapping. Suppose $\langle a, b\rangle \in \theta$. Then $\langle [a]_\Phi, [b]_\Phi\rangle \in \theta/\Phi$ thus h_Φ is a homomorphism of $\mathcal{A}$ onto $\mathcal{A}/\Phi$. Let $\langle [x]_\Phi, [y]_\Phi\rangle \in \theta/\Phi$. Then there exist $a \in [x]_\Phi, b \in [y]_\Phi$ such that $\langle a, b\rangle \in \theta$. Hence $\langle x, b\rangle \in \Phi \circ \theta = \theta \circ \Phi$, i.e. there exists $z \in A$ such that $\langle x, z\rangle \in \theta$ and $\langle z, b\rangle \in \Phi$, i.e. $h_\Phi(z) = h_\Phi(b)$. By (b) of Theorem 1, h_Φ is a class preserving mapping. □

Theorem 3 *Let $\mathcal{A} = (A, \theta), \mathcal{B} = (B, \theta')$ be equivalence systems and $h : A \to B$ a surjective strong homomorphism of $\mathcal{A}$ onto $\mathcal{B}$. Then h is a class preserving mapping if and only if θ and the h-induced equivalence Φ_h permute.*

Proof Let h be a class preserving mapping and suppose $\langle x, z\rangle \in \Phi_h \circ \theta$. Then there exists $y \in A$ with $\langle x, y\rangle \in \Phi_h$ and $\langle y, z\rangle \in \theta$. Thus $h(x) = h(y)$ and, as h is a homomorphism, $\langle h(x), h(z)\rangle \in \theta'$. By (b) of Theorem 1, there exists $u \in A$ with $\langle x, u\rangle \in \theta$ and $h(u) = h(z)$, i.e. $\langle u, z\rangle \in \Phi_h$. Hence $\langle x, z\rangle \in \theta \circ \Phi_h$ showing $\Phi_h \circ \theta \subseteq \theta \circ \Phi_h$. By Lemma 1, θ and Φ_h permute.

Conversely, let h be a surjective strong homomorphism and suppose $\theta \circ \Phi_h = \Phi_h \circ \theta$. Since h is a homomorphism we have $h([a]_\theta) \subseteq [h(a)]_{\theta'}$. Let $x \in [h(a)]_{\theta'}$. Then $\langle x, h(a)\rangle \in \theta'$. Since h is a strong homomorphism, there exist $b, c \in A$ such that $\langle b, c\rangle \in \theta$ and $h(b) = x, h(c) = h(a)$. Thus $\langle c, a\rangle \in \Phi_h$ and we have $\langle b, a\rangle \in \theta \circ \Phi_h = \Phi_h \circ \theta$. Hence, there exists $z \in A$ with $\langle b, z\rangle \in \Phi_h, \langle z, a\rangle \in \theta$. Thus $z \in [a]_\theta$ and $h(z) = h(b) = x$, i.e. h is a class preserving mapping. □

References

[1] Madarász, R., Crvenković, S.: Relacione algebre. *Matematički Institut, Beograd*, 1992.

[2] Maltsev, A. I.: Algebraic systems. *Nauka, Moskva*, 1970, (in Russian).

[3] Riguet, J.: *Relations binaires, fermetures, correspondances de Galois.* Bull. Soc. Math. France **76** (1948), 114–155.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 65–73

Local Versions of some Congruence Properties in Single Algebras *

Ivan CHAJDA [1], Gerhard DORFER [2], Helmut LÄNGER [3]

[1] *Department of Algebra and Geometry, Faculty of Science,
Palacký University, Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: chajda@risc.upol.cz*

[2,3] *Institute of Algebra and Computational Mathematics
Vienna University of Technology
Wiedner Hauptstr. 8-10/118, A-1040 Vienna, Austria
e-mail: g.dorfer@tuwien.ac.at
h.laenger@tuwien.ac.at*

(Received November 20, 2003)

Abstract

We investigate some local versions of congruence permutability, regularity, uniformity and modularity. The results are applied to several examples including implication algebras, orthomodular lattices and relative pseudocomplemented lattices.

Key words: Congruence permutability, congruence regularity, congruence uniformity, congruence modularity.

2000 Mathematics Subject Classification: 08A30, 08B05

Congruence permutability, regularity, uniformity and modularity are well studied concepts in universal algebra. For the convenience of the reader we refer to [4]. We introduce and study some local versions of these notions.

In the following let $\mathcal{A} = (A, F)$ be an arbitrary but fixed algebra and a, b arbitrary but fixed elements of A.

*Research of all three authors supported by ÖAD, Cooperation between Austria and Czech Republic in Science and Technology, grant number 2003/1; the second author acknowledges support by the Austrian science fund FWF under grant number S 8312.

Definition 1 For every positive integer n and every $i \in \{1, \ldots, n\}$ let C_{ni} denote the set of all n-ary functions on A which are compatible with all congruences on $\mathcal{A}$ with respect to the i-th variable, i.e. C_{ni} consists of all functions $f : A^n \to A$ satisfying the following condition: If $a_1, \ldots, a_n, \bar{a}_i \in A$, $\theta \in \mathrm{Con}(\mathcal{A})$ and $a_i \theta \bar{a}_i$ then

$$f(a_1, \ldots, a_i, \ldots, a_n) \theta f(a_1, \ldots, \bar{a}_i, \ldots, a_n).$$

Moreover, put $C_n := C_{n1} \cap \ldots \cap C_{nn}$ the set of all compatible n-ary functions on $\mathcal{A}$ for all positive integers n.

Definition 2 $\mathcal{A}$ is called (a, b)-*permutable* if for all $\theta, \phi \in \mathrm{Con}(\mathcal{A})$ the assertions $a(\theta \circ \phi)b$ and $a(\phi \circ \theta)b$ are equivalent. $\mathcal{A}$ is called (a, b)-*regular* if for all $\theta, \phi \in \mathrm{Con}(\mathcal{A})$, $[a]\theta = [a]\phi$ implies $[b]\theta = [b]\phi$. $\mathcal{A}$ is called (a, b)-*uniform* if $|[a]\theta| = |[b]\theta|$ for all $\theta \in \mathrm{Con}(\mathcal{A})$.

Remark 1 The following properties follow directly from Defintion 2:

- $\mathcal{A}$ is (a, b)-permutable if and only if $\mathcal{A}$ is (b, a)-permutable.
- $\mathcal{A}$ is permutable if and only if it is (c, d)-permutable for all $c, d \in A$.
- $\mathcal{A}$ is regular if and only if it is (c, d)-regular for all $c, d \in A$.
- $\mathcal{A}$ is (a, b)-uniform if and only if $\mathcal{A}$ is (b, a)-uniform.
- $\mathcal{A}$ is uniform if and only if it is (c, d)-uniform for all $c, d \in A$.

Theorem 1 *(i) If there exists an $f \in C_1$ with $f(b) = a$ and $f(a) = b$ then $\mathcal{A}$ is (a, b)-permutable.*

(ii) If there exist $f, g \in C_1$ with $f(b) = a$ and $g(f(x)) = x$ for all $x \in A$ then $\mathcal{A}$ is (a, b)-regular.

(iii) If there exist $f, g \in C_1$ such that $f(b) = a$ and $f(g(x)) = g(f(x)) = x$ for all $x \in A$ then $\mathcal{A}$ is (a, b)-uniform.

Proof Let $\theta, \phi \in \mathrm{Con}(\mathcal{A})$.

(i) If $a(\theta \circ \phi)b$ then there exists an element $c \in A$ with $a\theta c\phi b$ and hence $a = f(b)\phi f(c)\theta f(a) = b$ showing $a(\phi \circ \theta)b$, i.e. $a(\theta \circ \phi)b$ implies $a(\phi \circ \theta)b$. The converse implication follows by symmetry.

(ii) Assume $[a]\theta = [a]\phi$. If $c \in [b]\theta$ then $f(c) \in [f(b)]\theta = [a]\theta = [a]\phi$ and hence $c = g(f(c)) \in [g(a)]\phi = [g(f(b))]\phi = [b]\phi$ showing $[b]\theta \subseteq [b]\phi$. The converse inclusion follows by symmetry.

(iii) If $c \in [a]\theta$ then $g(c) \in [g(a)]\theta = [g(f(b))]\theta = [b]\theta$. If $d \in [b]\theta$ then $f(d) \in [f(b)]\theta = [a]\theta$. Moreover, $f(g(x)) = g(f(x)) = x$ for all $x \in A$. Hence $g|_{[a]\theta}$ and $f|_{[b]\theta}$ are mutually inverse bijections between $[a]\theta$ and $[b]\theta$ proving $|[a]\theta| = |[b]\theta|$. □

Example 1 An implication algebra (cf. [1]) is a groupoid $(A, \cdot)$ satisfying the identities

$$(xy)x = x, \quad (xy)y = (yx)x, \quad x(yz) = y(xz).$$

This implies $xx = yy$, i.e. xx is a constant denoted by 1 (if $A \neq \emptyset$ which we will assume). Moreover, $1x = (xx)x = x$ and $x1 = (1x)1 = 1$. With the partial order

$$x \leq y \text{ if and only if } xy = 1$$

$(A, \leq)$ is a $\vee$-semilattice with $x \vee y = (xy)y$ in which every interval $[c, 1]$ is a Boolean algebra. The element xy coincides with the complement of $x \vee y$ in the interval $[y, 1]$.

An implication algebra is (a, b)-permutable if and only if a and b have a common lower bound, i.e. if and only if there exists an interval $[c, 1]$ with $a, b \in [c, 1]$: Firstly suppose that such an element c exists. Let $+_c$ denote the symmetric difference in $[c, 1]$. $+_c$ can be represented as a polynomial function and thus $x +_c y$ makes sense for all $x, y \in A$ and is in C_2. Consequently $f(x) = x +_c (a +_c b)$ is in C_1 and obviously satisfies condition (i) of Theorem 1.

On the other hand, suppose a and b do not have a common lower bound. Let θ and ϕ be the principal congruences generated by $(a, 1)$ and $(b, 1)$, respectively. It can be verified easily that $(x, y) \in \theta$ if and only if $x \wedge y$ exists in A and $1 +_{x \wedge y} (x +_{x \wedge y} y) \geq a \vee (x \wedge y)$. Similarly ϕ can be characterized.

Obviously $(a, b) \in \theta \circ \phi$. Assume $(a, b) \in \phi \circ \theta$, i.e. there is $d \in A$ such that $(a, d) \in \phi$ and $(d, b) \in \theta$. $(a, d) \in \phi$ implies $(a, a \vee d) \in \phi$ which means $1 +_a (a +_a (a \vee d)) \geq b \vee a$ by the above characterization of ϕ. This implies $a \vee d \leq 1 +_a (a \vee b)$ and hence $(a \vee b) \wedge (a \vee d) = a$. $(d, b) \in \theta$ implies the existence of $b \wedge d$ and we infer $a \vee (b \wedge d) \leq (a \vee b) \wedge (a \vee d) = a$, hence $b \wedge d \leq a$. This is a contradiction to the assumption that a and b do not have a common lower bound.

One might suspect that (a, b)-regularity and (a, b)-uniformity can be characterized by the same condition as (a, b)-permutability. This is not the case: We consider the implication algebra $\mathcal{A}$ with $A = \{1, a, b, c, d\}$ consisting of the two Boolean subalgebras $\{1, a, b, c\}$ with $c \leq a, b \leq 1$ and $\{1, d\}$.

One can check easily that $\theta = \{a, c\}^2 \cup \{1, b, d\}^2$ and $\phi = \{a, c\}^2 \cup \{1, b\}^2 \cup \{d\}^2$ are congruences of $\mathcal{A}$. We have $c = a \wedge b$, $[a]\theta = [a]\phi$ but $[b]\theta \neq [b]\phi$, thus $\mathcal{A}$ is not (a, b)-regular. Moreover, $|[a]\theta| = 2$ and $|[b]\theta| = 3$, hence $\mathcal{A}$ is not (a, b)-uniform.

Example 2 Let $\mathcal{A}$ denote the algebra (A, s_1, s_2) with $A = \{a, b, c, d\}$ and unary operations s_1, s_2 defined as follows:

	a	b	c	d
s_1	d	c	c	d
s_2	b	a	d	c

$\mathcal{A}$ has exactly 3 non-trivial congruences, namely

$$\begin{aligned} \theta &= \{a\}^2 \cup \{b\}^2 \cup \{c, d\}^2, \\ \phi &= \{a, d\}^2 \cup \{b, c\}^2 \text{ and} \\ \psi &= \{a, b\}^2 \cup \{c, d\}^2. \end{aligned}$$

It follows

$$\begin{aligned}\theta \circ \phi &= \theta \cup \phi \cup \{(c,a),(d,b)\},\\ \phi \circ \theta &= \theta \cup \phi \cup \{(a,c),(b,d)\},\\ \theta \circ \psi &= \psi \circ \theta = \psi,\\ \phi \circ \psi &= \psi \circ \phi = A^2.\end{aligned}$$

$\mathcal{A}$ is (c,d)-permutable: For $f := s_1 \circ s_2$ it holds $f(c) = d$ and $f(d) = c$. Since $(b,d) \in (\phi \circ \theta) \setminus (\theta \circ \phi)$, $\mathcal{A}$ is not (b,d)-permutable.

$\mathcal{A}$ is (a,b)-regular: For $f = g := s_2$ it holds $f(b) = a$ and $g(f(x)) = x$ for all $x \in A$. Since $[a]\theta = [a]\omega$ (where ω denotes the least congruence on $\mathcal{A}$) and $[d]\theta \neq [d]\omega$, $\mathcal{A}$ is not (a,d)-regular.

$\mathcal{A}$ is (a,b)-uniform: In fact, for $f = g := s_2$ it holds $f(b) = a$ and $f(g(x)) = g(f(x)) = x$ for all $x \in A$. Since $|[a]\theta| \neq |[d]\theta|$, $\mathcal{A}$ is not (a,d)-uniform.

Corollary 1 *(i) If there exists $f \in C_{32}$ with $f(x,x,y) = f(y,x,x) = y$ for all $x,y \in A$ then $\mathcal{A}$ is permutable.*

(ii) If there exist $f,g \in C_{32}$ with $f(x,x,y) = y$ and $g(x,f(x,y,z),z) = y$ for all $x,y,z \in A$ then $\mathcal{A}$ is regular.

(iii) If there exist $f,g \in C_{32}$ with $f(x,x,y) = y$ and $f(x,g(x,y,z),z) = g(x,f(x,y,z),z) = y$ for all $x,y,z \in A$ then $\mathcal{A}$ is uniform.

Proof (i) Put $f_{cd}(x) := f(c,x,d)$ for all $c,d,x \in A$. Then $f_{cd} \in C_1$, $f_{cd}(c) = d$ and $f_{cd}(d) = c$ for all $c,d \in A$. According to Theorem 1, $\mathcal{A}$ is (c,d)-permutable for all $c,d \in A$ and hence permutable.

(ii) Put $f_{cd}(x) := f(d,x,c)$ and $g_{cd}(x) := g(d,x,c)$ for all $c,d,x \in A$. Then $f_{cd}, g_{cd} \in C_1$, $f_{cd}(d) = c$ and $g_{cd}(f_{cd}(x)) = g(d,f(d,x,c),c) = x$ for all $c,d,x \in A$. Hence $\mathcal{A}$ is (c,d)-regular for all $c,d \in A$ according to Theorem 1 and therefore regular.

(iii) With the same notation as in the proof of (ii) we now have $f_{cd}(d) = c$, $f_{cd}(g_{cd}(x)) = f(d,g(d,x,c),c) = x$ and $g_{cd}(f_{cd}(x)) = g(d,s(d,x,c),c) = x$ for all $c,d,x \in A$. By Theorem 1 $\mathcal{A}$ is (c,d)-uniform for all $c,d \in A$ and hence uniform. □

Example 3 Let $\mathcal{L} = (L,\vee,\wedge,',0,1)$ be an orthomodular lattice. For $x,y \in L$ we define

$$x + y := (x \vee (y \wedge x')) \wedge (x' \vee y').$$

Then it can be proved with standard methods:

$$x + 0 = 0 + x = x, \quad x + x = 0, \quad (x+y) + y = x.$$

Let $f(x,y,z) := (x+y)+z$, then we have $f(x,x,y) = (x+x)+y = 0+y = y$ and $f(y,x,x) = (y+x)+x = y$. Therefore $\mathcal{L}$ is permutable according to Corollary 1.

Now let $f(x,y,z) := (y+x)+z$ and $g(x,y,z) := (y+z)+x$. Then we have for all $x,y,z \in L$:

$$\begin{aligned}&f(x,x,y) = (x+x)+y = 0+y = y,\\ &f(x,g(x,y,z),z) = (((y+z)+x)+x)+z = (y+z)+z = y,\\ &g(x,f(x,y,z),z) = (((y+x)+z)+z)+x = (y+x)+x = y.\end{aligned}$$

By Corollary 1 $\mathcal{L}$ is both regular and uniform.

In the following let 0 be a fixed element of A. Recall that $\mathcal{A}$ is called

- permutable at 0 (cf. [2], [4], [6]) if $[0](\theta\circ\phi) = [0](\phi\circ\theta)$ for all $\theta, \phi \in \mathrm{Con}(\mathcal{A})$,
- weakly regular, (cf. [4], [5], [7]) if $\theta, \phi \in \mathrm{Con}(\mathcal{A})$ and $[0]\theta = [0]\phi$ imply $\theta = \phi$,
- locally regular (cf. [3], [4]) if $a \in A$, $\theta, \phi \in \mathrm{Con}(\mathcal{A})$ and $[a]\theta = [a]\phi$ imply $[0]\theta = [0]\phi$.

Corollary 2 *(i) If there exists $f \in C_{22}$ with $f(x, 0) = x$ and $f(x, x) = 0$ for all $x \in A$ then $\mathcal{A}$ is permutable at* 0.

(ii) If there exist $f, g \in C_{22}$ with $f(x, x) = 0$ and $g(x, f(x, y)) = y$ for all $x, y \in A$ then $\mathcal{A}$ is weakly regular.

(iii) If there exist $f, g \in C_{22}$ with $f(x, 0) = x$ and $g(x, f(x, y)) = y$ for all $x, y \in A$ then $\mathcal{A}$ is locally regular.

Proof It is easy to see that $\mathcal{A}$ is permutable at 0 if and only if $\mathcal{A}$ is $(c, 0)$-permutable for all $c \in A$, that $\mathcal{A}$ is weakly regular if and only if $\mathcal{A}$ is $(0, c)$-regular for all $c \in A$ and that $\mathcal{A}$ is locally regular if and only if $\mathcal{A}$ is $(c, 0)$-regular for all $c \in A$. Applying Theorem 1 to $f_c(x) := f(c, x)$ and $g_c(x) := g(c, x)$ the assertions follow immediately. □

Definition 3 $\mathcal{A}$ is called (a, b)-*semiuniform* if $|[a]\theta| \leq |[b]\theta|$ for all $\theta \in \mathrm{Con}(\mathcal{A})$. $\mathcal{A}$ is called 0-*semiuniform* if $\mathcal{A}$ is $(c, 0)$-semiuniform for all $c \in A$.

Theorem 2 *(i) If there exist $f, g \in C_1$ with $f(a) = b$ and $g(f(x)) = x$ for all $x \in A$ then $\mathcal{A}$ is (a, b)-semiuniform.*

(ii) If there exist $f, g \in C_{22}$ with $f(x, x) = 0$ and $g(x, f(x, y)) = y$ for all $x, y \in A$ then $\mathcal{A}$ is 0-semiuniform.

Proof (i) Let $\theta \in \mathrm{Con}(\mathcal{A})$. If $c \in [a]\theta$ then $f(c) \in [f(a)]\theta = [b]\theta$. If $d, e \in [a]\theta$ and $f(d) = f(e)$ then $d = g(f(d)) = g(f(e)) = e$. Hence $f|_{[a]\theta}$ is an injective mapping from $[a]\theta$ to $[b]\theta$ proving $|[a]\theta| \leq |[b]\theta|$.

(ii) Put $f_c(x) := f(c, x)$ and $g_c(x) := g(c, x)$ for all $c, x \in A$. Then $f_c, g_c \in C_1$, $f_c(c) = 0$ and $g_c(f_c(x)) = x$ for all $c, x \in A$. According to (i) $\mathcal{A}$ is $(c, 0)$-uniform for all $c \in A$, i.e. $\mathcal{A}$ is 0-semiuniform. □

Example 4 Every finite relatively pseudocomplemented lattice $\mathcal{L} = (L, \vee, \wedge, *, 0, 1)$ is 1-semiuniform: Let $\theta \in \mathrm{Con}(\mathcal{L})$. Since L is finite the class $[c]\theta$ contains the greatest element $\bar{c}$. Consider the function $\varphi_c(x) := \bar{c} * x$. For $x \in [c]\theta$ we have $\bar{c} * x\theta\bar{c} * \bar{c} = 1$, i.e. $\varphi_c(x) \in [1]\theta$. Suppose $x, y \in [c]\theta$ and $\varphi_c(x) = \varphi_c(y)$. Then

$$x = \bar{c} \wedge (\bar{c} * x) = \bar{c} \wedge \varphi_c(x) = \bar{c} \wedge \varphi_c(y) = \bar{c} \wedge (\bar{c} * y) = y.$$

This shows that φ_c is an injection from $[c]\theta$ into $[1]\theta$, i.e. $\mathcal{L}$ is $(c, 1)$-semiuniform for all $c \in L$.

Example 5 Every finite implication algebra $\mathcal{A} = (A, \cdot)$ is 1-semiuniform: Let $\theta \in \mathrm{Con}(\mathcal{A})$ and $c \in A$. Since A is finite, the class $[c]\theta$ has a greatest element $\bar{c}$. We consider $\varphi_c(x) := \bar{c}x$. Then for $x \in [c]\theta$ we have

$$(\bar{c}x)\theta\bar{c}\bar{c} = 1,$$

hence $\varphi_c(x) \in [1]\theta$. Suppose $\varphi_c(x) = \varphi_c(y)$ for $x, y \in [c]\theta$. We prove $\bar{c}x \wedge \bar{c} = x$: Since $x \in [c]\theta$ we have $x \leq \bar{c}$ and $x(\bar{c}x) = \bar{c}(xx) = 1$ implies $x \leq \bar{c}x$. Now suppose $z \leq \bar{c}x$ and $z \leq \bar{c}$, i.e. $z(\bar{c}x) = 1$ and $z\bar{c} = 1$. We have to show that $z \leq x$:

$$\begin{aligned} zx &= (z(\bar{c}x))(zx) = (\bar{c}(zx))(zx) = ((zx)\bar{c})\bar{c} = ((zx)((z\bar{c})\bar{c}))\bar{c} \\ &= ((zx)((\bar{c}z)z))\bar{c} = ((\bar{c}z)((zx)z))\bar{c} = ((\bar{c}z)z)\bar{c} = ((z\bar{c})\bar{c})\bar{c} = \bar{c}\bar{c} = 1. \end{aligned}$$

This proves $\bar{c}x \wedge \bar{c} = x$ and analogously we obtain $\bar{c}y \wedge \bar{c} = y$, thus we infer

$$x = (\bar{c}x) \wedge \bar{c} = (\bar{c}y) \wedge \bar{c} = y.$$

Consequently φ_c is an injection of $[c]\theta$ into $[1]\theta$, whence $|[c]\theta| \leq |[1]\theta|$. Thus $\mathcal{A}$ is 1-semiuniform.

Definition 4 Let $n > 1$. $\mathcal{A}$ is called *n-(a,b)-permutable* if $(a,b) \in \theta \circ \phi \circ \theta \circ \ldots$ (n factors) is equivalent to $(a,b) \in \phi \circ \theta \circ \phi \circ \ldots$ (n factors) for all $\theta, \phi \in \mathrm{Con}(\mathcal{A})$.

Theorem 3 *(i) If there exist functions $f_1 \in C_{31} \cap C_{33}$ and $f_2 \in C_{32} \cap C_{33}$ satisfying*

$$f_1(a, x, x) = a, \;\; f_1(x, x, b) = f_2(x, b, b), \;\; f_2(x, x, b) = b,$$

for all $x \in A$ then $\mathcal{A}$ is 3-(a,b)-permutable.

(ii) If there exists $f \in C_4$ satisfying

$$f(x, x, x, a) = a, \;\; f(x, x, x, b) = b, \;\; f(x, x, b, b) = f(b, x, b, x)$$

for all $x \in A$ then $\mathcal{A}$ is 3-(a,b)-permutable.

Proof (i) Let $\theta, \phi \in \mathrm{Con}(\mathcal{A})$ and $(a,b) \in \theta \circ \phi \circ \theta$. Then there are elements $c, d \in A$ with $a\theta c\phi d\theta b$. We infer

$$a = f_1(a, c, c)\phi f_1(a, c, d)\theta f_1(c, c, b) = f_2(c, b, b)\theta f_2(c, d, b)\phi f_2(c, c, b) = b,$$

whence $(a,b) \in \phi \circ \theta \circ \phi$.

(ii) Put $f_1(x, y, z) := f(z, y, z, x)$ and $f_2(x, y, z) := f(x, x, y, z)$. Then f_1, f_2 satisfy the conditions in (i). □

Definition 5 $\mathcal{A}$ is called *n-modular* (for $n \geq 2$) if for every $\theta, \phi, \psi \in \mathrm{Con}(\mathcal{A})$ with $\theta \subseteq \psi$ we have

$$(\underbrace{\theta \circ \phi \circ \theta \circ \ldots}_{n \text{ factors}}) \cap \psi \subseteq \theta \vee (\phi \cap \psi).$$

We remark that congruence modularity is equivalent to the condition $\theta \subseteq \psi$ implies $(\theta \vee \phi) \cap \psi \subseteq \theta \vee (\phi \cap \psi)$. Thus our concept of n-modularity is weaker than congruence modularity. Obviously $(n+1)$-modularity implies n-modularity.

Theorem 4 *Every algebra $\mathcal{A}$ is 3-modular (and hence 2-modular).*

Proof Suppose $\theta, \phi, \psi \in \mathrm{Con}(\mathcal{A})$ with $\theta \subseteq \psi$ and $(c, d) \in (\theta \circ \phi \circ \theta) \cap \psi$. Then there exist $e, f \in A$ with $c\theta e\phi f\theta d$ and we obtain $e\psi c\psi d\psi f$ and hence $c\theta e(\phi \cap \psi)f\theta d$. □

Example 6 Let $\mathcal{A} = (A, s_1, s_2, s_3)$ be an algebra with 3 unary operations and $A = \{a, b, \ldots, g\}$ with

	a	b	c	d	e	f	g
s_1	c	d	e	e	e	e	d
s_2	e	e	e	f	g	g	f
s_3	d	c	b	a	a	b	c

Then $\mathrm{Con}(\mathcal{A}) \cong \mathrm{N}_5$ since $\mathrm{Con}(\mathcal{A})$ consists of the trivial congruences and

$$\begin{aligned} \theta &= \{a, b\}^2 \cup \{c, d\}^2 \cup \{e, f\}^2 \cup \{g\}^2, \\ \phi &= \{a\}^2 \cup \{b, c\}^2 \cup \{d, e\}^2 \cup \{f, g\}^2, \\ \psi &= \{a, b, g\}^2 \cup \{c, d\}^2 \cup \{e, f\}^2, \end{aligned}$$

with $\theta \subseteq \psi$. Hence $\mathrm{Con}(\mathcal{A})$ is not modular.

However, $\mathrm{Con}(\mathcal{A})$ is 4-modular: The only non-trivial case to be checked refers to the triple (θ, ϕ, ψ) and we have

$$(\theta \circ \phi \circ \theta \circ \phi) \cap \psi = \theta \subseteq \theta \vee (\phi \cap \psi).$$

We remark that $\theta \circ \phi \circ \theta \circ \phi$ is not a congruence since $(a, e) \in \theta \circ \phi \circ \theta \circ \phi$ while $(e, a) \notin \theta \circ \phi \circ \theta \circ \phi$.

Definition 6 $\mathcal{A}$ is called (a, b)-*modular* if for all $\theta, \phi, \psi \in \mathrm{Con}(\mathcal{A})$ with $\theta \subseteq \psi$ we have $(a, b) \in (\theta \vee \phi) \cap \psi$ implies $(a, b) \in \theta \vee (\phi \cap \psi)$.

Remark 2 Of course, if for all $\theta, \phi \in \mathrm{Con}(\mathcal{A})$ it is true that

$$(a, b) \in \theta \vee \phi \text{ implies } (a, b) \in \theta \circ \phi \circ \theta \tag{1}$$

or

$$(a, b) \in \theta \vee \phi \text{ implies } (a, b) \in \theta \circ \phi \tag{2}$$

then, by Theorem 4, $\mathcal{A}$ is (a, b)-modular. Hence it is a natural to search for algebras satisfying the implications (1) or (2). Obviously (2) implies (a, b)-permutability and (1) implies 3-(a, b)-permutability. We are going to find sufficient conditions for (1) or (2).

Proposition 1 *Let R be a reflexive and compatible relation on $\mathcal{A}$.*

(i) If there exists an R-compatible unary function $f : A \to A$ such that $f(a) = b$ and $f(b) = a$ then $(a, b) \in R$ implies $(a, b) \in R^{-1}$.

(ii) If there exist a function $f : A^3 \to A$ compatible with R with respect to the first and third component such that $f(a, x, x) = a$ and $f(x, x, b) = b$ for all $x \in A$ then $(a, b) \in R \circ R$ implies $(a, b) \in R$.

Proof (i) If $(a, b) \in R$ then $(b, a) = (f(a), f(b)) \in R$ due to the compatibility of f with R.

(ii) Let $(a, b) \in R \circ R$. Then $aRcRb$ for some $c \in A$ and thus $a = f(a, c, c)$ R $f(c, c, b) = b$. □

For a binary relation R on A put $[a]R = \{x \in A \mid xRa\}$.

Definition 7 $\mathcal{A}$ is *n-permutable at a* $(n > 1)$ if for all $\theta, \phi \in \mathrm{Con}(\mathcal{A})$

$$[a](\theta \circ \phi \circ \ldots) = [a](\phi \circ \theta \circ \ldots)$$

(with n factors on both sides).

Theorem 5 *Let $\mathcal{A}$ be n-permutable at a. Then for all $\theta, \phi \in \mathrm{Con}(\mathcal{A})$ we have $(a, c) \in \theta \vee \phi$ if and only if $(a, c) \in \theta \circ \phi \circ \ldots$ (n factors).*

Proof Evidently, $(a, c) \in \theta \circ \phi \circ \ldots$ implies $(a, c) \in \theta \vee \phi$. Now, let $(a, c) \in \theta \vee \phi$. Then there exists an integer m such that $(a, b) \in \theta \circ \phi \circ \ldots$ (m factors). If $m \leq n$ we are done. We proof the assertion for $m = n+1$ and n even, the general proof works with the same idea. There exists an element $d \in A$ such that

$$a(\underbrace{\theta \circ \phi \circ \ldots \circ \phi}_{n \text{ factors}})d\theta c.$$

Hence $d \in [a](\phi \circ \theta \circ \ldots \circ \theta)$ (n factors). Due to n-permutability at a we have $d \in [a](\theta \circ \phi \circ \ldots \circ \phi)$ (n factors), i.e.

$$a(\underbrace{\phi \circ \theta \circ \ldots \circ \theta}_{n \text{ factors}})d\theta c,$$

hence

$$a(\underbrace{\phi \circ \theta \circ \ldots \circ \theta}_{n \text{ factors}})c,$$

and again by n-permutability at a we arrive at $(a, c) \in \theta \circ \phi \circ \ldots \circ \phi$ (n factors). □

Corollary 3 *If $\mathcal{A}$ is 3-permutable at a then $\mathcal{A}$ is (a, c)-modular for all $c \in A$.*

Proof Let $\theta, \phi, \psi \in \mathrm{Con}(\mathcal{A})$ with $\theta \subseteq \psi$ and $(a, c) \in (\theta \vee \phi) \cap \psi$. Then due to 3-permutability at a by Theorem 5 we have $(a, c) \in \theta \circ \phi \circ \theta$, i.e. there are $d, e \in A$ with $a\theta d\phi e\theta c$. Consequently we obtain $d\psi a\psi c\psi e$ and $a\theta d(\phi \cap \psi)e\theta c$. Thus $(a, c) \in \theta \vee (\phi \cap \psi)$ and we are done. □

References

[1] Abbott, J. C.: *Semi-boolean algebra.* Mat. Vesnik **4** (1967), 177–198.

[2] Chajda, I.: *A localization of some congruence conditions in varieties with nullary operations.* Annales Univ. Sci. Budapest **30** (1987), 17–23.

[3] Chajda, I.: *Locally regular varieties.* Acta Sci. Math. (Szeged) **64** (1998), 431–435.

[4] Chajda, I., Eigenthaler, G., Länger, H.: Congruence Classes in Universal Algebra. *Research and Exposition in Mathematics, Vol. 26, Heldermann Verlag, Lemgo*, 2003.

[5] Csákány, B.: *Characterizations of regular varieties.* Acta Sci. Math. (Szeged) **31** (1970), 187–189.

[6] Duda, J.: *Arithmeticity at 0.* Czechoslovak Math. J. **37** (1987), 197–206.

[7] Fichtner, K.: *Eine Bemerkung über Mannigfaltigkeiten universeller Algebren mit Idealen.* Monatsber. DAW **12** (1970), 21–25.

[8] Kalmbach, G.: Orthomodular Lattices. *Academic Press, London–New York*, 1983.

[9] Maltsev, A. I.: *On the general theory of algebraic systems.* Mat. Sbornik **35** (1954), 3–20 (in Russian).

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 75–86

Estimation of Dispersion in Nonlinear Regression Models with Constraints *

Lubomír KUBÁČEK [1], Eva TESAŘÍKOVÁ [2]

[1] *Department of Mathematical Analysis and Applications of Mathematics*
Faculty of Science, Palacký University
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kubacekl@aix.upol.cz

[2] *Department of Algebra and Geometry, Faculty of Science, Palacký University*
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: tesariko@aix.upol.cz

(Received April 14, 2004)

Abstract

Dispersion of measurement results is an important parameter that enables us not only to characterize not only accuracy of measurement but enables us also to construct confidence regions and to test statistical hypotheses. In nonlinear regression model the estimator of dispersion is influenced by a curvature of the manifold of the mean value of the observation vector. The aim of the paper is to find the way how to determine a tolerable level of this curvature.

Key words: Nonlinear regression model, linearization, estimation of dispersion.

2000 Mathematics Subject Classification: 62J05, 62F10

1 Introduction

The frequently used model in regression analysis is $\mathbf{Y} \sim N_n(\mathbf{f}(\boldsymbol{\beta}), \sigma^2\mathbf{V})$, $\boldsymbol{\beta} \in R^k$ (k-dimesional Euclidean space), where $\mathbf{Y}$ is an n-dimensional normally distributed observation vector, $\mathbf{f}(\boldsymbol{\beta})$ is its mean value, $\boldsymbol{\beta}$ is an unknown k-dimensional parameter, σ^2 is an unknown scalar parameter, $\sigma^2 \in (0, \infty)$, and $\mathbf{V}$ is a known $n \times n$ positive semidefinite matrix.

*Supported by the Council of Czech Government J14/98: 153 100011.

Sometimes the parameter $\boldsymbol{\beta}$ must satisfy a constraint $\mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}$.

The following text is devoted to the problem of determining of a tolerable level of a model curvature

$$\mathbf{Y} \sim N_n(\mathbf{f}(\boldsymbol{\beta}), \sigma^2\mathbf{V}), \quad \mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}, \tag{1}$$

how this curvature can be defined and how to use this measure of nonlinearity to a determination of a linearization region. This region will be defined as a set of such shifts of the parameter $\boldsymbol{\beta}$ arround the chosen value $\boldsymbol{\beta}_0$ which does not cause any essential deterioration of a quality of the estimator of σ^2 in the case that the actual value $\boldsymbol{\beta}^*$ of the parameter $\boldsymbol{\beta}$ is equal to $\boldsymbol{\beta}_0$.

2 Notation and auxiliary statement

Let in the model (1) the function $\mathbf{f}(\cdot)$ and $\mathbf{g}(\cdot)$ can be approximated as

$$\mathbf{f}(\boldsymbol{\beta}) = \mathbf{f}_0 + \mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) \quad \text{and} \quad \mathbf{g}(\boldsymbol{\beta}) = \mathbf{G}\delta\boldsymbol{\beta} + \frac{1}{2}\gamma(\delta\boldsymbol{\beta}),$$

where

$$\begin{aligned}
\mathbf{f}_0 &= \mathbf{f}(\boldsymbol{\beta}_0), \quad \mathbf{F} = \partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}'|_{u=\beta_0}, \quad \delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}_0,\\
\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) &= \Big(\kappa_1(\delta\boldsymbol{\beta}), \dots, \kappa_n(\delta\boldsymbol{\beta})\Big)',\\
\kappa_i(\delta\boldsymbol{\beta}) &= \delta\boldsymbol{\beta}'\mathbf{F}_i\delta\boldsymbol{\beta}, \quad i = 1, \dots, n,\\
\mathbf{F}_i &= \partial^2 f_i(\mathbf{u})/\partial\mathbf{u}\partial\mathbf{u}'|_{u=\beta_0}, \quad i = 1, \dots, n,\\
\mathbf{G} &= \partial\mathbf{g}(\mathbf{u})/\partial\mathbf{u}'|_{u=\beta_0},\\
\boldsymbol{\gamma}(\delta\boldsymbol{\beta}) &= \Big(\gamma_1(\delta\boldsymbol{\beta}), \dots, \gamma_q(\delta\boldsymbol{\beta})\Big)',\\
\gamma_i(\delta\boldsymbol{\beta}) &= \delta\boldsymbol{\beta}'\mathbf{G}_i\delta\boldsymbol{\beta}, \quad i = 1, \dots, q,\\
\mathbf{G}_i &= \partial^2 g_i(\mathbf{u})/\partial\mathbf{u}\partial\mathbf{u}'|_{u=\beta_0}, \quad i = 1, \dots, q.
\end{aligned}$$

The model

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n(\mathbf{F}\delta\boldsymbol{\beta}, \sigma^2\mathbf{V}), \quad \mathbf{G}\delta\boldsymbol{\beta} = \mathbf{0} \tag{2}$$

is a linearized version of the model (1) and

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n\Big(\mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}), \sigma^2\mathbf{V}\Big), \quad \mathbf{G}\delta\boldsymbol{\beta} + \frac{1}{2}\gamma(\delta\boldsymbol{\beta}) = \mathbf{0} \tag{3}$$

is a quadratic version of the model (1).

Assumption In the following text it is assumed that it is valid

$$r(\mathbf{F}_{n,k}) = k < n \quad \text{and} \quad r(\mathbf{G}_{q,k}) = q < k,$$

respectively, for the ranks of the matrices $\mathbf{F}$ and $\mathbf{G}$, respectively, and that the matrix $\mathbf{V}$ is positive definite.

Lemma 2.1 *The best quadratic estimator* $\hat{\sigma}^2$ *of the parameter* σ^2 *in the model (2) is*

$$\hat{\sigma}^2 = (\mathbf{Y} - \mathbf{f}_0)'\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^+(\mathbf{Y} - \mathbf{f}_0)/(n+q-k) \tag{4}$$

and $\hat{\sigma}^2 \sim \sigma^2\chi^2_{n+q-k}(0)/(n+q-k)$.

Here $\mathbf{M}_{G'} = \mathbf{I} - \mathbf{G}'(\mathbf{G}\mathbf{G}')^-\mathbf{G}, (\mathbf{G}'\mathbf{G})^-$ *is any g-inverse of the matrix* $\mathbf{G}\mathbf{G}'$*, the symbol*

$$\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^+$$

means the Moore–Penrose g-inverse of the matrix $\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}$ *(cf. [5]) and* $\chi^2_{n+q-k}(0)$ *is the random variable with the central chi-square distribution with* $n+q-k$ *degrees of freedom.*

Proof Cf. e.g. in [3].

3 Measure of nonlinearity

Lemma 3.1 *The estimator (4) in the model (3) is of the property*

$$(\mathbf{Y} - \mathbf{f}_0)'\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^+(\mathbf{Y} - \mathbf{f}_0)/(n+q-k) \sim \sigma^2\frac{\chi^2_{n+q-k}(\delta)}{n+q-k},$$

where $\delta = \frac{1}{4\sigma^2}\boxed{1}'\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^+\boxed{1}$,

$$\boxed{1} = \boldsymbol{\kappa}(\delta\boldsymbol{\beta}) - \mathbf{F}\mathbf{G}'(\mathbf{G}\mathbf{G}')^-\boldsymbol{\gamma}(\delta\boldsymbol{\beta}).$$

Proof It is sufficient to prove the equality

$$E(\mathbf{Y} - \mathbf{f}_0) = \mathbf{F}\mathbf{M}_{G'}\delta\boldsymbol{\beta} + \frac{1}{2}\left[\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) - \mathbf{F}\mathbf{G}'(\mathbf{G}\mathbf{G}')^-\boldsymbol{\gamma}(\delta\boldsymbol{\beta})\right].$$

Since

$$E(\mathbf{Y} - \mathbf{f}_0) = \mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = \mathbf{F}\mathbf{M}_{G'}\delta\boldsymbol{\beta} - \mathbf{F}\mathbf{G}'(\mathbf{G}\mathbf{G}')^-\mathbf{G}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) =$$
$$= \mathbf{F}\mathbf{M}_{G'}\delta\boldsymbol{\beta} + \frac{1}{2}\left[\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) - \mathbf{F}\mathbf{G}'(\mathbf{G}\mathbf{G}')^-\boldsymbol{\gamma}(\delta\boldsymbol{\beta})\right],$$

the statement is proved. □

Corollary 3.2 *Since* $E[\chi^2_f(\delta)] = f + \delta$*, the estimator (4) is biased and*

$$E(\hat{\sigma}^2) - \sigma^2 = \frac{1}{4(n+q-k)}\boxed{1}'\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^+\boxed{1}.$$

Now an analogy of the intrinsic curvature of the Bates and Watts [1] can be defined.

Definition 3.3 The quantity

$$K_{0,I}^{int}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boxed{1}'\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^{+}\boxed{1}}}{\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s}} : \delta\mathbf{s}\in R^{k-q}\right\},$$

where

$$\begin{aligned}\boxed{1} &= \boldsymbol{\kappa}(\mathbf{K}_G\delta\mathbf{s}) - \mathbf{F}\mathbf{G}^{-}\boldsymbol{\gamma}(\mathbf{K}_G\delta\mathbf{s}),\\ \mathcal{M}(\mathbf{K}_G) &= \mathcal{M}(\mathbf{M}_{G'}), \quad \mathbf{K}_G \text{ is } k\times(k-q) \text{ matrix},\\ \mathbf{C}_0 &= \mathbf{F}'\mathbf{V}^{-1}\mathbf{F},\end{aligned}$$

is intrinsic curvature at the point $\boldsymbol{\beta}_0$ for the model with constraints $\mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}$.

Remark 3.4 The Bates and Watts [1] intrinsic curvature for a regular model without constraints $\mathbf{Y}\sim N_n(\mathbf{f}(\boldsymbol{\beta}),\boldsymbol{\Sigma}), \boldsymbol{\beta}\in R^k$, is defined as

$$K^{int}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\left(\mathbf{M}_F^{\Sigma^{-1}}\right)'\boldsymbol{\Sigma}^{-1}\mathbf{M}_F^{\Sigma^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta}\in R^k\right\},$$

where $\mathbf{M}_F^{\Sigma^{-1}} = \mathbf{I} - \mathbf{F}(\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}$.

The model (3) can be reparametrized in the following way.

$$\begin{aligned}&\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \mathbf{K}_G\delta\mathbf{s} - \frac{1}{2}\mathbf{G}^{-}\boldsymbol{\gamma}(\mathbf{K}_G\delta\mathbf{s}) + \text{ terms of the higher order},\\ &\mathbf{Y}-\mathbf{f}_0 \sim N_n\Big(\mathbf{F}\mathbf{K}_G\delta\mathbf{s} - \frac{1}{2}\mathbf{F}\mathbf{G}^{-}\boldsymbol{\gamma}(\mathbf{K}_G\delta\mathbf{s}) + \frac{1}{2}\boldsymbol{\kappa}(\mathbf{K}_G\delta\mathbf{s}), \sigma^2\mathbf{V}\Big).\end{aligned}$$

Now, if the scheme

$$\begin{aligned}&\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) \to \boldsymbol{\kappa}(\mathbf{K}_G\delta\mathbf{s}) - \mathbf{F}\mathbf{G}^{-}\boldsymbol{\gamma}(\mathbf{K}_G\delta\mathbf{s}), \quad \mathbf{M}_F \to \mathbf{M}_{FM_{G'}},\\ &\left(\mathbf{M}_F^{\Sigma^{-1}}\right)'\boldsymbol{\Sigma}^{-1}\mathbf{M}_F^{\Sigma^{-1}} = \left(\mathbf{M}_F\boldsymbol{\Sigma}\mathbf{M}_F\right)^{+} \to (\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{F_{G'}}\Big)^{+}\end{aligned}$$

and the relationship

$$\delta\mathbf{s}'\mathbf{K}_G'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\mathbf{K}_G\delta\mathbf{s} = \delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s},$$

is taken into account, the expression for $K_{0,I}^{int}(\boldsymbol{\beta}_0)$ is obtained and its geometrical meaning can be seen.

Remark 3.5 If the model is linear, i.e. $\mathbf{Y}\sim N_n(\mathbf{F}\boldsymbol{\beta},\sigma^2\mathbf{V})$, however the constraints $\mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}$ are nonlinear, then $K_{0,I}^{int}(\boldsymbol{\beta}_0)$ is equal to

$$K_{0,I}^{int}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boldsymbol{\gamma}'(\mathbf{K}_G\delta\mathbf{s})\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^{+}\boldsymbol{\gamma}(\mathbf{K}_G\delta\mathbf{s})}}{\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s}} : \delta\mathbf{s}\in R^{k-q}\right\}.$$

The curvature of the manifold $\{\boldsymbol{\beta} : \mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}\}$ at the point $\boldsymbol{\beta}_0$ can be characterized as follows.

The parameter $\delta\boldsymbol{\beta}$ can be expressed as

$$\delta\boldsymbol{\beta} = \mathbf{K}_G\delta\mathbf{s} - \frac{1}{2}\mathbf{G}^{-}\boldsymbol{\gamma}(\mathbf{K}_G\delta\mathbf{s}) + \dots$$

The natural norm in the parametric space R^k can be assumed as

$$\|\delta\boldsymbol{\beta}\| = \sqrt{\delta\boldsymbol{\beta}\mathbf{F}'(\sigma^2\mathbf{V})^{-1}\mathbf{F}\delta\boldsymbol{\beta}},$$

since it is the Mahalanobis norm introduced by the estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0).$$

Thus the quantity $\sigma C_0^{constr}(\boldsymbol{\beta}_0)$, where

$$C_0^{constr}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boxed{2}'\mathbf{C}_0\boxed{2}}}{\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s}} : \delta\mathbf{s} \in R^{k-q}\right\},$$

$$\boxed{2} = \mathbf{M}_{M_{G'}}^{C_0}\mathbf{G}'(\mathbf{G}\mathbf{G}')^{-1}\boldsymbol{\gamma}(\mathbf{K}_G\delta\mathbf{s}),$$

can be considered as the intrinsic curvature of the constraints $\mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}$. However

$$(\mathbf{G}\mathbf{G}')^{-1}\mathbf{G}\left(\mathbf{M}_{M_{G'}}^{C_0}\right)'\mathbf{C}_0\mathbf{M}_{M_{G'}}^{C_0}\mathbf{G}'(\mathbf{G}\mathbf{G}')^{-1} =$$

$$= (\mathbf{G}\mathbf{G}')^{-1}\mathbf{G}\Big\{\mathbf{I} - \Big[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}_0^{-1}\Big]\mathbf{C}_0\Big\}'$$

$$\times\, \mathbf{C}_0\Big\{\mathbf{I} - \Big[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}_0^{-1}\Big]\mathbf{C}_0\Big\}\mathbf{G}'(\mathbf{G}\mathbf{G}')^{-1} = (\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}$$

and

$$(\mathbf{G}\mathbf{G}')^{-1}\mathbf{G}\mathbf{F}'\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^{+}\mathbf{F}\mathbf{G}'(\mathbf{G}\mathbf{G}')^{-1} =$$

$$= (\mathbf{G}\mathbf{G}')^{-1}\mathbf{G}\mathbf{F}'\Big\{\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{F}\mathbf{M}_{G'}\Big[\mathbf{M}_{G'}\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\mathbf{M}_{G'}\Big]^{+} \times \mathbf{M}_{G'}\mathbf{F}'\mathbf{V}^{-1}\Big\}\mathbf{F}\mathbf{G}'(\mathbf{G}\mathbf{G}')^{-1}$$

$$= (\mathbf{G}\mathbf{G}')^{-1}\mathbf{G}\Big\{\mathbf{C}_0 - \mathbf{C}_0\Big[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}_0^{-1}\Big]\mathbf{C}_0\Big\}\mathbf{G}'(\mathbf{G}\mathbf{G}')^{-1}$$

$$= (\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}.$$

Thus under the condition $\boldsymbol{\kappa}(\cdot) = \mathbf{0}$,

$$K_{0,I}^{int}(\boldsymbol{\beta}_0) = C_0^{constr}(\boldsymbol{\beta}_0).$$

Remark 3.6 If the model is nonlinear, i.e. $\mathbf{Y} \sim N_n(\mathbf{f}(\boldsymbol{\beta}), \sigma^2\mathbf{V})$, however the constraints are linear, i.e. $\mathbf{G}\delta\boldsymbol{\beta} = \mathbf{0}$, then $K_{0,I}^{int}(\boldsymbol{\beta}_0)$ is equal to

$$K_{0,I}^{int}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\mathbf{K}_G\delta\mathbf{s})\left(\mathbf{M}_{FM_{G'}}^{V^{-1}}\right)'\mathbf{V}^{-1}\mathbf{M}_{FM_{G'}}^{V^{-1}}\boldsymbol{\kappa}(\mathbf{K}_G\delta\mathbf{s})}}{\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s}} : \delta\mathbf{s} \in R^{k-q}\right\}.$$

Since

$$\left(\mathbf{M}_{FM_{G'}}^{V^{-1}}\right)'\mathbf{V}^{-1}\mathbf{M}_{FM_{G'}}^{V^{-1}} =$$

$$= \left(\mathbf{M}_{F}^{V^{-1}}\right)'\mathbf{V}^{-1}\mathbf{M}_{F}^{V^{-1}} + \mathbf{V}^{-1}\mathbf{F}\mathbf{C}_0\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1},$$

it can be written

$$K_{0,I}^{int}(\boldsymbol{\beta}_0) \leq K_0^{int}(\boldsymbol{\beta}_0),$$

in the case $\boldsymbol{\gamma}(\delta\boldsymbol{\beta}) = \mathbf{0}$, with respect to Remark 3.4. Here $K_0^{int}(\boldsymbol{\beta}_0) = K^{int}(\boldsymbol{\beta}_0)$ for $\sigma = 1$.

4 Linearization region

Definition 4.1 The ε-linearization region (at the point $\boldsymbol{\beta}_0$) for an estimation of the parameter σ^2 is

$$\mathcal{L}_\sigma = \left\{\boldsymbol{\beta}_0 + \delta\boldsymbol{\beta} : \delta\boldsymbol{\beta} = \mathbf{K}_G\delta\mathbf{s}, E(\hat{\sigma}^2) - \sigma^2 < \varepsilon^2\sigma^2\right\}.$$

Theorem 4.2 *The ε-linearization region from Definition 4.1 is*

$$\mathcal{L}_\sigma = \left\{\boldsymbol{\beta}_0 + \delta\boldsymbol{\beta} : \delta\boldsymbol{\beta} = \mathbf{K}_G\delta\mathbf{s}, \delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s} \leq \frac{2\sigma\varepsilon\sqrt{n+q-k}}{K_{0,I}^{int}(\boldsymbol{\beta}_0)}\right\}.$$

Proof The relationships

$$E(\hat{\sigma}^2) - \sigma^2 = \frac{1}{4(n+q-k)}\boxed{1}'\left(\mathbf{M}_{FM_{G'}}\mathbf{V}\mathbf{M}_{FM_{G'}}\right)^+\boxed{1} \leq$$

$$\leq \frac{1}{4(n+q-k)}\left(\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s}\right)^2\left(K_{0,I}^{int}(\boldsymbol{\beta}_0)\right)^2$$

are implied by a comparison of the bias from Corollary 3.2 and Definition 3.3. Thus

$$E(\hat{\sigma}^2) - \sigma^2 \leq \frac{1}{4(n+q-k)}\left(\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s}\right)^2\left(K_{0,I}^{int}(\boldsymbol{\beta}_0)\right)^2 \leq \sigma^2\varepsilon^2$$

$$\Leftrightarrow \delta\mathbf{s}\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s} \leq \frac{2\sigma\varepsilon\sqrt{n+q-k}}{K_{0,I}^{int}(\boldsymbol{\beta}_0)}. \qquad \square$$

Remark 4.3 The actual value $\boldsymbol{\beta}^*$ of the parameter $\boldsymbol{\beta}$ is unknown. However some information on $\boldsymbol{\beta}^*$ is given by the estimator

$$\hat{\hat{\boldsymbol{\beta}}} = \boldsymbol{\beta}_0 + \left\{\mathbf{I} - \mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\right)\hat{\boldsymbol{\beta}},$$

where $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}_0 + (\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0)$ and by the confidence region

$$\mathcal{E}_\beta = \left\{\boldsymbol{\beta}_0 + \mathbf{K}_G\mathbf{u} : (\mathbf{u} - \delta\hat{\hat{\boldsymbol{\beta}}})'\mathbf{C}_0(\mathbf{u} - \delta\hat{\hat{\boldsymbol{\beta}}}) \leq (k-q)\hat{\sigma}^2 F_{k-q,n+q-k}(1-\alpha)\right\}.$$

(The equalities

$$\mathrm{Var}(\hat{\hat{\boldsymbol{\beta}}}) = \sigma^2[\mathbf{C}_0 - \mathbf{C}_0\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]$$

and

$$[\mathbf{C}_0 - \mathbf{C}_0\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]^+ = \mathbf{C}_0$$

are utilized.)

Remark 4.4 With respect to Theorem 4.2 and the expression for the $(1-\alpha)$-confidence ellipsoid, it is clear that the values of the semiaxes of the ellipsoid depend on σ linearly, however the semiaxes of $\mathcal{L}_\sigma$ depend linearly on $\sqrt{\sigma}$. Thus the inclusion $\mathcal{E}_\beta \subset \mathcal{L}_\sigma$ can be attain by a smaller σ. It can be established by a proper design of experiment.

Remark 4.5 If $\mathcal{E}_\beta$ is significantly smaller than $\mathcal{L}_\sigma$ and $\mathcal{E}_\beta \subset \mathcal{L}_\sigma$, we can estimate parameter σ^2 by (4) and we can be sure that $E(\hat{\sigma}^2) - \sigma^2 < \varepsilon^2\sigma^2$.

Let $b(\hat{\sigma}^2) = E(\hat{\sigma}^2) - \sigma^2$ and $b(\hat{\sigma}) = E(\hat{\sigma}) - \sigma$. Then the approximation

$$b(\hat{\sigma}) \approx \sigma\frac{b(\hat{\sigma}^2)}{2} \leq \sigma\frac{\varepsilon^2}{2}$$

can be used. Thus, from the viewpoint of practice it seems to be important the validity of the following implication

$$\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s} \leq \frac{2\sigma\varepsilon\sqrt{n+q-k}}{K_{0,I}^{int}(\boldsymbol{\beta}_0)} \Rightarrow b(\hat{\sigma}) \leq \sigma\frac{\varepsilon^2}{2}\,.$$

5 Numerical example

In [4] the problem of linearization of the model with constraints with respect to the estimation of the parameter $\boldsymbol{\beta}$ was solved. The numerical example given there was chosen as follows.

$$\{\mathbf{f}\}_i(\boldsymbol{\beta}) = f_i(\boldsymbol{\beta}) = \begin{cases} l_1(x_i,\beta_1) = x_i\beta_1, & x_i \leq 5,\\ l_2(x_i,\beta_2,\beta_3) = \beta_1\exp(\beta_3 x_i), & x_i \geq 5 \end{cases}$$

and

$$g(\beta_1,\beta_2,\beta_3) = 5\beta_1 - \beta_2\exp(5\beta_3).$$

Measurement regarding this model was calculated at the points $x = 1,2,3,6,7,8$ and $\mathrm{Var}(\mathbf{Y}) = \sigma^2\mathbf{I}$. In [4] it is shown that for $\sigma = 0.5$ the model cannot be linearized with respect to the estimation of $\boldsymbol{\beta}$. The value of the parameter σ must be smaller than 0.01 in order for the linearization be admissible.

Quite different situation occurs in this example in the case that the estimator of σ^2 is under consideration. With the help of [7] we obtain the following results. Analogously as in [4], let the functions $\mathbf{f}(\cdot)$ and $g(\cdot)$ be those given at the beginning of the section, $x = 1, 2, 3, 6, 7, 8$, $\mathrm{Var}(\mathbf{Y}) = \sigma^2 \mathbf{I}$ and

$$\beta_1 = 1.473, \beta_2 = 33, \beta_3 = -0.29999, \alpha = 0.05, \varepsilon = 0.1, \sigma = 0.5.$$

Then the figures 1, 2 and 3 show that the $(1-\alpha)$-confidence ellipsoid is included into $\mathcal{L}_\sigma$ and the same is valid also for $\sigma = 1$; cf. figures 4,5,6.

$\sigma = 0.5$

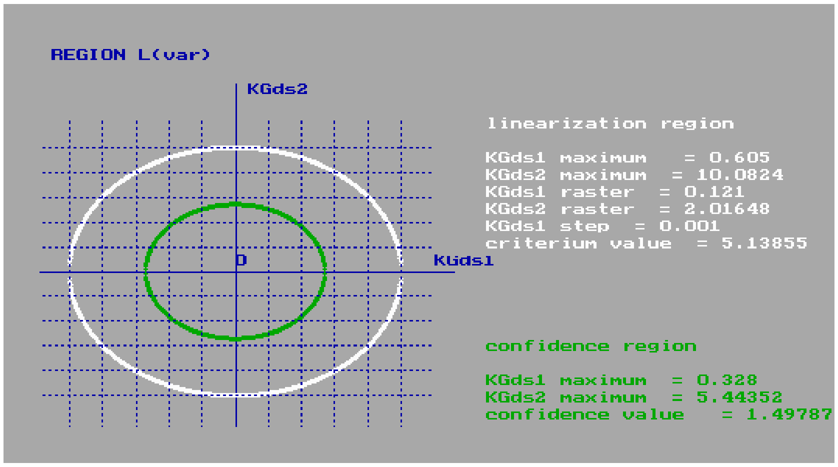

Figure 1 The sections of the confidence ellipsoid and the linearization region $\mathcal{L}_\sigma$ by the axes β_1 and β_2.

$\sigma = 0.5$

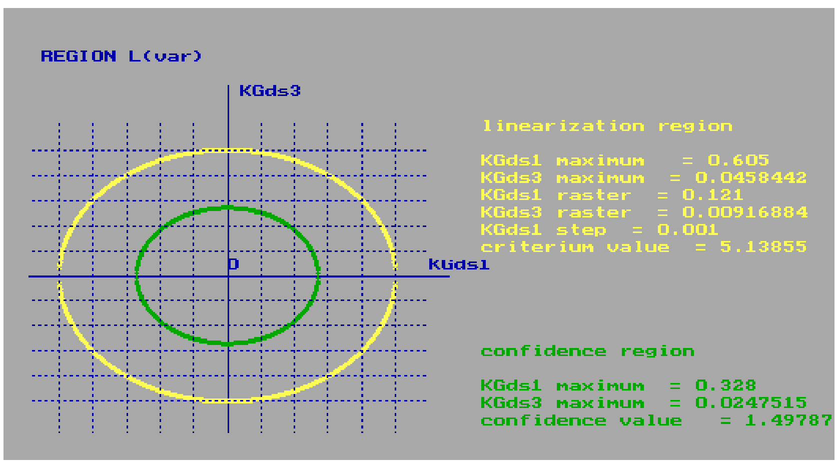

Figure 2 The sections of the confidence ellipsoid and the linearization region $\mathcal{L}_\sigma$ by the axes β_1 and β_3.

$\sigma = 0.5$

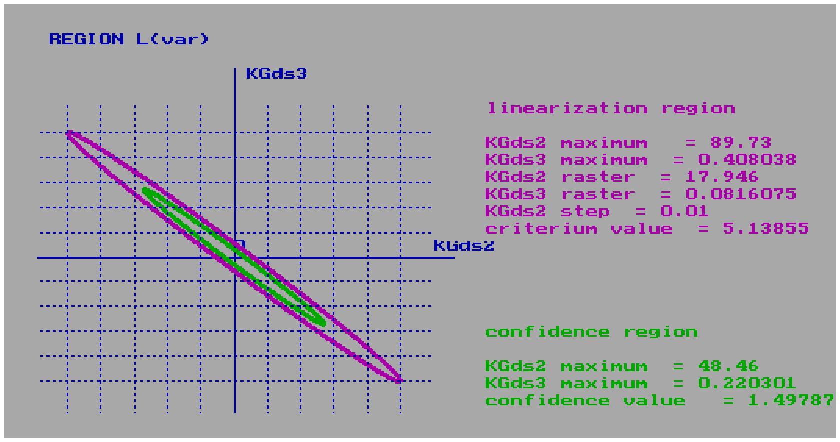

Figure 3 The sections of the confidence ellipsoid and the linearization region $\mathcal{L}_\sigma$ by the axes β_2 and β_3.

$\sigma = 1$

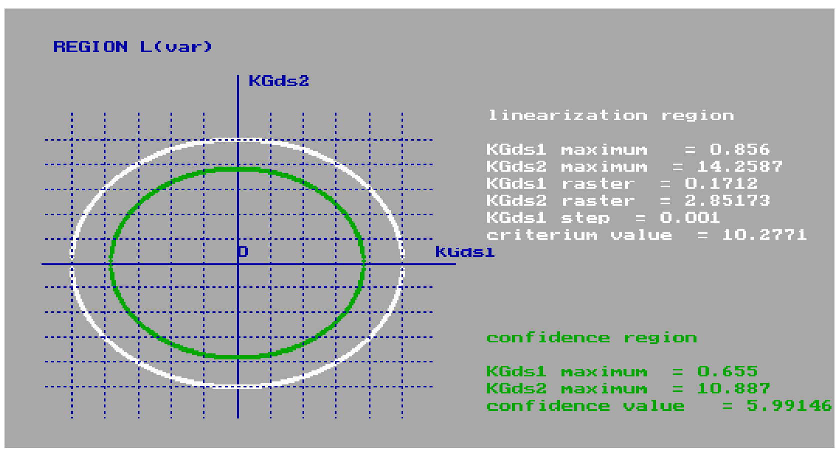

Figure 4 The sections of the confidence ellipsoid and the linearization region $\mathcal{L}_\sigma$ by the axes β_1 and β_2.

$\sigma = 1$

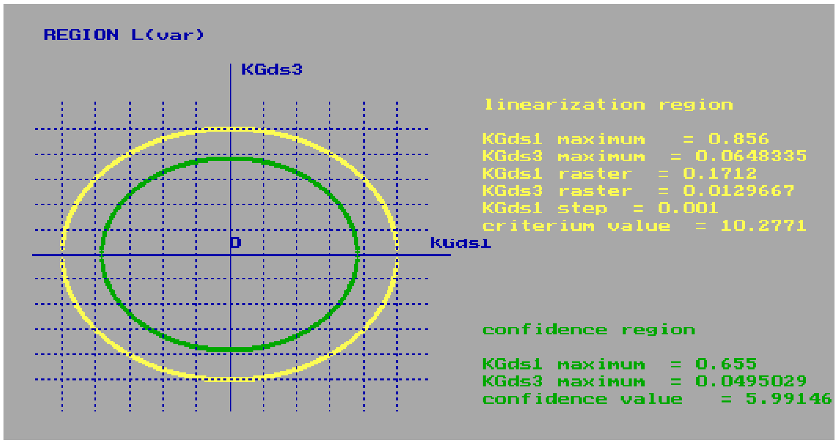

Figure 5 The sections of the confidence ellipsoid and the linearization region $\mathcal{L}_\sigma$ by the axes β_1 and β_3.

$\sigma = 1$

Figure 6 The sections of the confidence ellipsoid and the linearization region $\mathcal{L}_\sigma$ by the axes β_2 and β_3.

The empirical probabilty density function is given at figure 7 for $\sigma = 0.5$

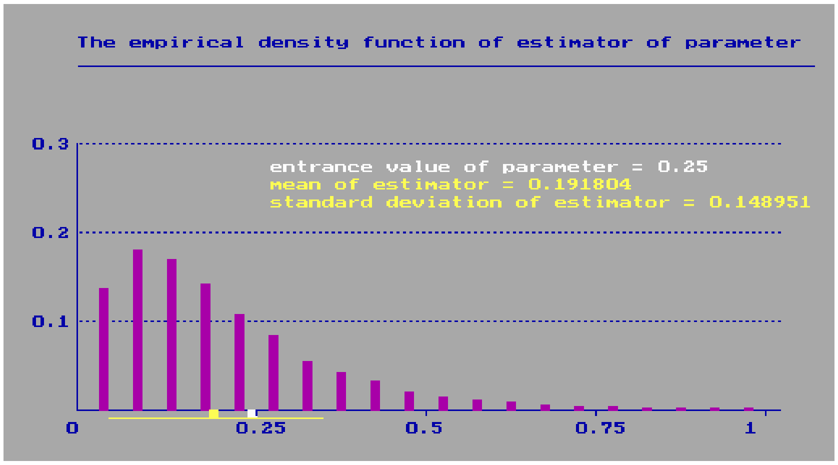

Figure 7 The empirical density function of the estimator $\hat{\sigma}^2$ (4) for $\sigma = 0.5$

The linearization is possible if the value of $K_{0,I}^{int}(\boldsymbol{\beta})$ is sufficiently small with respect to the quantile $F_{k-q,n+q-k}(1-\alpha)$ (cf. Remark 4.3). Therefore table 1 gives the different values of the parameter $\boldsymbol{\beta}$ for our example and table 2 gives the corresponding values $K_{0,I}^{int}(\boldsymbol{\beta})$; the values signed by the star are too large for the linearization of the model with respect to estimation of σ^2 if $\sigma = 0.5$.

β_1	$\beta_2(\beta_3 = -1)$	$\beta_2(\beta_3 = -0.5)$	$\beta_2(\beta_3 = 0.5)$	$\beta_2(\beta_3 = 1)$
0.5	371.032 898	30.456 235	0.205 212	0.016 845
1.0	742.065 796	60.912 470	0.410 425	0.033 690
1.5	1 113.098 693	91.368 705	0.615 637	0.050 535
2.0	1 484.131 591	121.824 940	0.820 850	0.067 379
2.5	1 855.164 488	152.281 174	1.026 062	0.084 224

Table 1 The values of the parameter $\boldsymbol{\beta}$ for Table 2

β_1	$\beta_3 = -1$	$\beta_3 = -0.5$	$\beta_3 = 0.5$	$\beta_3 = 1$
0.5	0.172199*	0.138345*	0.049 621	0.022 430
1.0	0.086 301	0.069 146	0.024 779	0.011 198
1.5	0.056 943	0.045 983	0.016 533	0.007 457
2.0	0.043 192	0.034 565	0.012 372	0.005 568
2.5	0.034 546	0.027 661	0.009 908	0.004 437

Table 2 The values of $K_{0,I}^{int}(\boldsymbol{\beta})$ for $\boldsymbol{\beta}$ given in Table 1

References

[1] Bates, D. M., Watts, D. G.: *Relative curvatures measures of nonlinearity.* J. Roy. Statist. Soc. **B 42** (1980), 1–25.

[2] Kubáček, L., Kubáčková, L.: *Regression models with a weak nonlinearity.* Technical report Nr. 1998.1, Universität Stuttgart, 1998, 1–67.

[3] Kubáček, L., Kubáčková, L.: Statistics and Metrology. *Vyd. Univ. Palackého, Olomouc*, 2000 (in Czech).

[4] Kubáček, L., Tesaříková, E.: *Confidence regions in nonlinear models with constraints.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **42** (2003), 43–58.

[5] Rao, C. R., Mitra, S. K.: Generalized Inverse of Matrices and its Applications. *J. Wiley & Sons, New York–London–Sydney–Toronto*, 1971.

[6] Scheffé, H.: The Analysis of Variance. *J. Wiley, New York*, 1959.

[7] Tesaříková, E., Kubáček, L.: *Estimators of dispersion in models with constraints (demo-program).* Department of Algebra and Geometry, Faculty of Science, Palacký University, Olomouc, 2003.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 87–104

On Eliminating Transformations for Nuisance Parameters in Multivariate Linear Model *

Pavla KUNDEROVÁ

Department of Mathematical Analysis and Applications of Mathematics,
Faculty of Science, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kunderov@inf.upol.cz

(Received January 27, 2004)

Abstract

The multivariate linear model, in which the matrix of the first order parameters is divided into two matrices: to the matrix of the useful parameters and to the matrix of the nuisance parameters, is considered. We examine eliminating transformations which eliminate the nuisance parameters without loss of information on the useful parameters and on the variance components.

Key words: Multivariate linear regression model, useful and nuisance parameters, LBLUE, eliminating transformation.

2000 Mathematics Subject Classification: 62J05

*Supported by the Council of Czech Government J14/98: 153 100011.

1 Notations, auxiliary statements

The following notations will be used throughout the paper:

R^n	the space of all n-dimensional real vectors;
$\boldsymbol{u}_p$	the real column p-dimensional vector;
$\boldsymbol{A}_{m,n}$, $Tr(\boldsymbol{A})$	the real $m \times n$ matrix, the trace of the matrix $\boldsymbol{A}$;
$\boldsymbol{A}', r(\boldsymbol{A})$	the transpose, the rank of the matrix $\boldsymbol{A}$;
$\boldsymbol{A}^{(j)}$	j-th column of the matrix $\boldsymbol{A}$;
$vec(\boldsymbol{A})$	the column vector $((\boldsymbol{A}^{(1)})', \ldots, (\boldsymbol{A}^{(n)})')'$;
$\boldsymbol{A} \otimes \boldsymbol{B}$	the Kronecker (tensor) product of the matrices $\boldsymbol{A}$,$\boldsymbol{B}$;
$\mathscr{M}(\boldsymbol{A})$	the range of the matrix $\boldsymbol{A}$;
$\boldsymbol{A}^-$	a generalized inverse of the matrix $\boldsymbol{A}$ (satisfying $\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A}$);
$\boldsymbol{A}^+$	the Moore-Penrose generalized inverse of the matrix $\boldsymbol{A}$ (satisfying $\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A} = \boldsymbol{A}, \boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^+, (\boldsymbol{A}\boldsymbol{A}^+)' = \boldsymbol{A}\boldsymbol{A}^+$, $(\boldsymbol{A}^+\boldsymbol{A})' = \boldsymbol{A}^+\boldsymbol{A}$);
$\boldsymbol{P}_A$	the orthogonal projector onto $\mathscr{M}(\boldsymbol{A})$;
$\boldsymbol{M}_A = \boldsymbol{I} - \boldsymbol{P}_A$	the orthogonal projector onto $\mathscr{M}^{\perp}(\boldsymbol{A}) = Ker(\boldsymbol{A}')$;
$\boldsymbol{I}_k$	the $k \times k$ identity matrix;
$0_{m,n}$	the $m \times n$ null matrix;
$\boldsymbol{o}$	the null element.

If $\mathscr{M}(\boldsymbol{A}) \subset \mathscr{M}(\boldsymbol{V})$, $\boldsymbol{V}$ p.s.d., then the symbol $\boldsymbol{P}_A^V$ denotes the projector on the subspace $\mathscr{M}(\boldsymbol{A})$ in the $\boldsymbol{V}$-seminorm given by the matrix $\boldsymbol{V}$, $||\boldsymbol{x}||_V = \sqrt{\boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}}$; $\boldsymbol{M}_A^V = \boldsymbol{I} - \boldsymbol{P}_A^V = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{V}\boldsymbol{A})^-\boldsymbol{A}'\boldsymbol{V}$. Let $\boldsymbol{N}_{n,n}$ is p.d. (p.s.d.) matrix and $\boldsymbol{A}_{m,n}$ an arbitrary matrix, then the symbol $\boldsymbol{A}^-_{m(N)}$ denotes the matrix satisfying $\boldsymbol{A}\boldsymbol{A}^-_{m(N)}\boldsymbol{A} = \boldsymbol{A}$ and $\boldsymbol{N}\boldsymbol{A}^-_{m(N)}\boldsymbol{A} = [\boldsymbol{N}\boldsymbol{A}^-_{m(N)}\boldsymbol{A}]'$. ($\boldsymbol{A}^-_{m(N)}\boldsymbol{y}$ is a solution of the consistent system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ whose N-seminorm is minimal, see [4], p. 151). $\boldsymbol{A}^-_{m(N)}$ is called a minimum N-seminorm g-inverse of the matrix $\boldsymbol{A}$. It holds

$$\mathscr{M}(\boldsymbol{A}') \subset \mathscr{M}(\boldsymbol{N}) \quad \Rightarrow \quad \boldsymbol{A}^-_{m(N)}\boldsymbol{N}^-\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{N}^-\boldsymbol{A}')^-.$$

Assertion 1 (see [3], Lemma 16)

$$(\boldsymbol{M}_S\Sigma\boldsymbol{M}_S)^+ = \Sigma^{-1} - \Sigma^{-1}\boldsymbol{S}(\boldsymbol{S}'\Sigma^{-1}\boldsymbol{S})^-\boldsymbol{S}'\Sigma^{-1} = \Sigma^{-1}\boldsymbol{M}_S^{\Sigma^{-1}},\ \textit{if } \Sigma \textit{ is p.d.},$$

$$(\boldsymbol{M}_S\Sigma\boldsymbol{M}_S)^+ = \Sigma^+ - \Sigma^+\boldsymbol{S}(\boldsymbol{S}'\Sigma^-\boldsymbol{S})^-\boldsymbol{S}'\Sigma^+,\ \textit{if } \Sigma \textit{ is p.s.d. and } \mathscr{M}(\boldsymbol{S}) \subset \mathscr{M}(\Sigma).$$

Assertion 2 *If Σ is p.d. matrix, $\boldsymbol{W}$ p.s.d. and $\boldsymbol{S}$ such matrices, that*

$$\mathscr{M}(\boldsymbol{S}') = \mathscr{M}(\boldsymbol{S}'\boldsymbol{W}\boldsymbol{S}),$$

then (see [6], Lemma 1)

$$(\boldsymbol{M}_S^W)'[\boldsymbol{M}_S^W\Sigma(\boldsymbol{M}_S^W)']^+\boldsymbol{M}_S^W = (\boldsymbol{M}_S\Sigma\boldsymbol{M}_S)^+.$$

2 Multivariate linear model with nuisance parameters

Let

$$\boldsymbol{Y}_{n,m} = \boldsymbol{X}_{n,k}\boldsymbol{B}_{k,l}\boldsymbol{Z}_{l,m} + \varepsilon_{n,m} \tag{1}$$

be a multivariate linear model under consideration. Here $\boldsymbol{Y}$ is an observation matrix, $\boldsymbol{X}$, $\boldsymbol{Z}$, are known nonzero matrices, ε is a random matrix and $\boldsymbol{B}$ is a matrix of unknown parameters

$$\boldsymbol{B} = (\boldsymbol{B}_1, \boldsymbol{B}_2),$$

where $\boldsymbol{B}_1$ is a $k \times r$ matrix of useful parameters which (or their functions) has to be estimated from the observation matrix $\boldsymbol{Y}$ and $\boldsymbol{B}_2$ is a $k \times s$ matrix of nuisance parameters. Thus we consider the model

$$\boldsymbol{Y} = \boldsymbol{X}(\boldsymbol{B}_1, \boldsymbol{B}_2)\begin{pmatrix} \boldsymbol{Z}_1 \\ \boldsymbol{Z}_2 \end{pmatrix} + \varepsilon. \tag{2}$$

Lemma 1 *The model (2) can be equivalently written in the form*

$$vec(\boldsymbol{Y}) = [\boldsymbol{Z}_1' \otimes \boldsymbol{X}, \boldsymbol{Z}_2' \otimes \boldsymbol{X}]\begin{pmatrix} vec(\boldsymbol{B}_1) \\ vec(\boldsymbol{B}_2) \end{pmatrix} + vec(\varepsilon). \tag{3}$$

where a $r \times m$ matrix $\boldsymbol{Z}_1$ and a $s \times m$ matrix $\boldsymbol{Z}_2$ are known nonzero matrices.

Proof is obvious by virtue of the following statement

$$vec(\boldsymbol{ABC}) = (\boldsymbol{C}' \otimes \boldsymbol{A})vec(\boldsymbol{B}), \tag{4}$$

valid for all matrices of corresponding types. □

Suppose that

1. the observation vector $vec(\boldsymbol{Y})$ has the mean value

$$E[vec(\boldsymbol{Y})] = [\boldsymbol{Z}_1' \otimes \boldsymbol{X}, \boldsymbol{Z}_2' \otimes \boldsymbol{X}]\begin{pmatrix} vec(\boldsymbol{B}_1) \\ vec(\boldsymbol{B}_2) \end{pmatrix},$$

and the covariance matrix

$$var[vec(\boldsymbol{Y})] = \Sigma_\vartheta \otimes \boldsymbol{I}_n,$$

where $m \times m$ matrix Σ_ϑ (the covariance matrix of any column of the matrix $\boldsymbol{Y}$) is such a matrix that

2. $\Sigma_\vartheta = \sum_{i=1}^{p} \vartheta_i V_i$, $\forall \vartheta = (\vartheta_1, \ldots, \vartheta_p)' \in \underline{\vartheta} \subset R^p$, $V_1, \ldots, V_p$ given symmetric matrices,
3. $\underline{\vartheta} \subset R^p$ contains an open sphere in R^p,
4. if $\vartheta \in \underline{\vartheta}$, the matrix Σ_ϑ is positive definite,
5. the matrix Σ_ϑ is not a function of the matrix $\boldsymbol{B} = (\boldsymbol{B}_1, \boldsymbol{B}_2)$,

6. suppose that

$$\mathscr{M}(\boldsymbol{Z}_1' \otimes \boldsymbol{X}, \boldsymbol{Z}_2' \otimes \boldsymbol{X}) \subset \mathscr{M}(\Sigma_\vartheta \otimes \boldsymbol{I}); \tag{5}$$

this condition is warranted by

$$\mathscr{M}(\boldsymbol{Z}_1) \subset \mathscr{M}(\Sigma_\vartheta) \quad \wedge \quad \mathscr{M}(\boldsymbol{Z}_2) \subset \mathscr{M}(\Sigma_\vartheta); \tag{6}$$

and it means that

$$vec(\boldsymbol{Y}) \in \mathscr{M}(\Sigma_\vartheta \otimes \boldsymbol{I}) \ (a.s.).$$

Remark 1 A parametric function $\boldsymbol{p}'vec(\boldsymbol{B}_1), \boldsymbol{p} \in R^{kr}$, is said to be unbiasedly estimable under the model (2) if there exists an estimator $\boldsymbol{L}'vec(\boldsymbol{Y})$, $\boldsymbol{L} \in R^{mn}$, such that $E[\boldsymbol{L}'vec(\boldsymbol{Y})] = \boldsymbol{p}'vec(\boldsymbol{B}_1)$, $\forall vec(\boldsymbol{B}_1), \forall vec(\boldsymbol{B}_2)$.

The equality

$$E[\boldsymbol{L}'vec(\boldsymbol{Y})] = \boldsymbol{L}'(\boldsymbol{Z}_1' \otimes \boldsymbol{X})vec(\boldsymbol{B}_1) + \boldsymbol{L}'(\boldsymbol{Z}_2' \otimes \boldsymbol{X})vec(\boldsymbol{B}_2) = \boldsymbol{p}'vec(\boldsymbol{B}_1),$$

$\forall vec(\boldsymbol{B}_1), \forall vec(\boldsymbol{B}_2)$, is fulfiled if and only if

$$\boldsymbol{p} = (\boldsymbol{Z}_1 \otimes \boldsymbol{X}')\boldsymbol{L} \quad \& \quad (\boldsymbol{Z}_2 \otimes \boldsymbol{X}')\boldsymbol{L} = \boldsymbol{o},$$

that is equivalent to

$$\boldsymbol{p} = (\boldsymbol{Z}_1 \otimes \boldsymbol{X}')\boldsymbol{M}_{Z_2' \otimes X}\boldsymbol{u}, \quad \boldsymbol{u} \in R^{mn}.$$

Thus the class of all unbiasedly estimable linear functions $\boldsymbol{p}'vec(\boldsymbol{B}_1)$ of the useful parameters in the model (2) is given by

$$\mathscr{E}_1 = \{\boldsymbol{p}'vec(\boldsymbol{B}_1) : \boldsymbol{p} \ \in \ \mathscr{M}[(\boldsymbol{Z}_1 \otimes \boldsymbol{X}')\boldsymbol{M}_{Z_2' \otimes X}] = \mathscr{M}[\boldsymbol{Z}_1\boldsymbol{M}_{Z_2'} \otimes \boldsymbol{X}']\}. \tag{7}$$

Obviously the class of all unbiasedly estimable linear functions $\boldsymbol{q}'vec(\boldsymbol{B}_2)$ of the nuisance parameters in the model (2) is given by

$$\mathscr{E}_2 = \{\boldsymbol{q}'vec(\boldsymbol{B}_2) : \boldsymbol{q} \in \mathscr{M}[(\boldsymbol{Z}_2 \otimes \boldsymbol{X}')\boldsymbol{M}_{Z_1' \otimes X}] = \mathscr{M}[\boldsymbol{Z}_2\boldsymbol{M}_{Z_1'} \otimes \boldsymbol{X}'].$$

Notation 1 Denote $\widehat{vec(\boldsymbol{B}_1)}$ and $\widehat{vec(\boldsymbol{B}_2)}$ an $(\Sigma_\vartheta^{-1} \otimes \boldsymbol{I})$-LS estimator of the vector parameter $vec(\boldsymbol{B}_1)$ and $vec(\boldsymbol{B}_2)$ respectively computed under the linear model (2) (see [1], p. 161). According to the assumption (6) $\boldsymbol{p}'\widehat{vec(\boldsymbol{B}_1)}$, $\boldsymbol{p} \in \mathscr{M}[(\boldsymbol{Z}_1 \otimes \boldsymbol{X}')\boldsymbol{M}_{Z_2' \otimes X}]$, and $\boldsymbol{q}'\widehat{vec(\boldsymbol{B}_2)}$, $\boldsymbol{q} \in \mathscr{M}[(\boldsymbol{Z}_2 \otimes \boldsymbol{X}')\boldsymbol{M}_{Z_1' \otimes X}]$, are the BLUEs of the function $\boldsymbol{p}'vec(\boldsymbol{B}_1)$ and $\boldsymbol{q}'vec(\boldsymbol{B}_2)$ respectively (see [1], Theorem 5.3.2., p. 162).

Theorem 1

$$\begin{pmatrix} \widehat{vec(\boldsymbol{B}_1)} \\ \widehat{vec(\boldsymbol{B}_2)} \end{pmatrix} =$$

$$= \begin{pmatrix} (\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta\boldsymbol{M}_{Z_2'}]^+\boldsymbol{Z}_1')^-\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta\boldsymbol{M}_{Z_2'}]^+ \otimes (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' \\ (\boldsymbol{Z}_2\Sigma_\vartheta^-\boldsymbol{Z}_2')^-\boldsymbol{Z}_2\Sigma_\vartheta^-\boldsymbol{M}_{Z_1'}^{(M_{Z_2'}\Sigma_\vartheta M_{Z_2'})^+} \otimes (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' \end{pmatrix} vec(\boldsymbol{Y}).$$

Proof According to [1], Theorem 5.3.1 we have under the model (2)

$$\begin{pmatrix} \widehat{vec(\boldsymbol{B}_1)} \\ \widehat{vec(\boldsymbol{B}_2)} \end{pmatrix} =$$

$$= [(\boldsymbol{Z}_1' \otimes \boldsymbol{X}, \boldsymbol{Z}_2' \otimes \boldsymbol{X})'(\Sigma_\vartheta \otimes \boldsymbol{I})^-(\boldsymbol{Z}_1' \otimes \boldsymbol{X}, \boldsymbol{Z}_2' \otimes \boldsymbol{X})]^- \begin{pmatrix} \boldsymbol{Z}_1 \otimes \boldsymbol{X}' \\ \boldsymbol{Z}_2 \otimes \boldsymbol{X}' \end{pmatrix} (\Sigma_\vartheta \otimes \boldsymbol{I})^- vec(\boldsymbol{Y})$$

$$= \begin{bmatrix} \boldsymbol{Z}_1 \Sigma_\vartheta^- \boldsymbol{Z}_1' \otimes \boldsymbol{X}'\boldsymbol{X}, & \boldsymbol{Z}_1 \Sigma_\vartheta^- \boldsymbol{Z}_2' \otimes \boldsymbol{X}'\boldsymbol{X} \\ \boldsymbol{Z}_2 \Sigma_\vartheta^- \boldsymbol{Z}_1' \otimes \boldsymbol{X}'\boldsymbol{X}, & \boldsymbol{Z}_2 \Sigma_\vartheta^- \boldsymbol{Z}_2' \otimes \boldsymbol{X}'\boldsymbol{X} \end{bmatrix}^- \begin{pmatrix} \boldsymbol{Z}_1 \Sigma_\vartheta^{-1} \otimes \boldsymbol{X}' \\ \boldsymbol{Z}_2 \Sigma_\vartheta^{-1} \otimes \boldsymbol{X}' \end{pmatrix} vec(\boldsymbol{Y}). \quad (8)$$

Using the following Rohde's formula for generalized inverse of partitioned p.s.d. matrix (see [3], Lemma 13, p.68)

$$\begin{pmatrix} \boldsymbol{A}, & \boldsymbol{B} \\ \boldsymbol{B}', & \boldsymbol{C} \end{pmatrix}^- = \begin{pmatrix} \boldsymbol{A}^- + \boldsymbol{A}^-\boldsymbol{B}(\boldsymbol{C} - \boldsymbol{B}'\boldsymbol{A}^-\boldsymbol{B})^-\boldsymbol{B}'\boldsymbol{A}^-, & -\boldsymbol{A}^-\boldsymbol{B}(\boldsymbol{C} - \boldsymbol{B}'\boldsymbol{A}^-\boldsymbol{B})^- \\ -(\boldsymbol{C} - \boldsymbol{B}'\boldsymbol{A}^-\boldsymbol{B})^-\boldsymbol{B}'\boldsymbol{A}^-, & (\boldsymbol{C} - \boldsymbol{B}'\boldsymbol{A}^-\boldsymbol{B})^- \end{pmatrix}$$

$$= \begin{pmatrix} (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^-\boldsymbol{B}')^-, & -(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^-\boldsymbol{B}')^-\boldsymbol{B}\boldsymbol{C}^- \\ -\boldsymbol{C}^-\boldsymbol{B}'(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^-\boldsymbol{B}')^-, & \boldsymbol{C}^- + \boldsymbol{C}^-\boldsymbol{B}'(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^-\boldsymbol{B}')^-\boldsymbol{B}\boldsymbol{C}^- \end{pmatrix},$$

we get the blocks of the g-inverse matrix in (8):

$$\boldsymbol{A}_{11} = (\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+\boldsymbol{Z}_1')^- \otimes (\boldsymbol{X}'\boldsymbol{X})^-,$$

$$\boldsymbol{A}_{12} = -[(\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+\boldsymbol{Z}_1')^-\boldsymbol{Z}_1\Sigma_\vartheta^-\boldsymbol{Z}_2'(\boldsymbol{Z}_2\Sigma_\vartheta^{=}\boldsymbol{Z}_2)^-$$
$$\otimes(\boldsymbol{X}'\boldsymbol{X})^-(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{X})^-,$$

$$\boldsymbol{A}_{21} = (\boldsymbol{A}_{12})',$$

$$\boldsymbol{A}_{22} = [(\boldsymbol{Z}_2\Sigma_\vartheta^-\boldsymbol{Z}_2')^- \otimes (\boldsymbol{X}'\boldsymbol{X})^-]$$
$$+[(\boldsymbol{Z}_2\Sigma_\vartheta^-\boldsymbol{Z}_2')^-\boldsymbol{Z}_2\Sigma_\vartheta^-\boldsymbol{Z}_1'(\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+\boldsymbol{Z}_1')^-\boldsymbol{Z}_1\Sigma_\vartheta^-\boldsymbol{Z}_2'(\boldsymbol{Z}_2\Sigma_\vartheta^-\boldsymbol{Z}_2')^-$$
$$\otimes(\boldsymbol{X}'\boldsymbol{X})^-(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{X})^-].$$

After some calculations we get

$$\widehat{vec(\boldsymbol{B}_1)} = [(\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+\boldsymbol{Z}_1')^-\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+ \otimes (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}']vec(\boldsymbol{Y}).$$

$$\widehat{vec(\boldsymbol{B}_2)} = [(\boldsymbol{Z}_2\Sigma_\vartheta^-\boldsymbol{Z}_2')^-\boldsymbol{Z}_2\Sigma_\vartheta^-\boldsymbol{M}_{Z_1'}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+} \otimes (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}']vec(\boldsymbol{Y}).$$

The estimates obtained by substitution $\widehat{vec(\boldsymbol{B}_1)}$ into unbiasedly estimable functions $\boldsymbol{p}'vec(\boldsymbol{B}_1)$ are given uniquely. It can be proved if we take the following assertion (see [3], Lemma 8, p.65)

$$\boldsymbol{A}\boldsymbol{B}^-\boldsymbol{C} \text{ is invariant to the choice of the g-inverse } \boldsymbol{B}^-$$
$$\iff \mathscr{M}(\boldsymbol{A}') \subset \mathscr{M}(\boldsymbol{B}') \quad \& \quad \mathscr{M}(\boldsymbol{C}) \subset \mathscr{M}(\boldsymbol{B}), \quad (9)$$

into account. □

Theorem 2 *Let us denote* $\Sigma_0 = \sum_{i=1}^{p} \vartheta_{0,i} \boldsymbol{V}_i$.

a) In model (2) the function $\boldsymbol{g}'\vartheta = \sum_{i=1}^{p} \boldsymbol{g}_i \vartheta_i$, $\vartheta \in \underline{\vartheta}$, *is unbiasedly, quadratically and invariantly estimable (i.e. the estimator has the form* $[vec(\boldsymbol{Y})]'\boldsymbol{A}[vec(\boldsymbol{Y})]$, *where* $\boldsymbol{A}_{mn,mn}$ *is symmetric matrix, the estimator is invariant with respect to the change of the matrix* $\boldsymbol{B}$*) if and only if*

$$\boldsymbol{g} \in \mathscr{M}\left(\boldsymbol{S}_{(M_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes I)M_{(Z_1'\otimes X, Z_2'\otimes X)})^+}\right),$$

where

$$\{\boldsymbol{S}_{(M_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes I)M_{(Z_1'\otimes X, Z_2'\otimes X)})^+}\}_{i,j} =$$

$$= Tr[(\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)})^+(\boldsymbol{V}_i\otimes \boldsymbol{I})(\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes \boldsymbol{I})$$

$$\times \boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)})^+(\boldsymbol{V}_j\otimes \boldsymbol{I})],\ \ i,j = 1,\dots,p.$$

b) If the function $\boldsymbol{g}'\vartheta$ *satisfies the condition from a), then the* ϑ_0*-MINQUE of* $\boldsymbol{g}'\vartheta$ *is given as*

$$\widehat{\boldsymbol{g}'\vartheta} = \sum_{i=1}^{p} \lambda_i (vec(\boldsymbol{Y}))'[\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)}]^+(\boldsymbol{V}_i\otimes \boldsymbol{I})$$

$$\times [\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)}]^+ vec(\boldsymbol{Y}),$$

where the vector $\lambda = (\lambda_1,\dots,\lambda_p)'$ *is a solution of the system of equations*

$$\boldsymbol{S}_{(M_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes I)M_{(Z_1'\otimes X, Z_2'\otimes X)})^+}\lambda = \boldsymbol{g}.$$

Proof see [4], Theorem IV.1.11.

Remark 2 The matrix $S_{(M_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes I)M_{(Z_1'\otimes X, Z_2'\otimes X)})^+}$ is called the criterional matrix for the estimability of the function $\boldsymbol{g}'\vartheta$.
As $\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)} = \boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{M_{(Z_2'\otimes X)}(Z_1'\otimes X)} = \boldsymbol{M}_{M_{(Z_2'\otimes X)}(Z_1'\otimes X)}\boldsymbol{M}_{Z_2'\otimes X}$, it holds

$$\{\boldsymbol{S}_{(M_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_0\otimes I)M_{(Z_1'\otimes X, Z_2'\otimes X)})^+}\}_{i,j}$$

$$= Tr[(\boldsymbol{M}_{M_{(Z_2'\otimes X)}(Z_1'\otimes X)}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{M_{(Z_2'\otimes X)}(Z_1'\otimes X)})^+(\boldsymbol{V}_i\otimes \boldsymbol{I})$$

$$\times(\boldsymbol{M}_{M_{(Z_2'\otimes X)}(Z_1'\otimes X)}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{M_{(Z_2'\otimes X)}(Z_1'\otimes X)})^+(\boldsymbol{V}_j\otimes \boldsymbol{I})]$$

$$= Tr[(\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X})^+(\boldsymbol{V}_i\otimes \boldsymbol{I})$$

$$\times(\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X})^+(\boldsymbol{V}_j\otimes \boldsymbol{I})], i,j = 1,\dots p,$$

where the equality

$$[\boldsymbol{M}_{M_{(Z_2'\otimes X)}(Z_1'\otimes X)}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{M_{Z_2'\otimes X}(Z_1'\otimes X)}]^+$$

$$= [\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X}]^+,$$

was used.

3 Eliminating transformations

There are situation in the practice, that the number of nuisance parameters is much more greater than the number of useful parameters. This fact could cause difficulties in the course of calculations.

There exist two approaches to the problem of nuisance parameters. One of them is to eliminate the nuisance parameters by a transformation of the observation vector provided this transformation is not allowed to cause a loss of information of the useful parameters.

Our task is to eliminate in the model (2) the matrix $\boldsymbol{Z}_2' \otimes \boldsymbol{X}$, belonging to the vector $vec(\boldsymbol{B}_2)$ of nuisance parameters, i.e. we consider the following class of eliminating matrices

$$\mathscr{T} = \{\boldsymbol{T} : \boldsymbol{T}(\boldsymbol{Z}_2' \otimes \boldsymbol{X}) = 0\},$$

that leads us to linear models

$$[\ Tvec(\boldsymbol{Y}), \boldsymbol{T}(\boldsymbol{Z}_1' \otimes \boldsymbol{X})vec(\boldsymbol{B}_1), \boldsymbol{T}(\Sigma_\vartheta \otimes \boldsymbol{I})\boldsymbol{T}'\ ]. \tag{10}$$

The general solution of the matrix equation $\boldsymbol{T}(\boldsymbol{Z}_2' \otimes \boldsymbol{X}) = 0$ is of the form

$$\boldsymbol{T} = \boldsymbol{A}[\boldsymbol{I} - (\boldsymbol{Z}_2' \otimes \boldsymbol{X})(\boldsymbol{Z}_2' \otimes \boldsymbol{X})^-],$$

where $\boldsymbol{A}$ is an arbitrary matrix of the corresponding type, $(\boldsymbol{Z}_2' \otimes \boldsymbol{X})^-$ is some version of generalized inverse of the matrix $\boldsymbol{Z}_2' \otimes \boldsymbol{X}$.

If we choose $(\boldsymbol{Z}_2' \otimes \boldsymbol{X})^- = [(\boldsymbol{Z}_2' \otimes \boldsymbol{X})'\boldsymbol{W}(\boldsymbol{Z}_2' \otimes \boldsymbol{X})]^-(\boldsymbol{Z}_2' \otimes \boldsymbol{X})'\boldsymbol{W}$, where $\boldsymbol{W} = \boldsymbol{W}_1 \otimes \boldsymbol{W}_2$ is an arbitrary p.s.d. matrix such that

$$\mathscr{M}(\boldsymbol{Z}_2 \otimes \boldsymbol{X}') = \mathscr{M}[(\boldsymbol{Z}_2 \otimes \boldsymbol{X}')\boldsymbol{W}(\boldsymbol{Z}_2' \otimes \boldsymbol{X})], \tag{11}$$

then $\boldsymbol{T} = \boldsymbol{A}\boldsymbol{M}^W_{Z_2' \otimes X}$, where $\boldsymbol{M}^W_{Z_2' \otimes X}$ is given uniquely.

First we consider the transformation matrix $\boldsymbol{T} = \boldsymbol{M}^W_{Z_2' \otimes X}$, i.e. we consider linear model

$$[\boldsymbol{M}^W_{Z_2' \otimes X} vec(\boldsymbol{Y}), \boldsymbol{M}^W_{Z_2' \otimes X}(\boldsymbol{Z}_1' \otimes \boldsymbol{X})vec(\boldsymbol{B}_1), \boldsymbol{M}^W_{Z_2' \otimes X}(\Sigma_\vartheta \otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2' \otimes X})'],\ \Sigma_\vartheta \text{ p.d.} \tag{12}$$

Remark 3 As $\boldsymbol{M}^{W_1 \otimes W_2}_{Z_2' \otimes X} vec(\boldsymbol{Y}) = (\boldsymbol{I}_m \otimes \boldsymbol{I}_n)vec(\boldsymbol{Y}) - (\boldsymbol{P}^{W_1}_{Z_2'} \otimes \boldsymbol{P}^{W_2}_X)vec(\boldsymbol{Y})$, we can write $\boldsymbol{Y}^{transf} = \boldsymbol{Y} - \boldsymbol{P}^{W_2}_X \boldsymbol{Y}(\boldsymbol{P}^{W_1}_{Z_2'})'$.

Lemma 2 *Let $\boldsymbol{W}$ is p.s.d. matrix such that (11) is valid. Then*

$$\mathscr{M}(\boldsymbol{M}_{Z_2' \otimes X}) = \mathscr{M}([\boldsymbol{M}^W_{Z_2' \otimes X}]').$$

Proof see [7], Lemma 2. □

Thus

$$\mathscr{M}[(\boldsymbol{Z}_1 \otimes \boldsymbol{X}')\boldsymbol{M}_{Z_2' \otimes X}] = \mathscr{M}[(\boldsymbol{Z}_1 \otimes \boldsymbol{X}')(\boldsymbol{M}^W_{Z_2' \otimes X})'],$$

i.e. the classes of the estimable functions $\boldsymbol{p}'vec(\boldsymbol{B}_1)$ in the model (2) and in the model (12) are identical.

Theorem 3 *The ϑ-LBLUE of the estimable function $\boldsymbol{p}'vec(\boldsymbol{B}_1)$, where $\boldsymbol{p} \in \mathscr{M}[(\boldsymbol{Z}_1 \otimes \boldsymbol{X}')\boldsymbol{M}_{Z_2'\otimes X}]$ in the model (12) is given as*

$$\boldsymbol{p}'\widehat{vec(\boldsymbol{B}_1)} =$$
$$= \boldsymbol{p}'[(\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+\boldsymbol{Z}_1')^-\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+ \otimes (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}']vec(\boldsymbol{Y}),$$

i.e. it is the same as in the model (2), (see Theorem 1).

Proof According to [2], Theorem 3.1.3 the ϑ-LBLUE in the model (12) is given as

$$\boldsymbol{p}'\widehat{vec(\boldsymbol{B}_1)} =$$
$$= \boldsymbol{p}'\left\{\left[\left(\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{Z}_1'\otimes \boldsymbol{X})\right)'\right]^-_{m(M^W_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)(M^W_{Z_2'\otimes X})')}\right\}'\boldsymbol{M}^W_{Z_2'\otimes X}vec(\boldsymbol{Y})$$
$$= \boldsymbol{p}'\{[\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{Z}_1'\otimes \boldsymbol{X})]'[(\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_\vartheta\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})']^-\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{Z}_1'\otimes \boldsymbol{X})\}^-$$
$$\times [\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{Z}_1'\otimes \boldsymbol{X})]'[\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_\vartheta\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})']^-\boldsymbol{M}^W_{Z_2'\otimes X}vec(\boldsymbol{Y})$$
$$= \boldsymbol{p}'\{(\boldsymbol{Z}_1\otimes \boldsymbol{X}')(\boldsymbol{M}^W_{Z_2'\otimes X})'[\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_\vartheta\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})']^+\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{Z}_1'\otimes \boldsymbol{X})\}^-$$
$$\times (\boldsymbol{Z}_1\otimes \boldsymbol{X}')(\boldsymbol{M}^W_{Z_2'\otimes X})'[\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_\vartheta\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})']^+\boldsymbol{M}^W_{Z_2'\otimes X}vec(\boldsymbol{Y}).$$

Using Assertion 2 and Assertion 1 we get

$$\boldsymbol{p}'\widehat{vec(\boldsymbol{B}_1)} = \boldsymbol{p}'\{(\boldsymbol{Z}_1\otimes \boldsymbol{X}')[\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_\vartheta\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}]^+(\boldsymbol{Z}_1'\otimes \boldsymbol{X})\}^-$$
$$\times(\boldsymbol{Z}_1\otimes \boldsymbol{X}')[\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_\vartheta\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}]^+vec(\boldsymbol{Y})$$
$$= \boldsymbol{p}'\{(\boldsymbol{Z}_1\otimes \boldsymbol{X}')(\Sigma_\vartheta^{-1}\otimes \boldsymbol{I})(\boldsymbol{Z}_1'\otimes \boldsymbol{X}) - (\boldsymbol{Z}_1\otimes \boldsymbol{X}')(\Sigma_\vartheta^{-1}\otimes \boldsymbol{I})(\boldsymbol{Z}_2'\otimes \boldsymbol{X})$$
$$\times [(\boldsymbol{Z}_2\otimes \boldsymbol{X}')(\Sigma_\vartheta^{-1}\otimes \boldsymbol{I})(\boldsymbol{Z}_2'\otimes \boldsymbol{X})]^-(\boldsymbol{Z}_2\otimes \boldsymbol{X}')(\Sigma_\vartheta^{-1}\otimes \boldsymbol{I})(\boldsymbol{Z}_1'\otimes \boldsymbol{X})\}^-$$
$$\times\{(\boldsymbol{Z}_1\otimes \boldsymbol{X}')(\Sigma_\vartheta^{-1}\otimes \boldsymbol{I}) - (\boldsymbol{Z}_1\otimes \boldsymbol{X}')(\Sigma_\vartheta^{-1}\otimes \boldsymbol{I})(\boldsymbol{Z}_2'\otimes \boldsymbol{X})$$
$$\times[(\boldsymbol{Z}_2\otimes \boldsymbol{X}')(\Sigma_\vartheta^{-1}\otimes \boldsymbol{I})(\boldsymbol{Z}_2'\otimes \boldsymbol{X})]^-(\boldsymbol{Z}_2\otimes \boldsymbol{X}')(\Sigma_\vartheta^{-1}\otimes \boldsymbol{I})\}vec(\boldsymbol{Y})$$
$$= \boldsymbol{p}'[(\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+\boldsymbol{Z}_1')^-\boldsymbol{Z}_1[\boldsymbol{M}_{Z_2'}\Sigma_\vartheta \boldsymbol{M}_{Z_2'}]^+ \otimes (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}']vec(\boldsymbol{Y}).$$

The validity of

$$\mathscr{M}[\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{Z}_1'\otimes \boldsymbol{X})] \subset \mathscr{M}[\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_\vartheta\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})']$$

follows from (5) and from regularity of Σ_ϑ. □

Lemma 3

$$(\boldsymbol{M}^W_{Z_2'\otimes X})'[\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}]^+\boldsymbol{M}^W_{Z_2'\otimes X}$$
$$= [\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X}]^+.$$

Proof Using Assertions 1,2 we have

$$(\boldsymbol{M}^W_{Z_2'\otimes X})'[\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}]^+\boldsymbol{M}^W_{Z_2'\otimes X}$$

$$= (\boldsymbol{M}^W_{Z_2'\otimes X})'[\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})']^+\boldsymbol{M}^W_{Z_2'\otimes X}$$

$$-(\boldsymbol{M}^W_{Z_2'\otimes X})'\left[\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\right]^+\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})$$

$$\times\{(\boldsymbol{Z}_1\otimes\boldsymbol{X}')(\boldsymbol{M}^W_{Z_2'\otimes X})'\left[\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\right]^+\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\}^-$$

$$\times(\boldsymbol{Z}_1\otimes\boldsymbol{X}')(\boldsymbol{M}^W_{Z_2'\otimes X})'\left[\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\right]^+\boldsymbol{M}^W_{Z_2'\otimes X}$$

$$= (\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X})^+ - (\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X})^+$$

$$\times(\boldsymbol{Z}_1'\otimes\boldsymbol{X})[(\boldsymbol{Z}_1\otimes\boldsymbol{X}')(\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X})^+(\boldsymbol{Z}_1'\otimes\boldsymbol{X})]^-$$

$$\times(\boldsymbol{Z}_1\otimes\boldsymbol{X}')(\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X})^+$$

$$= [\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X}]^+.$$

□

Theorem 4 *A linear function $\boldsymbol{g}'\vartheta$ of the vector parameter $\vartheta\in\underline{\vartheta}\subset R^p$, unbiasedly estimable in the model (2) before eliminating transformation is unbiasedly estimable in the transformed model (12).*

Proof The (i,j)-th element of the criterional matrix in the model (12) is given by

$$\left\{\boldsymbol{S}_{(M_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}M^W_{Z_2'\otimes X}(\Sigma_0\otimes I)(M^W_{Z_2'\otimes X})'M_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)})^+}\right\}_{i,j}$$

$$= Tr\Bigg\{\left[\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\right]^+$$

$$\times\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{V}_i\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'$$

$$\times\left[\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\right]^+$$

$$\times\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{V}_j\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\Bigg\}$$

$$= Tr\Bigg\{\left[\boldsymbol{M}^W_{Z_2'\otimes X})'[\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\right]^+$$

$$\times\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{V}_i\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'$$

$$\times\left[\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes\boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\right]^+\boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{V}_j\otimes\boldsymbol{I})\Bigg\}.$$

By Lemma 3 then

$$\left\{ \boldsymbol{S}_{(M_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)} M^W_{Z_2'\otimes X}(\Sigma_0\otimes I)(M^W_{Z_2'\otimes X})' M_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)})^+} \right\}_{i,j} =$$

$$= Tr\left\{ \left[\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X}\right]^+ (\boldsymbol{V}_i\otimes \boldsymbol{I})\right.$$

$$\left.\times \left[\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X}\right]^+ (\boldsymbol{V}_j\otimes \boldsymbol{I})\right\},\ i,j=1,\dots,p.$$

Due to the Remark 2 it is evident that the criterional matrices in the model (2) and in the model (12) are identical. □

Theorem 5 *Let $\boldsymbol{g}'\vartheta,\ \vartheta\in\underline{\vartheta}$ be an unbiasedly estimable function. Then the ϑ_0-MINQUE in the model (2) and the ϑ_0-MINQUE in the model (12) after elimination coincide.*

Proof We have seen that each function $\boldsymbol{g}'\vartheta$, that is unbiasedly estimable in the model (2) is unbiasedly estimable in the model (12).

According to Theorem 2 the ϑ_0-MINQUE in the model (12) is given by

$$\widehat{\boldsymbol{g}'\vartheta} = \sum_{i=1}^{p} \lambda_i (vec(\boldsymbol{Y}))'(\boldsymbol{M}^W_{Z_2'\otimes X})'$$

$$\left[\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\right]^+$$

$$\times\, \boldsymbol{M}^W_{Z_2'\otimes X}(\boldsymbol{V}_i\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'$$

$$\times \left[\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\boldsymbol{M}^W_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})(\boldsymbol{M}^W_{Z_2'\otimes X})'\boldsymbol{M}_{M^W_{Z_2'\otimes X}(Z_1'\otimes X)}\right]^+$$

$$\times\, \boldsymbol{M}^W_{Z_2'\otimes X} vec(\boldsymbol{Y})$$

$$= \sum_{i=1}^{p} \lambda_i (vec(\boldsymbol{Y}))' \left[\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X}\right]^+ (\boldsymbol{V}_i\otimes \boldsymbol{I})$$

$$\times \left[\boldsymbol{M}_{Z_1'\otimes X}\boldsymbol{M}_{Z_2'\otimes X}(\Sigma_0\otimes \boldsymbol{I})\boldsymbol{M}_{Z_2'\otimes X}\boldsymbol{M}_{Z_1'\otimes X}\right]^+ vec(\boldsymbol{Y}),$$

i.e. this estimator is identical to the estimator in the model (2)—see Remark 2. Lemma 3 has been taken into account. □

Lemma 4

$$[\boldsymbol{M}_{Z_1'\otimes X}(\Sigma_\vartheta\otimes \boldsymbol{I})\boldsymbol{M}_{Z_1'\otimes X}]^+ = (\Sigma_\vartheta^{-1}\otimes \boldsymbol{I}) - (\boldsymbol{P}^{\Sigma_\vartheta^{-1}}_{Z_1'}\otimes \boldsymbol{P}_X). \qquad (13)$$

Proof With respect to Assertion 1

$$[\boldsymbol{M}_{Z_1'\otimes X}(\Sigma_\vartheta \otimes \boldsymbol{I})\boldsymbol{M}_{Z_1'\otimes X}]^+ =$$

$$= (\Sigma_\vartheta^{-1} \otimes \boldsymbol{I}) - (\Sigma_\vartheta^{-1}\boldsymbol{Z}_1'(\boldsymbol{Z}_1\Sigma_\vartheta^{-1}\boldsymbol{Z}_1')^-\boldsymbol{Z}_1\Sigma_\vartheta^{-1} \otimes \boldsymbol{X}[\boldsymbol{X}'\boldsymbol{X}]^-\boldsymbol{X}')$$

$$= (\Sigma_\vartheta^{-1} \otimes \boldsymbol{I}) - (\Sigma_\vartheta^{-1}\boldsymbol{P}_{Z_1'}^{\Sigma_\vartheta^{-1}} \otimes \boldsymbol{P}_X).$$

□

Lemma 5

$$\boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}]^+} = \boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I}. \tag{14}$$

Proof With respect to $\boldsymbol{M}_A^V = \boldsymbol{I} - \boldsymbol{P}_A^V = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{V}\boldsymbol{A})^-\boldsymbol{A}'\boldsymbol{V}$ and using Lemma 4 we get

$$\boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\otimes X)(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}]^+}$$

$$= (\boldsymbol{I}\otimes\boldsymbol{I}) - (\boldsymbol{Z}_2'\otimes\boldsymbol{X})[(\boldsymbol{Z}_2\otimes\boldsymbol{X}')\{(\Sigma_\vartheta^{-1}\otimes\boldsymbol{I}) - (\Sigma_\vartheta^{-1}\boldsymbol{P}_{Z_1'}^{\Sigma_\vartheta^{-1}}\otimes\boldsymbol{P}_X)\}(\boldsymbol{Z}_2'\otimes\boldsymbol{X})]^-$$

$$\times (\boldsymbol{Z}_2\otimes\boldsymbol{X}')[(\Sigma_\vartheta^{-1}\otimes\boldsymbol{I}) - (\Sigma_\vartheta^{-1}\boldsymbol{P}_{Z_1'}^{\Sigma_\vartheta^{-1}}\otimes\boldsymbol{P}_X)]$$

$$= (\boldsymbol{I}\otimes\boldsymbol{I}) - (\boldsymbol{Z}_2'[\boldsymbol{Z}_2[\boldsymbol{M}_{Z_1'}\Sigma_\vartheta\boldsymbol{M}_{Z_1'}]^+\boldsymbol{Z}_2']^-\boldsymbol{Z}_2[\boldsymbol{M}_{Z_1'}\Sigma_\vartheta\boldsymbol{M}_{Z_1'}]^+\otimes\boldsymbol{X}[\boldsymbol{X}'\boldsymbol{X}]^-\boldsymbol{X}')$$

$$= (\boldsymbol{I}\otimes\boldsymbol{I}) - (\boldsymbol{P}_{Z_2'}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+}\otimes\boldsymbol{P}_X) = \boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I}.$$

□

Lemma 6

$$\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+}\cdot\boldsymbol{P}_{Z_2'\otimes X}^{[M_{Z_1'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}]^+} =$$

$$= \boldsymbol{P}_{Z_2'\otimes X}^{[M_{Z_1'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}]^+}\cdot\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+} = 0.$$

Proof With respect to Lemma 5

$$\boldsymbol{P}_{Z_2'\otimes X}^{[M_{Z_1'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}]^+} = \boldsymbol{P}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I} = \boldsymbol{P}_{Z_2'}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+}\otimes\boldsymbol{P}_X,$$

analogously $\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+} = \boldsymbol{P}_{Z_1'}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+}\otimes\boldsymbol{P}_X$. Since

$$\boldsymbol{P}_{Z_1'}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+}\cdot\boldsymbol{P}_{Z_2'}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+} = \boldsymbol{P}_{Z_2'}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+}\cdot\boldsymbol{P}_{Z_1'}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+} = 0,$$

we get the statements. □

Lemma 7

$$M^{\Sigma_\vartheta^{-1}\otimes I}_{(Z_1'\otimes X, Z_2'\otimes X)} = (I\otimes I) - (P_{Z_1'}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+}\otimes P_X) - (P_{Z_2'}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+}\otimes P_X)$$

$$= (I\otimes I) - P_{Z_1'\otimes X}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+} - P_{Z_2'\otimes X}^{[M_{Z_1'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}]^+}$$

$$= M_{Z_2'\otimes X}^{[M_{Z_1'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}]^+}\cdot M_{Z_1'\otimes X}^{[M_{Z_2'\otimes X}(\Sigma\otimes I)M_{Z_2'\otimes X}]^+}.$$

Proof

$$M^{\Sigma_\vartheta^{-1}\otimes I}_{(Z_1'\otimes X, Z_2'\otimes X)} = (I\otimes I) - (Z_1'\otimes X, Z_2'\otimes X)$$

$$\times\left[\begin{pmatrix} Z_1\otimes X' \\ Z_2\otimes X'\end{pmatrix}(\Sigma_\vartheta^{-1}\otimes I)(Z_1'\otimes X, Z_2'\otimes X)\right]^-\begin{pmatrix} Z_1\Sigma_\vartheta^{-1}\otimes X' \\ Z_2\Sigma_\vartheta^{-1}\otimes X'\end{pmatrix}$$

$$= (I\otimes I) - (Z_1'\otimes X, Z_2'\otimes X)$$

$$\times\begin{pmatrix} Z_1\Sigma_\vartheta^{-1}Z_1'\otimes X'X,\ Z_1\Sigma_\vartheta^{-1}Z_2'\otimes X'X \\ Z_2\Sigma_\vartheta^{-1}Z_1'\otimes X'X,\ Z_2\Sigma_\vartheta^{-1}Z_2'\otimes X'X\end{pmatrix}^-\begin{pmatrix} Z_1\Sigma_\vartheta^{-1}\otimes X' \\ Z_2\Sigma_\vartheta^{-1}\otimes X'\end{pmatrix}$$

$$= (I\otimes I) - (Z_1'\otimes X, Z_2'\otimes X)\begin{pmatrix} A_{11},\ A_{12} \\ A_{21},\ A_{22}\end{pmatrix}\begin{pmatrix} Z_1\Sigma_\vartheta^{-1}\otimes X' \\ Z_2\Sigma_\vartheta^{-1}\otimes X'\end{pmatrix},$$

where (using the second Rohde's formula)

$$A_{11} = (Z_1[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+Z_1')^-\otimes(X'X)^-,$$

$$A_{12} = -[(Z_1[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+Z_1')^-Z_1\Sigma_\vartheta^{-1}Z_2'(Z_2\Sigma_\vartheta^{-1}Z_2')^- \\ \otimes(X'X)^-(X'X)(X'X)^-],$$

and (using the first Rohde's formula)

$$A_{21} = -[(Z_2[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+Z_2')^-Z_2\Sigma_\vartheta^{-1}Z_1'(Z_1\Sigma_\vartheta^{-1}Z_1')^- \\ \otimes(X'X)^-(X'X)(X'X)^-],$$

$$A_{22} = (Z_2[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+Z_2')^-\otimes(X'X)^-.$$

Substituting these expressions we get the first assertion. The rest of the proof is evident (with respect to Lemma 5 and Lemma 6). □

If we use in the eliminating transformation $T = M^W_{Z_2'\otimes X}$ the following matrix

$$W = \left[M_{Z_1'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}\right]^+,$$

we get the transformation matrix (see (14))

$$T = M_{Z_2'\otimes X}^{[M_{Z_1'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_1'\otimes X}]^+} = M_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I},$$

that is very useful. It eliminates the nuisance parameters and does not change the design matrix belonging to the vector of useful parameters, i.e. this transformation yields the following model

$$\Bigg[\boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I} vec(\boldsymbol{Y}), (\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(\boldsymbol{B}_1),$$

$$\boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I}(\Sigma_\vartheta\otimes\boldsymbol{I})(\boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I})'\Bigg], \quad \Sigma_\vartheta \text{ p.d.} \tag{15}$$

Remark 4 a) The matrix $\boldsymbol{W} = [\boldsymbol{M}_{Z_1'\otimes X}(\Sigma_\vartheta\otimes\boldsymbol{I})\boldsymbol{M}_{Z_1'\otimes X}]^+$ satisfies the assumption (11), see [2], page 189.

b) Theorem 3, Theorem 4 and Theorem 5 are true in the model (15).

Let us consider the more general model

$$\Bigg[\boldsymbol{A}\boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I} vec(\boldsymbol{Y}), \boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(\boldsymbol{B}_1),$$

$$\boldsymbol{A}\boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I}(\Sigma_\vartheta\otimes\boldsymbol{I})(\boldsymbol{M}_{Z_2'\otimes X}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I})'\boldsymbol{A}'\Bigg], \quad \Sigma_\vartheta \quad \text{p.d.}, \tag{16}$$

where $\boldsymbol{A}$ is such that

$$\mathscr{M}[(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{A}'] = \mathscr{M}[(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{M}_{Z_2'\otimes X}], \tag{17}$$

i.e. the classes of the unbiasedly estimable functions in the model (2) and in the model (16) coincide.

It holds

$$E\left(\boldsymbol{A}\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+} vec(\boldsymbol{Y})\right) = E\left(\boldsymbol{A}\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+\otimes I} vec(\boldsymbol{Y})\right)$$

$$= \boldsymbol{A}\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+\otimes I}[(\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(\boldsymbol{B}_1) + (\boldsymbol{Z}_2'\otimes\boldsymbol{X})vec(\boldsymbol{B}_2)]$$

$$= \boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(\boldsymbol{B}_1),$$

i.e. $\boldsymbol{A}\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+\otimes I} vec(\boldsymbol{Y})$ is an unbiased estimator of the vector function $\boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(\boldsymbol{B}_1)$ for each matrix $\boldsymbol{A}$.

Lemma 8

$$\boldsymbol{A}\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+\otimes I} vec(\boldsymbol{Y}) = \boldsymbol{A}\boldsymbol{P}_{Z_1'\otimes X}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+} vec(\boldsymbol{Y})$$

is the best estimator of its mean value.

Proof We use the basic lemma on the locally best estimators (see [4], p. 84).

The class of the estimators of the null parametric function in the model (2) can be expressed in the form

$$\mathscr{U}_0 = \{\boldsymbol{u}'\boldsymbol{M}^{(\Sigma_\vartheta^{-1}\otimes I)}_{(Z_1'\otimes X, Z_2'\otimes X)} vec(\boldsymbol{Y}), \forall \boldsymbol{u} \in R^{mn}\},$$

as

$$E[\boldsymbol{L}'vec(\boldsymbol{Y})] = \boldsymbol{L}'(\boldsymbol{Z}_1'\otimes\boldsymbol{X}, \boldsymbol{Z}_2'\otimes\boldsymbol{X})\begin{pmatrix} vec(\boldsymbol{B}_1) \\ vec(\boldsymbol{B}_2)\end{pmatrix} = 0,$$

$$\forall vec(\boldsymbol{B}_1) \in R^{kr},\ \forall vec(\boldsymbol{B}_2) \in R^{ks},$$

$$\iff \boldsymbol{L}'(\boldsymbol{Z}_1'\otimes\boldsymbol{X}, \boldsymbol{Z}_2'\otimes\boldsymbol{X}) = \boldsymbol{o}'$$

$$\iff \boldsymbol{L} \in \mathscr{M}[\boldsymbol{M}_{(Z_1'\otimes X, Z_2'\otimes X)}] = \mathscr{M}[(\boldsymbol{M}^{\Sigma_\vartheta^{-1}\otimes I}_{(Z_1'\otimes X, Z_2'\otimes X)})'].$$

$$cov(\boldsymbol{A}\boldsymbol{P}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+}_{Z_1'\otimes X} vec(\boldsymbol{Y}), \boldsymbol{u}'\boldsymbol{M}^{\Sigma_\vartheta^{-1}\otimes I}_{(Z_1'\otimes X, Z_2'\otimes X)} vec(\boldsymbol{Y}))$$

$$= \boldsymbol{A}\boldsymbol{P}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+}_{Z_1'\otimes X}(\Sigma_\vartheta\otimes\boldsymbol{I})(\boldsymbol{M}^{\Sigma_\vartheta^{-1}\otimes I}_{(Z_1'\otimes X, Z_2'\otimes X)})'\boldsymbol{u}$$

$$= \boldsymbol{A}\boldsymbol{P}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+}_{Z_1'\otimes X}\boldsymbol{M}^{\Sigma_\vartheta^{-1}\otimes I}_{(Z_1'\otimes X, Z_2'\otimes X)}(\Sigma_\vartheta\otimes\boldsymbol{I})\boldsymbol{u} = \boldsymbol{o},\ \forall\boldsymbol{u}\in R^{mn},$$

for each matrix $\boldsymbol{A}$, as according to Lemma 6, Lemma 7

$$\boldsymbol{P}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+}_{Z_1'\otimes X}\boldsymbol{M}^{\Sigma_\vartheta^{-1}\otimes I}_{(Z_1'\otimes X, Z_2'\otimes X)} = 0. \qquad \square$$

Theorem 6 *In the model (16) the estimators* $\boldsymbol{A}\boldsymbol{P}^{[M_{Z_2'\otimes X}(\Sigma_\vartheta\otimes I)M_{Z_2'\otimes X}]^+}_{Z_1'\otimes X} vec(\boldsymbol{Y})$, *where* $\boldsymbol{A}$ *is an arbitrary matrix such that*

$$\mathscr{M}[(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{A}'] = \mathscr{M}[(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{M}_{Z_2'\otimes X}],$$

create the class of all optimal estimators of the vector function $\boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(\boldsymbol{B}_1)$.

Proof Let us denote $\boldsymbol{B} = \boldsymbol{M}^{[M_{Z_1'}\Sigma_\vartheta M_{Z_1'}]^+\otimes I}_{Z_2'\otimes X}$. According to [2], Theorem 3.1.3, the ϑ-LBLUE of the vector function $\boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(\boldsymbol{B}_1)$ in the model (16) is

$$\boldsymbol{A}\widehat{(\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec}(\boldsymbol{B}_1) = \boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\left\{[(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{A}']^-_{m(AB(\Sigma_\vartheta\otimes I)B'A')}\right\}'\boldsymbol{A}\boldsymbol{B}vec(\boldsymbol{Y})$$

$$= \boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\left\{(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{A}'\left[\boldsymbol{A}\boldsymbol{B}(\Sigma_\vartheta\otimes\boldsymbol{I})\boldsymbol{B}'\boldsymbol{A}'\right]^-\boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\right\}^-(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{A}'$$

$$\times\left[\boldsymbol{A}\boldsymbol{B}(\Sigma_\vartheta\otimes\boldsymbol{I})\boldsymbol{B}'\boldsymbol{A}'\right]^-\boldsymbol{A}\boldsymbol{B}vec(\boldsymbol{Y})$$

$$= \boldsymbol{A}\boldsymbol{B}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\{(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{B}'\boldsymbol{A}'[\boldsymbol{A}\boldsymbol{B}(\Sigma_\vartheta\otimes\boldsymbol{I})\boldsymbol{B}'\boldsymbol{A}']^-\boldsymbol{A}\boldsymbol{B}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\}^-(\boldsymbol{Z}_1\otimes\boldsymbol{X}')$$

$$\times\boldsymbol{B}'\boldsymbol{A}'\left[\boldsymbol{A}\boldsymbol{B}(\Sigma_\vartheta\otimes\boldsymbol{I})\boldsymbol{B}'\boldsymbol{A}'\right]^-\boldsymbol{A}\boldsymbol{B}vec(\boldsymbol{Y}) = \boldsymbol{P}^{[AB(\Sigma_\vartheta\otimes I)B'A']^-}_{AB(Z_1'\otimes X)}\boldsymbol{A}\boldsymbol{B}vec(\boldsymbol{Y}).$$

It is the best unbiased estimator. With respect to the basic lemma on the best estimators

$$cov\left\{\boldsymbol{P}_{AB(Z_1'\otimes X)}^{[AB(\Sigma_\vartheta\otimes I)B'A']^-}\boldsymbol{AB}vec(\boldsymbol{Y}),\boldsymbol{u}'\boldsymbol{M}_{(Z_1'\otimes X,Z_2'\otimes X)}^{\Sigma^{-1}\otimes I}vec(\boldsymbol{Y})\right\}=0,\ \forall\boldsymbol{u}\in R^{mn},$$

is valid, i.e.

$$\boldsymbol{P}_{AB(Z_1'\otimes X)}^{[AB(\Sigma_\vartheta\otimes I)B'A']^-}\boldsymbol{AB}(\Sigma_\vartheta\otimes\boldsymbol{I})(\boldsymbol{M}_{(Z_1'\otimes X,Z_2'\otimes X)}^{\Sigma_\vartheta^{-1}\otimes I})'\boldsymbol{u}'$$

$$=\boldsymbol{P}_{AB(Z_1'\otimes X)}^{[AB(\Sigma_\vartheta\otimes I)B'A']^-}\boldsymbol{ABM}_{(Z_1'\otimes X,Z_2'\otimes X)}^{\Sigma_\vartheta^{-1}\otimes I}(\Sigma_\vartheta\otimes\boldsymbol{I})\boldsymbol{u}'=0,\quad\forall u\ \in R^{mn}.$$

Thus

$$\boldsymbol{P}_{AB(Z_1'\otimes X)}^{[AB(\Sigma_\vartheta\otimes I)B'A']^-}\boldsymbol{ABM}_{Z_1'\otimes X}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+\otimes I}=0,$$

where Lemma 7 and Lemma 5 have been utilized. From this equality it follows

$$\boldsymbol{P}_{AB(Z_1'\otimes X)}^{[AB(\Sigma_\vartheta\otimes I)B'A']^-}\boldsymbol{AB}vec(\boldsymbol{Y})=\boldsymbol{P}_{AB(Z_1'\otimes X)}^{[AB(\Sigma_\vartheta\otimes I)B'A']^-}\boldsymbol{ABP}_{Z_1'\otimes X}^{[M_{Z_2'}\Sigma_\vartheta M_{Z_2'}]^+\otimes I}vec(\boldsymbol{Y}).$$

Let us denote

$$\boldsymbol{C}=(\boldsymbol{Z}_1\otimes\boldsymbol{X}')\boldsymbol{B}'\boldsymbol{A}'[\boldsymbol{AB}(\Sigma_\vartheta\otimes\boldsymbol{I})\boldsymbol{B}'\boldsymbol{A}']^-\boldsymbol{AB}(\boldsymbol{Z}_1'\otimes\boldsymbol{X}).$$

Then

$$\boldsymbol{A}(\widehat{\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(}\boldsymbol{B}_1)=\boldsymbol{P}_{AB(Z_1'\otimes X)}^{[AB(\Sigma_\vartheta\otimes I)B'A']^-}\boldsymbol{AB}vec(\boldsymbol{Y})$$

$$=\boldsymbol{AB}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\boldsymbol{C}^-\boldsymbol{C}[(\boldsymbol{Z}_1\otimes\boldsymbol{X}')([\boldsymbol{M}_{Z_2'}\Sigma_\vartheta\boldsymbol{M}_{Z_2'}]^+\otimes\boldsymbol{I})(\boldsymbol{Z}_1'\otimes\boldsymbol{X})]^-$$

$$\times(\boldsymbol{Z}_1\otimes\boldsymbol{X}')[(\boldsymbol{M}_{Z_2'}\Sigma_\vartheta\boldsymbol{M}_{Z_2'})^+\otimes\boldsymbol{I}]vec(\boldsymbol{Y})$$

$$=\boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})[(\boldsymbol{Z}_1\otimes\boldsymbol{X}')([\boldsymbol{M}_{Z_2'}\Sigma_\vartheta\boldsymbol{M}_{Z_2'}]^+\otimes\boldsymbol{I})(\boldsymbol{Z}_1'\otimes\boldsymbol{X})]^-$$

$$\times(\boldsymbol{Z}_1\otimes\boldsymbol{X}')[(\boldsymbol{M}_{Z_2'}\Sigma_\vartheta\boldsymbol{M}_{Z_2'})^+\otimes\boldsymbol{I}]vec(\boldsymbol{Y})$$

$$=\boldsymbol{AP}_{Z_1'\otimes X}^{(M_{Z_2'}\Sigma_\vartheta M_{Z_2'})^+\otimes I}vec(\boldsymbol{Y}),$$

(the best estimator of its mean value $\boldsymbol{A}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})vec(\boldsymbol{B}_1)$ according to Lemma 8).

The following equivalence has been taken into account

$$\boldsymbol{AM}_{Z_2'\otimes X}^{(M_{Z_2'}\Sigma_\vartheta M_{Z_2'})^+\otimes I}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\boldsymbol{C}^-\boldsymbol{C}=\boldsymbol{AM}_{Z_2'\otimes X}^{(M_{Z_2'}\Sigma_\vartheta M_{Z_2'})^+\otimes I}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})=\boldsymbol{AB}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})$$

$$\iff\mathscr{M}\left[\left(\boldsymbol{AM}_{Z_2'\otimes X}^{(M_{Z_2'}\Sigma_\vartheta M_{Z_2'})^+\otimes I}(\boldsymbol{Z}_1'\otimes\boldsymbol{X})\right)'\right]\subset\mathscr{M}(\boldsymbol{C}').$$

The g-inverse matrix in the matrix $\boldsymbol{C}$ can be chosen arbitrarily. If we chose it positive definite, the condition on the right side of the equivalence is obvious. □

Example 1 Let us consider following situation (see [5]). When laying the foundations for a large building it is necessary to determine the moment at which the subsoil (after large landscaping has been done) stabilizes to the point that it is possible to continue construction without risk of following damage.

There are n points chosen at the building site and their heights are repeatedly measured at the moments $t_1, \dots, t_m$. It is necessary to create a model describing the subsidence of the subsoil at the chosen points and to estimate the unknown parameters of this model on the basis of the results of the repeated measurements.

The result of the measurement at the i-th point in the j-th epoch could be described as follows:

$$\eta_i(t_j) = \kappa_i - \beta_1(1 - e^{-\beta_2 t_j}) + \varepsilon_{ij},\ i = 1, \dots, n,\ j = 1, \dots, m, \tag{18}$$

where κ_i is the height of the i-th point at time t_0, the function $\beta_1(1 - e^{-\beta_2 t})$ describes the movement of the earth-strata at each point. The parameters $\beta_1 > 0, \beta_2 > 0$ are the same at the different points, i.e. we suppose that the geological composition of the subsoil is homogenous. The aim is to estimate the unknown parameters β_1, β_2 and κ_i, $i = 1, \dots, n$.

The civil engineer needs to know when it is possible to continue the construction, i.e. when the subsidence of the subsoil at the points is insignificant. It means that it is necessary to determine such τ that

$$\beta_1(1 - e^{-\beta_2 \tau}) \ \geq \ C\beta_1,$$

where $0 < C < 1$ is a suitable constant which is sufficiently close to 1. It is possible to continue the construction at the time $t \geq \tau$.

The model (18) is not linear in parameters; we linearize it by using the first two members of the Taylor expansion of the function $\beta_1(1 - e^{-\beta_2 t})$ at the suitable point $(\beta_{1,0}, \beta_{2,0})$, $\beta_{1,0} > 0, \beta_{2,0} > 0$.

We get the model

$$\eta_i(t_j) =$$
$$= \kappa_i - [\beta_{1,0}(1 - e^{-\beta_{2,0} t_j}) + (1 - e^{-\beta_{2,0} t_j})(\beta_1 - \beta_{1,0}) + \beta_{1,0} t_j e^{-\beta_{2,0} t_j}(\beta_2 - \beta_{2,0})] + \varepsilon_{ij},$$
$$i = 1, \dots, n,\ j = 1, \dots, m.$$

Denote

$$Y_i^{(j)} = \eta_i(t_j) + \beta_{1,0}(1 - e^{-\beta_{2,0} t_j}),\ \varphi_1(t) = -(1 - e^{-\beta_{2,0} t}),\ \varphi_2(t) = -\beta_{1,0} t e^{-\beta_{2,0} t},$$
$$\delta\beta_1 = \beta_1 - \beta_{1,0}, \quad \delta\beta_2 = \beta_2 - \beta_{2,0},\ i = 1, \dots, n,\ j = 1, \dots, m.$$

Thus

$$Y_i^{(j)} = \kappa_i + \varphi_1(t_j)\delta\beta_1 + \varphi_2(t_j)\delta\beta_2 + \varepsilon_{ij},\ i = 1, \dots, n,\ j = 1, \dots, m.$$

Let us consider the observation vector

$$\boldsymbol{Y} = (Y^{(1)}, \dots, Y^{(m)}), \qquad \boldsymbol{Y}^{(j)} = (Y_1^{(j)}, \dots, Y_n^{(j)}).$$

The model described above could be rewritten in the form

$$\boldsymbol{Y} = \boldsymbol{X}(\boldsymbol{B}_1, \boldsymbol{B}_2)\begin{pmatrix}\boldsymbol{Z}_1\\ \boldsymbol{Z}_2\end{pmatrix} + \varepsilon,$$

where

$$\boldsymbol{X} = \boldsymbol{I}_k, \quad \boldsymbol{B}_1 = \begin{pmatrix}\delta\beta_1, & \delta\beta_2\\ \delta\beta_1, & \delta\beta_2\\ \vdots & \vdots\\ \delta\beta_1, & \delta\beta_2\end{pmatrix} \quad \boldsymbol{B}_2 = \begin{pmatrix}\kappa_1\\ \kappa_2\\ \vdots\\ \kappa_n\end{pmatrix},$$

$$\boldsymbol{Z}_1 = \begin{pmatrix}\varphi_1(t_1), & \varphi_1(t_2), & \dots & \varphi_1(t_m)\\ \varphi_2(t_1), & \varphi_2(t_2), & \dots & \varphi_2(t_m)\end{pmatrix} \quad \boldsymbol{Z}_2 = (1, 1, \dots, 1).$$

The $n \times 2$ matrix $\boldsymbol{B}_1$ is a matrix of useful parameters, the $n \times 1$ matrix $\boldsymbol{B}_2$ is a matrix of nuisance parameters.
Let us choose $n = 2, \quad m = 2, \quad t_1 = 1, \quad t_2 = 6, \quad \beta_{1,0} = 1, \quad \beta_{2,0} = 1,$

$$\boldsymbol{Z}_1 = \begin{pmatrix}-0,6321 & -0,9975\\ -0,3679 & -0,0149\end{pmatrix}, \quad \boldsymbol{Z}_2 = (1, 1).$$

$$\boldsymbol{B}_1 = \begin{pmatrix}\delta\beta_1 & \delta\beta_2\\ \delta\beta_1 & \delta\beta_2\end{pmatrix}, \quad \boldsymbol{B}_2 = \begin{pmatrix}\kappa_1\\ \kappa_2\end{pmatrix},$$

For the sake of simplicity let us choose $\boldsymbol{W} = \boldsymbol{I}, \quad \Sigma = \sigma^2\boldsymbol{I}$, then we have for $\boldsymbol{X} = \boldsymbol{I}$

$$\boldsymbol{M}^{W}_{Z_2'\otimes X} = \boldsymbol{I} - [\boldsymbol{Z}_2'(\boldsymbol{Z}_2\boldsymbol{Z}_2')^-\boldsymbol{Z}_2 \otimes \boldsymbol{I}] = \begin{pmatrix}0.5 & 0 & -0.5 & 0\\ 0 & 0.5 & 0 & -0.5\\ -0.5 & 0 & 0.5 & 0\\ 0 & -0.5 & 0 & 0.5\end{pmatrix},$$

$$\boldsymbol{M}^{[M_{Z_1'}\Sigma M_{Z_1'}]^+\otimes I}_{Z_2'\otimes X} = \boldsymbol{I} - [\boldsymbol{Z}_2'(\boldsymbol{Z}_2\boldsymbol{M}_{Z_1'}\boldsymbol{Z}_2')^-\boldsymbol{Z}_2\boldsymbol{M}_{Z_1'} \otimes \boldsymbol{I}] = \begin{pmatrix}1 & 0 & -1 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix},$$

$$\boldsymbol{P}^{[M_{Z_2'}\Sigma M_{Z_2'}]^+\otimes I}_{Z_1'} = \boldsymbol{Z}_1'(\boldsymbol{Z}_1\boldsymbol{M}_{Z_2'}\boldsymbol{Z}_1')^-\boldsymbol{Z}_1\boldsymbol{M}_{Z_2'} \otimes \boldsymbol{I}$$
$$= \begin{pmatrix}-0.3917 & 0 & 0.3917 & 0\\ 0 & -0.3917 & 0 & 0.3917\\ -1.3917 & 0 & 1.3917 & 0\\ 0 & -1.3917 & 0 & 1.3917\end{pmatrix}.$$

All these matrices eliminate the nuisance parameters.

Remark 5 Papers [3], [6] deal with univariate model, in [7] there is the multivariate linear model (2) with $var[vec(\boldsymbol{Y})] = \boldsymbol{I} \otimes \Sigma_\vartheta$ considered.

References

[1] Kubáček, L.: Foundations of estimation theory. *Elsevier, Amsterdam–Oxford–New York–Tokyo*, 1988.

[2] Kubáček, L., Kubáčková, L., Volaufová, J.: Statistical Models with Linear Structures. *Veda, Publishing House of the Slovak Academy of Sciences, Bratislava*, 1995.

[3] Kubáčková, L., Kubáček, L.: *Elimination Transformation of an Observation Vector preserving Information on the First and Second Order Parameters.* Technical Report, No 11, 1990, Institute of Geodesy, University of Stuttgart, 1–71.

[4] Kubáček, L., Kubáčková, L.: Statistika a metrologie. *Vydavatelství UP, Olomouc*, 2000.

[5] Kunderová, P.: *Locally best and uniformly best estimators in linear model with nuisance parameters.* Tatra Mt. Math. Publ. **22** (2001), 27–36.

[6] Kunderová, P.: *Eliminating transformations for nuisance parameters in linear model.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **42** (2003), 59–68.

[7] Kunderová, P.: *Eliminating transformations for nuisance parameters in multivariate linear model.* Folia Fac. Sci. Nat. Univ. Masarykianae Brunensis, Math. **12** (2004), 16 p. (to appear).

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 105–112

Remarks on Ideals in Lower-Bounded Dually Residuated Lattice-Ordered Monoids

Jan KÜHR

Department of Algebra and Geometry, Faculty of Science,
Palacký University, Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kuhr@inf.upol.cz

(Received November 25, 2003)

Abstract

Lattice-ordered groups, as well as GMV-algebras (pseudo MV-algebras), are both particular cases of dually residuated lattice-ordered monoids ($DR\ell$-monoids for short). In the paper we study ideals of lower-bounded $DR\ell$-monoids including GMV-algebras. Especially, we deal with the connections between ideals of a $DR\ell$-monoid A and ideals of the lattice reduct of A.

Key words: $DR\ell$-monoid, ideal, prime ideal.

2000 Mathematics Subject Classification: 06F05, 03G25

In 1965, K. L. N. Swamy [11] introduced the notion of a (commutative) dually residuated lattice-ordered semigroup in order to capture the common features of Abelian lattice-ordered groups and Brouwerian algebras. It turns out that well-known MV-algebras [1], an algebraic version of the Łukasiewicz infinite valued propositional logic, can be considered as certain bounded commutative $DR\ell$-monoids [7, 8]. The present concept of a (non-commutative) $DR\ell$-monoid is due to T. Kovář [3]:

Definition 1 An algebra $(A; +, 0, \vee, \wedge, \rightharpoonup, \leftharpoondown)$ of type $\langle 2, 0, 2, 2, 2, 2\rangle$ is said to be a *dually residuated lattice-ordered monoid* (simply, a *$DR\ell$-monoid*) if

(i) $(A; +, 0, \vee, \wedge)$ is an ℓ-monoid, i.e., $(A; +, 0)$ is a monoid, $(A; \vee, \wedge)$ is a lattice and the monoid operation distributes over the lattice operations;

(ii) for any $a, b \in A$, $a \rightharpoonup b$ is the least $x \in A$ such that $x + b \geqslant a$, and $a \leftharpoondown b$ is the least $y \in A$ such that $b + y \geqslant a$;

(iii) A fulfils the identities

$$((x \rightharpoonup y) \vee 0) + y \leqslant x \vee y, \quad y + ((x \leftharpoondown y) \vee 0) \leqslant x \vee y,$$
$$x \rightharpoonup x \geqslant 0, \quad x \leftharpoondown x \geqslant 0.$$

Recently, J. Rachůnek [10] established the notion of a GMV-algebra as a non-commutative generalization of MV-algebras. Non-commutative structures named pseudo MV-algebras extending MV-algebras were independently introduced also by G. Georgescu and A. Iorgulescu [2]. The relationship between GMV-algebras and $DR\ell$-monoids is similar to the commutative case [10, 6]: every GMV-algebra can be regarded as a bounded $DR\ell$-monoid satisfying certain additional conditions, and conversely, any bounded $DR\ell$-monoid that fulfils those conditions is in fact a GMV-algebra. Other examples come from lattice-ordered groups: every ℓ-group, as well as the positive cone of any ℓ-group, is a $DR\ell$-monoid. Therefore, dually residuated lattice-ordered monoids constitute a wide generalization of ℓ-groups and GMV-algebras. We should remark that there exist also other algebraic structures related to logic (for instance, pseudo BL-algebras) that are equivalent to particular $DR\ell$-monoids.

In this paper we deal with ideals of lower-bounded $DR\ell$-monoids (by [3], a $DR\ell$-monoid A is lower-bounded iff $0 \leqslant x$ for all $x \in A$). We will focus especially the connections between ideals in A and those in $\ell(A)$, the lattice reduct of A. The motivation is the following:

(1) When regarded to be a $DR\ell$-monoid, every GMV-algebra is a lower-bounded $DR\ell$-monoid;

(2) T. Kovář [3] proved that every $DR\ell$-monoid is isomorphic to the direct product of an ℓ-group and a $DR\ell$-monoid with 0 at the bottom.

Let us recall basic properties of dually residuated ℓ-monoids [3] and necessary facts about ideals [4].

Lemma 2 [3] *In any $DR\ell$-monoid we have:*

(i) $x \rightharpoonup x = 0 = x \leftharpoondown x$;

(ii) $((x \rightharpoonup y) \vee 0) + y = x \vee y = y + ((x \leftharpoondown y) \vee 0)$;

(iii) $x \rightharpoonup (y + z) = (x \rightharpoonup z) \rightharpoonup y$, $x \leftharpoondown (y + z) = (x \leftharpoondown y) \leftharpoondown z$;

(iv) *if* $x \leqslant y$ *then* $x \rightharpoonup z \leqslant y \rightharpoonup z$ *and* $x \leftharpoondown z \leqslant y \leftharpoondown z$;

(v) *if* $x \leqslant y$ *then* $z \rightharpoonup x \geqslant z \rightharpoonup y$ *and* $z \leftharpoondown x \geqslant z \leftharpoondown y$;

(vi) $x \leqslant y$ *iff* $x \rightharpoonup y \leqslant 0$ *iff* $x \leftharpoondown y \leqslant 0$;

(vii) $x \rightharpoonup (y \wedge z) = (x \rightharpoonup y) \vee (x \rightharpoonup z)$, $x \leftharpoondown (y \wedge z) = (x \leftharpoondown y) \vee (x \leftharpoondown z)$;

(viii) $(x \vee y) \rightharpoonup z = (x \rightharpoonup z) \vee (y \rightharpoonup z)$, $(x \vee y) \leftharpoondown z = (x \leftharpoondown z) \vee (y \leftharpoondown z)$.

Remark 3 In Definition 1, the condition (ii) can be equivalently replaced by the following identities [3, 10]:

$$\begin{array}{rl} (x \rightharpoonup y) + y \geqslant x, & y + (x \leftharpoondown y) \geqslant x, \\ x \rightharpoonup y \leqslant (x \vee z) \rightharpoonup y, & x \leftharpoondown y \leqslant (x \vee z) \leftharpoondown y, \\ (x + y) \rightharpoonup y \leqslant x, & (y + x) \leftharpoondown y \leqslant x. \end{array}$$

Letting $|x| = x \vee (0 \rightharpoonup x)$ we define the *absolute value* of $x \in A$. It is easily seen that $0 \leqslant x$ iff $x = |x|$, and hence in the special case that we are dealing with lower-bounded $DR\ell$-monoids, this concept is redundant.

Let $I \subseteq A$. Then I is said to be an *ideal* in A if (i) $0 \in I$, (ii) $x + y \in I$ for all $x, y \in I$, and (iii) $|y| \leqslant |x|$ implies $y \in I$ for all $x \in I$ and $y \in A$.

We use $\mathrm{Id}(A)$ to denote the set of all ideals in A; it is partially ordered by set-inclusion. Obviously, $\mathrm{Id}(A)$ is a complete lattice and for any $X \subseteq A$ there exists the smallest ideal, $I(X)$, including X. It can be easily shown that

$$I(X) = \{a \in A : |a| \leqslant |x_1| + \cdots + |x_n| \text{ for some } x_1, \ldots, x_n \in X, n \in \mathbb{N}\}.$$

In addition, the ideal lattice $\mathrm{Id}(A)$ is algebraic and distributive.

We define an ideal I to be *prime* if for all $J, K \in \mathrm{Id}(A)$, if $J \cap K \subseteq I$ then $J \subseteq I$ or $K \subseteq I$. Every ideal equals the intersection of all primes exceeding it, and $I \in \mathrm{Id}(A)$ is prime if and only if $|x| \wedge |y| \in I$ entails $x \in I$ or $y \in I$, for all $x, y \in A$.

An ideal I in A is called *normal* if $(x \rightharpoonup y) \vee 0 \in I$ iff $(x \leftharpoondown y) \vee 0 \in I$ for all $x, y \in A$. Equivalently, an ideal I is normal if and only if $x + I^+ = I^+ + x$ for every $x \in A$, where $I^+ = \{a \in I : 0 \leqslant a\}$. The normal ideals of any $DR\ell$-monoid correspond one-to-one to its congruence relations.

We shall write $\ell(A)$ for $(A; \vee, \wedge)$, the lattice reduct of A. As usual, for any $X \subseteq A$, $(X]$ denotes the lattice ideal generated by X. It is worth adding that by [3, Theorem 1.1.23], $\ell(A)$ is a distributive lattice.

From this moment on, A stands for a lower-bounded $DR\ell$-monoid!

Theorem 4 *For any $I \subseteq A$ such that $0 \in I$, the following conditions are equivalent:*

(i) *I is an ideal in A;*

(ii) *if $x \in I$ and $y \rightharpoonup x \in I$ then $y \in I$;*

(iii) *if $x \in I$ and $y \leftharpoondown x \in I$ then $y \in I$.*

Proof We are going to show (i) $\Leftrightarrow$ (ii); the proof of (i) $\Leftrightarrow$ (iii) is parallel.

(i) $\Rightarrow$ (ii): If $x \in I$ and $y \rightharpoonup x \in I$ then $y \leqslant x \vee y = (y \rightharpoonup x) + x \in I$, whence $y \in I$.

(ii) $\Rightarrow$ (i): For $x, y \in I$ we have

$$((x + y) \rightharpoonup y) \rightharpoonup x = (x + y) \rightharpoonup (x + y) = 0 \in I$$

which yields $(x + y) \rightharpoonup y \in I$ and therefore $x + y \in I$. If $y \leqslant x \in I$ then $y \rightharpoonup x = 0 \in I$, and so $y \in I$. □

Theorem 5 *Every ideal in A is an ideal in $\ell(A)$. Moreover, if I is a prime ideal in A then I is a prime ideal in $\ell(A)$.*

Proof Let $I \in \mathrm{Id}(A)$. Then clearly I is non-empty, $y \leqslant x$ entails $y \in I$ whenever $x \in I$, and we have also $x \vee y \in I$ for all $x, y \in I$ since $x \vee y \leqslant x + y$. The latter claim is evident. □

The converse statement fails to be true in general. However, we shall prove that if I is a lattice ideal generated by a set of additively idempotent elements or I is a minimal prime ideal in $\ell(A)$, then it is an ideal in A.

Let $Idem(A) = \{a \in A : a = a + a\}$.

Lemma 6 *For all $a \in Idem(A)$ and $x \in A$ we have:*

(i) $a + x = a \vee x = x + a$,

(ii) $x \rightharpoonup a = x \leftharpoondown a$.

Proof (i) To see that $a + x = a \vee x$, compute

$$\begin{aligned} a + x &= a \vee (a + x) = (a + a) \vee (a + x) \\ &= a + (a \vee x) = a + a + (x \leftharpoondown a) \\ &= a + (x \leftharpoondown a) = a \vee x. \end{aligned}$$

(ii) For every $y \in A$, $y \geqslant x \rightharpoonup a$ iff $a + y = y + a \geqslant x$ iff $y \geqslant x \leftharpoondown a$, so $x \rightharpoonup a = x \leftharpoondown a$. □

Theorem 7 *Let $X \subseteq Idem(A)$. Then $(X]$ is a normal ideal in A.*

Proof We have $a \in (X]$ iff $a \leqslant x_1 \vee ... \vee x_n$ for some $x_1, ..., x_n \in X$ and $a \in I(X)$ iff $a \leqslant x_1 + \cdots + x_m = x_1 \vee \ldots \vee x_m$ for some $x_1, \ldots, x_m \in X$, and therefore $I(X) = (X]$.

If $a \rightharpoonup b \in I(X)$ then $a \rightharpoonup b \leqslant x_1 + \cdots + x_n$, where $x_1, \ldots, x_n \in X$, which implies $a \leqslant x_1 + \cdots + x_n + b = b + x_1 + \cdots + x_n$, and so $a \leftharpoondown b \leqslant x_1 + \cdots + x_n$ proving $a \leftharpoondown b \in I(X)$. Similarly $a \leftharpoondown b \in I(X)$ entails $a \rightharpoonup b \in I(X)$, and consequently, $(X]$ is a normal ideal in A. □

We turn now to minimal prime ideals.

Theorem 8 *(i) Let I be a proper ideal in $\ell(A)$. For $x \in A \setminus I$, let us put*

$$\Phi(I, x) = \{a \in A : x \rightharpoonup a \notin I\}$$

and

$$\Phi(I) = \bigcap \{\Phi(I, x) : x \in A \setminus I\}.$$

Then $\Phi(I)$ is an ideal in A such that $\Phi(I) \subseteq I$. In addition, if I is prime then so is $\Phi(I)$.

(ii) Let I be a proper ideal in $\ell(A)$. For $x \in A \setminus I$, let us put

$$\Psi(I,x) = \{a \in A : x \leftharpoondown a \notin I\}$$

and

$$\Psi(I) = \bigcap\{\Psi(I,x) : x \in A \setminus I\}.$$

Then $\Psi(I)$ is an ideal in A such that $\Psi(I) \subseteq I$. In addition, if I is prime then so is $\Psi(I)$.

Proof (i) Let $a \in \Phi(I)$. If $a \notin I$ then $a \in \Phi(I,a)$, so $0 = a \rightharpoonup a \notin I$, a contradiction. Thus $a \in I$ and we have $\Phi(I) \subseteq I$.

We shall now prove that $\Phi(I) \in \mathrm{Id}(A)$. It is obvious that $0 \in \Phi(I)$ as $x \rightharpoonup 0 = x \notin I$ for all $x \in A \setminus I$. Further, let $a, b \in \Phi(I)$ and take any $x \in A \setminus I$. Then $x \rightharpoonup b \notin I$ and hence $x \rightharpoonup (a+b) = (x \rightharpoonup b) \rightharpoonup a \notin I$ since $a \in \Phi(I, x \rightharpoonup b)$; thus $a + b \in \Phi(I,x)$ for all $x \in A \setminus I$ and consequently, $a + b \in \Phi(I)$. If now $a \in \Phi(I)$ and $b \leqslant a$ then $x \rightharpoonup a \leqslant x \rightharpoonup b$ for every $x \in A \setminus I$, and therefore $x \rightharpoonup b \notin I$ since $x \rightharpoonup b \in I$ would imply $x \rightharpoonup a \in I$. Thus $b \in \Phi(I,x)$ for any $x \in A \setminus I$, i.e. $b \in \Phi(I)$.

For the latter statement we shall need two claims.

Claim A: If $x \leqslant y$ then $\Phi(I,x) \subseteq \Phi(I,y)$.

For every $a \in \Phi(I,x)$, $x \rightharpoonup a \leqslant y \rightharpoonup a$ entails $y \rightharpoonup a \notin I$, so $a \in \Phi(I,y)$.

Claim B: If $a \wedge b \in \Phi(I,x)$ then $a \in \Phi(I,x)$ or $b \in \Phi(I,x)$.

We have $a \wedge b \in \Phi(I,x)$ iff $(x \rightharpoonup a) \vee (x \rightharpoonup b) = x \rightharpoonup (a \wedge b) \notin I$ which yields $x \rightharpoonup a \notin I$ or $x \rightharpoonup b \notin I$.

Let now I be a prime ideal in $\ell(A)$ and assume that $a \wedge b \in \Phi(I)$ for $a, b \in A$. If neither a nor b belongs to $\Phi(I)$ then certainly $a \notin \Phi(I,x)$ and $b \notin \Phi(I,y)$ for some $x, y \in A \setminus I$. Since I a prime ideal in $\ell(A)$, it is obvious that $x \wedge y \notin I$. By Claim A we have $\Phi(I, x \wedge y) \subseteq \Phi(I,x) \cap \Phi(I,y)$, and so $a \wedge b \in \Phi(I)$ yields $a \wedge b \in \Phi(I, x \wedge y) \subseteq \Phi(I,x) \cap \Phi(I,y)$. Hence by Claim B, $a \in \Phi(I, x \wedge y) \subseteq \Phi(I,x) \cap \Phi(I,y)$ or $b \in \Phi(I, x \wedge y) \subseteq \Phi(I,x) \cap \Phi(I,y)$, a contradiction with $a \notin \Phi(I,x)$ and $b \notin \Phi(I,y)$. Thus $a \wedge b \in \Phi(I)$ implies $a \in \Phi(I)$ or $b \in \Phi(I)$.

By replacing "$\rightharpoonup$" by "$\leftharpoondown$" we obtain (ii). □

Remark 9 If $I \in \mathrm{Id}(A)$ then $I = \Phi(I) = \Psi(I)$. Indeed, by Theorem 4 (ii), $a \in I$ and $x \notin I$ yield $x \rightharpoonup a \notin I$. Thus $I \subseteq \Phi(I)$.

Corollary 10 *For every $I \subseteq A$, I is a minimal prime ideal in A if and only if it is a minimal prime ideal in $\ell(A)$.*

Proof If I is a minimal prime ideal in A, then it is a prime ideal in $\ell(A)$ by Theorem 5, and by Theorem 8, I is minimal prime.

Conversely, if I is a minimal prime ideal in $\ell(A)$ then, again by Theorem 8, $\Phi(I)$ is a minimal prime ideal in A and obviously $I = \Phi(I)$. □

Remark 11 Let I be an ideal in $\ell(A)$. If I is a normal subset of A, that is, $x \rightharpoonup y \in I$ iff $x \leftharpoondown y \in I$ for all $x, y \in A$, then one can easily show that $\Phi(I) = \Psi(I)$. Conversely, an ideal I in $\ell(A)$ satisfying $\Phi(I) = \Psi(I)$ need not be normal.

Lemma 12 *If $z \leqslant x + y$ then $z = x_1 + y_1$ for some $x_1 \leqslant x$ and $y_1 \leqslant y$.*

Proof Let $x_1 = x \wedge z \leqslant x$ and $y_1 = z \leftharpoondown x_1$. Then

$$x_1 + y_1 = x_1 + (z \leftharpoondown x_1) = z \vee x_1 = z,$$

where $y_1 = z \leftharpoondown (x \wedge z) = (z \leftharpoondown x) \vee (z \leftharpoondown z) = z \leftharpoondown x \leqslant y$ as desired. □

Corollary 13 *If I, J are normal ideals in A then*

$$I \vee J = \{a \in A : a = x + y \textit{ for some } x \in I, y \in J\}.$$

Proof Since I, J are normal ideals, $a \in I \vee J$ iff $a \leqslant x + y$ for $x \in I$ and $y \in J$, and so by Lemma 12, $a = x_1 + y_1$ for some $x_1 \leqslant x, y_1 \leqslant y$, i.e. $x_1 \in I$ and $y_1 \in J$. □

Let A be a bounded $DR\ell$-monoid with the greatest element 1. Let us denote by $B(A)$ the set of all $a \in A$ having the complement a' in $\ell(A)$.

Lemma 14 If $x \wedge y = 0$ then $x + y = x \vee y$.

Proof Let $x \wedge y = 0$. Then

$$x = x \rightharpoonup (x \wedge y) = (x \rightharpoonup x) \vee (x \rightharpoonup y) = x \rightharpoonup y$$

which yields $x + y = (x \rightharpoonup y) + y = x \vee y$. □

Lemma 15 *$B(A) \subseteq Idem(A)$.*

Proof Let $a \in B(A)$, i.e. $a \wedge a' = 0$ and $a \vee a' = 1$ for some $a' \in A$. Note that $a + a' = 1$ since $a \vee a' \leqslant a + a'$. Then

$$a = a + (a \wedge a') = (a + a) \wedge (a + a') = (a + a) \wedge 1 = a + a,$$

so $a \in Idem(A)$. □

Remark 16 Observe that if $a \in B(A)$ then $(a]$ and $(a']$ are normal ideals in A such that $(a] \cap (a'] = \{0\}$ and $(a] \vee (a'] = A$, and therefore we can easily see that A is isomorphic with the direct product of $(a]$ and $(a']$.

Theorem 17 *$B(A)$ is a $DR\ell$-submonoid of A in which $a + b = a \vee b$ and $a \rightharpoonup b = a \leftharpoondown b = a \wedge b'$.*

Proof One readily sees that $B(A)$ is a sublattice of $\ell(A)$ since $\ell(A)$ is a distributive lattice.

By Lemma 6, $a \rightharpoonup b = a \leftharpoondown b$ and $x \geqslant a \rightharpoonup b$ iff $x \vee b = x + b \geqslant a$, whence $a \wedge b' \leqslant (x \vee b) \wedge b' = x \wedge b' \leqslant x$. Conversely, if $x \geqslant a \wedge b'$ then $x + b = x \vee b \geqslant (a \wedge b') \vee b = a \vee b \geqslant a$, thus $x \geqslant a \rightharpoonup b$. Altogether, $x \geqslant a \rightharpoonup b$ iff $x \geqslant a \wedge b'$ for any $x \in A$. Therefore $(a \rightharpoonup b)' = a' \vee b$ and so $a \rightharpoonup b \in B(A)$. □

Corollary 18 $(B(A); \vee, \wedge, \ ', 0, 1)$ *is a Boolean algebra, where* $a' = 1 \rightharpoonup a$.

By [6, Theorem 2.3], A is a GMV-algebra if and only if the identities

$$x \wedge y = x \rightharpoonup (x \leftharpoondown y) = x \leftharpoondown (x \rightharpoonup y)$$

hold in A. Therefore, let

$$GMV(A) = \{a \in A : a \wedge x = x \rightharpoonup (x \leftharpoondown a) = x \leftharpoondown (x \rightharpoonup a) \text{ for all } x \in A\}.$$

Lemma 19 *The following identities hold in any DRℓ-monoid:*

(i) $y \geqslant x \rightharpoonup (x \leftharpoondown y)$, $y \geqslant x \leftharpoondown (x \rightharpoonup y)$,

(ii) $x \leftharpoondown (x \rightharpoonup (x \leftharpoondown y)) = x \leftharpoondown y$, $x \rightharpoonup (x \leftharpoondown (x \rightharpoonup y)) = x \rightharpoonup y$.

Proof (i) Obviously, $y \geqslant x \rightharpoonup (x \leftharpoondown y)$ iff $x \vee y = y + (x \leftharpoondown y) \geqslant x$.

(ii) From $y \geqslant x \rightharpoonup (x \leftharpoondown y)$ we obtain

$$x \leftharpoondown y \leqslant x \leftharpoondown (x \rightharpoonup (x \leftharpoondown y))$$

and by replacing y by $x \leftharpoondown y$ in (i) we immediately have

$$x \leftharpoondown y \geqslant x \leftharpoondown (x \rightharpoonup (x \leftharpoondown y)).$$ □

Theorem 20 $B(A) = Idem(A) \cap GMV(A)$.

Proof If $a \in Idem(A) \cap GMV(A)$ then

$$(1 \rightharpoonup a) \vee a = (1 \rightharpoonup a) + a = 1 \vee a = 1$$

and

$$\begin{aligned}(1 \rightharpoonup a) \wedge a &= (1 \rightharpoonup a) \leftharpoondown ((1 \rightharpoonup a) \rightharpoonup a) = (1 \rightharpoonup a) \leftharpoondown (1 \rightharpoonup (a + a)) \\ &= (1 \rightharpoonup a) \leftharpoondown (1 \rightharpoonup a) = 0,\end{aligned}$$

so $a \in B(A)$.

Conversely, let $a \in B(A) \subseteq Idem(A)$, that is, $a \wedge a' = 0$. In view of Lemma 19 (i) we have $x \rightharpoonup (x \leftharpoondown a) \leqslant x \wedge a$. However,

$$\begin{aligned}x \rightharpoonup (x \leftharpoondown a) &= (x \rightharpoonup (x \leftharpoondown a)) + (a \wedge a') \\ &= ((x \rightharpoonup (x \leftharpoondown a)) + a) \wedge ((x \rightharpoonup (x \leftharpoondown a)) + a') \geqslant a \wedge x\end{aligned}$$

since $(x \rightharpoonup (x \leftharpoondown a)) + a \geqslant a$ and $a' = 1 \leftharpoondown a \geqslant x \leftharpoondown a = x \leftharpoondown (x \rightharpoonup (x \leftharpoondown a))$ by Lemma 2 (iv) and Lemma 19 (ii), which implies $(x \rightharpoonup (x \leftharpoondown a)) + a' \geqslant x$. Therefore, $a \in Idem(A) \cap GMV(A)$. □

Lemma 21 $B(A) = \{a \in A : a \wedge (1 \rightharpoonup a) = 0\} = \{a \in A : a \wedge (1 \leftharpoondown a) = 0\}$.

Proof If $a \wedge (1 \rightharpoonup a) = 0$ then

$$(1 \rightharpoonup a) \vee a = (1 \rightharpoonup a) + a = 1 \vee a = 1$$

by Lemma 14. Thus $a' = 1 \rightharpoonup a$ is the complement of a in $\ell(A)$. □

Corollary 22 *Let I be a normal ideal in A. Then A/I is a Boolean algebra if and only if $a \wedge (1 \rightharpoonup a) \in I$ for all $a \in A$.*

References

[1] Cignoli, R. L. O., Mundici, D., D'Ottaviano, I. M. L.: Algebraic Foundations of Many-valued Reasoning. *Kluwer Acad. Publ., Dordrecht-Boston-London*, 2000.

[2] Georgescu, G., Iorgulescu, A.: *Pseudo MV-algebras.* Mult. Valued Log. **6** (2001), 95–135.

[3] Kovář, T.: *A General Theory of Dually Residuated Lattice Ordered Monoids.* Ph.D. Thesis, Palacký University, Olomouc, 1996.

[4] Kühr, J.: *Ideals of noncommutative $DR\ell$-monoids.* Czech. Math. J. (to appear).

[5] Kühr, J.: *Prime ideals and polars in $DR\ell$-monoids and pseudo BL-algebras.* Math. Slovaca **53** (2003), 233–246.

[6] Kühr, J.: *A generalization of GMV-algebras.* (submitted).

[7] Rachůnek, J.: *$DR\ell$-semigroups and MV-algebras.* Czech. Math. J. **48** (1998), 365–372.

[8] Rachůnek, J.: *MV-algebras are categorically equivalent to a class of $DR\ell_{1(i)}$-semigroups.* Math. Bohem. **123** (1998), 437–441.

[9] Rachůnek, J.: *Connections between ideals of non-commutative generalizations of MV-algebras and ideals of their underlying lattices.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **40** (2001), 195–200.

[10] Rachůnek, J.: *A non-commutative generalization of MV-algebras.* Czech. Math. J. **52** (2002), 255–273.

[11] Swamy, K. L. N.: *Dually residuated lattice ordered semigroups.* Math. Ann. **159** (1965), 105–114.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 113–117

Join-Closed and Meet-Closed Subsets in Complete Lattices *

František MACHALA [1], Vladimír SLEZÁK [2]

Department of Algebra and Geometry,
Faculty of Science,
Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: [1]*F.Machala@seznam.cz*
[2]*slezakv@seznam.cz*

(Received January 21, 2004)

Abstract

To every subset A of a complete lattice L we assign subsets $J(A)$, $M(A)$ and define join-closed and meet-closed sets in L. Some properties of such sets are proved. Join- and meet-closed sets in power-set lattices are characterized. The connections about join-independent (meet-independent) and join-closed (meet-closed) subsets are also presented in this paper.

Key words: Complete lattices, join-closed and meet-closed sets.

2000 Mathematics Subject Classification: 06B23, 08A02, 08A05

Let $(L, \leq)$ be a complete lattice in which $\bigvee A, \bigwedge A$ denote the supremum and the infimum of any subset $A \subseteq L$, respectively. The least and the greatest elements in $(L, \leq)$ are denoted by 0, 1, respectively. If $A \subseteq L$, $A \neq \emptyset$, then we put $A_x := A \setminus \{x\}$ for $x \in A$ and

$$J(A) = \Big\{\bigvee A_x \mid x \in A\Big\}, \quad M(A) = \Big\{\bigwedge A_x \mid x \in A\Big\}.$$

Instead of $M(J(A))$, $J(M(A))$ we write just $MJ(A), JM(A)$. If we put $P_x = (J(A))_{\bigvee A_x} = \{\bigvee A_a \mid a \in A_x\}$, then $MJ(A) = \{\bigwedge P_x \mid x \in A\}$. Dually, $R_x = (M(A))_{\bigwedge A_x} = \{\bigwedge A_a \mid a \in A_x\}$ and $JM(A) = \{\bigvee R_x \mid x \in A\}$. It is easy to see that $x \leq \bigwedge P_x$ and $\bigvee R_x \leq x$ for all $x \in A$, thus $\bigvee R_x \leq \bigwedge P_x$.

*Supported by the Council of Czech Government J14/98:153100011.

Proposition 1 *If $A \subseteq L$, $|A| > 2$, then $\bigvee M(A) \leq \bigwedge J(A)$.*

Proof Consider $x \in A$ and $z \in A_x$. By assumption, there exists an element $y \in A_x$ distinct from z. From $x, z \in A_y$ we get $\bigwedge A_z \leq \bigvee R_y$ and $\bigwedge P_y \leq \bigvee A_x$, thus $\bigwedge A_z \leq \bigvee A_x$. We also have $\bigwedge A_x \leq \bigvee A_x$ and hence $\bigwedge A_z \leq \bigvee A_x$ for all $z \in A$. We have obtained the relation $\bigvee M(A) \leq \bigvee A_x$ holding for all $x \in A$. Thus $\bigvee M(A) \leq \bigwedge J(A)$. □

Definition 1 A set $A \subseteq L$ is said to be *meet-closed* iff $MJ(A) = A$. Similarly, $A \subseteq L$ is *join-closed* iff $JM(A) = A$. In brief, we call them M-closed and J-closed, respectively.

Remark 1 A set $A = \{x\}$ is M-closed (J-closed) if and only if $x = 1$ ($x = 0$). If $A = \{x, y\}$, then $J(A) = A = M(A)$ and A is both M-closed and J-closed.

Proposition 2 *A subset $A \subseteq L$ is M-closed if and only if $x = \bigwedge P_x$ for all $x \in A$.*

Proof 1. If $x = \bigwedge P_x$ for all $x \in A$, then $MJ(A) = \{x \mid x \in A\} = A$.

2. Assume that $MJ(A) = A$ and consider $x \in A$. It follows from $\bigwedge P_x \in A$ that $\bigwedge P_x = y$ for a certain $y \in A$ and since $x \leq \bigwedge P_x$ we have $x \leq y$. Let us suppose that $x \neq y$. Then $\bigvee A_y \in P_x$ which yields $y \leq \bigvee A_y$. From $y \leq \bigwedge P_y$ we obtain $y \leq \bigwedge J(A)$. Consequently (with respect to $P_x \subseteq J(A)$), $\bigwedge J(A) \leq \bigwedge P_x = y$ and $y = \bigwedge J(A)$. There exists $z \in A$ such that $x = \bigwedge P_z$. Then $y \leq \bigwedge P_z$, i. e. $y \leq x$ which contradicts the assumption $x < y$. Thus $x = \bigwedge P_x$. □

Remark 2 The notions of M-closed and J-closed sets are dual, hence each assertion about M-closed and J-closed sets admits its corresponding dual one. Therefore, a set $A \subseteq L$ is J-closed iff $x = \bigvee R_x$ for all $x \in A$. In what follows the dual results will not be stated explicitly.

Proposition 3 *If $A \subseteq L$, then the set $M(A)$ is M-closed.*

Proof If we put $Q_x = (JM(A))_{\bigvee R_x} = \{\bigvee R_y \mid y \in A_x\}$, then $MJM(A) = \{\bigwedge Q_x \mid x \in A\}$. Consider $x \in A$. Then $\bigwedge Q_x \leq \bigvee R_y \leq y$ for all $y \in A_x$ which implies $\bigwedge Q_x \leq \bigwedge A_x$. Furthermore, $\bigwedge A_x \in R_y$, thus $\bigwedge A_x \leq \bigvee R_y$ and $\bigwedge A_x \leq \bigwedge Q_x$. We have obtained $\bigwedge Q_x = \bigwedge A_x$ and $MJM(A) = \{\bigwedge A_x \mid x \in A\} = M(A)$. □

Proposition 4 *If a set $A \subseteq L$, $|A| > 1$, is M-closed, then $\bigwedge J(A) = \bigwedge A$.*

Proof Let us consider $x \in A$. Then there exists $y \in A_x$ such that $\bigwedge A \leq y \leq \bigvee A_x$. Thus $\bigwedge A \leq \bigwedge J(A)$. We also have $P_x \subseteq J(A)$ and $x = \bigwedge P_x$ which yields $\bigwedge J(A) \leq x$ and $\bigwedge J(A) \leq \bigwedge A$. □

Remark 3 A set $A \subseteq L$ is M-closed if and only if $A \cup \{\bigwedge A\}$ is M-closed.

Proposition 5 *Every subset of an M-closed set containing at least two elements is M-closed.*

Proof Let X be a subset of an M-closed set $A \subseteq L$. If $|X| = 2$, then X is M-closed by Remark 1. Let $|X| > 2$. Consider $x \in X$ and denote $Q_x = \{\bigvee X_l \mid l \in X_x\}$, $y = \bigwedge Q_x$. Since $x \leq \bigvee X_l$ for all $l \in X_x$ we have $x \leq y$. Obviously, $X_l \subseteq A_l$ for all $l \in X_x$, which yields $y \leq \bigvee X_l \leq \bigvee A_l$. If $m \in A \setminus X$, then $X_l \subseteq X \subseteq A_m$ and $y \leq \bigvee X_l \leq \bigvee A_m$ for any $l \in X_x$. If $a \in A_x$, then either $a \in X_x$ or $a \in A \setminus X$. Thus $y \leq \bigvee A_a$ and $y \leq \bigwedge P_x = x$. It means that $x = \bigwedge Q_x$ and the set X is M-closed. □

Proposition 6 *Let $A \subseteq L$, $|A| > 1$, be an M-closed set, X_i, $i \in J$, be non-empty subsets of A such that $\bigcap_{i \in J} X_i = \emptyset$ and $\mathcal{X} = \{\bigvee X_i \mid i \in J\}$. Then $\bigwedge \mathcal{X} = \bigwedge A$.*

Proof It is easy to see that $\bigwedge A \leq \bigwedge \mathcal{X}$. For each $i \in J$ and $x \in A \setminus X_i$ we have $X_i \subseteq A_x$ and hence $\bigvee X_i \leq \bigvee A_x$. It follows from $\bigcap_{i \in J} X_i = \emptyset$ that $\bigcup_{i \in J}(A \setminus X_i) = A$ and $\bigwedge \mathcal{X} \leq \bigvee A_y$ for all $y \in A$. Thus $\bigwedge \mathcal{X} \leq \bigwedge J(A)$ and, according to Proposition 4, $\bigwedge \mathcal{X} \leq \bigwedge A$. □

Corollary 1 *Let $A \subseteq L$, $|A| > 1$, be an M-closed set. Then $\bigwedge X = \bigwedge A$ for any $X \subseteq A$, $|X| \geq 2$.*

Definition 2 A subset $A \subseteq L$ is said to be *join-independent (meet-independent)* if and only if $x \not\leq \bigvee A_x$ ($\bigwedge A_x \not\leq x$) for all $x \in A$.

Remark 4 The concept of independence have been studied in various types of lattices motivated by applications in algebra and geometry (refer to [1, 2, 3, 4, 8]). Definition 2 is given in [5] and some other related results are presented in [6, 7].

Remark 5 Join- and meet-independence are dual notions, hence each of the following results holds also dually.

Remark 6 If a set $A \subseteq L$ is join-independent, then $J(A)$ is meet-independent. (See [5, 6].)

Proposition 7 *If a set $A \subseteq L$, $|A| > 2$, is meet-independent, then it is not M-closed.*

Proof Let A be a meet-independent set. Suppose that it is also M-closed. Then $x = \bigwedge P_x$ for all $x \in A$. It follows from $P_x \subseteq J(A)$ that $\bigwedge J(A) \leq \bigwedge P_x$. Since $\bigvee M(A) \leq \bigwedge J(A)$ (Proposition 1) we have $\bigwedge A_x \leq \bigvee M(A) \leq \bigwedge J(A) \leq x$ which contradicts the meet-independence of A. □

Let A be a set. In what follows we denote the power set of A by $\mathcal{P}(A)$. Then $(\mathcal{P}(A), \subseteq)$ is a complete lattice with lattice operations $\cup, \cap$.

Proposition 8 *Let A be a set and $X = \{X_i \mid i \in J\} \subseteq \mathcal{P}(A)$ where $|J| > 1$. The set X is M-closed in $(\mathcal{P}(A), \subseteq)$ if and only if $X_k \cap X_l = \bigcap X$ for every two distinct elements k, l of J.*

Proof It is evident that $J(X) = \{\bigcup X_{X_i} \mid i \in J\} = \{\bigcup_{j \in J \setminus \{i\}} X_j \mid i \in J\}$, $P_{X_i} = \{\bigcup X_{X_j} \mid j \in J \setminus \{i\}\} = \{\bigcup_{m \in J \setminus \{j\}} X_m \mid j \in J \setminus \{i\}\}$ and $MJ(X) = \{\bigcap P_{X_i} \mid i \in J\}$.

1. Assume that $X = MJ(X)$. If $|J| = 2$, then $X = \{X_1, X_2\}$ and $\bigcap X = X_1 \cap X_2$. For $|J| > 2$ we have $X_i = \bigcap P_{X_i}$ for all $i \in J$ by Proposition 2. Consider any two distinct elements $k, l \in J$. Then $\bigcap X \subseteq X_k \cap X_l$. Let $x \in X_k \cap X_l$. If $i \in J$ is distinct from k, l, then for each $j \in J \setminus \{i\}$ either $X_k \subseteq \bigcup X_{X_j}$ or $X_l \subseteq \bigcup X_{X_j}$ and hence $x \in \bigcap P_{X_i}$ and $x \in X_i$. Since it holds for all $i \in J$ distinct from k, l we have $x \in \bigcap X$ which yields $\bigcap X = X_k \cap X_l$.

2. Assume that $\bigcap X = X_k \cap X_l$ for any $k, l \in J$, $k \neq l$. In case of $|J| = 2$ this equality always holds and X is M-closed by Remark 1. Let $|J| > 2$. Consider $i \in J$ and denote $X^j = \{X_m \mid m \in J \setminus \{i, j\}\}$ for all $j \in J \setminus \{i\}$. Then $P_{X_i} = \{X_i \cup (\bigcup X^j) \mid j \in J \setminus \{i\}\}$. Let $x \in \bigcap \{\bigcup X^j \mid j \in J \setminus \{i\}\}$, i. e. $x \in X_k$ for a certain $k \in J \setminus \{i\}$. However, x belongs to another set X_l, $l \in J \setminus \{i\}$, $l \neq k$. Indeed, otherwise we get $x \notin \bigcup X^k$ which is a contradiction. Thus $x \in X_k \cap X_l$ and, by assumption, also $x \in X_i$. It follows from $X_i \subseteq \bigcap P_{X_i}$ that $X_i = \bigcap P_{X_i}$ and the set X is M-closed by Proposition 2. □

Let $A \subseteq L$ be join-independent set. Consider a mapping $\psi : \mathcal{P}(A) \to L$ given by $\psi(X) = \bigvee X$ for all non-empty subsets $X \in \mathcal{P}(A)$ and $\psi(\emptyset) = \bigwedge A$. According to [5], $(\psi(\mathcal{P}(A)), \leq)$ is a complete lattice isomorphic to $(\mathcal{P}(A), \subseteq)$ which is also a complete join subsemilattice of $(L, \leq)$.

Proposition 9 *Let a set $A \subseteq L$ be join-independent and consider subsets $X = \{X_i \mid i \in J\} \subseteq \mathcal{P}(A)$, $\mathcal{X} = \{\psi(X_i) \mid i \in J\} \subseteq L$. The following statements are equivalent:*

(i) *X is join-independent in $(\mathcal{P}(A), \subseteq)$.*

(ii) *$X_i \nsubseteq \bigcup_{j \in J \setminus \{i\}} X_j$ for all $i \in J$.*

(iii) *$\mathcal{X}$ is join-independent in $(L, \leq)$.*

Proof It is obvious.

Proposition 10 *Let a join-independent set $A \subseteq L$, $|A| > 2$, be M-closed in $(L, \leq)$. The following statements are equivalent:*

(i) *The set $L_1 = \psi(\mathcal{P}(A))$ is a sublattice in $(L, \leq)$.*

(ii) *The image of any M-closed set in $(\mathcal{P}(A), \subseteq)$ of cardinality 3 under the mapping ψ is M-closed in $(L, \leq)$.*

(iii) *The image of any join-independent M-closed set in $(\mathcal{P}(A), \subseteq)$ of cardinality 3 under the mapping ψ is M-closed in $(L, \leq)$.*

Proof $(i) \Rightarrow (ii)$ Let $X = \{X_1, X_2, X_3\} \subseteq \mathcal{P}(A)$ be an M-closed set. According to Proposition 2, for each $i \in \{1,2,3\}$ we have $\bigcap P_{X_i} = (X_i \cup X_j) \cap (X_i \cup X_k) = X_i$ where $j,k \in \{1,2,3\}$ and i,j,k are pairwise distinct. If $\psi(X) = \{\psi(X_1), \psi(X_2), \psi(X_3)\}$, then in $(L, \leq)$ there we have

$$\bigwedge P_{\psi(X_i)} = (\psi(X_i) \vee \psi(X_j)) \wedge (\psi(X_i) \vee \psi(X_k)) = \psi(X_i \cup X_j) \wedge \psi(X_i \cup X_k)$$
$$= \psi((X_i \cup X_j) \cap (X_i \cup X_k)) = \psi(X_i).$$

Thus, by Proposition 2, the set $\psi(X)$ is M-closed in $(L, \leq)$.

$(ii) \Rightarrow (iii)$ Obvious.

$(iii) \Rightarrow (i)$ Since $\psi(\mathcal{P}(A))$ is a join subsemilattice in $(L, \leq)$ it suffices to prove that the infimum of any two elements of L_1 in $(L, \leq)$ belongs to L_1. Consider $\psi(X_1), \psi(X_2)$ for $X_1, X_2 \in \mathcal{P}(A)$. Let us put $Y = X_1 \cap X_2$. If for instance $Y = X_1$, then $X_1 \subseteq X_2$ and $\psi(X_1) = \psi(X_1) \wedge \psi(X_2)$. Further let us suppose that $Y \neq X_1, X_2$ which also means that $X_1, X_2 \neq \emptyset$. If $Y = \emptyset$, then $\psi(X_1) \wedge \psi(X_2) = \bigwedge A$ by Proposition 6. Assume that $Y \neq \emptyset$ and denote $X_1' = X_1 \setminus Y$, $X_2' = X_2 \setminus Y$, $X = \{Y, X_1', X_2'\}$. The set X is join-independent in $(\mathcal{P}(A), \subseteq)$ by Proposition 9. It follows from $Y \cap X_1' = Y \cap X_2' = X_1' \cap X_2' = \bigcap X = \emptyset$ that (by Proposition 8) X is M-closed in $(\mathcal{P}(A), \subseteq)$. According our assumption, the set $\psi(X) = \{\psi(Y), \psi(X_1'), \psi(X_2')\}$ is M-closed in $(L, \leq)$. Thus $\psi(X_1) \wedge \psi(X_2) = \psi(Y \cup X_1') \wedge \psi(Y \cup X_2') = (\psi(Y) \vee \psi(X_1')) \wedge (\psi(Y) \vee \psi(X_2')) = \bigwedge P_{\psi(Y)} = \psi(Y)$. □

References

[1] Crawley, P., Dilworth, R. P.: Algebraic Theory of Lattices. *Englewood Cliffs*, 1973.

[2] Czédli, G., Huhn, A. P., Schmidt, E. T.: *Weakly independent sets in lattices.* Algebra Univers. **20** (1985), 194–196.

[3] Dlab, V.: *Lattice formulation of general algebraic dependence.* Czech. Math. Journal **20**, 95 (1970), 603–615.

[4] Grätzer, G.: General Lattice Theory. *Birkhäuser Verlag*, 1998.

[5] Machala, F.: *Join-independent and meet-independent sets in complete lattices.* Order **18** (2001), 269–274.

[6] Machala, F., Slezák, V.: *Lattice-inadmissible incidence structures.* Discuss. Math. (submitted).

[7] Slezák, V.: *On the special context of independent sets.* Discuss. Math., Gen. Algebra and Appl. **21** (2001), 115–122.

[8] Szász, G.: Introduction to Lattice Theory. *Akadémiai Kiadó*, Budapest, 1963.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 119–131

Estimation in Connecting Measurements with Constraints of Type II *

Jaroslav MAREK

Department of Mathematical Analysis and Applications of Mathematics,
Faculty of Science, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: marekj@aix.upol.cz

(Received February 12, 2004)

Abstract

This paper is a continuation of the paper [6]. It dealt with parameter estimation in connecting two–stage measurements with constraints of type I. Unlike the paper [6], the current paper is concerned with a model with additional constraints of type II binding parameters of both stages.

The article is devoted primarily to the computational aspects of algorithms published in [5] and its aim is to show the power of $\mathbf{H}^*$-optimum estimators.

The aim of the paper is to contribute to a numerical solution of the estimation problem in the two stage model, where constraints of type II occur in the second stage.

Key words: Two stage regression models, uncertainty of the type A and B, BLUE, **H**–optimum estimators.

2000 Mathematics Subject Classification: 62J05

1 Introduction

In mathematical models of measurements "the connectedness syndrome" is very often encountered. This paper is concerned with a two–stage measurement with an additional condition of type II on parameters of both stages. The value $\hat{\boldsymbol{\Theta}}$

*Supported by the Council of Czech Government J14/98: 153 100011.

of the parameter may be known prior to the measurement, and may or may not be changed as a result of measurement in the second stage.

In relation to the uncertainty in the estimator $\hat{\mathbf{\Theta}}$ the notion of "the uncertainty of type B" is introduced, compared to "the uncertainty of type A", which is linked to the uncertainty in measurement in the second stage. In case these uncertainties are not neglected, certain difficulties arise.

During the search for statistical solutions of connecting measurement we define $\mathcal{U}_{\boldsymbol{\beta}}$ of unbiased estimators $\tilde{\boldsymbol{\beta}}$ of the parameters $\boldsymbol{\beta}$ in the regular model, where we respect errors in connecting points; and class $\widetilde{\mathcal{U}}_{\boldsymbol{\beta}}$ of unbiased estimators $\tilde{\boldsymbol{\beta}}$ of parameter $\boldsymbol{\beta}$ satisfying the constraints between parameters of the first and the second stage.

The estimators from the class $\mathcal{U}_{\boldsymbol{\beta}}$ need not fulfil the constraints between parameters of the first and the second stages. There does not exist any jointly efficient estimator in the class $\mathcal{U}_{\boldsymbol{\beta}}$. Therefore we study estimators from the class $\widetilde{\mathcal{U}}_{\boldsymbol{\beta}}$ which minimize a linear functional of the covariance matrix of the estimator $\tilde{\boldsymbol{\beta}}$.

2 Estimation in model of connecting measurements with constraints of type II

Definition 1 The two stage model of the second stage measurement is

$$\begin{pmatrix} \hat{\mathbf{\Theta}} \\ \mathbf{Y} - \mathbf{D}\hat{\mathbf{\Theta}} \end{pmatrix} \sim_n \left(\begin{pmatrix} \mathbf{\Theta} \\ \mathbf{X}\boldsymbol{\beta} \end{pmatrix}, \begin{pmatrix} \mathbf{\Sigma}_{1,1}, & -\mathbf{\Sigma}_{1,1}\mathbf{D}' \\ -\mathbf{D}\mathbf{\Sigma}_{1,1}, & \mathbf{\Sigma}_{2,2} + \mathbf{D}\mathbf{\Sigma}_{1,1}\mathbf{D}' \end{pmatrix} \right),$$

The parametric space of the two stage model with constraints of the type II is

$$\underline{\mathbf{\Theta}} = \{(\mathbf{\Theta}', \boldsymbol{\beta}') : \mathbf{B}^*\boldsymbol{\beta} + \mathbf{C}^*\mathbf{\Theta} + \mathbf{G}\boldsymbol{\gamma} + \mathbf{a} = \mathbf{0}\}$$

where $\mathbf{B}^*, \mathbf{C}^*, \mathbf{G}$ are given matrices with dimensions $q \times k_2, q \times k_1, q \times k_3$ and $\mathbf{a}$ is given q-dimensional vector, such that $\mathcal{M}(\mathbf{C}^*) \subset \mathcal{M}(\mathbf{B}^*)$, and $r(\mathbf{B}^*) = q < k_2$.

The vector $\mathbf{\Theta}$ is the parametr of the first stage (connecting stage).

The vector $\boldsymbol{\beta}$ is the parametr of the second stage (connected stage).

The estimator $\hat{\mathbf{\Theta}}$ of the parameter $\mathbf{\Theta}$ is given from the first stage.

$\mathbf{D}$ is the incidence matrix, which identify parameters of connecting network, that were used in the course of measurement in the second stage,

$\mathbf{X}$ is known matrix of the connecting network,

$\mathbf{\Theta}$ and $\boldsymbol{\beta}$ are effective values of the parameter from the first and second stage,

$\mathbf{\Sigma}_{1,1}$ is the covariance matrix of the estimator $\hat{\mathbf{\Theta}}$, $\mathbf{\Sigma}_{2,2}$ is the covariance matrix of the observation vector $\mathbf{Y}$.

The notation $\boldsymbol{\xi} \sim_n (\boldsymbol{\mu}, \mathbf{\Sigma}_{2,2})$ means, that the n-dimensional vector parameter $\boldsymbol{\xi}$ has the mean value equal to $\boldsymbol{\mu}$ and its covariance is $\mathbf{\Sigma}_{2,2}$.

From the first stage the unbiased estimator $\hat{\boldsymbol{\Theta}}$ and its covariance matrix $\boldsymbol{\Sigma}_{1,1}$ are at our disposal only.

The aim is to determine an estimator of the parameter $\boldsymbol{\beta}$ on the basis of random vector $\mathbf{Y} - \mathbf{D}\hat{\boldsymbol{\Theta}}$, where $\mathbf{Y}$ is the observation vector of the second stage and on the basis of the estimator $\hat{\boldsymbol{\Theta}}$.

Lemma 1 *If $\boldsymbol{\Theta}$ in the model from Definition 1 is known, then the BLUE of the parameter $(\boldsymbol{\beta}', \boldsymbol{\gamma}')'$ is*

$$\begin{aligned}
\hat{\hat{\boldsymbol{\beta}}} = \Bigg(\mathbf{I} &- \left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)'\Bigg\{\left[\mathbf{B}^*\left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1} \\
&- \left[\mathbf{B}^*\left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\mathbf{G}\Big\{\mathbf{G}'\Big[\mathbf{B}^*\left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1} \\
&\times (\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}\mathbf{G}\Big\}^{-1}\mathbf{G}'\left[\mathbf{B}^*\left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\Bigg\}\mathbf{B}^*\Bigg) \\
&\times \left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}(\mathbf{Y} - \mathbf{D}\boldsymbol{\Theta}) \\
&- \left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)'\Bigg\{\left[\mathbf{B}^*\left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1} \\
&- \left[\mathbf{B}^*\left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\mathbf{G}\Big\{\mathbf{G}'\Big[\mathbf{B}^*\left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)' \\
&+ \mathbf{G}\mathbf{G}'\Big]^{-1}\mathbf{G}\Big\}^{-1}\mathbf{G}'\left[\mathbf{B}^*\left(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X}\right)^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\Bigg\}(\mathbf{a}^* + \mathbf{C}^*\boldsymbol{\Theta}),
\end{aligned}$$

and

$$\begin{aligned}
\hat{\hat{\boldsymbol{\gamma}}} &= -\Big\{\mathbf{G}'\left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\mathbf{G}\Big\}^{-1}\mathbf{G}' \\
&\times \left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{Y} + \mathbf{a}^* + \mathbf{C}^*\boldsymbol{\Theta}\right].
\end{aligned}$$

Their covariance matrices and cross covariance matrix are

$$\begin{aligned}
\mathrm{Var}(\hat{\hat{\boldsymbol{\beta}}}) &= (\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1} - (\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)'\big[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1} \\
&\quad \times (\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\big]^{-1}\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1} + (\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1} \\
&\quad \times (\mathbf{B}^*)'\left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1} \\
&\quad \times \mathbf{G}\Big\{\mathbf{G}'\left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\mathbf{G}\Big\}^{-1}\mathbf{G}' \\
&\quad \times \left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}, \\
\mathrm{cov}(\hat{\hat{\boldsymbol{\beta}}}, \hat{\hat{\boldsymbol{\gamma}}}) &= -(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)'\left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1} \\
&\quad \times \mathbf{G}\Big\{\mathbf{G}'\left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\mathbf{G}\Big\}^{-1}, \\
\mathrm{Var}(\hat{\hat{\boldsymbol{\gamma}}}) &= \Big\{\mathbf{G}'\left[\mathbf{B}^*(\mathbf{X}'\boldsymbol{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\right]^{-1}\mathbf{G}\Big\}^{-1} - \mathbf{I}.
\end{aligned}$$

Proof [5], section 3.

Definition 2 The estimator from Lemma 1 obtained under the condition $\mathbf{\Sigma}_{1,1} = \mathbf{0}$ ($\Rightarrow \mathrm{Var}(\mathbf{\Theta}) = \mathbf{0}$) is called the standard estimator if in this estimator the vector $\mathbf{\Theta}$ is substituted by $\hat{\mathbf{\Theta}}$.

Remark 1 If $\mathbf{\Theta}$ in Lemma 1 is substituted by $\hat{\mathbf{\Theta}}$, the standard estimator is obtained. Its covariance matrix is given by the following relationships.

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \mathrm{Var}[\mathbf{N}_1(\mathbf{Y} - \mathbf{D}\hat{\mathbf{\Theta}})] + \mathrm{Var}[\mathbf{N}_2(\mathbf{C}^*\hat{\mathbf{\Theta}} + \mathbf{a})]$$
$$+ \mathrm{cov}[\mathbf{N}_1(\mathbf{Y} - \mathbf{D}\hat{\mathbf{\Theta}}), \mathbf{N}_2(\mathbf{C}^*\hat{\mathbf{\Theta}} + \mathbf{a})] + \mathrm{cov}[\mathbf{N}_2(\mathbf{C}^*\hat{\mathbf{\Theta}} + \mathbf{a}), \mathbf{N}_1(\mathbf{Y} - \mathbf{D}\hat{\mathbf{\Theta}})]$$
$$= \mathbf{N}_1(\mathbf{\Sigma}_{2,2} + \mathbf{D}\mathbf{\Sigma}_{1,1}\mathbf{D}')\mathbf{N}_1' + \mathbf{N}_2\mathbf{C}^*\mathbf{\Sigma}_{1,1}(\mathbf{C}^*)'\mathbf{N}_2'$$
$$- \mathbf{N}_1\mathbf{D}\mathbf{\Sigma}_{1,1}(\mathbf{C}^*)'\mathbf{N}_2' - \mathbf{N}_2\mathbf{C}^*\mathbf{\Sigma}_{1,1}\mathbf{D}'\mathbf{N}_1',$$

where

$$\mathbf{N}_1 = \Big(\mathbf{I} - (\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)'\Big\{\Big[\mathbf{B}^*(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}$$
$$- \Big[\mathbf{B}^*(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}\mathbf{G}\mathbf{G}'\Big[\mathbf{B}^*(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}\mathbf{G}'$$
$$\times \Big[\mathbf{B}^*(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}(\mathbf{B}^*)'(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1},$$
$$\mathbf{N}_2 = -(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)'\Big\{\Big[\mathbf{B}^*(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}$$
$$- \Big[\mathbf{B}^*(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}\mathbf{G}$$
$$\times \{\mathbf{G}'\Big[\mathbf{B}^*(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}\mathbf{G}\}^{-1}\mathbf{G}'$$
$$\times \Big[\mathbf{B}^*(\mathbf{X}'\mathbf{\Sigma}_{2,2}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}'\Big]^{-1}\Big\}.$$

Theorem 1 *In the model*

$$\mathbf{Y} - \mathbf{D}\hat{\mathbf{\Theta}} \sim_n (\mathbf{X}\boldsymbol{\beta}, \mathbf{\Sigma}_{2,2} + \mathbf{D}\mathbf{\Sigma}_{1,1}\mathbf{D}'), \qquad \mathbf{a}^* + \mathbf{C}^*\mathbf{\Theta} + \mathbf{B}^*\boldsymbol{\beta} + \mathbf{G}\boldsymbol{\gamma} = \mathbf{0},$$

the class of all unbiased linear estimators of $\binom{\boldsymbol{\beta}}{\boldsymbol{\gamma}}$ *based on the vectors* $\hat{\mathbf{\Theta}}$ *and* $\mathbf{Y} - \mathbf{D}\hat{\mathbf{\Theta}}$ *is*

$$\mathcal{U}_{\beta,\gamma} = \left\{ \begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\gamma}} \end{pmatrix} = \begin{pmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{K}_1, & \mathbf{K}_2 \\ \mathbf{K}_3, & \mathbf{K}_4 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{\Theta}} \\ \mathbf{Y} - \mathbf{D}\hat{\mathbf{\Theta}} \end{pmatrix} \right\},$$

where

$$\begin{pmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ -\mathbf{G}^- \end{pmatrix} \mathbf{a}^* + \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_3 \end{pmatrix} (\mathbf{I} - \mathbf{G}\mathbf{G}^-)\mathbf{a}^*,$$

$$\begin{pmatrix} \mathbf{K}_1, & \mathbf{K}_2 \\ \mathbf{K}_3, & \mathbf{K}_4 \end{pmatrix} = \begin{pmatrix} \mathbf{0}, & \mathbf{X}^- \\ -\mathbf{G}^-\mathbf{C}^*, & -\mathbf{G}^-\mathbf{B}^*\mathbf{X}^- \end{pmatrix}$$
$$+ \begin{pmatrix} \mathbf{Z}_1, & \mathbf{Z}_2 \\ \mathbf{Z}_3, & \mathbf{Z}_4 \end{pmatrix} \begin{pmatrix} (\mathbf{I}-\mathbf{G}\mathbf{G}^-)\mathbf{C}^*, & (\mathbf{I}-\mathbf{G}\mathbf{G}^-)\mathbf{B}^*\mathbf{X}^- \\ \mathbf{0}, & \mathbf{I}-\mathbf{X}\mathbf{X}^- \end{pmatrix},$$

where $\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{Z}_4$ *are arbitrary matrices with suitable dimensions.*

The covariance matrix of the estimator $\binom{\tilde{\boldsymbol{\beta}}}{\tilde{\boldsymbol{\gamma}}} \in \mathcal{U}_{\beta,\gamma}$ *is*

$$\operatorname{Var}\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\gamma}} \end{pmatrix} = \begin{pmatrix} \mathbf{K}_1, & \mathbf{K}_2 \\ \mathbf{K}_3, & \mathbf{K}_4 \end{pmatrix} \begin{pmatrix} \boldsymbol{\Sigma}_{1,1}, & -\boldsymbol{\Sigma}_{1,1}\mathbf{D}' \\ -\mathbf{D}\boldsymbol{\Sigma}_{1,1}, & \boldsymbol{\Sigma}_{2,2}+\mathbf{D}\boldsymbol{\Sigma}_{1,1}\mathbf{D}' \end{pmatrix} \begin{pmatrix} \mathbf{K}_1, & \mathbf{K}_2 \\ \mathbf{K}_3, & \mathbf{K}_4 \end{pmatrix}'.$$

Proof [5], section 3.

Lemma 2 *Let in Lemma 1* $\boldsymbol{\Theta}$ *be substituted by* $\hat{\boldsymbol{\Theta}}$. *Then such estimator (it is usually used in practice) belong to the class* $\mathcal{U}_{\beta,\gamma}$.

Proof [5], section 3.

Theorem 2 *The class* $\tilde{\mathcal{U}_{\beta,\gamma}}$ *of all linear unbiased estimators which in addition satisfy the constraints*

$$\mathbf{a}^* + \mathbf{C}^*\hat{\boldsymbol{\Theta}} + \mathbf{B}^*\tilde{\boldsymbol{\beta}} + \mathbf{G}\tilde{\boldsymbol{\gamma}} = \mathbf{0},$$

is given by such a choice of the matrices $\mathbf{Z}_1, \ldots, \mathbf{Z}_4$, *in Theorem 1, which satisfy the following equation*

$$\begin{pmatrix} \mathbf{Z}_1, & \mathbf{Z}_2 \\ \mathbf{Z}_3, & \mathbf{Z}_4 \end{pmatrix} = (\mathbf{B}^*, \mathbf{G})^-[-(\mathbf{I}-\mathbf{G}\mathbf{G}^-), \mathbf{0}] \begin{pmatrix} \mathbf{I}-\mathbf{G}\mathbf{G}^-, & \mathbf{0} \\ \mathbf{0}, & \mathbf{I}-\mathbf{X}\mathbf{X}^- \end{pmatrix}$$
$$+ \begin{pmatrix} \mathbf{W}_1, & \mathbf{W}_2 \\ \mathbf{W}_3, & \mathbf{W}_4 \end{pmatrix} - (\mathbf{B}^*, \mathbf{G})^-(\mathbf{B}^*, \mathbf{G}) \begin{pmatrix} \mathbf{W}_1, & \mathbf{W}_2 \\ \mathbf{W}_3, & \mathbf{W}_4 \end{pmatrix} \begin{pmatrix} \mathbf{I}-\mathbf{G}\mathbf{G}^-, & \mathbf{0} \\ \mathbf{0}, & \mathbf{I}-\mathbf{X}\mathbf{X}^- \end{pmatrix},$$

where the matrices $\mathbf{W}_1, \ldots, \mathbf{W}_4$ *are arbitrary.*

Proof [5], section 3.

3 H*-optimum estimator for constraints II

Definition 3 Let $\mathbf{H}^*$ be a given $(k+l)\times(k+l)$ positive semidefinite matrix. The estimator $\binom{\tilde{\boldsymbol{\beta}}}{\tilde{\boldsymbol{\gamma}}}$ from $\tilde{\mathcal{U}}_{\beta,\gamma}$ is said to be $\mathbf{H}^*$-optimum if it minimizes the value

$$\operatorname{Tr}\left[\mathbf{H}^* \operatorname{Var}\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\gamma}} \end{pmatrix}\right], \begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\gamma}} \end{pmatrix} \in \tilde{\mathcal{U}}_{\beta,\gamma}.$$

Theorem 3 *An estimator* $\binom{\tilde{\boldsymbol{\beta}}}{\tilde{\boldsymbol{\gamma}}}$ *is* $\mathbf{H}^*$*-optimum if the matrices* $\mathbf{W}_1, \mathbf{W}_2$, $\mathbf{W}_3, \mathbf{W}_4$ *(Theorem 2) are solution of the equation*

$$\left\{\mathbf{I} - (\mathbf{B}^*, \mathbf{G})'\left[(\mathbf{B}^*, \mathbf{G})^-\right]'\right\}\mathbf{H}^*\left[\mathbf{I} - (\mathbf{B}^*, \mathbf{G})^-(\mathbf{B}^*, \mathbf{G})\right]\mathbf{WSTS}' =$$
$$= -\left\{\mathbf{I} - (\mathbf{B}^*, \mathbf{G})'\left[(\mathbf{B}^*, \mathbf{G})^-\right]'\right\}\mathbf{H}^*(\mathbf{RTS}' + \mathbf{ASTS}'),$$

where

$$\mathbf{A} = (\mathbf{B}^*, \mathbf{G})^- \big[-(\mathbf{I} - \mathbf{G}\mathbf{G}^-, \mathbf{0}\big] \begin{pmatrix} \mathbf{I} - \mathbf{G}\mathbf{G}^-, & \mathbf{0} \\ \mathbf{0}, & \mathbf{I} - \mathbf{X}\mathbf{X}^- \end{pmatrix},$$

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_1, & \mathbf{W}_2 \\ \mathbf{W}_3, & \mathbf{W}_4 \end{pmatrix}, \qquad \mathbf{R} = \begin{pmatrix} \mathbf{0}, & \mathbf{X}^- \\ -\mathbf{G}^-\mathbf{C}^*, & -\mathbf{G}^-\mathbf{B}^*\mathbf{X}^- \end{pmatrix},$$

$$\mathbf{S} = \begin{pmatrix} (\mathbf{I} - \mathbf{G}\mathbf{G}^-)\mathbf{C}^*, & (\mathbf{I} - \mathbf{G}\mathbf{G}^-)\mathbf{B}^*\mathbf{X}^- \\ \mathbf{0}, & \mathbf{I} - \mathbf{X}\mathbf{X}^- \end{pmatrix},$$

$$\mathbf{T} = \mathrm{Var} \begin{pmatrix} \hat{\boldsymbol{\Theta}} \\ \mathbf{Y} - \mathbf{D}\hat{\boldsymbol{\Theta}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{1,1}, & -\boldsymbol{\Sigma}_{1,1}\mathbf{D}' \\ -\mathbf{D}\boldsymbol{\Sigma}_{1,1}, & \boldsymbol{\Sigma}_{2,2} + \mathbf{D}\boldsymbol{\Sigma}_{1,1}\mathbf{D}' \end{pmatrix}.$$

Proof [5], section 4.

Remark 2 Since the matrices $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, \mathbf{W}_4$ of the $\mathbf{H}^*$-optimum estimator are functions of the matrix $\mathbf{H}^*$, the joint efficient estimator does not exist in the class $\tilde{\mathcal{U}}_{\beta,\gamma}$.

4 Numerical studies—constraints type II

In this part we will concentrate on a numerical calculation of the estimator of parameters. In all following examples we need to construct a condition expressing a relation between parameters of the first and the second stages. From this condition we can always construct a vector function $\mathbf{g}$ of parameter β and Θ where $\mathbf{g}(\beta, \Theta, \gamma) = \mathbf{0}$. We apply the Taylor expansion at point (β_0, Θ_0) to this function. So for estimators of parameters we get the condition

$$\mathbf{g}(\boldsymbol{\beta}, \boldsymbol{\Theta}, \gamma) = \mathbf{g}(\beta_0, \Theta_0, \gamma_0) + \mathbf{C}\delta\theta + \mathbf{B}\delta\beta + \mathbf{G}\delta\gamma = \mathbf{0}.$$

Example 1 Let us consider the point A_1 from the first stage with the plane coordinates $(\Theta_1,\ \Theta_2)$, that were measured as $(\hat{\Theta}_1, \hat{\Theta}_2) = (59999.91, 41339.81)$. The accuracy of measurement was given by the dispersion $\omega_1^2 = 0.04^2$.

In the second stage we will assume the same dispersion $\omega_1^2 = 0.04^2$ for the measured coordinates $(y_1, y_2), \ldots, (y_7, y_8) = (54999.95, 40000.04, 49999.94, 41339.70, 54999.89, 60000.01, 65000.05, 49999.88)$ of the points $P_i = (\beta_{2i-1}, \beta_{2i})$ for $i = 1, 2, 3, 4$.

The aim is to make the estimator of the coordinates of the points P_1, P_2, P_3 and P_4 more accurate under the constraint that all these points together with the point A_1 are located on a circle, with a radius γ_3 and a center $[\gamma_1, \gamma_2]$ unknown.

Our constraints are

$$\begin{aligned} (\theta_1 - \gamma_1)^2 + (\theta_2 - \gamma_2)^2 - \gamma_3^2 &= 0\,, \\ (\beta_1 - \gamma_1)^2 + (\beta_2 - \gamma_2)^2 - \gamma_3^2 &= 0\,, \\ (\beta_3 - \gamma_1)^2 + (\beta_4 - \gamma_2)^2 - \gamma_3^2 &= 0\,, \\ (\beta_5 - \gamma_1)^2 + (\beta_6 - \gamma_2)^2 - \gamma_3^2 &= 0\,, \\ (\beta_7 - \gamma_1)^2 + (\beta_8 - \gamma_2)^2 - \gamma_3^2 &= 0\,. \end{aligned}$$

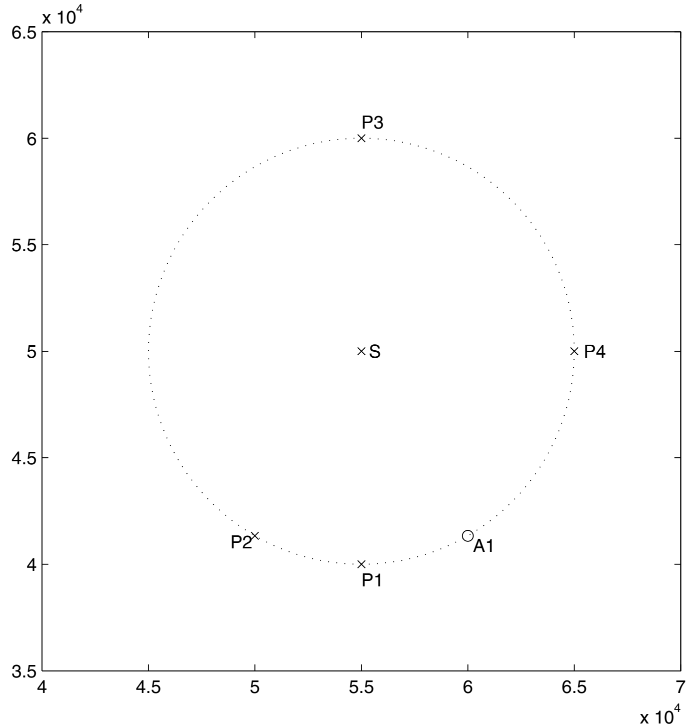

Figure 1: Situation (in S-JTSK)

In our linearized model we will determine numerically the estimator and the covariance matrix according to Lemma 1:

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} 54999.95 \\ 40000.14 \\ 50000.00 \\ 41339.80 \\ 54999.89 \\ 60000.11 \\ 64999.83 \\ 49999.88 \end{pmatrix} \text{ and } \hat{\boldsymbol{\gamma}} = \begin{pmatrix} 54999.84 \\ 50000.01 \\ 9999.98 \end{pmatrix} .$$

After that we will numerically determine $\mathbf{H}^*$–optimum estimator from Theorem 2 and 3 for the matrix

$$\mathbf{H}_1{}^* = \begin{pmatrix} 1&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0 \end{pmatrix} \text{ are } \tilde{\boldsymbol{\beta}} = \begin{pmatrix} 54999.95 \\ 40000.04 \\ 49999.91 \\ 41339.65 \\ 54999.89 \\ 60000.08 \\ 64999.69 \\ 49999.88 \end{pmatrix} \text{ and } \tilde{\boldsymbol{\gamma}} = \begin{pmatrix} 54999.67 \\ 50000.06 \\ 10000.02 \end{pmatrix}.$$

By chosen matrix $\mathbf{H}_1{}^*$ minimizing data errors in the process estimation of the vector $\tilde{\boldsymbol{\beta}}$ we got better estimator of the parameter β in comparison with the standard estimator $\hat{\boldsymbol{\beta}}$. It follows from the fact that for the chosen matrix $\mathbf{H}^*$ is $\mathrm{Tr}(\mathbf{H}^* \,\mathrm{Var}(\tilde{\boldsymbol{\beta}})) = 2.17 \cdot 10^{-3} < 2.66 \cdot 10^{-3} = \mathrm{Tr}(\mathbf{H}^* \,\mathrm{Var}(\hat{\boldsymbol{\beta}}))$.

Let us study the proportion accuracy of the standard estimator $\hat{\boldsymbol{\beta}}$ and the $\mathbf{H}_i{}^*$-optimum estimator $\tilde{\boldsymbol{\beta}}$ for $i = 2, 3, 4$. We will not determine the estimators from now, but we will only study the trace of the covariance matrix $\mathrm{Tr}(\mathbf{H}\,\mathrm{Var}(\tilde{\boldsymbol{\beta}}))$ for comparing it with the above mentioned $\mathrm{Tr}(\mathbf{H}\,\mathrm{Var}(\hat{\boldsymbol{\beta}}))$.

We get $\mathrm{Tr}(\mathbf{H}_2{}^* \,\mathrm{Var}(\tilde{\boldsymbol{\beta}})) = 2.59 \cdot 10^{-3} < \mathrm{Tr}(\mathbf{H}_2{}^* \,\mathrm{Var}(\hat{\boldsymbol{\beta}})) = 2.66 \cdot 10^{-3}$ for matrix

$$\mathbf{H}_2{}^* = \begin{pmatrix} 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0 \end{pmatrix},$$

we get $\mathrm{Tr}(\mathbf{H}_3{}^* \,\mathrm{Var}(\tilde{\boldsymbol{\beta}})) = 3.09 \cdot 10^{-3} < \mathrm{Tr}(\mathbf{H}_3{}^* \,\mathrm{Var}(\hat{\boldsymbol{\beta}})) = 3.21 \cdot 10^{-3}$ for matrix

$$\mathbf{H}_3{}^* = \begin{pmatrix} 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0\\ 0&0&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0 \end{pmatrix},$$

and we get $\mathrm{Tr}(\mathbf{H}_4{}^* \,\mathrm{Var}(\tilde{\boldsymbol{\beta}})) = 2.72 \cdot 10^{-3} < \mathrm{Tr}(\mathbf{H}_4{}^* \,\mathrm{Var}(\hat{\boldsymbol{\beta}})) = 3.49 \cdot 10^{-3}$ for matrix

$$\mathbf{H}_4{}^* = \begin{pmatrix} 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0 \end{pmatrix}.$$

It is evident that $\mathrm{Tr}(\mathbf{H}_i{}^* \,\mathrm{Var}(\tilde{\boldsymbol{\beta}})) < \mathrm{Tr}(\mathbf{H}_i{}^* \,\mathrm{Var}(\hat{\boldsymbol{\beta}}))$ for $i = 1, \dots 4$.

Now let us study the proportion of this values for different covariance matrices $\boldsymbol{\Sigma}_{1,1}$ and $\boldsymbol{\Sigma}_{2,2}$. In other numerical calculations we choose the matrix $\boldsymbol{\Sigma}_{1,1}$ as the fixed one and we change the matrix $\boldsymbol{\Sigma}_{2,2}$ by the multiplication by the number k.

The proportions in dependence on k are shown in the following table and graph.

The proportion $\mathrm{Tr}(\mathbf{H}_i{}^* \,\mathrm{Var}(\tilde{\boldsymbol{\beta}}))$ and $\mathrm{Tr}(\mathbf{H}_i{}^* \,\mathrm{Var}(\hat{\boldsymbol{\beta}}))$				
k	$i = 1, \mathbf{H}_1{}^*$	$i = 2, \mathbf{H}_2{}^*$	$i = 3, \mathbf{H}_3{}^*$	$i = 4, \mathbf{H}_4{}^*$
100	99.99 %	100.00 %	100.00 %	99.99%
64	99.98 %	100.00 %	100.00 %	99.98%
50	99.97 %	100.00 %	100.00 %	99.97%
25	99.90 %	99.99 %	99.98 %	99.87%
16	99.77 %	99.97 %	99.96 %	99.70%
9	99.33 %	99.92 %	99.89 %	99.12%
5	98.14 %	99.78 %	99.67 %	97.57%
4	97.30 %	99.67 %	99.52 %	96.50%
3	95.74 %	99.47 %	99.22 %	94.54%
2	92.29 %	98.99 %	98.52 %	90.27%
1	81.73 %	97.24 %	96.00 %	77.81%
1/2	65.01 %	93.41 %	90.64 %	59.52%
1/4	45.48 %	86.18 %	81.00 %	39.92%
1/10	23.72 %	69.58 %	61.02 %	19.92%
1/16	16.02 %	58.29 %	48.90 %	13.26%
1/25	10.77 %	46.87 %	37.66 %	8.83%
1/50	5.64 %	30.34 %	22.98 %	4.58%
1/64	4.45 %	25.34 %	18.86 %	3.60%
1/100	2.89 %	17.79 %	12.91 %	2.33%
1/400	0.74 %	5.11 %	3.55 %	0.59%

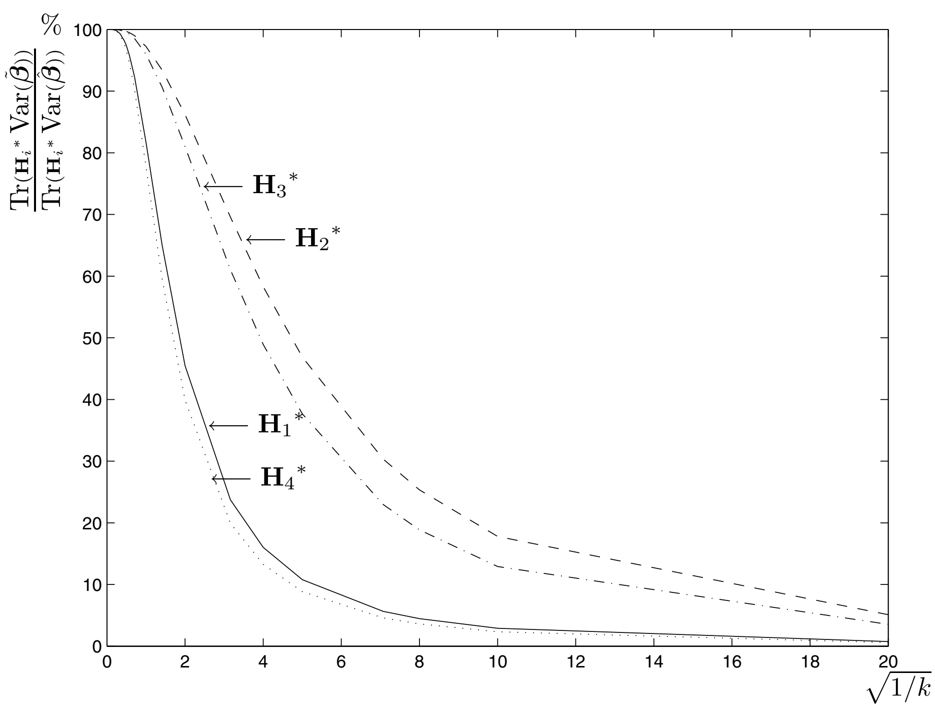

Figure 2: The proportion $\mathrm{Tr}(\mathbf{H}_i{}^* \,\mathrm{Var}(\tilde{\boldsymbol{\beta}}))$ and $\mathrm{Tr}(\mathbf{H}_i{}^* \,\mathrm{Var}(\hat{\boldsymbol{\beta}}))$

Example 2 We have at our disposal the coordinates of the points A_1, A_2, A_3, A_4, A_5 that are given from the first stage – from the connecting measurement.

All the angles $\angle y_1 = (A_2A_1P_1)$, $\angle y_2 = (A_1P_1P_2)$, $\angle y_3 = (P_1P_2P_3)$, $\angle y_4 = (P_2P_3P_4)$, $\angle y_5 = (P_3P_4A_5)$, $\angle y_6 = (P_4A_5A_4)$, $\angle y_7 = (P_1A_3P_2)$ and distances $y_8 = A_1P_1$, $y_9 = P_1P_2$, $y_{10} = P_2P_3$, $y_{11} = P_3P_4$, $y_{12} = P_4A_5$ were measured in the second stage—in the connecting stage.

The aim is to find an estimator for the plane coordinates $(\beta_1, \beta_2), \dots, (\beta_7, \beta_8)$ of the points P_1, P_2, P_3 and P_4 from the second stage, in such a way so as the distance between the points P_1 and P_3 would be determined as accurately as possible.

Values of plane-coordinates and distances will be given in meters, values of angles will be given in radians.

The accuracy of measurement is given by the dispersion or covariance matrices. We suppose that the points from the first stage are determined with the dispersion 0.06^2 m. Measurement of angles in the second stage was performed with the standard deviation $\omega_a = 10/206265$. Measurement of distances in the second stage was performed with the standard deviation $\omega_d = 0.005$ m.

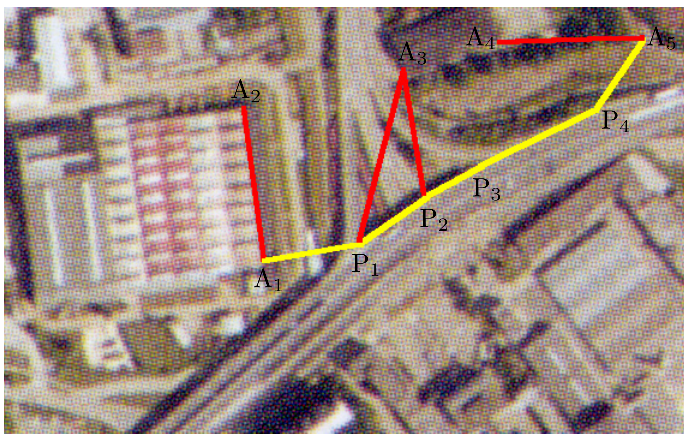

Figure 3: The aerial photograph of the Tovární Street, Olomouc

We carry out numerical studies in this example for the plane coordinates of points A_i

	Y	X	
A_1	543330,15	1121488,64	corner of the assembly hall
A_2	544347,49	1121390,53	corner of the assembly hall
A_3	544246,27	1121374,30	corner of the assembly hall
A_4	544187,59	1121350,71	corner of the assembly hall
A_5	544101,01	1121357,58	plastic point

and for measured values from the second stage

$y_1 = 1.6091000$, $y_2 = 2.7466880$, $y_3 = 3.2469781$, $y_4 = 3.2134906$,
$y_5 = 2.5395759$, $y_6 = 1.1120582$, $y_7 = 4.4991793$,
$y_8 = 56.515$, $y_9 = 50.889$, $y_{10} = 43.064$, $y_{11} = 80.486$, $y_{12} = 41.524$.

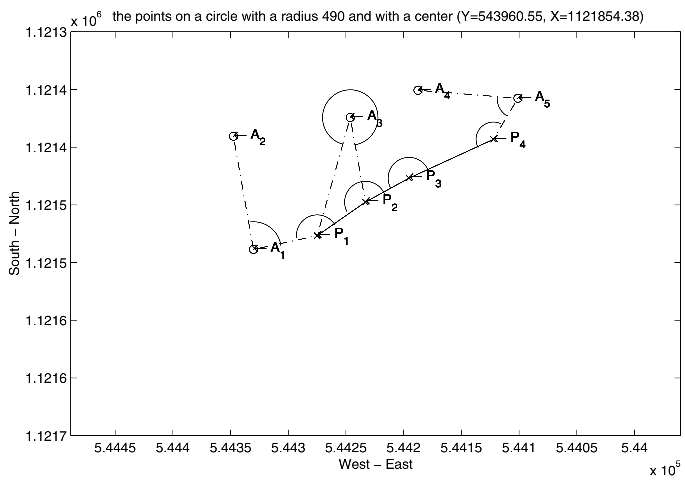

Figure 4: Situation (S-JTSK)

Now we make nonlinear model of our example $\mathbf{Y} = f(\boldsymbol{\Theta}, \boldsymbol{\beta}) + \varepsilon$.

$$\mathbf{Y} = \begin{pmatrix} f_1(\Theta,\beta) \\ f_2(\Theta,\beta) \\ f_3(\Theta,\beta) \\ f_4(\Theta,\beta) \\ f_5(\Theta,\beta) \\ f_6(\Theta,\beta) \\ f_7(\Theta,\beta) \\ f_8(\Theta,\beta) \\ f_9(\Theta,\beta) \\ f_{10}(\Theta,\beta) \\ f_{11}(\Theta,\beta) \\ f_{12}(\Theta,\beta) \end{pmatrix} = \begin{bmatrix} \arctan(\frac{\beta_2-\theta_2}{\beta_1-\Theta_1}) - \arctan(\frac{\Theta_4-\Theta_2}{\Theta_3-\Theta_1}) \\ \pi - \arctan(\frac{\beta_4-\beta_2}{\beta_3-\beta_1}) - \arctan(\frac{\Theta_2-\beta_2}{\Theta_1-\beta_1}) \\ \pi - \arctan(\frac{\beta_6-\beta_4}{\beta_5-\beta_3}) + \arctan(\frac{\beta_2-\beta_4}{\beta_1-\beta_3}) \\ \pi - \arctan(\frac{\beta_8-\beta_6}{\beta_7-\beta_5}) + \arctan(\frac{\beta_4-\beta_6}{\beta_3-\beta_5}) \\ \pi - \arctan(\frac{\Theta_{10}-\beta_8}{\Theta_9-\beta_7}) + \arctan(\frac{\beta_6-\beta_8}{\beta_5-\beta_7}) \\ -\arctan(\frac{\Theta_8-\theta_{10}}{\Theta_7-\Theta_9}) + \arctan(\frac{\beta_8-\Theta_{10}}{\beta_7-\Theta_9}) \\ \pi - \arctan(\frac{\beta_2-\Theta_6}{\beta_1-\Theta_5}) + \arctan(\frac{\beta_4-\Theta_6}{\beta_3-\Theta_5}) \\ \sqrt{{\beta_1}^2 - 2\,\beta_1\,\Theta_1 + {\Theta_1}^2 + {\beta_2}^2 - 2\,\beta_2\,\Theta_2 + {\Theta_2}^2} \\ \sqrt{{\beta_3}^2 - 2\,\beta_3\,\beta_1 + {\beta_1}^2 + {\beta_4}^2 - 2\,\beta_4\,\beta_2 + {\beta_2}^2} \\ \sqrt{{\beta_5}^2 - 2\,\beta_5\,\beta_3 + {\beta_3}^2 + {\beta_6}^2 - 2\,\beta_6\,\beta_4 + {\beta_4}^2} \\ \sqrt{{\beta_7}^2 - 2\,\beta_7\,\beta_5 + {\beta_5}^2 + {\beta_8}^2 - 2\,\beta_8\,\beta_6 + {\beta_6}^2} \\ \sqrt{{\beta_7}^2 - 2\,\beta_7\,\Theta_9 + {\Theta_9}^2 + {\beta_8}^2 - 2\,\beta_8\,\Theta_{10} + {\Theta_{10}}^2} \end{bmatrix}$$

Points from the second stage are situated on a circle—additional constraints are

$$\begin{aligned} g_1(\Theta,\beta,\beta) &= (\beta_1-\gamma_1)^2 + (\beta_2-\gamma_2)^2 - \gamma_3^2 \,=\, 0\,, \\ g_2(\Theta,\beta,\beta) &= (\beta_3-\gamma_1)^2 + (\beta_4-\gamma_2)^2 - \gamma_3^2 \,=\, 0\,, \\ g_3(\Theta,\beta,\beta) &= (\beta_5-\gamma_1)^2 + (\beta_6-\gamma_2)^2 - \gamma_3^2 \,=\, 0\,, \\ g_4(\Theta,\beta,\beta) &= (\beta_7-\gamma_1)^2 + (\beta_8-\gamma_2)^2 - \gamma_3^2 \,=\, 0\,. \end{aligned}$$

Now we need to use the Taylor expansion $\mathbf{Y} = f_0 + \mathbf{B}\delta\beta + \mathbf{D}\delta\Theta + \mathbf{G}\delta\gamma = \mathbf{0}$, where the matrices $\mathbf{B} = \frac{\partial f_i(\Theta^0,\beta^0)}{\partial\beta'}$, $\mathbf{D} = \frac{\partial f_i(\Theta^0,\beta^0)}{\partial\Theta'}$, $\mathbf{G} = \frac{\partial g_j(\Theta^0,\beta^0,\gamma^0)}{\partial\gamma'}$ and $f_0 = \mathbf{Y}(\Theta^0, \beta^0)$.

In the linearized model, we calculate the estimator $\hat{\boldsymbol{\beta}}$ and the distance estimator $\hat{y}_{10}$.

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} 544274.921 \\ 1121476.680 \\ 544233.140 \\ 1121447.619 \\ 544195.417 \\ 1121426.860 \\ 544122.290 \\ 1121393.236 \end{pmatrix}, \qquad \hat{y}_{10} = 43.0569.$$

Next by the same procedure as in the preceding example we calculate, by Lemma 2.2 , the $\mathbf{H}^*$-optimum estimator $\tilde{\boldsymbol{\beta}}$. The matrix $\mathbf{H}^*$ of the type

$$\mathbf{H}^* = \mathbf{p}\mathbf{p}', \quad \mathbf{p}' = \frac{\partial\sqrt{(\beta_5 - \beta_3)^2 + (\beta_6 - \beta_4)^2}}{\partial\beta'}$$

is chosen in such a way, so that the resulting estimator would be optimal for determining the distance between the points P_2 and P_3. We arrive at the estimator

$$\tilde{\boldsymbol{\beta}} = \begin{pmatrix} 544274.916 \\ 1121476.678 \\ 544233.150 \\ 1121447.604 \\ 544195.416 \\ 1121426.854 \\ 544122.286 \\ 1121393.239 \end{pmatrix}, \qquad \widetilde{y}_{10} = 43.0637.$$

Let us conclude with the comparison of the resulting estimators of distance between P_2 and P_3. The mesurement distance y_{10} was 43.064 m, the distance determined by the standard estimator $\hat{y}_{10}$ was 43.0569 m and by the $\mathbf{H}^*$-optimum estimator $\widetilde{y}_{10} = 43.0637$ m (the difference between estimators is 6.8 mm).

References

[1] Kubáček, L.: *Two stage regression models with constraints.* Math. Slovaca **43** (1993), 643–658.

[2] Kubáček, L.: *Štatistické modely pripojovacích meraní.* In. Celoslovenský seminár "Modernizácia geodetických základov Slovenska", zborník prednášok (ed. M. Petrovič), VÚGK Bratislava, 1993, 28–40.

[3] Kubáček, L., Kubáčková, L.: *Dvouetapové sítě s podmínkami typu I a II.* In: Sborník příspěvků spolupracovníků k devadesátinám pana profesora Josefa Vykutila. (Ed. D. Dušátko), Praha–Brno, Vojenský zeměpisný ústav Praha 2002, 58–72.

[4] Kubáčková, L.: Foundations of Experimental Data Analysis. *CRC-Press, Boca Raton–Ann Arbor–London–Tokyo*, 1992.

[5] Kubáček, L., Marek, J.: *Partial optimum estimator in two stage regression model with constraints and a problem of equivalence.* Math. Slovaca **54** (2004), (to appear).

[6] Marek, J.: *Estimation in connecting measurement.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **43** (2003), 69–86.

[7] Marek, J.: *Statistical view on measurement in Czech and world literature.* (submitted to proceedings of Datastat'03).

[8] Kubáček, L., Kubáčková, L.: Statistics and Metrology (in Czech). Olomouc. *Publishing House of Palacký University*, 2000.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 133–136

A Characterization of Almost Continuity and Weak Continuity

Chrisostomos PETALAS[1], Theodoros VIDALIS[2]

Department of Mathematics
University of Ioannina, 451 10 Ioannina, Greece
e-mail: [1]*cpetalas@cc.uoi.gr*
[2]*tvidalis@cc.uoi.gr*

(Received May 18, 2004)

Abstract

It is well known that a function f from a space X into a space Y is continuous if and only if, for every set K in X the image of the closure of K under f is a subset of the closure of the image of it.

In this paper we characterize almost continuity and weak continuity by proving similar relations for the subsets K of X.

Key words: Almost continuous function, weakly continuous function.

2000 Mathematics Subject Classification: 54C10

1 Introduction and notations

The term almost continuous function is defined in different ways by several authors [3, 4, 5, 7]. In this paper we adopt the following definition due to Singal and Singal [7].

Definition 1 *A function $f : X \to Y$ is said to be almost continuous if for each point $x \in X$ and each open set V in Y containing $f(x)$, there exists an open set U in X containing x, such that $f(U) \subset \overline{V}^0$.*

The following definition of weak continuity is due to N. Levine [2].

Definition 2 *A mapping $f : X \to Y$ is said to be weakly continuous if for each point $x \in X$ and each open set V in Y containing $f(x)$, there exists an open set U in X containing x such that $f(U) \subset \overline{V}$.*

It is well known the following:

Proposition 1 *A function $f : X \to Y$ is continuous iff for every set K in X we have*

$$f(\overline{K}) \subset \overline{f(K)}. \tag{1}$$

Comparing definitions 1, 2 and this proposition, it is natural to look for similar to (1) relations for the other two kinds of continuity.

The aim of this paper is to prove two theorems which give such relations.

We will use the following definitions:

A set A is said to be regularly open if $A = \overline{A}^0$. Since $\overline{\overline{A}^0}^0 = \overline{A}^0$, we have that for every set A, the set $\overline{A}^0$ is regularly open.

For a set $B \subset X$ we denote by $\overline{B}^{reg}$, the regular closure of B, that is the set of points of X for which, every regularly open set which contains x, intersects B.

2 Almost continuous functions

From definitions 1, 2 we have that continuity implies almost continuity implies weak continuity, see also [1, 6].

If Y is a regular space then all these kinds of continuity coincide.

The analogous relation to (1) for the almost continuity is given by the following:

Theorem 1 *A function $f : X \to Y$ is almost continuous iff for every $K \subset X$ we have*

$$f(\overline{K}) \subset \overline{f(K)}^{reg}.$$

Proof ($\Leftarrow$) Suppose, in contrary, that f is not almost continuous. Then there exists an open set V containing $f(x)$ such that for every open set U containing x we have

$$f(U) \not\subset \overline{V}^0.$$

Therefore in every such U, there exists a point y_U such that $f(y_U) \notin \overline{V}^0$. These points y_U define a net, which converges to x, and such that $f(y_U) \notin \overline{V}^0$, for every U.

Let $K = \{y_U : U \text{ is a neighborhood of } x, f(y_U) \notin \overline{V}^0\}$. Since $x \in \overline{K}$, it follows that $f(x) \in f(\overline{K})$. Now there exists a regularly open set containing $f(x)$, namely the set $\overline{V}^0$, which does not intersect $f(K)$, i.e.

$$f(x) \notin \overline{f(K)}^{reg}.$$

So it follows that,

$$f(\overline{K}) \not\subset \overline{f(K)}^{reg},$$

a contradiction.

($\Rightarrow$) Suppose, in contrary, that there exists a subset K of X with $f(\overline{K}) \not\subset \overline{f(K)}^{reg}$.

Then we can find a point $x \in \overline{K}$ such that $f(x) \notin \overline{f(K)}^{reg}$. It follows that there is a regularly open set $V = \overline{V}^0$ containing $f(x)$ with

$$V \cap f(K) = \emptyset. \tag{2}$$

Let U be an open neighborhood of x. Since $x \in \overline{K}$ it follows that $U \cap K \neq \emptyset$, which gives that

$$f(U) \cap f(K) \neq \emptyset. \tag{3}$$

By (2) $f(K) \subset V^c$, so (3) imply that

$$f(U) \cap V^c \neq \emptyset$$

i.e.

$$f(U) \not\subset V = \overline{V}^0.$$

Therefore f is not almost continuous, a contradiction. □

3 Weak continuity

The corresponding to (1) characterization of weakly continuous functions is given by the following:

Theorem 2 *A function $f : X \to Y$ is weakly continuous if and only if for every subset K of X we will have*

$$f(\overline{K}) \subset \cap\{\overline{U(f(K))} : U(f(K)) \text{ is an open subset of } Y \text{ containing } f(K)\}. \tag{4}$$

Proof ($\Leftarrow$) Suppose that f is not weakly continuous. Then there exists an x and an open subset V of Y containing $f(x)$, such that for every open subset U of X containing x we have

$$f(U) \not\subset \overline{V}.$$

Choose a y_U in every member U of an open base at x, such that

$$f(y_U) \notin \overline{V}. \tag{5}$$

In this way we take a net y_U of X converging to x. If K is the set all of these y_U, then $x \in \overline{K}$ and so $f(x) \in f(\overline{K})$.

We will show that there is an open set W containing $f(K)$ such that $f(x) \notin \overline{W}$, which contradicts (4).

Actually (5) implies that

$$f(K) \subset \overline{V}^c \subset V^c. \tag{6}$$

Put $W = \overline{V}^c$. Then $\overline{W} = \overline{\overline{V}^c} \subset V^c$ because of V^c is a closed set. Now since $f(x) \in V$ it follows that

$$f(x) \notin \overline{W}.$$

($\Rightarrow$) Suppose that (4) does not hold. Then there exists a set $K \subset X$ and an open set U containing $f(K)$, with

$$f(\overline{K}) \not\subset \overline{U}.$$

So for some point x in $\overline{K}$ we have $f(x) \notin \overline{U}$, i.e.

$$f(x) \in \overline{U}^c.$$

Put

$$V = \overline{U}^c.$$

We assert that for the open set V, which contains $f(x)$, there does not exist a W containing x with

$$f(W) \subset \overline{V} \tag{7}$$

which contradicts the weak continuity of f.

Suppose, in contrary, that such a W exists.

Since $x \in \overline{K}$ we have $W \cap K \neq \emptyset$, so $f(W) \cap f(K) \neq \emptyset$.

By (7) this implies that

$$f(K) \cap \overline{V} \neq \emptyset. \tag{8}$$

But $V = \overline{U}^c \subset U^c$ and the last set is closed, so $\overline{V} \subset U^c$. Since U contains $f(K)$, it follows that $\overline{V} \cap f(K) = \emptyset$ which contradicts (8). Thus such a W does not exist, and this proves our assertion, which completes the proof. □

References

[1] Dontchev, J., Noiri, T.: *A note on Saleh's paper "Almost continuity implies closure continuity"*. Glaskow Math. J. **40** (1988), 473.

[2] Levine, N.: *A decomposition of continuity in topological spaces.* Amer. Math. Monthly **68** (1961), 44–46.

[3] Long, P. E., McGehee, E. E.: *Properties of almost continuous functions.* Proc. Amer. Math. Soc. **24** (1970), 175–180.

[4] Long, P. E., Carnahan, D. A.: *Comparing almost continuous functions.* Proc. Amer. Math. Soc. **38** (1973), 413–418.

[5] Noire, T. : *On weakly continuous mappings.* Proc. Amer. Math. Soc. **46** (1974), 120–124.

[6] Saleh, M.: *Almost continuity implies closure continuity.* Glaskow Math. J. **40** (1998), 263–264.

[7] Singal, M. K., Singal, A. R.: *Almost continuous mappings.* Yokohama Math. J. **16** (1968), 63–73.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 137–141

Zeros of Derivatives of Solutions to Singular $(p, n-p)$ Conjugate BVPs *

Irena RACHŮNKOVÁ [1], Svatoslav STANĚK [2]

Department of Mathematical Analysis and Applications of Mathematics,
Faculty of Science, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: [1] *rachunko@inf.upol.cz*
[2] *stanek@inf.upol.cz*

(Received January 10, 2004)

Abstract

Positive solutions of the singular $(p, n-p)$ conjugate BVP are studied. The set of all zeros of their derivatives up to order $n-1$ is described. By means of this, estimates from below of the solutions and the absolute values of their derivatives up to order $n-1$ on the considered interval are reached. Such estimates are necessary for the application of the general existence principle to the BVP under consideration.

Key words: Singular conjugate BVP, positive solutions, zeros of derivatives, estimates from below.

2000 Mathematics Subject Classification: 34B15, 34B16, 34B18

1 Introduction

Let $n, p \in \mathbb{N}$, $n > 2$, $p \leq n-1$, and T be a positive number. In [3] (for $p = 1$) and [6], the authors have considered the singular $(p, n-p)$ conjugate boundary value problem (BVP)

$$(-1)^p x^{(n)}(t) = f(t, x(t), \dots, x^{(n-1)}(t)), \tag{1.1}$$

*Supported by Grant No. 201/04/1077 of the Grant Agency of the Czech Republic and by the Council of Czech Government J14/98 153100011

$$x^{(i)}(0) = 0, \quad x^{(j)}(T) = 0 \quad 0 \le i \le n-p-1, \ 0 \le j \le p-1, \tag{1.2}$$

where f satisfies the local Carathéodory conditions on the set $\mathcal{D} = [0,T] \times ((0,\infty) \times \mathbb{R}_0^{n-1})$ with $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$ and f is singular at the value 0 of each its phase variable. They have given conditions on f guaranteeing the existence of a positive (on $(0,T)$) solution to BVP (1.1), (1.2). The singularities of the function f in (1.1) 'appear' in any positive solution of BVP (1.1), (1.2) and some its derivatives at the fixed points $t = 0$, $t = T$, and all its derivatives up to order $n-1$ 'pass through' singularities of f also at inner points of the interval $(0,T)$ which are not fixed. Therefore for proving the solvability of BVP (1.1), (1.2) in the class of positive functions on $(0,T)$ it is very important to give a localization analysis of zeros of derivatives up to order $n-1$ of positive solutions to BVP (1.1), (1.2). This analysis have been presented for $p = 1$ in [3] and for $p = 2$ in [6] under the assumption that $f \ge c$ on $\mathcal{D}$ with a positive constant c. The aim of this paper is to complete this analysis for all values of p. We note that the singular differential equation

$$(-1)^p x^{(n)}(t) = \phi(t) g(t, x(t)) \tag{1.3}$$

together with the boundary conditions (1.2) have been discussed for $\phi(t)g(t,x)$: $(0,1) \times (0,\infty) \to (0,\infty)$ continuous in [1], [2], [4] and [5] (in [4] and [5] with $\phi = 1$). But for BVP (1.3), (1.2) singularities of g 'appear' in its positive solutions only at the fixed points $t = 0$ and $t = 1$.

2 Localization analysis of zeros to solutions of BVP (1.1), (1.2)

Let c be a positive constant and let f in (1.1) satisfy $f \ge c$ on $\mathcal{D}$. Then the localization analysis of zeros to solutions of BVP (1.1), (1.2) and their derivatives up to order $n-1$ can be studied by the localization analysis of zeros to solutions of the differential inequality

$$(-1)^p x^{(n)}(t) \ge c \tag{2.1}$$

satisfying the boundary conditions (1.2). By a *solution of problem* (2.1), (1.2) we understand a function $x \in AC^{n-1}([0,T])$ (functions having absolutely continuous $(n-1)^{\text{st}}$ derivative on $[0,T]$) satisfying (2.1) for a.e. $t \in [0,T]$ and fulfilling (1.2).

Having a solution x of problem (2.1), (1.2) we are interested in zeros of $x^{(k)}$, $0 \le k \le n-1$, belonging to $(0,T)$. Without loss of generality we can suppose

$$p - 1 \le n - p - 1 \tag{2.2}$$

that is $p \le n/2$, because by replacing t by $T - t$ we can transform the case $n/2 < p$ to (2.2).

For $p = 1, 2$ we have already studied zeros of $x^{(k)}$ and we have proved the following results:

Lemma 2.1 *Let x be a solution of problem (2.1), (1.2) for $p = 1$. Then $x > 0$ on $(0, T)$ and $x^{(k)}$ has just one zero in $(0, T)$, $1 \le k \le n-1$.*

Proof Lemma follows from [3], Lemmas 2.7 and 2.9. □

Lemma 2.2 *Let x be a solution of problem (2.1), (1.2) for $p = 2$. Then*

(i) $x > 0$ *on* $(0, T)$,

(ii) $x^{(k)}$ *has just one zero in* $(0, T)$ *for* $k = 1$ *and* $k = n-1$,

(iii) $x^{(k)}$ *has just two zeros in* $(0, T)$ *for* $2 \le k \le n-2$.

Proof See [6], Lemmas 2.2. □

Decomposition analysis of zeros to solutions of BVP (2.1), (1.2) with $p \ge 3$ is described in the next theorem.

Theorem 2.3 *Let x be a solution of problem (2.1), (1.2) for $p \ge 3$ and let (2.2) hold. Then*

(i) $x > 0$ *on* $(0, T)$,

(ii) $x^{(k)}$ *has just* j *zeros in* $(0, T)$ *for* $k = j$ *and* $k = n-j$ *where* $j = 1, 2, \dots, p-1$,

(iii) $x^{(k)}$ *has just* p *zeros in* $(0, T)$ *for* $p \le k \le n-p$.

Proof The proof is divided into three parts.

I. *Lower bounds for zeros.* By (1.2) we see that x' has at least one zero $t_1^1 \in (0, T)$. Hence $x'(0) = x'(t_1^1) = x'(T) = 0$, which implies that x'' has at least two zeros $t_1^2, t_2^2 \in (0, T)$. So, we have $x''(0) = x''(t_1^2) = x''(t_2^2) = x''(T) = 0$. By induction we conclude that $x^{(j)}$, $j = 3, \dots, p-1$, has at least j zeros $t_1^j, \dots, t_j^j \in (0, T)$ and, due to (1.2) and (2.2) $x^{(j)}(0) = x^{(j)}(t_1^j) = \dots = x^{(j)}(t_j^j) = x^{(j)}(T) = 0$, $j = 3, \dots, p-1$. Therefore $x^{(p)}$ hat at least p zeros in $(0, T)$. Now we will distinguish two cases: $p < n/2$ and $p = n/2$.

1. Let $p < n/2$. Then $p \le n-p-1$ and, by (1.2),

$$x^{(j)}(0) = 0, \quad j = p, \dots, n-p-1.$$

Thus $x^{(k)}$ has at least p zeros in $(0, T)$ for $k = p+1, \dots, n-p$.

2. Let $p = n/2$ (clearly n is even in this case). Then $p = n-p$ and $x^{(n-p)}$ has at least p zeros in $(0, T)$.

Hence we have shown that $x^{(n-p)}$ has at least p zeros in $(0, T)$ in the both cases. Since for $x^{(n-j)}$, $1 \le j \le p-1$, we cannot already use (1.2), we deduce that $x^{(n-j)}$ has at least j zeros in $(0, T)$ for $j = 1, \dots, p-1$. Particularly $x^{(n-1)}$ has at least one zero in $(0, T)$.

II. *Exact number of zeros.* By (2.1), $x^{(n-1)}$ is strictly monotonous and hence it has just one zero in $(0, T)$. Therefore, by I, we deduce that $x^{(n-k)}$ has just k zeros in $(0, T)$ for $2 \le k \le p-1$ and $x^{(k)}$ has just p zeros in $(0, T)$ for

$p \le k \le n-p$. Similarly, $x^{(k)}$ has just k zeros in $(0,T)$ for $1 \le k \le p-1$ and x has no zero in $(0,T)$.

III. *Positivity of* x. Denote by t_1^k the first zero of $x^{(k)}$ in $(0,T)$, $1 \le k \le n-1$. Inequality (2.1) implies that $(-1)^p x^{(n-1)} < 0$ on $[0, t_1^{n-1})$ and $(-1)^p x^{(n-2)} > 0$ on $[0, t_1^{n-2})$. Therefore $(-1)^{p+j} x^{(n-j)} > 0$ on $(0, t_1^{n-j})$ for $j = 3, \dots, p$. Particularly we have $x^{(n-p)} > 0$ on $(0, t_1^p)$, wherefrom, by virtue of (1.2), we obtain $x^{(k)} > 0$ on $(0, t_1^k)$, $1 \le k \le n-p-1$, and consequently $x > 0$ on $(0,T)$. □

Our next theorem provides estimates from below of solutions to problem (2.1), (1.2) and of the absolute value of their derivatives up to order $n-1$ on the interval $[0,T]$. These estimations are necessary to apply the general existence principle of [6] to problem (1.1), (1.2) with f in (1.1) satisfying the inequality $f \ge c$ on $\mathcal{D}$.

Theorem 2.4 *Let x be a solution of problem (2.1), (1.2). Then for any $i \in \{1, \dots, n-1\}$ there are $p_i + 1$ disjoint intervals (a_k, a_{k+1}), $0 \le k \le p_i$, $p_i < (n-1)p$, such that*

$$\bigcup_{k=0}^{p_i} [a_k, a_{k+1}] = [0,T] \tag{2.3}$$

and for each $k \in \{0, \dots, p_i\}$ one of the inequalities

$$|x^{(n-i)}(t)| \ge \frac{c}{i!}(t - a_k)^i \quad \text{for } t \in [a_k, a_{k+1}] \tag{2.4}$$

or

$$|x^{(n-i)}(t)| \ge \frac{c}{i!}(a_{k+1} - t)^i \quad \text{for } t \in [a_k, a_{k+1}] \tag{2.5}$$

is satisfied.

Proof Let x be a solution of problem (2.1), (1.2) and let $t_i^j \in (0,T)$ be zeros of $x^{(j)}$ described in Lemmas 2.1, 2.2 and Theorem 2.3. Integrating (2.1) we get

$$\begin{aligned} &(-1)^{p+1} x^{(n-1)}(t) \ge c(t_1^{n-1} - t) && \text{for } t \in [0, t_1^{n-1}] \\ &(-1)^p x^{(n-1)}(t) \ge c(t - t_1^{n-1}) && \text{for } t \in [t_1^{n-1}, T]. \end{aligned} \tag{2.6}$$

Now, integrate the first inequality in (2.6) from $t \in [0, t_1^{n-2})$ to t_1^{n-2}, we have

$$(-1)^p x^{n-2}(t) \ge \frac{c}{2}\Big(-(t_1^{n-1} - t_1^{n-2})^2 + (t_1^{n-1} - t)^2 \Big) \ge \frac{c}{2!}(t_1^{n-2} - t)^2.$$

Hence, we get in such a way

$$\begin{aligned} &(-1)^p x^{(n-2)}(t) \ge \tfrac{c}{2!}(t_1^{n-2} - t)^2 && \text{for } t \in [0, t_1^{n-2}] \\ &(-1)^{p+1} x^{(n-2)}(t) \ge \tfrac{c}{2!}(t - t_1^{n-1})^2 && \text{for } t \in [t_1^{n-2}, t_1^{n-1}] \\ &(-1)^{p+1} x^{(n-2)}(t) \ge \tfrac{c}{2!}(t_2^{n-2} - t)^2 && \text{for } t \in [t_1^{n-1}, t_2^{n-2}] \\ &(-1)^p x^{(n-2)}(t) \ge \tfrac{c}{2!}(t - t_2^{n-2})^2 && \text{for } t \in [t_2^{n-2}, T]. \end{aligned} \tag{2.7}$$

Choose $i \in \{1, \dots, n-1\}$ and take all different zeros of functions $x^{(n-1)}, \dots,$ $x^{(n-i)}$, which are in $(0, T)$. By Lemmas 2.1, 2.2 and Theorem 2.3, there is a finite number $p_i < (n-1)p$ of these zeros. Let us put them in order and denote by $a_1, \dots, a_{p_i}$. Set $a_0 = 0$, $a_{p_i+1} = T$. In this way we get $p_i + 1$ disjoint intervals (a_k, a_{k+1}), $0 \le k \le p_i$, satisfying (2.3).

If $i = 1$, then for $a_1 = t_1^{n-1}$, $a_2 = T$, we get by (2.6) that $|x^{(n-1)}(t)| \ge c(a_1 - t)$ for $t \in [a_0, a_1]$ and $|x^{(n-1)}(t)| \ge c(t - a_1)$ for $t \in [a_1, a_2]$.

If $i = 2$, we put $t_1^{n-1} = a_1$, $t_1^{n-2} = a_2$, $t_2^{n-2} = a_3$, $T = a_4$, and then (2.7) gives (2.4) or (2.5).

If $i > 2$ and we integrate the inequalities in (2.7) $(i-2)$-times, we get that on each $[a_k, a_{k+1}]$, $k \in \{0, \dots, p_i\}$ either (2.4) or (2.5) has to be fulfilled. □

References

[1] Agarwal, R. P., O'Regan, D.: *Positive solutions for (p,n-p) conjugate boundary value problems.* J. Differential Equations **150** (1998), 462–473.

[2] Agarwal, R. P., O'Regan, D.: *Multiplicity results for singular conjugate, focal, and (n, p) problems.* J. Differential Equations **170** (2001), 142–156.

[3] Agarwal, R. P., O'Regan, D., Rachůnková, I., Staněk, S.: *Two-point higher order BVPs with singularities in phase variables.* Computers & Mathematics with Applications **46** (2003), 1799–1826.

[4] Eloe, P. W., Henderson, J.: *Singular nonlinear $(n-1, 1)$ conjugate boundary value problems.* Georgian Math. J. **4** (1997), 401–412.

[5] Eloe, P. W., Henderson, J.: *Singular nonlinear $(k, n-k)$ conjugate boundary value problems.* J. Differential Equations **133** (1997), 136–151.

[6] Rachůnková, I., Staněk, S.: *General existence principle for singular BVPs and its applications.* Georgian Math. J. **11** (2004), 549–565.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 143–148

Semirings Embedded in a Completely Regular Semiring

M. K. SEN & S. K. MAITY [1]

Department of Pure Mathematics, University of Calcutta
35, Ballygunge Circular Road, Kolkata-700019, India
e-mail: [1]*senmk@cal3.vsnl.net.in*
[2]*sskmaity@yahoo.com*

(Received July 17, 2003)

Abstract

Recently, we have shown that a semiring S is completely regular if and only if S is a union of skew-rings. In this paper we show that a semiring S satisfying $a^2 = na$ can be embedded in a completely regular semiring if and only if S is additive separative.

Key words: Completely regular semiring, skew-ring, b-lattice, archimedean semiring, additive separative semiring.

2000 Mathematics Subject Classification: 16A78, 20M10, 20M07

1 Introduction

Recall that a semiring $(S, +, \cdot)$ is a type (2,2) algebra whose semigroup reducts $(S, +)$ and $(S, \cdot)$ are connected by ring like distributivity, that is,

$$a(b + c) = ab + ac \quad \text{and} \quad (b + c)a = ba + ca$$

for all $a, b, c \in S$. A semiring $(S, +, \cdot)$ is called a Boolean semiring if $a^2 = a$ for all $a \in S$. A semiring S is called additive cancellative if the additive reduct $(S, +)$ is a cancellative semigroup, i.e., for $a, b, c \in S$, $a + b = a + c$ implies $b = c$.

[1]The research is supported by CSIR, India.

In this paper, we call an element a of a semiring $(S,+,\cdot)$ completely regular if there exists an element $x \in S$ satisfying the following conditions:

$$(i) \qquad a = a + x + a$$
$$(ii) \qquad a + x = x + a$$
$$(iii) \qquad a(a + x) = a + x$$

Naturally, a semiring $(S,+,\cdot)$ is a completely regular semiring if every element a of S is completely regular. There are plenty of examples of completely regular semirings, for example, every ring is a completely regular semiring and every distributive lattice is also a completely regular semiring. By definition, if $(S,+,\cdot)$ is a completely regular semiring then its additive reduct $(S,+)$ is a completely regular semigroup but the converse may not be true. For example, if we let $(S,+,\cdot)$ be a semiring whose additive reduct $(S,+)$ is an idempotent semigroup and the multiplicative reduct $(S,\cdot)$ is not a band, then we can immediately see that $(S,+)$ is completely regular but the semiring $(S,+,\cdot)$ itself is not completely regular. Throughout this paper, we denote the set of all inverse elements of a in the regular semigroup $(S,+)$ by $V^+(a)$. As usual, we denote the Green's $\mathcal{H}$-relations on $(S,+)$ by $\mathcal{H}^+$

The following useful concept is due to M. P. Grillet [2].

Definition 1.1 A semiring $(S,+,\cdot)$ is called a skew-ring if its additive reduct $(S,+)$ is a group, not necessarily an abelian group.

We have obtained the following result in [4].

Theorem 1.2 *The following statements on a semiring S are equivalent.*
(I) S is completely regular.
(II) Every $\mathcal{H}^+$-class is a skew-ring.
(II) S is union (disjoint) of skew-rings.

Corollary 1.3 *An additive commutative semiring S is completely regular if and only if S is union of rings.*

2 b-lattice decomposition

We consider the additive commutative semiring $(S,+,\cdot)$ such that for each $a \in S$ there exists a positive integer n such that

$$a^2 = na. \qquad (A)$$

Clearly, every Boolean semiring is a semiring which satisfies condition (A). Also the semiring of all natural numbers is a semiring of this kind which is not Boolean.

We now consider the following examples:

Example 2.1 Let $S = \mathbb{N}\times\{1,2,3\}$. On S we define addition and multiplication by

$$(a,i)+(b,j)=(a+b,\max\{i,j\})$$

and

$$(a,i)\cdot(b,j)=(ab,\min\{i,j\}).$$

Then $(S,+,\cdot)$ is a semiring satisfying condition (A).

Example 2.2 Let $S=\{0,a,b\}$ be a semiring with the following Cayley tables:

$+$	0	a	b
0	0	a	b
a	a	0	b
b	b	b	b

$\cdot$	0	a	b
0	0	0	0
a	0	0	0
b	0	0	b

Then $(S,+,\cdot)$ is a semiring which satisfies condition (A) but not Boolean.

Definition 2.3 A semiring $(S,+,\cdot)$ is called a b-lattice if $(S,+)$ is a semilattice and $(S,\cdot)$ is a band. Moreover, a congruence ρ on a semiring S is called a b-lattice congruence if S/ρ is a b-lattice. A semiring S is called a b-lattice Y of semirings $S_\alpha\,(\alpha\in Y)$ if S admits a b-lattice congruence ρ on S such that $Y=S/\rho$ and each S_α is a ρ-class.

Definition 2.4 Let $(S,+,\cdot)$ be a semiring. We define a relation η on S by $a\,\eta\,b$ if and only if there exist $x,y\in S^0$ and positive integers m,n such that $a+x=mb$ and $b+y=na$. Also, we define a relation σ on S by $a\,\sigma\,b$ if and only if there exists a positive integer n such that $a+nb=(n+1)b$ and $b+na=(n+1)a$.

It should be noted that if there exist positive integers m,n such that $a+mb=(m+1)b$ and $b+na=(n+1)a$ then $a\,\sigma\,b$. For if, say $m<n$, then we can add $a+mb=(m+1)b$ by $(n-m)b$ and obtain $a+nb=(n+1)b$.

Definition 2.5 A semiring S is called archimedean if $(S,+)$ is an archimedean semigroup i.e., for any $a,b\in S$ there exist $x,y\in S$ and positive integers m,n such that $a+x=mb$ and $b+y=na$.

Lemma 2.6 *Let S be a semiring satisfying (A). Then*

(i) η is a congruence on S and S/η is the maximal b-lattice homomorphic image of S.

(ii) S is uniquely expressible as a b-lattice T of archimedean semirings $S_\alpha(\alpha\in T)$. The b-lattice T is isomorphic with the maximal b-lattice homomorphic image S/η of S and $S_\alpha(\alpha\in T)$ are equivalent classes of η in S.

Proof (i) From Theorem 4.12 in [1], it follows that η is a semilattice congruence on $(S,+)$. Let $a\,\eta\,b$ and $c\in S$. Then there exist $x,y\in S^0$ and positive integers m,n such that $a+x=mb$ and $b+y=na$. This leads to $ac+xc=m(bc)$ and $bc+yc=n(ac)$. Thus $ac\,\eta\,bc$. Similarly, we can show that $ca\,\eta\,cb$. Hence η is a congruence on the semiring S.

Since S satisfies $a^2 = na$ so $a^2 \,\eta\, na$. Again since η is a semilattice congruence on $(S,+)$, it follows that $na \,\eta\, a$. Thus, $a^2 \,\eta\, a$ and hence η is a b-lattice congruence on the semiring S.

S/η is the maximal homomorphic image of S follows from Theorem 4.12 in [1].

(ii) By (i) of this Lemma, η is a b-lattice congruence on S. By Theorem 4.13 in [1], each η-class $S_\alpha (\alpha \in S/\eta)$ is archimedean semigroup under addition. We show that each S_α is a semiring. For this let $b, c \in \eta(a)$, where $\eta(a)$ is the η-class of $a \in S$. Then $b \,\eta\, a$ and $c \,\eta\, a$. This leads to $bc \,\eta\, a^2 \,\eta\, a$. So $bc \in \eta(a)$ and hence $(S_\alpha, +, \cdot)$ is an archimedean semiring. Thus, S is a b-lattice T of archimedean semirings. Unique expression of S as a b-lattice of archimedean semirings follows from Theorem 4.13 in [1].

The last part of the theorem follows from the Theorem 4.13 in [1].

Definition 2.7 A congruence ρ on a semiring S is said to be additive separative (AS-congruence) if S/ρ is an additive separative semiring (AS-semiring) i.e., $(a+b)\,\rho\,(a+a)\,\rho\,(b+b)$ implies $a \,\rho\, b$.

Lemma 2.8 *The relation σ defined in Definition 2.4 is a congruence on a semiring S and S/σ is the maximal additive separative homomorphic image of S.*

Proof By Theorem 4.14 in [1], σ is a congruence on $(S,+)$. Let $a \,\sigma\, b$ and $c \in S$. Then there exist positive integers m, n such that $a + nb = (n+1)b$ and $b+ma = (m+1)a$. This leads to $ac+n(bc) = (n+1)bc$ and $bc+m(ac) = (m+1)ac$. Hence $ac \,\sigma\, bc$. Similarly, one can show that $ca \,\sigma\, cb$. Thus, σ is a congruence on S.

Last part follows from Theorem 4.14 in [1].

Corollary 2.9 *Let S be an additive separative semiring. If a, b are elements of S such that $a + mb = (m+1)b$ and $b + na = (n+1)a$ for some positive integers m and n, then $a = b$.*

Theorem 2.10 *A semiring S satisfying the condition (A) can be embedded in a completely regular semiring if and only if S is additive separative.*

Proof First suppose that S can be embedded in a completely regular semiring. Then the additive reduct $(S,+)$ of the semiring S can be embedded in a completely regular semigroup. Then by Theorem 4.19 in [1], we have the semigroup reduct $(S,+)$ is separative, i.e., S is additive separative semiring.

Conversely, assume that S is additive separative. Since the semiring S satisfies the condition $a^2 = na$ so S can be expressed as a b-lattice of archimedean semirings. Let $S = \bigcup_{\alpha \in T} S_\alpha$ be the expression of S as a b-lattice T of its archimedean components $S_\alpha (\alpha \in T)$. Since S is additive separative, by Theorem 4.16 in [1] we have S_α is additive cancellative. So by Theorem 5.11 in [3] S_α can be embedded in a ring R_α. Since S_α are mutually disjoint, we can assume that R_α are mutually disjoint. Now every element of R_α can be expressed in the form $a_1 - a_2$ with $a_1, a_2 \in S_\alpha$ and that $a_1 - a_2 = c_1 - c_2$ if and only if $a_1 + c_2 = a_2 + c_1$.

Let $S' = \bigcup_{\alpha \in T} R_\alpha$. On S' we define $\oplus$ and $\odot$ as follows:

$$a \oplus b = (a_1 + b_1) - (a_2 + b_2)$$

and

$$a \odot b = (a_1 b_1 + a_2 b_2) - (a_1 b_2 + b_2 a_1),$$

where $a = a_1 - a_2$ and $b = b_1 - b_2$.

We first show that the operations are well defined. For this let $a = a_1 - a_2 = c_1 - c_2$ and $b = b_1 - b_2 = d_1 - d_2$. So $a_1 + c_2 = a_2 + c_1$ and $b_1 + d_2 = b_2 + d_1$. Now,

$$(a_1+b_1)+(c_2+d_2) = (a_1+c_2)+(b_1+d_2) = (a_2+c_1)+(b_2+d_1) = (a_2+b_2)+(c_1+d_1)$$

This leads to,

$$(a_1 + b_1) - (a_2 + b_2) = (c_1 + d_1) - (c_2 + d_2),$$
$$(a_1 - a_2) \oplus (b_1 - b_2) = (c_1 - c_2) \oplus (d_1 - d_2).$$

So $\oplus$ is well defined.

Again,

$$a_1 b_1 + c_2 b_1 + a_2 b_2 + c_1 b_2 = a_2 b_1 + c_1 b_1 + a_1 b_2 + c_2 b_2,$$
$$(a_1 b_1 + a_2 b_2) + (c_2 b_1 + c_1 b_2) = (c_1 b_1 + c_2 b_2) + (a_2 b_1 + a_1 b_2),$$
$$(a_1 b_1 + a_2 b_2) - (a_2 b_1 + a_1 b_2) = (c_1 b_1 + c_2 b_2) - (c_2 b_1 + c_1 b_2),$$
$$(a_1 - a_2) \odot (b_1 - b_2) = (c_1 - c_2) \odot (b_1 - b_2).$$

Similarly, we can show that

$$(c_1 - c_2) \bigodot (b_1 - b_2) = (c_1 - c_2) \bigodot (d_1 - d_2).$$

Thus,

$$(a_1 - a_2) \bigodot (b_1 - b_2) = (c_1 - c_2) \bigodot (d_1 - d_2).$$

Hence $\bigodot$ is well defined.

Clearly, if $a \in R_\alpha$ and $b \in R_\beta$ $(\alpha, \beta \in T)$ then $a \oplus b \in R_{\alpha+\beta}$ and $a \bigodot b \in R_{\alpha\beta}$.

The associativity under $\oplus$ and $\bigodot$ is easily verified. Also, we can show the distributivity. Hence S' is indeed a semiring which contains S. Since S' is union of rings so by Corollary 1.3, S' is a completely regular semiring.

We now show that if a and b are elements of S then $a \oplus b$ and $a \bigodot b$ are respectively the same as the original operation $a + b$ and $a.b$ respectively in S. Let $a \in R_\alpha$ and $b \in R_\beta (\alpha, \beta \in T)$. Then $a = 2a - a$ and $b = 2b - b$ so that $a \oplus b = (2a - a) \oplus (2b - b) = (2a + 2b) - (a + b) = 2(a + b) - (a + b) = a + b$ and $a \bigodot b = (2a - a) \bigodot (2b - b) = \big((2a)(2b) + ab\big) - (2ab + 2ab) = 5ab - 4ab = a \cdot b$, as desired.

Acknowledgements The authors express their sincere thanks to the learned referee for his valuable suggestions.

References

[1] Clifford, A. H., Preston, G. B.: The algebraic theory of semigroups. *Amer. Math. Soc., Providence Rhode Island, Vol. I*, 1961.

[2] Grillet, M. P.: *Semirings with a completely simple additive semigroup.* J. Austral. Math. Soc. **A 20** (1975), 257–267.

[3] Hebisch, U, Weinert, H. J.: Semirings, Algebra Theory and Applications in Computer Science, Series in Algebra. *Vol. 5, World Scientific, Singapore*, 1998.

[4] Sen, M. K., Maity, S. K., Shum, K. P.: *On completely regular semirings.* (submitted).

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 149–154

A Note on Orthodox Additive Inverse Semirings

M. K. SEN & S. K. MAITY

Department of Pure Mathematics, University of Calcutta
35, Ballygunge Circular Road, Kolkata-700019, India
e-mail: senmk@cal3.vsnl.net.in

(Received November 20, 2003)

Abstract

We show in an additive inverse regular semiring $(S, +, \cdot)$ with $E^{\bullet}(S)$ as the set of all multiplicative idempotents and $E^{+}(S)$ as the set of all additive idempotents, the following conditions are equivalent:
(i) For all $e, f \in E^{\bullet}(S)$, $ef \in E^{+}(S)$ implies $fe \in E^{+}(S)$.
(ii) $(S, \cdot)$ is orthodox.
(iii) $(S, \cdot)$ is a semilattice of groups.
This result generalizes the corresponding result of regular ring.

Key words: Additive inverse semirings, regular semirings, orthodox semirings.

2000 Mathematics Subject Classification: 16A78, 20M10, 20M07

1 Introduction

A semiring $(S, +, \cdot)$ is a nonempty set S on which operations of addition, $+$, and multiplication, $\cdot$, have been defined such that the following conditions are satisfied:

(1) $(S, +)$ is a semigroup.

(2) $(S, \cdot)$ is a semigroup.

(3) Multiplication distributes over addition from either side.

A semiring $(S, +, \cdot)$ is called an additive inverse semiring if $(S, +)$ is an inverse semigroup, that is for each $a \in S$ there exists a unique element $a' \in S$ such that $a + a' + a = a$ and $a' + a + a' = a'$. Additive inverse semirings were first studied by Karvellas [4] in 1974. Karvellas [4] proved the following: (Karvellas (1974), Theorem 3(ii) and Theorem 7) Take any additive inverse semiring $(S, +, \cdot)$.

(i) For all $x, y \in S$, $(x \cdot y)' = x' \cdot y = x \cdot y'$ and $x' \cdot y' = x \cdot y$

(ii) If $a \in aS \cap Sa$ for all $a \in S$ then S is additively commutative.

A semiring $(S, +, \cdot)$ is called regular if for each $a \in S$ there exists $x \in S$ such that $axa = a$. In a regular semiring S, for any element $a \in S$, $V^\bullet(a) = \{x \in S : axa = a \text{ and } xax = x\}$. A regular semiring S contains element e such that $e \cdot e = e$. We denote the set of such elements by $E^\bullet(S)$. If in a regular semiring S, $E^\bullet(S)$ is a subsemigroup of the semigroup of $(S, \cdot)$, then the semiring S is called an orthodox semiring.

Chaptal [1] proved the following result in 1966.

Result 1.1 For a ring $(R, +, \cdot)$ the following conditions are equivalent.

(i) $(R, \cdot)$ is a union of groups.

(ii) $(R, \cdot)$ is an inverse semigroup.

(iii) $(R, \cdot)$ is a semilattice of groups.

Latter J. Zeleznekow [5] proved the following result.

Result 1.2 In a regular ring $(R, +, \cdot)$ the following conditions are equivalent.

(i) $(R, \cdot)$ is orthodox.

(ii) $(R, \cdot)$ is a union of groups.

(iii) $(R, \cdot)$ is an inverse semigroup.

(iv) $(R, \cdot)$ is a semilattice of groups.

These results do not hold in arbitrary semiring [see Example 2.1.]. The aim of this paper is to generalize these results in an additive inverse semiring with some conditions. For notations and terminologies not given in this note, the reader is referred to the monograph of Golan [2] and Howie [3].

2 Orthodox additive inverse semiring

An additive inverse semiring S is called orthodox if $(S, \cdot)$ is an orthodox semigroup.

Example 2.1 [5] Let S be the set of all binary relations on a two element set. Under the operations of union and composition of binary relations, S becomes a semiring in which $(S, \cdot)$ is regular but neither orthodox nor a union of groups.

Example 2.2 Let $(S, +)$ be a semilattice with more than one element. On S, define the multiplication, $\cdot$, by $a \cdot b = a$ for all $a, b \in S$. Then $(S, +, \cdot)$ is a semiring such that $(S, +)$ is an inverse semigroup, $(S, \cdot)$ is orthodox. Hence this semiring is an orthodox additive inverse semiring. In this semiring we find that $(S, \cdot)$ is not an inverse semigroup.

¿From the above example we find that J. Zeleznekow's result is not true in an orthodox additive inverse semiring. Let S be an additive inverse semiring. We say that S satisfies conditions (A) and (B) if for all $a, b \in S$

(A) $a(b + b') = (b + b')a$.

(B) $a + a(b + b') = a$.

Clearly rings, distributive lattices and direct products of distributive lattice and ring are natural examples of such additive inverse semiring. We consider the following example.

Example 2.3 Let $S = \{0, a, b\}$. Define addition and multiplication on S by the following Cayley tables:

$+$	0	a	b
0	0	a	b
a	a	0	b
b	b	b	b

$\cdot$	0	a	b
0	0	0	0
a	0	0	0
b	0	0	b

It is easy to see that $(S, +, \cdot)$ is a semiring such that $(S, +)$ is an additive inverse semiring with conditions (A) and (B).

In the remaining part of this section we assume that S denotes an additive commutative and additive inverse semiring satisfying conditions (A) and (B). Also we assume that $E^+(S) = \{a \in S : a + a = a\}$. Note that $E^+(S)$ is an ideal of S.

We now prove the following Lemma.

Lemma 2.4 *Let $a, b \in S$ be such that $a + b' \in E^+(S)$ and $a + a' = b + b'$. Then $a = b$.*

Proof Since $a + b' \in E^+(S)$ so we have

$$a + b' = (a + b') + (a + b')' = a + b' + b + a' = a + a' + b + b' = b + b'.$$

This leads to, $a + b' + b = b + b' + b$, i.e., $a + a' + a = b$. Hence $a = b$. □

Nest we prove the following important lemma.

Lemma 2.5 *If the semiring S is multiplicatively regular then the following conditions are equivalent.*

(i) For all $e, f \in E^\bullet(S)$, $ef \in E^+(S)$ implies $fe \in E^+(S)$.

(ii) For all $e \in E^\bullet(S)$, for all $x \in S$, $ex \in E^+(S)$ implies $xe \in E^+(S)$.

(iii) For all $n \in \mathbb{N}$, for all $x \in S$, $x^n \in E^+(S)$ implies $x \in E^+(S)$.

(iv) For all $x \in S$, $x^2 \in E^+(S)$ implies $x \in E^+(S)$.

(v) For all $x, y \in S$, $xy \in E^+(S)$ implies $yx \in E^+(S)$.

Furthermore, each is implied by

(vi) $(S, \cdot)$ is orthodox.

Proof (i) $\Rightarrow$ (ii): Let $e \in E^\bullet(S)$ and $x \in S$ be such that $ex \in E^+(S)$. Then $ex = ex + (ex)' = ex + ex'$. Now,

$$\begin{aligned}
(e+xe')^2 &= e(e+xe') + xe'(e+xe') \\
&= e^2 + exe' + xe'e + xe'xe' \\
&= e^2 + exe' + xe'e + xe'x'e \\
&= e^2 + exe' + xe'e + x(e')'xe \\
&= e^2 + exe' + xe'e + xexe \\
&= e + (ex+ex')e' + xe' + x(ex+ex')e \\
&= e + (exe' + ex'e') + xe' + x(e'x' + e'x)e \\
&= e + e(xe' + xe) + xe' + xe'(x'e + xe) \\
&= e + xe' \text{ (by condition (B))}.
\end{aligned}$$

Thus $e + xe' \in E^\bullet(S)$. Let $x^* \in V^\bullet(x)$. Now,

$$\begin{aligned}
(e+xe')(xx^*) &= exx^* + xe'xx^* \\
&= exx^* + (xexx^*)' \\
&= exx^* + x'(ex)x^* \in E^+(S) \text{ (as } E^+(S) \text{ is an ideal of } S).
\end{aligned}$$

But $e + xe', xx^* \in E^\bullet(S)$. So by (i), $xx^*(e + xe') \in E^+(S)$ and thus $xx^*e + xx^*xe' = xx^*e + xe' \in E^+(S)$. Also,

$$\begin{aligned}
xx^*e + xx^*e' &= xx^*e' + xx^*e \\
&= xx^*e' + xx^*e + xx^*e(x+x') \text{ (by condition (B))} \\
&= xx^*(e'+e) + xx^*(x+x')e \text{ (by condition (A))} \\
&= x(e'+e)x^* + xx^*x(e+e') \text{ (by condition (A))} \\
&= x(e'x^* + ex^*) + xe + xe' \\
&= xe(x^{*\prime} + x^*) + xe + xe' \\
&= xe + xe' \text{ (by condition (B))}
\end{aligned}$$

Hence by Lemma 2.4., we have $xx^*e = xe$. Now, $exx^* \in E^+(S)$ [as $ex \in E^+(S)$] and hence $xe = xx^*e \in E^+(S)$.

(ii) $\Rightarrow$ (iii): Take any $x \in S$ with $x^n \in E^+(S)$ for some $n > 1$. Let $x^* \in V^\bullet(x)$. Then $x^*x^n \in E^+(S)$ and so $(x^*x)x^{n-1} \in E^+(S)$. But $x^*x \in E^\bullet(S)$ and thus $x^{n-1}x^*x \in E^+(S)$. This leads to $x^{n-2}xx^*x = x^{n-1} \in E^+(S)$. Continuing this process, we have $x \in E^+(S)$.

(iii) $\Rightarrow$ (iv): This is obvious.

(iv) $\Rightarrow$ (v): Let $x, y \in S$ be such that $xy \in E^+(S)$. Now $(yx)^2 = y(xy)x \in E^+(S)$. Hence by given condition we have $yx \in E^+(S)$.

(v) $\Rightarrow$ (i): This is obvious.

Thus (i), (ii), (iii), (iv) and (v) are equivalent.

(vi) $\Rightarrow$ (i): Let $e, f \in E^\bullet(S)$ be such that $ef \in E^+(S)$. Because $(S, \cdot)$ is orthodox we have $fe \in E^\bullet(S)$. Then $fe = (fe)^2 = f(ef)e \in E^+(S)$. Thus the proof is completed.

We now generalize Chaptal's Theorem in S.

Theorem 2.6 In a semiring S the following conditions are equivalent.

(i) $(S, \cdot)$ is a union of groups.

(ii) $(S, \cdot)$ is an inverse semigroup.

(iii) $(S, \cdot)$ is a semilattice of groups.

Proof (i) $\Rightarrow$ (ii): Let $(S, \cdot)$ be a union of groups $(G_\alpha, \cdot)(\alpha \in I)$ where I is an index set. Let $e \in E^\bullet(S)$ and $y \in S$. Then,

$$\begin{aligned}(ye + ey'e)^2 &= ye(ye + ey'e) + ey'e(ye + ey'e)\\ &= yeye + yey'e + ey'eye + ey'ey'e \in E^+(S).\end{aligned}$$

Let $(ye + ey'e)^2$ be in the group G_α for some $\alpha \in I$ and let z be the inverse of $(ye + ey'e)$ in G_α. Then $ye + ey'e = (ye + ey'e)(ye + ey'e)z = (ye + ey'e)^2 z \in E^+(S)$, because $E^+(S)$ is an ideal of S. Also, $eye + ey'e = e(ye + y'e) = (ye + y'e)e$ (by condition (A)) $= ye + y'e$. Thus, by Lemma 2.4., we at once have $ye = eye$. Similarly, we have $ey = eye$. Hence $ey = ye$. Thus idempotents in $(S, \cdot)$ are central. Hence, $(S, \cdot)$ is an inverse semigroup.

(ii) $\Rightarrow$ (iii): Let $(S, \cdot)$ be an inverse semigroup. Let $e \in E^\bullet(S)$ and $y \in S$. Now,

$$\begin{aligned}(ye + ey'e)^2 &= ye(ye + ey'e) + ey'e(ye + ey'e)\\ &= yeye + yey'e + ey'eye + ey'ey'e \in E^+(S).\end{aligned}$$

So by (iv) of Lemma 2.5., we have $ye + ey'e \in E^+(S)$. Also, $eye + ey'e = e(ye + y'e) = (ye + y'e)e = ye + y'e$. Hence by Lemma 2.4., we at once have $ye = eye$. Similarly, $ey = eye$. Hence $ey = ye$. Thus idempotents in $(S, \cdot)$ are central. Thus $(S, \cdot)$ is a Clifford semigroup. Hence $(S, \cdot)$ is a semilattice of groups.

(iii) $\Rightarrow$ (i):This is obvious. □

We now prove the following theorem.

Theorem 2.7 If the semiring S is multiplicatively regular then the following conditions are equivalent.

(i) $(S, \cdot)$ is orthodox.

(ii) $(S, \cdot)$ is an inverse semigroup.

Proof (i) $\Rightarrow$ (ii): Let $(S, \cdot)$ be orthodox. Let $e, f \in E^\bullet(S)$. Then $e(f + e'f) = ef + ee'f = ef + eef' = ef + ef' \in E^+(S)$. So by (ii) of Lemma 2.5., we have

$(f + e'f)e \in E^+(S), i.e., fe + e'fe \in E^+(S)$. Also, $efe + e'fe = efe + ef'e = e(fe + f'e) = (fe + f'e)e$ (by condition (A)) $= fe + f'e$. Thus, by Lemma 2.4., we have $efe = fe$. Similarly, we can show that $efe = ef$. Thus, $ef = fe$. So idempotents in $(S, \cdot)$ commutes. Hence, $(S, \cdot)$ is an inverse semigroup.

(ii) $\Rightarrow$ (i): This is obvious. □

We now generalize Zeleznekow's Theorem in a semiring S.

Theorem 2.8 If the semiring S is multiplicatively regular then the following conditions are equivalent.

(i) $(S, \cdot)$ is orthodox.

(ii) $(S, \cdot)$ is a union of groups.

(iii) $(S, \cdot)$ is an inverse semigroup.

(iv) $(S, \cdot)$ is a semilattice of groups.

Proof Follows from Theorem 2.6. and Theorem 2.7. □

References

[1] Chaptal, N.: *Anneaux dont le demi groupe multiplicatif est inverse.* C. R. Acad. Sci. Paris, Ser. A-B, 262 (1966), 247–277.

[2] Golan, J. S.: The Theory of Semirings with Applications in Mathematics and Theoretical Computer Science. *Pitman Monographs and Surveys in Pure and Applied Mathematics 54, Longman Scientific*, 1992.

[3] Howie, J. M., Introduction to the theory of semigroups. *Academic Press*, 1976.

[4] Karvellas, P. H., *Inverse semirings.* J. Austral. Math. Soc. **18** (1974), 277–288.

[5] Zeleznekow, J.: *Regular semirings.* Semigroup Forum **23** (1981), 119–136.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **43** (2004) 155–169

Further Results for some Third Order Differential Systems with Nonlinear Dissipation *

Awar Simon UKPERA

Department of Mathematics,
Obafemi Awolowo University, Ile-Ife, Nigeria
e-mail: aukpera@oauife.edu.ng

(Received June 9, 2003)

Abstract

We formulate nonuniform nonresonance criteria for certain third order differential systems of the form $X^{'''} + AX^{''} + G(t, X^{'}) + CX = P(t)$, which further improves upon our recent results in [12], given under sharp nonresonance considerations. The work also provides extensions and generalisations to the results of Ezeilo and Omari [5], and Minhós [9] from the scalar to the vector situations.

Key words: Nonlinear dissipation, sharp and nonuniform nonresonance.

2000 Mathematics Subject Classification: 34B15, 34C15, 34C25

1 Introduction

An investigation of the solvability circumstances for the nonlinear differential system

$$X^{'''} + AX^{''} + G(t, X^{'}) + CX = P(t) \tag{1.1}$$

subject to the T-periodic boundary conditions

$$X(0) - X(T) = X^{'}(0) - X^{'}(T) = X^{''}(0) - X^{''}(T) = 0 \tag{1.2}$$

*Research supported by Obafemi Awolowo University Research Grant Code 1425TK.

on $[0,T]$ with $T>0$, was initiated in our recent paper [12]. Our basic motivation has been to provide vector analogues to some existing results in the literature for several scalar prototypes such as those contained in [1], [2], [4] and [5]. For instance, Ezeilo and Omari [5] studied firstly the 2π-periodic solutions associated with the scalar version of (1.1), with $g=g(x^{'})$, satisfying the sharp nonresonance conditions

$$(g_1) \qquad k^2+\alpha^-(|y|)<\frac{g(y)}{y}<(k+1)^2-\alpha^+(|y|), \quad k\in\mathbb{N},$$

where $\alpha^{\pm}:(0,+\infty)\to\mathbb{R}$ are two nonincreasing functions such that

$$\lim_{|y|\to+\infty}|y|\,\alpha^{\pm}(|y|)=+\infty,$$

This result has been improved by Minhós [9] by weakening the condition on the oscillation of g, with the condition (g_1) replaced by the two conditions

$$(g_2) \qquad k^2\le\liminf_{|y|\to\pm\infty}\frac{g(y)}{y}\le\limsup_{|y|\to\pm\infty}\frac{g(y)}{y}\le(k+1)^2$$

and

$$(\mathcal{G}) \qquad k^2<\limsup_{y\to+\infty}\frac{2\mathcal{G}(y)}{y^2}, \qquad \liminf_{y\to+\infty}\frac{2\mathcal{G}(y)}{y^2}<(k+1)^2)$$

where $\mathcal{G}$ denotes the primitive of the nonlinear function g, that is,

$$\mathcal{G}(y)=\int_0^y g(\tau)\,d\tau$$

Here, the ratio $\frac{g(y)}{y}$ may interact with the spectrum $\{k^2,\ k\in\mathbb{N}\}$, although $(\mathcal{G})$ imposes some 'density' control given by the asymptotic behaviour of the primitive of g.

Moreover, when $g=g(t,x^{'})$, nonuniform assumptions

$$(g_3) \qquad k^2\le\gamma^-(t)\le\liminf_{|y|\to\infty}\frac{g(t,y)}{y}\le\limsup_{|y|\to\infty}\frac{g(t,y)}{y}\le\gamma^+(t)\le(k+1)^2$$

uniformly in $y\in\mathbb{R}$ for a.e. $t\in[0,2\pi]$, where $\gamma^{\pm}\in L^1(0,2\pi)$ such that strict inequalities hold on subsets of $[0,2\pi]$ of positive measure; were also established in [5] for the existence of 2π-periodic solutions, with accompanying uniqueness results given by appropriate modification of these conditions.

Our earlier objective, in [12], to generalise some of these results has been partially addressed with the generation of the sharp nonresonance hypotheses

$$(\mathcal{G}_1) \qquad k^2\omega^2+\alpha^-(\|Y\|)\le\frac{\langle G(t,Y),Y\rangle}{\|Y\|^2}\le(k+1)^2\omega^2-\alpha^+(\|Y\|)\ ,$$

uniformly in $Y \in \mathbb{R}^n$ with $\|Y\| \geq r > 0$, and a.e. $t \in [0, T]$, where $k \in \mathbb{N}$, $\omega = \frac{2\pi}{T}$, and $\alpha^{\pm} : \mathbb{R}^n_+ \to \mathbb{R}$ are two functions which are such that

$$(\mathcal{G}_2) \qquad \lim_{\|Y\| \to +\infty} \|Y\| \, \alpha^{\pm}(\|Y\|) = +\infty$$

for the existence of T-periodic solutions to (1.1)–(1.2). These relations clearly generalise the sharp nonresonance conditions prescribed in [5].

There are however, certain equations of type (1.1) with G not satisfying $(\mathcal{G}_1)$–$(\mathcal{G}_2)$, for which, nevertheless, T-periodic solvability results appear to be provable, subject to some other generalisations on G. An example is the system

$$X^{'''} + AX^{''} + \frac{1}{2}\big((k+1)^2\omega^2 + k^2\omega^2 + (2k+1)\omega^2 \cos t\big)X^{'} + CX = P(t) \quad (1.3)$$

with the ratio

$$\frac{\langle G(t,Y), Y\rangle}{\|Y\|^2} = \frac{1}{2}\big((k+1)^2\omega^2 + k^2\omega^2 + (2k+1)\omega^2 \cos t\big)$$

lying in the open interval $\big(k^2\omega^2, (k+1)^2\omega^2\big)$ for a.e. $t \in [0, T]$, but for which there do not exist functions $\alpha^{\pm}$ satisfying $(\mathcal{G}_2)$ for which $(\mathcal{G}_1)$ holds (since the ratio touches both (possible) eigenvalues as $(k+1)^2 - k^2 = 2k+1$). This justifies a further treatment of (1.1) incorporating g_2 and g_3 along the lines of [3], [7], [8] and [10], which clearly specifies the growth pattern and asymptotic conditions on G, unlike the rather arbitrary assumptions employed in [11]. This article proposes some generalisations in this direction.

Note also that condition $(\mathcal{G}_2)$ cannot be dropped as shown by the nonlinear system

$$X^{'''} + AX^{''} + k^2\omega^2 X^{'} + \tan^{-1}(X^{'}) + CX = P(t) \quad (1.4)$$

Here, the ratio

$$\frac{\langle G(t,Y), Y\rangle}{\|Y\|^2} = k^2\omega^2 + \|Y\|^{-1} \tan^{-1}(Y) \ ,$$

with

$$\alpha^{-}(\|Y\|) = \|Y\|^{-1} \tan^{-1}(Y) \quad \text{and} \quad \alpha^{+}(\|Y\|) = 2k\omega^2$$

but

$$\lim_{\|Y\| \to \infty} \|Y\| \alpha^{-}(\|Y\|) = \frac{\pi}{2} \neq +\infty,$$

so that $(\mathcal{G}_2)$ is not fulfilled by α^- and therefore, the system has no T-periodic solution.

Accordingly, $X \in \mathbb{R}^n$, A and C are constant real $n \times n$ nonsingular matrices, and $G : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $P : [0,T] \to \mathbb{R}^n$ are n-vectors, which are T-periodic in t. We shall assume further that G satisfies the Carathéodory conditions, that is, $G(\cdot, X^{'})$ is measurable for every $X^{'} \in \mathbb{R}^n$; $G(t, \cdot)$ is continuous for a.e. $t \in [0,T]$, and for each $r > 0$, there exists an integrable function $\gamma_r \in L^1([0,T], \mathbb{R})$ such that $||G(t, X^{'})|| \leq \gamma_r(t)$, for $||X^{'}|| \leq r$ and a.e. $t \in [0,T]$.

Let X be a point of the Euclidean space $\mathbb{R}^n$ equipped with the usual norm $\|X\|$. For any pair $X, Y \in \mathbb{R}^n$, we shall write $\langle X, Y\rangle$ for the usual scalar product of X and Y so that in particular, $\langle X, X\rangle = \|X\|^2$.

It is standard result that if D is a real $n \times n$ symmetric matrix, then for any $X \in \mathbb{R}^n$,

$$\delta_d ||X||^2 \leq \langle DX, X\rangle \leq \Delta_d ||X||^2, \tag{1.6}$$

where δ_d and Δ_d are respectively the least and greatest eigenvalues of D. In general, $\lambda_i(D)$ shall denote the eigenvalues of any matrix D, and $\|D\|_2$ its spectral norm.

The following Banach spaces will also be frequently refered to:

(i) the classical spaces of k times continuously differentiable functions $C^k([0, T], \mathbb{R}^n), k \geq 0$ an integer, where $C^0 = C$ and $C^\infty = \cap_{k\geq 0} C^k$ with norms $\|X\|_{C^k}$ and $\|X\|_\infty$ respectively;

(ii) the space of T-periodic functions $C^k_T([0, T], \mathbb{R}^n)$ defined by

$$C^k_T = \{X : [0, T] \to \mathbb{R}^n : X \in C^k \text{ and } X \text{ is } T\text{-periodic}\}$$

with the norm on C^k ;

(iii) $L^p([0, T], \mathbb{R}^n)$, $1 \leq p < +\infty$, the usual Lebesgue spaces with the norms $\|X\|_{L^p}$ and $\|X\|_\infty$ for $p = +\infty$;

(iv) the Sobolev space $W^{k,p}_T([0, T], \mathbb{R}^n)$, of T-periodic functions of order k, defined by

$$\begin{aligned} W^{k,p}_T = \{X : [0, T] \to \mathbb{R}^n : X, X', \ldots, X^{(k-1)} &\text{ are absolutely continuous} \\ \text{on } [0, T],\ X^{(k)} \in L^p(0, T) &\text{ and } X^{(i)}(0) - X^{(i)}(T) = 0, \\ &i = 0, 1, 2, \ldots, k-1,\ k \in \mathbb{N}\} \end{aligned}$$

with corresponding norm $\|X\|_{W^{k,p}_T}$;

(v) The Hilbert space $H^1([0, T], \mathbb{R}^n)$ defined by

$$\begin{aligned} H^1(0, T) = \{X : [0, T] \to \mathbb{R}^n : X, &\text{ is absolutely continuous on } [0, T], \\ X' \in L^2(0, T) &\text{ and } X^{(i)}(0) - X^{(i)}(T) = 0,\ i = 0, 1\} \end{aligned}$$

with norm

$$\|X\|_{H^1} = \left\{\sum_{i=1}^{n}\left[\left(\frac{1}{T}\int_0^T x_i(t)\,dt\right)^2 + \frac{1}{T}\int_0^T \big(x_i(t)\big)^2 dt + \frac{1}{T}\int_0^T \big(x_i'(t)\big)^2 dt\right]\right\}^{\frac{1}{2}}.$$

Let

$$\widetilde{H}^1(0, T) = \left\{X \in H^1(0, T) \,\Big|\, \frac{1}{T}\int_0^T X(t)\,dt = 0\right\}$$

2 Previous investigations and some preliminary results

Consider the eigenvalue problem

$$X^{'''} + AX^{''} + CX = -\lambda X^{'} \tag{2.1}$$

together with (1.2), with A, C nonsingular, and λ a real parameter. It has been shown in [5] that

(i) any $\lambda \neq k^2\omega^2$, for each $k = 1, 2, \ldots$, is not an eigenvalue; and

(ii) $\lambda = k^2\omega^2$, for some $k = 1, 2, \ldots$, is an eigenvalue if and only if $C = k^2\omega^2 A$.

Let $\mathcal{E}_k$ be the eigenspace corresponding to the unique eigenvalue $k^2\omega^2$, when it exists. Then we deduce from [9] the following result:

For every $X \in W_T^{3,2}(0, 2\pi)$, we have

$$\int_0^T \langle X^{'''} + AX^{''} + k^2\omega^2 X^{'} + CX,\ X^{'''} + AX^{''} + (k+1)^2\omega^2 X^{'} + CX\rangle\, dt \geq 0, \tag{2.2}$$

and the equality holds if and only if $X = 0$ or either $k^2\omega^2$ or $(k+1)^2\omega^2$ is an eigenvalue of (2.1) and $X \in \mathcal{E}_k$ or $X \in \mathcal{E}_{k+1}$, respectively.

Each of the statements (i) or (ii) has an important bearing on the solvability of the PBVP for the non-autonomous system

$$X^{'''} + AX^{''} + \lambda X^{'} + CX = P(t) \tag{2.3}$$

with $P \in L^1$.

It is clear for instance, from (i) and the Fredholm alternative, that a solution for (1.1)–(1.2) can be expected if the ratio $\langle G(t, X^{'}), X^{'}\rangle / \|X^{'}\|^2$ is such that

$$k^2\omega^2 < \frac{\langle G(t, X^{'}), X^{'}\rangle}{\|X^{'}\|^2} < (k+1)^2\omega^2,$$

for $\|X^{'}\|$ sufficiently large, and a.e. $t \in [0, T]$, provided that some control is put on the closeness of the ratio to $k^2\omega^2$ and $(k+1)^2\omega^2$. This expectation has resulted in the evolution of conditions $(\mathcal{G}_1) - (\mathcal{G}_2)$.

The main role of statement (ii) is to provide an adequate background against which the sharpness of our conditions on G can be tested. Observe that $\alpha^{\pm}$ considered in $(\mathcal{G}_1)$ can be infinitesimal as $\|Y\| \to +\infty$, but by $(\mathcal{G}_2)$ their order must be less than one. This implies that the ratio can approach the (possible) eigenvalues $k^2\omega^2$ and $(k+1)^2\omega^2$, provided that the approach is not too fast. For instance, conditions $(\mathcal{G}_1) - (\mathcal{G}_2)$ admit functions G such as

$$G(Y) = k^2 Y - \|Y\|^{\alpha}\,\mathrm{sgn}(Y), \quad m \in \mathbb{N},\ 0 < \alpha < 1,$$

satisfying

$$\lim_{\|Y\| \to +\infty} \frac{\langle\, G(Y), Y\,\rangle}{\|Y\|^2} = k^2,$$

and yet by the statement (ii), (2.3)–(1.2) with $\lambda = k^2$, does not have a solution in general, that is, for unrestricted A and C nonsingular. Thus for (1.1), we seek conditions on $G(t, Y)$ allowing $\lim_{\|Y\|\to+\infty} \frac{\langle G(t,Y),Y\rangle}{\|Y\|^2}$ (if it exists) to touch k^2, $k \in \mathbb{N}$, for many values of t.

In the sequel, we shall require some preliminary lemmas.

Lemma 2.1 *Consider the linear homogeneous system*

$$X^{'''}(t) + AX^{''}(t) + B(t)X^{'}(t) + CX(t) = 0 \tag{2.4}$$

where A is an arbitrary matrix, C is a nonsingular matrix and $B(t) \equiv (b_{ij}(t))$ is such that $b_{ij} \in L^1(0, T)$ and

$$(\mathcal{B}_1) \qquad k^2\omega^2 \le \lambda_i(B(t)) \le (k+1)^2\omega^2$$

for a.e. $t \in [0, T]$, $i = 1, \dots, n$, $k \in \mathbb{N}$, with the strict inequality holding on subsets of $[0, T]$ of positive measure.

Then, (2.4)–(1.2) has no non-trivial solution.

Proof Let the solution $X(t) = \overline{X}(t) + \widetilde{X}(t)$ have the Fourier expansion

$$X(t) \sim \sum_{i=1}^{n}\left(c_{0,i} + \sum_{k=1}^{\infty}(c_{k,i}\cos k\omega t + d_{k,i}\sin k\omega t)\right),$$

such that

$$\overline{X} = \sum_{i=1}^{n}\left(c_{0,i} + \sum_{k=1}^{N}(c_{k,i}\cos k\omega t + d_{k,i}\sin k\omega t)\right)$$

and

$$\widetilde{X} = \sum_{i=1}^{n}\sum_{k=N+1}^{\infty}(c_{k,i}\cos k\omega t + d_{k,i}\sin k\omega t\,),$$

for some integer $N > 0$ with $N^2\omega^2 < \lambda < (N+1)^2\omega^2$, where $\omega = \frac{2\pi}{T}$.

Then, multiplying (2.4) by $\overline{X}^{'}(t) - \widetilde{X}^{'}(t)$ and integrating over $[0.T]$ gives,

$$\int_0^T \left(\left(\widetilde{X}^{''}(t)\right)^2 - \left\langle B(t)\widetilde{X}^{'}(t), \widetilde{X}^{'}(t)\right\rangle\right) dt$$
$$- \int_0^T \left(\left(\overline{X}^{''}(t)\right)^2 - \left\langle B(t)\overline{X}^{'}(t), \overline{X}^{'}(t)\right\rangle\right) dt = 0. \tag{2.5}$$

Let δ be a constant defined by

$$\delta = \quad \frac{1}{2}\left(\min \lambda_i\left(B(t)\right) + \max \lambda_i\left(B(t)\right)\right) \tag{2.6}$$

for a.e. $t \in [0, T]$. Then in fact,

$$\begin{array}{ll} k^2\omega^2 \le \delta \le (k+1)^2\omega^2, & \text{for a.e. } t \in [0,T], \text{ and} \\ k^2\omega^2 < \delta < (k+1)^2\omega^2, & \text{on subsets of } [0,T] \text{ of positive measure.} \end{array} \tag{2.7}$$

Thus, combining $(\mathcal{B}_1)$, (2.6) and (2.7), (2.5) becomes

$$0 \geq \int_0^T \left[\left(\widetilde{X}''(t)\right)^2 - \delta \left(\widetilde{X}'(t)\right)^2 \right] dt - \int_0^T \left[\left(\overline{X}''(t)\right)^2 - \delta \left(\overline{X}'(t)\right)^2 \right] dt = 0. \tag{2.8}$$

By Parseval's identity given by

$$\int_0^T \|X\|^2 \, dt = \sum_{i=1}^{n} \left(c_{0,i}^2 T + \frac{T}{2} \sum_{k=1}^{\infty} (c_{k,i}^2 + d_{k,i}^2) \right),$$

(2.8) becomes

$$\frac{T}{2} \sum_{i=1}^{n} \left[\sum_{k=N+1}^{\infty} k^2\omega^2(k^2\omega^2 - \delta)(c_{k,i}^2 + d_{k,i}^2) + \sum_{k=1}^{N} k^2\omega^2(\delta - k^2\omega^2)(c_{k,i}^2 + d_{k,i}^2) \right] = 0. \tag{2.9}$$

It follows from (2.7) that $c_{k,i} = 0$ $(k = 0, 1, 2, \ldots)$ and $d_{k,i} = 0$ $(k = 1, 2, \ldots)$, for all $i = 1, \ldots, n$. Thus, $X \equiv 0$, and the lemma follows. □

Lemma 2.2 *Let C be nonsingular, and assume that $M, N \in L^1([0,T], \mathbb{R}^{n^2})$ are nonsingular matrices which satisfy the following conditions*

$$k^2\omega^2\|Y\|^2 \leq \langle M(t)Y, Y\rangle \leq \langle N(t)Y, Y\rangle \leq (k+1)^2\omega^2\|Y\|^2 \tag{2.10}$$

uniformly in $Y \in \mathbb{R}^n$, for a.e. $t \in [0,T]$, $k \in \mathbb{N}$, $\omega = \frac{2\pi}{T}$, and

$$k^2\omega^2\|Y\|^2 < \langle M(t)Y, Y\rangle, \qquad \langle N(t)Y, Y\rangle < (k+1)^2\omega^2\|Y\|^2 \tag{2.11}$$

on subsets of $[0,T]$ of positive measure.

Then, there exists constants $\epsilon = \epsilon(M, N, C) > 0$ and $\delta_0 = \delta_0(M, N, C) > 0$ uniformly a.e. on $[0,T]$, such that for all $B(t) \equiv (b_{ij}(t))$ with $b_{ij} \in L^1([0,T], \mathbb{R})$ satisfying

$$(\mathcal{B}_2) \qquad \langle M(t)Y, Y\rangle - \epsilon\|Y\|^2 \leq \langle B(t)Y, Y\rangle \leq \langle N(t)Y, Y\rangle + \epsilon\|Y\|^2$$

uniformly in $Y \in \mathbb{R}^n$, a.e. on $[0,T]$, and all $X \in W_T^{3,1}([0,T], \mathbb{R}^n)$, one has

$$\|X''' + AX'' + B(\cdot)X' + CX\|_{L^1} \geq \delta_0 \|X\|_{W_T^{3,1}} \tag{2.12}$$

Proof Let us assume that the conclusion of the Lemma does not hold, that is, ϵ and δ_0 do not exist. Then, there exists a sequence $(X_n) \in W^{3,1}([0,T], \mathbb{R}^n)$ with $\|X_n\|_{W^{3,1}} = 1$, and a sequence $(B_n) \in L^1([0,T], \mathbb{R}^{n^2})$ of nonsingular matrices with

$$\langle M(t)Y, Y\rangle - \frac{1}{n}\|Y\|^2 \leq \langle B_n(t)Y, Y\rangle \leq \langle N(t)Y, Y\rangle + \frac{1}{n}\|Y\|^2, \quad n \in \mathbb{N}, \tag{2.13}$$

uniformly in $Y \in \mathbb{R}^n$, for a.e. $t \in [0,T]$, such that for all $X \in W^{3,1}$, one has

$$\int_0^T \|X_n'''(t) + AX_n''(t) + B_n(t)X_n'(t) + CX_n\| \, dt < \frac{1}{n}\,. \tag{2.14}$$

Let $\|B_n\|$ denote the norm of B_n. Then, by (2.13), there exists some $\beta \in L^1([0,T],\mathbb{R})$ such that

$$\|B_n(t)\| \leq \beta(t), \quad n = 1, 2, \ldots \tag{2.15}$$

for a.e. $t \in [0,T]$, $n \in \mathbb{N}$. For example, one can take

$$\beta(t) \equiv \frac{1}{\|Y\|^2}\left[\|\langle M(t)Y, Y\rangle - \langle Y, Y\rangle\| + \|\langle N(t)Y, Y\rangle + \langle Y, Y\rangle\|\right].$$

Now, by the compact embedding of $W^{3,1}([0,T],\mathbb{R}^n)$ into $W^{2,1}([0,T],\mathbb{R}^n)$ and the continuous embedding of $W^{2,1}([0,T],\mathbb{R}^n)$ into $C^1([0,T],\mathbb{R}^n)$ imply that by going to subsequences if neccessary, we can assume that

$$X_n \to X \text{ in } C^1([0,T],\mathbb{R}^n), \quad X_n^{''} \to X^{''} \text{ in } L^\infty([0,T],\mathbb{R}^n) \subset L^1([0,T],\mathbb{R}^n). \tag{2.16}$$

Moreover, by (2.15), we deduce that

$$B_n \rightharpoonup B \text{ in } L^1([0,T],\mathbb{R}^{n^2}) \tag{2.17}$$

so that by (2.13),

$$\langle M(t)Y, Y\rangle \leq \langle B(t)Y, Y\rangle \leq \langle N(t)Y, Y\rangle \tag{2.18}$$

for a.e. $t \in [0,T]$.

On the other hand, for every $\Phi \in L^\infty([0,T],\mathbb{R}^n)$, we have by Schwarz inequality

$$\begin{aligned} &\left\|\int_0^T \left\langle B_n(t)X_n^{'}(t) - B(t)X^{'}(t), \Phi(t)\right\rangle dt\right\| \\ \leq &\left\|\int_0^T \left\langle B_n(t)\big(X_n^{'}(t) - X^{'}(t)\big), \Phi(t)\right\rangle dt\right\| + \left\|\int_0^T \left\langle \big(B_n(t) - B(t)\big)X^{'}(t), \Phi(t)\right\rangle dt\right\| \\ &\leq \|\Phi\|_\infty \|\beta\|_{L^1} \|X_n^{'} - X^{'}\|_\infty + \left\|\int_0^T \left\langle \big(B_n(t) - B(t)\big)X^{'}(t), \Phi(t)\right\rangle dt\right\|. \end{aligned} \tag{2.19}$$

The right hand side of (2.19) tends to zero by (2.16) and (2.17), and we deduce that

$$B_n X_n^{'} \rightharpoonup BX^{'} \text{ in } L^1([0,T],\mathbb{R}^n). \tag{2.20}$$

By (2.14), (2.16) and (2.20), it follows that

$$X_n^{'''} = -AX_n^{''} - B_n(\cdot)X_n^{'} - CX_n \rightharpoonup -AX^{''} - B(\cdot)X^{'} - CX \text{ in } L^1([0,T],\mathbb{R}^n). \tag{2.21}$$

Since the operator

$$\frac{d^3}{dt^3} : W^{3,1}([0,T],\mathbb{R}^n) \subset L^1([0,T],\mathbb{R}^n) \to L^1([0,T],\mathbb{R}^n)$$

is weakly closed, this implies (by (2.16) and (2.21)) that $X \in W_T^{3,1}([0,T],\mathbb{R}^n)$, and $X^{'''} = -AX^{''} - B(\cdot)X^{'} - CX$, that is,

$$X^{'''}(t) + AX^{''}(t) + B(t)X^{'}(t) + CX(t) = 0, \tag{2.22}$$

for a.e. $t \in [0,T]$ and $X \in W^{3,1}([0,T],\mathbb{R}^n)$.

It follows from (2.9), (2.10), (2.18), (2.22) and Lemma 2.1 that $X \equiv 0$, that is, $X_n \to 0$ in $W^{3,1}([0,T],\mathbb{R}^n)$ as $n \to \infty$. But this clearly contradicts the initial assumption that $\|X_n\|_{W^{3,1}} = 1$ for all n, and the proof is complete. □

Lemma 2.3 *Let $D \in L^1\big([0,T],\mathbb{R}^{n^2}\big)$ be a nonsingular matrix such that $0 \le \lambda_i(D(t)) \le \omega^2$ a.e. on $[0,T]$, with the strict inequality holding on a subset of $[0,T]$ of positive measure. Then, there exists a constant $\eta = \eta(D) > 0$ such that for all $\widetilde{X} \in \widetilde{H}^1([0,T],\mathbb{R}^n)$, we have*

$$\frac{1}{T}\int_0^T \left[\left(\widetilde{X}^{\,'}(t)\right)^2 - \langle D(t)\widetilde{X}(t), \widetilde{X}(t)\rangle\right] dt \ge \eta\|\widetilde{X}\|_{H^1}^2 \tag{2.23}$$

Proof This is clearly the same as in the proof of Lemma 1 of [8] by setting $\lambda_i(D(t)) \equiv \Gamma_i(t)$, $i = 1,2,\ldots,n$, where $\Gamma_i \in L^1([0,T],\mathbb{R})$ satisfies $\Gamma_i(t) \le \omega^2$ a.e. on $[0,T]$, with the strict inequality holding on a subset of $[0,T]$ of positive measure, and replacing the period 2π by T. □

3 The main results

We now present our main results:

Theorem 3.1 *Let C be a nonsingular matrix. Suppose that G is L^1-Carathéodory and satisfies*

$$(\mathcal{G}_3) \qquad k^2\omega^2 \le \frac{\langle M(t)Y,Y\rangle}{\|Y\|^2} \le \liminf_{\|Y\|\to\infty}\frac{\langle G(t,Y),Y\rangle}{\|Y\|^2} \le \limsup_{\|Y\|\to\infty}\frac{\langle G(t,Y),Y\rangle}{\|Y\|^2} \le \frac{\langle N(t)Y,Y\rangle}{\|Y\|^2} \le (k+1)^2\omega^2$$

uniformly in $Y \in \mathbb{R}^n$ for a.e. $t \in [0,T]$, $k \in \mathbb{N}$ and $M,N \in L^1([0,T],\mathbb{R}^{n^2})$ are such that $k^2\omega^2\|Y\|^2 < \langle M(t)Y,Y\rangle$, $\langle N(t)Y,Y\rangle < (k+1)^2\omega^2\|Y\|^2$ on subsets of $[0,T]$ of positive measure. Then, for any arbitrary matrix A, the system (1.1)–(1.2) has at least one solution for every $P \in L^1([0,T],\mathbb{R}^n)$.

Proof Let $\epsilon > 0$ be as in Lemma 2.2. Then, by $(\mathcal{G}_3)$, we can fix a constant vector $\rho = \rho(\epsilon)$ with each $\rho_i > 0$ such that

$$\langle M(t)Y,Y\rangle - \epsilon\|Y\|^2 \le \langle G(t,Y),Y\rangle \le \langle N(t)Y,Y\rangle + \epsilon\|Y\|^2 \tag{3.1}$$

for a.e. $t \in [0,T]$ and all $Y \in \mathbb{R}^n$ with $|y_i| \ge \rho_i$.

Now define $\nu(t,Y) \equiv (\nu_i(t,Y))_{1\le i\le n} : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ by

$$\nu_i(t,Y) = \begin{cases} y_i^{-1} g_i(t,Y), \text{ if } |y_i| \ge \rho_i; \\ y_i\rho_i^{-2} g_i(t,y_1,\dots,y_{i-1},\rho_i,y_{i+1},\dots,y_n) + (1-\frac{y_i}{\rho_i})\beta(t), \\ \text{if } 0 \le y_i < \rho_i; \\ y_i\rho_i^{-2} g_i(t,y_1,\dots,y_{i-1},-\rho_i,y_{i+1},\dots,y_n) + (1+\frac{y_i}{\rho_i})\beta(t), \\ \text{if } -\rho_i \le y_i < 0. \end{cases}$$

for a.e. $t \in [0,T]$, where β is given by

$$\beta(t) \equiv \frac{1}{\|Y\|^2}\left[\,\|\langle M(t)Y,Y\rangle - \langle Y,Y\rangle\| + \|\langle N(t)Y,Y\rangle + \langle Y,Y\rangle\|\,\right], \tag{3.2}$$

so that by construction and (3.1), we deduce that

$$\langle M(t)Y,Y\rangle - \epsilon\|Y\|^2 \le \langle \nu(t,Y),Y\rangle \le \langle N(t)Y,Y\rangle + \epsilon\|Y\|^2 \tag{3.3}$$

for a.e. $t \in [0,T]$ and $Y \in \mathbb{R}^n$.

The function $\tilde{G} \equiv (\tilde{g}_i(t,Y))_{1\le i\le n}[0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ defined by $\tilde{g}_i(t,Y) = \nu_i(t,Y)y_i$ satisfies the Carathéodory conditions, by construction. Hence, setting $\Psi(t,Y) = G(t,Y) - \tilde{G}(t,Y)$, then $\Psi(t,Y)$ is also L^1-Carathéodory with

$$\|\Psi(t,Y)\| \le \sup_{|y_i|\le\rho_i} \|G(t,Y) - \tilde{G}(t,Y)\| \le \varphi(t) \tag{3.4}$$

for a.e. $t \in [0,T]$ and $Y \in \mathbb{R}^n$, for some $\varphi \in L^1([0,T],\mathbb{R})$ depending only on M, N and γ_r mentioned at the beginning in association with G. Then, the problem (1.1) is equivalent to

$$X^{'''}(t) + AX^{''}(t) + \tilde{G}(t,X^{'}(t)) + \Psi(t,X^{'}(t)) + CX(t) = P(t) \tag{3.5}$$

By the Leray–Schauder technique (see Mawhin [6]), the proof of the Theorem now follows by showing that there is a constant $K > 0$, independent of $\lambda \in (0,1)$, such that $\|X\|_{C^2} < K$, for all possible solutions X of the homotopy

$$X^{'''} + AX^{''} + (1-\lambda)N(t)X^{'} + \lambda\tilde{G}(t,X^{'}) + \lambda\Psi(t,X^{'}) + CX = \lambda P(t) \tag{3.6}$$

We observe from (3.3) that

$$\langle M(t)Y,Y\rangle - \epsilon\|Y\|^2 \le \langle\,(1-\lambda)N(t)Y + \lambda\tilde{G}(t,Y),Y\rangle \le \langle N(t)Y,Y\rangle + \epsilon\|Y\|^2 \tag{3.7}$$

for a.e. $t \in [0,T]$, $Y \in \mathbb{R}^n$ and $\lambda \in [0,1]$.

Thus, we may set $(1-\lambda)N(t)X^{'} + \lambda\tilde{G}(t,X^{'}) \equiv B(t)X^{'}$, for a.e. $t \in [0,T]$, $X^{'} \in \mathbb{R}^n$ and $\lambda \in [0,1]$, where, by (3.7), $B(t)$ is such that

$$\langle M(t)X^{'},X^{'}\rangle - \epsilon\|X^{'}\|^2 \le \langle B(t)X^{'},X^{'}\rangle \le \langle N(t)X^{'},X^{'}\rangle + \epsilon\|X^{'}\|^2 \tag{3.8}$$

for a.e. $t \in [0,T]$, $X^{'} \in \mathbb{R}^n$ and $\lambda \in [0,1]$.

Thus (3.6) becomes

$$0 \geq \|X^{'''} + AX^{''} + B(\cdot)X^{'} + CX\|_{L^1} - \|\Psi(\cdot, X^{'})\|_{L^1} - \|P(\cdot)\|_{L^1} \tag{3.9}$$

Using Lemma 2.2 and (3.4) finally gives

$$0 \geq \delta_0 \|X\|_{W^{3,1}} - \|\delta\|_{L^1} - \|P\|_{L^1} \tag{3.10}$$

which yields a constant $K_0 > 0$ such that $\|X\|_{W^{3,1}} \leq K_0$. Hence, we obtain the required constant $K > 0$ such that $\|X\|_{C^2} < K$, following a standard procedure just as in [2], and the conclusion follows. □

Remark 3.1 The result of Theorem 3.1 can be extended to nonlinear systems of the form

$$X^{'''} + \frac{d}{dt} \operatorname{grad} f(X^{'}) + G(t, X^{'}) + H(X) = P(t), \tag{3.11}$$

under suitable assumptions on G satisfying some requirements in respect of the first (possible) eigenvalue $\lambda = \omega^2$ of (2.1)–(1.2).

Here, $f : \mathbb{R}^n \to \mathbb{R}$ is a C^2-function, $H : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and satisfies a sign condition, while G and P are as specified earlier.

Theorem 3.2 *Assume that G satisfies*

$$(\mathcal{G}_4) \qquad \lim_{\|Y\| \to +\infty} \frac{\langle G(t, Y), Y \rangle}{\|Y\|^2} \leq \frac{\langle N(t)Y, Y \rangle}{\|Y\|^2} \leq \omega^2$$

uniformly in $Y \in \mathbb{R}^n$ for a.e. $t \in [0, T]$, where $N \in L^1([0, T], \mathbb{R}^{n^2})$ is such that $\langle N(t)Y, Y \rangle < \omega^2 \|Y\|^2$ on subsets of $[0, T]$ of positive measure.

Moreover, suppose that H satisfies

$$(\mathcal{H}) \qquad \lim_{\|X\| \to +\infty} \operatorname{sgn}(X)\, H(X) = +\infty.$$

Then, (3.11)–(1.2) has at least one solution for every $P \in L^1([0, T], \mathbb{R}^n)$.

Proof As in the preceding proof, for each $\epsilon > 0$, there exists $\rho = \rho(\epsilon) > 0$ such that

$$\langle G(t, Y), Y \rangle \leq \langle N(t)Y, Y \rangle + \epsilon \|Y\|^2$$

for a.e. $t \in [0, T]$ and all $Y \in \mathbb{R}^n$ with $|y_i| \geq \rho_i$.

Then, define $\tilde{G}(t, Y)$ and $\Psi(t, Y)$ as before, so that the relations

$$\langle (1 - \lambda) N(t) Y + \lambda \tilde{G}(t, Y), Y \rangle \leq \langle N(t)Y, Y \rangle + \epsilon \|Y\|^2, \quad \lambda \in [0, 1]$$

and

$$\|\Psi(t, Y)\| \leq \varphi(t)$$

hold, for a.e. $t \in [0, T]$ and every $Y \in \mathbb{R}^n$.

It suffices to establish the neccessary (or appropriate) a-priori bounds for the λ-dependent family of systems

$$X^{'''} + \lambda\frac{d}{dt}\operatorname{grad} f(X^{'}) + (1-\lambda)N(t)X^{'} + \lambda\tilde{G}(t,X^{'}) + \lambda\Psi(t,X^{'}) \\ + (1-\lambda)CX + \lambda H(X) = \lambda P(t), \tag{3.12}$$

for $\lambda \in [0,1]$, where C is a fixed nonsingular and positive definite matrix.

Let X be a solution of (3.12)–(1.2). Taking the scalar product of (3.12) with $X^{'}(t)$ and integrating over $[0,T]$ using (1.2) gives

$$\int_0^T \|X^{''}\|^2\,dt = \int_0^T \langle (1-\lambda)N(t)X^{'} + \lambda\tilde{G}(t,X^{'}),X^{'}\rangle\,dt + \langle \Psi(\cdot,X^{'}) - P(\cdot),X^{'}\rangle_{L^2} \tag{3.13}$$

That is, from above

$$\|X^{''}\|_{L^2}^2 \le \int_0^T \langle\, N(t)X^{'}(t),X^{'}(t)\,\rangle\,dt + \epsilon\|X^{'}\|_{L^2}^2 + (\|\varphi\|_{L^1} + \|P\|_{L^1})\|X^{'}\|_\infty \tag{3.14}$$

Noting that by Lemma 2.3,

$$\|X^{''}\|_{L^2}^2 - \int_0^T \langle\, N(t)X^{'}(t),X^{'}(t)\rangle\,dt =$$

$$= \int_0^T ((X^{''}(t))^2 - \langle\, N(t)X^{'}(t),X^{'}(t)\rangle)\,dt \ge \eta\|X^{'}\|_{H^1}^2 = \frac{\eta}{T}\|X^{''}\|_{L^2}^2,$$

for some constant $\eta = \eta(\Gamma) > 0$, we obtain from (3.14)

$$\eta\|X^{''}\|_{L^2}^2 \le \frac{\epsilon T}{\omega^2}\|X^{''}\|_{L^2}^2 + (\|\varphi\|_{L^1} + \|P\|_{L^1})T^{\frac{3}{2}}\|X^{''}\|_{L^2} \tag{3.15}$$

by the Wirtinger and other standard inequalities. Hence, taking $0 < \epsilon T < \omega^2\eta$, we deduce that

$$\|X^{''}\|_{L^2} \le c_1 \quad , \tag{3.16}$$

for some $c_1 > 0$. Thus, we have

$$\|X^{'}\|_\infty \le \sqrt{T}\|X^{''}\|_{L^2} \le \sqrt{T}c_1 \tag{3.17}$$

This implies that

$$\|X - X(t_0)\| \le T\|X^{'}\|_\infty \le T^{\frac{3}{2}}c_1 \tag{3.18}$$

where $t_0 \in [0,T]$ is arbitrarily fixed.

Now observe that

$$\int_0^T (1-\lambda)N(t)X^{'} + \lambda\tilde{G}(t,X^{'})\,dt \le \int_0^T (\,N(t)X^{'} + \epsilon X^{'}\,)\,dt = 0 \tag{3.19}$$

Then, taking the average of (3.12) on $[0,T]$, we obtain by the Mean Value Theorem,

$$\begin{aligned} &\big\|(1-\lambda)X(t^\star)+\lambda C^{-1}H(X(t^\star))\,\big\| = \\ &= \Big\|(1-\lambda)\Big(\frac{1}{T}\int_0^T X(t)\,dt\Big)+\lambda\Big(\frac{1}{T}\int_0^T C^{-1}H(X(t))\,dt\Big)\Big\| \\ &\le \|C^{-1}\|\left(\frac{1}{T}\|\delta\|_{L^1}+\frac{1}{T}\|P\|_{L^1}\right) := c_2 \end{aligned} \tag{3.20}$$

for some $t^\star \in [0,T]$.

Now by hypothesis $(\mathcal{H})$, it follows that for any $k>0$, there exists a $q=q(k)>0$ such that

$$\|C^{-1}H(X)\| = \|\widetilde{H}(X)\| = \operatorname{sgn}(X)\widetilde{H}(X) > k, \tag{3.21}$$

for every $\|X\|>\max\{k,q\}$, and all positive definite C. Hence, for any $\lambda\in(0,1]$, we have

$$\big\|(1-\lambda)X+\lambda C^{-1}H(X)\,\big\| = \operatorname{sgn}(X)\big((1-\lambda)X+\lambda C^{-1}H(X)\big) \ge (1-\lambda)k+\lambda k = k \tag{3.22}$$

for every $\|X\|>\max\{k,q\}$. Thus, choosing $k>c_2$, it follows that

$$\|X(t^\star)\| \le \max\{k,q\} := c_3 \tag{3.23}$$

Combining (3.18) and (3.23) with $t_0=t^\star$, we obtain

$$\|X\|_\infty \le T^{\frac{3}{2}}c_1 + c_3 := c_4 \tag{3.24}$$

Lastly, integrating (3.12) and using the continuity of H and (3.24), we deduce the existence of a constant $c_5>0$, such that

$$\|X^{'''}\|_{L^1} \le c_5, \tag{3.25}$$

so that

$$\|X^{''}\|_\infty \le T\|X^{'''}\|_{L^1} = Tc_5 \tag{3.26}$$

Therefore, by (3.17), (3.24) and (3.26),

$$\|X\|_{C^2} = \|X\|_\infty + \|X^{'}\|_\infty + \|X^{''}\|_\infty \le c_6, \tag{3.27}$$

for some $c_6>0$, and we are done. □

As pointed out earlier, Theorem 3.2 admits solutions for periodic systems associated with

$$X^{'''} + \frac{d}{dt}\operatorname{grad} f(X^{'}) + \frac{\omega^2}{2}(1+\sin t)X^{'} + H(X) = P(t). \tag{3.28}$$

Finally, we conclude this study with a uniqueness criterion for the system (1.1)–(1.2). The following result holds:

Theorem 3.3 *Let C be nonsingular and suppose that G satisfies, for some $k \in \mathbb{N}$,*

$$(\mathcal{G}_5) \qquad k^2\omega^2 \leq \frac{\langle M(t)(Y_1 - Y_2), Y_1 - Y_2\rangle}{\|Y_1 - Y_2\|^2} \leq \frac{\langle G(t, Y_1) - G(t, Y_2), Y_1 - Y_2\rangle}{\|Y_1 - Y_2\|^2}$$

$$\leq \frac{\langle N(t)(Y_1 - Y_2), Y_1 - Y_2\rangle}{\|Y_1 - Y_2\|^2} \leq (k+1)^2\omega^2,$$

or

$$(\mathcal{G}_6) \qquad \frac{\langle G(t, Y_1) - G(t, Y_2), Y_1 - Y_2\rangle}{\|Y_1 - Y_2\|^2} < \omega^2,$$

uniformly for a.e. $t \in [0, T]$ and $Y_1, Y_2 \in \mathbb{R}^n$ with $Y_1 \neq Y_2$.

Then, (1.1)–(1.2) has at most one solution.

Proof *Case (i)* G subject to $(\mathcal{G}_5)$: The PBVP satisfied by $V = Y_1 - Y_2$, for any two solutions Y_1, Y_2 of (1.1)–(1.2) is of the form

$$V^{'''}(t) + AV^{''}(t) + B^\star(t, V^{'})V^{'}(t) + CV(t) = 0, \tag{3.28}$$

with

$$V(0) - V(T) = V^{'}(0) - V^{'}(T) = V^{''}(0) - V^{''}(T) \tag{3.29}$$

where the matrix $B^\star \in L^1(0, T)$ is defined by

$$B^\star(t, V(t))V(t) = \begin{cases} G(t, V + Y_2) - G(t, Y_2), & \text{if } V \neq 0 \\ M(t), & \text{if } V = 0 \end{cases}$$

and by $(\mathcal{G}_5)$ satisfies

$$\lambda_i(M(t)) \leq \lambda_i\left(B^\star(t, V(t))\right) \leq \lambda_i(N(t))$$

uniformly in $V \in \mathbb{R}^n$ for a.e. $t \in [0, T]$.

Hence, using the arguments of Lemma 2.1, we see that $V \equiv 0$, and the uniqueness, subject to $(\mathcal{G}_5)$, is thus proved.

Case (ii) G subject to $(\mathcal{G}_6)$: We consider the PBVP (3.28)-(3.29) as before except that this time $B^\star$ is defined by

$$B^\star(t, V(t))V(t) = \begin{cases} G(t, V + Y_2) - G(t, Y_2), & \text{if } V \neq 0 \\ 0, & \text{if } V = 0 \end{cases}$$

so that by $(\mathcal{G}_6)$, $\lambda_i\left(B^\star(t, V(t))\right) < \omega^2$ uniformly in $V \in \mathbb{R}^n$ for $t \in [0, T]$.

Multiply now (3.28) scalarly by $V^{'}(t)$ and integrate over $[0, T]$ using (3.29) and we get

$$\int_0^T \|V^{''}(t)\|^2\, dt = \int_0^T \langle\, B^\star(t, V(t))V^{'}(t), V^{'}(t)\rangle\, dt \leq \int_0^T \langle \widetilde{B}(t)V^{'}(t), V^{'}(t)\rangle\, dt, \tag{3.30}$$

where we set $\lambda_i(\widetilde{B}(t)) = \max\{0, \lambda_i(B^\star(t, V(t)))\}$ uniformly in V for a.e. $t \in [0, T]$.

Clearly then, $\widetilde{B}(t) \in L^1(0,T)$ is such that $0 \le \lambda_i(\widetilde{B}(t)) < \omega^2$ for a.e. $t \in [0,T]$. Thus using Lemma 2.3 setting $\widetilde{X} = V'$, (3.30) becomes

$$0 \ge \int_0^T \|V''(t)\|^2\, dt - \int_0^T \langle \widetilde{B}(t)V'(t), V'(t)\rangle\, dt \ge \eta \|V'\|^2_{H^1} \tag{3.31}$$

from which we deduce that $V' \equiv 0$, leading to $V \equiv 0$, and the proof is complete. □

References

[1] Afuwape, A. U., Omari, P., Zanolin, F.: *Nonlinear perturbations of differential operators with nontrivial kernel and applications to third-order periodic boundary value problems.* J. Math. Anal. Appl. **143**, 1 (1989), 35–56.

[2] Andres, J., Vlček, V.: *Periodic solutions of the third order parametric differential equations involving large nonlinearities.* Math. Slovaca **41**, 4 (1991), 337–349.

[3] Conti, G., Iannacci, R., Nkashama, M. N.: *Periodic solutions of Liénard systems at resonance.* Ann. Math. Pura. Appl. **141**, 4 (1985), 313–327.

[4] Ezeilo, J. O. C., Nkashama, M. N.: *Resonant and nonresonant oscillations for third-order nonlinear ordinary differential equations.* Nonlinear Analysis, T.M.A. **12**, 10 (1988), 1029–1046.

[5] Ezeilo, J. O. C., Omari, P.: *Nonresonant oscillations for some third-order Differential equations II.* J. Nigerian Math. Soc. **8** (1989), 25–48.

[6] Mawhin, J.: *Topological Degree Methods in Nonlinear Boundary Value Problems.* In: CBMS Regional Conference Series in Mathematics **40** (1979), American Math. Soc., Providence, R.I.

[7] Mawhin, J., Ward, J. R.: *Nonuniform nonresonance conditions at the two first eigenvalues for periodic solutions of forced Liénard and Duffing equations.* Rocky Mount. J. Math. **12**, 4 (1982), 643–653.

[8] Mawhin, J., Ward, J. R.: *Periodic solutions of some forced Liénard differential equations at resonance.* Arch. Math. **41** (1983), 337–351.

[9] Minhós, F.: *Periodic solutions for a third order differential equation under conditions on the potential.* Portugaliae Math. **55**, 4 (1998), 475–484.

[10] Tejumola, H. O., Afuwape, A. U.: *Periodic solutions of certain third-order nonlinear differential systems with delay.* I.C.T.P. Trieste, Preprint IC/90/418, (1990).

[11] Ukpera, A.S.: *Periodicity results for strongly nonlinear systems of third order boundary value problems.* Anal. Stiint. ale Univ. "Al. I. Cuza" **46** (2000), 215–230.

[12] Ukpera, A.S.: *Periodic solutions of certain third order differential systems with nonlinear dissipation.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **41** (2002), 147–159.

ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
MATHEMATICA 43 (2004)

Published by the Palacký University Olomouc

First edition

ISBN 80-244-0965-8
ISSN 0231-9721