BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 3–18
ISSN 1024–7696

Optimal control for one complex dynamic system, II

Alla Albu, Vladimir Zubov

Abstract. The optimal control problem of the metal solidification in casting is considered. The process is modeled by a three-dimensional two-phase initial-boundary value problem of the Stefan type. A numerical algorithm for solving the direct problem was presented in the first part of this article, published in [1]. The optimal control problem was solved numerically using the gradient method. The gradient of the cost function was found with the help of conjugate problem. The discreet conjugate problem was posed with the help of Fast Automatic Differentiation technique.

Mathematics subject classification: 49J20, 93C20.
Keywords and phrases: Heat conduction, Stefan problem, finite-difference scheme, optimal control, gradient method, Fast Automatic Differentiation technique.

6 Calculation of the gradient in the optimal control problem

6.1 The canonical form of the discrete version of the direct problem

The variational problem formulated in Section 2 (part I) was solved numerically by gradient methods. To calculate the gradient of the function the Fast Automatic Differentiation (FAD) methodology [2] was used.

In accordance with the FAD-methodology, all equations, that approximate the direct problem, have to be presented in a special, so-called, canonical form that we will give below.

For this canonical form to be more compact, let us introduce the following designations.

For all $i=\overline{0,I}$, $\quad l=\overline{0,L}$ let us designate as (X_m) and (X_f) these $(N+2)$-dimentional vectors:

$$(X_m)^j_{0il}=-\left(r_1(\beta^j_{0il})\beta^j_{0il}+q^j_1\right)\Big|_{S^{1x-}_{0il}},\quad (X_f)^j_{0il}=-\left(r_2(\beta^j_{0il})\beta^j_{0il}+q^j_2\right)\Big|_{S^{2x-}_{0il}},$$

$$(X_m)^j_{nil}=R^j_{n-1}\frac{\beta^j_{nil}-\beta^j_{n-1,il}}{h^x_{n-1}},\qquad (X_f)^j_{nil}=B^j_{n-1}\frac{\beta^j_{nil}-\beta^j_{n-1,il}}{h^x_{n-1}},\quad (n=\overline{1,N}),$$

$$(X_m)^j_{N+1,il}=\left(r_1(\beta^j_{Nil})\beta^j_{Nil}+q^j_1\right)\Big|_{S^{1x+}_{Nil}},\ (X_f)^j_{N+1,il}=\left(r_2(\beta^j_{Nil})\beta^j_{Nil}+q^j_2\right)\Big|_{S^{2x+}_{Nil}}.$$

For all $n=\overline{0,N}$, $\quad l=\overline{0,L}$ let us designate as (Y_m) and (Y_f) these $(I+2)$-dimentional vectors:

$$(Y_m)^j_{n0l}=-\left(r_1(\beta^j_{n0l})\beta^j_{n0l}+q^j_1\right)\Big|_{S^{1y-}_{n0l}},\quad (Y_f)^j_{n0l}=-\left(r_2(\beta^j_{n0l})\beta^j_{n0l}+q^j_2\right)\Big|_{S^{2y-}_{n0l}},$$

$$(Y_m)^j_{nil} = \widehat{R}^j_{i-1}\frac{\beta^j_{nil}-\beta^j_{n,i-1,l}}{h^y_{i-1}}, \qquad (Y_f)^j_{nil} = \widehat{B}^j_{i-1}\frac{\beta^j_{nil}-\beta^j_{n,i-1,l}}{h^y_{i-1}}, \quad i=\overline{1,I},$$

$$(Y_m)^j_{n,I+1,l}=\Big(r_1(\beta^j_{nIl})\beta^j_{nIl}+q^j_1\Big)\Big|_{S^{1y+}_{nIl}},\ (Y_f)^j_{n,I+1,l}=\Big(r_2(\beta^j_{nIl})\beta^j_{nIl}+q^j_2\Big)\Big|_{S^{2y+}_{nIl}}.$$

For all $n=\overline{0,N}$, $i=\overline{0,I}$ let us designate as (Z_m) and (Z_f) these $(L+2)$-dimentional vectors:

$$(Z_m)^j_{ni0}=-\left(r_1(\beta^j_{ni0})\beta^j_{ni0}+q^j_1\right)\Big|_{S^{1z-}_{ni0}}, \quad (Z_f)^j_{ni0}=-\left(r_2(\beta^j_{ni0})\beta^j_{ni0}+q^j_2\right)\Big|_{S^{2z-}_{ni0}},$$

$$(Z_m)^j_{nil} = \widetilde{R}^j_{l-1}\frac{\beta^j_{nil}-\beta^j_{ni,l-1}}{h^z_{l-1}}, \qquad (Z_f)^j_{nil} = \widetilde{B}^j_{l-1}\frac{\beta^j_{nil}-\beta^j_{ni,l-1}}{h^z_{l-1}}, \quad l=\overline{1,L},$$

$$(Z_m)^j_{ni,L+1}=\Big(r_1(\beta^j_{niL})\beta^j_{niL}+q^j_1\Big)\Big|_{S^{1z+}_{niL}},\ (Z_f)^j_{ni,L+1}=\Big(r_2(\beta^j_{niL})\beta^j_{niL}+q^j_2\Big)\Big|_{S^{2z+}_{niL}}.$$

Here and further the subscripts m and f indicate the belonging of the variable to the metal or to the form respectively.

Taking into account the introduced designations the three subproblems that approximate the direct problem can be written for all $j=\overline{0,J-1}$ in the following form:

x − direction

$$E^{j+\frac{1}{3}}_{nil} = E^j_{nil} + \omega^{j+1}_{nil}\Big[S^{1x+}_{nil}(X_m)^{j+\frac{1}{3}}_{n+1,il} - S^{1x-}_{nil}(X_m)^{j+\frac{1}{3}}_{nil} + S^{2x+}_{nil}(X_f)^{j+\frac{1}{3}}_{n+1,il} -$$
$$-S^{2x-}_{nil}(X_f)^{j+\frac{1}{3}}_{nil} + S^{1y+}_{nil}(Y_m)^j_{n,i+1,l} - S^{1y-}_{nil}(Y_m)^j_{nil} + S^{2y+}_{nil}(Y_f)^j_{n,i+1,l} - S^{2y-}_{nil}(Y_f)^j_{nil} +$$
$$+ S^{1z+}_{nil}(Z_m)^j_{ni,l+1} - S^{1z-}_{nil}(Z_m)^j_{nil} + S^{2z+}_{nil}(Z_f)^j_{ni,l+1} - S^{2z-}_{nil}(Z_f)^j_{nil}\Big],$$

y − direction

$$E^{j+\frac{2}{3}}_{nil} = E^{j+\frac{1}{3}}_{nil} + \omega^{j+1}_{nil}\Big[S^{1y+}_{nil}(Y_m)^{j+\frac{2}{3}}_{n,i+1,l} - S^{1y-}_{nil}(Y_m)^{j+\frac{2}{3}}_{nil} + S^{2y+}_{nil}(Y_f)^{j+\frac{2}{3}}_{n,i+1,l} -$$
$$-S^{2y-}_{nil}(Y_f)^{j+\frac{2}{3}}_{nil} + S^{1x+}_{nil}(X_m)^{j+\frac{1}{3}}_{n+1,il} - S^{1x-}_{nil}(X_m)^{j+\frac{1}{3}}_{nil} + S^{2x+}_{nil}(X_f)^{j+\frac{1}{3}}_{n+1,il} - S^{2x-}_{nil}(X_f)^{j+\frac{1}{3}}_{nil} +$$
$$+S^{1z+}_{nil}(Z_m)^{j+\frac{1}{3}}_{ni,l+1} - S^{1z-}_{nil}(Z_m)^{j+\frac{1}{3}}_{nil} + S^{2z+}_{nil}(Z_f)^{j+\frac{1}{3}}_{ni,l+1} - S^{2z-}_{nil}(Z_f)^{j+\frac{1}{3}}_{nil}\Big],$$

z − direction

$$E^{j+1}_{nil} = E^{j+\frac{2}{3}}_{nil} + \omega^{j+1}_{nil}\Big[S^{1z+}_{nil}(Z_m)^{j+1}_{ni,l+1} - S^{1z-}_{nil}(Z_m)^{j+1}_{nil} + S^{2z+}_{nil}(Z_f)^{j+1}_{ni,l+1} -$$
$$-S^{2z-}_{nil}(Z_f)^{j+1}_{nil} + S^{1x+}_{nil}(X_m)^{j+\frac{2}{3}}_{n+1,il} - S^{1x-}_{nil}(X_m)^{j+\frac{2}{3}}_{nil} + S^{2x+}_{nil}(X_f)^{j+\frac{2}{3}}_{n+1,il} - S^{2x-}_{nil}(X_f)^{j+\frac{2}{3}}_{nil} +$$
$$+ S^{1z+}_{nil}(Z_m)^{j+\frac{2}{3}}_{ni,l+1} - S^{1z-}_{nil}(Z_m)^{j+\frac{2}{3}}_{nil} + S^{2z+}_{nil}(Z_f)^{j+\frac{2}{3}}_{ni,l+1} - S^{2z-}_{nil}(Z_f)^{j+\frac{2}{3}}_{nil}\Big],$$

$$n=\overline{0,N}; \qquad i=\overline{0,I}; \qquad l=\overline{0,L}.$$

Let us introduce the following two-dimensional vectors:

$$S^{x+}_{nil} = \begin{bmatrix} S^{1x+}_{nil} \\ S^{2x+}_{nil} \end{bmatrix}, \qquad S^{x-}_{nil} = \begin{bmatrix} S^{1x-}_{nil} \\ S^{2x-}_{nil} \end{bmatrix}, \qquad S^{y+}_{nil} = \begin{bmatrix} S^{1y+}_{nil} \\ S^{2y+}_{nil} \end{bmatrix},$$

$$S^{y-}_{nil} = \begin{bmatrix} S^{1y-}_{nil} \\ S^{2y-}_{nil} \end{bmatrix}, \qquad S^{z+}_{nil} = \begin{bmatrix} S^{1z+}_{nil} \\ S^{2z+}_{nil} \end{bmatrix}, \qquad S^{z-}_{nil} = \begin{bmatrix} S^{1z-}_{nil} \\ S^{2z-}_{nil} \end{bmatrix},$$

$$n = \overline{0, N}; \qquad i = \overline{0, I}; \qquad l = \overline{0, L};$$

$$(X_{mf})^j_{nil} = \begin{bmatrix} (X_m)^j_{nil} \\ (X_f)^j_{nil} \end{bmatrix} \qquad n = \overline{0, N+1}; \qquad i = \overline{0, I}; \qquad l = \overline{0, L};$$

$$(Y_{mf})^j_{nil} = \begin{bmatrix} (Y_m)^j_{nil} \\ (Y_f)^j_{nil} \end{bmatrix} \qquad n = \overline{0, N}; \qquad i = \overline{0, I+1}; \qquad l = \overline{0, L};$$

$$(Z_{mf})^j_{nil} = \begin{bmatrix} (Z_m)^j_{nil} \\ (Z_f)^j_{nil} \end{bmatrix} \qquad n = \overline{0, N}; \qquad i = \overline{0, I}; \qquad l = \overline{0, L+1}.$$

Note that $S^{x+}_{nil} = S^{x-}_{n+1,il}$, $n = \overline{0, N-1}$;

$$S^{y+}_{nil} = S^{y-}_{n,i+1,l}, \ i = \overline{0, I-1}; \quad S^{z+}_{nil} = S^{z-}_{ni,l+1}, \ l = \overline{0, L-1}.$$

Let us introduce also designations for the following scalar products:

$$\widetilde{X}^j_{nil} = \left(S^{x-}_{nil}, (X_{mf})^j_{nil}\right), \qquad \widetilde{X}^j_{N+1,il} = \left(S^{x+}_{Nil}, (X_{mf})^j_{N+1,il}\right),$$

$$\widetilde{Y}^j_{nil} = \left(S^{y-}_{nil}, (Y_{mf})^j_{nil}\right), \qquad \widetilde{Y}^j_{n,I+1,l} = \left(S^{y+}_{nIl}, (Y_{mf})^j_{n,I+1,l}\right),$$

$$\widetilde{Z}^j_{nil} = \left(S^{z-}_{nil}, (Z_{mf})^j_{nil}\right), \qquad \widetilde{Z}^j_{ni,L+1} = \left(S^{z+}_{niL}, (Z_{mf})^j_{ni,L+1}\right),$$

$$n = \overline{0, N}; \qquad i = \overline{0, I}; \qquad l = \overline{0, L}.$$

Note that $\widetilde{X}^j_{nil}$ for all $n = \overline{1, N}$ is a function of two variables: E^j_{nil} and $E^j_{n-1,il}$; $\widetilde{X}^j_{0il}$ is a function of one variable E^j_{0il}, and $\widetilde{X}^j_{N+1,il}$ is also a function of one variable E^j_{Nil}. Similar statements are valid for $\widetilde{Y}^j_{nil}$ and $\widetilde{Z}^j_{nil}$.

With the aid of introduced designations the last three subproblems can be for $j = \overline{0, J-1}$ written in this compact form:

x − direction

$$E^{j+\frac{1}{3}}_{nil} = E^j_{nil} + \omega^{j+1}_{nil}\left(\widetilde{X}^{j+\frac{1}{3}}_{n+1,il} - \widetilde{X}^{j+\frac{1}{3}}_{nil} + \widetilde{Y}^j_{n,i+1,l} - \widetilde{Y}^j_{nil} + \widetilde{Z}^j_{ni,l+1} - \widetilde{Z}^j_{nil}\right), \tag{24}$$

y − direction

$$E^{j+\frac{2}{3}}_{nil} = E^{j+\frac{1}{3}}_{nil} + \omega^{j+1}_{nil}\left(\widetilde{Y}^{j+\frac{2}{3}}_{n,i+1,l} - \widetilde{Y}^{j+\frac{2}{3}}_{nil} + \widetilde{X}^{j+\frac{1}{3}}_{n+1,il} - \widetilde{X}^{j+\frac{1}{3}}_{nil} + \widetilde{Z}^{j+\frac{1}{3}}_{ni,l+1} - \widetilde{Z}^{j+\frac{1}{3}}_{nil}\right), \tag{25}$$

z − direction

$$E^{j+1}_{nil}=E^{j+\frac{2}{3}}_{nil}+\omega^{j+1}_{nil}\left(\widetilde{Z}^{j+1}_{ni,l+1}-\widetilde{Z}^{j+1}_{nil}+\widetilde{X}^{j+\frac{2}{3}}_{n+1,il}-\widetilde{X}^{j+\frac{2}{3}}_{nil}+\widetilde{Y}^{j+\frac{2}{3}}_{n,i+1,l}-\widetilde{Y}^{j+\frac{2}{3}}_{nil}\right), \qquad (26)$$

$$n=\overline{0,N} \qquad i=\overline{0,I} \qquad l=\overline{0,L}.$$

The cost functional $I(U)$ is approximated by the function $F(U)$ with the aid of the trapezoids method:

$$I(U)\cong F(U)=\frac{1}{2(t_2-t_1)}\left(\tau^{j_1+1}f^{j_1}+\sum_{j=j_1+1}^{j_2-1}(\tau^j+\tau^{j+1})f^j+\tau^{j_2}f^{j_2}\right).$$

Here j_1 is the ordinal number of the mesh point of the temporal grid which corresponds to the moment t_1, j_2 is the ordinal number of the mesh point of the temporal grid which corresponds to the moment t_2,

$$f^j=\sum_{n=n_1}^{n_2}\sum_{i=i_1}^{i_2}\left(Z^j_{ni}-z^j_*\right)^2 h^x_n h^y_i,$$

n_1, n_2 and i_1, i_2 are the ordinal numbers of the mesh points of the three-dimensional spacial grid along the Ox and Oy axes respectively which define the boundaries of the section S (i.e. $mesS=(x_{n_2}-x_{n_1})\times(y_{i_2}-y_{i_1})$), $Z^j_{ni}=Z_{pl}(x_n,y_i,t^j)$, $z^j_*=z_*(t^j)$.

Matrix elements Z^j_{ni} $\left(n=\overline{n_1,n_2},\quad i=\overline{i_1,i_2}\right)$ for each temporal layer j are defined by linear interpolation of the temperature field, obtained as a result of solving the direct problem. Let x_n,y_i,z_l be the coordinates of the mesh point of the spacial grid. For each mesh point $(x_n,y_i)\in S$, $(n=\overline{n_1,n_2},\ i=\overline{i_1,i_2})$ we find such index l_* for which one of the following conditions is valid: either $\beta\left(E^j_{ni,l_*+1}\right)\le T_{pl}\le\beta\left(E^j_{nil_*}\right)$, or $\beta\left(E^j_{ni,l_*}\right)\le T_{pl}\le\beta\left(E^j_{ni,l_*+1}\right)$.

Then

$$Z^j_{ni}=\frac{(z_{l_*+1}-z_{l_*})T_{pl}+(z_{l_*}\beta^j_{ni,l_*+1}-z_{l_*+1}\beta^j_{nil_*})}{\beta^j_{ni,l_*+1}-\beta^j_{nil_*}}.$$

Each equation of the selected discrete version of the direct problem (24)–(26) is presented in the canonical form (27) in accordance with the FAD-methodology:

$$E^j_{nil}=\Psi\left((n,i,l,j),\Lambda_{(n,i,l,j)},U_{(n,i,l,j)}\right). \qquad (27)$$

Here $\Lambda_{(n,i,l,j)}$ is the set of all $E^\nu_{\alpha\beta\gamma}$ with such indices α, β, γ and ν that corresponding elements occur in the right side of the equality (27); $U_{(n,i,l,j)}$ is the set of all components of the control vector U^ν $\quad(U^\nu=U(t^\nu))$ that occur in the right side of the equality (27). In spite of the fact that the control depends only on temporal index j the set $U_{(n,i,l,j)}$ is marked also by the spacial indices n,i, and l in order to emphasize the fact that the influence of this control is different at different spacial points.

To calculate the components of the gradient of the function $F(U)$ along the components of the vector U^j we will use the following relation, which is the generalization of a similar relation in [**?**]:

$$\frac{dF}{dU^j}=\frac{\partial F}{\partial U^j}+\sum_{(\alpha,\beta,\gamma,\nu)\in\overline{K}_{(n,i,l,j)}}\Psi^T_{u^j}\left((\alpha,\beta,\gamma,\nu),\Lambda_{(\alpha,\beta,\gamma,\nu)},U_{(\alpha,\beta,\gamma,\nu)}\right)p^{\nu}_{\alpha\beta\gamma},\ j=\overline{1,J}, \quad (28)$$

where $p^{\nu}_{\alpha\beta\gamma}$ are the conjugate variables (impulses), determined from solving the following system of linear algebraic equations

$$p^j_{nil}=\frac{dF}{dE^j_{nil}}+\sum_{(\alpha,\beta,\gamma,\nu)\in\overline{Q}_{(n,i,l,j)}}\Psi^T_{E^j_{nil}}\left((\alpha,\beta,\gamma,\nu),\Lambda_{(\alpha,\beta,\gamma,\nu)},U_{(\alpha,\beta,\gamma,\nu)}\right)p^{\nu}_{\alpha\beta\gamma}, \quad (29)$$

$$n=\overline{0,N},\qquad i=\overline{0,I},\qquad l=\overline{0,L},\qquad j=\overline{1,J}.$$

Index sets $\overline{Q}_{(n,i,l,j)}$ and $\overline{K}_{(n,i,l,j)}$ are determined by the following relations:

$$\overline{Q}_{(n,i,l,j)}=\left\{(\alpha,\beta,\gamma,\nu):E^j_{nil}\in\Lambda_{(\alpha,\beta,\gamma,\nu)}\right\},\overline{K}_{(n,i,l,j)}=\left\{(\alpha,\beta,\gamma,\nu):u^j\in U_{(\alpha,\beta,\gamma,\nu)}\right\}.$$

The system of linear algebraic equations (29) for determining the impulses p^j_{nil} is usually called the conjugate problem.

Let us introduce the following designations for a number of derivatives that will be used to write our conjugate problem in a compact form:

$$\forall i=\overline{0,I}\quad\text{and}\quad\forall l=\overline{0,L}$$

$$(D_{x+})^j_{nil}=\frac{\partial\widetilde{X}^j_{nil}}{\partial E^j_{nil}},\quad(D_{x-})^j_{nil}=\frac{\partial\widetilde{X}^j_{nil}}{\partial E^j_{n-1,il}},\quad(n=\overline{1,N}),\quad(D_{x+})^j_{0il}=\frac{\partial\widetilde{X}^j_{0il}}{\partial E^j_{0il}},$$

$$(D_{x-})^j_{0il}=0,\quad(D_{x+})^j_{N+1,il}=0,\quad(D_{x-})^j_{N+1,il}=\frac{\partial\widetilde{X}^j_{N+1,il}}{\partial E^j_{Nil}};$$

$$\forall n=\overline{0,N}\quad\text{and}\quad\forall l=\overline{0,L}$$

$$(D_{y+})^j_{nil}=\frac{\partial\widetilde{Y}^j_{nil}}{\partial E^j_{nil}},\quad(D_{y-})^j_{nil}=\frac{\partial\widetilde{Y}^j_{nil}}{\partial E^j_{n,i-1,l}},\quad(i=\overline{1,I}),\quad(D_{y+})^j_{n0l}=\frac{\partial\widetilde{Y}^j_{n0l}}{\partial E^j_{n0l}},$$

$$(D_{y-})^j_{n0l}=0,\quad(D_{y+})^j_{n,I+1,l}=0,\quad(D_{y-})^j_{n,I+1,l}=\frac{\partial\widetilde{Y}^j_{n,I+1,l}}{\partial E^j_{nIl}};$$

$$\forall n=\overline{0,N}\quad\text{and}\quad\forall i=\overline{0,I}$$

$$(D_{z+})^j_{nil}=\frac{\partial\widetilde{Z}^j_{nil}}{\partial E^j_{nil}},\quad(D_{z-})^j_{nil}=\frac{\partial\widetilde{Z}^j_{nil}}{\partial E^j_{ni,l-1}},\quad(l=\overline{1,L}),\quad(D_{z+})^j_{ni0}=\frac{\partial\widetilde{Z}^j_{ni0}}{\partial E^j_{ni0}},$$

$$(D_{z-})^j_{ni0}=0,\quad(D_{z+})^j_{ni,L+1}=0,\qquad(D_{z-})^j_{ni,L+1}=\frac{\partial\widetilde{Z}^j_{ni,L+1}}{\partial E^j_{niL}}.$$

For the differentiation to be valid, the functions $\beta(E^j_{nil})$, $\Omega_1(E^j_{nil})$ and $\Omega_2(E^j_{nil})$ were smoothed out in the neighborhood of their salient points.

Usage of the FAD-methodology leads us to the following systems of equations for determining the impulses.

6.2 The conjugate problem

6.2.1 Initial conditions for the impulses

In order to obtain the adjoint variables on the last temporal layer $j = J$, it is necessary for all $n = \overline{0,N}$ and $i = \overline{0,I}$ to solve the following system of $(L+1)$ linear algebraic equations for the variables p^j_{nil} $(l = \overline{0,L})$:

$$p^J_{ni0} = \omega^J_{ni0}\left((D_{z-})^J_{ni1} - (D_{z+})^J_{ni0}\right)p^J_{ni0} - \omega^J_{ni1}(D_{z-})^J_{ni1}p^J_{ni1} + \frac{\partial F}{\partial E^J_{ni0}},$$

$$p^J_{nil} = \omega^J_{ni,l-1}(D_{z+})^J_{nil}p^J_{ni,l-1} + \omega^J_{nil}\left((D_{z-})^J_{ni,l+1} - (D_{z+})^J_{nil}\right)p^J_{nil} -$$
$$-\omega^J_{ni,l+1}(D_{z-})^J_{ni,l+1}p^J_{ni,l+1} + \frac{\partial F}{\partial E^J_{nil}}, \quad (l = \overline{1, L-1}),$$

$$p^J_{niL} = \omega^J_{ni,L-1}(D_{z+})^J_{niL}p^J_{ni,L-1} + \omega^J_{niL}\left((D_{z-})^J_{ni,L+1} - (D_{z+})^J_{niL}\right)p^J_{niL} + \frac{\partial F}{\partial E^J_{niL}}.$$

It is possible to give to this system a more compact form if for all $n = \overline{0,N}$ and $i = \overline{0,I}$ to assume that

$$\omega^J_{ni,-1} = \omega^J_{ni,L+1} = 0 \quad \text{and} \quad p^J_{ni,-1} = p^J_{ni,L+1} = 0.$$

As a result we will obtain:

$$p^J_{nil} = \omega^J_{ni,l-1}(D_{z+})^J_{nil}p^J_{ni,l-1} + \omega^J_{nil}\left((D_{z-})^J_{ni,l+1} - (D_{z+})^J_{nil}\right)p^J_{nil} -$$
$$-\omega^J_{ni,l+1}(D_{z-})^J_{ni,l+1}p^J_{ni,l+1} + \frac{\partial F}{\partial E^J_{nil}}, \quad l = \overline{0,L}. \tag{30}$$

6.2.2 First subproblem for the impulses (y-direction)

In order to calculate the impulses $p^{j+\frac{2}{3}}_{nil}$ on the temporal sublayer $(j + 2/3)$ $(j = \overline{J-1,0})$ it is necessary to solve a linear algebraic system of $(I + 1)$ equations for all $n = \overline{0,N}$ and $l = \overline{0,L}$. This system can be written down more compactly if we make the following assumption:

$$\omega^{j+1}_{n,-1,l} = \omega^{j+1}_{n,I+1,l} = \omega^{j+1}_{-1,il} = \omega^{j+1}_{N+1,il} = 0,$$

$$p^{j+\frac{2}{3}}_{n,-1,l} = p^{j+\frac{2}{3}}_{n,I+1,l} = p^{j+1}_{n,-1,l} = p^{j+1}_{n,I+1,l} = p^{j+1}_{-1,il} = p^{j+1}_{N+1,il} = 0,$$
$$n = \overline{0,N}, \qquad i = \overline{0,I}, \qquad l = \overline{0,L}, \qquad j = \overline{J-1,0}.$$

As a result we will have:

$$p_{nil}^{j+\frac{2}{3}} = \omega_{n,i-1,l}^{j+1}(D_{y+})_{nil}^{j+\frac{2}{3}} p_{n,i-1,l}^{j+\frac{2}{3}} + \omega_{nil}^{j+1}\left((D_{y-})_{n,i+1,l}^{j+\frac{2}{3}} - (D_{y+})_{nil}^{j+\frac{2}{3}}\right) p_{nil}^{j+\frac{2}{3}} -$$

$$-\omega_{n,i+1,l}^{j+1}(D_{y-})_{n,i+1,l}^{j+\frac{2}{3}} p_{n,i+1,l}^{j+\frac{2}{3}} + \xi_{nil}^{j+\frac{2}{3}}, \qquad (31)$$

where

$$\xi_{nil}^{j+\frac{2}{3}} = p_{nil}^{j+1} + \omega_{n-1,il}^{j+1}(D_{x+})_{nil}^{j+\frac{2}{3}} p_{n-1,il}^{j+1} + \omega_{nil}^{j+1}\left((D_{x-})_{n+1,il}^{j+\frac{2}{3}} - (D_{x+})_{nil}^{j+\frac{2}{3}}\right) p_{nil}^{j+1} -$$

$$-\omega_{n+1,il}^{j+1}(D_{x-})_{n+1,il}^{j+\frac{2}{3}} p_{n+1,il}^{j+1} + \omega_{n,i-1,l}^{j+1}(D_{y+})_{nil}^{j+\frac{2}{3}} p_{n,i-1,l}^{j+1} + \omega_{nil}^{j+1}\Big((D_{y-})_{n,i+1,l}^{j+\frac{2}{3}} -$$

$$- (D_{y+})_{nil}^{j+\frac{2}{3}}\Big) p_{nil}^{j+1} - \omega_{n,i+1,l}^{j+1}(D_{y-})_{n,i+1,l}^{j+\frac{2}{3}} p_{n,i+1,l}^{j+1} + \frac{\partial F}{\partial E_{nil}^{j+\frac{2}{3}}}, \quad i = \overline{0,I}.$$

The formulation of other two subproblems for calculating the impulses will be provided only in the final compact form. If we assume that

$$\omega_{n,-1,l}^{j} = \omega_{n,I+1,l}^{j} = \omega_{-1,il}^{j} = \omega_{N+1,il}^{j} = \omega_{ni,-1}^{j} = \omega_{ni,L+1}^{j} = 0,$$

$$p_{-1,il}^{j+\frac{1}{3}} = p_{N+1,il}^{j+\frac{1}{3}} = p_{-1,il}^{j+\frac{2}{3}} = p_{N+1,il}^{j+\frac{2}{3}} = p_{ni,-1}^{j+\frac{2}{3}} = p_{ni,L+1}^{j+\frac{2}{3}} = 0,$$

$$p_{ni,-1}^{j} = p_{ni,L+1}^{j} = p_{ni,-1}^{j+\frac{1}{3}} = p_{ni,L+1}^{j+\frac{1}{3}} = p_{n,-1,l}^{j+\frac{1}{3}} = p_{n,I+1,l}^{j+\frac{1}{3}} = 0,$$

$$n = \overline{0,N}, \qquad i = \overline{0,I}, \qquad l = \overline{0,L}, \qquad j = \overline{0,J},$$

it is similar to how this was done for the first subproblem.

6.2.3 Second subproblem for the impulses (x-direction)

In order to calculate the adjoint variables $p_{nil}^{j+\frac{1}{3}}$ on the temporal sublayer $j+1/3 \quad (j = \overline{J-1,0})$ it is necessary to solve the following linear algebraic system of $(N+1)$ equations for all $i = \overline{0,I}$ and $l = \overline{0,L}$:

$$p_{nil}^{j+\frac{1}{3}} = \omega_{n-1,il}^{j+1}(D_{x+})_{nil}^{j+\frac{1}{3}} p_{n-1,il}^{j+\frac{1}{3}} + \omega_{nil}^{j+1}\left((D_{x-})_{n+1,il}^{j+\frac{1}{3}} - (D_{x+})_{nil}^{j+\frac{1}{3}}\right) p_{nil}^{j+\frac{1}{3}} -$$

$$-\omega_{n+1,il}^{j+1}(D_{x-})_{n+1,il}^{j+\frac{1}{3}} p_{n+1,il}^{j+\frac{1}{3}} + \xi_{nil}^{j+\frac{1}{3}}, \qquad (32)$$

where

$$\xi_{nil}^{j+\frac{1}{3}} = p_{nil}^{j+\frac{2}{3}} + \omega_{n-1,il}^{j+1}(D_{x+})_{nil}^{j+\frac{1}{3}} p_{n-1,il}^{j+\frac{2}{3}} + \omega_{nil}^{j+1}\left((D_{x-})_{n+1,il}^{j+\frac{1}{3}} - (D_{x+})_{nil}^{j+\frac{1}{3}}\right) p_{nil}^{j+\frac{2}{3}} -$$

$$-\omega_{n+1,il}^{j+1}(D_{x-})_{n+1,il}^{j+\frac{1}{3}} p_{n+1,il}^{j+\frac{2}{3}} + \omega_{ni,l-1}^{j+1}(D_{z+})_{nil}^{j+\frac{1}{3}} p_{ni,l-1}^{j+\frac{2}{3}} + \omega_{nil}^{j+1}\Big((D_{z-})_{ni,l+1}^{j+\frac{1}{3}} -$$

$$- (D_{z+})_{nil}^{j+\frac{1}{3}}\Big) p_{nil}^{j+\frac{2}{3}} - \omega_{ni,l+1}^{j+1}(D_{z-})_{ni,l+1}^{j+\frac{1}{3}} p_{ni,l+1}^{j+\frac{2}{3}} + \frac{\partial F}{\partial E_{nil}^{j+\frac{1}{3}}}, \quad n = \overline{0,N}.$$

6.2.4 Third subproblem for the impulses (z-direction)

In order to calculate the adjoint variables p^j_{nil} on temporal layer j ($j = \overline{J-1,1}$) it is necessary to solve the following linear algebraic system of $(L+1)$ equations for all $n = \overline{0,N}$ and $i = \overline{0,I}$:

$$p^j_{nil} = \omega^j_{ni,l-1}(D_{z+})^j_{nil}p^j_{ni,l-1} + \omega^j_{nil}\left((D_{z-})^j_{ni,l+1} - (D_{z+})^j_{nil}\right)p^j_{nil} -$$

$$-\omega^j_{ni,l+1}(D_{z-})^j_{ni,l+1}p^j_{ni,l+1} + \xi^j_{nil}, \tag{33}$$

where

$$\xi^j_{nil} = p^{j+\frac{1}{3}}_{nil} + \omega^{j+1}_{n,i-1,l}(D_{y+})^j_{nil}p^{j+\frac{1}{3}}_{n,i-1,l} + \omega^{j+1}_{nil}\left((D_{y-})^j_{n,i+1,l} - (D_{y+})^j_{nil}\right)p^{j+\frac{1}{3}}_{nil} -$$

$$-\omega^{j+1}_{n,i+1,l}(D_{y-})^j_{n,i+1,l}p^{j+\frac{1}{3}}_{n,i+1,l} + \omega^{j+1}_{ni,l-1}(D_{z+})^j_{nil}p^{j+\frac{1}{3}}_{ni,l-1} + \omega^{j+1}_{nil}\left((D_{z-})^j_{ni,l+1} -\right.$$

$$\left. - (D_{z+})^j_{nil}\right)p^{j+\frac{1}{3}}_{nil} - \omega^{j+1}_{ni,l+1}(D_{z-})^j_{ni,l+1}p^{j+\frac{1}{3}}_{ni,l+1} + \frac{\partial F}{\partial E^j_{nil}}, \quad l = \overline{0,L}.$$

Systems (30)–(33) approximate the initial-boundary value problem for the reverse thermal conductivity equation.

Each of systems (30)–(33) is solved with the aid of tridiagonal Gaussian elimination. Solving these three subproblems successively for all $j = \overline{J,0}$ allows us to obtain the values of the adjoint variables in the following order: p^J_{nil}, $p^{(J-1)+2/3}_{nil}$, $p^{(J-1)+1/3}_{nil}$, $p^{(J-1)}_{nil}$,...,$p^{1+1/3}_{nil}$, p^1_{nil}, $p^{0+2/3}_{nil}$, $p^{0+1/3}_{nil}$, ($n = \overline{0,N}, i = \overline{0,I}, l = \overline{0,L}$).

In the first two subproblems (i.e. in the systems of equations (31)–(32)) all derivatives $\frac{\partial F}{\partial E^{j+2/3}_{nil}}$ and $\frac{\partial F}{\partial E^{j+1/3}_{nil}}$ ($j = \overline{J-1,0}$, $n = \overline{0,N}$, $i = \overline{0,I}$, $l = \overline{0,L}$) are equal to zero. In the last subproblem (33) only derivatives $\frac{\partial F}{\partial E^j_{nil_*}}$ and $\frac{\partial F}{\partial E^j_{ni,l_*+1}}$ are not equal to zero. They are calculated using the following formulas:

$$\frac{\partial F}{\partial E^j_{nil_*}} = \frac{\mu^j}{t_2 - t_1}\left(Z^j_{ni} - z^j_*\right)\frac{\partial \beta(E^j_{nil_*})}{\partial E^j_{nil_*}} \cdot \frac{(z_{l_*+1} - z_{l_*})(T_{pl} - \beta(E^j_{ni,l_*+1}))}{\left(\beta(E^j_{ni,l_*+1}) - \beta(E^j_{nil_*})\right)^2}h^x_n h^y_i,$$

$$\frac{\partial F}{\partial E^j_{ni,l_*+1}} = \frac{\mu^j}{t_2 - t_1}\left(Z^j_{ni} - z^j_*\right)\frac{\partial \beta(E^j_{ni,l_*+1})}{\partial E^j_{ni,l_*+1}} \cdot \frac{(z_{l_*} - z_{l_*+1})(T_{pl} - \beta(E^j_{nil_*}))}{\left(\beta(E^j_{ni,l_*+1}) - \beta(E^j_{nil_*})\right)^2}h^x_n h^y_i,$$

where $\mu^{j_1} = \tau^{j_1+1}$, $\mu^j = \tau^j + \tau^{j+1}$($j = \overline{j_1+1, j_2-1}$), $\mu^{j_2} = \tau^{j_2}$.

6.3 Gradient of the objective function of the discrete optimal control problem

Let us examine the first case, when the control function $U(t)$ is selected as the dependence on time of the displacement of the foundry mold in the melting furnace, namely, the z-coordinate of the lower bound of the wall of the furnace $Z_{Sou}(t)$. This parameter enters into the expressions that determine the functions $q_1(t)$ and $q_2(t)$ when the considered cell is located outside of the liquid aluminum. The control function $U(t)$ is approximated by a piecewise constant function that has constant values in each time interval $[t^j, t^{j+1}]$. Namely, we assume that on this time interval control equals to $U(t) = Z_{Sou}(t^{j+1}) = Z_{Sou}^{j+1}$. Consequently, $q_1^{j+1/3} = q_1^{j+2/3} = q_1^{j+1}$ and $q_2^{j+1/3} = q_2^{j+2/3} = q_2^{j+1}$.

According to the FAD-methodology, the components of the gradient of the objective function are calculated from the following formula:

$$
\begin{aligned}
\frac{dF}{dU^j} = \frac{\partial F}{\partial U^j} &+ \sum_{n=0}^{N}\sum_{i=0}^{I}\left(\omega_{niL}^j \frac{\partial \widetilde{Z}_{ni,L+1}^j}{\partial U^j} p_{niL}^j - \omega_{ni0}^j \frac{\partial \widetilde{Z}_{ni0}^j}{\partial U^j} p_{ni0}^j\right) + \\
&+ \sum_{n=0}^{N}\sum_{l=0}^{L}\left(\omega_{nIl}^j \frac{\partial \widetilde{Y}_{n,I+1,l}^{j-\frac{1}{3}}}{\partial U^j} p_{nIl}^j - \omega_{n0l}^j \frac{\partial \widetilde{Y}_{n0l}^{j-\frac{1}{3}}}{\partial U^j} p_{n0l}^j\right) + \\
&+ \sum_{i=0}^{I}\sum_{l=0}^{L}\left(\omega_{Nil}^j \frac{\partial \widetilde{X}_{N+1,il}^{j-\frac{1}{3}}}{\partial U^j} p_{Nil}^j - \omega_{0il}^j \frac{\partial \widetilde{X}_{0il}^{j-\frac{1}{3}}}{\partial U^j} p_{0il}^j\right) + \\
&+ \sum_{n=0}^{N}\sum_{l=0}^{L}\left(\omega_{nIl}^j \frac{\partial \widetilde{Y}_{n,I+1,l}^{j-\frac{1}{3}}}{\partial U^j} p_{nIl}^{j-\frac{1}{3}} - \omega_{n0l}^j \frac{\partial \widetilde{Y}_{n0l}^{j-\frac{1}{3}}}{\partial U^j} p_{n0l}^{j-\frac{1}{3}}\right) + \\
&+ \sum_{i=0}^{I}\sum_{l=0}^{L}\left(\omega_{Nil}^j \frac{\partial \widetilde{X}_{N+1,il}^{j-\frac{2}{3}}}{\partial U^j} p_{Nil}^{j-\frac{1}{3}} - \omega_{0il}^j \frac{\partial \widetilde{X}_{0il}^{j-\frac{2}{3}}}{\partial U^j} p_{0il}^{j-\frac{1}{3}}\right) + \qquad (34) \\
&+ \sum_{n=0}^{N}\sum_{i=0}^{I}\left(\omega_{niL}^j \frac{\partial \widetilde{Z}_{ni,L+1}^{j-\frac{2}{3}}}{\partial U^j} p_{niL}^{j-\frac{1}{3}} - \omega_{ni0}^j \frac{\partial \widetilde{Z}_{ni0}^{j-\frac{2}{3}}}{\partial U^j} p_{ni0}^{j-\frac{1}{3}}\right) + \\
&+ \sum_{i=0}^{I}\sum_{l=0}^{L}\left(\omega_{Nil}^j \frac{\partial \widetilde{X}_{N+1,il}^{j-\frac{2}{3}}}{\partial U^j} p_{Nil}^{j-\frac{2}{3}} - \omega_{0il}^j \frac{\partial \widetilde{X}_{0il}^{j-\frac{2}{3}}}{\partial U^j} p_{0il}^{j-\frac{2}{3}}\right) + \\
&+ \sum_{n=0}^{N}\sum_{l=0}^{L}\left(\omega_{nIl}^j \frac{\partial \widetilde{Y}_{n,I+1,l}^{j-1}}{\partial U^j} p_{nIl}^{j-\frac{2}{3}} - \omega_{n0l}^j \frac{\partial \widetilde{Y}_{n0l}^{j-1}}{\partial U^j} p_{n0l}^{j-\frac{2}{3}}\right) + \\
&+ \sum_{n=0}^{N}\sum_{i=0}^{I}\left(\omega_{niL}^j \frac{\partial \widetilde{Z}_{ni,L+1}^{j-1}}{\partial U^j} p_{niL}^{j-\frac{2}{3}} - \omega_{ni0}^j \frac{\partial \widetilde{Z}_{ni0}^{j-1}}{\partial U^j} p_{ni0}^{j-\frac{2}{3}}\right), \quad j = \overline{1, J}.
\end{aligned}
$$

Since the functional $F(U)$ does not depend explicitly on the control vector $\{U^j\}$, all components $\frac{\partial F}{\partial U^j} = 0$.

Let us give an example of calculation of one of the derivatives that occur in formula (34):

$$\frac{\partial \widetilde{X}^{j-\frac{2}{3}}_{N+1,il}}{\partial U^j} = S^{2x+}_{Nil} \frac{\partial \left((X_f)^{j-\frac{2}{3}}_{N+1,il} \right)}{\partial U^j} = S^{2x+}_{Nil} \frac{\partial \left(\left(q_2^j \right) \Big|_{S^{2x+}_{Nil}} \right)}{\partial U^j}.$$

It's taken into account here that cells with the indices (N, i, l), $(i = \overline{0, I},\ l = \overline{0, L})$ don't contain metal. Therefore $(X_m)^{j-\frac{2}{3}}_{N+1,il} = 0$. If at the moment $t = t^j$ cell with number (N, i, l) is located in the liquid aluminum, then $\frac{\partial \left((q_2^j)|_{S^{2x+}_{Nil}} \right)}{\partial U^j} = \frac{\partial \left((q_2^j)|_{S^{2x+}_{Nil}} \right)}{\partial Z^j_{Sou}} = 0$ and, therefore $\frac{\partial \widetilde{X}^{(j-1)+1/3}_{N+1,il}}{\partial U^j} = 0$. But if at the moment $t = t^j$ this cell is located outside of the liquid aluminum, then (according to (20) and (21))

$$\frac{\partial \left(\left(q_2^j \right) \Big|_{S^{2x+}_{Nil}} \right)}{\partial Z^j_{Sou}} = \frac{\partial \left(\varphi_s + \varphi_a \right)}{\partial Z^j_{Sou}} =$$

$$= \frac{\partial (q_s(X_s, Y_{Sou} - y_i + L_{Sou}, Z_{Sou} - z_l + H_{Sou}) - q_s(X_s, Y_{Sou} - y_i, Z_{Sou} - z_l + H_{Sou}))}{\partial Z^j_{Sou}} +$$

$$+ \frac{\partial \left(q_s \left(X_s, Y_{Sou} - y_i, Z_{Sou} - z_l \right) - q_s \left(X_s, Y_{Sou} - y_i + L_{Sou}, Z_{Sou} - z_l \right) \right)}{\partial Z^j_{Sou}} +$$

$$+ \frac{\partial (q_a(Z_a, Y_{al} - y_i + L_{al}, X_{al} - X_b + H_{al}) - q_a(Z_a, Y_{al} - y_i, X_{al} - X_b + H_{al}))}{\partial Z^j_{Sou}}.$$

The third argument of the function q_s and the first argument of the function q_a depend on the value Z^j_{Sou}. According to formulas (22) and (23) (see part I) we have:

$$\widetilde{q}_s(\xi, l, h) \equiv \frac{\partial q_s(\xi, l, h)}{\partial h} = M_S \left[\frac{\xi^2}{\eta^3} \arctan \left(\frac{l}{h} \right) + \frac{l \xi^2}{\eta^2 (\eta^2 + l^2)} \right],$$

$$\widetilde{q}_a(\xi, l, h) \equiv \frac{\partial q_a(\xi, l, h)}{\partial \xi} = -M_a \cdot \frac{l}{\xi^2 + l^2} - \frac{M_a}{\eta^2} \left[\left(\eta - \frac{\xi^2}{\eta} \right) \arctan \left(\frac{l}{\eta} \right) - \frac{l \xi^2}{\eta^2 + l^2} \right],$$

where $\eta = \sqrt{\xi^2 + h^2}$. Thus,

$$\frac{\partial \left(\left(q_2^j \right) \Big|_{S^{2x+}_{Nil}} \right)}{\partial Z^j_{Sou}} = \widetilde{q}_s \left(X_s, Y_{Sou} - y_i + L_{Sou}, Z_{Sou} - z_l + H_{Sou} \right) -$$

$$- \widetilde{q}_s \left(X_s, Y_{Sou} - y_i, Z_{Sou} - z_l + H_{Sou} \right) +$$

$$+ \widetilde{q}_s \left(X_s, Y_{Sou} - y_i, Z_{Sou} - z_l \right) - \widetilde{q}_s \left(X_s, Y_{Sou} - y_i + L_{Sou}, Z_{Sou} - z_l \right) +$$

$$+\left[\widetilde{q}_a(Z_a, Y_{al}-y_i+L_{al}, X_{al}-X_b+H_{al})-\widetilde{q}_a(Z_a, Y_{al}-y_i, X_{al}-X_b+H_{al})\right]\cdot\frac{\partial Z_a}{\partial Z^j_{Sou}},$$

$$\frac{\partial Z_a}{\partial Z^j_{Sou}}=\begin{cases}-1, & \text{object did not reach the surface of aluminum,}\\ -1-\dfrac{X_b\cdot Y_b}{L_{al}\cdot H_{al}-X_b\cdot Y_b}, & \text{object reached the surface of aluminum.}\end{cases}$$

There is a special practical interest in the dependence of the solidification front on the speed $\widetilde{u}(t)$ of the displacement of the object. In this case the speed of the displacement of the foundry mold in the melting furnace is selected as the control function. Z-coordinate of the lower bound of the wall of the furnace $Z_{Sou}(t)$ is determined with the aid of the speed $\widetilde{u}(t)$ as follows:

$$Z_{Sou}(t^j)=Z_{Sou}(t^{j-1})-\tau^j\widetilde{u}(t^j),\quad \text{or}\quad Z_{Sou}(t^j)=\widetilde{z}-\sum_{k=1}^{j}\tau^k\widetilde{u}(t^k),$$

where $\widetilde{z}$ is the z-coordinate of the lower bound of the wall of the furnace at the initial time. In this case the component of the gradient of the function $F(\widetilde{u})$ along the components of vector $\{\widetilde{u}^j\}$, $\big(\widetilde{u}^j=\widetilde{u}(t^j)\big)$, are calculated using the following formula:

$$\frac{dF}{d\widetilde{u}^j}=\frac{\partial F}{\partial\widetilde{u}^j}-\tau^j\sum_{k=j}^{J}\left(\frac{dF}{dU^k}-\frac{\partial F}{\partial U^k}\right), j=\overline{1,J}, \tag{35}$$

where $\frac{dF}{dU^k}(k=\overline{1,J})$ are calculated using the formula (34). Due to the specific character of the functional in the considered problem, $\frac{\partial F}{\partial\widetilde{u}^j}=\frac{\partial F}{\partial U^j}=0$, $(j=\overline{1,J})$.

Let us give the formula for calculating the gradient of the functional in the case, when the speed function $\widetilde{u}(t)$ in the temporal section $[0,t^J]$ was approximated by piecewise constant function with an arbitrary number of segments.

The time interval $[0,t^J]$ is divided in Θ "large" subintervals. The function $\widetilde{u}(t)$ has a constant value on each subinterval. Each of these subintervals contains β elementary intervals $[t^{j-1},t^j]$. Thus, $\widetilde{u}^{(s-1)\beta+\alpha}=\widetilde{v}^s,(\alpha=\overline{1,\beta})$, where $\widetilde{v}^s$ $(s=\overline{1,\Theta})$ is given. Then the component of the gradient of the objective function $F(U)$ along the components of vector $\{\widetilde{v}^s\},(s=\overline{1,\Theta})$, are calculated using the following formula:

$$\frac{dF}{d\widetilde{v}^s}=\sum_{\alpha=1}^{\beta}\frac{dF}{d\widetilde{u}^{(s-1)\beta+\alpha}},\quad s=\overline{1,\Theta}, \tag{36}$$

where derivatives $\frac{dF}{d\widetilde{u}^j}$ are determined with the aid of relation (35).

Let us point out also that the systems of equations (30)–(33) don't depend on the choice of the control function.

Let us especially note that the value of the gradient of the objective function, calculated according to formulas (34)–(36), is precise for the selected approximation of the optimal control problem.

The calculation of the approximate value of the gradient of the objective function with the aid of the finite-difference method in this optimal control problem is connected with enormous difficulties [3].

The machine time needed for calculation of the gradient components using the approach presented here (based on the FAD-methodology) is not more than half of machine time needed for solving the direct problem.

Therefore, in spite of the difficulties connected with obtaining the discrete version of the conjugate problem and the gradient, it seems unavoidable finding the precise value of the gradient of the objective function using the FAD-methodology while solving complex problems of optimal control.

7 Numerical results of solving the optimal control problem

The speed $\widetilde{u}(t)$ of the displacement of foundry mold in the melting furnace was chosen as the control $U(t)$. The formulated optimal control problem was solved numerically using the gradient method. During the solution of the optimal control problem the time interval $[0, t^J]$ was divided into N parts (subintervals). The control function $U(t)$ was approximated by piecewise constant function, so that for each of subintervals it was constant. The components of the gradient of the objective function are calculated using the formula (36).

The optimal control problem was studied for a rectangular parallelepiped. The previous parameters of the problem, indicated in the fifth section (part I), were used, with the exception of some given below:

$$T_{Sou} = 1900.15, \qquad T_{al} = 1033.15, \qquad L_{Sou} = 0.350, \qquad H_{Sou} = 0.380,$$

$$X_b = 0.040, \quad Y_b = 0.060, \quad Z_b = 0.180.$$

The parallelepiped was immersed into the liquid aluminum to 5/6 of its height. The number t^J, which determines the length of the time interval $[0, t^J]$, was equal to 3299 s. Z-coordinate $z_*(t)$ of the desired solidification front changed with a constant velocity $U_*(t) = 0.1mm/s$. Calculations were performed for different numbers N of subintervals, on which the control function $U(t)$ was constant.

In Fig. 16a the dependence of the optimal cost functional $J(U)$ upon the number N of subintervals is represented. It is obtained as the result of the solution of optimization problem. Here N has the following values: 1, 2, 4, 12, 24, 600. As shown in Fig. 16a, the optimal value of the functional decreases noticeably for the small values of N, and for the great values of $N (N > 30)$ it weakly diminishes and comes out to a certain constant asymptotical value. Fig. 16b is a fragment of Fig. 16a in which there is no point corresponding to the value $N = 600$. This makes it possible to examine more precisely the dependence of the optimal value of the cost functional upon the number of subintervals for low values of N.

In Figures 17 the optimal trajectories of the foundry mold are shown. These are those trajectories with which optimum values of functional examined above are obtained (see Fig. 16), namely, for $N = 1, 2, 4, 12, 24, 600$. Numbers near the curves indicate the number N of subintervals used. The convergence of the optimal trajectories to a certain limit function when the number N increase is visible in Figures 17. Let us note that the qualitatively correct structure of optimal trajectory is already obtained for $N = 12$. Further increase of the number N only smoothes the optimal trajectory.

Figure 18 shows the behavior of the standard deviation of the real solidification front from the desired one for several control functions. Standard deviation is determined by the formula

$$D(t) = \sqrt{\frac{1}{|S|} \iint\limits_{S} \left[Z_{pl}(x,y,t) - z_*(t)\right]^2 dxdy}, \tag{37}$$

where $|S|$ is the area of the cross section S. Curve 1 in Fig. 18 corresponds to the regime when the foundry mold is moved with a small constant velocity $\widetilde{u}(t) = 0.083 mm/s$ relative to the furnace. Curve 2, just as curve 1, corresponds to the regime with a constant velocity of the displacement of the foundry mold, but $\widetilde{u}(t) = 0.150 mm/s$. Curve 3 corresponds to such displacement of foundry mold when the functional (3) reaches the minimum value. All these calculations were performed for $N = 24$.

The advantages of the optimal process of metal crystallization are vividly shown by the figures given below. Figures 19-21 illustrate isotherms for different times in two cross sections through the object's vertical axis of symmetry parallel to the parallelepiped faces. Since the object is symmetric about the vertical axis, the figures present only halves of the cross sections. Figures 19a, 20a, 21a (first experiment) illustrate the process of metal solidification in a mold moving relative to the furnace with the constant speed $\widetilde{u}(t) = 0.417 mm/min$. Figures 19b, 20b, 21b (second experiment) correspond to a mold moving with the optimal speed, corresponding $N = 4$.

Figures 19–21 show that the isotherms are concentrated within the mold. Moreover, the results of the second experiment are superior to those of the first one. First, the phase boundary in the second experiment is closer to a horizontal plane. Second, bubbles of liquid metal form and collapse inside the casting in the first experiment (Fig. 21a), which results in a casting of poor quality, whereas no bubbles are observed in the second experiment. Third, the process of solidification in the first experiment proceeds too quickly (for about 962 s.), which also degrades the casting. In the second experiment, the solidification process lasts roughly twice as long as in the first (1930 s.).

Acknowledgments. This work was supported by the Russian Foundation for Basic Research (project no. 08-01-90100-Mol_a) and the program "Leading Scientific Schools" (project no. NSh-5073.2008.1).

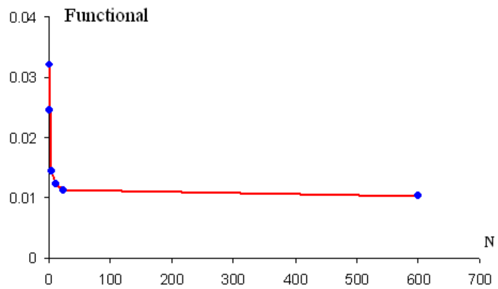


Fig. 16a

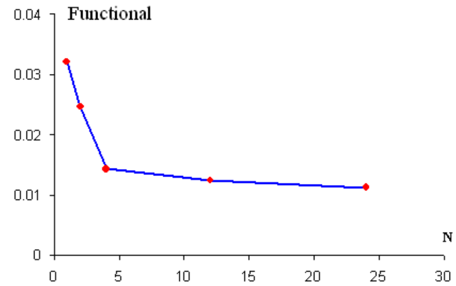


Fig. 16b

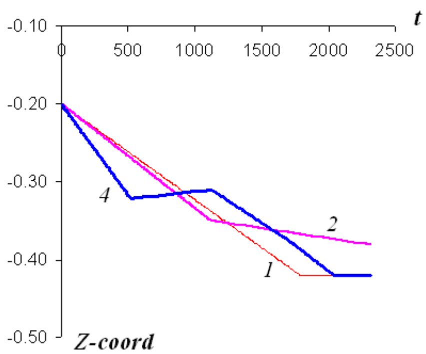


Fig. 17a

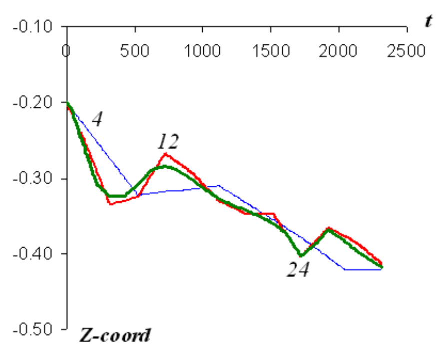


Fig. 17b

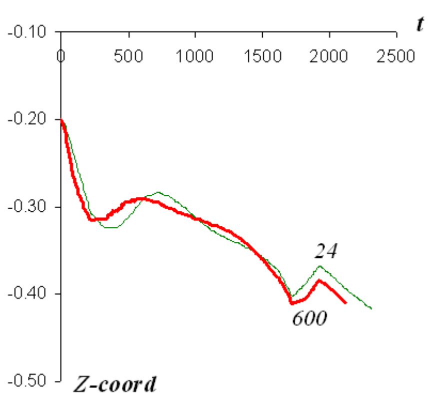


Fig. 17c

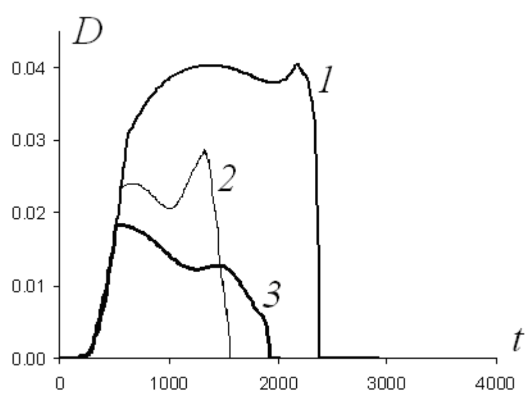


Fig. 18

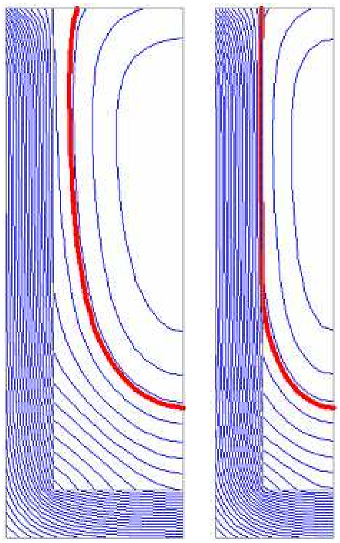


Fig. 19a

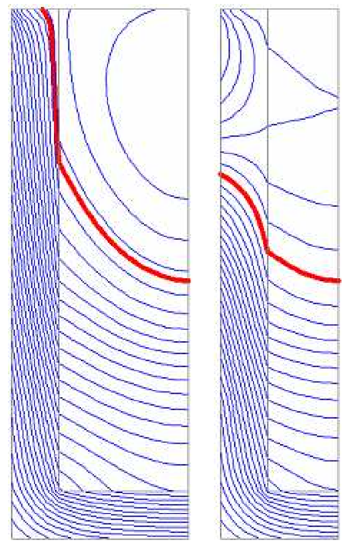


Fig. 19b

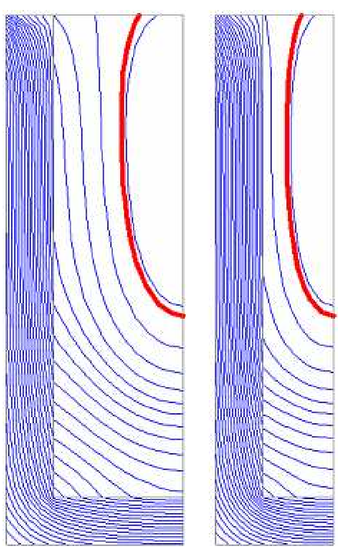


Fig. 20a

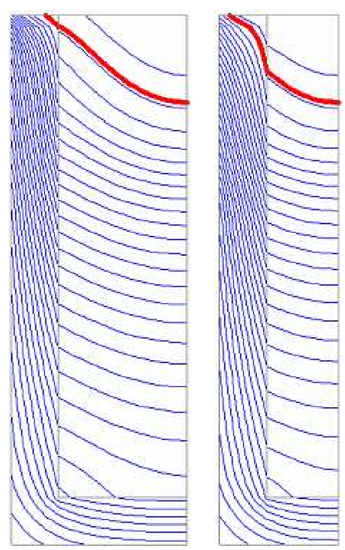


Fig. 20b

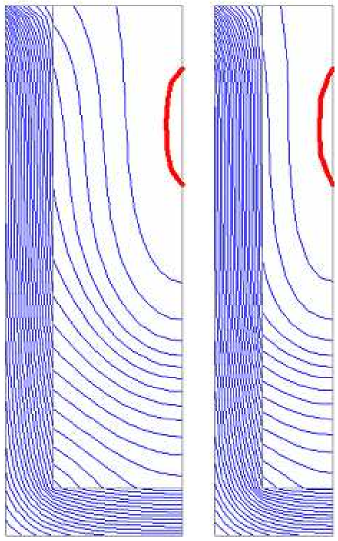


Fig. 21a

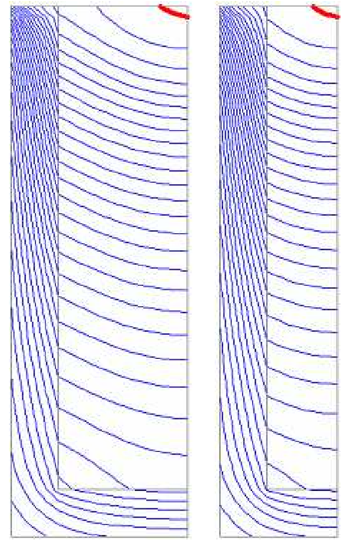


Fig. 21b

References

[1] Albu A., Zubov V. *Optimal control for one complex dynamic system, I.* Bul. Acad. de Ştiinţe a Republicii Moldova, Matematica, 2009, No. 1(59), 3–21.

[2] Evtushenko Y.G. *Computation of Exact Gradients in Distributed Dynamic Systems*. Optimizat. Methods and Software, 1998, **9**, 45–75.

[3] Albu A.F., Zubov V.I. *Calculation of the gradient of functional in one optimal control problem*. Zh. Vychisl. Mat. Mat. Fiz., 2009, **49**, No. 1 (Comp. Math. Math. Phys., 2009, **49**).

V. I. Zubov
Computing Centre of Russian Academy of Sciences
40, Vavilova Str., 119333 Moscow
Russia
E-mail: *zubov@ccas.ru*

Received October 9, 2008

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 19–28
ISSN 1024–7696

About group topologies of the primary Abelian group of finite period which coincide on a subgroup and on the factor group *

V. I. Arnautov

Abstract. Let G be any Abelian group of the period p^n and $G_1 = \{g \in G | pg = 0\}$, $G_2 = \{g \in G | p^{n-1}g = 0\}$. If τ and τ' are a metrizable, linear group topologies such that G_2 is a closed subgroup in each of topological groups (G, τ) and (G, τ'), then $\tau|_{G_2} = \tau'|_{G_2}$ and $(G, \tau)/G_1 = (G, \tau')/G_1$ if and only if there exists a group isomorphism $\varphi : G \to G$ such that the following conditions are true:
1. $\varphi(G_2) = G_2$;
2. $g - \varphi(g) \in G_1$ for any $g \in G$;
3. $\varphi : (G, \tau) \to (G, \tau')$ is a topological isomorphism.

Mathematics subject classification: 22A05.
Keywords and phrases: Topological group, natural homomorphism, topological isomorphism, subgroup of topological group, factor group of topological group, basis of neighborhoods of zero..

Wyen studying properties of lattices of all group topologies[1] on Abelians groups or their sublattices there is a need to establish the interconnections between group topologies which coincide on some subgroups and on some factor groups.

A partial answer to this question is given in the present article.

The main result of this article is Theorem 9.

1. Notations. During all this work, if it is not stipulated opposite, we shall adhere to the following notations;

1.1. p is some fixed prime number;

1.2. n is some fixed natural number;

1.3. $\mathbb{N}$ is the set of all natural numbers;

1.4. G is an Abelian group of the period p^n;

1.5. G' is a subgroup of the group G;

1.6. $\omega : G \to G/G'$ is the natural homomorphism (i.e. $\omega(g) = g + G'$ for any $g \in G$);

1.7. If $A \subseteq G$ then we denote by $< A >$ the subgroup in G, generated by the subset A. In particular we denote by $< g >$ the subgroup in G generated by the element g;

1.8. If $\{A_\gamma | \gamma \in \Gamma\}$ is some set of groups, then we denote by $\bigoplus\limits_{\gamma\in\Gamma} A_\gamma$ the direct sum of these groups;

*The author was partially supported by the grant 08.820.08.12 RF.

[1]The considered topologies are not necessarily Hausdorff

1.9. If τ is a group topology on G, then we denote by $\tau|_{G'}$ the induced topology on G', i.e. $\tau|_{G'} = \{U \bigcap G' | U \in \tau\}$;

1.10. If (G, τ) is a topological group, then we denote by $(G, \tau)/G'$ the topological group $(G/G', \bar{\tau})$, where $\bar{\tau} = \{\omega(U) | U \in \tau\}$.

2. Proposition. *If τ and τ' are group topologies on G then the following statements are true:*

2.1. *If $\tau|_{G'} = \tau'|_{G'}$ then topological groups (G, τ) and (G, τ') possess such bases $\{W_\gamma | \gamma \in \Gamma\}$ and $\{W'_\gamma | \gamma \in \Gamma\}$ of the neighborhoods of zero respectively, that $W_\gamma \bigcap G' = W_\gamma \bigcap G'$ for any $\gamma \in \Gamma$. Moreover if topologies τ and τ' are linear, then both W_γ and W'_γ are subgroups of the group G;*

2.2. *If $(G, \tau)/G' = (G, \tau')/G'$, then topological groups (G, τ) and (G, τ') possess such bases $\{W_\gamma | \gamma \in \Gamma\}$ and $\{W'_\gamma | \gamma \in \Gamma\}$ of the neighborhoods of zero, respectively, that $\omega(W_\gamma) = \omega(W'_\gamma)$ for any $\gamma \in \Gamma$. Moreover if topologies τ and τ' are linear, then both W_γ and W'_γ are subgroups of the group G;*

2.3. *Let G_1 and G_2 be such subgroups of group G that $G_1 \subseteq G_2$ or $G_2 \subseteq G_1$ and $\tau|_{G_1} = \tau'|_{G_1}$. If $(G, \tau)/G_2 = (G, \tau')/G_2$, then topological groups (G, τ) and (G, τ') possess such bases $\{U_\gamma | \gamma \in \Gamma\}$ and $\{U'_\gamma | \gamma \in \Gamma\}$ of the neighborhoods of zero, respectively, that $U_\gamma \bigcap G_1 = U'_\gamma \bigcap G_1$ and $G_2 + U_\gamma = G_2 + U'_\gamma$ for any $\gamma \in \Gamma$. Moreover if topologies τ and τ' are linear, then U_γ and U'_γ are subgroups of the group G.*

Proof. Let $\{V_\alpha | \alpha \in \Omega\}$ and $\{V'_\beta | \beta \in \Delta\}$ be some bases of the neighborhoods of zero in topological groups (G, τ) and (G, τ'), respectively, moreover, if topological groups (G, τ) and (G, τ') are linear, then V_α and V'_β are subgroups of the group G.

Proof of the statement 2.1. For any $\alpha \in \Omega$ and $\beta \in \Delta$ we shall consider sets $W_{\alpha,\beta} = V_\alpha + (V'_\beta \bigcap G')$ and $W'_{\alpha,\beta} = V'_\beta + (V_\alpha \bigcap G')$ and we shall show that sets $\{W_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ and $\{W'_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ are required bases of the neighborhoods of zero in topological groups (G, τ) and (G, τ'), respectively.

As G' is a subgroup and $V'_\beta \bigcap G' \subseteq G'$ and $V_\alpha \bigcap G' \subseteq G'$, then

$$W_{\alpha,\beta} \bigcap G' = \big(V_\alpha + (V'_\beta \bigcap G')\big) \bigcap G' = (V_\alpha \bigcap G') + (V'_\beta \bigcap G') =$$

$$\big(V'_\beta + (V_\alpha \bigcap G')\big) \bigcap G' = W'_{\alpha,\beta} \bigcap G'.$$

Let's check up now that the sets $\{W_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ and $\{W'_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ are bases of the neighborhoods of zero in topological groups (G, τ) and (G, τ'), accordingly.

As $V_\alpha = V_\alpha + 0 \subseteq V_\alpha + (V'_\beta \bigcap G') = W_{\alpha,\beta}$, then the set $W_{\alpha,\beta}$ is a neighborhood of zero in the topological group (G, τ).

If U is an arbitrary neighborhood of zero in (G, τ), then $V_{\alpha_0} \subseteq U$ for some $\alpha_0 \in \Omega$. As (G, τ) is a topological group, then there exists such $\alpha_1 \in \Omega$ that $V_{\alpha_1} + V_{\alpha_1} \subseteq V_{\alpha_0}$, and as $\tau|_{G'} = \tau'|_{G'}$ there exists such $\beta_1 \in \Delta$ that $V'_{\beta_1} \bigcap G' \subseteq V_{\alpha_1} \bigcap G'$. Then

$$W_{\alpha_1,\beta_1} = V_{\alpha_1} + (V'_{\beta_1} \bigcap G') \subseteq V_{\alpha_1} + V_{\alpha_1} \subseteq V_{\alpha_0} \subseteq U.$$

Hence $\{W_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ is a basis of the neighborhoods of zero in the topological group (G, τ).

It is similarly checked that the set $\{W'_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ is a basis of the neighborhoods of zero in the topological group (G, τ').

It is easy to see that if V_α and V'_β are subgroups of the group G, then $W_{\alpha,\beta}$ and $W'_{\alpha,\beta}$ will be subgroups in the group G.

The statement 2.1 is completely proved.

Proof of the statement 2.2. For any $\alpha \in \Omega$ and $\beta \in \Delta$ we shall consider sets $W_{\alpha,\beta} = V_\alpha \bigcap (\omega)^{-1}(\omega(V'_\beta))$ and $W'_{\alpha,\beta} = V'_\beta \bigcap \omega^{-1}(\omega(V_\alpha))$. Also we shall show that sets $\{W_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ and $\{W'_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ are required bases of the neighborhoods of zero in topological groups (G, τ) and (G, τ'), accordingly.

Let's check up in the beginning that $\omega(W_{\alpha,\beta}) = \omega(W'_{\alpha,\beta})$.

If $\overline{g} \in \omega(W_{\alpha,\beta})$, then $\overline{g} = \omega(g)$ for some $g \in W_{\alpha,\beta} = V_\alpha \bigcap \omega^{-1}(\omega(V'_\beta))$ and hence there exists such $g' \in V'_\beta$ that $g - g' \in G'$. Then $g' \in V_\alpha + G' = \omega^{-1}(\omega(V_\alpha))$ and hence $g' \in \omega^{-1}(\omega(V_\alpha)) \bigcap V'_\beta = W'_{\alpha,\beta}$, and $\overline{g} = \omega(g) = \omega(g') \in \omega(W'_{\alpha,\beta})$.

From the arbitrarity of the element $\overline{g}$ it follows that $\omega(W_{\alpha,\beta}) \subseteq \omega(W'_{\alpha,\beta})$.

It is similarly proved that $\omega(W'_{\alpha,\beta}) \subseteq \omega(W_{\alpha,\beta})$, and hence $\omega(W_{\alpha,\beta}) = \omega(W'_{\alpha,\beta})$.

Let's check up now that the sets $\{W_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ and $\{W'_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ are bases of the neighborhoods of zero in topological groups (G, τ) and (G, τ'), accordingly.

Let $\alpha \in \Omega$ and $\beta \in \Delta$. As $\omega : (G, \tau') \to (G, \tau')/G' = (G, \tau)/G'$ is an open homomorphism, then for any $\beta \in \Delta$ the set $\omega(V_\beta)$ is a neighborhood of zero in the topological group $(G, \tau)/G'$, and hence $\omega^{-1}(\omega(V_\beta))$ will be a neighborhood of zero in topological group (G, τ). Then the set $W_{\alpha,\beta} = V_\alpha \bigcap \omega^{-1}(\omega(V'_\beta))$ will also be a neighborhood of zero in the topological group (G, τ).

Besides, if U is a neighborhood of zero in the topological group (G, τ), then $V_\alpha \subseteq U$ for some $\alpha \in \Omega$, and hence $W_{\alpha,\beta} = V_\alpha \bigcap \omega^{-1}(\omega(V'_\beta) \subseteq V_\alpha \subseteq U$.

Hence the set $\{W_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ is a basis of a neighborhoods of zero in topological group (G, τ).

It is similarly checked that the set $\{W'_{\alpha,\beta} | \alpha \in \Omega,\ \beta \in \Delta\}$ is a basis of the neighborhoods of zero in topological group (G, τ').

It is easy to see that if V_α and V'_β are subgroups of the group G, then $W_{\alpha,\beta}$ and $W'_{\alpha,\beta}$ will be subgroups in the group G.

The statement 2.2 is completely proved.

Proof of the statement 2.3. Let $\psi : G \to G/G_2$ be the natural homomorphism.

If $G_2 \subseteq G_1$, then with accordance to the statement 2.1 topological groups (G, τ) and (G, τ') possess such bases $\{W_\gamma | \gamma \in \Gamma\}$ and $\{W'_\gamma | \gamma \in \Gamma\}$ of the neighborhoods of zero, respectively, that $G_1 \bigcap W_\gamma = G_1 \bigcap W'_\gamma$ for any $\gamma \in \Gamma$, moreover if topologies τ and τ' are linear, then W_γ and $W'_\gamma|$ will be subgroups of the group G.

For every $\gamma \in \Gamma$ we shall consider sets $U_\gamma = W_\gamma \bigcap (W'_\gamma + G_2)$ and $U'_\gamma = W'_\gamma \bigcap (W_\gamma + G_2))$. In the proof of the statement 2.2 it is demonstrated that sets $\{U_\gamma | \gamma \in \Gamma\}$ and $\{U'_\gamma | \gamma \in \Gamma\}$ are bases of the neighborhoods of zero in topological groups (G, τ) and (G, τ'), respectively, and $\psi(U_\gamma) = \psi(U'_\gamma)$ for any $\gamma \in \Gamma$. Moreover if topological groups (G, τ) and (G, τ') are linear, then U_γ and U'_γ will be subgroups of group G.

As $G_2 \subseteq G_1$, then $(W'_\gamma + G_2) \bigcap G_1 = (W'_\gamma \bigcap G_1) + G_2 = (W_\gamma \bigcap G_1) + G_2 = (W_\gamma + G_2) \bigcap G_1$. Then $U_\gamma \bigcap G_1 = W_\gamma \bigcap (W'_\gamma + G_2) \bigcap G_1 =$

$$W_\gamma \bigcap G_1 \bigcap (W'_\gamma + G_2) = W'_\gamma \bigcap G_1 \bigcap (W_\gamma + G_2) = U'_\gamma \bigcap G_1.$$

The statement 2.3 in this case is proved.

Let now $G_1 \subseteq G_2$. Then in accordance with the statement 2.2 topological groups (G, τ) and (G, τ') possess such bases $\{W_\gamma | \gamma \in \Gamma\}$ and $\{W'_\gamma | \gamma \in \Gamma\}$ of the neighborhoods of zero, respectively, that $\psi(W_\gamma) = \psi(W'_\gamma)$ for any $\gamma \in \Gamma$, moreover if topologies τ and τ' are linear, then W_γ and $W'_\gamma|$ will be subgroups of the group G.

For every $\gamma \in \Gamma$ we shall consider sets $U_\gamma = W_\gamma + (W'_\gamma \bigcap G_1)$ and $U'_\gamma = W_\gamma + (W_\gamma \bigcap G_1)$.

In the proof of the statement 2.2 it is demonstrated that sets $\{U_\gamma | \gamma \in \Gamma\}$ and $\{U'_\gamma | \gamma \in \Gamma\}$ are bases of the neighborhoods of zero in topological groups (G, τ) and (G, τ'), respectively, and $U_\gamma \bigcap G_1 = U_\gamma \bigcap G_1$.

As $G_1 \subseteq G_2$ and $\psi(G_2) = \{0\}$, then

$$\psi(U_\gamma) = \psi(W_\gamma + (W'_\gamma \bigcap G_1)) = \psi(W_\gamma) = \psi(W'_\gamma) =$$

$$\psi(W'_\gamma + (W_\gamma \bigcap G_1)) = \psi(U'_\gamma),$$

moreover if topological groups (G, τ) and (G, τ') are linear then U_γ and U'_γ will be subgroups of the group G for any $\gamma \in \Gamma$..

So, the proposition is completely proved.

3. Definition. *As usual, we shall name a subgroup A of the Abelian groups G a serving subgroup in G if for any natural number k and any element $a \in A$ from the resolvability of the equation $kx = a$ in the group G its resolvability in A follows.*

4. Remark. From the definition of the serving subgroup the following statements follow:

4.1. If $g \in G$ is such an element of group G that $p^{n-1} \cdot g \neq 0$ then the subgroup $< g >= \{kg | k \in \mathbb{N}\}$ is a serving subgroup in the group G;

4.2. The direct sum of any number of serving subgroups of the group G is a serving subgroup in the group G.

5. Theorem (Priufer–Kulikov, see [2, p. 154]). *Every serving subgroup A of a group G is a direct summand in the group G.*

6. Proposition. *Let C be a serving subgroup of the group G. If C is the direct sum of cyclic subgroups of the period p^n and B is such subgroup of the group G that $C \bigcap B = \{0\}$, then there exists such subgroup A of the group G that $B \subseteq A$ and G is the direct sum of subgroups C and A.*

Proof. We shall consider the set Δ of all such subgroups D of the group G that $B \subseteq D$ and $D \bigcap C = \{0\}$. As the sum of ascendent chain of subgroups from Δ belongs to Δ, then Δ contains maximal elements. If A is some of these maximal element, then $B \subseteq A$ and $A \bigcap C = \{0\}$.

For finishing the proof of the proposition it is necessary to check up that $G = C + A$.

We assume the contrary, i.e. that $G \neq C + A$, and let $g \notin C + A$. As the period of the group G is equal to p^n, then there exists such natural number $1 \leq s \leq n$ that $p^s \cdot g \in C + A$ and $p^{s-1} \cdot g \notin C + A$. Let $p^s \cdot g = c + a$, where $c \in C$ and $a \in A$. As $A \bigcap C = \{0\}$ and

$$0 = p^n \cdot g = p^{n-s} \cdot (p^s \cdot g) = p^{n-s} \cdot c + p^{n-s} \cdot a,$$

then $p^{n-s} \cdot c = 0$ and as C is the direct sum of cyclic subgroups of the period p^n, then $c = p \cdot c_1$ for some element $c_1 \in C$. Then $a_1 = p^{s-1} \cdot g - c_1 \in G$. As $p^{s-1} \cdot g = a_1 + c_1 \notin A + C$, then $a_1 \notin A$, and $p \cdot a_1 = p \cdot (p^{s-1} \cdot g - c_1) = p^s \cdot g - p \cdot c_1 = p^s \cdot g - c = a \in A$. Then $A_1 = \{0, a_1, 2 \cdot a_1, \ldots, (p-1) \cdot a_1\} + A$ is a subgroup of the group G, and $B \subsetneq A \subsetneq A_1$.

From the definition of the subgroup A it follows that $A_1 \bigcap C \neq \{0\}$, and hence $0 \neq k \cdot g + a_1 \in C$ for some natural number $k \leq p - 1$ and some element $a_2 \in A$.

As numbers k and p^n are coprime numbers, then there exist such integers l and m that $l \cdot k + m \cdot p^n = 1$. Then $g = (l \cdot k + m \cdot p^n) \cdot g = l \cdot k \cdot g + p^n \cdot g = l \cdot k \cdot g \in l \cdot (a_2 + C) \subseteq A + C$. We arrived at the contradiction with the choice of the element g.

So the proposition is completely proved.

7. Proposition. *Let $\{g_\gamma | \gamma \in \Gamma\}$ be a set of elements of the group G of order p^n and $G' = \{g \in G | p^{n-1} \cdot g = 0\}$. If the set $\{\omega(g_\gamma) | \gamma \in \Gamma\}$ is linear independent in the linear space G/G', then $A =< \{g_\gamma | \gamma \in \Gamma\}$ is a serving subgroup in the group G and $A = \bigoplus\limits_{\gamma \in \Gamma} < g_\gamma >$.*

Proof. From the Remark 4 it follows that for the proof of the proposition it is enough to prove that $A = \bigoplus\limits_{\gamma \in \Gamma} < g_\gamma >$.

We assume the contrary, i.e. that $A \neq \bigoplus\limits_{\gamma \in \Gamma} < g_\gamma >$. As $\sum\limits_{\gamma \in \Gamma} < g_\gamma >= A$, then there exist such subsets $\{g_{\gamma_1}, \ldots, g_{\gamma_k}\} \subseteq \{g_\gamma | \gamma \in \Gamma\}$ and $\{t_1, \ldots, t_k\} \subseteq \mathbb{N}$ that $\sum\limits_{i=1}^{k} t_i \cdot g_{\gamma_i} = 0$ and $t_i \cdot g_{\gamma_i} \neq 0$ for $i = 1, \ldots, k$.

Let $t_i = s_i \cdot p^{j_i}$, where $0 < s_i$ and s_i are not divisible by p for $i = 1, \ldots, k$.

If $j = min\{j_1, \dots, j_k\}$ and $S = \{i | j_i = j\}$, then $p^{n-1-j} \cdot t_i$ are divisible by p^n for $i \notin S$. Then

$$0 = p^{n-1-j} \cdot 0 = p^{n-1-j} \cdot (\sum_{i=1}^{k} t_i \cdot g_{\gamma_i}) = \sum_{i=1}^{k} (s_i \cdot p^{j_i - j}) \cdot p^{n-1} g_{\gamma_i} =$$

$$\sum_{i=1}^{k} (s_i \cdot p^{j_i - j}) \cdot \omega(g_{\gamma_i}) = \sum_{i \in S} s_i \cdot \omega(g_{\gamma_i}).$$

We arrived at the contradiction with the fact that the set $\{\omega(g_\gamma) | \gamma \in \Gamma\}$ is linear independent in the linear space G/G'.

8. Proposition. *Let $G' = \{g \in G | p^{n-1} \cdot g = 0\}$ and $\{g_\gamma | \gamma \in \Gamma\}$ and $\{g'_\gamma | \gamma \in \Gamma\}$ be such sets of elements of the group G of the order p^n that $\omega(g_\gamma) = \omega(g'_\gamma)$ for any $\gamma \in \Gamma$ and the set $\{\omega(g_\gamma) | \gamma \in \Gamma\}$ is linear independent in the linear space G/G'. If $A = \bigoplus\limits_{\gamma | \gamma \in \Gamma} < g_\gamma >$ and $A' = \bigoplus\limits_{\gamma | \gamma \in \Gamma} < g'_\gamma >$, then for any subgroup B of the group G' are true the following statements:*

8.1. *If $A \bigcap B = \{0\}$, then $A' \bigcap B = \{0\}$;*

8.2. *If $G = A \bigoplus B$, then $G = A' \bigoplus B$.*

Proof 8.1. Assume the contrary, and let $0 \neq b \in A' \bigcap B$, i.e. $b = \sum\limits_{i=1}^{k} r_i \cdot g'_{\gamma_i}$. As $\omega(g_\gamma) = \omega(g'_\gamma)$ for any $\gamma \in \Gamma$, then $h_{\gamma_i} = g_{\gamma_i} - g'_{\gamma_i} \in G'$.

If $r_i = p^{s_i} \cdot q_i$, where q_i are not divisible by p and $s = min\{s_1, \dots, s_k\}$, then $p^{n-1-s} \cdot r_i \cdot g_{\gamma_i} \neq 0$ for some number $1 \leq i \leq k$. As $A = \bigoplus\limits_{\gamma | \gamma \in \Gamma} < g_\gamma >$, then $\sum\limits_{i=1}^{k} p^{n-1-s} \cdot r_i \cdot g_{\gamma_i} \neq 0$.

Subsequently

$$p^{n-1-s} \cdot b = p^{n-1-s} \cdot (\sum_{i=1}^{k} r_i \cdot \gamma'_i) = p^{n-1-s} \cdot (\sum_{i=1}^{k} r_i \cdot g_{\gamma_i} - h_{\gamma_i}) =$$

$$\sum_{i=1}^{k} p^{n-1-s} \cdot r_i \cdot g_{\gamma_i} - \sum_{i=1}^{k} p^{n-1-s} \cdot r_i \cdot h_{\gamma_i} = \sum_{i=1}^{k} p^{n-1-s} \cdot r_i \cdot g_{\gamma_i} \neq 0.$$

But this contradicts the equality $A \bigcap B = \{0\}$.

The statement 8.1 is proved.

Proof 8.2. As $G = A \bigoplus B$, then $A \bigcap B = \{0\}$. Then, according to the statement 8.1, $A' \bigcap B = \{0\}$ and according to Proposition 6, there exists such subgroup B' that $B \subseteq B'$ and $G = A' \bigoplus B'$. And according to the statement 8.1, $A \bigcap B' = \{0\}$.

So, we have obtained that $B \subseteq B'$ and $A \bigcap B' = \{0\}$. As $G = A \bigoplus B$, then $B = B'$.

The statement 8.2 is proved.

9. Theorem *Let G be any Abelian group of the period p^n and $G_2 = \{g \in G | p \cdot g = 0\}$. If τ and τ' are such metrizable, linear, group topologies that the subgroup $G_1 = \{g \in G | p^{n-1} \cdot g = 0\}$ is a closed subgroup in each of topological groups (G, τ) and (G, τ'), then $\tau|_{G_1} = \tau'|_{G_1}$ and $(G, \tau)/G_2 = (G, \tau')/G_2$ if and only if there exist such group isomorphism $\varphi : G \to G$ that the following conditions are satisfied:*

1. $\varphi(G_1) = G_1$;

2. $g - \varphi(g) \in G_2$ for any $g \in G$;

3. $\varphi : (G, \tau) \to (G, \tau')$ is a topological isomorphism (i.e. open and continuous isomorphism).

Proof. Sufficiency. Let $\varphi : G \to G$ be a group isomorphism such that conditions 1 - 3 are executed.

If $V \in \tau|_{G_1}$, then there exists such $U \in \tau$ that $U \bigcap G_1 = V$. As $\varphi : (G, \tau) \to (G, \tau')$ is a topological isomorphism, then $U' = \varphi(U) \in \tau'$. Because $\varphi : G \to G$ is a bijection mapping and $\varphi(G_1) = G_1$, it follows

$$\varphi(V) = \varphi(U \bigcap G_1) = \varphi(U) \bigcap \varphi(G_1) = U' \bigcap G_1 \in \tau'|_{G_1}.$$

From the arbitrarity of the set V it follows that $\tau|_{G_1} \subseteq \tau'|_{G_1}$.

It is similarly proved that $\tau'|_{G_1} \subseteq \tau|_{G_1}$, and hence $\tau|_{G_1} = \tau'|_{G_1}$.

Now we consider the following commutative diagram:

$$\begin{array}{ccc}
(G, \tau) & \xrightarrow{\varphi} & (G, \tau') \\
\omega \big\downarrow & & \omega \big\downarrow \\
(G, \tau)/G_2 & \xrightarrow{\bar{\varphi}} & (G, \tau')/G_2, \\
\bar{\omega} \big\downarrow & & \bar{\omega} \big\downarrow \\
(G, \tau)/G_1 & \xrightarrow{\widetilde{\varphi}} & (G, \tau')/G_1
\end{array}$$

here ω and $\bar{\omega}$ are natural homomorphisms, and $\bar{\varphi}$ and $\widetilde{\varphi}$ are such isomorphisms that $\bar{\varphi}(g + G_2) = \varphi(g) + G_2$ and $\widetilde{\varphi}(g + G_1) = \widetilde{\varphi}(g) + G_1$.

As $g - \varphi(g) \in G_2$, then $g + G_2 = \varphi(g) + G_2$. Hence $\bar{\varphi}(g + G_2) = \varphi(g) + G_2 = g + G_2$ and $\widetilde{\varphi}(g + G_1) = \widetilde{\varphi}(g) + G_1$, i.e. $\bar{\varphi} : G/G_2) = G/G_2$ and $\widetilde{\varphi} : G/G_1) = G/G_1$ are identical mappings.

From the fact that $\omega : (G, \tau) \to (G, \tau)/G_2$ and $\omega : (G, \tau') \to (G, \tau')/G_2$ are open and continuous homomorphisms it follows that $\bar{\varphi} : (G, \tau)/G_2 \to (G, \tau')/G_2$ is an open and continuous isomorphism, i.e. $(G, \tau)/G_2 = (G, \tau')/G_2)$.

Sufficiency is completely proved.

Necessity. Let τ and τ' be such metrizable, linear, group topologies that $\tau|_{G_1} = \tau'|_{G_1}$ and $(G,\tau)/G_2 = (G,\tau')/G_2$. If $\omega : G \to G/G_2$ and $\bar{\omega} : G/G_2 \to G/G_1 = (G/G_2)/(G_1/G_2)$ are natural homomorphisms, then according to the statement 2.3, there exist sets $\{V_i | i \in \mathbb{N}\bigcup\{0\}\}$ and $\{V'_i | i \in \mathbb{N}\bigcup\{0\}\}$ of subgroups which are bases of the neighborhoods of zero in topological groups (G,τ) and (G,τ'), respectively, and $V_i \bigcap G_1 = V'_i \bigcap G_1$ and $\omega(V_i) = \omega(V'_i)$ for any $i \in \mathbb{N}\bigcup\{0\}$. Without loss of generality, we can consider that $V_0 = V'_0 = G$.

For every $i \in \mathbb{N}$ let $\bar{V}_i = \omega(V_i) = \omega(V'_i)$ and $\widetilde{V}_i = \bar{\omega}(\bar{V}_i)$.

As $\bar{G} = G/G_1$ is a linear space over the field $F_p = \mathbb{Z}/p\cdot\mathbb{Z}$ and $\widetilde{V}_i$ is a subspace of the linear space $\bar{G}$, then for every $i \in \mathbb{N}\bigcup\{0\}$ there exists a set $\{\widetilde{U}_i | i \in \mathbb{N}\bigcup\{0\}\}$ of subspaces of the linear space $\bar{G}$ such that $\widetilde{V}_i = \widetilde{U}_i \bigoplus \widetilde{V}_{i+1}$ for any $i \in \mathbb{N}\bigcup\{0\}$. Then $\widetilde{V}_k = (\bigoplus\limits_{i=k}^{n} \widetilde{U}_i)\bigoplus(\widetilde{V}_{n+1})$ for any $k \le n \in \mathbb{N}\bigcup\{0\}$. As G_1 is a closed subgroup in the topological groups (G,τ) and (G,τ'), then (se [1], theorem 1.3.2) $\bigcap\limits_{k\in\mathbb{N}} \widetilde{V}_k = \{0\}$ and hence $\widetilde{V}_k = \bigoplus\limits_{i=k}^{\infty} \widetilde{U}_i$.

For every $k \in \mathbb{N}\bigcup\{0\}$ we shall consider a basis $\{\widetilde{x}_{k,\gamma} | \gamma \in \Gamma_k\}$ of the linear space $\widetilde{U}_k$.

As $\widetilde{U}_i \subseteq \widetilde{V}_i = \bar{\omega}(\omega(V_i)))$ for any $i \in \mathbb{N}\bigcup\{0\}$, then for any $k \in \mathbb{N}\bigcup\{0\}$ and any $\gamma \in \Gamma_k$ there exists an element $x_{k,\gamma} \in V_k$ such that $\bar{\omega}(\omega(x_{k,\gamma})) = \widetilde{x}_{k,\gamma}$.

As $\omega(V_i) = \omega(V'_i)$ for any $i \in \mathbb{N}\bigcup\{0\}$, then for any $i \in \mathbb{N}\bigcup\{0\}$ and any $\gamma \in \Gamma$ there exists an element $x'_{i,\gamma} \in V'_i$ such that $\omega(x_{i,\gamma}) = \omega(x'_{i,\gamma})$.

According to Proposition 7, the subgroups $A =< \{x_{k,\gamma} | k \in \mathbb{N}\bigcup\{0\}, \gamma \in \Gamma\} >$ and $A' =< \{x'_{k,\gamma} | k \in \mathbb{N}\bigcup\{0\}, \gamma \in \Gamma\} >$ are serving subgroups of the group G and they are direct sums of cyclic groups of the order p^n.

According to the Prufer-Kulikov theorem (see Theorem 5) there exists a subgroup B of the group G such that $G = B \bigoplus A$. Then, according to the statement 8.2, $G = B \bigoplus A'$. As $\bar{\omega}(\omega(A)) = \bar{\omega}(\omega(V_0)) = G/G_1$, then $B \subseteq G_1$.

If $f : \{x_{k,\gamma} | k \in \mathbb{N}\bigcup\{0\}, \gamma \in \Gamma\} \to \{x'_{k,\gamma} | k \in \mathbb{N}\bigcup\{0\}, \gamma \in \Gamma\}$ is a mapping such that $f(x_{k,\gamma}) = x'_{k,\gamma}$ for any $k \in \mathbb{N}\bigcup\{0\}$ and $\gamma \in \Gamma$ then it can be extended to a group isomorphism $\widehat{f} : A \to A'$.

We suppose $\varphi(a + b) = \widehat{f}(a) + b$ for any $a \in A$ and any $b \in B$. Then $\varphi : G \to G$ is a group isomorphism.

As $\omega(x_{k,\gamma}) = \omega(x'_{k,\gamma}) = \bar{x}_{k,\gamma}$ for any $k \in \mathbb{N}\bigcup\{0\}$ and any $\gamma \in \Gamma$, then $h_{k,\gamma} = x_{k,\gamma} - x'_{k,\gamma} \in G_2$.

Let now $g \in G_1$. Then $g = \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b$, where $b \in B \subseteq G_1$. As

$$0 = \bar{\omega}(\omega(g)) = \sum_{i=1}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot \bar{\omega}(\omega(x_{i,\gamma_j})) + \bar{\omega}(\omega(b)) = \sum_{i=1}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot \bar{x}_{i,\gamma_j},$$

then all t_{i,γ_j} are divisible by p, and hence

$$\varphi(g) = \sum_{i=1}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x'_{i,\gamma_j} + \varphi(b) = \sum_{i=1}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot (x_{i,\gamma_j} - h_{i,\gamma_j}) + \varphi(b) =$$

$$\sum_{i=1}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} - \sum_{i=1}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot h_{i,\gamma_j} + b = \sum_{i=1}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b = g.$$

So we have proved that $\varphi(g) = g$ for any $g \in G_1$. Then $\varphi(G_1) = G_1$, i.e. the first statement of the theorems is true.

Let now $g \in G$. Then $g = \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b$, where $b \in B \subseteq G_1$, and hence

$$g - \varphi(g) = \sum_{i=0}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b - (\sum_{i=0}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x'_{i,\gamma_j} + b) =$$

$$\sum_{i=0}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot (x_{i,\gamma_j} - x'_{i,\gamma_j}) = \sum_{i=0}^{k}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot h_{i,\gamma_j} \in G_2,$$

i.e. the second statement of the theorem is also true.

For finishing the proof of the theorem it remained to check up that the isomorphism $\varphi : (G.\tau) \to (G.\tau')$ is a topological isomorphism. For this purpose it is enough to verify that $\varphi(V_{k,\gamma}) = V'_{k,\gamma}$ for any $k \in \mathbb{N}$ and any $\gamma \in \Gamma$.

So, let $g \in V_{k,\gamma}$. Then $g = \sum\limits_{i=0}^{m}\sum\limits_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b$, where $b \in B \subseteq G_1$.

As (see definition of elements $x_{i,\gamma}$) $\sum\limits_{i=k}^{m}\sum\limits_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} \in V_k$, then

$$\sum_{i=0}^{k-1}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b = g - \sum_{i=k}^{m}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} \in V_k.$$

Besides that as

$$\sum_{i=0}^{m}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot \bar{x}_{i,\gamma_j} = \omega(\sum_{i=0}^{m}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b) =$$

$\omega(g) \in \omega(V_k) = \bar{V}_k = \bigoplus\limits_{i=k}^{\infty} \bar{U}_i$, then for any $i < k$ all numbers t_{i,γ_j} are divided by p, and hence $\sum\limits_{i=0}^{k-1}\sum\limits_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} \in G_1$. Then $\sum\limits_{i=0}^{k-1}\sum\limits_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b \in G_1 \bigcap V_k = G_1 \bigcap V'_k$, and hence, $\varphi(\sum\limits_{i=0}^{k-1}\sum\limits_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b) = \sum\limits_{i=0}^{k-1}\sum\limits_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b \in V'_k$. Then

$$\varphi(g) = \varphi(\sum_{i=0}^{k-1}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j} + b) + \varphi(\sum_{i=k}^{m}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x_{i,\gamma_j}) \in$$

$$V'_k + \sum_{i=k}^{m}\sum_{j=1}^{s} t_{i,\gamma_j} \cdot x'_{i,\gamma_j} \subseteq V'_k + V'_k = V'_k.$$

From the arbitrarity of the element g it follows that $\varphi(V_k) \subseteq V'_k$.

In a similar way it can be proved that $\varphi^{-1}(V'_k) \subseteq V_k$, and hence $\varphi(V_k) = V'_k$.

The theorem is completely proved.

References

[1] Arnautov V.I., Glavatsky S.T., Mikhalev A.V. *Introduction to the theory of topological rings and modules*. Marcel Dekker, Inc. New York-Basel-Hong Kong 1996.

[2] Kurosh A.G. *The theory of groups*. Moscow, Nauka, 1967.

Institute of Mathematics and Computer Science
Academy of Sciences of Moldova
Academiei str. 5, MD-2028 Chişinau
Moldova

E-mail: *arnautov@math.md*

Received May 12, 2009

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 29–54
ISSN 1024–7696

A complete classification of quadratic differential systems according to the dimensions of $\mathit{Aff}(2,\mathbb{R})$–orbits

N. Gherstega*, V. Orlov*, N. Vulpe

Abstract. In this article we consider the action of the group $\mathit{Aff}(2,\mathbb{R})$ of affine transformations and time rescaling on real planar quadratic differential systems. Via affine invariant conditions we give a complete stratification of this family of systems according to the dimension $\mathcal{D}$ of affine orbits proving that $3 \le \mathcal{D} \le 6$. Moreover we give a complete topological classification of all the systems located on the orbits of dimension $\mathcal{D} \le 5$ constructing the affine invariant criteria for the realization of each of 49 possible topologically distinct phase portraits.

Mathematics subject classification: 34C14.
Keywords and phrases: Quadratic differential system, Lie algebra of operators, $\mathit{Aff}(2,\mathbb{R})$–orbit, affine invariant polynomial.

We consider here real planar differential systems of the form

$$\begin{aligned}\frac{dx}{dt} &= p_0 + p_1(x,y) + p_2(x,y) \equiv P(x,y),\\ \frac{dy}{dt} &= q_0 + q_1(x,y) + q_2(x,y) \equiv Q(x,y)\end{aligned} \tag{1}$$

with

$$\begin{aligned} p_0 &= a, \quad p_1(x,y) = cx + dy, \quad p_2(x,y) = gx^2 + 2hxy + ky^2,\\ q_0 &= b, \quad q_1(x,y) = ex + fy, \quad q_2(x,y) = lx^2 + 2mxy + ny^2.\end{aligned}$$

We say that these systems are quadratic if $|p_2(x,y)| + |q_2(x,y)| \neq 0$.

Consider also the group $\mathit{Aff}(2,\mathbb{R})$ of affine transformations given by the equalities:

$$\bar{x} = \alpha x + \beta y + \nu, \quad \bar{y} = \gamma x + \delta y + \varkappa, \quad \det\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix} \neq 0, \quad \alpha,\beta,\gamma,\delta,\nu,\varkappa \in \mathbb{R}.$$

According to [1] the operators of the linear representation of the group $\mathit{Aff}(2,\mathbb{R})$

*This work is supported partially by CSSDT grant Ref. Nr.07.411.05INDF.

in the space of the coefficients and variables of systems (1) will take the form

$$\begin{aligned}
X_1 &= x\frac{\partial}{\partial x}+a\frac{\partial}{\partial a}+d\frac{\partial}{\partial d}-e\frac{\partial}{\partial e}-g\frac{\partial}{\partial g}+k\frac{\partial}{\partial k}-2l\frac{\partial}{\partial l}-m\frac{\partial}{\partial m},\\
X_2 &= y\frac{\partial}{\partial x}+b\frac{\partial}{\partial a}+e\frac{\partial}{\partial c}+(f-c)\frac{\partial}{\partial d}-e\frac{\partial}{\partial f}+l\frac{\partial}{\partial g}+\\
&+(m-g)\frac{\partial}{\partial h}+(n-2h)\frac{\partial}{\partial k}-l\frac{\partial}{\partial m}-2m\frac{\partial}{\partial n},\\
X_3 &= x\frac{\partial}{\partial y}+a\frac{\partial}{\partial b}-d\frac{\partial}{\partial c}+(c-f)\frac{\partial}{\partial e}+d\frac{\partial}{\partial f}-2h\frac{\partial}{\partial g}-k\frac{\partial}{\partial h}+\\
&+(g-2m)\frac{\partial}{\partial l}+(h-n)\frac{\partial}{\partial m}+k\frac{\partial}{\partial n},\\
X_4 &= y\frac{\partial}{\partial y}+b\frac{\partial}{\partial b}-d\frac{\partial}{\partial d}+e\frac{\partial}{\partial e}-h\frac{\partial}{\partial h}-2k\frac{\partial}{\partial k}+l\frac{\partial}{\partial l}-n\frac{\partial}{\partial n},\\
X_5 &= \frac{\partial}{\partial x}-c\frac{\partial}{\partial a}-e\frac{\partial}{\partial b}-2g\frac{\partial}{\partial c}-2h\frac{\partial}{\partial d}-2l\frac{\partial}{\partial e}-2m\frac{\partial}{\partial f},\\
X_6 &= \frac{\partial}{\partial y}-d\frac{\partial}{\partial a}-f\frac{\partial}{\partial b}-2h\frac{\partial}{\partial c}-2k\frac{\partial}{\partial d}-2m\frac{\partial}{\partial e}-2n\frac{\partial}{\partial f}.
\end{aligned}\tag{2}$$

These operators form a six-dimensional Lie algebra [1].

Let $\tilde{a}=(a,b,c,d,e,f,g,h,k,\,l,m,n)$ be the 12-tuple of the coefficients of systems (1), i.e. each particular system (1) yields a point in $E^{12}(\tilde{a})$, where $E^{12}(\tilde{a})$ is the Euclidean space of the coefficients of the right-hand sides of systems (1). We denote by $\tilde{a}(q)\in E^{12}(\tilde{a})$ the point which corresponds to the system, obtained from a system (1) with coefficients $\tilde{a}$ via a transformation $q\in \mathit{Aff}(2,\mathbb{R})$.

Definition 1. Consider a system (1) and its corresponding point $\tilde{a}\in E^{12}(\tilde{a})$. We call the set $O(\tilde{a})=\{\tilde{a}(q)|q\in \mathit{Aff}\ (2,\mathbb{R})\}$ the $\mathit{Aff}(2,\mathbb{R})$ - *orbit* of this system.

It is known from [1] that

$$\mathfrak{D}\overset{def}{=}\dim_{\mathbb{R}} O(\tilde{a})=rank\mathcal{M},$$

where $\mathcal{M}$ is the matrix constructed on the coordinate vectors of operators (2):

$$\mathcal{M}=\begin{pmatrix}
a & 0 & 0 & d & -e & 0 & -g & 0 & k & -2l & -m & 0\\
b & 0 & e & -c+f & 0 & -e & l & -g+m & -2h+n & 0 & -l & -2m\\
0 & a & -d & 0 & c-f & d & -2h & -k & 0 & g-2m & h-n & k\\
0 & b & 0 & -d & e & 0 & 0 & -h & -2k & l & 0 & -n\\
-c & -e & -2g & -2h & -2l & -2m & 0 & 0 & 0 & 0 & 0 & 0\\
-d & -f & -2h & -2k & -2m & -2n & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

We denote by $\Delta_{i,j,k,l,m,n}$ the minor of the 6th order of the matrix M_1, constructed on the columns i,j,k,l,m,n $(1\le i<j<k<l<m<n\le 12)$ and by $\Delta^{i_1,i_2,\dots,i_s}_{j_1,j_2,\dots,j_s}$ the minor of the order s $(s=5,4,3)$ constructed on the lines $i_1,i_2,\dots,i_s$ $(1\le i_1<i_2<\dots<i_s\le 6\})$ and on the columns $j_1,j_2,\dots,j_s$ $(1\le j_1<j_2<\dots<j_s\le 12\})$.

In [2] a minimal polynomial basis of $GL(2,\mathbb{R})$-invariant polynomials (which are also named center-affine comitants and invariants) is constructed. We shall use here the following elements of this basis, defined in tensorial form (we keep the notations from [2]):

$$\begin{gathered}
I_1 = a^\alpha_\alpha,\ I_2 = a^\alpha_\beta a^\beta_\alpha,\ I_3 = a^\alpha_p a^\beta_{\alpha q} a^\gamma_{\beta\gamma}\varepsilon^{pq},\ I_5 = a^\alpha_p a^\beta_{\gamma q} a^\gamma_{\alpha\beta}\varepsilon^{pq},\\
I_7 = a^\alpha_{pr} a^\beta_{\alpha q} a^\gamma_{\beta s} a^\delta_{\gamma\delta}\varepsilon^{pq}\varepsilon^{rs},\ I_8 = a^\alpha_{pr} a^\beta_{\alpha q} a^\gamma_{\delta s} a^\delta_{\beta\gamma}\varepsilon^{pq}\varepsilon^{rs},\ I_9 = a^\alpha_{pr} a^\beta_{\beta q} a^\gamma_{\gamma s} a^\delta_{\alpha\delta}\varepsilon^{pq}\varepsilon^{rs},\\
I_{18} = a^\alpha a^q a^p_\alpha \varepsilon_{pq},\ I_{21} = a^\alpha a^\beta a^q a^p_{\alpha\beta}\varepsilon_{pq},\ K_1 = a^\alpha_{\alpha\beta}x^\beta,\ K_2 = a^p_\alpha x^\alpha x^q \varepsilon_{pq},\\
K_3 = a^\alpha_\beta a^\beta_{\alpha\gamma}x^\gamma,\ K_4 = a^\alpha_\gamma a^\beta_{\alpha\beta}x^\gamma,\ K_5 = a^p_{\alpha\beta}x^\alpha x^\beta x^q\varepsilon_{pq},\ K_6 = a^\alpha_{\alpha\beta}a^\beta_{\gamma\delta}x^\gamma x^\delta,\\
K_7 = a^\alpha_{\beta\gamma}a^\beta_{\alpha\delta}x^\gamma x^\delta,\ K_{11} = a^p_\alpha a^\alpha_{\beta\gamma}x^\beta x^\gamma x^q\varepsilon_{pq},\ K_{13} = a^\alpha_\gamma a^\beta_{\alpha\beta}a^\gamma_{\delta\mu}x^\delta x^\mu,\\
K_{21} = a^p x^q \varepsilon_{pq},\ K_{23} = a^p a^q_{\alpha\beta}x^\alpha x^\beta \varepsilon_{pq}.
\end{gathered} \tag{3}$$

Here the following notations are used:

$$\begin{gathered}
a^1 = a,\quad a^2 = b,\quad a^1_1 = c,\quad a^1_2 = d,\quad a^2_1 = e,\quad a^2_2 = f,\quad a^1_{11} = g,\quad a^1_{12} = h,\\
a^1_{22} = k,\quad a^2_{11} = l,\quad a^2_{12} = m,\quad a^2_{22} = n,\quad x^1 = x,\quad x^2 = y,
\end{gathered}$$

and the unit bi-vectors ε^{pq} and ε_{pq} have the coordinates: $\varepsilon^{11} = \varepsilon^{22} = \varepsilon_{11} = \varepsilon_{22} = 0$, $\varepsilon^{12} = -\varepsilon^{21} = \varepsilon_{12} = -\varepsilon_{21} = 1$,

We consider the polynomials

$$\begin{aligned}
C_i(\tilde a, x, y) &= y p_i(\tilde a, x, y) - x q_i(\tilde a, x, y) \in \mathbb{R}[\tilde a, x, y],\ i = 0,1,2,\\
D_i(\tilde a, x, y) &= \frac{\partial}{\partial x} p_i(\tilde a, x, y) + \frac{\partial}{\partial y} q_i(\tilde a, x, y) \in \mathbb{R}[\tilde a, x, y],\ i = 1,2,
\end{aligned} \tag{4}$$

which are the only GL-comitants of degree one with respect to the coefficients of systems (1) that could exist for these systems. Comparing (3) with (4) we have the following identities: $C_0 \equiv K_{21}$, $C_1 \equiv K_2$, $C_2 \equiv K_5$, $D_1 \equiv I_1$, $D_2 \equiv 2K_1$.

Using the so-called *transvectant of index k* (see [3]) of two real polynomials f and g:

$$(f,g)^{(k)} = \sum_{h=0}^{k} (-1)^h \binom{k}{h} \frac{\partial^k f}{\partial x^{k-h}\partial y^h}\ \frac{\partial^k g}{\partial x^h \partial y^{k-h}}$$

we shall construct the following GL–comitants of the second degree with respect to the coefficients of initial systems:

$$\begin{aligned}
&T_1 = (C_0, C_1)^{(1)}, \quad &&T_2 = (C_0, C_2)^{(1)}, \quad &&T_3 = (C_0, D_2)^{(1)},\\
&T_4 = (C_1, C_1)^{(2)}, \quad &&T_5 = (C_1, C_2)^{(1)}, \quad &&T_6 = (C_1, C_2)^{(2)},\\
&T_7 = (C_1, D_2)^{(1)}, \quad &&T_8 = (C_2, C_2)^{(2)}, \quad &&T_9 = (C_2, D_2)^{(1)}.
\end{aligned}$$

According to [4] the transvectant $(f,g)^{(k)}$ of two GL–comitants (respectively T–comitants) of systems (1) is a GL–comitant (respectively T–comitant) of these systems too.

In what follows we shall construct the following T–comitants (and CT–comitants, see [5] for detailed definitions), which are responsible for the dimensions of the affine orbits for systems (1):

$$\beta(\tilde{a}) = 27I_8 - I_9 - 18I_7 = -2\mathrm{Discrim}\big(C_2(\tilde{a},x,y)\big),$$
$$M(\tilde{a},x,y) = 2\,\mathrm{Hess}\big(C_2(x,y)\big), \quad \widehat{H}(\tilde{a},x,y) = (-T_8 + 8T_9 + 2D_2^2)/72,$$
$$D(\tilde{a},x,y) = \big[2C_0(T_8 - 8T_9 - 2D_2^2) + C_1(6T_7 - T_6) - (C_1,T_5)^{(1)} +$$
$$+6D_1(C_1D_2 - T_5) - 9D_1^2C_2\big]/36, \quad U_1(\tilde{a},x,y) = (C_2,D)^{(1)},$$
$$U_2(\tilde{a},x,y) = I_1K_1^2(2K_1^2K_2 - 2K_2K_6 - K_1K_{11}) - 2K_1^3(K_2K_4 + 2K_7K_{21}) +$$
$$+4K_1K_6^2K_{21} + K_1^2[2K_4K_{11} + K_2K_{13} + 2K_{23}(K_6 + K_7)] - 4K_6^2K_{23},$$
$$U_3 = K_2^2 - 4K_5K_{21}.$$

However we also need several affine comitants which we shall construct here following [6]. Denote by $\tilde{a}(\tau)$ the point from the space $E^{12}(\tilde{a})$ that corresponds to the system, obtained from a system (1) with coefficients $\tilde{a}$ via a translation $\tau : x = \bar{x} + x_0,\ y = \bar{y} + y_0$. It is evident that $\tilde{a}(\tau) = \tilde{a}(x_0,y_0)$. According to [6] if $I(\tilde{a})$ is a center-affine invariant of systems (1), then the polynomial

$$\bar{K}(\tilde{a},x,y) = I(\tilde{a}(x_0,y_0))|_{\{x_0=x,y_0=y\}}$$

is an affine comitant of these systems. So, considering (3) we obtain the following affine comitants of systems (1):

$$Af_i(\tilde{a},x,y) = I_i(\tilde{a}(x_0,y_0))|_{\{x_0=x,y_0=y\}}, \quad (i = 1,2,5,18,21).$$

We shall use the notations

$$\begin{array}{ll} W_1 = Af_1^2 - Af_2, & W_2 = Af_1Af_{18} - Af_{21}, \\ V_1 = Af_1^2 - 2Af_2, & V_2 = Af_1Af_{18} - 4Af_{21}, \\ \mathcal{V}_1 = \widehat{H}^2 + Af_{21}^2, & \mathcal{U} = Af_5^2 + U_1^2 + U_2^2, \\ \mathcal{V}_2 = D^2 + U_3^2. & \end{array} \tag{5}$$

In what follows by $\bar{U}(\tilde{a},x,y) = 0$ (where $\bar{U}(\tilde{a},x,y)$ is an arbitrary comitant) we shall understand $\bar{U}(\tilde{a},x,y) = 0$ in $\mathbb{R}[x,y]$ (i.e. this comitant identically vanishes as a polynomial in x and y).

Taking into consideration Remark 1 (see below) according to [7] we have the next result.

Proposition 1. *A system* (1) *is located on the affine orbit of the dimension six if and only if one of the following three sets of conditions holds:*

$$(i)\ \beta \neq 0; \quad (ii)\ \beta = 0,\ K_5\mathcal{U} \neq 0; \quad (iii)\ \beta = 0,\ K_5 = 0, Af_5 \neq 0.$$

Remark 1. In Proposition 1 we use the set of conditions $\beta = 0,\ K_5 = 0, Af_5 \neq 0$ which is equivalent to the set of conditions $\beta = 0,\ K_5 = 0,\ Af_4(Af_4 - Af_3) \neq 0$ from [7, Theorem, page 126]. We note also that $\mathcal{U}$ here denotes the expression $Af_5^2 + T_{16}^2 + Kom^2$ from [7].

Considering Proposition 1 it remains to construct the affine invariant criteria for a system (1) to be located on the orbit of a given dimension $s \le 5$.

Lemma 1. *The rank of the matrix $\mathcal{M}$ is equal to five if and only if $\beta = 0$, $\mathcal{U} = 0$ and one of the following four sets of conditions holds:*

$$\begin{array}{llll} (i) & M\mathcal{V}_1 \neq 0; & (ii) & M = 0,\ K_5 W_1 V_2 \neq 0; \\ (iii) & M = W_1 = 0,\ K_5 W_2 \neq 0; & (iv) & K_5 = 0,\ W_2 \neq 0. \end{array}$$

Proof. By Proposition 1 a system (1) is located on the affine orbit of the dimension less than six (i.e. the rank of the matrix $\mathcal{M}$ is ≤ 5) if and only if $\beta = 0$ and either

$$(\alpha_1)\ K_5 \neq 0 \text{ and } \mathcal{U} = 0, \quad \text{or} \quad (\alpha_2)\ K_5 = Af_5 = 0. \tag{6}$$

1) Assume first $K_5 \neq 0$. As $\beta = 0$ following [2] we could use a center-affine transformation which brought the binary form $K_5(x, y)$ to the canonical form: $K_5(x, y) = x^2(x + \delta y)$ with $\delta \in \{0, 1\}$. Moreover, the same transformation will bring systems (1) in this case to the form (excluding also the linear term x in the second equation via an additional translation):

$$\begin{aligned} \dot{x} &= a + cx + dy + (2m + \delta)x^2 + 2hxy, \\ \dot{y} &= b + fy - x^2 + 2mxy + 2hy^2, \quad \delta \in \{0, 1\}. \end{aligned} \tag{7}$$

For these systems we calculate $M = -8x^2\delta^2$ and we shall consider two subcases: $M \neq 0$ and $M = 0$.

a) If $M \neq 0$ then $\delta = 1$. Since according to (5) the condition $\mathcal{U} = 0$ implies $Af_5 = U_1 = U_2 = 0$ we have: Coefficient$[Af_5, y] = -2h^2 = 0$. So we obtain $h = 0$ and then $Af_5 = -d(5m^2 + 4m + 1) = 0$ and this evidently yields $d = 0$. Therefore we obtain the systems

$$\dot{x} = a + cx + (2m + 1)x^2, \quad \dot{y} = b + fy - x^2 + 2mxy, \tag{8}$$

for which calculations yield:

$$\begin{aligned} U_1 &= 6m(cf - f^2 - 2am - 2bm)x^4, \\ U_2 &= (1 + 3m)^3(cf - f^2 - 2am - 2bm)x^6. \end{aligned} \tag{9}$$

We note that the GL-invariant U_2 keeps the value, indicated in (9), after any translation $(x, y) \mapsto (\tilde{x} + x_0, \tilde{y} + y_0)$ with arbitrary $(x_0, y_0) \in \mathbb{R}^2$ applied to systems (8). In other words for any system located in the orbit under the translation group action of a system (8), i.e. for systems of the form

$$\begin{aligned} \dot{x} &= a + x_0(c + x_0 + 2mx_0) + (c + 2x_0 + 4mx_0)x + (2m + 1)x^2, \\ \dot{y} &= b - x_0(x_0 - 2my_0) + fy_0 - 2x(x_0 - my_0) + (f + 2mx_0)y - x^2 + 2mxy, \end{aligned}$$

we have $U_2 = (1 + 3m)^3(cf - f^2 - 2am - 2bm)x^6$. This means that the polynomial U_2 is a CT-comitant [5] for the family of systems (8) and hence the condition $U_2 = 0$ is affine invariant.

Clearly the conditions $U_1 = U_2 = 0$ (i.e. $\mathcal{U} = 0$) imply $(cf - f^2 - 2am - 2bm) = 0$ and then we can convince ourself that all the minors of order 6 of the matrix M_1 vanish. We claim that the existence of at least one nonzero minor of order 5 is equivalent to the condition $\mathcal{V}_1 \neq 0$, i.e. considering (5) to the condition $\widehat{H}^2 + Af_{21}^2 \neq 0$.

Indeed, for systems (8) we calculate $\widehat{H} = m^2x^2$. On the other hand we obtain $\Delta^{1,2,4,5,6}_{5,6,10,11,12} = -8m^4$, i.e. if $\widehat{H} \neq 0$ then $rank(M_1) = 5$.

Assume $\widehat{H} = 0$, i.e. $m = 0$. Then for systems (8) we have $Af_{21} = [a + b + cx + fy](a+cx+x^2)^2$. At the same time we calculate $\Delta^{1,2,4,5,6}_{2,5,7,8,10} = 2f$, $\Delta^{1,2,3,4,5}_{3,5,7,8,10} = 2(f-c)$ and $\Delta^{1,2,3,4,5}_{2,3,7,10,11} = 2(a + b)$. As $Af_{21} \neq 0$ is equivalent to $(a + b)^2 + c^2 + f^2 \neq 0$ we conclude that in this case there exist non-zero minors of order 5. It remains to observe that in the case $Af_{21} = 0$ (i.e. $f = c = a + b = 0$) all the minors of order 5 vanish. Thus, our claim is proved.

b) Assume now $M = 0$, i.e. for systems (7) we have $\delta = 0$. In order to examine the condition $\mathcal{U} = 0$ (i.e. $Af_5^2 + U_1^2 + U_2^2 = 0$, see (5)) for these systems we calculate: Coefficient$[Af_5, x] = -6h^2 = 0$ and this yields $h = 0$. Therefore we have $Af_5 = -5dm^2 = 0$, i.e. $dm = 0$ and then we obtain Coefficient$[U_1, x^3y] = -6d^2$. So the condition $U_1 = 0$ yields $d = 0$ and this leads to the systems

$$\dot{x} = a + cx + 2mx^2, \quad \dot{y} = b + fy - x^2 + 2mxy, \tag{10}$$

for which calculations yield:

$$U_1 = 6m(cf - f^2 - 2am)x^4, \quad U_2 = 27m^3(cf - f^2 - 2am)x^6. \tag{11}$$

We note that the GL-invariant U_2 keeps the value, indicated above, after any translation $(x, y) \mapsto (\tilde{x} + x_0, \tilde{y} + y_0)$ with arbitrary $(x_0, y_0) \in \mathbb{R}^2$ applied to systems (10). This means that the polynomial U_2 is a CT-comitant [5] for the family of systems (10) and hence the condition $U_2 = 0$ is invariant under the affine transformation.

Evidently the conditions $U_1 = U_2 = 0$ (i.e. $\mathcal{U} = 0$) imply $m(cf - f^2 - 2am) = 0$ and then all the minors of order 6 for the matrix M_1 vanish. We shall consider two subcases: $m \neq 0$ and $m = 0$.

b_1) If $m \neq 0$ then without loss of generality for systems (10) we may assume $f = 0$ due to the translation $(x, y) \mapsto \left(x - \frac{f}{2m}, y - \frac{f}{2m^2}\right)$. Therefore considering (11) in this case the conditions $U_1 = U_2 = 0$ yield $a = 0$. So we get the systems

$$\dot{x} = cx + 2mx^2, \quad \dot{y} = b - x^2 + 2mxy, \tag{12}$$

for which $W_1 = 4mx(c + 4mx)$. We note that all the minors of order 6 for the matrix M_1 corresponding to these systems vanish. On the other hand we have $\Delta^{1,2,4,5,6}_{5,6,8,10,11} = -4m^4 \neq 0$, i.e. $rank(M_1) = 5$. It remains to observe that for systems (12) Coefficient$[V_2, x^6] = -8m^3$. Therefore the condition $m \neq 0$ implies $V_2 \neq 0$ and hence the conditions *(ii)* of Lemma 1 are valid.

b_2) Assuming $m = 0$ and considering (10) we get the family of systems

$$\dot{x} = a + cx, \quad \dot{y} = b + fy - x^2, \tag{13}$$

for which we calculate

$$W_1 = 2cf, \quad \text{Coefficient}[V_2, xy] = cf(c^2 - f^2). \tag{14}$$

We shall examine two cases: $W_1 \neq 0$ and $W_1 = 0$.

γ_1) Admit first $W_1 \neq 0$, i.e. $cf \neq 0$. If $c^2 - f^2 \neq 0$ (this implies $V_2 \neq 0$) by (14) we obtain $cf(c-f) \neq 0$. Therefore $rank(M_1) = 5$ as we have $\Delta^{2,3,4,5,6}_{1,2,5,10,11} = cf(c-f)$.

Assume $c^2 - f^2 = 0$. If $f = c$ (respectively $f = -c$) for systems (13) we calculate $V_2 = -4a(a+cx)^2$ (respectively $V_2 = -4(a+cx)^3$). On the other hand we have $\Delta^{1,2,3,4,5}_{1,2,5,10,11} = 2a^2$ (respectively $\Delta^{2,3,4,5,6}_{1,2,4,5,10} = 4c^4$) and evidently if $V_2 \neq 0$ then $rank(M_1) = 5$ (we note that in the second case the condition $W_1 \neq 0$ yields $c \neq 0$ and this implies $V_2 \neq 0$). Moreover straightforward calculations show us that the condition $V_2 = 0$ (and this happens only in the first case) implies $a = 0$ and all the minors of order 5 vanish. So, the conditions *(ii)* of Lemma 1 are true.

γ_2) Suppose now $W_1 = 0$, i.e. $cf = 0$. In this case for systems (13) we obtain:

$$\begin{array}{lll} (\beta_1) & W_2 = -a(a^2 + bf^2 - 2afx - f^2x^2 + f^3y) & \text{if} \quad c = 0; \\ (\beta_2) & W_2 = (bc^2 - a^2)(a + cx) & \text{if} \quad f = 0. \end{array} \tag{15}$$

On the other hand in the case (β_1) (respectively (β_2)) we have that $\Delta^{1,2,3,4,5}_{1,2,5,10,11}$ equals $2a^2$ (respectively $2(a^2 - bc^2)$. So if $W_2 \neq 0$ then $rank(M_1) = 5$, and it can be easily verified that in the case (β_1) as well as in the case (β_2) the condition $W_2 = 0$ implies the vanishing of all the minors of order 5. This completes the proof of the conditions *(iii)* of Lemma 1.

Remark 2. It follows from the reasons above that in the case $m \neq 0$ for systems (10) we have $V_2 \neq 0$. Hence we decide that in the case $\mathcal{U} = V_2 = 0$ and $W_1 \neq 0$ for systems (10) the relations $m = 0$, $f = c \neq 0$ and $a = 0$ hold.

2) Assume finally $K_5 = 0$ (see condition (α_2) from (6)). As systems (1) are quadratic (i.e. there exists at least one quadratic term) then via an affine transformation systems (1) can be brought to the systems (see for example, [10])

$$\dot{x} = a + cx + dy + x^2, \quad \dot{y} = b + xy. \tag{16}$$

Straightforward calculations show us that for these systems $U_1 = U_2 = 0$. Moreover, the GL-invariant U_2 vanishes after any translation $(x, y) \mapsto (\tilde{x} + x_0, \tilde{y} + y_0)$ with arbitrary $(x_0, y_0) \in \mathbb{R}^2$ applied to systems (16). So according to (5) the condition $\mathcal{U} = 0$ is equivalent to $Af_5 = 0$. Since for systems (16) we have $Af_5 = -5d/4$ then we obtain $d = 0$ and for these systems we calculate:

$$W_2 = (c + 3x)(a + cx + x^2)(bc + bx - ay). \tag{17}$$

On the other hand for the matrix M_1 corresponding to systems (16) with $d = 0$ we have $\Delta^{1,2,4,5,6}_{2,3,5,7,12} = -2b$ and $\Delta^{1,2,3,5,6}_{2,5,6,7,12} = a$. It is clear that if $W_2 \neq 0$ then $rank(M_1) = 5$. Moreover straightforward calculations show that for $a = b = 0$ (i.e. when $W_2 = 0$) all the minors of order 5 vanish and hence the conditions *(iv)* of the lemma are proved. This completes the proof of Lemma 1. $\square$

Lemma 2. *The rank of the matrix $\mathcal{M}$ is equal to four if and only if $\beta = 0$, $\mathcal{U} = 0$ and one of the following four sets of conditions holds:*

(i) $M \neq 0,\ \mathcal{V}_1 = 0$; *(ii)* $M = V_2 = 0,\ K_5 W_1 \neq 0$;
(iii) $M = W_1 = W_2 = 0,\ K_5 \mathcal{V}_2 \neq 0$; *(iv)* $K_5 = W_2 = 0$.

Proof. According to the hypothesis of the lemma we assume $\beta = 0$, $\mathcal{U} = 0$ and we shall consider step by step the sets of conditions *(i)*- *(iv)*.

Conditions (i). As it was proved earlier (see the proof of Lemma 1, page 33) when $\beta = 0$, $M \neq 0$ and $\mathcal{U} = 0$ systems (1) will be brought via an affine transformation to systems (8) for which the conditions (9) hold. Moreover, it was proved that the condition $\mathcal{V}_1 = 0$ (i.e. $\widehat{H} = Af_{21} = 0$) yields for systems (8) $m = f = c = a + b = 0$ (see page 34). So we get the family of systems:

$$\dot{x} = a + x^2, \quad \dot{y} = -a - x^2,$$

for which without loss of generality we may assume $a \in \{0, -1, 1\}$ due to the transformation $(x, y, t) \mapsto \big(|a|^{-1/2}x, |a|^{-1/2}y, |a|^{1/2}t\big)$ if $a \neq 0$.

It remains to observe that for the matrix M_1 corresponding to these systems all the minors of order 6 and 5 vanish and $\Delta^{1,2,3,5}_{5,7,8,10} = -2$. Thus the systems of this family could be located only on the orbit of dimension 4.

Conditions (ii). In this case the condition $V_2 = 0$ holds. Therefore according to Remark 2 when $M = 0$, $\mathcal{U} = 0$ and $K_5 W_1 \neq 0$ systems (1) could be brought via an affine transformation to systems (13), for which $f = c \neq 0$ and $a = 0$. In other words when the conditions *(ii)* of Lemma 2 are satisfied, then we get the family of systems

$$\dot{x} = cx, \quad \dot{y} = b + cy - x^2. \tag{18}$$

As $c \neq 0$ (since $W_1 \neq 0$) we may assume $b = 0$ and $c = 1$ due to the transformation $(x, y, t) \mapsto \left(x, \dfrac{1}{c}(y - b), \dfrac{t}{c}\right)$. It remains to note that all the minors of order 6 and 5 for the matrix M_1 corresponding to these systems vanish. On the other hand $\Delta^{2,4,5,6}_{1,2,7,10} = 1$, i.e. system (18) (with $b = 0$ and $c = 1$) is located on the orbit of dimension 4.

Conditions (iii). In this case the condition $W_2 = 0$. As $M = W_1 = 0$ and $K_5 \neq 0$ it was proved earlier (see the proof of Lemma 1, page 34) that in this case systems (1) will be brought via an affine transformation to systems (13) with $cf = 0$ (i.e. $W_1 = 0$, see (14)). We shall examine two subcases: $c = 0$ and $c \neq 0$.

a) Assume first $c = 0$. Then considering (15) the condition $W_2 = 0$ yields $a = 0$ and we get the systems

$$\dot{x} = 0, \quad \dot{y} = b + fy - x^2, \tag{19}$$

for which $D = -f^2x^3$, $U_3 = x^2(4bx^2 + f^2y^2)$. Moreover, for a system located in the orbit under the translation group action of a system (19), i.e. for systems of the form

$$\dot{x} = 0, \quad \dot{y} = b - x_0^2 + fy_0 - 2xx_0 + fy - x^2, \tag{20}$$

for the GL-comitant U_3 we have $U_3 = x^2(4bx^2 + f^2y^2) + 4fx^3(xy_0 - yx_0)$. We recall that by (5) the condition $\mathcal{V}_2 \neq 0$ is equivalent to $D^2 + U_3^2 \neq 0$.

We note that all the minors of order 5 for the matrix M_1 corresponding to these systems vanish. On the other hand $\Delta^{1,2,3,6}_{2,5,7,10} = -2f^2$. Hence, if $D \neq 0$ (i.e. $f \neq 0$) we obtain $rank(M_1) = 4$.

Assume $D = 0$. Then $f = 0$ and then for any point (x_0, y_0) for system (20) we have $U_3 = 4bx^4$, i.e. for these systems the GL-comitant U_3 is a CT-comitant.

On the other hand we calculate $\Delta^{1,2,4,5}_{2,5,7,10} = 4b$. It is clear that if $U_3 \neq 0$ (this implies $\mathcal{V}_2 \neq 0$) then $rank(M_1) = 4$. Moreover straightforward calculations show that for $f = b = 0$ (i.e. when $\mathcal{V}_2 = 0$) all the minors of order 4 vanish and hence the conditions *(iii)* of Lemma 2 are proved in this case.

It remains to note that without loss of generality we may assume either $f = 1$ and $b = 0$ if $f \neq 0$ (due to the change $(x, y, t) \mapsto (x, (y - b)/f, t/f)$) or $f = 0$ and $b \in \{0, \pm 1\}$ (due to the change $(x, y, t) \mapsto (|b|^{1/2}x, |b|y, t)$).

b) Supposing $c \neq 0$ the condition $W_1 = 0$ yields $f = 0$. Moreover we may assume $c = 1$ due to the change $(x, y, t) \mapsto (x, y/c, t/c)$. Then the condition $W_2 = 0$ (see case (β_2) from (15)) gives $b - a^2 = 0$, i.e. $b = a^2$. Therefore we get the family of systems

$$\dot{x} = a + x, \quad \dot{y} = a^2 - x^2, \tag{21}$$

for the respective matrix M_1 of which we have $\Delta^{2,3,4,5}_{1,5,7,10} = -1$, i.e. $rank(M_1) = 4$. It remains to note that we may assume $a = 0$ due to the affine transformation $\bar{x} = x + a$, $\bar{y} = -2ax + y$. We also observe that for the obtained system as well as for any system located on its orbit under the translation group action we have $U_3 = x^2y^2 \neq 0$.

Conditions (iv). As it was shown in the proof of Lemma 1 (see page 35) when $K_5 = \mathcal{U} = 0$ systems (1) will be brought via an affine transformation to systems (16) with $d = 0$. Moreover, if $W_2 = 0$ according to (17) we obtain $a = b = 0$. So we arrive at the systems

$$\dot{x} = cx + x^2, \quad \dot{y} = xy,$$

for which all the minors of order 5 of the corresponding matrix M_1 vanish. But for this matrix we have $\Delta^{1,2,5,6}_{3,5,7,8} = 1$. Thus the systems of this family could be located only on the orbit of dimension 4. It remains to note that we may assume $c \in \{0, 1\}$ due to the change $(x, y, t) \mapsto (cx, y, t/c)$ if $c \neq 0$. □

Lemma 3. *The rank of the matrix $\mathcal{M}$ is equal to three if and only if the following conditions hold: $M = W_2 = \mathcal{V}_2 = 0$, $K_5 \neq 0$.*

Proof. Necessity. Assume that a system (1) is located on the orbit of dimension 3. As it follows from the proof of Lemma 2 this system could be located on the orbit of dimension less than or equal to 3 if and only if $\mathcal{U} = 0$ and the conditions *(iii)* of Lemma 2 with $\mathcal{V}_2 = 0$ instead of $\mathcal{V}_2 \neq 0$ are fulfilled. Moreover in this case via an affine transformation we arrive at a system of the form (19) with $b = f = 0$. So we get the system

$$\dot{x} = 0, \quad \dot{y} = -x^2, \tag{22}$$

for which the conditions provided by Lemma 3 are verified.

Sufficiency. Assume that the hypothesis of Lemma 3 is fulfilled. As $M = 0$ and $K_5 \neq 0$ then there exists an affine transformation which will brought systems (1) to the form (7) with $\delta = 0$, i.e. to the systems:

$$\begin{aligned} \dot{x} &= a + cx + dy + 2mx^2 + 2hxy, \\ \dot{y} &= b + fy - x^2 + 2mxy + 2hy^2, \end{aligned} \tag{23}$$

for which $\beta = 0$. We claim that for these systems the condition $W_2 = 0$ implies $W_1 = \mathcal{U} = 0$. Indeed, for systems (23) calculations yield:

$$\text{Coefficient}[W_2, x^3y^3] = 16h^3, \quad \text{Coefficient}[W_2\big|_{\{h=0\}}, x^6] = 16m^3,$$
$$\text{Coefficient}[W_2\big|_{\{h=m=0\}}, y^3] = -d^3, \ \text{Coefficient}[W_2\big|_{\{h=m=d=0\}}, x^3] = cf(2c+f).$$

We remark that if $h = m = d = 0$ then for systems above we obtain

$$Af_5 = U_1 = U_2 = 0, \quad W_1 = 2cf$$

and considering (5) this leads to the identity $\mathcal{U} = 0$. We also observe, that $W_2 = 0$ yields $cf(2c+f) = 0$. If $cf = 0$ then evidently $W_1 = 0$. In the case $f = -2c$ (considering also the conditions $h = m = d = 0$) we calculate Coefficient$[W_2, xy] = 6c^4$. Thus we get $c = 0$ and again we obtain $W_1 = 0$. Our claim is proved.

So the hypothesis of Lemma 2 corresponding to the conditions *(iii)* is verified except the condition $\mathcal{V}_2 \neq 0$. According to Lemma 2 in this case we have $rank(M_1) \leq 4$ and we obtain the equality if and only if $\mathcal{V}_2 \neq 0$.

Suppose now $\mathcal{V}_2 = 0$. As it was proved above the condition $W_2 = 0$ implies $h = m = d = cf = 0$ and we get two possibilities: $c = 0$ and $c \neq 0$. As it was shown in the proof of Lemma 2 (see page 36) in the case $c \neq 0$ the CT-comitant U_3 (and hence, $\mathcal{V}_2$) could not vanish. So the condition $c = 0$ has to be fulfilled and then we arrive at the systems (19) for which the condition $\mathcal{V}_2 = 0$ yields $b = f = 0$. Thus we get the system (22) for the corresponding matrix $\mathcal{M}$ of which we have found $rank(\mathcal{M}) = 3$ (see above). This completes the proof of Lemma 3. $\square$

In order to formulate and prove the Main Theorem we need some more invariant polynomials constructed in [11] as follows (we keep the respective notations).

We consider the differential operator $\mathcal{L} = x \cdot \mathbf{L}_2 - y \cdot \mathbf{L}_1$ acting on $\mathbb{R}[a, x, y]$ constructed in [13], where

$$\begin{aligned} \mathbf{L}_1 &= 2a_{00}\frac{\partial}{\partial a_{10}} + a_{10}\frac{\partial}{\partial a_{20}} + \frac{1}{2}a_{01}\frac{\partial}{\partial a_{11}} + 2b_{00}\frac{\partial}{\partial b_{10}} + b_{10}\frac{\partial}{\partial b_{20}} + \frac{1}{2}b_{01}\frac{\partial}{\partial b_{11}}, \\ \mathbf{L}_2 &= 2a_{00}\frac{\partial}{\partial a_{01}} + a_{01}\frac{\partial}{\partial a_{02}} + \frac{1}{2}a_{10}\frac{\partial}{\partial a_{11}} + 2b_{00}\frac{\partial}{\partial b_{01}} + b_{01}\frac{\partial}{\partial b_{02}} + \frac{1}{2}b_{10}\frac{\partial}{\partial b_{11}}. \end{aligned}$$

Then setting $\mu_0(a) = \text{Res}_{\,x}(p_2, q_2)/y^4$ we construct the following polynomials:

$$
\begin{aligned}
\mu_i(a,x,y) &= \frac{1}{i!}\mathcal{L}^{(i)}(\mu_0),\ i=1,..,4; \quad \kappa(a) = (M,K)^{(2)}/4; \quad \kappa_1(a) = (M,C_1)^{(2)};\\
L(a,x,y) &= 4K(a,x,y) + 8H(a,x,y) - M(a,x,y);\\
R(a,x,y) &= L(a,x,y) + 8K(a,x,y);\\
K_1(a,x,y) &= p_1(x,y)q_2(x,y) - p_2(x,y)q_1(x,y);\\
K_2(a,x,y) &= 4\,Jacob(J_2,\xi) + 3\,Jacob(C_1,\xi)D_1 - \xi(16J_1 + 3J_3 + 3D_1^2);\\
K_3(a,x,y) &= 2C_2^2(2J_1 - 3J_3) + C_2(3C_0K - 2C_1J_4) + 2K_1(3K_1 - C_1D_2),
\end{aligned}
$$

where $\mathcal{L}^{(i)}(\mu_0) = \mathcal{L}(\mathcal{L}^{(i-1)}(\mu_0))$ and

$$
J_1 = Jacob(C_0,D_2),\ J_2 = Jacob(C_0,C_2),\ J_3 = Discrim(C_1),
$$
$$
J_4 = Jacob(C_1,D_2),\ \xi = M - 2K.
$$

To distinguish topologically different phase portraits we also need the following invariant polynomials (constructed also in [11]):

$$
\begin{aligned}
B_3(\tilde{a},x,y) &= (C_2,D)^{(1)} = Jacob\,(C_2,D)\,,\\
B_2(\tilde{a},x,y) &= (B_3,B_3)^{(2)} - 6B_3(C_2,D)^{(3)},\\
B_1(\tilde{a}) &= \text{Res}_{\,x}\,(C_2,D)\,/y^9 = -2^{-9}3^{-8}\,(B_2,B_3)^{(4)}\,,\\
H(\tilde{a},x,y) &= (T_8 - 8T_9 - 2D_2^2)/18\ (= -4\widehat{H}(\tilde{a},x,y));\\
N(\tilde{a},x,y) &= K(\tilde{a},x,y) + H(\tilde{a},x,y);\\
\theta(\tilde{a}) &= \text{Discrim}\big(N(\tilde{a},x,y)\big);\\
H_1(\tilde{a}) &= -\big((C_2,C_2)^{(2)},C_2)^{(1)},D\big)^{(3)};\\
H_2(\tilde{a},x,y) &= (C_1,\ 2H - N)^{(1)} - 2D_1N;\\
H_3(\tilde{a},x,y) &= (C_2,D)^{(2)};\\
H_4(\tilde{a}) &= \big((C_2,D)^{(2)},(C_2,D_2)^{(1)}\big)^{(2)};\\
H_5(\tilde{a}) &= \big((C_2,C_2)^{(2)},(D,D)^{(2)}\big)^{(2)} + 8\big((C_2,D)^{(2)},(D,D_2)^{(1)}\big)^{(2)};\\
H_6(\tilde{a},x,y) &= 16N^2(C_2,D)^{(2)} + H_2^2(C_2,C_2)^{(2)};\\
H_7(\tilde{a}) &= (N,C_1)^{(2)};\\
H_8(\tilde{a}) &= 9\big((C_2,D)^{(2)},(D,D_2)^{(1)}\big)^{(2)} + 2\Big[(C_2,D)^{(3)}\Big]^2;\\
H_9(\tilde{a}) &= -\Big(\big((D,D)^{(2)},D\big)^{(1)}D\Big)^{(3)};\\
H_{10}(\tilde{a}) &= \big((N,D)^{(2)},\ D_2\big)^{(1)};\\
H_{11}(\tilde{a},x,y) &= 8H\big[(C_2,D)^{(2)} + 8(D,D_2)^{(1)}\big] + 3H_2^2;
\end{aligned}
$$

$$\begin{aligned}
N_1(\tilde{a},x,y) &= C_1(C_2,C_2)^{(2)} - 2C_2(C_1,C_2)^{(2)},\\
N_2(\tilde{a},x,y) &= D_1(C_1,C_2)^{(2)} - \big((C_2,C_2)^{(2)},C_0\big)^{(1)},\\
N_3(\tilde{a},x,y) &= (C_2,C_1)^{(1)},\\
N_4(\tilde{a},x,y) &= 4\,(C_2,C_0)^{(1)} - 3C_1D_1,\\
N_5(\tilde{a},x,y) &= \big[(D_2,C_1)^{(1)} + D_1D_2\big]^2 - 4\big(C_2,C_2\big)^{(2)}\big(C_0,D_2\big)^{(1)},\\
N_6(\tilde{a},x,y) &= 8D + C_2\left[8(C_0,D_2)^{(1)} - 3(C_1,C_1)^{(2)} + 2D_1^2\right].
\end{aligned}$$

Some important geometric propriety of the constructed above polynomials $\mu_i(\tilde{a},x,y)$ $(i = 0,1,\dots,4$ is revealed by the next lemma proved in [13].

Lemma 4 ([13]). *A system* (1) *is degenerate (i.e.* $\gcd(P,Q) \neq 1$*) if and only if* $\mu_i = 0$ *for all* $i = 0,1,..,4$.

Main Theorem *(i) A system* (1) *is located on an affine orbit of the given above dimension if and only if one of the respective sets of the conditions holds:*

$$\begin{aligned}
&6 \;\Leftrightarrow\; \beta \neq 0 \quad \textit{or} \quad \beta = 0 \quad \textit{and} \quad \mathcal{U} \neq 0;\\
&5 \;\Leftrightarrow\; \beta = 0,\ \mathcal{U} = 0 \ \textit{and either} \left\{\begin{array}{l} M\mathcal{V}_1 \neq 0,\ \textit{or}\\ M = 0,\ K_5W_1V_2 \neq 0,\ \textit{or}\\ M = W_1 = 0,\ K_5W_2 \neq 0,\ \textit{or}\\ K_5 = 0,\ W_2 \neq 0;\end{array}\right.\\
&4 \;\Leftrightarrow\; \beta = 0,\ \mathcal{U} = 0 \ \textit{and either} \left\{\begin{array}{l} M \neq 0,\ \mathcal{V}_1 = 0,\ \textit{or}\\ M = V_2 = 0,\ K_5W_1 \neq 0,\ \textit{or}\\ M = W_1 = W_2 = 0,\ K_5\mathcal{V}_2 \neq 0,\ \textit{or}\\ K_5 = W_2 = 0;\end{array}\right.\\
&3 \;\Leftrightarrow\; M = W_2 = \mathcal{V}_2 = 0,\ K_5 \neq 0.
\end{aligned}$$

(ii) Assume that a quadratic system is located on the affine orbit of the dimension less than or equal to 5. Then the phase portrait of this system is topologically equivalent to one of the 49 topologically distinct phase portraits given in Fig. 1. Moreover in Table 1 we give necessary and sufficient conditions, invariant with respect to the action of the affine group and time rescaling, for the realization of each one of the phase portraits corresponding to a system located on an orbit of the given dimension (≤ 5)*. The first column of Table 1 contains dimension of the orbit. In the second column we list the necessary and sufficient affine invariant conditions for a system to be located on the orbit of the respective dimension. In the third column the additional conditions needed for the realization of the corresponding phase portrait in the last column are listed.*

Proof. The proof of the statement *(i)* of Main Theorem follows immediately from Proposition 1 and Lemmas 1 – 3. So we shall concentrate our attention on the proof of the statement *(ii)*.

Table 1

$\mathfrak{D}$	*Necessary and sufficient conditions*	*Additional conditions for phase portraits*				*Phase portrait*
5	$\beta=0,\ \mathcal{U}=0,$ $M\mathcal{V}_1\neq 0$	$H\neq 0$	$H_6\neq 0$	$K\neq 0$	$H_{11}>0,\mu_2>0,L>0$	$\mathcal{P}_1$
					$H_{11}>0,\mu_2>0,L<0$	$\mathcal{P}_2$
					$H_{11}>0,\mu_2<0,K<0$	$\mathcal{P}_3$
					$H_{11}>0,\ \mu_2<0,$ $K>0,\ L>0$	$\mathcal{P}_4$
					$H_{11}>0,\ \mu_2<0,$ $K>0,\ L<0$	$\mathcal{P}_5$
					$H_{11}>0,\mu_2=0,K<0$	$\mathcal{P}_6$
					$H_{11}>0,\ \mu_2=0,$ $K>0,\ L<0$	$\mathcal{P}_7$
					$H_{11}>0,\ \mu_2=0,$ $K>0,\ L>0$	$\mathcal{P}_8$
					$H_{11}<0,\ L>0$	$\mathcal{P}_9$
					$H_{11}<0,\ L<0$	$\mathcal{P}_{10}$
					$H_{11}=0,\ L>0$	$\mathcal{P}_{11}$
					$H_{11}=0,\ L<0$	$\mathcal{P}_{12}$
				$K=0$	$H_{11}\neq 0,\ K_1\mu_3<0$	$\mathcal{P}_{13}$
					$H_{11}\neq 0,\ K_1\mu_3>0$	$\mathcal{P}_{14}$
					$H_{11}\neq 0,\ \mu_3=0$	$\mathcal{P}_{15}$
					$H_{11}=0,\ \varkappa_2<0$	$\mathcal{P}_9$
					$H_{11}=0,\ \varkappa_2>0$	$\mathcal{P}_{10}$
			$H_6=0$	$N\neq 0$	$H_{11}\neq 0,\ L>0$	$\mathcal{P}_1$
					$H_{11}\neq 0,\ L<0$	$\mathcal{P}_5$
					$H_{11}\neq 0,\ L=0$	$\mathcal{P}_{14}$
					$H_{11}=0,\ L<0$	$\mathcal{P}_{16}$
					$H_{11}=0,L>0,K<0$	$\mathcal{P}_{17}$
					$H_{11}=0,L>0,K>0$	$\mathcal{P}_{18}$
					$H_{11}=0,\ L=0$	$\mathcal{P}_{19}$
				$N=0$	$H_3>0$	$\mathcal{P}_5$
					$H_3<0$	$\mathcal{P}_{10}$
					$H_3=0$	$\mathcal{P}_{16}$

Table 1 *(continued)*

<table>
<tr><th>$\mathfrak{D}$</th><th>Necessary and sufficient conditions</th><th colspan="3">Additional conditions for phase portraits</th><th>Phase portrait</th></tr>
<tr><td rowspan="28">5</td><td rowspan="10">$\beta = 0,\ \mathcal{U} = 0,$
$M\mathcal{V}_1 \neq 0$</td><td rowspan="10">$H = 0$</td><td rowspan="3">$D \neq 0$</td><td>$N_5 > 0$</td><td>$\mathcal{P}_1$</td></tr>
<tr><td>$N_5 < 0$</td><td>$\mathcal{P}_9$</td></tr>
<tr><td>$N_5 = 0$</td><td>$\mathcal{P}_{11}$</td></tr>
<tr><td rowspan="7">$D = 0$</td><td>$N_5 < 0$</td><td>$\mathcal{P}_9$</td></tr>
<tr><td>$N_5 > 0,\ \mu_4 > 0$</td><td>$\mathcal{P}_{20}$</td></tr>
<tr><td>$N_5 > 0,\ \mu_4 < 0$</td><td>$\mathcal{P}_{21}$</td></tr>
<tr><td>$N_5 > 0,\ \mu_4 = 0,\ N_1 \neq 0$</td><td>$\mathcal{P}_{22}$</td></tr>
<tr><td>$N_5 > 0,\ \mu_4 = 0,\ N_1 = 0$</td><td>$\mathcal{P}_{23}$</td></tr>
<tr><td>$N_5 = 0,\ \mu_4 \neq 0$</td><td>$\mathcal{P}_{24}$</td></tr>
<tr><td>$N_5 = 0,\ \mu_4 = 0$</td><td>$\mathcal{P}_{25}$</td></tr>
<tr><td rowspan="9">$M = \mathcal{U} = 0,$
$K_5W_1V_2 \neq 0$</td><td rowspan="6">$H \neq 0$</td><td rowspan="3">$H_{11} \neq 0$</td><td>$K_3 > 0$</td><td>$\mathcal{P}_{26}$</td></tr>
<tr><td>$K_3 < 0$</td><td>$\mathcal{P}_{27}$</td></tr>
<tr><td>$K_3 = 0$</td><td>$\mathcal{P}_{28}$</td></tr>
<tr><td rowspan="3">$H_{11} = 0$</td><td>$K_3 > 0$</td><td>$\mathcal{P}_{29}$</td></tr>
<tr><td>$K_3 < 0$</td><td>$\mathcal{P}_{30}$</td></tr>
<tr><td>$K_3 = 0$</td><td>$\mathcal{P}_{31}$</td></tr>
<tr><td rowspan="3">$H = 0$</td><td colspan="2">$\mu_3K_1 > 0,\ K_3 \geq 0$</td><td>$\mathcal{P}_{32}$</td></tr>
<tr><td colspan="2">$\mu_3K_1 > 0,\ K_3 < 0$</td><td>$\mathcal{P}_{33}$</td></tr>
<tr><td colspan="2">$\mu_3K_1 < 0$</td><td>$\mathcal{P}_{34}$</td></tr>
<tr><td rowspan="3">$M = \mathcal{U} = 0,$
$W_1 = 0, K_5W_2 \neq 0$</td><td colspan="3">$D \neq 0$</td><td>$\mathcal{P}_{35}$</td></tr>
<tr><td rowspan="2">$D = 0$</td><td colspan="2">$\mu_4 < 0$</td><td>$\mathcal{P}_{36}$</td></tr>
<tr><td colspan="2">$\mu_4 > 0$</td><td>$\mathcal{P}_{29}$</td></tr>
<tr><td rowspan="6">$K_5 = 0,\ W_2 \neq 0$</td><td colspan="3">$H_{11} < 0$</td><td>$\mathcal{P}_{37}$</td></tr>
<tr><td rowspan="3">$H_{11} > 0$</td><td colspan="2">$\mu_2 > 0$</td><td>$\mathcal{P}_{38}$</td></tr>
<tr><td colspan="2">$\mu_2 < 0$</td><td>$\mathcal{P}_{39}$</td></tr>
<tr><td colspan="2">$\mu_2 = 0$</td><td>$\mathcal{P}_{40}$</td></tr>
<tr><td rowspan="2">$H_{11} = 0$</td><td colspan="2">$\mu_2 \neq 0$</td><td>$\mathcal{P}_{41}$</td></tr>
<tr><td colspan="2">$\mu_2 = 0$</td><td>$\mathcal{P}_{42}$</td></tr>
</table>

Table 1 *(continued)*

$\mathfrak{D}$	*Necessary and sufficient conditions*	*Additional conditions for phase portraits*	*Phase portrait*
4	$\beta=0,\ \mathcal{U}=0,$ $M\neq0, \mathcal{V}_1=0$	$N_5>0$	$\mathcal{P}_{23}$
		$N_5<0$	$\mathcal{P}_{9}$
		$N_5=0$	$\mathcal{P}_{43}$
	$M=\mathcal{U}=V_2=0, K_5W_1\neq0$	–	$\mathcal{P}_{32}$
	$M=\mathcal{U}=W_1=0,$ $W_2=0, K_5\mathcal{V}_2\neq0$	$D\neq0$	$\mathcal{P}_{44}$
		$D=0$: $N_6\neq0$	$\mathcal{P}_{45}$
		$D=0$: $N_6=0,\ U_3<0$	$\mathcal{P}_{29}$
		$D=0$: $N_6=0,\ U_3>0$	$\mathcal{P}_{46}$
	$K_5=0,\ W_2=0$	$H_{11}\neq0$	$\mathcal{P}_{47}$
		$H_{11}=0$	$\mathcal{P}_{48}$
3	$M=W_2=\mathcal{V}_2=0, K_5\neq0$	–	$\mathcal{P}_{49}$

In other words we assume that a quadratic system is located on an affine orbit of dimension ≤ 5 and we shall determine the phase portrait of this system as well as the respective affine invariant conditions for its realization.

According to Lemmas 1 – 3 for a system located on an orbit of the dimension ≤ 5 the conditions $\beta=0$ and $\mathcal{U}=0$ have to be fulfilled and in what follows we assume that these conditions hold.

1) The case $M\neq0$. In this case via an affine transformation a quadratic system (1) could be brought to the form (8) (see page 33), i.e.

$$\dot{x}=a+cx+(2m+1)x^2,\quad \dot{y}=b+fy-x^2+2mxy, \tag{24}$$

for which the condition $cf-f^2-2am-2bm=0$ holds. For these systems we have $H=-4m^2x^2$.

a) The subcase $H\neq0$. Then $m\neq0$ and we may assume $f=0$ due to the translation $(x,y)\mapsto\left(x-\frac{f}{2m}, y-\frac{f}{2m^2}\right)$. Then for these systems the condition above yields $m(a+b)=0$ and as $m\neq0$ we get $b=-a$. Thus we obtain the systems

$$\dot{x}=a+cx+(2m+1)x^2,\quad \dot{y}=-a-x^2+2mxy, \tag{25}$$

for which we calculate the needed invariant polynomials applied in [11] (keeping the respective notations):

$$\begin{gathered} B_3=\theta=\eta=\mu_0=H_7=0,\quad H_6=-2048m^4(-4a+c^2-8am-4am^2)x^6,\\ H_{11}=768m^4(-4a+c^2-8am)x^4,\quad H=-4m^2x^2,\quad K=4m(1+2m)x^2,\\ L=8(1+2m)x^2,\quad \mu_2=4am^2(1+2m)x^2,\quad N=4m(1+m)x^2. \end{gathered} \tag{26}$$

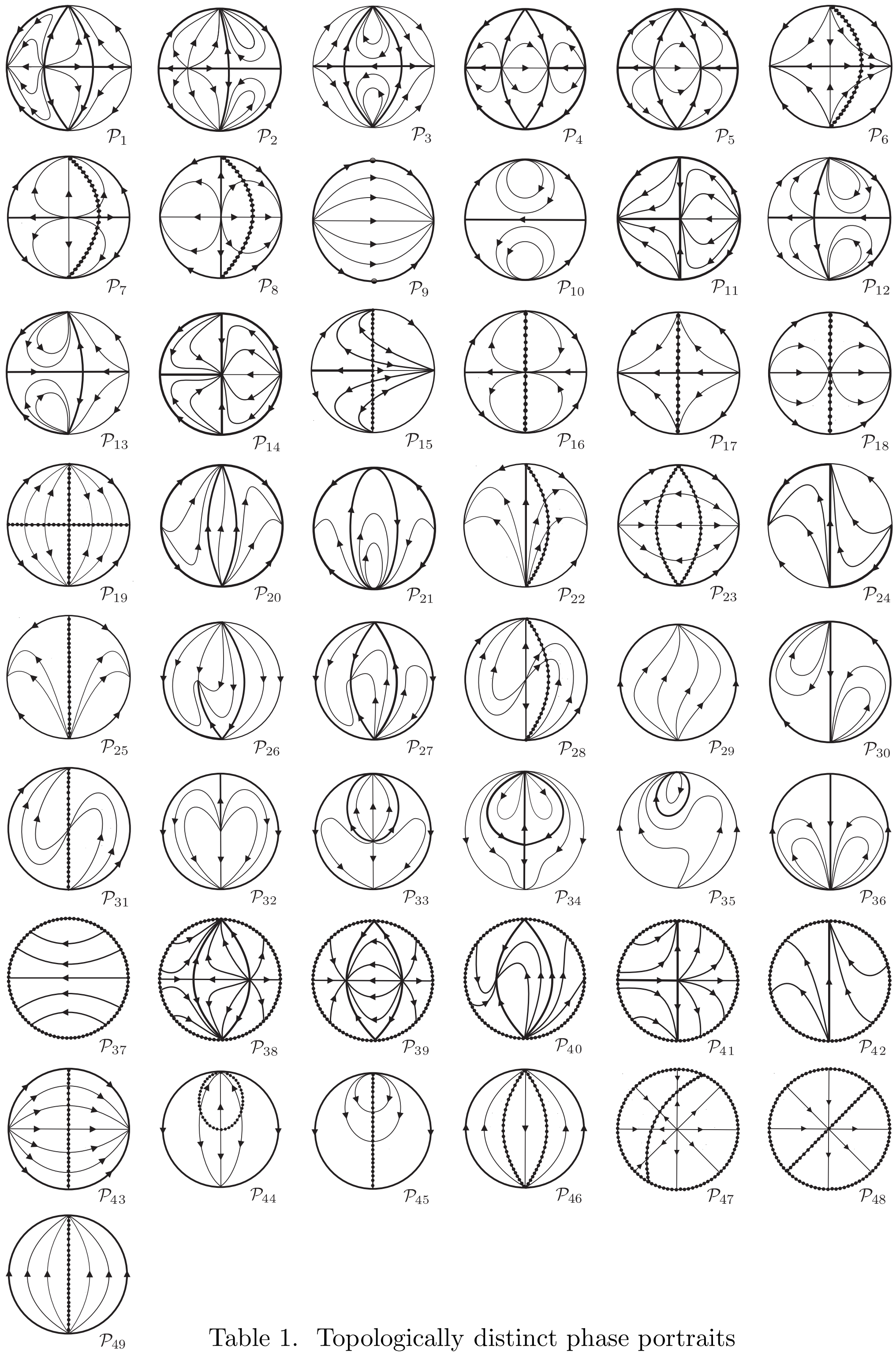


Table 1. Topologically distinct phase portraits

As $H = -4\widehat{H}$ by (5) the condition $H \neq 0$ implies $\mathcal{V}_1 \neq 0$, i.e. for $m \neq 0$ a system (25) is located on the orbit of dimension 5.

On the other hand as $\beta = 0$ (which is equivalent to $\eta = 0$ from [11]), $M \neq 0$ and $B_3 = \theta = \mu_0 = H_7 = 0$, according to [8] and [9] the family of non-degenerate systems (25) possesses invariant lines (considered with multiplicity and including the line at infinity) of total multiplicity four if $H_6 \neq 0$ and at least five if $H_6 = 0$.

a_1) *The possibility* $H_6 \neq 0$. According to [11] for the non-degenerate systems (25) we get the following phase portraits (we keep the respective notations from [11]):

Picture 4.12(a) $[\mathcal{P}_1]$ if $K \neq 0,\ H_{11} > 0,\ \mu_2 > 0,\ L > 0$;
Picture 4.12(b) $[\mathcal{P}_2]$ if $K \neq 0,\ H_{11} > 0,\ \mu_2 > 0,\ L < 0$;
Picture 4.12(c) $[\mathcal{P}_3]$ if $K \neq 0,\ H_{11} > 0,\ \mu_2 < 0,\ K < 0$;
Picture 4.12(d) $[\mathcal{P}_4]$ if $K \neq 0,\ H_{11} > 0,\ \mu_2 < 0,\ K > 0,\ L > 0$;
Picture 4.12(e) $[\mathcal{P}_5]$ if $K \neq 0,\ H_{11} > 0,\ \mu_2 < 0,\ K > 0,\ L < 0$;
Picture 4.15(a) $[\mathcal{P}_9]$ if $K \neq 0,\ H_{11} < 0,\ L > 0$;
Picture 4.15(b) $[\mathcal{P}_{10}]$ if $K \neq 0,\ H_{11} < 0,\ L < 0$;
Picture 4.24(a) $[\mathcal{P}_{11}]$ if $K \neq 0,\ H_{11} = 0,\ L > 0$;
Picture 4.24(b) $[\mathcal{P}_{12}]$ if $K \neq 0,\ H_{11} = 0,\ L < 0$;
Picture 4.19(a) $[\mathcal{P}_{13}]$ if $K = 0,\ N \neq 0,\ H_{11} \neq 0,\ K_1\mu_3 < 0$;
Picture 4.19(b) $[\mathcal{P}_{14}]$ if $K = 0,\ N \neq 0,\ H_{11} \neq 0,\ K_1\mu_3 > 0$;
Picture 4.36(a) $[\mathcal{P}_9]$ if $K = 0,\ N \neq 0,\ H_{11} = 0,\ \varkappa_2 < 0$;
Picture 4.36(b) $[\mathcal{P}_{10}]$ if $K = 0,\ N \neq 0,\ H_{11} = 0,\ \varkappa_2 > 0$.

Here in square brackets the respective phase portraits from Figure 1 which are topologically equivalent to those from [11], respectively are indicated. As by (26) the conditions $H \neq 0$ and $K = 0$ imply $N \neq 0$ we arrive at the respective conditions from Table 1.

It remains to look for degenerate systems which could belong to the family (25) when the conditions $H \neq 0$ (i.e. $m \neq 0$) and $H_6 \neq 0$ (i.e. $c^2 - 4a(1+m)^2 \neq 0$) hold. According to Lemma 4 we calculate the polynomials μ_i for this family:

$$\mu_0 = \mu_1 = 0, \quad \mu_2 = 4am^2(1+2m)x^2,\ \ \mu_3 = 4acmx^2(x + mx - my),$$
$$\mu_4 = ax^2[(c^2 + 4am^2)x^2 - 2m(c^2 - 4am)xy + 4am^2y^2].$$

Evidently the conditions $\mu_i = 0$ $(i = 0, 1, .., 4)$ (see Lemma 4) are equivalent to $a = 0$.

Assume first $K \neq 0$. Considering (26) we have $m(1 + 2m) \neq 0$ and hence the condition $a = 0$ is equivalent to $\mu_2 = 0$. So we get the family of degenerate systems

$$\dot{x} = x[c + (2m + 1)x], \quad \dot{y} = x(-x + 2my), \tag{27}$$

for which $H_6 = -2048c^2m^4x^6 \neq 0$. For the respective family of linear systems

$$\dot{x} = c + (2m + 1)x, \quad \dot{y} = -x + 2my$$

we have $\lambda_1 = 2m + 1$, $\lambda_2 = 2m$ and therefore $\operatorname{sign}(\lambda_1\lambda_2) = \operatorname{sign}(K)$. Moreover by (26) we have $\operatorname{sign}(\lambda_1) = \operatorname{sign}(L)$.

We observe that due to $H_6 \neq 0$ (i.e. $c \neq 0$) the invariant line $c + (2m+1)x = 0$ does not coincide with line $x = 0$ (filled with singularities).

Thus after some standard investigations we decide that the phase portrait of a degenerate system (27) corresponds to picture $\mathcal{P}_6$ if $K < 0$; $\mathcal{P}_7$ if $K > 0$ and $L < 0$ and to picture $\mathcal{P}_8$ if $K > 0$ and $L > 0$.

Suppose now $K = 0$. As $H \neq 0$ (i.e. $m \neq 0$) considering (26) we have $m = -1/2$ and then systems (25) become

$$\dot{x} = a + cx, \quad \dot{y} = -a - x^2 - xy, \tag{28}$$

for which we calculate:

$$\mu_2 = 0, \quad \mu_3 = -acx^2(x+y), \quad H_{11} = 48c^2x^4, \quad H_6 = 128(a - c^2)x^6.$$

We observe that the systems above could be degenerate (i.e. $a = 0$) only if $H_{11} \neq 0$ as $H_{11} = 0$ gives $c = 0$ and then $H_6 = 128ax^6 \neq 0$. On the other hand if $H_{11} \neq 0$ then the condition $a = 0$ is equivalent to $\mu_3 = 0$.

So, setting $a = 0$ in systems (28) we may assume $c = 1$ (due to the rescaling $(x, y, t) \mapsto (cx, cy, t/c)$) and we easily get the phase portrait $\mathcal{P}_{15}$.

a_2) *The possibility* $H_6 = 0$. Then we obtain $c^2 - 4a(1+m)^2 = 0$ and we need to distinguish 2 cases: $N \neq 0$ and $N = 0$.

If $N \neq 0$ then by (26) we have $m+1 \neq 0$ and therefore we get $a = c^2/(4(1+m)^2)$. We observe that due to $H \neq 0$ by (26) the condition $K = 0$ is equivalent to $L = 0$. So, according to [12] the phase portrait of such a system corresponds to *Picture 5.14(a)* [$\mathcal{P}_1$] if $L > 0$; *Picture 5.14(b)* [$\mathcal{P}_5$] if $L < 0$ and to *Picture 5.18* [$\mathcal{P}_{14}$] if $L = 0$.

On the other hand for systems (25) we have

$$\mu_3 = c^3mx^2[(1+m)x - my]/(1+m)^2, \quad H_{11} = 768c^2m^6x^4/(1+m)^2.$$

Therefore according to Lemma 4 the necessary condition $\mu_3 = 0$ for a system to be degenerate yields $c = 0$ and this condition is given by $H_{11} = 0$. In this case evidently systems (25) (with $c = a = 0$) become degenerate systems

$$\dot{x} = (2m+1)x^2, \quad \dot{y} = x(-x + 2my),$$

which are a subfamily of systems (27) (corresponding to $c = 0$). So in the case $2m+1 \neq 0$ (i.e. $L \neq 0$) the singular invariant line $x = 0$ coincides with the invariant line of the respective linear systems and hence we get the phase portrait $\mathcal{P}_{16}$ if $L < 0$; $\mathcal{P}_{17}$ if $L > 0$ and $K < 0$ and $\mathcal{P}_{18}$ if $L > 0$ and $K > 0$. If $L = 0$ we have $m = -1/2$ and we get two singular lines. This evidently leads to picture $\mathcal{P}_{19}$.

Assume $N = 0$. As $H \neq 0$ and $H_6 = 0$ by (26) we get $m + 1 = c = 0$. In this case for systems (25) we have

$$H_6 = H_2 = 0, \quad H_3 = 32ax^2$$

and according to [8] systems (25) possess invariant line of total multiplicity 6.

On the other hand in the case $H_3 \neq 0$ (i.e. $a \neq 0$ and systems are non-degenerate) according to [12] the phase portrait corresponds to *Picture 6.8* $[\mathcal{P}_5]$ if $H_3 > 0$ and to *Picture 6.9* $[\mathcal{P}_{10}]$ if $H_3 < 0$.

Assuming $H_3 = 0$ (i.e. $a = 0$) we get the degenerate system

$$\dot{x} = -x^2, \quad \dot{y} = -x(x + 2y),$$

the phase portrait of which corresponds to $[\mathcal{P}_{16}]$.

b) The subcase $H = 0$. Then $m = 0$ and considering (24) we get the family of systems

$$\dot{x} = a + cx + x^2, \quad \dot{y} = b + fy - x^2, \tag{29}$$

for which the condition $(c - f)f = 0$ must hold. For these systems we calculate:

$$B_3 = \theta = \mu_0 = N = K = H = 0, \quad D = -f^2x^2(x + y),$$
$$N_1 = 8(c - f)x^4, \quad N_2 = 4(4a - c^2 + f^2)x, \quad N_5 = -16(4a - c^2)x^2,$$
$$Af_{21} = (a + cx + x^2)^2(a + b + cx + fy).$$

So according to [9] and [8] these systems possess invariant line of total multiplicity at least 4.

As $H = 0$ the condition $\mathcal{V}_1 = 0$ is equivalent to $Af_{21} = 0$, i.e. a system (29) is located on the orbit of dimension four if and only if $a + b = c = f = 0$. In this case we obtain the degenerate system

$$\dot{x} = a + x^2, \quad \dot{y} = -(a + x^2),$$

where $a \in \{-1, 0, 1\}$ due to the rescaling $(x, y, t) \mapsto (|a|^{-1/2}x, |a|^{-1/2}y, |a|^{1/2}t)$ if $a \neq 0$. For these systems $N_5 = -64ax^2$ and we obviously obtain the phase portrait $\mathcal{P}_{23}$ (respectively $\mathcal{P}_9$; $\mathcal{P}_{43}$) if $N_5 > 0$ (respectively $N_5 < 0$; $N_5 = 0$).

Assuming $\mathcal{V}_1 \neq 0$ we shall consider two possibilities: $D \neq 0$ and $D = 0$.

b_1) The possibility $D \neq 0$. In this case $f \neq 0$ and then for systems (29) we obtain $f = c \neq 0$. Then we may consider $b = 0$ and $c = 1$ due to the transformation $(x, y, t) \mapsto (cx, (c^2y - b)/c, t/c)$. So we arrive at the family of systems

$$\dot{x} = a + x + x^2, \quad \dot{y} = y - x^2,$$

for which we calculate: $N_1 = 0$, $N_2 = 16ax$, $N_5 = 16(1 - 4a)x^2$. So according to [8] these systems possess invariant line of total multiplicity at least 5. Moreover following [12] for non-degenerate systems we get the phase portraits

- *Picture 5.13* $[\mathcal{P}_1]$ if $N_2 \neq 0$, $N_5 > 0$;
- *Picture 5.15* $[\mathcal{P}_9]$ if $N_2 \neq 0$, $N_5 < 0$;
- *Picture 5.17* $[\mathcal{P}_{11}]$ if $N_2 \neq 0$, $N_5 = 0$;
- *Picture 6.7* $[\mathcal{P}_1]$ if $N_2 = 0$.

We observe that the condition $N_2 = 0$ implies $N_5 > 0$ and that the systems above could not be degenerate due to $\mu_2 = x^2 \neq 0$ (see Lemma 4).

b_2) *The possibility* $D = 0$ In this case $f = 0$ and we arrive at the family of systems

$$\dot{x} = a + cx + x^2, \quad \dot{y} = b - x^2, \tag{30}$$

for which we calculate:

$$\begin{gathered} N_1 = 8cx^4,\ N_2 = 4(4a - c^2)x,\ N_5 = 16(c^2 - 4a)x^2,\ \mu_0 = \mu_1 = \mu_2 = \mu_3 = 0, \\ \mu_4 = [(a+b)^2 - bc^2]x^4,\ Af_{21} = (a + b + cx)(a + cx + x^2)^2. \end{gathered} \tag{31}$$

So according to [9] and [8] these systems possess invariant line of total multiplicity at least 4. On the other hand by [11] and [12] for non-degenerate systems we get the following phase portraits:

- *Picture 4.29(a)* $[\mathcal{P}_{20}]$ if $N_1 \neq 0,\ N_2 \neq 0,\ N_5 > 0,\ \mu_4 > 0$;
- *Picture 4.29(b)* $[\mathcal{P}_{21}]$ if $N_1 \neq 0,\ N_2 \neq 0,\ N_5 > 0,\ \mu_4 < 0$;
- *Picture 4.33* $[\mathcal{P}_9]$ if $N_1 \neq 0,\ N_2 \neq 0,\ N_5 < 0$;
- *Picture 5.28* $[\mathcal{P}_{24}]$ if $N_1 \neq 0,\ N_2 = 0$;
- *Picture 5.20* $[\mathcal{P}_{20}]$ if $N_1 = 0,\ N_2 \neq 0,\ N_5 > 0$;
- *Picture 5.24* $[\mathcal{P}_9]$ if $N_1 = 0,\ N_2 \neq 0,\ N_5 < 0$;
- *Picture 6.10* $[\mathcal{P}_{24}]$ if $N_1 = 0,\ N_2 = 0$.

It remains to determine the phase portraits for degenerate systems (30), i.e. by Lemma 4 the condition $(a+b)^2 - bc^2 = 0$ must hold. On the other hand the condition $Af_{21} \neq 0$ gives $(a + b)^2 + c^2 \neq 0$ and this implies $b \geq 0$.

If $b > 0$ then we may assume $b = 1$ due to the rescaling $(x, y, t) \mapsto (b^{-1/2}x, b^{-1/2}y, b^{1/2}t)$. Therefore we get $c = \pm(a + 1)$ and it is sufficient to consider only the case $c = a + 1$ due to the change $(x, y, t, c) \mapsto (-x, -y, -t, -c)$ which keeps systems (30). Thus we obtain the family of degenerate systems

$$\dot{x} = (1 + x)(a + x), \quad \dot{y} = 1 - x^2. \tag{32}$$

Taking into consideration the two invariant lines $x = -1$ (singular) and $x = a$ which could coincide if $a = 1$ as well as the critical value $a = -1$ (when the respective linear system is also degenerate) we arrive at phase portrait $\mathcal{P}_{22}$ if $a^2 - 1 \neq 0$; $\mathcal{P}_{25}$ if $a = 1$ and $\mathcal{P}_{23}$ if $a = -1$.

It remains to note that for systems (32) we have $\mu_4 = 0$, $N_1 = 8(1 + a)x^4$, $N_2 = -4(a - 1)^2x$ and $N_5 = 16(a - 1)^2x^2$.

If $b = 0$ then the condition $\mu_4 = 0$ gives $a = 0$ and then $Af_{21} = cx^3(c + x)^2 \neq 0$. Hence we can assume $c = 1$ due to the rescaling $(x, y, t) \mapsto (cx, cy, t/c)$. So we get the degenerate system

$$\dot{x} = x(1 + x), \quad \dot{y} = -x^2,$$

for which $N_1 = 8x^4$, $N_2 = -4x$ and $N_5 = 16x^2$.

Considering (31) and the case $\mu_4 = 0$ examined above we observe that the condition $N_5 < 0$ implies $\mu_4 \neq 0$ and $N_5 = 0$ if and only if $N_2 = 0$. Moreover the condition $N_1 = 0$ implies $\mu_4 \geq 0$.

Considering this observation we could unite the conditions for the realization of topologically distinct phase portraits in the considered case (including the degenerate systems) as follows:

- $\mathcal{P}_9$ if $N_5 < 0$;
- $\mathcal{P}_{20}$ if $N_5 > 0$, $\mu_4 > 0$;
- $\mathcal{P}_{21}$ if $N_5 > 0$, $\mu_4 < 0$;
- $\mathcal{P}_{22}$ if $N_5 > 0$, $\mu_4 = 0$, $N_1 \neq 0$;
- $\mathcal{P}_{23}$ if $N_5 > 0$, $\mu_4 = 0$, $N_1 = 0$;
- $\mathcal{P}_{24}$ if $N_5 = 0$, $\mu_4 \neq 0$;
- $\mathcal{P}_{25}$ if $N_5 = 0$, $\mu_4 = 0$.

Thus we arrive at the respective conditions from Table 1.

2) The case $M = 0$ and $K_5 \neq 0$. As it was shown in the proof of Lemma 1 (see page 34) in this case via an affine transformation a quadratic system (1) could be brought to the form (10), for which the condition $m(cf - f^2 - 2am) = 0$ holds. For these systems we have $H = -4m^2x^2$.

a) The subcase $H \neq 0$. Then $m \neq 0$ and we may assume $f = 0$ due to the translation $(x, y) \mapsto \left(x - \frac{f}{2m}, y - \frac{f}{2m^2}\right)$. Then for these systems the condition above gives $am = 0$ and as $m \neq 0$ we get $a = 0$. Then we arrive at the family of systems

$$\dot{x} = cx + 2mx^2, \quad \dot{y} = b - x^2 + 2mxy, \tag{33}$$

for which calculations yield:

$$\begin{gathered} B_3 = \theta = 0,\ N = 4m^2x^2,\ K_5 = x^3 \neq 0,\ H_{11} = 768c^2m^4x^4, \\ K_3 = -24bm^2x^6,\ D = 4bm^2x^3, \quad N_6 = 8(c^2 + 4bm^2)x^3, \\ \text{Coefficient}[V_2, x^6] = -8m^3, \quad W_1 = 4mx(c + 4mx). \end{gathered} \tag{34}$$

So according to [8] and [9] the family of non-degenerate systems (33) possesses invariant lines of total multiplicity at least four.

The condition $H \neq 0$ implies $V_2W_1 \neq 0$, i.e. for $m \neq 0$ a system (33) is located on the orbit of dimension 5. As for these systems we have

$$\mu_0 = \mu_1 = \mu_2 = 0, \quad \mu_3 = -4bcm^2x^3, \quad \mu_4 = bx^3[(4bm^2 - c^2)x + 2c^2my]$$

according to Lemma 4 a system (33) becomes degenerate if and only if $b = 0$ and this is equivalent to $K_3 = 0$.

As $M=0$, $B_3=\theta=0$ and $N\neq 0$, according to [11] and [12] for a non-degenerate system (33) we obtain the following phase portraits:

Picture 4.31(a) $[\mathcal{P}_{26}]$ $\Leftrightarrow$ $H_{11}\neq 0,\ N_6\neq 0,\ K_3>0;$
Picture 4.31(b) $[\mathcal{P}_{27}]$ $\Leftrightarrow$ $H_{11}\neq 0,\ N_6\neq 0,\ K_3<0;$
Picture 5.23 $[\mathcal{P}_{26}]$ $\Leftrightarrow$ $H_{11}\neq 0,\ N_6=0;$
Picture 4.44(a) $[\mathcal{P}_{29}]$ $\Leftrightarrow$ $H_{11}=0,\ K_3>0;$
Picture 4.44(b) $[\mathcal{P}_{30}]$ $\Leftrightarrow$ $H_{11}=0,\ K_3<0;$

Assume now that systems (33) are degenerate, i.e. $K_3=0$ (that implies $b=0$). As $x=0$ is an invariant line filled with singularities and the condition $c=0$ is equivalent to $H_{11}=0$, we obviously obtain the phase portrait $\mathcal{P}_{28}$ if $H_{11}\neq 0$ and $\mathcal{P}_{31}$ if $H_{11}=0$. Taking into account that the conditions $H_6=0$ and $H_{11}\neq 0$ by (34) imply $K_3>0$ an that the condition $K_3<0$ implies $H_6\neq 0$ we obviously arrive at the respective conditions from Table 1.

b) The subcase $H=0$. Then $m=0$ and systems (10) become

$$\dot{x}=a+cx,\quad \dot{y}=b+fy-x^2. \tag{35}$$

For these systems we calculate

$$\begin{gathered}
M=B_3=N=0,\quad K_5=x^3\neq 0,\quad N_3=3(c-f)x^3,\\
N_6=8c(c-f)x^3,\quad K_1=-cx^3,\quad K_3=6(2c-f)fx^6,\\
D_1=c+f,\quad D=-f^2x^3,\quad \mu_3=-c^2fx^3,\quad W_1=2cf,\quad V_2=(a+cx)\times\\
\big[b(c^2-f^2)-4a^2-2a(3c-f)x-(c-f)(3c+f)x^2+f(c^2-f^2)y\big]
\end{gathered}\tag{36}$$

So according to [9] and [8] non-degenerate systems (35) possess invariant straight lines of total multiplicity at least four.

b_1) Assume first $W_1\neq 0$, i.e. $cf\neq 0$. Then $\mu_3\neq 0$ and according to Lemma 4 the family of systems (35) does not contain degenerate systems.

If $V_2\neq 0$ then by statement (i) of Main Theorem any system (35) is located on an orbit of dimension 5. Moreover from (36) it follows that the condition $W_1N_6\neq 0$ implies $K_1DN_3\neq 0$ and $N_6=0$ gives $N_3=0$ (due to $W_1\neq 0$). So as $M=0$ and $B_3=N=0$, according to [11] and [12] a non-degenerate system could have one of the following phase portraits:

Picture 4.37(a) $[\mathcal{P}_{32}]$ $\Leftrightarrow$ $N_6\neq 0,\ \mu_3K_1>0,\ K_3\geq 0;$
Picture 4.37(b) $[\mathcal{P}_{33}]$ $\Leftrightarrow$ $N_6\neq 0,\ \mu_3K_1>0,\ K_3<0;$
Picture 4.37(c) $[\mathcal{P}_{34}]$ $\Leftrightarrow$ $N_6\neq 0,\ \mu_3K_1<0;$
Picture 5.27 $[\mathcal{P}_{32}]$ $\Leftrightarrow$ $N_6=0.$

We remark that when $N_6=0$ (i.e. $f=c$) we have $N_4=12ax^2\neq 0$ due to $V_2\neq 0$. We observe also that *Picture 5.27* is topologically equivalent to *Picture 4.37(a)* and the condition $N_6=0$ implies $\mu_3K_1>0$ and $K_3>0$. So we get the respective conditions given in Table 1.

Suppose $V_2=0$. By (36) due to $cf\neq 0$ we obtain

$$b(c^2-f^2)-4a^2=a(3c-f)=(c-f)(3c+f)=(c^2-f^2)=0.$$

This implies $c = f$ (otherwise $c+f = 3c+f = 0$ gives $c = f = 0$). So if $W_1 \neq 0$ and $V_2 = 0$ we obtain $a = c - f = 0$ and $cf \neq 0$ and then $W_2 = -3c^3x^3 \neq 0$. Then by statement (i) of Main Theorem any system (35) in this case is located on an orbit of dimension 4.

Thus for $W_1 \neq 0$ and $V_2 = 0$ we arrive at systems

$$\dot{x} = cx, \quad \dot{y} = b + cy - x^2,$$

for which we have

$$M = B_3 = N = N_3 = N_4 = 0, \quad \mu_3 = -c^3x^3, \quad W_1 = 2c^2 \neq 0.$$

According to [8] these systems possess invariant straight lines of total multiplicity six and by [12] we get the phase portrait *Picture 6.11* which is topologically equivalent to $\mathcal{P}_{32}$.

b_2) Assume now $W_1 = 0$, i.e. $cf = 0$. If $D \neq 0$ by (36) we have $f \neq 0$ and this implies $c = 0$. Moreover we may assume $f = 1$ and $b = 0$ due to the change $(x, y, t) \mapsto (x, (y-b)/f, t/f)$. So we get the family of systems

$$\dot{x} = a, \quad \dot{y} = y - x^2, \tag{37}$$

for which we calculate:

$$W_2 = -a(a^2 - 2ax - x^2 + y), \quad M = B_3 = N = N_6 = 0, N_3 = -3x^3, \ D = -x^3.$$

If $W_2 \neq 0$ then $a \neq 0$ and we may assume $a = 1$ due to the rescaling $(x, y, t) \mapsto (ax, a^2y)$. By statement ($i$) of Main Theorem this system is located on an orbit of dimension 5. According to [11] it possesses invariant lines of total multiplicity 4 (more exactly the infinite line is of multiplicity 4) and its phase portrait corresponds to *Picture 4.46* which is topologically equivalent to $\mathcal{P}_{35}$.

Assume $W_2 = 0$. Then $a = 0$ and system (37) is degenerate having the parabola $y = x^2$ filled with singularities. Obviously we get the phase portrait $\mathcal{P}_{44}$. On the other hand as $D \neq 0$ we have $\mathcal{V}_2 \neq 0$ and by statement (i) of Main Theorem this system is located on an orbit of dimension 4.

Suppose now $D = 0$, i.e. $f = 0$. Then systems (35) become

$$\dot{x} = a + cx, \quad \dot{y} = b - x^2, \tag{38}$$

for which calculations yield

$$M = 0, \quad K_5 = x^3, \quad W_1 = D = 0, \ U_3 = x^2(4bx^2 - 4axy + c^2y^2),$$
$$W_2 = -(a^2 - bc^2)(a + cx).$$

As $D = 0$ the condition $\mathcal{V}_2 = 0$ is equivalent to $U_3 = 0$. So by statement (i) of Main Theorem a system of this family is located on an orbit of dimension 5 (respectively 4; 3) if $W_2 \neq 0$ (respectively $W_2 = 0$, $U_3 \neq 0$; $W_2 = U_3 = 0$).

On the other hand for systems (38) we have

$$M = B_3 = N = D = 0, \quad N_3 = 3cx^3, \quad N_6 = 8c^2x^3,$$
$$D_1 = c, \quad \mu_0 = \mu_1 = \mu_2 = \mu_3 = 0, \quad \mu_4 = (a^2 - bc^2)x^4$$

and according to [9] and [8] non-degenerate systems (38) possess invariant straight lines of total multiplicity at least four.

α) Assume first $W_2 \neq 0$. Then $\mu_4 \neq 0$ and by Lemma 4 systems (38) are non-degenerate. According to [11, Table 2] in the case $N_3 \neq 0$ (then $D_1 \neq 0$) we get the phase portrait *Picture 4.38(a)* if $\mu_4 > 0$ and *Picture 4.38(b)* if $\mu_4 < 0$. However there is a missprint in [11, Table 2].

Remark 3. Assume that a quadratic systems has a configuration of invariant lines given by *Config. 4.38*. Then its phase portrait corresponds to *Picture 4.38(a)* [$\mathcal{P}_{36}$] if $\mu_4 < 0$ and *Picture 4.38(b)* [$\mathcal{P}_{29}$] if $\mu_4 > 0$.

If $N_6 = 0$ (i.e. $c = 0$) we have $N_3 = D_1 = 0$ and $N_4 = 12ax^2 \neq 0$ due to $W_2 \neq 0$. According to [12] in this case the phase portrait corresponds to *Picture 5.30* which is topologically equivalent to $\mathcal{P}_{29}$. As the condition $c = 0$ implies $\mu_4 = a^2x^4 > 0$ we could unite the cases $N \neq 0$ and $N = 0$ as it is given in Table 1.

β) Suppose now $W_2 = 0$, i.e. $bc^2 - a^2 = 0$. Then $\mu_4 = 0$ and by Lemma 4 systems (38) become degenerate.

If $N_6 \neq 0$ then $c \neq 0$ (this implies $U_3 \neq 0$) and we may assume $c = 1$ due to the rescaling $(x, y, t) \mapsto (cx, cy, t/c)$. So we obtain $b = a^2$ and this leads to the degenerate systems

$$\dot{x} = a + x, \quad \dot{y} = (a + x)(a - x),$$

with $a \in \{0, 1\}$ due to the rescaling $(x, y) \mapsto (ax, a^2y)$ in the case $a \neq 0$. Obviously in both cases we obtain the same phase portrait $\mathcal{P}_{45}$.

If $N_6 = 0$ then $c = 0$ and the condition $W_2 = 0$ gives $a = 0$. So we get the systems

$$\dot{x} = 0, \quad \dot{y} = b - x^2,$$

where $b \in \{0, \pm1\}$ due to the rescaling $(x, y) \mapsto (|b|^{-1/2}x, |b|^{-1}y)$, in the case $b \neq 0$. Evidently we obtain the phase portrait $\mathcal{P}_{29}$ if $b < 0$, $\mathcal{P}_{46}$ if $b > 0$ and $\mathcal{P}_{49}$ if $b = 0$. We observe that for the systems above we have $U_3 = 4bx^4$ and hence $\text{sign}\,(b) = \text{sign}\,(U_3)$ (if $U_3 \neq 0$). We recall also that for $U_3 = 0$ (i.e. $b = 0$) the respective system is located on the orbit of dimension 3.

It remains to note that for systems (10) we have $\text{Coefficient}[W_1, x^2] = 16m^2$ and hence the condition $W_1 = 0$ implies $H = 0$.

3) The case $K_5 = 0$**.** It was shown earlier in the proof of Lemma 1 (see page 35) that in this case a system can be brought via an affine transformation to form (16), for which the condition $\mathcal{U} = 0$ gives $d = 0$. So we get the family

$$\dot{x} = a + cx + x^2, \quad \dot{y} = b + xy, \tag{39}$$

for which we calculate:

$$W_2 = (c+3x)(a+cx+x^2)(bc+bx-ay), \quad H_{10} = 0, \quad H_{12} = -8a^2x^2,$$
$$H_{11} = 48(c^2-4a)x^4, \quad \mu_2 = ax^2.$$

By statement (i) of Main Theorem a system of this family is located on an orbit of dimension 5 if $W_2 \neq 0$ and of dimension 4 if $W_2 = 0$. We observe that the condition $W_2 = 0$ is equivalent to $a = b = 0$ and then systems (39) become degenerate.

Assume $W_2 \neq 0$. According to [10] a non-degenerate system (39) could possess one of the following phase portraits:

Picture $C_2.5(a)$	$[\mathcal{P}_{39}]$	$\Leftrightarrow$	$H_{12} \neq 0,\ H_{11} > 0,\ \mu_2 < 0$;
Picture $C_2.5(b)$	$[\mathcal{P}_{38}]$	$\Leftrightarrow$	$H_{12} \neq 0,\ H_{11} > 0,\ \mu_2 > 0$;
Picture $C_2.6$	$[\mathcal{P}_{37}]$	$\Leftrightarrow$	$H_{12} \neq 0,\ H_{11} < 0$;
Picture $C_2.7$	$[\mathcal{P}_{41}]$	$\Leftrightarrow$	$H_{12} \neq 0,\ H_{11} = 0$;
Picture $C_2.8$	$[\mathcal{P}_{40}]$	$\Leftrightarrow$	$H_{12} = 0,\ H_{11} \neq 0$;
Picture $C_2.9$	$[\mathcal{P}_{42}]$	$\Leftrightarrow$	$H_{12} = 0,\ H_{11} = 0$;

We observe that the condition $\mu_2 \neq 0$ implies $H_{12} \neq 0$. Moreover if $H_{11} < 0$ then $\mu_2 > 0$. So we arrive at the respective conditions given in Table 1.

Suppose now $W_2 = 0$, i.e. $a = b = 0$. In this case we get the family of degenerate systems

$$\dot{x} = cx + x^2, \quad \dot{y} = xy.$$

Obviously we obtain the phase portrait given by picture $\mathcal{P}_{47}$ if $c \neq 0$ and $\mathcal{P}_{48}$ if $c = 0$. It remains to observe that for $a = b = 0$ we have $H_{11} = 48c^2x^4$ and this polynomial gives the condition $c = 0$. $\square$

References

[1] Popa M.N. *Algebraic methods for differential systems.* Editura the Flower Power, Universitatea din Piteşti, Seria Matematică Aplicată şi Industrială, 2004, **15** (in Romanian).

[2] Sibirsky K.S. *Introduction to algebraic theory of invariants of differential equations*, Chisinau, Stiinta, 1982 (in Russian, published in English in 1988).

[3] Olver P.J. *Classical Invariant Theory*, London Mathematical Society student texts: **44**, Cambridge University Press, 1999.

[4] Boularas D., Calin Iu., Timochouk L., Vulpe N. *T-comitants of quadratic systems: a study via the translation invariants*, Delft University of Technology, Report **96-90**, Delft 1996

[5] Schlomiuk D., Vulpe N. *Geometry of quadratic differential systems in the neighborhood of infinity*, J. Differential Equations, 2005, **215**, 357–400.

[6] Boularas D. *Computation of affine covariants of quadratic bivariate differential systems*, J. Complexity, 2000, **16**, No. 4, 691–715.

[7] Gherstega N., Orlov V. *Classification of Aff$(2,\mathbb{R})-$orbit's dimensions for quadratic differential system*, Buletinul Academiei de Ştiinţe a Republicii Moldova, Matematica, 2008, No. 2(57), 122-126.

[8] Schlomiuk D., Vulpe N. *Planar quadratic differential systems with invariant straight lines of at least five total multiplicity*, Qualitative Theory of Dynamical Systems, 2004, **5**, 135–194.

[9] Schlomiuk D., Vulpe N. *Planar quadratic differential systems with invariant straight lines of total multiplicity four*, Nonlinear Anal., 2008, **68**, No. 4, 681–715.

[10] Schlomiuk D., Vulpe N. *The full study of planar quadratic differential systems possessing a line of singularities at infinity*, J. Dynam. Differential Equations, 2008, **20**, 737–775.

[11] Schlomiuk D., Vulpe N. *Integrals and phase portraits of planar quadratic systems with invariant lines of total multiplicity four*, Bul. Acad. Ştiinţe Repub. Mold. Mat., 2008, No. 1(56), 27–83.

[12] Schlomiuk D., Vulpe N. *Integrals and phase portraits of planar quadratic differential systems with invariant lines of at least five total multiplicity*, Rocky Mountain J. Math., 2008, **38**, 2015–2076.

[13] Baltag V. A., Vulpe N. I. *Total multiplicity of all finite critical points of the polynomial differential system,* Planar nonlinear dynamical systems (Delft, 1995). Differential Equations & Dynam. Systems, 1997, **5**, No. 3-4, 455–471.

Institute of Mathematics and Computer Sciences
Academy of Sciences of Moldova
str. Academiei 5, MD-2028 Chisinau
Moldova,

E-mail: *gherstega@gmail.com*, *orlovictor@gmail.com*, *nvulpe@gmail.com*

Received June 18, 2009

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 55–61
ISSN 1024–7696

On stability and quasi-stability radii for a vector combinatorial problem with a parametric optimality principle

Vladimir A. Emelichev, Evgeny E. Gurevsky, Andrey A. Platonov

Abstract. A vector combinatorial linear problem with a parametric optimality principle that allows us to relate the well-known choice functions of jointly-extremal and Pareto solution is considered. A quantitative analysis of stability for the set of generalized efficient trajectories under the independent perturbations of coefficients of linear functions is performed. Formulas of stability and quasi-stability radii are obtained in the l_∞-metric. Some results published earlier are derived as corollaries.

Mathematics subject classification: 90C29, 90C31.
Keywords and phrases: Multiobjectivity, combinatorial optimization, Pareto optimality, jointly-extremal optimality, stability radius, quasi-stability radius.

1 Problem statement

Let us consider a typical vector (n-criteria) combinatorial problem. Assume that, on the system of subsets (trajectories) $T \subseteq 2^{N_m}$, $|T| \geq 2$, $N_m = \{1, 2, \ldots, m\}$, $m \geq 2$, a vector criterion

$$f(t, A) = (f_1(t, A), f_2(t, A), \ldots, f_n(t, A)) \to \min_{t \in T}$$

is defined. Here

$$f_i(t, A) = \sum_{j \in t} a_{ij}, \ i \in N_n, \ n \geq 1$$

are the linear partial criteria, where $A = [a_{ij}]_{n \times m} \in \mathbf{R}^{n \times m}$, $n, m \in \mathbf{N}$. Assume that $f_i(\emptyset, A) = 0$.

Now we introduce the binary relation $\succ$, in the space $\mathbf{R}^d$ of any dimension $d \in \mathbf{N}$, which generates the Pareto optimality principle [1], assuming that, for any different vectors $y = (y_1, y_2, \ldots, y_d)$ and $y' = (y'_1, y'_2, \ldots, y'_d)$ of the space the formula

$$y \succ y' \ \Leftrightarrow \ y \geq y' \ \& \ y \neq y'$$

holds.

Let $s \in N_n$, $N_n = \bigcup_{r \in N_s} J_r$ be the partitioning of the set N_n into s nonempty nonintersecting groups, i. e. $J_r \neq \emptyset$, $r \in N_s$; $p \neq q \Rightarrow J_p \cap J_q = \emptyset$. For this partitioning define the set $T^n(A, J_1, J_2, \dots, J_s)$ of *generalized efficient*, or, in other words of $(J_1, J_2, \dots, J_s)$-*efficient* trajectories according to the formula

$$T^n(A, J_1, J_2, \dots, J_s) = \{t \in T : \exists k \in N_s \ \ \forall t' \in T \ \ (f_{J_k}(t, A) \overline{\succ} f_{J_k}(t', A))\},$$

where $\overline{\succ}$ denotes the negation of relation $\succ$, $f_{J_k}(t, A)$ is the projection of the vector $f(t, A)$ onto the coordinate axes of the space $\mathbf{R}^n$ with the numbers of group J_k.

It is evident that N_n-efficient trajectory $t \in T^n(A, N_n)$ $(s = 1)$ is a Pareto optimal trajectory on the set of trajectories T. Therefore, it is easy to see that the set of N_n-efficient trajectories $T^n(A, N_n)$ is Pareto set

$$P^n(A) = \{t \in T : \ \forall t' \in T \ \ (f(t, A) \overline{\succ} f(t', A))\}.$$

Clearly, in another extreme case, where $s = n$, the set of trajectories $T^n(A, \{1\}, \{2\}, \dots, \{n\})$ is the set of *jointly-extremal* trajectories

$$C^n(A) = \{t \in T : \ \exists k \in N_n \ \ \forall t' \in T \ \ (f_k(t, A) \leq f_k(t', A))\}$$

(see, for example, [2,3]).

In this context, by the parametrization of the principle of optimality we mean introducing a characteristic of binary relation of preference that allows us to relate the well-known choice functions of jointly-extremal and Pareto solution.

Denote the vector problem of finding $T^n(A, J_1, J_2, \dots, J_s)$ by $Z^n(A, J_1, J_2, \dots, J_s)$. It is evident that $T^1(A, N_1)$ is the set of optimal trajectories of the scalar (single criterion) linear combinatorial problem $Z^1(A, N_1)$, where $A \in \mathbf{R}^m$, in scheme of which many extremal graph, boolean programming and scheduling theory problems are put in.

2 Stability radius

Following [4–10], the *stability radius* of $Z^n(A, J_1, J_2, \dots, J_s)$ is the number

$$\rho_1^n(A, J_1, J_2, \dots, J_s) = \begin{cases} \sup \Xi_1 & \text{if } \ \Xi_1 \neq \emptyset, \\ 0 & \text{in other cases,} \end{cases}$$

where $\Xi_1 = \{\varepsilon > 0 : \ \forall B \in \Omega(\varepsilon) \ \ (T^n(A+B, J_1, J_2, \dots, J_s) \subseteq T^n(A, J_1, J_2, \dots, J_s))\}$, $\Omega(\varepsilon) = \{B \in \mathbf{R}^{n \times m} : \ ||B|| < \varepsilon\}$, $||B|| = \max\{|b_{ij}| : \ (i, j) \in N_n \times N_m\}$, $B = [b_{ij}]_{n \times m}$.

In other words, the stability radius of $Z^n(A, J_1, J_2, \dots, J_s)$ determines the limiting level of perturbations of elements of A of payoff function in the l_∞-metric, for which new $(J_1, J_2, \dots, J_s)$-efficient trajectories do not appear. Obviously, $Z^n(A, J_1, J_2, \dots, J_s)$ is stable and the stability radius is infinite if the equality $T^n(A, J_1, J_2, \dots, J_s) = T$ holds. If the set

$$\overline{T^n}(A, J_1, J_2, \dots, J_s) := T \setminus T^n(C, J_1, J_2, \dots, J_s)$$

is nonempty, then we say that $Z^n(A, J_1, J_2, \ldots, J_s)$ is *non-trivial.*

For any nonempty set $J \subseteq N_n$ we introduce the notation

$$P(A, J) = \{t \in T : \ \forall t' \in T \ \ (f_J(t, A) \overline{\succ} f_J(t', A))\}.$$

Then we have

$$P(A, N_n) = P^n(A),$$

$$T^n(A, J_1, J_2, \ldots, J_s) = \{t \in T : \ \exists k \in N_s \ \ (t \in P(A, J_k))\}. \tag{1}$$

Suppose

$$\Delta(t, t') = |(t \cup t') \setminus (t \cap t')|,$$

$$g_i(t, t', A) = f_i(t, A) - f_i(t', A).$$

Henceforth we will use the following evident inequality

$$g_i(t, t', A) \leq ||A||\Delta(t, t'). \tag{2}$$

Theorem 1. *For the stability radius $\rho_1^n(A, J_1, J_2, \ldots, J_s)$ of the nontrivial problem $Z^n(A, J_1, J_2, \ldots, J_s)$, $n \geq 1$, $s \geq 1$, the following formula is valid*

$$\rho_1^n(A, J_1, J_2, \ldots, J_s) = \min_{k \in N_s} \ \min_{t \in \overline{T^n}(A, J_1, J_2, \ldots, J_s)} \ \max_{t' \in T^n(A, J_1, J_2, \ldots, J_s)} \ \min_{i \in J_k} \frac{g_i(t, t', A)}{\Delta(t, t')}. \tag{3}$$

Proof. Note that due to the nontriviality of $Z^n(A, J_1, J_2, \ldots, J_s)$ the set $\overline{T^n}(A, J_1, J_2, \ldots, J_s)$ is nonempty.

Let us introduce the notations: ρ_1 and φ are accordingly the left-hand and the right-hand sides of equality (3).

It is easy to see that $\varphi \geq 0$. At first we prove the inequality $\rho_1 \geq \varphi$. If $\varphi = 0$, then this inequality is obvious. Let $\varphi > 0$, $B \in \Omega(\varphi)$, $t \in \overline{T^n}(A, J_1, J_2, \ldots, J_s)$. Let us show that $t \in \overline{T^n}(A + B, J_1, J_2, \ldots, J_s)$.

It follows directly from the definition of φ that for any $k \in N_s$ and $t \in \overline{T^n}(A, J_1, J_2, \ldots, J_s)$ there exists trajectory $t^* \in T^n(A, J_1, J_2, \ldots, J_s)$ such that for any indices $i \in J_k$ the inequality $g_i(t, t^*, A) \geq \varphi\Delta(t, t^*)$ holds.

Hence, taking into account (2), we derive

$$g_i(t, t^*, A + B) = g_i(t, t^*, A) + g_i(t, t^*, B) \geq \varphi\Delta(t, t^*) - ||B||\Delta(t, t^*) > 0, \ i \in J_k.$$

Therefore we have $f_{J_k}(t, A + B) \succ f_{J_k}(t^*, A + B)$, $k \in N_s$, i.e. $t \in \overline{T^n}(A + B, J_1, J_2, \ldots, J_s)$.

Thus, the formula

$$\forall B \in \Omega(\varphi) \ \ (T^n(A + B, J_1, J_2, \ldots, J_s) \subseteq T^n(A, J_1, J_2, \ldots, J_s))$$

holds, and as consequence, $\rho_1 \geq \varphi$.

Now let us show that $\rho_1 \leq \varphi$. According to the definition of φ there exist $k \in N_s$ and $t^0 \in \overline{T^n}(A, J_1, J_2, \dots, J_s)$ such that for any trajectory $t' \in T^n(A, J_1, J_2, \dots, J_s)$ there exists an index $p = p(t') \in J_k$ such that

$$g_p(t^0, t', A) \leq \varphi\Delta(t^0, t').$$

Then, assuming $\varepsilon > \varphi$, $B^0 = [b^0_{ij}]_{n\times m} \in \Omega(\varepsilon)$, where

$$b^0_{ij} = \begin{cases} \alpha & \text{if } \ i \in J_k, \ j \notin t^0, \\ -\alpha & \text{if } \ i \in J_k, \ j \in t^0, \\ 0 & \text{in other cases,} \end{cases}$$

$$\varphi < \alpha < \varepsilon,$$

and using (2), we derive

$$g_p(t^0, t', A + B^0) = g_p(t^0, t', A) + g_p(t^0, t', B^0) \leq \varphi\Delta(t^0, t') - \alpha\Delta(t^0, t') < 0,$$

i. e. we have

$$\forall\varepsilon > \varphi \ \ \exists B^0 \in \Omega(\varepsilon) \ \ \forall t' \in T^n(A, J_1, J_2, \dots, J_s) \ \ (f_{J_k}(t^0, A + B^0) \overline{\succ} f_{J_k}(t', A + B^0)). \tag{4}$$

Consider two possible cases.

Case 1. $t^0 \in T^n(A + B^0, J_1, J_2, \dots, J_s)$. Then, using the inclusion $t^0 \in \overline{T^n}(A, J_1, J_2, \dots, J_s)$, we derive

$$\forall\varepsilon > \varphi \ \ \exists B^0 \in \Omega(\varepsilon) \ \ (T^n(A + B^0, J_1, J_2, \dots, J_s) \not\subseteq T^n(A, J_1, J_2, \dots, J_s)). \tag{5}$$

Case 2. $t^0 \in \overline{T^n}(A + B^0, J_1, J_2, \dots, J_s)$. Then $t^0 \notin P(A + B^0, J_k)$ and due to the external stability [11] of Pareto set $P(A + B^0, J_k)$ there exists a trajectory $t^* \in P(A + B^0, J_k)$, such that $f_{J_k}(t^0, A + B^0) \succ f_{J_k}(t^*, A + B^0)$. Hence, according to (4) we have $t^* \in \overline{T^n}(A, J_1, J_2, \dots, J_s)$ and taking into account (1) we obtain $t^* \in T^n(A + B^0, J_1, J_2, \dots, J_s)$. Therefore formula (5) holds.

Summarizing these two cases, we conclude that for any $\varepsilon > \varphi$ we have $\rho_1 < \varepsilon$. Consequently, $\rho_1 \leq \varphi$. $\square$

Theorem 1 implies the following results known earlier.

Corollary 1 [5]. *For the stability radius of the nontrivial problem $Z^n(A, N_n)$ with Pareto optimality principle the following formula*

$$\rho^n_1(A, N_n) = \min_{t\in\overline{P^n}(A)} \ \max_{t'\in P^n(A)} \ \min_{i\in N_n} \ \frac{g_i(t, t', A)}{\Delta(t, t')} \tag{6}$$

holds, where $\overline{P^n}(A) = T \setminus P^n(A)$.

Corollary 2 [12]. *For the stability radius of the nontrivial problem* $Z^n(A, \{1\}, \{2\}, \dots, \{n\})$ *with jointly-extremal optimality principle the following formula*

$$\rho_1^n(A, \{1\}, \{2\}, \dots, \{n\}) = \min_{i \in N_n} \min_{t \in \overline{C^n}(A)} \max_{t' \in C^n(A)} \frac{g_i(t, t', A)}{\Delta(t, t')} \tag{7}$$

holds, where $\overline{C^n}(A) = T \setminus C^n(A)$.

The partial case of the formulas (6) and (7) is the well-known formula of the stability radius of the scalar ($n = 1$) linear trajectory problem [4, 6].

3 Quasi-stability radius

As usual (see [5, 7, 13, 14]), the *quasi-stability radius* of $Z^n(A, J_1, J_2, \dots, J_s)$ is defined as

$$\rho_2^n(A, J_1, J_2, \dots, J_s) = \begin{cases} \sup \Xi_2 & \text{if } \ \Xi_2 \neq \emptyset, \\ 0 & \text{in other cases,} \end{cases}$$

where

$$\Xi_2 = \{\varepsilon > 0 : \ \forall B \in \Omega(\varepsilon) \ \ (T^n(A, J_1, J_2, \dots, J_s) \subseteq T^n(A + B, J_1, J_2, \dots, J_s))\}.$$

Thus, the quasi-stability radius of $Z^n(A, J_1, J_2, \dots, J_s)$ is the limit level of independent perturbations of elements of A, for which the generalized efficient trajectories of initial problem do not disappear.

Theorem 2. *For the quasi-stability radius* $\rho_2^n(A, J_1, J_2, \dots, J_s)$ *of the problem* $Z^n(A, J_1, J_2, \dots, J_s)$, $n \geq 1$, $s \geq 1$, *the following formula is valid*

$$\rho_2^n(A, J_1, J_2, \dots, J_s) = \min_{t' \in T^n(A, J_1, J_2, \dots, J_s)} \max_{k \in N_s} \min_{t \in T \setminus \{t'\}} \max_{i \in J_k} \frac{g_i(t, t', A)}{\Delta(t, t')}. \tag{8}$$

Proof. Let us introduce the notations: ρ_2 and ξ are accordingly the left-hand and the right-hand sides of equality (8).

It is easy to see that $\xi \geq 0$. At first we prove the inequality $\rho_2 \geq \xi$. If $\xi = 0$, then this inequality is obvious. Let $\xi > 0$, $B \in \Omega(\xi)$.

It follows from the definition of ξ that for any trajectory $t' \in T^n(A, J_1, J_2, \dots, J_s)$ there exists $k \in N_s$ such that for any trajectory $t \in T \setminus \{t'\}$ there exists $p = p(t) \in J_k$ such that

$$g_p(t, t', A) \geq \xi \Delta(t, t').$$

Hence, taking into account (2), we derive

$$g_p(t, t', A + B) = g_p(t, t', A) + g_p(t, t', B) \geq \xi \Delta(t, t') - ||B|| \Delta(t, t') > 0.$$

Therefore, we have $f_{J_k}(t', A + B) \overline{\succ} f_{J_k}(t, A + B)$. Thus, we have proved the formula

$$\forall B \in \Omega(\xi) \ \forall t' \in T^n(A, J_1, J_2, \dots, J_s) \ \exists k \in N_s \ \forall t \in T \ \ (f_{J_k}(t', A{+}B) \overline{\succ} f_{J_k}(t, A{+}B)),$$

which implies

$$\forall B \in \Omega(\xi) \quad (T^n(A, J_1, J_2, \dots, J_s) \subseteq T^n(A + B, J_1, J_2, \dots, J_s)),$$

and therefore the inequality $\rho_2 \geq \xi$ holds.

Now we show that $\rho_2 \leq \xi$. According to the definition of ξ there exists a trajectory $t^0 \in T^n(A, J_1, J_2, \dots, J_s)$ such that for any $k \in N_s$ there exists a trajectory $t^* \in T \setminus \{t^0\}$ such that

$$\forall i \in J_k \quad (g_i(t^*, t^0, A) \leq \xi \Delta(t^*, t^0)).$$

Then, assuming $\varepsilon > \xi$, $\widehat{B} = [\widehat{b}_{ij}]_{n \times m} \in \Omega(\varepsilon)$, where

$$\widehat{b}_{ij} = \begin{cases} \alpha & \text{if } \ i \in N_n, \ j \in t^0, \\ -\alpha & \text{if } \ i \in N_n, \ j \notin t^0, \end{cases}$$

$$\xi < \alpha < \varepsilon,$$

and taking into account (2), we derive

$$g_i(t^*, t^0, A + \widehat{B}) = g_i(t^*, t^0, A) + g_i(t^*, t^0, \widehat{B}) \leq \xi \Delta(t^*, t^0) - \alpha \Delta(t^*, t^0) < 0, \ i \in J_k,$$

i. e. $f_{J_k}(t^0, A + \widehat{B}) \succ f_{J_k}(t^*, A + \widehat{B})$. Thus, we have proved the following formula

$$\forall \varepsilon > \xi \quad \exists \widehat{B} \in \Omega(\varepsilon) \quad \forall k \in N_s \quad \exists t^* \in T \quad (f_{J_k}(t^0, A + \widehat{B}) \succ f_{J_k}(t^*, A + \widehat{B})),$$

which implies

$$T^n(A, J_1, J_2, \dots, J_s) \not\subseteq T^n(A + \widehat{B}, J_1, J_2, \dots, J_s).$$

It follows that the quasi-stability radius ρ_2 does not exceed ξ. □

Corollary 3 [13]. *For the quasi-stability radius of the problem $Z^n(A, N_n)$ with Pareto optimality principle the following formula is valid*

$$\rho_2^n(A, N_n) = \min_{t' \in P^n(A)} \ \min_{t \in T \setminus \{t'\}} \ \max_{i \in N_n} \ \frac{g_i(t, t', A)}{\Delta(t, t')}.$$

Corollary 4. *For the quasi-stability radius of the problem $Z^n(A, \{1\}, \{2\}, \dots, \{n\})$ with jointly-extremal optimality principle the following formula is valid*

$$\rho_2^n(A, \{1\}, \{2\}, \dots, \{n\}) = \min_{t' \in C^n(A)} \ \max_{i \in N_n} \ \min_{t \in T \setminus \{t'\}} \ \frac{g_i(t, t', A)}{\Delta(t, t')}.$$

In conclusion we note that the analogous quantitative characteristics of different stability types of discrete and game theory problems with another kinds of parametrization of optimality principles were considered in the works [8–10, 14–16].

References

[1] Pareto V. *Manuel d'économie politique*. Paris, Giard, 1909.

[2] Sholomov L.A. *Logical methods for investigating of discrete models of choice*. Moskow, Nauka, 1989 (in Russian).

[3] Yudin D.B *Computational methods of desicion making theory*. Moskow, Nauka, 1989 (in Russian).

[4] Leontev V.K. *Stability of linear discrete problems*. Cybernetics Problems, 1979, **35**, 169–184 (in Russian).

[5] Emelichev V.A., Kravtsov M.K. *Stability in vector optimization path problems*. Cybernetics and Systems Analysis, 1995, **31**, No. 4. 599–604.

[6] Gordeev E.N., Leontev V.K. *General approach of stability solution investigating of discrete optimization problems*. Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki, 1996, **36**, No. 1, 66–72 (in Russian).

[7] Emelichev V.A., Girlich E., Nikulin Yu.V., Podkopaev D.P. *Stability and regularization of vector problems of integer linear programming*. Optimization, 2002, **51**, No. 4, 645–676.

[8] Bukhtoyarov S.E., Emelichev V.A. *Optimality principle parametrization (from Pareto to Slater) and stability of multicriteria trajectory problems*. Diskretnyi Analiz i Issledovanie Operatsii, Seriya 2, 2003, **10**, No. 2, 3–18 (in Russian).

[9] Bukhtoyarov S.E., Emelichev V.A. *Measure of stability for a finite cooperative game with a parametric optimality principle (from Pareto to Nash)*. Computational Mathematics and Mathematical Physics, 2006, **46**, No. 7, 1193–1199.

[10] Emelichev V.A., Gurevsky E.E., Platonov A.A. *Measure of stability for a finite cooperative game with a generalized concept of equilibrium*. Buletinul Academiei de Ştiinţe a Republicii Moldova, Matematica, 2006, No. 3(52), 17–26.

[11] Podinovskii V.V., Nogin V.D. *Pareto-optimal solutions of multiobjective problems*. Moscow, Nauka, 1982 (in Russian).

[12] Bukhtoyarov S.E., Emelichev V.A. *Stability measure of a vector linear trajectory problem with jointly-extremal optimality principle*. Vestnik BGU, Seriya 1, 2002, No. 3, 84–86 (in Russian).

[13] Emelichev V.A., Kravtsov M.K., Podkopaev D.P. *On the quasistability of trajectory problems of vector optimization*. Mathematical Notes, 1998, **63**, No. 1, 19–24.

[14] Emelichev V.A., Kuzmin K.G. *Quasi-stability measure of a vector combinatorial problem with a parametric optimality principle in the l_1-metric*. Izvestiya Vuzov, Matematika, 2005, No. 12, 3–10 (in Russian).

[15] Bukhtoyarov S.E., Emelichev V.A., Stepanishina Yu.V. *Stability of discrete vector problems with the parametric principle of optimality*. Cybernetics and Systems Analysis, 2003, **39**, No. 4, 604–614.

[16] Emelichev V.A., Kuzmin K.G. *Finite cooperative games with a parametric concept of equilibrium under uncertainty conditions*. Journal of Computer and Systems Sciences International, 2006, **45**, No. 2, 276–281.

Belarusian State University
av. Nezavisimosti, 4, 220030 Minsk
Belarus

E-mail: *emelichev@bsu.by; evgeny_gurevsky@tut.by; platonov@scnsoft.com*

Received April 03, 2009

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 62–72
ISSN 1024–7696

On preradicals associated to principal functors of module categories, I

A. I. Kashu*

Abstract. The preradicals associated to the functor $Hom_R(U,\text{-}) : R\text{-}Mod \to \mathcal{A}b$ are revealed, their properties and the relations between these preradicals are studied.

Mathematics subject classification: 16D90, 16S90, 16D40.
Keywords and phrases: Preradical, torsion, cotorsion, functor, torsion class, torsionfree class.

Introduction

The radicals and torsions associated to adjoint situations and Morita contexts were studied in a series of papers, which were totalized in the book [1]. The aim of this article is the generalization, supplement and specification of some results of [1] concerning the preradicals in module categories which are determined by principal functors of module categories:

$$H = H^U = Hom_R(U,\text{-}) : R\text{-}Mod \to \mathcal{A}b \quad ({}_RU \in R\text{-}Mod),$$

$$T = T^U = U \otimes_S \text{-} : S\text{-}Mod \to \mathcal{A}b \quad (U_S \in Mod\text{-}S),$$

$$H' = H_U = Hom_R(\text{-}, U) : R\text{-}Mod \to \mathcal{A}b \quad ({}_RU \in R\text{-}Mod),$$

where $\mathcal{A}b$ is the category of abelian groups. In particular, it will be shown that some results which were proved for adjoint situations and Morita contexts are valid in general case (without supplementary restrictions). The preradicals associated to each of functors H, T and H' will be elucidated, the properties of these preradicals, as well as the relations between them and the conditions of coincidence of some preradicals will be shown.

The part I of this work is dedicated to the study of indicated above questions for the functor $H = Hom_R(U,\text{-})$ for an arbitrary module ${}_RU \in R$-Mod. In the following parts the functors T and H' will be investigated from the same aspect.

1 Preliminary notions and results

The basic notions and results of radical theory in modules can be found in the books [2–5]. For specification of terminology and notations we will remind some of them.

*The author was partially supported by the grant 08.820.08.12 RF.

Let R be a ring with unity and R-Mod is the category of unitary left R-modules. A *preradical* r of R-Mod is a subfunctor of identic functor of R-Mod, i.e. r associates to every module $M \in R$-Mod a submodule $r(M) \subseteq M$ such that $f(r(M)) \subseteq r(M')$ for every R-morphism $f: M \to M'$.

Now we remind the principal types of preradicals [2, 4].

A preradical r of R-Mod is called:

– *idempotent preradical* if $r(r(M)) = r(M)$ for every $M \in R$-Mod;

– *radical* if $r(M/r(M)) = 0$ for every $M \in R$-Mod;

– *idempotent radical* if both previous conditions are fulfilled;

– *pretorsion* if $r(N) = N \cap r(M)$ for every $N \subseteq M$;

– *torsion* if r is radical and pretorsion;

– *cohereditary preradical* if $r(M/N) = (r(M) + N)/N$ for every $N \subseteq M$;

– *cotorsion* if r is idempotent and cohereditary.

Every preradical r of R-Mod defines two classes of modules:

1) the class of r-*torsion* modules

$$\mathcal{R}(r) = \{M \in R\text{-}Mod \,|\, r(M) = M\};$$

2) the class of r-*torsionfree* modules

$$\mathcal{P}(r) = \{M \in R\text{-}Mod \,|\, r(M) = 0\}.$$

The special types of preradicals indicated above can be described by associated classes of modules. More exactly:

- every *idempotent preradical* r is described by the class $\mathcal{R}(r)$, which is closed under homomorphic images and direct sums; such classes are called *pretorsion classes*;
- every *radical* r is described by the class $\mathcal{P}(r)$, which is closed under submodules and direct products; the classes with such properties are called *pretorsionfree classes*;
- every *idempotent radical* r can be restored both by the class $\mathcal{R}(r)$ and $\mathcal{P}(r)$; the class $\mathcal{R}(r)$ is pretorsion and closed under extensions – such classes are called *torsion classes*; the class $\mathcal{P}(r)$ is pretorsionfree and closed under extensions – such classes are called *torsionfree classes*.

If r is a torsion then $\mathcal{R}(r)$ is a *hereditary torsion class* and $\mathcal{P}(r)$ is a *stable torsionfree class*. If r is a cotorsion, then $\mathcal{P}(r)$ is simultaneously a torsion class and a torsionfree class; such classes are called *TTF-classes*.

If r is an idempotent preradical of R-Mod, then it can be restored by the class $\mathcal{R}(r)$ in the following way:

$$r(M) = \sum \{N \subseteq M \,|\, N \in \mathcal{R}(r)\}.$$

Dually, if r is a radical of R-Mod, then it can be expressed by the class $\mathcal{P}(r)$ as follows:

$$r(M) = \cap \{N \subseteq M \,|\, M/N \in \mathcal{P}(r)\}.$$

In the theory of radicals in modules an essential role is played by the following two *operators of Hom-orthogonality*. For an arbitrary class of modules $\mathcal{K} \subseteq R$-Mod we define:

$$\mathcal{K}^{\uparrow} = \{M \in R\text{-}Mod \,|\, Hom_R(M, N) = 0 \;\text{ for every }\; N \in \mathcal{K}\},$$

$$\mathcal{K}^{\downarrow} = \{N \in R\text{-}Mod \,|\, Hom_R(M,N) = 0 \ \text{ for every } \ \ M \in \mathcal{K}\}.$$

The following facts are well known. For every class $\mathcal{K} \subseteq R$-Mod we have:

– $\mathcal{K}^{\uparrow}$ is a torsion class;

– $\mathcal{K}^{\downarrow}$ is a torsionfree class;

– $\mathcal{K}^{\downarrow\uparrow}$ is the least torsion class containing $\mathcal{K}$;

– $\mathcal{K}^{\uparrow\downarrow}$ is the least torsionfree class containing $\mathcal{K}$.

If r is an idempotent radical then:

$$\mathcal{R}(r) = \mathcal{P}(r)^{\uparrow}, \quad \mathcal{P}(r) = \mathcal{R}(r)^{\downarrow}.$$

In the family of all preradicals of the category R-Mod the relation of *partial order* can be defined as follows:

$$r_1 \leq r_2 \overset{\text{def}}{\Longleftrightarrow} r_1(M) \subseteq r_2(M) \text{ for every } M \in R\text{-Mod}.$$

For the preradicals of special types this relation can be expressed by associated classes of modules. In particular:

– for idempotent preradicals

$$r_1 \leq r_2 \iff \mathcal{R}(r_1) \subseteq \mathcal{R}(r_2);$$

– for radicals

$$r_1 \leq r_2 \iff \mathcal{P}(r_1) \supseteq \mathcal{P}(r_2);$$

– for idempotent radicals

$$r_1 \leq r_2 \iff \mathcal{R}(r_1) \subseteq \mathcal{R}(r_2) \Longleftrightarrow \mathcal{P}(r_1) \supseteq \mathcal{P}(r_2).$$

2 Preradicals associated to functor *H*

Let $U \in R$-Mod be an arbitrary left R-module and consider the functor $H = Hom_R(U,\text{-}) : R\text{-}Mod \to \mathcal{A}b$, where $\mathcal{A}b$ is the category of abelian groups. We denote:

$$Gen\,({}_RU) = \{M \in R\text{-}Mod \,|\, \text{there exists an epi } U^{(\mathfrak{A})} \to M \to 0\},$$

i.e. $Gen\,({}_RU)$ is the class of modules generated by the fixed module ${}_RU$. It is clear that the class $Gen\,({}_RU)$ is closed under homomorphic images and direct sums, so it is a pretorsion class. We define by ${}_RU$ the function r^U as follows:

$$r^U(M) = \sum_{f:U\to M} Im\, f, \quad M \in R\text{-}Mod,$$

i.e. $r^U(M)$ is the *trace* of ${}_RU$ in ${}_RM$ for every $M \in R$-Mod. The following fact is obvious.

Proposition 2.1. *For every module $U \in R$-Mod the function r^U is an idempotent radical of R-Mod, determined by the class of r^U-torsion modules*: $\mathcal{R}(r^U) = Gen\,({}_RU)$. $\square$

For the functor H we denote:

$$Ker\,H = \{M \in R\text{-}Mod \,|\, H(M) = 0\}.$$

From the definition of operator $(\ )^{\downarrow}$ it follows for the class $\mathcal{K} = \{{}_RU\}$ that $\mathrm{Ker}\, H = \{{}_RU\}^{\downarrow}$. From the properties of the functor H we have

Proposition 2.2. *$Ker\, H$ is a torsionfree class, i.e. it is closed under submodules, direct products and extensions.* $\square$

Therefore the class $\mathrm{Ker}\, H$ defines an *idempotent radical* $\overline{r}^U$ such that $\mathcal{P}(\overline{r}^U) \stackrel{\text{def}}{=\!=} Ker\, H \ \ (= \{{}_RU\}^{\downarrow})$. The respective torsion class for $\overline{r}^U$ is:

$$\mathcal{R}(\overline{r}^U) = (Ker\, H)^{\uparrow} = \{{}_RU\}^{\downarrow\uparrow} = (Gen\,({}_RU))^{\downarrow\uparrow}.$$

Since $\mathcal{R}(r^U) = Gen\,({}_RU)$, it follows that $\mathcal{R}(\overline{r}^U)$ is the least torsion class containing $\mathcal{R}(r^U)$. In the language of preradicals this means the following.

Proposition 2.3. *For every module $U \in R\text{-}Mod$ we have $r^U \leq \overline{r}^U$ and $\overline{r}^U$ is the least idempotent radical, containing r^U.* $\square$

Now we will investigate the question when these preradicals coincide: $r^U = \overline{r}^U$. For that we introduce the following notion.

Definition 1. A module ${}_RU$ will be called *weakly projective* if the functor $H = Hom_R(U, \text{-}) : R\text{-}Mod \to \mathcal{A}b$ preserves the exactness of the short exact sequences of the form:

$$0 \to r^U(M) \underset{\subseteq}{\xrightarrow{i}} M \underset{\text{nat}}{\xrightarrow{\pi}} M\,/\,r^U(M) \to 0$$

for every module $M \in R$-Mod, where i is the inclusion and π is the natural epimorphism.

In other words, ${}_RU$ is weakly projective if for every $M \in R$-Mod and every R-morphism $f : U \to M\,/\,r^U(M)$ there exists an R-morphism $g : U \to M$ such that $\pi g = f$ (π is natural morphism):

$$\begin{tikzcd} & {}_RU \arrow[dl, dashed, "g"'] \arrow[dr, "f"] & \\ M \arrow[rr, "\pi"] & & M/r^u(M) \end{tikzcd}$$

Fig. 1

Proposition 2.4. *For the module ${}_RU$ the following conditions are equivalent:*

1) $r^U = \overline{r}^U$;
2) r^U *is an (idempotent) radical*;
3) $Gen({}_RU) = (Ker\, H)^{\uparrow} \ \ (= \{{}_RU\}^{\downarrow\uparrow})$;
4) ${}_RU$ *is a weakly projective module.*

Proof. 1) $\Longleftrightarrow$ 2) $\Longleftrightarrow$ 3) follows from Proposition 2.3.
2) $\Rightarrow$ 4). If r^U is a radical, then for every $M \in R$-Mod we have:

$$M\,/\,r^U(M) \in \mathcal{P}(r^U) = \mathcal{P}(\overline{r}^U) = Ker\, H,$$

therefore $Hom_R(U, M) \,/\, r^U(M) = 0$ and that implies immediately that ${}_RU$ is weakly projective ($f = 0 \Rightarrow g = 0$).
4) $\Rightarrow$ 2). Let ${}_RU$ be weakly projective and we verify that $r^U(M \,/\, r^U(M)) = 0$ for every $M \in R$-Mod. Consider an arbitrary R-morphism $f : U \to M \,/\, r^U(M)$. From condition 4) it follows that there exists a morphism $g : U \to M$ such that $\pi\, g = f$. Since $Im\, g \subseteq r^U(M)$ by definition of $r^U(M)$, we have $\pi\, g = 0$ and $f = 0$. So $Hom_R(U, M \,/\, r^U(M)) = 0$, i.e. $r^U(M \,/\, r^U(M)) = 0$ and r^U is a radical. □

Examples. 1) If ${}_RU$ is a projective module, then it is weakly projective, therefore $r^U = \overline{r}^U$.
2) If ${}_RU$ is a generator of R-Mod, then $Gen\,({}_RU) = R$-Mod, so $r^U = \overline{r}^U = \mathbf{1}$, where $\mathbf{1}$ is the greatest trivial preradical of R-Mod ($\mathbf{1}(M) = M$ for every $M \in R$-Mod).

More strong than the conditions of Proposition 2.4 is the request that the idempotent preradical r^U must be a cotorsion. To indicate when such situation takes place we need the

Definition 2 [1]. A module ${}_RU$ will be called *cohereditary below* if the class $\{{}_RU\}^{\downarrow}$ is cohereditary (*i.e. a TTF-class*).

This means that if $Hom_R(U, M) = 0$ for a module $M \in R$-Mod, then $Hom_R(U, M/N) = 0$ for every submodule $N \subseteq M$.

From Proposition 2.4 and definitions follows

Proposition 2.5. *For a module ${}_RU$ the following conditions are equivalent:*
1) *r^U is a cotorsion*;
2) *$r^U = \overline{r}^U$ and the class $\mathcal{P}(\overline{r}^U) = Ker\, H$ is cohereditary*;
3) *${}_RU$ is weakly projective and cohereditary below.* □

It is obvious that if the module ${}_RU$ is projective, then r^U is a cotorsion.

3 Preradicals defined by trace-ideal $\boldsymbol{I = Trace_U({}_RR)}$

For a fixed module $U \in R$-Mod we consider its trace in ${}_RR$:

$$I = r^U({}_RR) = \sum_{f : U \to R} Im\, f,$$

which is a two-sided ideal of R. It defines the following three classes of modules (see [6]):

$${}_I\mathcal{T} = \{M \in R\text{-Mod} \mid IM = M\},$$

$${}_I\mathcal{F} = \{M \in R\text{-Mod} \mid m \in M, Im = 0 \Rightarrow m = 0\},$$

$$\mathcal{A}(I) = \{M \in R\text{-Mod} \mid IM = 0\},$$

i.e. ${}_I\mathcal{T}$ is the class of I-accessible modules, ${}_I\mathcal{F}$ is the class of modules without nonzero elements annihilated by I, and $\mathcal{A}(I)$ consists of the modules annihilated by I.

It is easy to verify the following properties of these classes.

Proposition 3.1. 1) *${}_I\mathcal{T}$ is a torsion class;*
2) *${}_I\mathcal{F}$ is a torsion free and stable class;*
3) *$\mathcal{A}(I)$ is closed under submodules, homomorphic images and direct products (hence also under direct sums). So the class $\mathcal{A}(I)$ is simultaneously a pretorsion and a pretorsionfree class.* □

Therefore the class ${}_I\mathcal{T}$ defines an *idempotent radical* r^I such that:

$$\mathcal{R}(r^I) \stackrel{\text{def}}{=\!=} {}_I\mathcal{T},$$

while the class ${}_I\mathcal{F}$ determines a *torsion* r_I such that:

$$\mathcal{P}(r_I) \stackrel{\text{def}}{=\!=} {}_I\mathcal{F},$$

which is the ideal torsion, defined by I (see [4]).

The class $\mathcal{A}(I)$ as pretorsion (and hereditary) class determines a *pretorsion* $r_{(I)}$ by the rule:

$$\mathcal{R}(r_{(I)}) \stackrel{\text{def}}{=\!=} \mathcal{A}(I).$$

For every $M \in R$-mod we have:

$$r_{(I)}(M) = \{m \in M \mid I \cdot m = 0\}.$$

From the other hand, $\mathcal{A}(I)$ as pretorsionfree (and cohereditary) class defines the *cohereditary radical* $r^{(I)}$ such that:

$$\mathcal{P}(r^{(I)}) \stackrel{\text{def}}{=\!=} \mathcal{A}(I),$$

which acts by the rule:

$$r^{(I)}(M) = IM, \ M \in R\text{-Mod} \ \ (\text{see [2, 4, 6]}).$$

Thus by definitions the idempotent radical r^I has the associated classes:

$$\big({}_I\mathcal{T} = \mathcal{R}(r^I), \quad {}_I\mathcal{T}^{\downarrow} = \mathcal{P}(r^I)\big),$$

while the torsion r_I is defined by the classes:

$$\big({}_I\mathcal{F}^{\uparrow} = \mathcal{R}(r_I), \quad {}_I\mathcal{F} = \mathcal{P}(r_I)\big).$$

In continuation we will indicate a series of relations between the classes of modules mentioned above. They imply the respective connexions between the preradicals defined by these classes.

Proposition 3.2. 1) $\mathcal{A}(I)^{\uparrow} = {}_I\mathcal{T}$; 2) $\mathcal{A}(I)^{\downarrow} = {}_I\mathcal{F}$.

Proof. 1) $(\subseteq)$. Let $M \in \mathcal{A}(I)^{\uparrow}$. Since $M/IM \in \mathcal{A}(I)$, we have $Hom_R(M, M/IM) = 0$, hence $M/IM = 0$ and $M = IM$.

$(\supseteq)$. Let $M \in {}_I\mathcal{T}$. Then for every $N \in \mathcal{A}(I)$ and $f : M \to N$ we have $f(M) = f(IM) = I \cdot f(M) \subseteq I \cdot N = 0$, so $f = 0$. Thus $Hom_R(M, N) = 0$ for every $N \in \mathcal{A}(I)$, i.e. $M \in \mathcal{A}(I)^{\uparrow}$.

2) $(\subseteq)$. Let $M \in \mathcal{A}(I)^{\downarrow}$. If $m \in M$ and $I \cdot m = 0$, then since $Rm \in \mathcal{A}(I)$ we have $Hom_R(Rm, M) = 0$, therefore $Rm = 0$ and $m = 0$. This means that $M \in {}_I\mathcal{F}$.

$(\supseteq)$. Let $M \in {}_I\mathcal{F}$. We consider an arbitrary module $N \in \mathcal{A}(I)$ and an R-morphism $f : N \to M$. For every element $n \in N$ we have:

$$I \cdot f(n) = f(I \cdot n) \subseteq f(IN) = f(0) = 0,$$

and from the assumption $M \in {}_I\mathcal{F}$ now follows $f(n) = 0$, thus $f = 0$. In that way $Hom_R(M, N) = 0$ for every $N \in \mathcal{A}(I)$ and so $M \in \mathcal{A}(I)^{\downarrow}$. □

From the relations of Proposition 3.2 the corresponding connexions between the preradicals defined by ideal I follow. Namely, from ${}_I\mathcal{T} = \mathcal{A}(I)^{\uparrow}$ we obtain ${}_I\mathcal{T}^{\downarrow} = \mathcal{A}(I)^{\uparrow\downarrow}$, therefore the class ${}_I\mathcal{T}^{\downarrow} \big(\stackrel{\text{def}}{=} \mathcal{P}(r^I)\big)$ is the least torsionfree class containing $\mathcal{A}(I)$ $\big(\stackrel{\text{def}}{=} \mathcal{P}(r^{(I)})\big)$.

Similarly, from ${}_I\mathcal{F} = \mathcal{A}(I)^{\downarrow}$ we have ${}_I\mathcal{F}^{\uparrow} = \mathcal{A}(I)^{\downarrow\uparrow}$, therefore ${}_I\mathcal{F}^{\uparrow}$ $\big(\stackrel{\text{def}}{=} \mathcal{R}(r_I)\big)$ is the least torsion class containing $\mathcal{A}(I)$ $\big(\stackrel{\text{def}}{=} \mathcal{R}(r_{(I)})\big)$. Translating this facts in the language of preradicals, associated to these classes, we obtain the following results.

Proposition 3.3. 1) $r^I \leq r^{(I)}$ *and* r^I *is the greatest idempotent radical contained in* $r^{(I)}$.
2) $r_I \geq r_{(I)}$ *and* r_I *is the least idempotent radical containing* $r_{(I)}$. □

Thus we have two pairs of "near" preradicals: $r^I \leq r^{(I)}$ and $r_I \geq r_{(I)}$. It is natural to search the conditions of its coincidence.

Proposition 3.4. *The following conditions are equivalent:*

1) $r^I = r^{(I)}$;
2) $r^{(I)}$ *is idempotent*;
3) $\mathcal{A}(I) = {}_I\mathcal{T}^{\downarrow}$;
4) $r_I = r_{(I)}$;
5) $r_{(I)}$ *is a radical*;
6) $\mathcal{A}(I) = {}_I\mathcal{F}^{\uparrow}$;
7) $I = I^2$.

Proof. Consists in the direct verification (see, for example,[4], p. 22). □

If the equivalent conditions of Proposition 3.4 are fulfilled, then $\mathcal{A}(I)$ is TTF-class, r^I is a cotorsion defined by the classes $\big({}_I\mathcal{T}, \mathcal{A}(I)\big)$ and r_I is a jansian torsion with the associated classes $(\mathcal{A}(I), {}_I\mathcal{F})$.

4 Relations between preradicals defined by H and preradicals defined by I

In this section we will show that there exists some remarkable connexions between the preradicals $r^U, \overline{r}^U$ of Section 2 and preradicals defined by ideal I (Section 3). For that we clarify firstly the relations between the respective classes of modules. We start by the following remark.

Lemma 4.1. *For every module $M \in R$-Mod we have $IM \subseteq r^U(M)$ (where ${}_RU$ is a fixed module and $I = r^U({}_RR)$).*

Proof. We must verify that $\big(\sum\limits_{f:U\to R} Im\, f\big)\, M \subseteq \sum\limits_{f:U\to M} Im\, g$. For every $f : U \to R$ and $m \in M$ we have the R-morphism

$$g_{(f,m)} : U \to M, \quad g_{(f,m)}(u) \stackrel{\text{def}}{=\!=} f(u)\cdot m, \; u \in U.$$

Since $Im\, g_{(f,m)} = (Im\, f)\cdot m \subseteq \sum\limits_{f:U\to M} Im\, g = r^U(M)$ for every $f : U \to M$ and $m \in M$, we obtain $IM \subseteq r^U(M)$. □

Lemma 4.2. $Ker\, H \subseteq \mathcal{A}(I)$ $\big($*i.e.* $\mathcal{P}(\overline{r}^U) \subseteq \mathcal{P}\big(r^{(I)}\big)$, *hence* $\overline{r}^U \geq r^{(I)}\big)$.

Proof. Let $M \in Ker H$, i.e. $Hom_R(U, M) = 0$. Then $r^U(M) = \sum\limits_{f:U\to M} Im\, g = 0$ and by Lemma 4.1 we have $IM \subseteq r^U(M) = 0$, so $M \in \mathcal{A}(I)$. □

Lemma 4.3. ${}_I\mathcal{T} \subseteq Gen({}_RU)$ $\big($*i.e.* $\mathcal{R}(r^I) \subseteq \mathcal{R}(r^U)$, *hence* $r^I \leq r^U\big)$.

Proof. Let $M \in {}_I\mathcal{T}$, i.e. $IM = M$. From Lemma 4.1 we have $M = IM \subseteq r^U(M)$, thus $M = r^U(M)$. Therefore, $M \in \mathcal{R}(r^U) = Gen({}_RU)$. □

In a schematic form the relations between the preradicals indicated above can be presented as follows:

$$\begin{array}{ccc} & & \overline{r}^U \\ & \swarrow & \downarrow \\ r^U & & r^{(I)} \\ \downarrow & \swarrow & \\ r^I & & \end{array}$$

Fig.2

where the arrow $r_1 \leftarrow r_2$ means $r_1 \leq r_2$.

In the Propositions 2.4 and 3.4 the criterions of coincidences $r^U = \overline{r}^U$ and $r^I = \overline{r}^I$ are indicated. Now we will consider the case when all four preradicals of Fig. 2 coincide.

Proposition 4.4. *The following conditions are equivalent:*

1) $r^U = r^I$ (i.e. $Gen\,({}_RU) = {}_I\mathfrak{T}$);

2) $\overline{r}^U = r^I$ (i.e. $Ker\,H)^{\uparrow} = {}_I\mathfrak{T}$);

3) $\overline{r}^U = r^{(I)}$ (i.e. $Ker\,H = \mathcal{A}(I)$);

4) $r^U = r^{(I)}$;

5) $I\,U = U$.

Proof. We will prove that every condition 1)–4) implies the coincidence of all four preradicals.

1) If $r^U = r^I$, then since r^I is a radical we have that r^U is a radical, so $r^U = \overline{r}^U$ (Proposition 2.4). Therefore $r^I = \overline{r}^U$ and $r^I = r^{(I)} = \overline{r}^U$.

2) If $\overline{r}^U = r^I$ then is obvious that all preradicals coincide.

3) If $\overline{r}^U = r^{(I)}$, then since $\overline{r}^U$ is idempotent, follows that $r^{(I)}$ is idempotent, hence $\overline{r}^I = r^{(I)}$ (Proposition 3.4) and then $\overline{r}^U = r^I$.

4) If $r^U = r^{(I)}$, then r^U is a radical and $r^{(I)}$ is idempotent, therefore $r^U = \overline{r}^U$ and $r^I = r^{(I)}$.

From the previous arguments follows that the conditions 1)–4) are equivalent.

1) $\Rightarrow$ 5). If $r^U = r^I$, then $\mathcal{R}(r^U) = \mathcal{R}(r^I)$, i.e. $Gen\,({}_RU) = {}_I\mathfrak{T}$. Since ${}_RU \in Gen\,({}_RU)$, we have ${}_RU \in {}_I\mathfrak{T}$, i.e. $I\,U = U$.

5) $\Rightarrow$ 1). Let $I\,U = U$, i.e. ${}_RU \in {}_I\mathfrak{T}$. Then $Gen\,({}_RU) \subseteq {}_I\mathfrak{T}$ (because ${}_I\mathfrak{T}$ is a torsion class). From Lemma 4.3 we obtain $Gen\,({}_RU) = {}_I\mathfrak{T}$, thus $r^U = r^I$. □

Corollary 4.5. *If $I\,U = U$, then module ${}_RU$ is weakly projective and $I = I^2$.*

Proof. The conditions of Proposition 4.4 implies in particular $r^U = \overline{r}^U$ and $r^{(I)} = r^I$, therefore ${}_RU$ is weakly projective (Proposition 2.4) and $I = I^2$ (Proposition 3.4). □

Remark. In the previous study do not participate the pair of preradicals $(r_I, r_{(I)})$. In general case the relation between preradicals $\overline{r}^U$ and r_I can be expressed by inclusion $\mathcal{P}(\overline{r}^U) \subseteq \mathcal{R}(r_I)$ (i.e. $Ker\,H \subseteq \mathcal{A}(I)^{\downarrow\uparrow}$). In the case when $I\,U = U$ (Proposition 4.4) we have $\mathcal{P}(\overline{r}^U) = \mathcal{R}(r_I)$, since then $Ker\,H = \mathcal{A}(I) = \mathcal{A}(I)^{\downarrow\uparrow}$.

In conclusion we totalize by the following scheme the relations between all classes of modules studied above (Fig. 3).

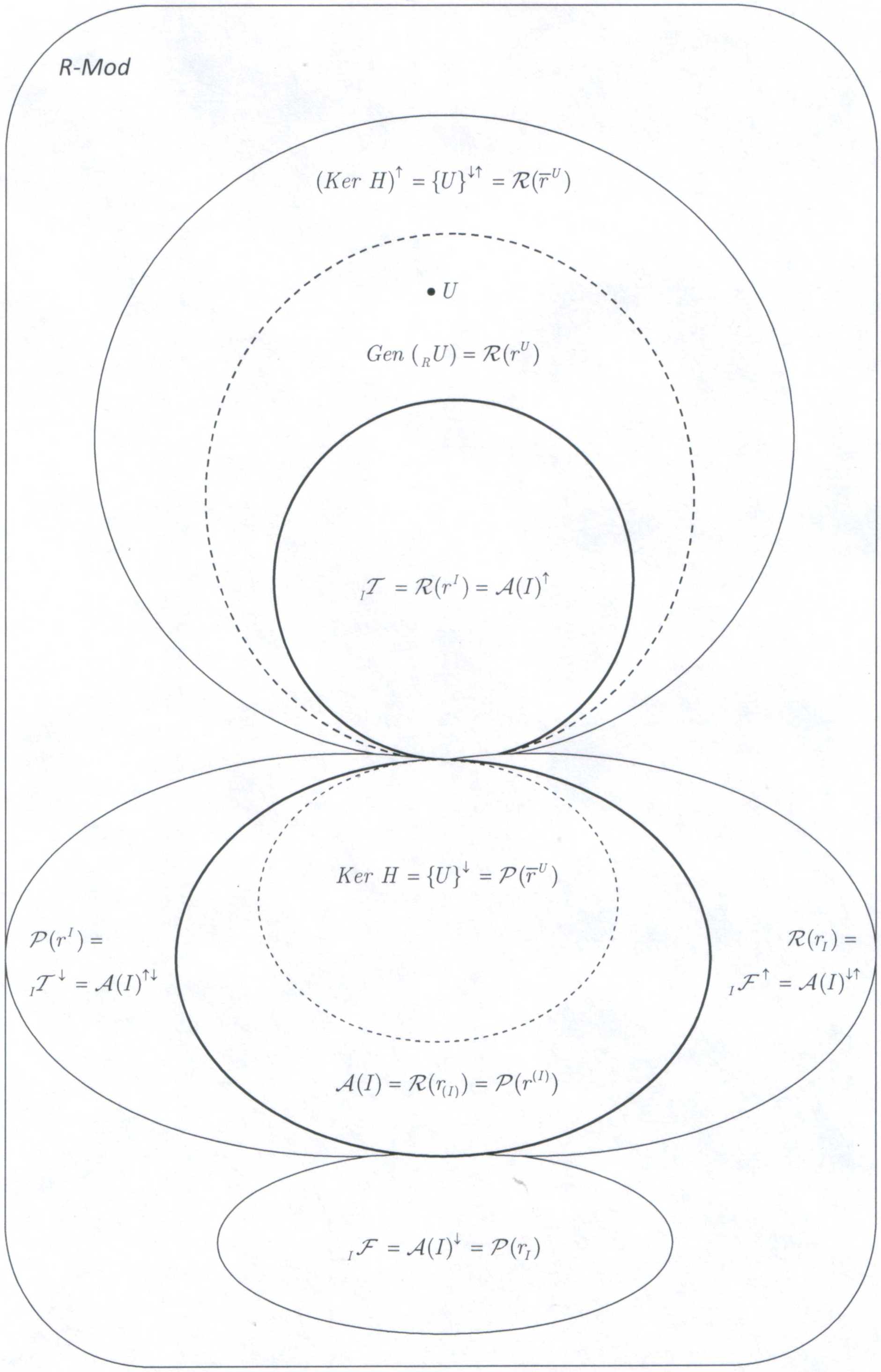

R-Mod
$(Ker\ H)^{\uparrow} = \{U\}^{\downarrow\uparrow} = \mathcal{R}(\overline{r}^{U})$
$\bullet\ U$
$Gen\ ({}_RU) = \mathcal{R}(r^{U})$
${}_I\mathcal{T} = \mathcal{R}(r^{I}) = \mathcal{A}(I)^{\uparrow}$
$Ker\ H = \{U\}^{\downarrow} = \mathcal{P}(\overline{r}^{U})$
$\mathcal{P}(r^{I}) =$
${}_I\mathcal{T}^{\downarrow} = \mathcal{A}(I)^{\uparrow\downarrow}$
$\mathcal{R}(r_I) =$
${}_I\mathcal{F}^{\uparrow} = \mathcal{A}(I)^{\downarrow\uparrow}$
$\mathcal{A}(I) = \mathcal{R}(r_{(I)}) = \mathcal{P}(r^{(I)})$
${}_I\mathcal{F} = \mathcal{A}(I)^{\downarrow} = \mathcal{P}(r_I)$

Fig. 3.

This general situation will be completed after the study of preradicals, associated to the functor of tensor product T, adding two preradicals t^V and $\overline{t}^V$ (dual to r^U and $\overline{r}^U$), connected with r_I and $r_{(I)}$ similar as the pairs $(r^U, \overline{r}^U)$ and $(r^I, \overline{r}^{(I)})$ are connected (see Fig. 2).

References

[1] Kashu A.I., *Functors and torsions in categories of modules.* Acad. of Sciences of RM, Inst. of Mathem., Chişinău, 1997 (in Russian).

[2] Bican L., Kepka P., Nemec P. *Rings, modules and preradicals.* Marcel Dekker, New York, 1982.

[3] Golan J.S. *Torsion theories.* Longman Sci. Techn., New York, 1986.

[4] Kashu A.I. *Radicals and torsions in modules.* Ştiinţa, Chişinău, 1983 (In Russian).

[5] Stenström B. *Rings of quotients.* Springer Verlag, Berlin, 1975.

[6] Kashu A.I. *On some bijections between ideals, classes of modules and preradicals of R-Mod.* Bulet. A.Ş.R.M., Matematica, 2001, No. 2(36), 101–110.

Institute of Mathematics and Computer Science
Academy of Sciences of Moldova
5 Academiei str., Chişinău MD−2028
Moldova

E-mail: *kashuai@math.md*

Received February 20, 2009

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 73–90
ISSN 1024–7696

Dynamic Programming Algorithms for Solving Stochastic Discrete Control Problems

Dmitrii Lozovanu, Stefan Pickl

Abstract. The stochastic versions of classical discrete optimal control problems are formulated and studied. Approaches for solving the stochastic versions of optimal control problems based on concept of Markov processes and dynamic programming are suggested. Algorithms for solving the problems on stochastic networks using such approaches and time-expended network method are proposed.

Mathematics subject classification: 93E20, 49L20.
Keywords and phrases: Time-Discrete System, Optimal Control, Stochastic Networks, Markov Processes, Dynamic Programming.

1 Introduction

The paper is concerned with studying and solving the stochastic versions of the classical discrete optimal control problems from [1,2,5]. In the deterministic control problems from [1,2] the choosing of the vector of control parameters from the corresponding feasible set at every moment of time for an arbitrary state is assumed to be at our disposition, i.e each dynamical state of the system is assumed to be controllable. In this paper we consider the control problems for which the discrete system in the control process may meet dynamical states where the vector of control parameters is changing in a random way according to given distribution functions of the probabilities on given feasible dynamical stats. We call such states of dynamical system uncontrollable dynamical states. So, we consider the control problems for which the dynamics may contain controllable states as well uncontrollable ones. We show that in general form these versions of the problems can be formulated on stochastic networks and new approaches for their solving based on concept of Markov processes and dynamic programming from [3,4] can be suggested. Algorithms for solving the considered stochastic versions of the problems using the mentioned concept and the time-expended network method from [5,6] are proposed and grounded.

2 Problems Formulations and the Main Concept

We consider a time-discrete system L with a finite set of states $X \subset R^n$. At every time-step $t = 0, 1, 2, \dots,$ the state of the system L is $x(t) \in X$. Two states x_0 and x_f are given in X, where $x_0 = x(0)$ represents the starting state of system

L and x_f is the state in which the system L must be brought, i.e. x_f is the final state of L. We assume that the system L should reach the final state x_f at the time-moment $T(x_f)$ such that $T_1 \leq T(x_f) \leq T_2$, where T_1 and T_2 are given. The dynamics of the system L is described as follows

$$x(t+1) = g_t(x(t), u(t)), \quad t = 0, 1, 2, \ldots, \tag{1}$$

where

$$x(0) = x_0 \tag{2}$$

and $u(t) = (u_1(t), u_2(t), \ldots, u_m(t)) \in R^m$ represents the vector of control parameters. For any time-step t and an arbitrary state $x(t) \in X$ a feasible finite set $U_t(x(t)) = \{u^1_{x(t)}, u^2_{x(t)}, \ldots, u^{k(x(t))}_{x(t)}\}$, for the vector of control parameters $u(t)$ is given, i.e.

$$u(t) \in U_t(x(t)), \quad t = 0, 1, 2, \ldots. \tag{3}$$

We assume that in (1) the vector functions $g_t(x(t), u(t))$ are determined uniquely by $x(t)$ and $u(t)$, i.e. the state $x(t+1)$ is determined uniquely by $x(t)$ and $u(t)$ at every time-step $t = 0, 1, 2, \ldots$. In addition we assume that at each moment of time t the cost $c_t\big(x(t), x(t+1)\big) = c_t\big(x(t), g_t(x(t), u(t))\big)$ of system's transaction from the state $x(t)$ to the state $x(t+1)$ is known.

Let $x_0 = x(0), x(1), x(2), \ldots, x(t), \ldots$ be a trajectory generated by given vectors of control parameters $u(0), u(1), \ldots, u(t-1), \ldots$. Then either this trajectory passes through the state x_f at the time-moment $T(x_f)$ or it does not pass through x_f.We denote

$$F_{x_0 x_f}(u(t)) = \sum_{t=0}^{T(x_f)-1} c_t\big(x(t), g_t(x(t), u(t))\big) \tag{4}$$

the integral-time cost of system's transactions from x_0 to x_f if $T_1 \leq T(x_f) \leq T_2$; otherwise we put $F_{x_0 x_f}(u(t)) = \infty$. In [1, 2, 5] have been formulated and studied the following problem: to determine the vectors of control parameters $u(0), u(1), \ldots, u(t), \ldots$ which satisfy conditions (1)-(3) and minimize functional (4). This problem can be regarded as a control model with controllable states of dynamical system because for an arbitrary state $x(t)$ at every moment of time the choosing of vector of control parameter $u(t) \in U_t(x(t))$ is assumed to be at our disposition. In the following we consider the stochastic versions of the control model formulated above. We assume that the dynamical system L may contain uncontrollable states, i.e. for the system L there exist dynamical states in which we are not able to control the dynamics of the system and the vector of control parameters $u(t) \in U_t(x(t))$ for such states is changing in the random way according to given distribution function

$$p : U_t(x(t)) \to [0, 1], \qquad \sum_{i=1}^{k(x(t))} p(u^i_{x(t)}) = 1 \tag{5}$$

on the corresponding dynamical feasible sets $U_t(x(t))$. If we regard arbitrary dynamic state $x(t)$ of system L at given moment of time t as position (x, t) then the set of positions

$$Z = \{(x,t)\,|\, x \in X,\ t = 0,1,2,\dots,T_2\}$$

of dynamical system can be divided into two disjoint subsets

$$Z = Z^C \cup Z^N \quad (Z^C \cap Z^N = \emptyset),$$

where Z^C represents the set of controllable positions of L and Z^N represents the set of positions $(x,t) = x(t)$ for which the distribution function (5) of the vectors of control parameters $u(t) \in U_t(x(t))$ are given. This mean that the dynamical system L works as follows. If the starting point belongs to controllable positions then the decision maker fixes a vector of control parameter and we obtain the state $x(1)$. If the starting state belongs to the set of uncontrollable positions then the system passes to the next state in a random way. After that if at the time-moment $t = 1$ the state $x(1)$ belong to the set of controllable positions then the decision maker fixes the vector of control parameter $u(t) \in U_t(x(t))$ and we obtain the state $x(2)$. If $x(1)$ belongs to the set of uncontrollable positions then the system passes to the next state in a random way and so on. In this dynamic process the final state may be reached at given moment of time with a probability which depend on the control of the system in the deterministic states as well as the expectation of integral time cost by trajectory depends on control of the system in these states. The main results of this paper are concerned with studying and solving the following problems.

1. For given vectors of control parameters $u(t) \in U_t(x(t))$, $x(t) \in Z^C$, to determine the probability that the dynamical system L with given starting state $x_0 = x(0)$ will reach the final state x_f at the moment of time $T(x_2)$ such that $T_1 \le T(x_f) \le T_2$. We denote this probability $P_{x_0}\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big)$; if $T_1 = T_2 = T$ then we use the the notation $P_{x_0}(u(t), x_f, T)$.

2. To find the vectors of control parameters $u^*(t) \in U_t(x(t))$, $x(t) \in Z^C$ for which the probability in problem 1 is maximal. We denote this probability we denote $P_{x_0}\big(u^*(t), x_f, T_1 \le T(x_f) \le T_2\big)$; in the case $T_1 = T_2 = T$ we shall use the notation $P_{x_0}(u^*(t), x_f, T)$.

3. For given vectors of control parameters $u(t) \in U_t(x(t))$, $x(t) \in Z^C$ and given number of stages T to determine the expectation of integral-time of system's transactions within T stages for system L with staring state $x_0 = x(0)$. We denote this expectation $C_{x_0}(u(t), T)$.

4. To determine the vectors of control parameters $u^*(t) \in U_t(x(t))$, $x(t) \in Z^C$ for which the expectation of integral-time cost for dynamical system in problem 3 is minimal. We denote this expectation $C_{x_0}(u^*(t), x_f, T)$.

5. For given vectors of control parameters $u(t) \in U_t(x(t))$, $x(t) \in Z^C$, to determine the expectation of integral-time cost of system's transactions from starting state x_0 to final state x_f when the final state is reached at the time-moment $T(x_f)$ such that $T_1 \le T(x_f) \le T_2$. We denote this expectation $C_{x_0}\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big)$; if $T_1 = T_2 = T$ then we denote $C_{x_0}(u(t), x_f, T)$.

6. To determine the vectors of control parameters $u^*(t) \in U_t(x(t))$, $x(t) \in Z^C$ for which the expectation of integral-time cost of system's transactions in problem 5 is minimal. We denote this expectation $C_{x_0}\big(u^*(t), x_f, T_1 \le T(x_f) \le T_2\big)$; in the case $T_1 = T_2 = T$ we shall use the notation $C_{x_0}(u^*(t), x_f, T)$.

It is easy to observe that problems 1-6 extend and generalize a large class of deterministic and stochastic dynamic problems including problems from [1,2,4]. The problems from [4] related to finite Markov processes became problems 1–3 in the case when $Z^C = \emptyset$ and the probabilities $p(u^i_{x(t)})$ do not depend on time but depend only on states; the discrete optimal control problems from [1, 2] became problems 4-6 in the case $Z^N = \emptyset$. In the following we propose algorithms for solving the problem formulated above based on results from [1, 2, 4] and time-expended method from [5, 6].

3 Some Auxiliary Results and Definitions of the Basic Notions

In this section we describe some auxiliary results concerned with calculation of the state probabilities in a simple finite Markov processes and make more precise some basic definitions for our control problems. We shall use these results and the specification of the basic notion we shall use in next sections for a strict argumentation of the algorithms for solving problems 1-6.

3.1 Determining the State Probabilities of the Dynamical System in Finite Markov Processes

We consider a dynamical system with the set of states X where for every state $x \in X$ are given the probabilities $p_{x,y}$ of system's passage from x to another states $y \in X$ such that $\sum\limits_{y \in X} p_{x,y} = 1$. Here the probabilities $p_{x,y}$ do not depend on time, i.e. we have a simple Markov process determined by the stochastic matrix of probabilities $P = (p_{x,y})$ and the starting state x_0 of dynamical system. The probability $P_{x_0}(x, t)$ of system's passage from the state x_0 to an arbitrary state $x \in X$ by using given t unite of time is defined and calculated on the basis of the following recursive formula [4]

$$P_{x_0}(x, \tau + 1) = \sum_{y \in X} P_{x_0}(y, \tau) p_{y,x}, \quad \tau = 0, 1, 2, \dots, t$$

where $P_{x_0}(x_0, 0) = 1$ and $P_{x_0}(x, 0) = 0$ for $x \in X \setminus \{x_0\}$. In the case when the probabilities of system's passage from one state to another depend on time we have a non-stationary process defined by a dynamic matrix $P(t) = (p_{x,y}(t))$ which describe this process. If this matrix is stochastic for every moment of time $t = 0, 1, 2, \dots,$ then the state probabilities $P_{x_0}(x, t)$ can be defined and calculated by using a formula obtained similarly from written one changing $p_{x,y}$ by $p_{x,y}(\tau)$.

Now let us show how to calculate the probability of systems passage from the state x_0 to the state x when x is reached at the time moment $T(x)$ such that $T_1 \le T(x) \le T_2$ where T_1 and T_2 are given. So, we are seeking for the probability that the system L will reach the state x at least at one of the moments of time

$T_1, T_1+1, \dots, T_2$. We denote this probability $P_{x_0}(x, T_1 \leq T(x) \leq T_2)$. For this reason we shall give the graphical interpretation of the simple Markov processes by using the random graph of state transitions $GR = (X, ER)$. In this graph each vertex $x \in X$ corresponds to a state of dynamical system and a possible system passage from one state x to another state y with positive probability $p_{x,y}$ is represented by the directed edge $e = (x, y) \in ER$ from x to y; to directed edges $(x, y) \in ER$ in G the corresponding probabilities $p_{x,y}$ are associated. It is evident that in the graph GR each vertex x contains at least one leaving edge (x, y) and $\sum\limits_{y \in X} p_{x,y} = 1$. In general we will consider also the stochastic process which may stop if one of the states from given subset of states of dynamical system is reached. This means that the random graph of such process may contain the deadlock vertices. So, we consider the stochastic process for which the random graph may contains the deadlock vertices and $\sum\limits_{y \in X} p_{x,y} = 1$ for the vertices $x \in X$ which contain at least one leaving directed edge. Such random graphs do not correspond to Markov processes and the matrix of probability P contains rows with zero components. Nevertheless the probabilities $P_{x_0}(x, t)$ in the both cases of the considered processes can be calculated on the basis of recursive formula given above. In the next sections we can see that the state probabilities of the system can be also calculated starting from final state by using the backward dynamic procedure.In the following the random graph with given probability function $p : ER \to R$ on edge set ER and given distinguished vertices which correspond to starting and final states of dynamical system we be called the stochastic network. Further we shall use the stochastic networks for calculation of the probabilities $P_{x_0}(x, T_1 \leq T(x) \leq T_2)$.

Lemma 1. *Let be given a Markov process determined by stochastic matrix of probabilities $P = (p_{x,y})$ and the starting state x_0. Then the following formula holds:*

$$P_{x_0}(x, T_1 \leq T(x) \leq T_2) = P_{x_0}(x, T_1) + P_{x_0}^{T_1}(x, T_1+1) +$$
$$+ P_{x_0}^{T_1, T_1+1}(x, T_1+1) + \dots + P_{x_0}^{T_1, T_1+1, \dots, T_2-1}(x, T_2) \tag{6}$$

where $P_{x_0}^{T_1, T_1+1, \dots, T_1+i-1}(x, T_1+i), i = 1, 2, \dots, T_2 - T_1$, is the probability that the system L will reach the state x from x_0 by using $T_1 + i$ transactions and it does not pass through x at the moments of times $T_1, T_1+1, T_1+2, \dots, T_1+i-1$.

Proof. Taking into account that $P_{x_0}(x, T_1 \leq T(x) \leq T_1 + i$ expresses the probability of the system L to reach from x_0 the state x at least at one of the moments of time $T_1, T_1+1, \dots, T_1+i$ we can write the following recursive formula

$$P_{x_0}(x, T_1 \leq T(x) \leq T_1 + i) = P_{x_0}(x, T_1 \leq T(x) \leq T_1 + i - 1) +$$

$$+ P_{x_0}^{T_1, T_1+1, \dots, T_1+i-1}(x, T_1+i). \tag{7}$$

Applying $T_2 - T_1$ times this formula for $i = 1, 2, \dots, T_2 - T_1$ we obtain the equality (6). $\square$

Note that formula 6 and 7 couldn't be used directly for calculation of the probability $P_{x_0}(x, T_1 \leq T(x) \leq T_2)$. Nevertheless we can see that such representation of the probability $P_{x_0}(x, T_1 \leq T(x) \leq T_2)$ in the time expended network method will allow to ground a suitable algorithms for calculation of this probability and to develop new algorithms for solving problems from Section 2.

Corollary 1. *If the state x of dynamical system L in random graph $GR = (X, ER)$ corresponds to a deadlock vertex then*

$$P_{x_0}\big(x, T_1 \leq T(x) \leq T_2\big) = \sum_{t=T_1}^{T_2} P_{x_0}(x, t). \tag{8}$$

Let X_f be a subset of X and assume that at the moment of time $t = 0$ the system L is in the state x_0. Denote by $P_{x_0}\big(X_f, T_1 \leq T(X_f) \leq T_2\big)$ the probability that at least one of the states $x \in X_f$ will be reached at the time moment $T(x)$ such that $T_1 \leq T(x) \leq T_2$. Then the following corollary holds.

Corollary 2. *If the subset of states $X_f \subset X$ of dynamical system L in the random graph $GR = (X, ER)$ corresponds to the subset of deadlock vertices then for the probability $P_{x_0}\big(X_f, T_1 \leq T(X_f) \leq T_2\big)$ the following formula holds*

$$P_{x_0}\big(X_f, T_1 \leq T(X_f) \leq T_2\big) = \sum_{x \in X_f} \sum_{t=T_1}^{T_2} P_{x_0}(x, t). \tag{9}$$

3.2 Determining the Expectation of Integral-time cost of system's transactions in Finite Markov Processes

In order to define strictly the expectation of integral-time cost for dynamical system in problems 3-6 we need to introduce the notion of expectation of integral-time cost for finite Matkov processes with cost function on the set of state's transaction of dynamical system. We introduce this notion we introduce in the same way as the total expected earning in the Markov processes with rewards introduced in [4]. We consider a simple Marcov process determined by the stochastic matrix $p = (p_{x,y})$ and starting state x_0 of system L. Assume that for arbitrary two states $x, y \in X$ of the dynamical system is given the value $c_{x,y}$ which we treat as the cost of system L to pass from the state x to the state y. The matrix $C = (c_{x,y})$ is called the matrix of the costs of system's transactions for the dynamical system. Note that in [4] the values $c_{x,y}$ for given x are treated as the "earning" of the system's transaction from the state x to the states $y \in X$ and the corresponding Markov process with associated matrix C is called Markov process with reward. The Markov process with associated cost matrix C generates a sequence of costs when the system makes transactions from one state to another. Thus the cost is a random variable with a probability distribution induced by the probability relations of the Markov process. This means that for the system L the integral-time cost during T transactions is a random variable for which the expectation can be defined. We denote the expectation of integral-time cost in such process by $C_{x_0}(T)$. So, $C_{x_0}(T)$ expresses the expected integral-time cost of the system in the next T transactions if the system at the starting moment of time is in the state $x_0 = x(0)$. For an arbitrary $x \in X$ the values $C_x(\tau)$ are defined strictly and calculated on the basis of the following recursive formula

$$C_x(\tau) = \sum_{y \in Y} p_{x,y}(c_{x,y} + C_y(\tau - 1)), \quad \tau = 1, 2, \ldots, t$$

where $C_x(0) = 0$ for every $x \in X$. This formula can be treated in the similar way as formula for calculation the total earning in the Markov processes with rewards [4]. The expression $c_{x,y} + C_y(\tau - 1)$ means that if the system makes transaction from the state x to the state y then it spends the amount $c_{x,y}$ plus the amount it expects to spend during the next $\tau - 1$ transactions when the system start transactions in the state y at the moment of time $\tau = 1$. Taking into account that in the state x the system makes transactions in the random way with the probability distribution $p_{x,y}$ we obtain that the values $c_{x,y} + C_y(\tau - 1)$ should be weighted by the probabilites of transactions $p_{x,y}$. In the case of the non-stationary process, i.e. when the probabilities and the costs are changing in time, the expectation of integral-time cost of dynamical system is defined and calculated in similar way; in formula written above we should change $p_{x,y}$ by $p_{x,y}(\tau)$ and $c_{x,y}$ by $c_{x,y}(t)$.

3.3 Definition of the State Probability and The Expectation of Integral time cost in Control Problems 1–6

Using the definitions from previous subsections we can now specify the notions of state probabilities $P_{x(0)}(u(t), x, T)$, $P_{x(0)}\big(u(t), x_f, T_1 \leq T(x_f) \leq T_2\big)$ and the expectations of integral-time cost $C_{x(0)}(u(t), T)$, $C_{x(0)}\big(u(t), x_f, T\big)$, $C_{x(0)}\big(u(t), x_f, T_1 \leq T(x_f) \leq T_2\big)$ in problems 1-6. First of all we stress our attention to the definition of probability $P_{x_0}(u(t), x, T)$. For given starting state x_0, given time-moment T and fixed control $u(t)$ we define this probability in the following way. We consider that each system passage from an controllable state $x = x(t)$ to the next state $y = x(t + 1)$ generated by the control $u(t)$ is made with probability $p_{x,y} = 1$ and the rest of probabilities of system's passages from x at the moment of time t to the next states are equal to zero. Thus we obtain a finite Markov process for which the probability of system passage from starting state x_o to final state x by using T unites of time can be defined. We denote this probability $P_{x_0}(u(t), x, T)$. We define the probability $P_{x_0}\big(u(t), x, T_1 \leq T(x) \leq T_2\big)$ for given T_1 and T_2 as probability of the dynamical system L to reach the state x at least at one of the moments of time $T_1, T_1 + 1, \ldots, T_2$. In order to define strictly the expectation of integral-time cost of dynamical system in problems 3-6 we shall use the notion of expectation of integral-time cost for Markov processes with costs defined on system's transactions. The expectation of integral-time cost $C_{x_0}(u(t), T)$ of system L in problem 3 for fixed control $u(t)$ is defined as the expectation of the integral-time cost during

T transitions of dynamical system in the Markov process generated by the control $u(t)$ and the corresponding costs of state's transactions of dynamical system. The expectation $C_{x_0}\big(u(t), x, T_1 \le T(x) \le T_2\big)$ in the problems 5 and 6 will made more precise in more detail form in Section 5.

4 The Main Approach and Algorithms for Determining the State Probabilities in the Control Problems on Stochastic Networks

In order to provide a better understanding of the main approach and to ground the algorithms for solving the problems formulated in Section 2 we shall use the network representation of the dynamics of the system and will formulate these problems on stochastic network. Note that in our control problems the probabilities and the costs of system's passage from one state to another depend on time. Therefore here we develop time-expended network method from [5, 6] for the stochastic versions of control problems and reduce them to the static cases of the problems. At first we show how to construct the stochastic network and how to solve the problems with fixed number of stages, i.e. we consider the case $T_1 = T_2 = T$.

4.1 Construction of Stochastic Network and Algorithms for Solving the Problems in the Case $T_1 = T_2 = T$

If the dynamics of discrete system L and the information related to the feasible sets $U_t(x(t))$ and the cost functions $c_t(x(t), g_t(x(t), u(t)))$ in the problems with $T_1 = T_2 = T$ are known then our stochastic network can be obtained in the following way. We identify each position (x, t) which correspond to a dynamic state $x(t)$ with a vertex $z = (x, t)$ of the network. So, the set of vertices Z of the network can be represented as follows $Z = Z_1 \cup Z_2 \cup \cdots \cup Z_T$ where $Z_t = \{(x, t)\,|\, x \in X\}, t = 0, 1, 2, \ldots, T$. To each vector of control parameters $u(t) \in U_t(x(t)), t = 1, 2, \ldots, T-1$ which provide a system passage from the state $x(t) = (x, t)$ to the state $x(t+1) = (y, t+1)$ we associate in our network a directed edge $e(z, w) = \big((x, t), (y, t+1)\big)$ from the vertex $z = (x, t)$ to the vertex $w = (y, t+1)$, i.e., the set of edges E of the network is determined by the feasible sets $U_t(x(t))$. After that to each directed edge $e = (z, w) = \big((x, t), (y, t+1)\big)$ originating in uncontrollable positions (x, t) we put in correspondence the probability $p(e) = p\,(u(t))$, where $u(t)$ is the vector of control parameter which provide the passage of the system from the state $x = x(t)$ to the state $x(t+1) = (y, t+1)$. Thus if we distinguish in E the subset of edges $E_N = \big\{e = (z, w) \in E | z \in Z^N\big\}$ originating in uncontrollable positions Z^N then on E_N we obtain the probability function $p : E \to R$ which satisfies the condition

$$\sum_{e \in E^+(z)} p(e) = 1, \quad z \in Z^N \setminus Z_T$$

where $E^+(z)$ is the set of edges originating in z. In addition in the network we add to the edges $e = (z, w) = \big((x, t), (y, t+1)\big)$ the costs $c(z, w) = c\big((x, t), (y, t+1)\big) = c_t\big(x(t), x(t+1)\big)$ which correspond to the costs of system's passage from states $x(t)$ to the states $x(t+1)$. The subset of edges of the graph G originating in vertices

$z \in Z^C$ is denoted E_C, i.e. $E_C = E \setminus E_N$. So, our network is determined by the tuple $(G, Z^C, Z^N, z_0, z_f, c, p, T)$, where $G = (Z, E)$ is the graph which describes the dynamics of the system; the vertices $z_0 = (x, 0)$ and $z_f = (x_f, 0)$ correspond to the starting and the final states of the dynamical system, respectively; c represents the cost function defined on the set of edges E and p is the probability function defined on the set of edges E_N which satisfy condition (5). Note that $Z = Z^C \cup Z^N$, where Z^C is a subset of vertices of G which correspond to the set of controllable positions of dynamical system and Z^N is a subset of vertices of G which correspond to the set of uncontrollable positions of system L. In addition we shall use the notation Z_t^C and Z_t^N, where $Z_t^C = \big\{(x,t) \in Z_t | (x,t) \in Z^C\big\}$ and $Z_t^N = \big\{(x,t) \in Z_t | (x,t) \in Z^C\big\}$.

It is easy to observe that after the construction described above the problem 1 in the case $T_1 = T_2 = T$ can be formulated and solved on stochastic network $(G, Z^C, Z^N, z_0, z_f, p, T)$. A control $u(t)$ of system L in this network means a fixed passage from each controllable position $z = (x, t)$ to the next position $z = (x, t)$ through a leaving edge $e = (z, w) = \big((x,t),(y,t+1)\big)$ generated by $u(t)$; this is equivalent with the prescription to these leaving edges the probability $p(e) = 1$ of the system's passage from the state (x, t) to the state $(y, t+1)$ considering $p(e) = 0$ for the rest of leaving edges. In other words a control on stochastic network means an extension of the probability function p from E_N to E by adding to the edges $e \in E \setminus E_N$ the probabilities $p(e)$ according to the mentioned above rule. We denote this probability function on E by p_u and will keep in mind that $p_u(e) = p(e)$ for $e \in E \setminus E_N$ and on E_C this function satisfies the following property

$$p_u : E_C \to \{0,1\}\,, \qquad \sum_{e \in E_C^+(z)} p_u(e) = 1 \text{ for } z \in Z^C$$

induced by the control $u(t)$ in the problems 1–6. If for the problems from section 2 the control $u(t)$ is given then we denote the stochastic network $(G, Z^C, Z^N, z_0, z_f, c, p_u, T)$; If the control $u(t)$ is not fixed then for the stochastic network we shall use the notation $(G, Z^C, Z^N, z_o, z_f, c, p, T)$. For the state probabilities of the system L on this stochastic network we shall use similar notations $P_{z_0}(u(t), z, T), P_{z_0}\big(u(t), z, T_1 \le T(z) \le T_2\big)$ and each time we will specify on which network they are calculated, i.e. will take into account that these probabilities are calculated by using the probability function on edges p_u which already do not depend on time.

Algorithm 1: *Determining the state probabilities of the system in Problem 1*

Preliminary step (Step 0): Put $P_{z_0}(u(t), z_0, 0) = 1$ for the position $z_0 \in Z$ and $P_{z_0}(u_p, z, t) = 0$ for the positions $z \in Z \setminus \{z_0\}$.

General step (Step $\tau, \tau \ge 1$): For every $z \in Z_\tau$ calculate

$$P_{z_0}(u(t), z, \tau) = \sum_{(w,z) \in E^-(z)} P_{z_0}(u(t), w, \tau - 1) p_u(w, z)$$

where $E^-(z) = \{(w, z) \in E \,|\, w \in Z_{\tau-1}\}$. If $\tau = T$ then stop; otherwise go to the next step.

The correctness of the algorithm follows from definition and network interpretation of the dynamics of system L. In the following we will consider that for the control problems from Section 2 the condition $U_t(x(t)) \neq \emptyset$ for every $x(t) \in X, t = 0, 1, 2, \dots, T_2 - 1$ holds.

Algorithm 2: *Determining the state probability of the system based on backward dynamic programming procedure*

Preliminary step (Step 0): Put $P_{z_f}(u(t), z_f, T) = 1$ for the position $z_f = (x_f, T)$ and $P_z(u(t), z, T) = 0$ for the the positions $z \in Z_T \setminus \{(x_f, T)\}$.

General step (Step $\tau, \tau \geq 1$): For every $z \in Z_{T-\tau}$ calculate

$$P_z(u(t), z_f, T) = \sum_{(z,w)\in E^+(z)} P_w(u(t), z_f, T) p_u(z, w)$$

where $E^+(z) = \{(z, w) \in E \,|\, w \in Z_{\tau+1}\}$. If $\tau = T$ then stop; otherwise go to next step.

Theorem 1. *For given control $u(t)$ Algorithm 2 correctly finds the state probabilities $P_{(x,T-\tau)}(u(t), x_f, T)$ for every $x \in X$ and $\tau = 0, 1, 2, \dots, T$. The running time of the algorithm is $O(|X|^2 T)$.*

Proof. The preliminary step of the algorithm is evident. The correctness of the general step of the algorithm follow from recursive formula at this general step which reflects dynamic programming principle for the state probabilities in simple stochastic process. In order to estimate the running time of the algorithm it is sufficient to estimate the number of elementary operations of general step of the algorithm. It is easy to see that the number of elementary operations for tabulation of state probabilities at the general step is $O(|X|^2)$. Taking into account that the number of steps of the algorithms is T we obtain that the running time of the algorithm is $O(|X|^2 L)$. □

Algorithm 3: *Determining the optimal control for Problem 1 with $T_1 = T_2 = T$*

We describe the algorithm for finding the optimal control $u^*(t)$ and the probabilities $P_{x(T-\tau)}(u^*(t), x_f, T)$ of system's passage from the states $x \in X$ at the moment of time $T-\tau$ to the state x_f by using τ units of time for $\tau = 0, 1, 2, \dots, T$. The algorithm consists of the preliminary, general and final steps. The preliminary and general steps of the algorithm find the values $\pi_{(x_f,T-\tau)}(z_f, T)$ of positions $(x, T-\tau) \in Z$ which correspond to probabilities $P_{x(T-\tau)}(u^*(t), z_f, T)$ of system passages from the state $x(T-\tau) \in X$ at the moment of time $T-\tau$ to the state $x_f(T) \in X$ at the moment of time T when the optimal control $u^*(t)$ is applied. At the end of the last iteration of general step of the algorithm 2 gives the subset of edges $E_C(u^*)$ of E_C which determines the optimal controls. The final step of the algorithm constructs an optimal control $u^*(t)$ of the problem.

Preliminary step (Step 0): Put $\pi_{z_f}(x_f, T) = 1$ for the position $z_f = (x_f, T)$ and $\pi_z(z_f, T) = 0$ for the positions $z \in Z_T \setminus \{(x_f, T)\}$; in addition put $E_C(u^*) = \emptyset$.

General step (Step $\tau \geq 1, \tau \geq 1$): For given τ do items a) and b):

a) For each uncontrollable position $z \in Z_\tau^N$ calculate

$$\pi_z(z_f, T) = \sum_{(z,w)\in E^+(z)} \pi_w(x_f, T) p(z, w);$$

b) For each controllable position $z \in Z^C_\tau$ calculate

$$\pi_z(z_f, T) = \max_{(z,w)\in E^+(x,T-\tau)} \pi_w(z_f, T)$$

and include in E^*_C edges (z, w) which satisfy the condition $\pi_z(z_f, T) = \pi_w(z_f, T)$. If $\tau = T$ then go to Final step; otherwise go to step $\tau + 1$.

Final Step: Form the graph $G^* = \big(Z, E^*_C \cup (E \setminus E_C)\big)$ and fix in G^* a map

$$u^* : (x, t) \to (y, t + 1) \in X_{G^*}(x, t) \text{ for } (x, t) \in Z^C$$

where $X_{G^*} = \big\{(y, t + 1) \in Z | \big((x, t), (y, t + 1)\big) \in E^*_C\big\}$.

Theorem 2. *Algorithm 3 correctly finds the optimal control $u^*(t)$ and the state probability $P_{x(0)}(u^*(t), x_f, T)$ for an arbitrary starting position $x(0) \in X$ in problem 1 with $T = T_1 = T_2$. The running time of the algorithm is $O(|X|^2T)$.*

Proof. The general step of the algorithm reflects the principle of optimality of dynamic programming for the problem of determining the control with maximal probabilities $P_{x(T-\tau)}(u^*(t), x_f, T) = \pi_{(x,T-\tau)}(x_f, T)$ of system's passages from the states $x \in X$ at the moment of time $T - \tau$ to the final state at the moment of time T. For each controllable position $(x, T - \tau) \in Z$ the values $\pi_{(x,T-\tau)}(z_f, T)$ are calculated on stochastic network in consideration that for given moment of time $T - \tau$ and given state $x \in X$ the optimal control $u^*(T - \tau) \in U_t(x(T - \tau))$ is applied. The computational complexity of the algorithm can be estimated in the same way as in Algorithm 2. Algorithm makes T steps and at each step uses $O(|X|^2)$ elementary operations. Therefore the running time of the algorithm is $O(|X|^2T)$ □

4.2 Algorithm for determining the state probabilities in the case $T_1 \neq T_2$

We construct our network using the network $(G, Z^C, Z^N, z_0, z_f, c, p, T)$ with $T = T_2$ obtained according to the construction from Subsection 3.1. In this network we delete all edges originating in vertices (x, t) for $t = T_1, T_1+1, \dots, T_2-1$ preserving edges originating in vertices (x, t) for $t = 0, 1, 2, \dots, T_1-1$. We denote the stochastic network in this case $(G^0, Z^C, Z^N, z_0, Y, c, p, T_1, T_2)$, where $Y = \big\{(x_f, T_1), (x_f, T_1 + 1), \dots, (x_f, T_2)\big\}$ and $G^0 = (Z, E^0)$ is the graph obtained from G dy deleting all edges which originate in vertices from Y, i.e $E^0 = E \setminus \{(z, w) \in E | z \in Y\}$. Let $P_{z_0}\big(u(t), Y, T_1 \le T(Y) \le T_2\big)$ be the probability of dynamical system to reach at least one of the states $(x_f, T_1), (x_f, T_1 + 1), \dots, (x_f, T_2)$ at the moment of time t such that $T_1 \le t \le T_2$ if the dynamical system at the moment of time $\tau = 0$ has the state x_0.

Theorem 3. *For an arbitrary feasible control $u(t)$ and given staring state x_0 of dynamical system L the following formula holds*

$$P_{x_0}(u(t), x_f, T_1 \le T(x_f) \le T_2) = P_{z_0}(u(t), Y, T_1 \le T(Y) \le T_2). \tag{10}$$

Proof. We prove the theorem by using induction principle on the number $k = T_2 - T_1$. Let us prove formula (10) for $k = 1$. In this case our network $(G^0, Z^C, Z^N, z_0, Y, c, p_u, T_1, T_2)$ is obtained from $(G, Z^C, Z^N, z_0, z_f, c, p, T_2)$ by deleting the edges $((x_f, T_1), (x_f, T_1 + 1))$ originating in (x_f, T_1). For this network we have $T_2 = T_2 + 1$ and $Y = (x_f, T_1), (x_f, T_1 + 1)$. Basing on formula (6) we can write the following equality

$$P_{x_0}(u(t), x_f, T_1 \leq T(x_f) \leq T_2) = P_{x_0}u(t), x_f, T_1) + P_{x_0}^{T_1, T_1+1}(u(t), x_f, T_1 + 1),$$

where $P_{x_0}^{T_1, T_1+1}(u(t), x_f, T_1 + 1)$ for given control $u(t)$ represents the probability of the system L to reach the state x_f from x_0 such that it does not pass at the moment of time T_1 through x_f. Taking into account that in our network all edges originating in (x_f, T_1) are deleted we obtain

$$P_{x_0}^{T_1, T_1+1}(u(t), x_f, T_1 + 1 = P_{z_0}(u(t), (x_f, T_1 + 1), T_1 + 1).$$

This means that

$$P_{x_0}(u(t), x_f, T_1 \leq T(x_f) \leq T_2) =$$
$$= P_{z_0}(u(t), (x_f, T_1), T_1) + P_{z_0}(u(t), (x_f, T_1 + 1), T_1 + 1).$$

If we use the property from Corollary 2 then we obtain formula (10) for $k = 1$.

Now assume that formula (10) holds for an arbitrary $k \geq 1$ an let us prove that it is true for $k + 1$.

We apply formula (7) for $P_{x_0}(u(t), x_f, T_1 \leq T(x) \leq T_2)$. Then we obtain

$$P_{x_o}(u(t), x_f, T_1 \leq T(x_f) \leq T_1 + k + 1) =$$

$$= P_{x_0}(u(t), x_f, T_1 \leq T(x_f) \leq T_1 + k) + P_{x_o}^{T_1, T_1+1, \dots, T_1+k}(u(t), x_f, T_1 + k + 1)$$

where$P_{x_o}^{T_1, T_1+1, \dots, T_1+k}(u(t), x_f, T_1 + k + 1)$ expresses the probability for the system L to reach the state x_f and it does not passes at the moment of time $T_1, T_1+1, \dots, T_1+k$ through the state x_f. According to the assumption of induction principle we can write

$$P_{x_o}(u(t), x_f, T_1 \leq T(x_f) \leq T_1 + k + 1) =$$
$$= P_{z_0}(u(t), Y \setminus (x_f, T_1 + k + 1), T_1 \leq T(Y \setminus (x, T_1 + k + 1) \leq T_1 + k) +$$
$$+ P_{x_o}^{T_1, T_1+1, \dots, T_1+k}(u(t), x_f, T_1 + k + 1).$$

Here in a similar way as in the case $k = 1$ holds

$$P_{x_o}^{T_1, T_1+1, \dots, T_1+k}(u(t), x_f, T_1 + k + 1) = P_{z_o}(u(t), (x_f, T_1 + k + 1), T_1 + k + 1)$$

because the stochastic network $(G^0, Z^C, Z^N, z_0, Y, c, p_u, T_1, T_2)$ is obtained from $(G, Z^C, Z^N, z_0, z_f, c, p, T_2)$ by deleting all edges originating in the vertices (x, T_1), $(x_f, T_1 + 1), \dots, (x_f, T_1 + k)$. So, the following formula holds

$$P_{x_o}(u(t), x_f, T_1 \leq T(x_f) \leq T_1 + k + 1) =$$

$$= P_{z_0}(u(t), Y \setminus (x_f, T_1 + k + 1), T_1 \le T(Y \setminus (x_f, T_1 + k + 1) \le T_1 + k) +$$
$$+ P_{z_o}(u(t), (x_f, T_1 + k + 1), T_1 + k + 1).$$

Now if we use the property from Corollary 2 of Lemma 1 then we obtain formula (10). □

Corollary 3. *For an arbitrary feasible control $u(t)$ and given staring state x_0 of dynamical system L the following formula holds*

$$P_{x_0}\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big) = \sum_{k=0}^{T_2 - T_1} P_{z_0}\big(u(t), (x_f, T_1 + k), T_1 + k\big). \tag{11}$$

Basing on this result we can calculate $P_{x_0}\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big)$ in the following way. We apply Algorithm 1 on network $\big(G_f, Z^C, Z^N, z_0, Y, c, p_u, T_1, T_2\big)$ and determine the state probabilities $P_{z_0}(u(t), (x, \tau), \tau)$ for every $(x, \tau) \in Z$ and $\tau = 0, 1, 2, \dots, T_2$. Then on the basis of formula (11) we find the probability $P_{x_o}\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big)$. We can use this fact for an another algorithm for finding the probability $P_x\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big)$. The algorithm finds the probabilities $P_z\big(u(t), Z_f, T_1 \le T(Y) \le T_2\big)$ on stochastic network $\big(G_f, Z^C, Z^N, z_0, Y, c, p_u, T_1, T_2\big)$ for every $z = (x, T - \tau) \in Z$. Then for $\tau = T$ we obtain $P_{x(T-\tau)}\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big) = P_{(x, T_2 - \tau)}\big(u(t), Y, T_1 \le T(Y) \le T_2\big)$ for every $\tau = 0, 1, 2, \dots, T_2$; if we fix $\tau = T_2$ then we find the probabilities $P_x\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big)$.

Algorithm 4: *Determining the solution of Problem 1 in the case $T_1 \ne T_2$*

Preliminary step (Step 0): Put $P_z\big(u(t), Y, T_1 \le T(Y) \le T_2\big) = 1$ for every position $z \in Y$ and $P_z\big(u(t), Y, T_1 \le T(y) \le T_2\big) = 0$ for the positions $z \in Z_{T_2} \setminus \{(x_f, T_2)\}$.

General step (Step $\tau, \tau \ge 1$): Calculate

$$P_z\big(u(t), Y, T_1 \le T(Y) \le T_2\big) = \sum_{(z,w) \in E^0(z)} P_w\big(u(t), Y, T_1 \le T(Y) \le T_2\big) p_u(z, w)$$

for every $z \in Z_{T_2 - \tau} \setminus Y$ where $E^0(z) = \big\{(z, w) \in E^0 \mid w \in Z_{\tau+1}\big\}$. If $\tau = T$ then go to final step; otherwise go to step $\tau + 1$.

Theorem 4. *Algorithm 4 correctly finds the state probability $P_{x(0)}(u^*(t), x_f, T)$ for an arbitrary starting position $x(0) \in X$ in problem 1 with $T_1 \le T_2$. The running time of the algorithm is $O(|X|^2 T_2)$.*

Proof. In algorithm 4 the value $P_{x_0}\big(u(t), x_f, T_1 \le T(x_f) \le T_2\big)$ is calculated on the basis of formula (11) applying Algorithm 2 for finding $P_{(x_0,0)}\big(u(t), (x_f, T_1 + k), k\big)$ for $k = 0, 1, 2, \dots, T_2 - T_1$. The application of Algorithm 2 on network with respect to each final position is equivalent with the specification of the preliminary step as it is described in Algorithm 4. So, the algorithm correctly finds the probability for the problem 1 with $T_1 \ne T_2$. The general step of the algorithm is made T_2 times. Therefore the running time of the algorithm is $O(|X|^2 T_2)$. □

Now let us show that the network $(G^0, Z^C, Z^N, z_0, Y, c, p_u, T_1, T_2)$ can be modified such that Algorithm 4 becomes Algorithm 2 on an auxiliary stochastic network. We make the following non-essential transformations of the structure of the network. In $G^0 = (Z, E^0)$ we add directed edges

$$\big((x_f, T_1), (x_f, T_1+1)\big), \big((x_f, T_1+1), (x_f, T_1+2)\big), \dots, \big((x_f, T_2-1), (x_f, T_2)\big).$$

To each directed edge $e_i = \big((x_f, T_1+i), (x_f, T_1+i+1)\big), i = 0, 1, 2, \dots, T_2 - T_1 - 1$ we define the values $p(e_i) = 1$ and $c(e_i) = 0$ which express respectively the probabilities and the costs of system's passage from the positions $(x_f, T_1 + i)$ to the position (x_f, T_1+i+1). We denote the network obtained after this construction by $(G^*, Z^C, Z^N, z_0, z_f, c^*, p_u^*, T_1, T_2)$, where $G^* = (Z, E^*)$ is the graph obtained from G^0 by using the construction described above, i.e. $E^* = E \cup \big\{\big((x_f, T_1+i), (x_f, T_1+i+1)\big), \quad i = 0, 1, 2, \dots, T_2 - T_1 - 1\big\}$; the probability and the cost functions $p_u^*, c^*$are obtained from p_u and c, respectively, according to given above additional construction. It is easy to see that if on this network we apply Algorithm 2 considering $T = T_2$ and $(x_f, T) = (x_f, T_2)$ then we find the state probabilities $P_{x_f, T_2-\tau}\big(u(t), (x_f, T_2), T_2\big)$ which coincide with the state probabilities $P_{x_f, T_2-\tau}\big(u(t), Y, T_1 \le T(Y) \le T_2\big)$.

Algorithm 5: *Determining the optimal control for Problem 2 with $T_1 \neq T_1$*

The algorithm consists of the preliminary, general and final steps. The preliminary and general steps find the values $\pi_{(x, T_2-\tau)}\big(Y, T_1 \le T(Y) \le T_2\big)$ which correspond to probabilities $P_{(x, T-\tau)}\big(u^*(t), Y, T_1 \le T(Y)\big) \le T_2$ when the optimal control is taken into account. So, these values represent the probabilities $P_{x(T-\tau)}\big(u^*(t), x_f, T_1 \le T(x_f) \le T_2\big)$ of system transactions from the states $x(T-\tau) \in X$ to the state x_f when the optimal control $u^*(t)$ is applied. At the end of the last iteration of general step the subset $E_C(u^*)$ from E_C is constructed. This subset determines the set of optimal controls for Problem 2. The final step of the algorithm fixes an optimal control $u^*(t)$.

Preliminary step (Step 0): Put $\pi_{(z,T)}(Y, T_1 \le T(Y) \le T_2) = 1$ for every position $z \in Y$ and $\pi_z(Y, T_1 \le T(Y) \le T_2) = 0$ for every positions $z \in Z_{T_2} \setminus \{(x_f, T_2)\}$; in addition put $E_C(u^*) = \emptyset$.

General step (Step $\tau, \tau \ge 1$): For given τ do the following items $a)$ and $b)$:

$a)$ For each position $z \in Z_\tau^N$ calculate

$$\pi_z\big(Y, T_1 \le T(Y) \le T_2\big) = \sum_{(z,w) \in E^+(z)} \pi_w\big(Y, T_1 \le T(Y) \le T_2\big) p(z, w);$$

$b)$ For each position $z \in Z_\tau^C$ calculate

$$\pi_z\big(Y, T_1 \le T(y) \le T_2\big) = \max_{(z,w) \in E^+(z)} \pi_w\big(Y, T_1 \le T(y) \le T_2\big)$$

and include in the set E_C^* each edge $e^* = (z, w)^*$ which satisfy the condition

$$\pi_z\big(Y, T_1 \le T(Y) \le T_2\big) = \pi_w\big(Y, T_1 \le T(Y) \le T_2\big).$$

If $\tau = T$ then go to Final step; otherwise go to step $\tau + 1$.

Final Step: Form the graph $G^* = \big(Z, E_C^* \cup (E \setminus E_C)\big)$ and fix in G^* a map

$$u^* : (x,t) \to (y, t+1) \in X_{G^*}(x,t) \quad \text{for} \quad (x,t) \in Z^C$$

where $X_{G^*} = \{(y,t+1) \in Z | ((x,t),(y,t+1)) \in E_C^*\}$.

Theorem 5. *Algorithm 5 correctly finds the optimal control $u^*(t)$ and the state probability $P_{x(0)}(u^*(t), x_f, T)$ for an arbitrary starting position $x(0) \in X$ in problem 1 with fixed final state $x_f \in X$ and given $T = T_1 = T_2$. The running time of the algorithm is $O(|X|^2 T)$.*

Proof. The proof of this theorem is similar to the prove of Theorem 2. The general step of the algorithm reflects the principle of optimality of dynamic programming for the problem of finding the probabilities $P_{x(T-\tau)}\big(u^*(t), x_f, T_1 \le T(x_f) \le T_2\big)$. These probabilities in stochastic networks correspond to the probabilities $P_{(x,T_2-\tau)}\big(u^*(t), Y, T_1 \le T(Y) \le T_2\big) = \pi_{(x,T_2-\tau)}(Y, T_1 \le T(Y) \le T_2)$. For each controllable position $(x, T-\tau)$ the values $\pi_{(x,T_2-\tau)}\big(Y, T_1 \le T(Y) \le T_2\big)$ are calculated in consideration that for given moment of time $T - \tau$ and given state x the optimal control $u^*(T_2 - \tau) \in U_t(x(T_2 - \tau))$ is applied. Therefore $\pi_{(x,T_2-\tau)}(Y, T_1 \le T(Y) \le T_2) = P_{x(T_2-\tau)}\big(u^*(t), x_f, T_1 \le T(Y) \le T_2\big)$ for every $x \in X$ and $\tau = 0, 1, 2, \ldots, T_2$. Taking into account that at each step the directed edges e^* correspond to the optimal control for the corresponding positions on stochastic network, we obtain at the final step the set of edges E^* which give the optimal control for arbitrary state x and arbitrary moment of time t. In the same way as in previous algorithms we can show that the running time of the algorithm is $O(|X|^2 T_2)$. □

5 Algorithms for Determining the Expectation of Integral-Time Cost in Problems 3–6

In this section we describe algorithms for calculation of the expected integral-time costs of state transactions of dynamical system in problems 3-6.

5.1 Calculation of the Expectation of Integral-Time cost in Problem 3

The expectation of integral-time cost for dynamical system L on stochastic network $(G, Z^C, Z^N, z_o, c, p_u, T)$ in problem 3 is defined in analogues way as in Subsection 3.2 using the following recursive formula:

$$C_z(u(t), T) = \sum_{(z,w) \in E^+(z)} p_u(z,w)\big(c(z,w) + C_w(u(t), T)\big),$$

$$z \in Z_{T-\tau}, \tau = 1, 2, \ldots, T,$$

where $E^+(z) = \{(z,w) \in E | w \in Z_{T-\tau+1}\}$. This formula can be treated in the following way. Assume that we should estimate the expected integral-time cost of system's transactions during τ units of time when the system starts transactions in position $z = (x, T - \tau)$ at the moment of time $T - \tau$. If the system makes a transition from the position $z = (x, T - \tau)$ to the position $w = (y, T - \tau + 1)$ it will spend the amount $c(z, w)$ plus the amount it expects to spend if the system starts the remained $\tau - 1$ transactions in the position $w = (y, T - \tau + 1)$. Therefore if the system L at the moment of time $T - \tau$ is in position $z = (x, T - \tau)$ then the expected integral-cost of system's transitions from z must be weighted by he probabilities of such transactions $p_u(z, w)$ to obtain the total expected integral-time costs.

Algorithm 6: *Determining the expectation of integral-time cost in Problem 3*
Preliminary Step (Step 0): Put $C_z(u(t), T) = 0$ for every $z \in Z_T$.
General Step (Step τ,$\tau \geq 1$): For each $z \in Z_{T-\tau}$ calculate

$$C_z(u(t), T) = \sum_{(z,w)\in E^+(z)} p_u(z, w)\big(c(z, w) + C_w(u(t), T)\big).$$

If $\tau = T$ then stop; otherwise go to step $\tau + 1$.

Algorithm 6 uses the backward dynamic procedure and finds $C_z(u(t), T)$ for every position $z \in Z$. For a fixed position $z = (x, T - \tau) \in Z$ the value $C_z(u(t), T)$ corresponds to the expected integral-time cost $C_{x(T-\tau)}(u(t), T)$ of the system in the next τ transactions when it starts in the state $x = x(T - \tau)$ at the moment of time $T - \tau$, i.e. $C_{(x,T-\tau)}(u(t), T) = C_{x(o)}(u(t), T)$.

Algorithm 7: *Determining the optimal control for problem 4*

The algorithm consists of the preliminary, general and final steps. At the preliminary and general steps the algorithm finds the optimal values of the expectation of integral-time costs $C_z(u(t), T)$ which in algorithm are denoted by $Exp_z(T)$. For a position $z = (x, T - \tau)$ the value $Exp_z(T)$ expresses the expected integral-time cost during τ transactions of the system when it starts transactions in the state $x = x(T - \tau)$ at the moment of time $T - \tau$.This value is calculated in the consideration that the optimal control $u(t)$ is applied. In addition at the general step of the algorithm the possible directed edges $e^* = \big((x, T - \tau), (y, T - \tau + 1)\big)^*$ which correspond to optimal control in the state $x = x(T - \tau)$ at the moment of time $T - \tau$ are cumulated in the set $E_C(u^*)$. The set of optimal controls is determined by $E_C(u^*)$; at the final step an optimal control is fixed.

Preliminary step (Step 0): Put $Exp_z(T) = 0$ for $z \in Z_T$ and $E_C(u) = \emptyset$.

General step (Step τ, $\tau \geq 1$): For given τ do the following items a) and b):

a) For each uncontrollable position $z \in Z^N_{T-\tau}$ calculate

$$Exp_z(T) = \sum_{(z,w)\in E^+(z)} p(z, w)\big(c(z, w) + Exp_w(T)\big);$$

b) For each controllable position $z \in Z^C_{T-\tau}$ calculate

$$Exp_z(T) = \max_{(z,w)\in E^+(z)} \big(c(z, w) + Exp_w(T)\big)$$

and include in the set E_C^* each edge $e^* = (z,w)^*$ which satisfies the condition

$$c((z,w)^*) + Exp_{w^*}(T) = \max_{(z,w)\in E^+(z)} \big(c(z,w) + Exp_w(T)\big).$$

If $\tau = T$ then go to Final step; otherwise go to step $\tau + 1$.

Final Step: Form the graph $G^* = (Z, E_C^* \cup (E \setminus E_C))$ and fix in G^* a map $u^* : (x,t) \to (y,t+1) \in X_{G^*}(x,t) \quad for \quad (x,t) \in Z^C$.

Theorem 6. *Algorithm 7 correctly finds the optimal control $u^*(t)$ and the expected integral-time costs $C_{x(0)}(T)$ of the system's transactions during T units of time from an arbitrary starting position $x = x(0) \in X$ in problem 4. The running time of the algorithm is $O(|X|^2 T)$.*

This theorem can be proved in analogues way as Theorem 2.

5.2 Determining the Expectations of Integral-Time cost in Problems 5 and 6

For problems 5 and 6 we need to precise what is meant by the expectation of integral-time cost for dynamical system when the state x_f is reached at the moment of time $T(x)$ such that $T_1 \le T(x) \le T_2$. At first let us analyze the case $T_1 = T_2 = T$. We consider this problem on stochastic network $(G, Z^C, Z^N, z_0, z_f, c, p_u, T)$. If we assume that the final position $z_f = (x_f, T)$ is reached at the moment of time T then we should consider that the probability of system transaction from an arbitrary starting position $z = (x, 0)$ to the position z_f is equal to 1. This means that the probabilities $p_u(e)$ on edges $e \in E$ should be redefined or transformed in such way that the mentioned above condition on stochastic network holds.We denote these redefined values by $p'_u(e)$ and call them conditional probabilities. It is evident that if the system never can meet a directed edge $e \in E$ during transition from a position $(x, 0)$ to the position z_f then the conditional probability $p'_u(e)$ of this edge is equal to zero. So, the first step we should do in the transformation is to delete all such edges from the graph G. After such transformation we obtain a new graph $G' = (Z, E')$ in which for some positions $z \in Z$ the condition $\sum\limits_{(z,w)\in E'(z)} p(z,w) = 1$ is not satisfied (here $E'(z)$ represents the subset of edges from E' which in vertex z, i.e. $E'(z) = \{(z,w) | (z,w) \in E'\}$). Then we find for each position $z \in Z$ the value $\pi(z) = \sum\limits_{(z,w)\in E'(z)} p_u(z,w)$ and after that for an arbitrary position $z \in Z$ with $\pi(z) \neq 0$ we make the transformation

$$p'_u(z,w) = \frac{1}{\pi(z)} p_u(z,w)$$

for every $(z,w) \in E'(z)$. After these transformations we can apply Algorithm 6 on stochastic network $(G, Z^C, Z^N, z_0, c, p'_u, T)$ with conditional probabilities $p'_u(e)$ of edges $e \in E$; here $p'_u(e) = 0$ for $e \in E \setminus E'$. If for this network we find $C_{z_0}(T)$

then fix $C_{z_0}(T) = C_{x_0}(u(t), x_f, T) = C_{z_0}(T)$, i. e. this value represents the expected integral-time cost of dynamical system L in problem 5. In the case $T_1 \neq T_2$ the expected integral-time cost $C_{x_o}(u(t), x_f, T_1 \leq T(x) \leq T_2)$ can be found in analogues way if we consider problem 5 on stochastic network $(G^*, Z^C, Z^N, z_0, z_f, c^*, p^*, T_1, T_2)$ and will make a similar transformation. It is evident that the control problem 6 can be reduced to control problem 4 using the approach described above. This allows us to find the optimal control $u^*(t)$ which provides a maximal expected integral-time cost $C_{x_0}(u^*(t), x_f, T_1 \leq T(x_f) \leq T_2)$ of system transactions from starting state x_0 to final state x_f such that $T_1 \leq T(x_f) \leq T_2$.

References

[1] Bellman R., Kalaba R. *Dynamic programming and modern control theory.* Academic Press, New York and London (1965).

[2] Boltjanski W.G. *Optimale Steuerung diskreter Systeme.* Leipzig Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1976.

[3] Feller W. *An Introduction to Probability Theory and its Applications*, vol. 1, New York, John Wiley, 2nd edition, 1957.

[4] Howard R.A. *Dynamic Programming and Markov Processes.* Wiley, 1960.

[5] Lozovanu D., Pickl S. *Optimization and Multiobjective Control of Time-Discrete Systems.* Springer, 2008.

[6] Lozovanu D., Pickl S. *Allgorithms and the calculation of Nash equilibria for multi-objective control of time-discrete systems and polynomial-time algorithms for dynamic c-games on networks.* European Journal of Operational Research, 2007, **181**, 1214–1232.

Dmitrii Lozovanu
Institute of Mathematics and Computer Science
Academy of Sciences of Moldova
5 Academiei str. Chişinău, MD−2028
Moldova
E-mail: *lozovanu@math.md*

Received January 25, 2009

Stefan Pickl
Institut fur Angewandte Systemwissenschaften und
Wirtschaftsinformatik
Fakultät fur Informatik, Universität der Bundeswehr,
München
E-mail: *stefan.pickl@unibw.de*

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 91–110
ISSN 1024–7696

On a Generalization of Hardy-Hilbert's Integral Inequality

Namita Das, Srinibas Sahoo

Abstract. A generalization of Hardy-Hilbert's integral inequality was given by B.Yang in [18]. The main purpose of the present article is to generalize the inequality. As applications, the reverse, the equivalent form of the inequality, some particular results and the generalization of Hardy-Littlewood inequalities are derived.

Mathematics subject classification: 26D15.
Keywords and phrases: Hardy-Hilbert's integral inequality; β-function; Hölder's inequality; Hardy-Littlewood inequality.

1 Introduction

Let $\frac{1}{p}+\frac{1}{q}=1$ $(p>1)$, $f,g\geq 0$. Suppose $0<\int_0^\infty f^p(x)dx<\infty$ and $0<\int_0^\infty g^q(x)dx<\infty$. The well known Hardy-Hilbert's integral inequality (see [1]) is given by

$$\int_0^\infty\int_0^\infty \frac{f(x)g(y)}{x+y}dxdy<\frac{\pi}{sin(\pi/p)}\left(\int_0^\infty f^p(x)dx\right)^{\frac{1}{p}}\left(\int_0^\infty g^q(x)dx\right)^{\frac{1}{q}}, \quad (1)$$

and an equivalent form is given by

$$\int_0^\infty\left(\int_0^\infty \frac{f(x)}{x+y}dx\right)^p dy<\left[\frac{\pi}{sin(\pi/p)}\right]^p\int_0^\infty f^p(x)dx, \quad (2)$$

where the constant factor $\pi/sin(\pi/p)$ and $[\pi/sin(\pi/p)]^p$ are the best possible. Recently many generalizations and refinements of these inequalities were also obtained. Some of them are given in [4]–[27]. One of the generalizations given by Yang [18] is the following:

Theorem 1. *If $p>1$, $\frac{1}{p}+\frac{1}{q}=1$, $\phi_r>0$ $(r=p,q)$, $\phi_p+\phi_q=\lambda$, $u(x)$ is a differentiable strictly increasing function in (a,b) $(-\infty\leq a<b\leq\infty)$ such that $u(a+)=0$ and $u(b-)=\infty, f,g\geq 0$ satisfy $0<\int_a^b\frac{(u(x))^{p(1-\phi_q)-1}}{(u'(x))^{p-1}}f^p(x)dx<\infty$ and $0<\int_a^b\frac{(u(x))^{q(1-\phi_p)-1}}{(u'(x))^{q-1}}g^q(x)dx<\infty$ then*

$$\begin{aligned}\int_a^b\int_a^b\frac{f(x)g(y)}{(u(x)+u(y))^\lambda}dxdy<B(\phi_p,\phi_q)\left(\int_a^b\frac{(u(x))^{p(1-\phi_q)-1}}{(u'(x))^{p-1}}f^p(x)dx\right)^{\frac{1}{p}}\times\\ \times\left(\int_a^b\frac{(u(x))^{q(1-\phi_p)-1}}{(u'(x))^{q-1}}g^q(x)dx\right)^{\frac{1}{q}};\end{aligned} \quad (3)$$

where the constant factor $B(\phi_p, \phi_q)$ is the best possible. If $p < 1 (p \neq 0), \{\lambda : \phi_r > 0, (r = p, q), \phi_p + \phi_q = \lambda\} \neq \Phi$, with the above assumption, the reverse of (3) *holds and the constant factor is still the best possible.*

In this paper, we have generalized the inequality (3), where we have weakened the normalized condition $\phi_p + \phi_q = \lambda$ and considered two different functions $u(x)$ and $v(x)$, which is more generalized inequality and from which most of the recent results are obtained by specialising the parameters and the functions $u(x)$ and $v(x)$. We have also given the generalization of Hardy–Littlewood inequality.

2 Some Lemmas

We first set the following notations: Suppose $p \notin \{0, 1\}$, $\frac{1}{p} + \frac{1}{q} = 1$, $0 < \phi_r < \lambda$ $(r = p, q)$, $u(x)$ and $v(x)$ are differentiable strictly increasing function in (a, b) $(-\infty \leq a < b \leq \infty)$ and (c, d) $(-\infty \leq c < d \leq \infty)$ respectively such that $u(a+) = v(c+) = 0$ and $u(b-) = v(d-) = \infty$.

We need the formula of the $\beta-$function as (cf. Wang et al. [3]):

$$B(p, q) = \int_0^\infty \frac{1}{(1+t)^{p+q}} t^{p-1} dt = B(q, p) \tag{4}$$

Lemma 1. (cf. Kuang [2]). *If $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, $\omega(t) > 0$, $f, g \geq 0, f \in L^p_\omega(E)$ and $g \in L^q_\omega(E)$, then one has the Hölder's inequality with weight as:*

$$\int_E \omega(t) f(t) g(t) dt \leq \left\{ \int_E \omega(t) f^p(t) dt \right\}^{\frac{1}{p}} \left\{ \int_E \omega(t) g^q(t) dt \right\}^{\frac{1}{q}}. \tag{5}$$

If $p < 1 (p \neq 0)$, with the above assumption, the reverse of (5) *holds, where the equality in the above two cases holds if and only if there exists non-negative real numbers c_1 and c_2 such that they are not all zero and*

$$c_1 f^p(t) = c_2 g^q(t), \quad a.e. \ in \ \ E.$$

Lemma 2. *Define $\omega_\lambda(u, v, p, x)$ and $\omega_\lambda(v, u, q, y)$ as*

$$\omega_\lambda(u, v, p, x) = \int_c^d \frac{(v(y))^{\phi_p - 1} v'(y)}{(u(x) + v(y))^\lambda} dy, \ x \in (a, b), \tag{6}$$

$$\omega_\lambda(v, u, q, y) = \int_a^b \frac{(u(x))^{\phi_q - 1} u'(x)}{(u(x) + v(y))^\lambda} dx, \ y \in (c, d). \tag{7}$$

Then

$$\omega_\lambda(u, v, p, x) = B(\phi_p, \lambda - \phi_p)(u(x))^{\phi_p - \lambda}, \ x \in (a, b), \tag{8}$$

$$\omega_\lambda(v, u, q, y) = B(\phi_q, \lambda - \phi_q)(v(y))^{\phi_q - \lambda}, \ y \in (c, d). \tag{9}$$

Proof. Setting $t = \frac{v(y)}{u(x)}$ in (6), we have

$$\omega_\lambda(u, v, p, x) = \int_0^\infty \frac{(tu(x))^{\phi_p - 1} u(x)}{(u(x) + tu(x))^\lambda} dt = \\ = (u(x))^{\phi_p - \lambda} \int_0^\infty \frac{1}{(1+t)^\lambda} t^{\phi_p - 1} dt.$$

By (4), we get (8). Similarly, (9) can be proved. The lemma is proved. □

Lemma 3. *Suppose* $\phi_p + \phi_q = \lambda$. *Take* $a_1 = u^{-1}(1),\ c_1 = v^{-1}(1)$.

(i) *If* $p > 1,\ \frac{1}{p} + \frac{1}{q} = 1$ *and* $0 < \varepsilon < q\phi_p$, *then*

$$I := \int_{a_1}^b \int_{c_1}^d \frac{(u(x))^{\phi_q - \frac{\varepsilon}{p} - 1} u'(x) (v(y))^{\phi_p - \frac{\varepsilon}{q} - 1} v'(y)}{(u(x) + v(y))^\lambda} dxdy > \\ > \frac{1}{\varepsilon} B\left(\phi_p - \frac{\varepsilon}{q}, \phi_q + \frac{\varepsilon}{q}\right) - \bigcirc(1). \tag{10}$$

(ii) *If* $0 < p < 1$ *(or* $p < 0$*) and* $0 < \varepsilon < -q\phi_q$ *(or* $0 < \varepsilon < q\phi_p$*), then*

$$I < \frac{1}{\varepsilon} B\left(\phi_p - \frac{\varepsilon}{q}, \phi_q + \frac{\varepsilon}{q}\right). \tag{11}$$

Proof. For fixed $x \in (a_1, b)$, setting $t = \frac{v(y)}{u(x)}$ in (10), we have

$$I := \int_{a_1}^b (u(x))^{-1-\varepsilon} u'(x) \left(\int_{\frac{1}{u(x)}}^\infty \frac{1}{(1+t)^\lambda} t^{\phi_p - \frac{\varepsilon}{q} - 1} dt\right) dx = \\ = \int_{a_1}^b \frac{u'(x)}{(u(x))^{1+\varepsilon}} dx \int_0^\infty \frac{1}{(1+t)^\lambda} t^{\phi_p - \frac{\varepsilon}{q} - 1} dt - \\ - \int_{a_1}^b \frac{u'(x)}{(u(x))^{1+\varepsilon}} \left(\int_0^{\frac{1}{u(x)}} \frac{1}{(1+t)^\lambda} t^{\phi_p - \frac{\varepsilon}{q} - 1} dt\right) dx > \\ > \frac{1}{\varepsilon} \int_0^\infty \frac{1}{(1+t)^\lambda} t^{\phi_p - \frac{\varepsilon}{q} - 1} dt - \int_{a_1}^b \frac{u'(x)}{u(x)} \left(\int_0^{\frac{1}{u(x)}} t^{\phi_p - \frac{\varepsilon}{q} - 1} dt\right) dx = \\ = \frac{1}{\varepsilon} \int_0^\infty \frac{1}{(1+t)^\lambda} t^{\phi_p - \frac{\varepsilon}{q} - 1} dt - \left(\phi_p - \frac{\varepsilon}{q}\right)^{-2}. \tag{12}$$

By (4), inequality (10) is valid. If $0 < p < 1$ (or $p < 0$), by (12) we get

$$I < \int_{a_1}^b \frac{u'(x)}{(u(x))^{1+\varepsilon}} dx \int_0^\infty \frac{1}{(1+t)^\lambda} t^{\phi_p - \frac{\varepsilon}{q} - 1} dt$$

and then by (4), inequality (11) is valid. The lemma is proved. □

3 Main Results

Theorem 2. *If* $p>1,\ \frac{1}{p}+\frac{1}{q}=1,\ 0<\phi_r<\lambda\ (r=p,q)$ *and* $f,g\geq 0$ *satisfy* $0<\int_a^b \frac{(u(x))^{\phi_p-\lambda+(p-1)(1-\phi_q)}}{(u'(x))^{p-1}}f^p(x)dx<\infty$ *and* $0<\int_c^d \frac{(v(x))^{\phi_q-\lambda+(q-1)(1-\phi_p)}}{(v'(x))^{q-1}}g^q(x)dx<\infty$ *then*

$$\int_a^b\int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^\lambda}dxdy < H_\lambda(\phi_p,\phi_q)\left(\int_a^b \frac{(u(x))^{\phi_p-\lambda+(p-1)(1-\phi_q)}}{(u'(x))^{p-1}}f^p(x)dx\right)^{\frac{1}{p}}\times$$
$$\times\left(\int_c^d \frac{(v(x))^{\phi_q-\lambda+(q-1)(1-\phi_p)}}{(v'(x))^{q-1}}g^q(x)dx\right)^{\frac{1}{q}} \tag{13}$$

where $H_\lambda(\phi_p,\phi_q)=B^{\frac{1}{p}}(\phi_p,\lambda-\phi_p)B^{\frac{1}{q}}(\phi_q,\lambda-\phi_q)$.

If $p<1(p\neq 0),\{\lambda:0<\phi_r<\lambda,r=p,q\}\neq\Phi$, *with the above assumption, the reverse of* (13) *holds.*

Proof. By (5), we have

$$\begin{aligned}\int_a^b\int_c^d &\frac{f(x)g(y)}{(u(x)+v(y))^\lambda}dxdy=\\ &=\int_a^b\int_c^d\frac{1}{(u(x)+v(y))^\lambda}\left[\frac{(v(y))^{(\phi_p-1)/p}(v'(y))^{1/p}}{(u(x))^{(\phi_q-1)/q}(u'(x))^{1/q}}f(x)\right]\times\\ &\times\left[\frac{(u(x))^{(\phi_q-1)/q}(u'(x))^{1/q}}{(v(y))^{(\phi_p-1)/p}(v'(y))^{1/p}}g(y)\right]dxdy\leq \\ &\leq\left\{\int_a^b\left[\int_c^d\frac{(v(y))^{\phi_p-1}v'(y)}{(u(x)+v(y))^\lambda}dy\right]\frac{(u(x))^{(p-1)(1-\phi_q)}}{(u'(x))^{p-1}}f^p(x)dx\right\}^{\frac{1}{p}}\times\\ &\times\left\{\int_c^d\left[\int_a^b\frac{(u(x))^{\phi_q-1}u'(x)}{(u(x)+v(y))^\lambda}dx\right]\frac{(v(y))^{(q-1)(1-\phi_p)}}{(v'(y))^{q-1}}g^q(y)dy\right\}^{\frac{1}{q}}.\end{aligned} \tag{14}$$

If (14) takes the form of equality, then by (5) there exist non negative numbers c_1 and c_2 such that they are not all zero and

$$\begin{aligned}c_1&\frac{(v(y))^{\phi_p-1}v'(y)(u(x))^{(p-1)(1-\phi_q)}}{(u'(x))^{p-1}}f^p(x)=\\ &=c_2\frac{(u(x))^{\phi_q-1}u'(x)(v(y))^{(q-1)(1-\phi_p)}}{(v'(y))^{q-1}}g^q(y),\quad \text{a. e. in } (a,b)\times(c,d).\end{aligned}$$

It follows that

$$c_1\frac{(u(x))^{p(1-\phi_q)}}{(u'(x))^{p-1}}f^p(x)=c_2\frac{(v(y))^{q(1-\phi_p)}}{(v'(y))^{q-1}}g^q(y)=c_3,\quad \text{a. e. in } (a,b)\times(c,d)$$

where c_3 is a constant. Without loss of generality, suppose that $c_1 \neq 0$. Then we have

$$\int_a^b \frac{(u(x))^{\phi_p-\lambda+(p-1)(1-\phi_q)}}{(u'(x))^{p-1}} f^p(x)dx = \frac{c_3}{c_1}\int_a^b (u(x))^{\phi_p+\phi_q-\lambda-1}u'(x)dx =$$
$$= \frac{c_3}{c_1}\left\{\int_0^1 t^{\phi_p+\phi_q-\lambda-1}dt + \int_1^\infty t^{\phi_p+\phi_q-\lambda-1}dt\right\} = \infty$$

which contradicts to

$$0 < \int_a^b \frac{(u(x))^{\phi_p-\lambda+(p-1)(1-\phi_q)}}{(u'(x))^{p-1}} f^p(x)dx < \infty.$$

Then by (6) and (7), we have

$$\int_a^b\int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^\lambda}dxdy < \left\{\int_a^b \omega_\lambda(u,v,p,x)\frac{(u(x))^{(p-1)(1-\phi_q)}}{(u'(x))^{p-1}} f^p(x)dx\right\}^{\frac{1}{p}} \times$$
$$\times \left\{\int_c^d \omega_\lambda(v,u,q,y)\frac{(v(y))^{(q-1)(1-\phi_p)}}{(v'(y))^{q-1}} g^q(y)dy\right\}^{\frac{1}{q}} \tag{15}$$

and in view of (8) and (9), it follows that (13) is valid.

For $0 < p < 1$ (or $p < 0$), by the reverse of (5) and using the same procedure, we can obtain the reverse of (13). The theorem is proved. $\square$

Theorem 3. *Let the assumptions of Theorem 2 hold.*

(i) If $p > 1, 1/p + 1/q = 1$, we obtain the equivalent inequality of (13) as follows:

$$\begin{aligned}\int_c^d \frac{v'(y)}{(v(y))^{1-\phi_p+(p-1)(\phi_q-\lambda)}}\left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right]^p dy <\\ < [H_\lambda(\phi_p,\phi_q)]^p \int_a^b \frac{(u(x))^{\phi_p-\lambda+(p-1)(1-\phi_q)}}{(u'(x))^{p-1}} f^p(x)dx;\end{aligned} \tag{16}$$

(ii) If $0 < p < 1$, we obtain the reverse of (16) equivalent to the reverse of (13);

(iii) If $p < 0$, we obtain inequality (16) equivalent to the reverse of (13).

Proof. Set $g(y) = \frac{v'(y)}{(v(y))^{1-\phi_p+(p-1)(\phi_q-\lambda)}}\left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right]^{p-1}$. By (13), we have

$$0 < \int_c^d \frac{(v(y))^{\phi_q-\lambda+(q-1)(1-\phi_p)}}{(v'(y))^{q-1}} g^q(y)dy =$$
$$= \int_c^d \frac{v'(y)}{(v(y))^{1-\phi_p+(p-1)(\phi_q-\lambda)}}\left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right]^p dy =$$
$$= \int_a^b\int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^\lambda}dxdy \leq$$

$$\leq H_\lambda(\phi_p,\phi_q)\left(\int_a^b \frac{(u(x))^{\phi_p-\lambda+(p-1)(1-\phi_q)}}{(u'(x))^{p-1}} f^p(x)dx\right)^{\frac{1}{p}} \times$$
$$\times \left(\int_c^d \frac{(v(x))^{\phi_q-\lambda+(q-1)(1-\phi_p)}}{(v'(x))^{q-1}} g^q(x)dx\right)^{\frac{1}{q}}, \tag{17}$$

then

$$0 < \left\{\int_c^d \frac{(v(y))^{\phi_q-\lambda+(q-1)(1-\phi_p)}}{(v'(y))^{q-1}} g^q(y)dy\right\}^{\frac{1}{p}} =$$
$$= \left\{\int_c^d \frac{v'(y)}{(v(y))^{1-\phi_p+(p-1)(\phi_q-\lambda)}}\left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right]^p dy\right\}^{\frac{1}{p}} \leq \tag{18}$$
$$\leq H_\lambda(\phi_p,\phi_q)\left\{\int_a^b \frac{(u(x))^{\phi_p-\lambda+(p-1)(1-\phi_q)}}{(u'(x))^{p-1}} f^p(x)dx\right\}^{\frac{1}{p}} < \infty.$$

It follows that (17) takes the form of strict inequality by using (13); so, does (18). Hence we can get (16).

On the other hand, if (16) holds, then by (5), we have

$$\int_a^b\int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^\lambda}dxdy =$$
$$= \int_c^d \left[\frac{(v'(y))^{1/p}}{(v(y))^{(1-\phi_p+(p-1)(\phi_q-\lambda))/p}}\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right] \times$$
$$\times \left[\frac{(v(y))^{(1-\phi_p+(p-1)(\phi_q-\lambda))/p}}{(v'(y))^{1/p}} g(y)\right] dy \leq$$
$$\leq \left\{\int_c^d \frac{v'(y}{(v(y))^{1-\phi_p+(p-1)(\phi_q-\lambda)}}\left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right]^p dy\right\}^{\frac{1}{p}} \times$$
$$\times \left\{\int_c^d \frac{(v(y))^{\phi_q-\lambda+(q-1)(1-\phi_p)}}{(v'(y))^{q-1}} g^q(y)dy\right\}^{\frac{1}{q}}.$$

Hence by (16), (13) yields. Thus it follows that (13) and (16) are equivalent. The theorem is proved. □

Theorem 4. *If* $p>1,\ 1/p+1/q=1,\ \phi_r>0(r=p,q),\ \phi_p+\phi_q=\lambda$ *and* $f,g\geq 0$ *satisfy* $0<\int_a^b \frac{(u(x))^{p(1-\phi_q)-1}}{(u'(x))^{p-1}}f^p(x)dx<\infty$ *and* $0<\int_c^d \frac{(v(x))^{q(1-\phi_p)-1}}{(v'(x))^{q-1}}g^q(x)dx<\infty$ *then*

$$\begin{aligned}\int_a^b\int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^\lambda}dxdy &< B(\phi_p,\phi_q)\left(\int_a^b \frac{(u(x))^{p(1-\phi_q)-1}}{(u'(x))^{p-1}}f^p(x)dx\right)^{\frac{1}{p}}\times\\ &\times\left(\int_c^d \frac{(v(x))^{q(1-\phi_p)-1}}{(v'(x))^{q-1}}g^q(x)dx\right)^{\frac{1}{q}}\end{aligned}\tag{19}$$

where the constant factor $B(\phi_p,\phi_q)$ *is the best possible.*

If $p<1(p\neq 0),\{\lambda:\phi_r>0,(r=p,q),\phi_p+\phi_q=\lambda\}\neq\Phi$, *with the above assumption, the reverse of* (19) *holds and the constant is still the best possible.*

Proof. Since $\phi_p+\phi_q=\lambda$, then by Theorem 2, (19) and its inverse are valid. For $0<\varepsilon<q\phi_p$, setting

$$f_\varepsilon(x)=\begin{cases}0 & \text{if } x\in(a,a_1)\ (a_1=u^{-1}(1)),\\ (u(x))^{\phi_q-\frac{\varepsilon}{p}-1}u'(x) & \text{if } x\in[a_1,b),\end{cases}$$

$$g_\varepsilon(x)=\begin{cases}0 & \text{if } x\in(c,c_1)\ (c_1=v^{-1}(1)),\\ (v(x))^{\phi_p-\frac{\varepsilon}{q}-1}v'(x) & \text{if } x\in[c_1,d),\end{cases}$$

we have

$$\left(\int_a^b \frac{(u(x))^{p(1-\phi_q)-1}}{(u'(x))^{p-1}}f_\varepsilon^p(x)dx\right)^{\frac{1}{p}}\left(\int_c^d \frac{(v(x))^{q(1-\phi_p)-1}}{(v'(x))^{q-1}}g_\varepsilon^q(x)dx\right)^{\frac{1}{q}}=\frac{1}{\varepsilon}\tag{20}$$

If the constant factor $B(\phi_p,\phi_q)$ in (19) is not the best possible, then there exists a positive constant $K<B(\phi_p,\phi_q)$ such that (19) is still valid if we replace $B(\phi_p,\phi_q)$ by K. In particular, by (10) and (20), we have

$$\begin{aligned}&B\left(\phi_p-\frac{\varepsilon}{q},\phi_q+\frac{\varepsilon}{q}\right)-\varepsilon\bigcirc(1)<\\ &<\varepsilon\int_a^b\int_c^d\frac{f_\varepsilon(x)g_\varepsilon(y)}{(u(x)+v(y))^\lambda}dxdy<\\ &<\varepsilon K\left(\int_a^b \frac{(u(x))^{p(1-\phi_q)-1}}{(u'(x))^{p-1}}f_\varepsilon^p(x)dx\right)^{\frac{1}{p}}\left(\int_c^d \frac{(v(x))^{q(1-\phi_p)-1}}{(v'(x))^{q-1}}g_\varepsilon^q(x)dx\right)^{\frac{1}{q}}=K,\end{aligned}$$

and then $B(\phi_p,\phi_q)\leq K\ (\varepsilon\to 0^+)$. This contradiction leads to the conclusion that the constant factor in (19) is the best possible.

For the best constant factor in the reverse of (19), for $0<p<1$ ($p<0$), we set $f_\varepsilon(x)$ and $g_\varepsilon(x)$, for $0<\varepsilon<-q\phi_q$ (or $0<\varepsilon<q\phi_p$), as the above; we still have (20).

If the constant factor $B(\phi_p,\phi_q)$ in the reverse of (19) is not the best possible, then there exists a positive constant $K > B(\phi_p,\phi_q)$ such that the reverse of (19) is still valid if we replace $B(\phi_p,\phi_q)$ by K. In particular, by (11) and (20), we have

$$\begin{aligned}
&B\left(\phi_p-\frac{\varepsilon}{q},\phi_q+\frac{\varepsilon}{q}\right) > \\
&> \varepsilon\int_a^b\int_c^d \frac{f_\varepsilon(x)g_\varepsilon(y)}{(u(x)+v(y))^\lambda}dxdy > \\
&> \varepsilon K\left(\int_a^b \frac{(u(x))^{p(1-\phi_q)-1}}{(u'(x))^{p-1}}f_\varepsilon^p(x)dx\right)^{\frac{1}{p}}\left(\int_c^d \frac{(v(x))^{q(1-\phi_p)-1}}{(v'(x))^{q-1}}g_\varepsilon^q(x)dx\right)^{\frac{1}{q}} = K,
\end{aligned}$$

and then $B(\phi_p,\phi_q) \geq K$ $(\varepsilon \to 0^+)$. This contradiction leads to the conclusion that the constant factor in the reverse of (19) is the best possible. The theorem is proved. $\square$

Corollary 1. *For* $f=g$, $u=v$, $\lambda=1$, $\phi_r=\frac{1}{r}$ $(r=p,q)$, *if* $0<\int_a^b(u'(x))^{1-r}f^r(x)dx<\infty$ $(r=p,q)$ *then*

$$\begin{aligned}
&\int_a^b\int_a^b \frac{f(x)f(y)}{u(x)+u(y)}dxdy < \\
&< \frac{\pi}{\sin\frac{\pi}{p}}\left(\int_a^b (u'(x))^{1-p}f^p(x)dx\right)^{\frac{1}{p}}\left(\int_a^b (u'(x))^{1-q}f^q(x)dx\right)^{\frac{1}{q}},
\end{aligned} \tag{21}$$

where the constant factor $\frac{\pi}{\sin(\pi/p)}$ *is the best possible.*

Corollary 2. *For* $f=g$, $u=v$, $\lambda=1$, $\phi_r=\frac{1}{2}$ $(r=p,q)$, *if* $0<\int_a^b\frac{(u(x))^{\frac{r}{2}-1}}{(u'(x))^{r-1}}f^r(x)dx<\infty$ $(r=p,q)$ *then*

$$\begin{aligned}
&\int_a^b\int_a^b \frac{f(x)f(y)}{u(x)+u(y)}dxdy < \\
&< \pi\left(\int_a^b \frac{(u(x))^{\frac{p}{2}-1}}{(u'(x))^{p-1}}f^p(x)dx\right)^{\frac{1}{p}}\left(\int_a^b \frac{(u(x))^{\frac{q}{2}-1}}{(u'(x))^{q-1}}f^q(x)dx\right)^{\frac{1}{q}},
\end{aligned} \tag{22}$$

where the constant factor π *is the best possible.*

Theorem 5. *Let the assumptions of Theorem 4 hold.*

(i) If $p>1, 1/p+1/q=1$, *we obtain the equivalent inequality of* (19) *as follows:*

$$\begin{aligned}
&\int_c^d \frac{v'(y)}{(v(y))^{1-p\phi_p}}\left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right]^p dy < \\
&< [B(\phi_p,\phi_q)]^p\int_a^b \frac{(u(x))^{p(1-\phi_q)-1}}{(u'(x))^{p-1}}f^p(x)dx;
\end{aligned} \tag{23}$$

(ii) If $0 < p < 1$, we obtain the reverse of (23) equivalent to the reverse of (19);

(iii) If $p < 0$, we obtain inequality (23) equivalent to the reverse of (19), where the constants in the above inequalities are all the best possible.

Proof. Since $\phi_p + \phi_q = \lambda$, then by Theorem 3, we get inequality (23) and its inverse which are equivalent to (19) and its inverse accordingly. By Theorem-4, the constants in (19) and its inverse are best possible, hence the constants in (23) and it's inverse are best possible. The theorem is proved. $\square$

4 Some Particular Inequalities

Theorem 6. *If $p > 1, \frac{1}{p} + \frac{1}{q} = 1, \lambda > \max\{\frac{1}{p}, \frac{1}{q}\}, 0 < \int_a^b \frac{(u(x))^{1-\lambda}}{(u'(x))^{p-1}} f^p(x)dx < \infty$ and $0 < \int_c^d \frac{(v(x))^{1-\lambda}}{(v'(x))^{q-1}} g^q(x)dx < \infty$, then we have the following two equivalent inequalities:*

$$\begin{aligned} &\int_a^b \int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^\lambda} dxdy < \\ &< \tilde{H}_\lambda \left(\frac{1}{p}, \frac{1}{q}\right) \left(\int_a^b \frac{(u(x))^{1-\lambda}}{(u'(x))^{p-1}} f^p(x)dx\right)^{\frac{1}{p}} \left(\int_c^d \frac{(v(x))^{1-\lambda}}{(v'(x))^{q-1}} g^q(x)dx\right)^{\frac{1}{q}}, \end{aligned} \tag{24}$$

$$\begin{aligned} &\int_c^d \frac{v'(y)}{(v(y))^{(p-1)(1-\lambda)}} \left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda} dx\right]^p dy < \\ &< \left[\tilde{H}_\lambda \left(\frac{1}{p}, \frac{1}{q}\right)\right]^p \int_a^b \frac{(u(x))^{1-\lambda}}{(u'(x))^{p-1}} f^p(x)dx, \end{aligned} \tag{25}$$

where $\tilde{H}_\lambda \left(\frac{1}{p}, \frac{1}{q}\right) = B^{\frac{1}{p}} \left(\frac{1}{p}, \lambda - \frac{1}{p}\right) B^{\frac{1}{q}} \left(\frac{1}{q}, \lambda - \frac{1}{q}\right)$.

Proof. Setting $\phi_r = \frac{1}{r}$ $(r = p, q)$, in Theorem 2 and Theorem 3, we get the inequalities (24) and (25) respectively. $\square$

We discuss a number of special cases of inequality (24). Similar examples apply also to inequality (25).

Example 1. Set $u(x) = Ax + C$ $(A > 0)$, $x \in (-C/A, \infty)$ and $v(x) = Bx + C$ $(B > 0)$, $x \in (-C/B, \infty)$ in Theorem 6. Then (24) becomes

$$\begin{aligned} &\int_{-\frac{C}{A}}^\infty \int_{-\frac{C}{B}}^\infty \frac{f(x)g(y)}{(Ax+By+2C)^\lambda} dxdy < \\ &< \frac{1}{A^{1/q}B^{1/p}} \tilde{H}_\lambda \left(\frac{1}{p}, \frac{1}{q}\right) \left(\int_{-\frac{C}{A}}^\infty (Ax+C)^{1-\lambda} f^p(x)dx\right)^{\frac{1}{p}} \times \\ &\times \left(\int_{-\frac{C}{B}}^\infty (Bx+C)^{1-\lambda} g^q(x)dx\right)^{\frac{1}{q}}. \end{aligned} \tag{26}$$

For $A = B = 1, C = -\alpha$, we recover the result of Yang [7].

Example 2. Set $u(x) = x^{\alpha}$ $(\alpha > 0)$, $x \in (0,\infty)$ and $v(x) = x^{\beta}$ $(\beta > 0)$, $x \in (0,\infty)$ in Theorem 6. Then (24) becomes

$$\begin{aligned}\int_0^{\infty}\int_0^{\infty} \frac{f(x)g(y)}{(x^{\alpha}+y^{\beta})^{\lambda}} dxdy &< \\ &< \frac{1}{\alpha^{\frac{1}{q}}\beta^{\frac{1}{p}}}\tilde{H}_{\lambda}\left(\frac{1}{p},\frac{1}{q}\right)\left(\int_0^{\infty} x^{\alpha(2-\lambda-p)+p-1} f^p(x)dx\right)^{\frac{1}{p}} \times \\ &\times \left(\int_0^{\infty} x^{\beta(2-\lambda-q)+q-1} g^q(x)dx\right)^{\frac{1}{q}}.\end{aligned} \tag{27}$$

Taking $\lambda = 1$ and $\alpha = \beta$, we get the result of Yang [12].

Example 3. Set $u(x) = v(x) = lnx$, $x \in (1,\infty)$ in Theorem 6. Then (24) becomes

$$\begin{aligned}\int_1^{\infty}\int_1^{\infty} \frac{f(x)g(y)}{(\ln x+\ln y)^{\lambda}} dxdy < \tilde{H}_{\lambda}\left(\frac{1}{p},\frac{1}{q}\right)\left(\int_1^{\infty}(\ln x)^{1-\lambda}x^{p-1}f^p(x)dx\right)^{\frac{1}{p}} \times \\ \times\left(\int_1^{\infty}(\ln x)^{1-\lambda}x^{q-1}g^q(x)dx\right)^{\frac{1}{q}}.\end{aligned} \tag{28}$$

Theorem 7. *Suppose $f,\ g \geq 0$ satisfy $0 < \int_a^b \frac{(u(x))^{1-\lambda}}{(u'(x))^{p-1}} f^p(x)dx < \infty$ and $0 < \int_c^d \frac{(v(x))^{1-\lambda}}{(v'(x))^{q-1}} g^q(x)dx < \infty$.*

(i) *If $p > 1, \frac{1}{p}+\frac{1}{q} = 1, \lambda > 2-\min\{p,q\}$, then we have the following two equivalent inequalities:*

$$\begin{aligned}\int_a^b\int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^{\lambda}} dxdy &< \\ &< k_{\lambda}(p)\left(\int_a^b \frac{(u(x))^{1-\lambda}}{(u'(x))^{p-1}} f^p(x)dx\right)^{\frac{1}{p}}\left(\int_c^d \frac{(v(x))^{1-\lambda}}{(v'(x))^{q-1}} g^q(x)dx\right)^{\frac{1}{q}},\end{aligned} \tag{29}$$

$$\begin{aligned}\int_c^d \frac{v'(y)}{(v(y))^{(p-1)(1-\lambda)}}\left[\int_a^b \frac{f(x)}{(u(x)+v(y))^{\lambda}}dx\right]^p dy &< \\ &< [k_{\lambda}(p)]^p \int_a^b \frac{(u(x))^{1-\lambda}}{(u'(x))^{p-1}} f^p(x)dx\ ,\end{aligned} \tag{30}$$

where $k_{\lambda}(p) = B\left(\frac{p+\lambda-2}{p},\frac{q+\lambda-2}{q}\right)$.

(ii) *If $0 < p < 1$ and $2-p < \lambda < 2-q$, we have two equivalent reverses of* (29) *and* (30).

(iii) *If $p < 0$ and $2-q < \lambda < 2-p$, we have reverse of* (29) *and the inequality* (30)*, which are equivalent; where the constants in the above inequalities are all the best possible.*

Proof. Setting $\phi_r = 1 + (1 - \frac{1}{r})(\lambda - 2)$ $(r = p, q)$, in Theorem 4 and Theorem 5, we get the inequalities (29) and (30) respectively. $\square$

Example 4. Set $u(x) = Ax + C$ $(A > 0)$, $x \in (-C/A, \infty)$ and $v(x) = Bx + C$ $(B > 0)$, $x \in (-C/B, \infty)$ in Theorem 7. Then (29) becomes

$$\begin{aligned} &\int_{-\frac{C}{A}}^{\infty} \int_{-\frac{C}{B}}^{\infty} \frac{f(x)g(y)}{(Ax + By + 2C)^{\lambda}} dxdy < \\ &< \frac{k_{\lambda}(p)}{A^{1/q} B^{1/p}} \left(\int_{-\frac{C}{A}}^{\infty} (Ax + C)^{1-\lambda} f^p(x) dx \right)^{\frac{1}{p}} \left(\int_{-\frac{C}{B}}^{\infty} (Bx + C)^{1-\lambda} g^q(x) dx \right)^{\frac{1}{q}}. \end{aligned} \tag{31}$$

For $C = 0$ and $p = q = 2$ this is the result of Yang [11] and for $C = 0$ we get the result of Yang and Debnath [15]. Setting $A = B = 1, C = -\alpha$, we recover the result of Yang [9].Taking $A = B = 1, C = 0$, we get the result of Yang [10].

Example 5. Set $u(x) = x^{\alpha}$ $(\alpha > 0)$, $x \in (0, \infty)$ and $v(x) = x^{\beta}$ $(\beta > 0)$, $x \in (0, \infty)$, in Theorem 7. Then (29) becomes

$$\begin{aligned} &\int_0^{\infty} \int_0^{\infty} \frac{f(x)g(y)}{(x^{\alpha} + y^{\beta})^{\lambda}} dxdy < \frac{k_{\lambda}(p)}{\alpha^{\frac{1}{q}} \beta^{\frac{1}{p}}} \left(\int_0^{\infty} x^{\alpha(2-\lambda-p)+p-1} f^p(x) dx \right)^{\frac{1}{p}} \times \\ &\quad \times \left(\int_0^{\infty} x^{\beta(2-\lambda-q)+q-1} g^q(x) dx \right)^{\frac{1}{q}}. \end{aligned} \tag{32}$$

For $\alpha = \beta = 1$, this is the result of Yang [10]. Taking $\lambda = 1$ and $\alpha = \beta$, we get the result of Yang [12, Theorem 3].

Example 6. Set $u(x) = v(x) = lnx$, $x \in (1, \infty)$ in Theorem 7. Then (29) becomes

$$\begin{aligned} &\int_1^{\infty} \int_1^{\infty} \frac{f(x)g(y)}{(\ln x + \ln y)^{\lambda}} dxdy < k_{\lambda}(p) \left(\int_1^{\infty} (\ln x)^{1-\lambda} x^{p-1} f^p(x) dx \right)^{\frac{1}{p}} \times \\ &\quad \times \left(\int_1^{\infty} (\ln x)^{1-\lambda} x^{q-1} g^q(x) dx \right)^{\frac{1}{q}}. \end{aligned} \tag{33}$$

For $\lambda = 1$ this is the result of Yang [16, Theorem 3.1].

Theorem 8. *Suppose* $f,\ g \geq 0$ *satisfy* $0 < \int_a^b \frac{(u(x))^{p(1-\lambda)/2}}{(u'(x))^{p-1}} f^p(x) dx < \infty$ *and* $0 < \int_c^d \frac{(v(x))^{q(1-\lambda)/2}}{(v'(x))^{q-1}} g^q(x) dx < \infty$.

(i) *If* $p>1, \frac{1}{p}+\frac{1}{q}=1, \lambda>1-2\min\{\frac{1}{p},\frac{1}{q}\}$, *then we have the following two equivalent inequalities:*

$$\int_a^b\int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^\lambda}dxdy < \tilde{k}_\lambda(p)\left(\int_a^b \frac{(u(x))^{p(1-\lambda)/2}}{(u'(x))^{p-1}}f^p(x)dx\right)^{\frac{1}{p}} \times \times\left(\int_c^d \frac{(v(x))^{q(1-\lambda)/2}}{(v'(x))^{q-1}}g^q(x)dx\right)^{\frac{1}{q}}, \tag{34}$$

$$\int_c^d \frac{v'(y)}{(v(y))^{p(1-\lambda)/2}}\left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right]^p dy < < \left[\tilde{k}_\lambda(p)\right]^p \int_a^b \frac{(u(x))^{p(1-\lambda)/2}}{(u'(x))^{p-1}}f^p(x)dx\ , \tag{35}$$

where $\tilde{k}_\lambda(p)=B\left(\frac{\lambda-1}{2}+\frac{1}{p},\frac{\lambda-1}{2}+\frac{1}{q}\right)$.

(ii) *If* $0<p<1$ *and* $1-\frac{2}{p}<\lambda<1-\frac{2}{q}$, *we have two equivalent reverses of* (34) *and* (35).

(iii) *If* $p<0$ *and* $1-\frac{2}{q}<\lambda<1-\frac{2}{p}$, *we have reverse of* (34) *and the inequality* (35), *which are equivalent; where the constants in the above inequalities are all the best possible.*

Proof. Setting $\phi_r=\frac{\lambda-1}{2}+\frac{1}{r}$ $(r=p,q)$, in Theorem 4 and Theorem 5, we get the inequalities (34) and (35) respectively. □

Example 7. Set $u(x)=Ax+C$ $(A>0)$, $x\in(-C/A,\infty)$ and $v(x)=Bx+C$ $(B>0)$, $x\in(-C/B,\infty)$ in Theorem 8. Then (34) becomes

$$\int_{-\frac{C}{A}}^\infty\int_{-\frac{C}{B}}^\infty \frac{f(x)g(y)}{(Ax+By+2C)^\lambda}dxdy < < \frac{\tilde{k}_\lambda(p)}{A^{1/q}B^{1/p}}\left(\int_{-\frac{C}{A}}^\infty (Ax+C)^{p(1-\lambda)/2}f^p(x)dx\right)^{\frac{1}{p}} \times \times\left(\int_{-\frac{C}{B}}^\infty (Bx+C)^{q(1-\lambda)/2}g^q(x)dx\right)^{\frac{1}{q}}. \tag{36}$$

For $C=0$ and $p=q=2$ this is the result of Yang [11].

Example 8. Set $u(x)=x^\alpha$ $(\alpha>0)$, $x\in(0,\infty)$ and $v(x)=x^\beta$ $(\beta>0)$, $x\in(0,\infty)$ in Theorem 8. Then (34) becomes

$$\int_0^\infty\int_0^\infty \frac{f(x)g(y)}{(x^\alpha+y^\beta)^\lambda}dxdy < \frac{\tilde{k}_\lambda(p)}{\alpha^{\frac{1}{q}}\beta^{\frac{1}{p}}}\left(\int_0^\infty x^{p-1+\alpha(2-p\lambda-p)/2}f^p(x)dx\right)^{\frac{1}{p}} \times \times\left(\int_0^\infty x^{q-1+\beta(2-q\lambda-q)/2}g^q(x)dx\right)^{\frac{1}{q}}. \tag{37}$$

Taking $\lambda = 1$ and $\alpha = \beta$, we get the result of Yang [12, Theorem 3].

Example 9. Set $u(x) = v(x) = lnx,\ x \in (1, \infty)$ in Theorem 8. Then (34) becomes

$$\int_1^\infty \int_1^\infty \frac{f(x)g(y)}{(\ln x + \ln y)^\lambda} dxdy < \tilde{k}_\lambda(p) \left(\int_1^\infty (\ln x)^{p(1-\lambda)/2} x^{p-1} f^p(x) dx \right)^{\frac{1}{p}} \times \times \left(\int_1^\infty (\ln x)^{q(1-\lambda)/2} x^{q-1} g^q(x) dx \right)^{\frac{1}{q}}. \tag{38}$$

For $\lambda = 1$ this is the result of Yang [16, Theorem 3.1].

Theorem 9. *If* $p > 1, \frac{1}{p} + \frac{1}{q} = 1, \lambda > 0, f, g \geq 0$ *satisfy* $0 < \int_a^b \frac{(u(x))^{(p-1)(1-\lambda)}}{(u'(x))^{p-1}} f^p(x) dx < \infty$ *and* $0 < \int_c^d \frac{(v(x))^{(q-1)(1-\lambda)}}{(v'(x))^{q-1}} g^q(x) dx < \infty$, *then we have the following two equivalent inequalities :*

$$\int_a^b \int_c^d \frac{f(x)g(y)}{(u(x) + v(y))^\lambda} dxdy < B\left(\frac{\lambda}{p}, \frac{\lambda}{q}\right) \left(\int_a^b \frac{(u(x))^{(p-1)(1-\lambda)}}{(u'(x))^{p-1}} f^p(x) dx \right)^{\frac{1}{p}} \times \times \left(\int_c^d \frac{(v(x))^{(q-1)(1-\lambda)}}{(v'(x))^{q-1}} g^q(x) dx \right)^{\frac{1}{q}}, \tag{39}$$

$$\int_c^d \frac{v'(y)}{(v(y))^{1-\lambda}} \left[\int_a^b \frac{f(x)}{(u(x) + v(y))^\lambda} dx \right]^p dy < < \left[B\left(\frac{\lambda}{p}, \frac{\lambda}{q}\right) \right]^p \int_a^b \frac{(u(x))^{(p-1)(1-\lambda)}}{(u'(x))^{p-1}} f^p(x) dx, \tag{40}$$

where the constants in the above inequalities are all the best possible.

Proof. Setting $\phi_r = \frac{\lambda}{r}$ $(r = p, q)$, in Theorem 4 and Theorem 5, we get the inequalities (39) and (40) respectively. $\square$

Example 10. Set $u(x) = Ax + C$ $(A > 0)$, $x \in (-C/A, \infty)$ and $v(x) = Bx + C$ $(B > 0)$, $x \in (-C/B, \infty)$ in Theorem 9. Then (39) becomes

$$\int_{-\frac{C}{A}}^\infty \int_{-\frac{C}{B}}^\infty \frac{f(x)g(y)}{(Ax + By + 2C)^\lambda} dxdy < < \frac{1}{A^{1/q} B^{1/p}} B\left(\frac{\lambda}{p}, \frac{\lambda}{q}\right) \left(\int_{-\frac{C}{A}}^\infty (Ax + C)^{(p-1)(1-\lambda)} f^p(x) dx \right)^{\frac{1}{p}} \times \times \left(\int_{-\frac{C}{B}}^\infty (Bx + C)^{(q-1)(1-\lambda)} g^q(x) dx \right)^{\frac{1}{q}}. \tag{41}$$

For $C = 0$ and $p = q = 2$ this is the result of Yang [11].

Example 11. Set $u(x) = x^{\alpha}$ $(\alpha > 0)$, $x \in (0, \infty)$ and $v(x) = x^{\beta}$ $(\beta > 0)$, $x \in (0, \infty)$ in Theorem 9. Then (39) becomes

$$\int_0^{\infty}\int_0^{\infty} \frac{f(x)g(y)}{(x^{\alpha}+y^{\beta})^{\lambda}}dxdy < \frac{1}{\alpha^{\frac{1}{q}}\beta^{\frac{1}{p}}}B\left(\frac{\lambda}{p},\frac{\lambda}{q}\right)\left(\int_0^{\infty} x^{(p-1)(1-\alpha\lambda)}f^p(x)dx\right)^{\frac{1}{p}} \times \times \left(\int_0^{\infty} x^{(q-1)(1-\beta\lambda)}g^q(x)dx\right)^{\frac{1}{q}}. \tag{42}$$

For $\alpha = \beta = 1$, this is the result of Yang [17]; for $\alpha = \beta, \lambda = 1$ this gives the result of Yang [14]; for $\alpha = \beta = 2, \lambda = \frac{1}{2}$ this gives the result of Hong [5].

Example 12. Set $u(x) = ax^{1+x}, v(x) = bx^{1+x},\ x \in (0, \infty)$ and $\lambda = 1$ in Theorem 9. Then (39) becomes

$$\int_0^{\infty}\int_0^{\infty} \frac{f(x)g(y)}{ax^{1+x}+by^{1+y}}dxdy < < \frac{1}{a^{\frac{1}{q}}b^{\frac{1}{p}}}B\left(\frac{1}{p},\frac{1}{q}\right)\left(\int_0^{\infty} (x^x(1+x+xlnx))^{1-p}f^p(x)dx\right)^{\frac{1}{p}} \times \times \left(\int_0^{\infty} (x^x(1+x+xlnx))^{1-q}g^q(x)dx\right)^{\frac{1}{q}}. \tag{43}$$

This is the result of Jia and Gao [19].

Example 13. Set $u(x) = v(x) = lnx,\ x \in (1, \infty)$ in Theorem 9. Then (39) becomes

$$\int_1^{\infty}\int_1^{\infty} \frac{f(x)g(y)}{(\ln x+\ln y)^{\lambda}}dxdy < B\left(\frac{\lambda}{p},\frac{\lambda}{q}\right)\left(\int_1^{\infty} (\ln x)^{(p-1)(1-\lambda)}x^{p-1}f^p(x)dx\right)^{\frac{1}{p}} \times \times \left(\int_1^{\infty} (\ln x)^{(q-1)(1-\lambda)}x^{q-1}g^q(x)dx\right)^{\frac{1}{q}}. \tag{44}$$

For $\lambda = 1$ this is the result of Yang [16, Theorem 3.1].

Theorem 10. *If* $p > 1,\ \frac{1}{p}+\frac{1}{q} = 1,\ \lambda > 0,\ f, g \geq 0$ *satisfy* $0 < \int_a^b \frac{(u(x))^{p-\lambda-1}}{(u'(x))^{p-1}}f^p(x)dx < \infty$ *and* $0 < \int_c^d \frac{(v(x))^{q-\lambda-1}}{(v'(x))^{q-1}}g^q(x)dx < \infty$, *then we have the following two equivalent inequalities:*

$$\int_a^b\int_c^d \frac{f(x)g(y)}{(u(x)+v(y))^{\lambda}}dxdy < B\left(\frac{\lambda}{p},\frac{\lambda}{q}\right)\left(\int_a^b \frac{(u(x))^{p-\lambda-1}}{(u'(x))^{p-1}}f^p(x)dx\right)^{\frac{1}{p}} \times \times \left(\int_c^d \frac{(v(x))^{q-\lambda-1}}{(v'(x))^{q-1}}g^q(x)dx\right)^{\frac{1}{q}}, \tag{45}$$

$$\int_c^d \frac{v'(y)}{(v(y))^{1-\lambda(p-1)}} \left[\int_a^b \frac{f(x)}{(u(x)+v(y))^\lambda}dx\right]^p dy < \\ < \left[B\left(\frac{\lambda}{p},\frac{\lambda}{q}\right)\right]^p \int_a^b \frac{(u(x))^{p-\lambda-1}}{(u'(x))^{p-1}} f^p(x)dx, \tag{46}$$

where the constants in the above inequalities are all the best possible.

Proof. Setting $\phi_r = \lambda\left(1-\frac{1}{r}\right)$ $(r=p,q)$, in Theorem 4 and Theorem 5, we get the inequalities (45) and (46) respectively. □

Example 14. Set $u(x)=Ax+C$ $(A>0)$, $x\in(-C/A,\infty)$ and $v(x)=Bx+C$ $(B>0)$, $x\in(-C/B,\infty)$ in Theorem-10. Then (45) becomes

$$\int_{-\frac{C}{A}}^\infty \int_{-\frac{C}{B}}^\infty \frac{f(x)g(y)}{(Ax+By+2C)^\lambda}dxdy < \\ < \frac{1}{A^{1/q}B^{1/p}} B\left(\frac{\lambda}{p},\frac{\lambda}{q}\right)\left(\int_{-\frac{C}{A}}^\infty (Ax+C)^{p-\lambda-1}f^p(x)dx\right)^{\frac{1}{p}} \times \\ \times \left(\int_{-\frac{C}{B}}^\infty (Bx+C)^{q-\lambda-1}g^q(x)dx\right)^{\frac{1}{q}}. \tag{47}$$

For $C=0$ and $p=q=2$ this is the result of Yang [11].

Example 15. Set $u(x)=x^\alpha\,(\alpha>0)$, $x\in(0,\infty)$ and $v(x)=x^\beta\,(\beta>0)$, $x\in(0,\infty)$ in Theorem 10. Then (45) becomes

$$\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x^\alpha+y^\beta)^\lambda}dxdy < \frac{1}{\alpha^{\frac{1}{q}}\beta^{\frac{1}{p}}} B\left(\frac{\lambda}{p},\frac{\lambda}{q}\right) \times \\ \times \left(\int_0^\infty x^{p-\alpha\lambda-1}f^p(x)dx\right)^{\frac{1}{p}}\left(\int_0^\infty x^{q-\beta\lambda-1}g^q(x)dx\right)^{\frac{1}{q}}. \tag{48}$$

This is the result of Azar [23, Theorem 1], with the constant factor $\frac{1}{\alpha^{1/q}\beta^{1/p}}B\left(\frac{\lambda}{p},\frac{\lambda}{q}\right)$ is the best possible for $\alpha=\beta$, but we proved that the constant factor is the best possible for all α and β . For $\alpha=\beta=1$, we get the result of Yang [17].

Example 16. Set $u(x)=v(x)=lnx$, $x\in(1,\infty)$ in Theorem 10. Then (45) becomes

$$\int_1^\infty \int_1^\infty \frac{f(x)g(y)}{(\ln x+\ln y)^\lambda}dxdy < B\left(\frac{\lambda}{p},\frac{\lambda}{q}\right)\left(\int_1^\infty (\ln x)^{p-\lambda-1}x^{p-1}f^p(x)dx\right)^{\frac{1}{p}} \times \\ \times \left(\int_1^\infty (\ln x)^{q-\lambda-1}x^{q-1}g^q(x)dx\right)^{\frac{1}{q}}. \tag{49}$$

Remark 1. For $\lambda = 1, u(x) = v(x) = x^\alpha, \phi_r = \frac{1}{\alpha r}$ $(r = p, q)$, Theorem 2 gives

$$\begin{aligned} \int_0^\infty \int_0^\infty \frac{f(x)g(y)}{x^\alpha + y^\alpha} dxdy &< \frac{\pi}{\alpha \sin^{\frac{1}{p}}(\pi/\alpha p) \sin^{\frac{1}{q}}(\pi/\alpha q)} \times \\ &\times \left(\int_0^\infty x^{1-\alpha} f^p(x) dx \right)^{\frac{1}{p}} \left(\int_0^\infty x^{1-\alpha} g^q(x) dx \right)^{\frac{1}{q}}, \end{aligned} \tag{50}$$

which is the result of Kuang[4].

Remark 2. For $\lambda = b + c + 1, u(x) = v(x) = x, \phi_p = c + 1 - \frac{1}{p}, \phi_q = b + 1 - \frac{1}{q}$, Theorem 2 gives

$$\begin{aligned} \int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x+y)^{b+c+1}} dxdy &< B\left(b + \frac{1}{p}, c + \frac{1}{q}\right) \times \\ &\times \left(\int_0^\infty x^{p(1-b)-2} f^p(x) dx \right)^{\frac{1}{p}} \left(\int_0^\infty x^{q(1-c)-2} g^q(x) dx \right)^{\frac{1}{q}}, \end{aligned} \tag{51}$$

which is given by Peachey [24].

Remark 3. For $u(x) = v(x) = x^\alpha$, $\phi_p = \frac{1-mp}{\alpha}$, $\phi_q = \frac{1-nq}{\alpha}$, Theorem 2 gives the following results:

If $p > 1, 1/p + 1/q = 1, \alpha > 0, \lambda > 0, m, n \in \mathbb{R}$ such that $0 < 1 - mp < \alpha\lambda$, $0 < 1 - nq < \alpha\lambda$ and $f \geq 0, g \geq 0$ satisfy $0 < \int_0^\infty x^{(1-\alpha\lambda)+p(n-m)} f^p(x) dx < \infty$ and $0 < \int_0^\infty y^{(1-\alpha\lambda)+q(m-n)} g^q(x) dx < \infty$ then

$$\begin{aligned} \int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x^\alpha + y^\alpha)^\lambda} dxdy &< H_{\lambda,\alpha}(m,n,p,q) \left(\int_0^\infty x^{(1-\alpha\lambda)+p(n-m)} f^p(x) dx \right)^{\frac{1}{p}} \times \\ &\times \left(\int_0^\infty y^{(1-\alpha\lambda)+q(m-n)} g^q(y) dy \right)^{\frac{1}{q}}, \end{aligned} \tag{52}$$

where $H_{\lambda,\alpha}(m,n,p,q) = \frac{1}{\alpha} B^{\frac{1}{p}}(\frac{1-mp}{\alpha}, \lambda - \frac{1-mp}{\alpha}) B^{\frac{1}{q}}(\frac{1-nq}{\alpha}, \lambda - \frac{1-nq}{\alpha})$.

Further if $mp + nq = 2 - \alpha\lambda$, then Theorem 4 gives

$$\begin{aligned} \int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x^\alpha + y^\alpha)^\lambda} dxdy &< \frac{1}{\alpha} B\left(\frac{1-mp}{\alpha}, \frac{1-nq}{\alpha}\right) \left(\int_0^\infty x^{n(p+q)-1} f^p(x) dx \right)^{\frac{1}{p}} \times \\ &\times \left(\int_0^\infty y^{m(p+q)-1} g^q(y) dy \right)^{\frac{1}{q}}, \end{aligned} \tag{53}$$

where the constant factor $\frac{1}{\alpha} B(\frac{1-mp}{\alpha}, \frac{1-nq}{\alpha})$ is the best possible. These two inequalities are given by Hong [6].

Remark 4. Replacing $u(x)$ by $xu(x)$ and $v(x)$ by $xv(x)$ and taking $\phi_p = 1 - A_2 p$, $\phi_q = 1 - A_1 q$ in Theorem 2, we get the following result given by Mario Krnic et. al [27]:

If $p > 1, 1/p + 1/q = 1, \lambda > 0, A_1 \in \left(\frac{1-\lambda}{q}, \frac{1}{q}\right)$ *and* $A_2 \in \left(\frac{1-\lambda}{p}, \frac{1}{p}\right)$ then

$$\begin{aligned} &\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(xu(x)+yv(y))^\lambda} dxdy < \\ &< L\left(\int_0^\infty (xu(x))^{1-\lambda+p(A_1-A_2)}(u(x)+xu'(x))^{1-p} f^p(x)dx\right)^{1/p} \times \\ &\times \left(\int_0^\infty (xv(x))^{1-\lambda+q(A_2-A_1)}(v(x)+xv'(x))^{1-q} g^q(x)dx\right)^{1/q}, \end{aligned} \tag{54}$$

where $L = \left(B\left(1-A_2p, \lambda-1+A_2p\right)\right)^{\frac{1}{p}} \left(B\left(1-A_1q, \lambda-1+A_1q\right)\right)^{\frac{1}{q}}$.

Remark 5. For $u(x) = Aa^x, v(x) = Bb^x, \phi_r = 1+(1-\frac{1}{r})(\lambda-2)\,(r = p, q)$, Theorem 4 gives the following inequality:

If $p > 1, \frac{1}{p} + \frac{1}{q} = 1, \lambda > 2 - \min\{p, q\}, A > 0, B > 0, a > 1, b > 1$ and $f, g \geq 0$ satisfy $0 < \int_{-\infty}^\infty a^{(2-\lambda-p)x} f^p(x)dx < \infty$ and $0 < \int_{-\infty}^\infty b^{(2-\lambda-q)x} g^q(x)dx < \infty$, then

$$\begin{aligned} &\int_{-\infty}^\infty \int_{-\infty}^\infty \frac{f(x)g(y)}{(Aa^x+Bb^y)^\lambda} dxdy < \\ &< C\left(\int_{-\infty}^\infty a^{(2-\lambda-p)x} f^p(x)dx\right)^{\frac{1}{p}} \left(\int_{-\infty}^\infty b^{(2-\lambda-q)x} g^q(x)dx\right)^{\frac{1}{q}}, \end{aligned} \tag{55}$$

where the constant factor $C = \left(\frac{A^{1-\lambda}}{B\ln b}\right)^{\frac{1}{p}} \left(\frac{B^{1-\lambda}}{A\ln a}\right)^{\frac{1}{q}} B\left(\frac{p+\lambda-2}{p}, \frac{q+\lambda-2}{q}\right)$ is the best possible. This inequality is an extension of the result of Zhou et.al[22], where they consider the parameter $p \geq q > 1, 1 - \frac{q}{p} < \lambda \leq 2$.

Remark 6. For $u(x) = v(x)$, Theorem 4 gives (52).

For other appropriate values of $\lambda, \phi_p, \phi_q, u(x)$ and $v(x)$ taken in Theorem 2–5, many new inequalities can be obtained.

5 Applications

In this section, we will give the generalizations of Hardy-Littlewood's inequality. Let $f \in L^2(0,1)$ and $f(x) \neq 0$. If

$$a_n = \int_0^1 x^n f(x)dx, \quad n = 0, 1, 2, 3, \ldots$$

then we have the Hardy-Littlewood's inequality (see [1]) of the form

$$\sum_{n=0}^\infty a_n^2 < \pi \int_0^1 f^2(x)dx \tag{56}$$

where the constant factor π is the best possible.

In [20, 21], Gao gave the integral version of Hardy-Littlewood's inequality as follows :

Let $h \in L^2(0,1)$ and $h \neq 0$. If

$$f(x) = \int_0^1 t^x |h(t)| dt, \quad x \in [0, \infty),$$

then

$$\int_0^\infty f^2(x) dx < \pi \int_0^1 h^2(t) dt \tag{57}$$

and

$$\int_0^\infty f^2(x) dx < \pi \int_0^1 t h^2(t) dt. \tag{58}$$

Theorem 11. *Let $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, $h \in L^2(0,1)$ and $h(t) \neq 0$. Define a function $f(x)$ by*

$$f(x) = (u'(x))^{\frac{1}{p}} \int_0^1 t^{u(x)} |h(t)| dt, \quad x \in (a, b).$$

If $0 < \int_a^b (u'(x))^{2-p} f^{p(p-1)}(x) dx < \infty$, then

$$\left(\int_a^b f^p(x) dx \right)^{1+\frac{1}{p}} < \frac{\pi}{\sin \frac{\pi}{p}} \left(\int_a^b (u'(x))^{2-p} f^{p(p-1)}(x) dx \right)^{\frac{1}{p}} \int_0^1 t h^2(t) dt. \tag{59}$$

Proof. We can write

$$f^p(x) = f^{p-1}(x) u'(x)^{\frac{1}{p}} \int_0^1 t^{u(x)} |h(t)| dt.$$

Now applying, Schwartz inequality and Corollary-1, we have

$$\begin{aligned}
&\left(\int_a^b f^p(x) dx \right)^2 = \\
&= \left\{ \int_0^1 \left(\int_a^b f^{p-1}(x) u'(x)^{\frac{1}{p}} t^{u(x)-\frac{1}{2}} dx \right) t^{\frac{1}{2}} |h(t)| dt \right\}^2 \leq \\
&\leq \int_0^1 \left(\int_a^b f^{p-1}(x) u'(x)^{\frac{1}{p}} t^{u(x)-\frac{1}{2}} dx \right)^2 dt \times \int_0^1 t h^2(t) dt = \\
&= \left(\int_a^b \int_a^b \frac{f^{p-1}(x) u'(x)^{\frac{1}{p}} f^{p-1}(y) u'(y)^{\frac{1}{p}}}{u(x) + u(y)} dx dy \right) \times \int_0^1 t h^2(t) dt \leq \\
&\leq \frac{\pi}{\sin \frac{\pi}{p}} \left(\int_a^b (u'(x))^{1-p} f^{p(p-1)}(x) u'(x) dx \right)^{\frac{1}{p}} \times \\
&\times \left(\int_a^b (u'(x))^{1-q} f^{q(p-1)}(x) (u'(x))^{\frac{q}{p}} dx \right)^{\frac{1}{q}} \times \int_0^1 t h^2(t) dt = \\
&= \frac{\pi}{\sin \frac{\pi}{p}} \left(\int_a^b (u'(x))^{2-p} f^{p(p-1)}(x) dx \right)^{\frac{1}{p}} \left(\int_a^b f^p(x) dx \right)^{\frac{1}{q}} \times \int_0^1 t h^2(t) dt.
\end{aligned} \tag{60}$$

Since $h(t) \neq 0$, so, $f(x) \neq 0$. Hence it is impossible for equality in (60) and then we get the inequality (59). This completes the theorem. □

Theorem 12. *Let $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, $h \in L^2(0,1)$ and $h(t) \neq 0$. Define a function $f(x)$ by*

$$f(x) = (u(x))^{\frac{1}{2}-\frac{1}{p}} (u'(x))^{\frac{1}{p}} \int_0^1 t^{u(x)} |h(t)| dt, \quad x \in (a,b).$$

If $0 < \int_a^b \left(\frac{u'(x)}{u(x)}\right)^{2-p} f^{p(p-1)}(x) dx < \infty$, *then*

$$\left(\int_a^b f^p(x) dx \right)^{1+\frac{1}{p}} < \pi \left(\int_a^b \left(\frac{u'(x)}{u(x)} \right)^{2-p} f^{p(p-1)}(x) dx \right)^{\frac{1}{p}} \int_0^1 t h^2(t) dt. \tag{61}$$

Proof. Proceeding as in Theorem refhlthm1 and using Corollary 2, we complete the theorem. □

Remark 7. Taking $p = 2$ in Theorem 11 and Theorem 12, we get

$$\int_0^\infty f^2(x) dx < \pi \int_0^1 t h^2(t) dt \tag{62}$$

which is a generalization of Hardy-Littlewood inequality (58).

References

[1] Hardy G.H., Littlewood J.E., Polya G. *Inequalities.* Cambridge University Press, Cambrige, 1952.

[2] Kuang J. *Applied Inequalities.* Shangdong Science and Technology Press, Jinan, 2004.

[3] Wang Z., Guo D. *An Introduction to Special Functions.* Science Press, Beijing, 1979.

[4] Kuang J. *On new extensions of Hilbert's Inequalities.* J. Math. Anal. Appl., 1999, **235**, No. 2, 608–614.

[5] Hong Y. *All-sided Generalization about Hardy-Hilbert's integral inequalities.* Acta Math. Sinica (China), 2001, **44**, 619–626.

[6] Hong Y. *On Hardy-Hilbert's integral inequalities with some parameters.* J. Ineq. Pure and Appl. Math., 2005, **6**, No. 4, Art-92.

[7] Yang B. *On generalization of Hardy-Hilbert's integral inequalities.* Acta Math. Sinica (China),, 1998, **41**, No. 4, 839–844.

[8] Yang B., Debnath L. *On a new generalization of Hilbert's integral inequality and it's application.* RGMIA Research Report Collection, 1999, **2**, No. 5, Article 13.

[9] Yang B. *On new generalizations of Hilbert's inequality.* J. Math. Anal. Appl., 2000, **248**, 29–40.

[10] Yang B. *On a general Hardy-Hilbert's integral inequality with a best value.* Chin. Ann. of Math., 2000, **21A**, 401–408.

[11] Yang B. *On Hardy-Hilbert's integral inequality.* J. Math. Anal. Appl., 2001, **261**, 295–306.

[12] Yang B. *On an extension of Hardy-Hilbert's inequality.* RGMIA Research Report Collection, 2001, **4**, No. 3, Article 7.

[13] Yang B. *On Generalization of the Hardy-Hilbert's integral inequality.* RGMIA Research Report Collection, 2001, **4**, No. 4, Article 1.

[14] Yang B. *On a generalization of Hardy-Hilbert's inequality.* Ann. of Math. (China), 2002, **23**, 247–254.

[15] Yang B., Debnath L. *On the extended Hardy-Hilbert's inequality.* J. Math. Anal. Appl., 2002, **272**, 187–199.

[16] Yang B. *On a new inequality similar to Hardy-Hilbert's inequality.* Math. Ineq. and Appl., 2003, **6**, No. 1, 37–44.

[17] Yang B. *On the extended Hilbert's integral inequality.* J. Ineq. Pure and Appl. Math., 2004, **5**, No. 4, Art-96.

[18] Yang B. *A relation to Hardy-Hilbert's integral inequality and Mulholland's inequality.* J. Ineq. Pure and Appl. Math., 2005, **6**, No. 4, Art-112.

[19] Jia W., Gao M. *An extended Hardy-Hilbert's inequality and it's applications.* J. Ineq. Pure and Appl. Math., 2006, **7**, No. 1, Art-30.

[20] Gao M. *On Hilbert's Inequality and Its Applications.* J. Math. Anal. Appl., 1997, **212**, 316–323.

[21] Gao M., Tan Li, Debnath L. *Some Improvements on Hilbert's Integral Inequality.* J. Math. Anal. Appl., 1999, **229**, 682–689.

[22] Zhou Yu, Gao M. *An Extension on the Hardy-Hilbert's Integral Inequality and its Application.* J. Math. Ineq., 2007, **1**, No. 2, 149–158.

[23] Azar L.E. *On an Extension of Hardy-Hilbert's Integral Inequality with Three Parameters.* Int. Math. Forum, 2007, **2**, No. 48, 2369–2378.

[24] Peachey T.C. *Some integral inequalities related to Hilbert's.* J. Ineq. Pure and Appl. Math., 2003, **4**, No. 1, Art-19.

[25] Brnetic I., Pecaric J. *Generalization of Hilbert's integral inequality.* Math. Ineq. and Appl., 2004, **7**, No. 2, 199–205.

[26] Brnetic I., Pecaric J. *Generalization of inequalities of Hardy-Hilbert type.* Math. Ineq. and Appl., 2004, **7**, No. 2, 217–225.

[27] Mario K., Gao M., Pecaric J., Gao X. *On the Best Constant in Hilbert's inequality.* Math. Ineq. and Appl., 2005, **8**, No. 2, 317–29.

Namita Das
P.G. Department of Mathematics
Utkal University, Vani Vihar
Bhubaneswar, Orissa, India 751004
E-mail: *namitadas440@yahoo.co.in*

Received November 29, 2008

Srinibas Sahoo
Department of Mathematics
Banki Autonomous College, Banki
Orissa, India 754008
E-mail: *sahoosrinibas@yahoo.co.in*

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 111–130
ISSN 1024–7696

The cubic differential system with six real invariant straight lines along three directions

V. Puţuntică, A. Şubă

Abstract. We classify all cubic systems possessing exactly six real invariant straight lines along three directions taking into account their degree of invariance. We prove that there are 6 affine different classes of such systems. For every class we carried out the qualitative investigation in the Poincaré disc.

Mathematics subject classification: 34C05.
Keywords and phrases: Cubic differential system, invariant line.

1 Introduction

We consider the real polynomial system of differential equations

$$\frac{dx}{dt} = P(x,y), \quad \frac{dy}{dt} = Q(x,y), \tag{1}$$

and the vector field

$$X = P(x,y)\frac{\partial}{\partial x} + Q(x,y)\frac{\partial}{\partial y}$$

associated to system (1).

Denote $n = max\{deg(P),\, deg(Q)\}$. If $n = 2$ ($n = 3$) then system (1) is called quadratic (cubic).

An algebraic curve $f(x,y) = 0$, $f \in \mathbb{C}[x,y]$ (a function $f = exp(g/h)$; $g, h \in \mathbb{C}[x,y]$) is called invariant algebraic curve (invariant exponential function) of the system (1) if there exists a polynomial $K_f \in \mathbb{C}[x,y]$, $deg(K_f) \leq n-1$ such that the identity holds

$$X(f) \equiv f(x,y)K_f(x,y). \tag{2}$$

It should be observed that if in (2) for invariant algebraic curve $f(x,y) = 0$ we have $K_f(x,y) \equiv f^m(x,y)K(x,y)$ for any natural number $m \in \mathbb{N}$ and polynomial $K(x,y)$, then $exp(1/f), ..., exp(1/f^m)$ are invariant exponential functions. If, in addition, the polynomial $f(x,y)$ does not divide $K(x,y)$, then we say that the invariant algebraic curve $f(x,y) = 0$ has the degree of invariance equal to $m+1$.

Let $f \in \mathbb{C}[x,y]$ and $f = f_1^{n_1} \cdots f_s^{n_s}$ be its factorization in irreducible factors over $\mathbb{C}[x,y]$. Then $f(x,y) = 0$ is an invariant algebraic curve for (1) if and only if each of the algebraic curves $f_j(x,y) = 0$, $j = \overline{1,s}$, has this property.

It is easy to see that there is no correlation between the degree of invariance of the invariant algebraic curve $f(x,y)=0$ and the degree of invariance of its factors $f_j(x,y)=0$, $j=\overline{1,s}$, in general case. For example, for a system $\dot{x}=x^3$, $\dot{y}=y(2x^2+y^2)$, we have that $x^2+y^2=0$ is an algebraic curve with the degree of invariance equal to two, while for each of its factors $x\pm iy=0$, $i^2=-1$, the degree of invariance is equal to one. For the system [5]: $\dot{x}=2x^3$, $\dot{y}=y(3x^2+y^2)$, each of the invariant straight lines $x\pm iy=0$ has the degree of invariance equal to two, and their product $x^2+y^2=0$ has the degree of invariance equal to one.

Let $f_1(x,y)=0,...,f_k(x,y)=0$ be some irreducibles invariant algebraic curves; $f_{k+1}(x,y)=exp(g_{k+1}/h_{k+1}),...,f_s(x,y)=exp(g_s/h_s)$ be some invariant exponential functions of the system (1) and let $\lambda_1,...,\lambda_s$ be some real or complex numbers. We compose the function

$$F=f_1^{\lambda_1}\cdots f_s^{\lambda_s}. \tag{3}$$

If $F\not\equiv const$ and $X(F)\equiv 0$ $(X(F)\equiv -F(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}))$, i.e. $F(x,y)=const$ is a first integral (F is an integrating factor) for (1), then we say that the system (1) is Darboux integrable. In order that (3) be a first integral (an integrating factor) for (1), it is necessary and sufficient that cofactors $K_{f_1},...,K_{f_s}$ and numbers $\lambda_1,...,\lambda_s$ verify the identity

$$\lambda_1K_{f_1}(x,y)+\cdots+\lambda_sK_{f_s}(x,y)\equiv 0$$

$$\Big(\lambda_1K_{f_1}(x,y)+\cdots+\lambda_sK_{f_s}(x,y)\equiv -\frac{\partial P}{\partial x}-\frac{\partial Q}{\partial y}\Big).$$

Later on, we will be interested in invariant algebraic curve of degree one, that is invariant straight lines $\alpha x+\beta y+\gamma=0$.

A set of invariant straight lines can be infinite, finite or empty. Systems with infinite number of invariant straight lines will not be considered.

At present a great number of works are dedicated to the investigation of polynomial differential systems with invariant straight lines. Here we indicate some problems and works concerning the polynomial differential system with invariant straight lines.The problem of estimation for the number of invariant straight lines which can have a polynomial differential system was considered in [2]; the problem of coexistence of the invariant straight lines and limit cycles in {[9]: $n=2$}; {[4]: $n=3$}; [10]; the problem of coexistence of the invariant straight lines and the singular points of a center type for the cubic system in [3, 11] An interesting relation between the number of invariant straight lines and the possible number of directions for them is established in [1].

The classification of all cubic systems possessing the maximum number of invariant straight lines taking into account their multiplicities is given in [5].

The cubic system with exactly eight and exactly seven invariant straight lines has been studied in [5-7] and with six invariant straight lines along two directions in [8].

In this paper a qualitative investigation of cubic systems with exactly six real invariant straight lines along three direction is given.

The main obtained results are shown in the following theorem:

Theorem. *Any cubic system having real invariant straight lines along three directions with total degree of invariance six via affine transformation and time rescaling can be written as one of the following six systems. The bifurcation diagrams in the space of parameters and the phase portraits in the Poincaré disc are presented in the figures for each system.*

$$\begin{cases} \dot{x} = x(x+1)(x-a), \ a > 0, \\ \dot{y} = y(y+1)(-a+dx+(1-d)y), & Fig.1 \\ d(d-1)(a+d-1)(a-d+2) \neq 0; & (Fig.4.1; Tab.4.1, 4.2) \end{cases} \tag{4}$$

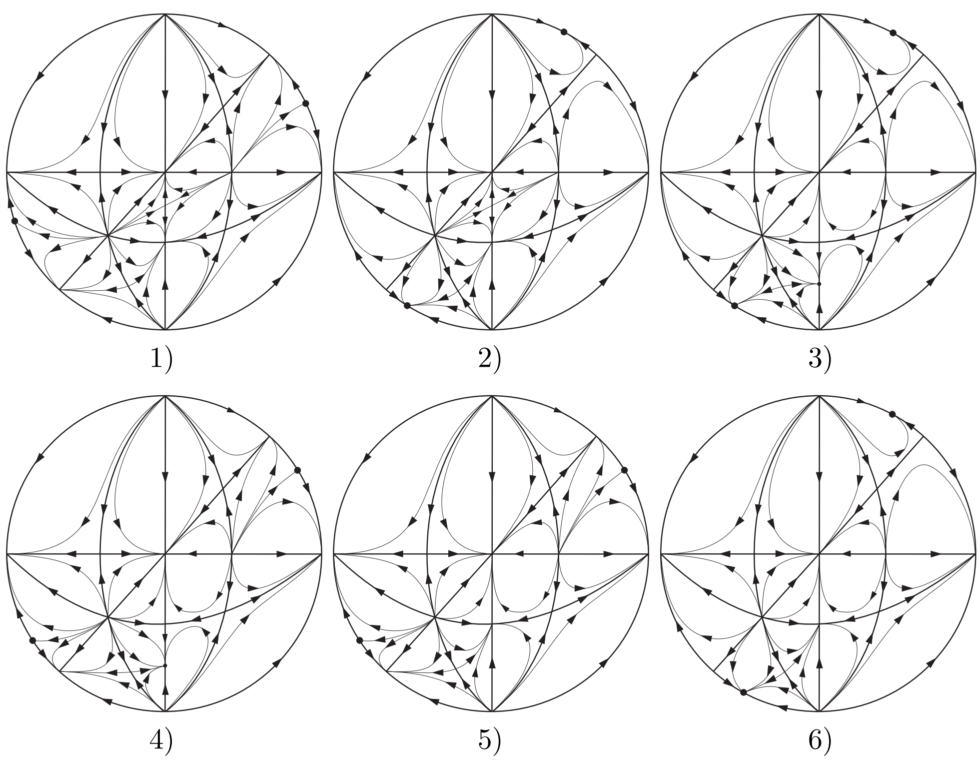
$$\begin{cases} \dot{x} = x(x+1)(x-a), \ a > 0, \\ \dot{y} = y(y+1)(a(a-d+1)+dx+(a+1)(a-d+1)y), & Fig.2 \\ d(a-d+1)(a-d+2)(2a-d+1) \neq 0; & (Fig.4.2; Tab.4.3, 4.4) \end{cases} \tag{5}$$

$$\begin{cases} \dot{x} = x^2(x+1), \ d(d-1) \neq 0, \\ \dot{y} = y(y+1)(dx+(1-d)y); & Fig.3\,(Tab.4.5) \end{cases} \tag{6}$$

$$\begin{cases} \dot{x} = x^2(x+1), \ d(d-1) \neq 0, \\ \dot{y} = y^2(1+dx+(1-d)y); & Fig.4\,(Tab.4.6) \end{cases} \tag{7}$$

$$\begin{cases} \dot{x} = x^3, \\ \dot{y} = y^2(2x-y); & Fig.5 \end{cases} \tag{8}$$

$$\begin{cases} \dot{x} = x(x+1)(a-ax+y), \\ \dot{y} = y(y+1)(a+(2-3a)x+(2a-1)y), & Fig.6 \\ a(3a-1)(2a-1)(2-3a)(a-1) \neq 0. & (Tab.4.7, 4.8) \end{cases} \tag{9}$$



1) 2) 3)

4) 5) 6)

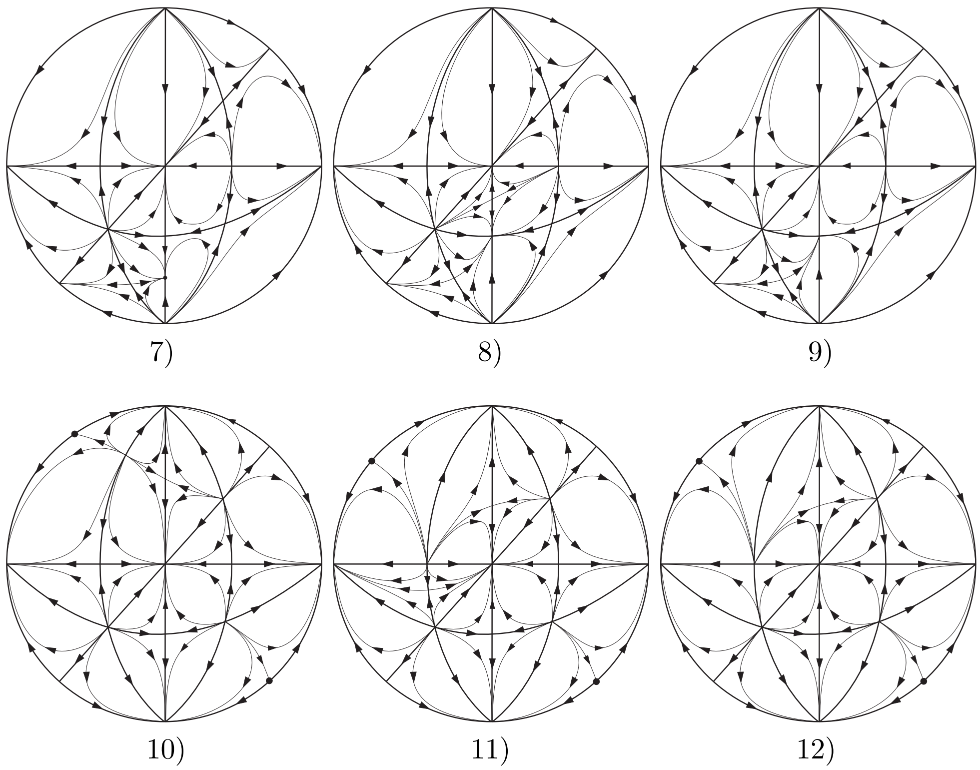
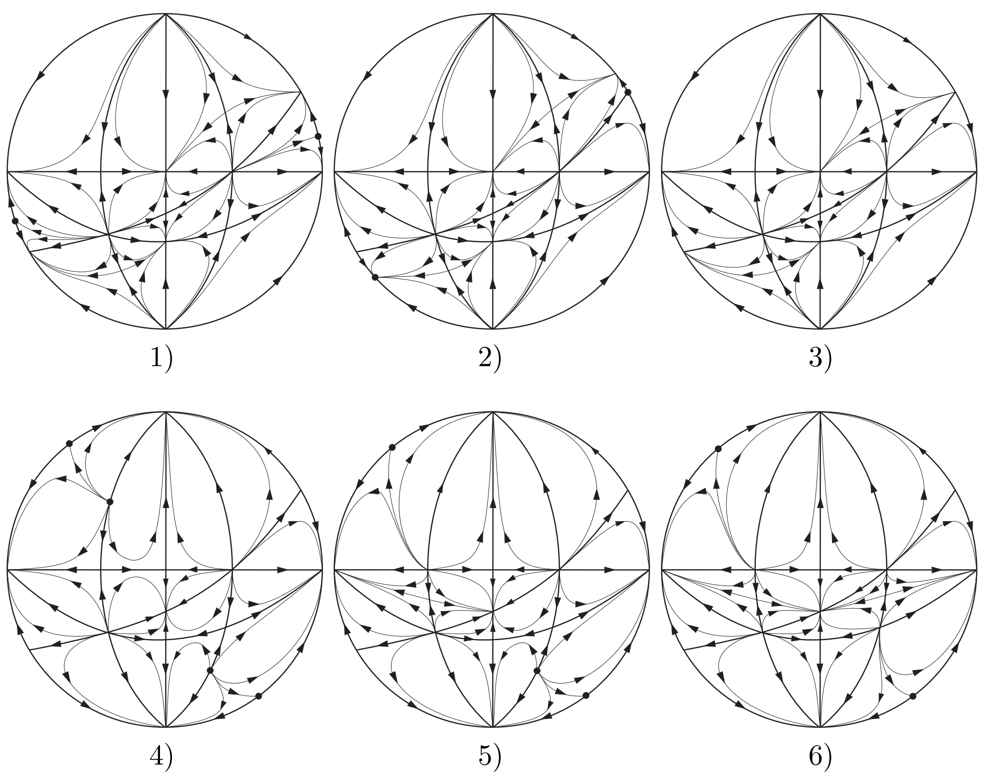
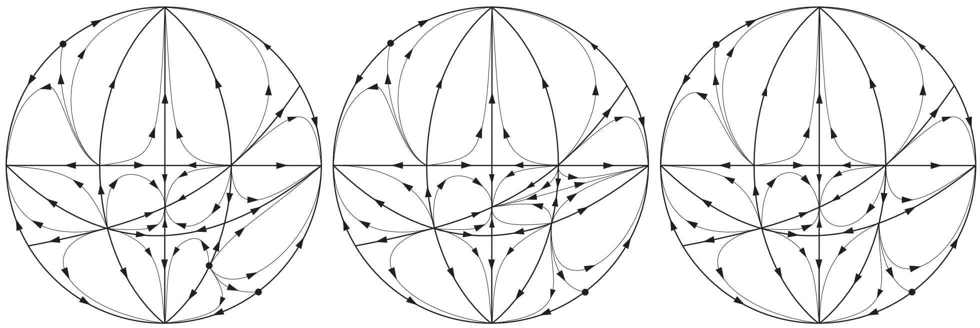


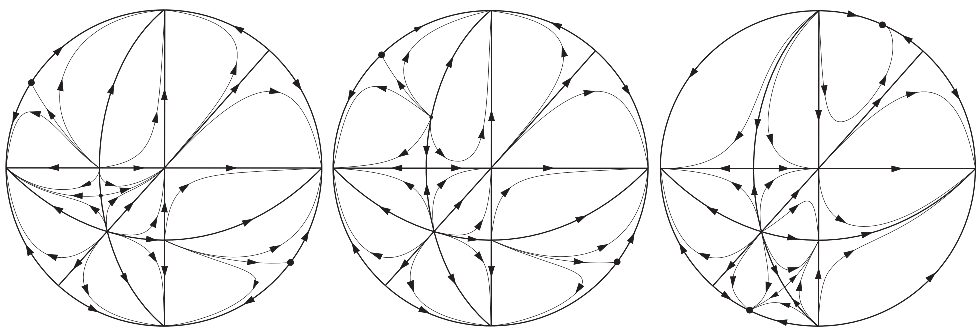
Fig. 1.



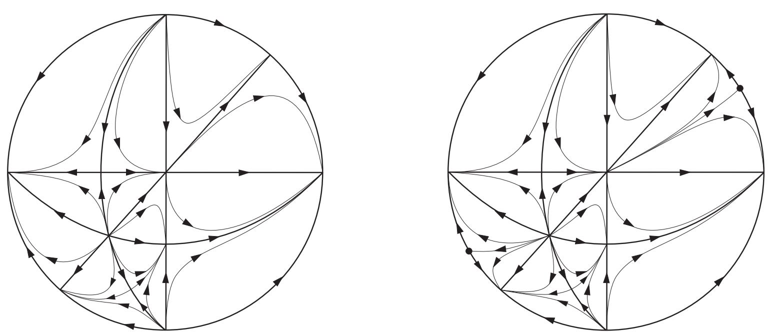


7) 8) 9)

Fig. 2.

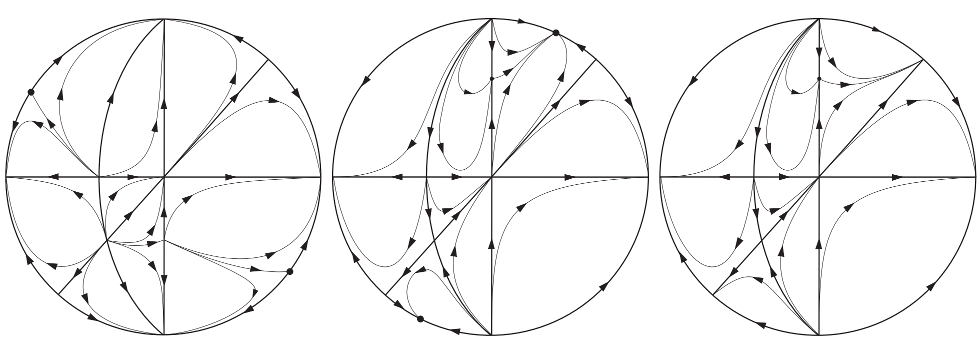


1) 2) 3)

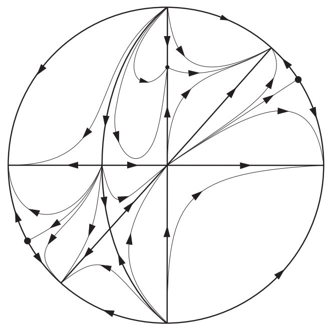


4) 5)

Fig. 3.



1) 2) 3)



4)
Fig. 4

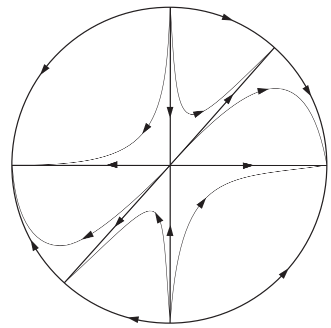
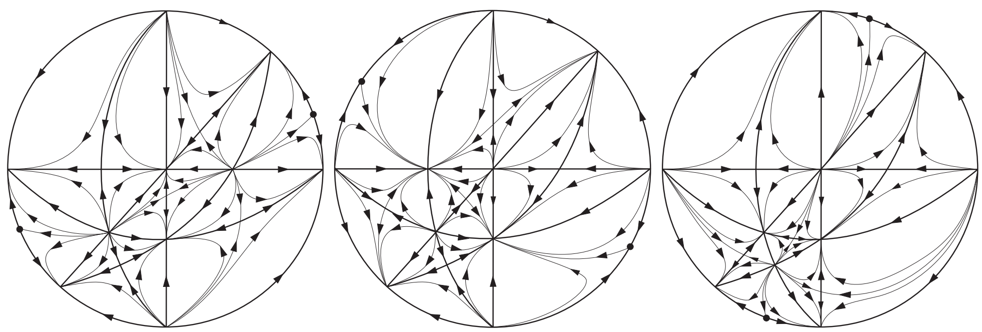


Fig. 5.



1) ⇄ 2) 3)

Fig. 6.

2 Preliminaries

We consider the real cubic differential system

$$\begin{cases} \dfrac{dx}{dt} = \sum_{r=0}^{3} P_r(x,y) \equiv P(x,y), \\ \dfrac{dy}{dt} = \sum_{r=0}^{3} Q_r(x,y) \equiv Q(x,y), \ GCD(P,Q)=1, \end{cases} \tag{10}$$

where $P_r(x,y) = \sum\limits_{j+l=r} a_{jl}x^jy^l$, $Q_r(x,y) = \sum\limits_{j+l=r} b_{jl}x^jy^l$. It is assumed that the right-hand sides of the system (10) have not a non-constant common factor.

We will mention some properties of the system (10):

2.1) in the finite part of the phase plane the system (10) has at most nine singular points;

2.2) at infinity the system (10) has at most four singular points if $yP_3(x,y)$ $-xQ_3(x,y) \not\equiv 0$. In the case $yP_3(x,y) - xQ_3(x,y) \equiv 0$ the infinity is degenerate, i.e. consists only of singular points;

2.3) in the finite part of the phase plane the system (10) can not have more than three colinear singular points;

2.4) in the finite part of the phase plane the system (10) has no more than eight invariant straight lines [5, 6];

2.5) the infinity for (10) represents an invariant straight line;

2.6) the system (10) has invariant straight lines along at most six different directions [1, 12];

2.7) the system (10) can not have more than three invariant straight lines parallel among themselves.

Let $a_j x + b_j y + c_j = 0,\ j = 1, 2,\ a_1 b_2 - a_2 b_1 \neq 0$ be two real invariant straight lines of the system (10). The transformation $X = a_1 x + b_1 y + c_1$, $Y = a_2 x + b_2 y + c_2$ reduces (10) to a system of the Lotka-Volterra form

$$\begin{cases} \dot{x} = x(a_{10} + a_{20}x + a_{11}y + a_{30}x^2 + a_{21}xy + a_{12}y^2), \\ \dot{y} = y(b_{01} + b_{11}x + b_{02}y + b_{21}x^2 + b_{12}xy + b_{03}y^2) \end{cases} \tag{11}$$

(we preserved the old notations).

The property **2.7)** says that every cubic system with at least four real invariant straight lines can be written in the form (11).

For system (11) a straight line $y = Ax + B,\ A \neq 0$ is invariant if and only if A and B are the solutions of the system:

$$\begin{aligned} &B(b_{01} + b_{02}B + b_{03}B^2) = 0, \\ &b_{11}B + b_{12}B^2 + [b_{01} - a_{10} + (2b_{02} - a_{11})B + (3b_{03} - a_{12})B^2] \cdot A = 0, \\ &b_{21}B + [b_{11} - a_{20} + (2b_{12} - a_{21})B] \cdot A + [b_{02} - a_{11} + (3b_{03} - 2a_{12})B] \cdot A^2 = 0, \\ &b_{21} - a_{30} + (b_{12} - a_{21}) \cdot A + (b_{03} - a_{12}) \cdot A^2 = 0. \end{aligned} \tag{12}$$

Its cofactor is

$$K(x, y) = c_{00} + c_{10}x + c_{01}y + c_{20}x^2 + c_{11}xy + c_{02}y^2,$$

where

$$c_{00} = b_{01} + b_{02}B + b_{03}B^2,\ \ c_{01} = b_{02} + b_{03}B,$$

$$c_{10} = b_{11} + b_{12}B + (b_{02} - a_{11})A + (2b_{03} - a_{12})AB,$$

$$c_{20} = b_{21} + (b_{12} - a_{21})A + (b_{03} - a_{12})A^2,\ \ c_{11} = b_{12} + (b_{03} - a_{12})A,\ \ c_{02} = b_{03}.$$

The invariant straight line $Ax - y + B = 0,\ A \neq 0$, of (11) has the degree of invariance not less than two if and only if A and B verify the following seven relations:

$$\begin{aligned} &B(b_{02} + 2b_{03}B) = 0, \quad b_{01} + 2b_{02}B + 3b_{03}B^2 = 0, \\ &a_{10}A + 2b_{02}AB + (b_{12} + 6b_{03}A - a_{12}A)B^2 = 0, \\ &a_{20} + b_{02}A + 2(b_{12} + 3b_{03}A - a_{12}A)B = 0, \\ &b_{11} - a_{20} + (b_{02} - a_{11})A = 0,\ \ a_{30} + b_{12}A + (2b_{03} - a_{12})A^2 = 0, \\ &b_{21} + (2b_{12} - a_{21})A + (3b_{03} - 2a_{12})A^2 = 0. \end{aligned} \tag{13}$$

In this case, the cofactor of invariant straight line is $K(x, y) = c_{00} + c_{10}x + c_{01}y$, where

$$c_{00} = -b_{02} - 2b_{03}B, \quad c_{10} = -b_{12} + (a_{12} - 2b_{03})A, \quad c_{01} = -b_{03}.$$

Proposition 1. *Let the cubic system have two real not parallel invariant straight lines l_1 and l_2, of which l_1 has the degree of invariance equal to m, $1 \leq m \leq 3$. Then the number of singular points lying on $l_2 \setminus l_1$ is at most $3 - m$.*

Proof. In hypothesis of Proposition 1 via affine transformation, system (10) can be written in the form:

$$\dot{x} = x^m \tilde{P}_{3-m}(x, y), \qquad \dot{y} = y\tilde{Q}_2(x, y), \tag{14}$$

where $\tilde{P}_i$, $\tilde{Q}_i$ are polynomials of degree at most i and $\tilde{P}_{3-m}(x, 0) \not\equiv 0$, $\tilde{P}_{3-m}(0, y) \not\equiv 0$. The system (14) has the invariant straight lines $l_1 : x = 0$ and $l_2 : y = 0$ of which l_1 has the degree of invariance equal to m. The assertion of Proposition 1 follows from the fact that the equation $\tilde{P}_{3-m}(x, 0) = 0$ can not have more than $3 - m$ roots. $\square$

We say that the straight lines l_1, l_2 and l_3 are of generic position (”triangle” position) if $l_i \cap l_j \neq \varnothing$ and $l_1 \cap l_2 \cap l_3 = \varnothing$.

Proposition 2. *If cubic system (10) has three real invariant straight lines of generic position, then the sum of their degrees of invariance is at most four.*

Proof. We mention that any invariant straight line of the cubic system (10) can not have the degree of invariance more than three.

As the point of intersection of two invariant straight lines is a singular point for (10), Proposition 1 does not allow that any of these three straight lines l_1, l_2 and l_3 to have the degree of invariance equal to three.

Let each of the invariant straight lines l_1 and l_2 has the degree of invariance equal to two. By affine transformation and time rescaling the system (10) can be written in the form:

$$\begin{cases} \dot{x} = x^2(a + bx + y) \equiv P(x, y), \\ \dot{y} = y^2(c + dx + ey) \equiv Q(x, y), \ GCD(P, Q) = 1, \end{cases} \tag{15}$$

for which $l_1 = x$ and $l_2 = y$, and equalities (12) have the form

$$\begin{gathered} B^2(c + eB) = 0, \ 2cA + dB + 3eAB = 0, \\ a + (1 - 2d)B - cA - 3eAB = 0, \ eA^2 + (d - 1)A - b = 0. \end{gathered} \tag{16}$$

Let $l_3 = y - Ax - B$, $AB \neq 0$. The points $(0, B) = l_1 \cap l_3$ and $(-B/A, 0) = l_2 \cap l_3$ are singular points for (15). Therefore, $P(-B/A, 0) = Q(0, B) = 0$, yielding $A = -c/a$ and $B = -c/b$. Substituting these values of A and B in the first three equalities of (16), we get that $c = ab$, $d = b^2$ and $e = b$. In this case, $GCD(P, Q) = a + bx + y$. So, the assumption that system (15) can have invariant straight lines not passing through the origin of coordinates is false. $\square$

3 Canonical forms and Darboux integrability

There are the following possible configurations of six invariant straight lines along three directions:

$$1)\,(3,2,1);\quad 2)\,(3(2),2,1);\quad 3)\,(3(3),2,1);\quad 4)\,(3,2(2),1);$$

$$5)\,(3(2),2(2),1);\quad 6)\,(3(3),2(2),1);\quad 7)\,(2,2,2);\quad 8)\,(2(2),2,2);$$

$$9)\,(2(2),2(2),2);\quad 10)\,(2(2),2(2),2(2)).$$

Notation $(3,2,1)$ means that along one direction there are three distinct straight lines, along the second direction there are two distinct invariant straight lines and along the third direction there is one invariant straight line; $(3(2),2,1)$ means that along one direction the differential system has two distinct straight lines from which one is double (i.e. has the degree of invariance equal to two), along the second direction there are two distinct invariant straight lines and along the third direction there is one invariant straight line and so on.

3.1) Configuration $(3,2,1)$. We note that the point of intersection of two real invariant straight lines of the system (10) is a singular point for this system.

Assume that the cubic system (10) has six distinct invariant straight lines, including one couple Then, taking into account the property **2.3)** from Section 2, the given straight lines can have (up to some affine transformation) one of the following 2 geometric positions given in Fig. 3.1.

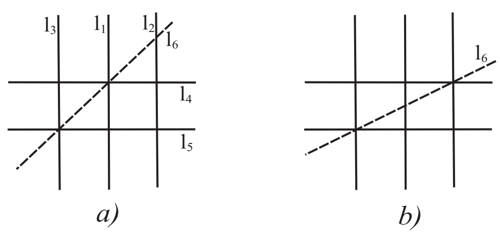


Fig. 3.1

The cubic system which includes both configurations, via affine transformation and time rescaling can be written in the form

$$\begin{cases} \dot{x} = x(x+1)(x-a), \ a>0, \\ \dot{y} = y(y+1)(c+dx+ey), \ d(|e|+|c(c-d)(c+ad)|) \neq 0. \end{cases} \tag{17}$$

The system (17) has the invariant straight lines

$$l_1 \equiv x=0, \ l_2 \equiv y=0, \ l_3 \equiv x+1=0, \ l_4 \equiv y+1=0, \ l_5 \equiv x-a=0.$$

We have to determine the conditions on parameters c, d and e such that (17) has only one invariant straight line of the form $l_6 \equiv y-Ax-B=0, \ A \neq 0$.

For (17) the equalities (12) look as:

$$\begin{array}{l} B(B+1)(eB+c)=0,\ dB+dB^2+[a+c+2(c+e)B+3eB^2]\cdot A=0, \\ A\cdot[a+d-1+(c+e)A+2dB+3eAB]=0,\ eA^2+dA-1=0. \end{array} \tag{18}$$

Otherwise, we observe that the fourth equation of (18) doesn't allow for cubic system of $\dot{x}=x(x+1)(x-a),\ \dot{y}=cy(y+1),\ a|c|>0$ the configuration $(3,2,1)$ to be realized.

In the cases $a)$ the straight line l_6 has the equation $y=x$. Putting in (18) $A=1$ and $B=0$, we obtain

$$c=-a,\ e=1-d. \tag{19}$$

In conditions (19) the equalities (18) show that the straight line $y=-x/a$ $(y=(x-a)/(a+1))$ is invariant for (17) if $a+d-1=0$ $(a-d+2=0)$.

Equalities (19) and inequality $(a+d-1)(a-d+2)\neq 0$ show that for (17) the case $a)$ is realized, excluding, at the same time, the cases when (17) can has more than 6 invariant straight lines. In these conditions, (17) can be written in the form (4).

In the cases $b)$ the straight line $l_6:\ y=(x-a)/(a+1)$ is invariant for (17) if

$$c=a(1+a-d),\ e=(a+1)(1+a-d). \tag{20}$$

If $a-d+2=0$ $(2a-d+1=0)$ then (17) has the invariant straight line $l_7=x-y$ $(l_7=x-ay-a)$.

The conditions (20) and $(a-d+2)(2a-d+1)\neq 0$ reduce (17) to the system (5).

The systems (4) and (5) are Darboux integrable and have respectively the integrating factors:

$$\mu(x,y)=x^{a/\delta}(x+1)^{-(a+1)/\delta}(x-a)^{-2}y^{(d-a-2)/\delta}(y+1)^{(d+a-1)/\delta}(y-x)^{d/\delta},$$

$$\mu(x,y)=x^{-2}(x+1)^{-\sigma}(x-a)^{-a\sigma}y^{-(1+\sigma)}(y+1)^{-(1+a\sigma)}\left(y-\frac{x-a}{a+1}\right)^{d\sigma},$$

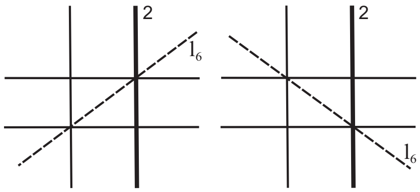
where $\delta=1-d,\ \sigma=1/(a-d+1)$.

3.2) Configuration $(3(2),2,1)$. The cubic system (10), with invariant straight lines of configuration $(3(2),2)$, via affine transformation and time rescaling, can be written in the form

$$\begin{cases} \dot{x}=x^2(x+1), \\ \dot{y}=y(y+1)(c+dx+ey),\ d(|e|+|c(c-d)|)\neq 0. \end{cases} \tag{21}$$

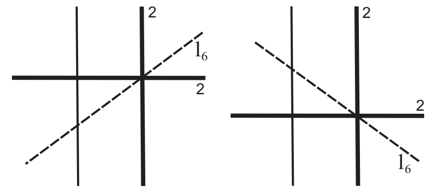
For this system the conditions (12) for the existence of invariant straight lines are of the form (18) with $a=0$.

For (21), the invariant straight line $x=0$ has the degree of invariance equal to two. Taking into account the propriety **2.3)** and Proposition 1, the system (21) can have invariant straight lines along three directions only of one of the following two geometric positions indicated in Fig. 3.2.



a) b)

Fig. 3.2



a) b)

Fig. 3.3

It is obvious that geometrical position of the straight lines in a) and b) are affine equivalent. We will examine only the case a). In order the straight line which passes through singular points $(-1,-1)$ and $(0,0)$, i.e. the straight line $y = x$, to be invariant for (21), it is necessary that $c = 0$ and $e = 1 - d$. In this conditions, (21) is reduced to the form (6). This system is Darboux integrable and has an integrating factor

$$\mu(x,y) = x^{-2}(x+1)^{-1/\delta}y^{-1-1/\delta}(y+1)^{-1}(y-x)^{d/\delta},$$

where $\delta = 1 - d$.

3.3) Configuration $(3(3),2,1)$ **and** $(3.2(2),1)$. The property **2.3)** and Proposition 1 do not allow the realization of these configurations.

3.4) Configuration $(3(2),2(2),1)$. Considering the configuration $(3(2),2(2))$ of invariant straight lines we obtain the system

$$\begin{cases} \dot{x} = x^2(x+1), \\ \dot{y} = y^2(c+dx+ey), \ d(|e|+|c(c-d)|) \neq 0, \end{cases} \tag{22}$$

which has the invariant straight lines $l_1 = x$, $l_2 = x+1$, $l_3 = y$ and the invariant exponential functions $l_4 = \exp(1/x)$, $l_5 = \exp(1/y)$. The straight lines l_1 and l_3 have the degree of invariance equal to two.

Proposition 2 allows only the positions from Fig.3.3 of the straight lines l_1, l_2, l_3 and $l_6 = y - Ax - B$, $A \neq 0$.

For (22) the equations (12) with condition $A \neq 0$ can be written as:

$$\begin{array}{l} B^2(c+eB) = 0, \ (dB + (2c+3eB)A)B = 0, \\ cA + (2d+3eA)B - 1 = 0, \ eA^2 + dA - 1 = 0. \end{array} \tag{23}$$

On the straight line $l_3 = x + 1$ the system (22) can have only the singular points $(-1,0)$ and $(-1,(d-c)/e)$. The straight line which passes through the points $(0,0)$ and $(-1,(d-c)/e)$ is described by the equation $y = (c-d)x/e$. Putting in (23) $A = (c-d)/e$ and $B = 0$, we obtain that $e = cd(c-d)$. This leads to the system

$$\dot{x} = x^2(x+1), \ \dot{y} = y^2(c+dx+c(c-d)y), \ c(c-d) \neq 0,$$

which by substitutions $d \to cd, x \to x, y \to y/c$ can be reduced to a system (7).

The system (7) is Darboux integrable and has an integrating factor

$$\mu(x,y) = x^{-1/\delta}\exp(\delta/x)(x+1)^{-2}y^{(2d-3)/\delta}\exp(-\delta/y)(y-x)^{d/\delta},$$

where $\delta = 1 - d$.

3.5) Configuration $(3(3), 2(2), 1)$. For first step we consider the system

$$\dot{x} = x^3,\ \dot{y} = y^2(c + dx + ey),\ d(|c| + |e|) \neq 0. \tag{24}$$

For (24) the equalities (12) look as:

$$\begin{array}{l} B^2(c + eB) = 0,\ (dB + (2c + 3eB)A)B = 0, \\ cA + (2d + 3eA)B = 0,\ eA^2 + dA - 1 = 0. \end{array} \tag{25}$$

Proposition 1 allows for differential system (24) to have besides the straight lines $l_{1,2,3} = x$, $l_{4,5} = y$ also the invariant straight lines of the form $y = Ax$, $A \neq 0$. Putting in (25) $B = 0$, we obtain that $c = 0$ and $A_{1,2} = (-d \pm \sqrt{d^2 + 4e})/(2e)$. If $d^2 + 4e > 0$ ($d^2 + 4e < 0$), the system (24) has seven (five) real straight lines, and if $d^2 + 4e = 0$, i.e. $e = -d^2/4$, after a transformation $y \to 2y/d$ we come to the system (8) with invariant straight line $l_6 = x - y$. This system has an integrating factor $\mu(x, y) = 1/(xy(x - y)^2)$.

3.6) Configuration $(2, 2, 2)$. Taking into account the propriety **2.3)**, the system (10) with such configuration has at least two singular points through which three invariant straight lines of different directions pass. By a translation one of these points can be brought at the origin. The system (10) realizing this configuration via an affine transformation and time rescaling can be brought to the form

$$\begin{cases} \dot{x} = x(x + 1)(a + bx + y) \equiv P(x, y), \\ \dot{y} = y(y + 1)(c + dx + ey) \equiv Q(x, y),\ GCD(P, Q) = 1. \end{cases} \tag{26}$$

For (26) the equalities (12) look as:

$$\begin{cases} B(B + 1)(c + eB) = 0, \\ (c - a)A + dB + dB^2 + (2c + 2e - 1 + 3eB)AB = 0, \\ d - a - b + (c + e - 1)A + (2d - 1)B + 3eAB = 0, \\ eA^2 + (d - 1)A - b = 0. \end{cases} \tag{27}$$

Besides the invariant straight lines $l_1 = x$, $l_2 = x + 1$, $l_3 = y$, $l_4 = y + 1$, we will seek the conditions on parameters of (27) such that it has exactly two more invariant straight lines of the form $y = Ax$, $y = Ax + B$, $AB \neq 0$. For this, we put $B = 0$ in (27). The second equation of (27) gives $c = a$, and the third one becomes

$$d - a - b + (a + e - 1)A = 0. \tag{28}$$

In assumption that $AB \neq 0$ and $c = a$, the system of equations ((27), (28)) has the following solutions:

1) $b = -a$, $c = a$, $d = 2 - 3a$, $e = 2a - 1$, $A = 1$, $B = -1$.

System (26) with the conditions above has the invariant straight lines $l_5 = y - x$, $l_6 = y - x + 1$. The condition $GCD(P, Q) = 1$ implies the inequality $a(2a - 1)(a - 1) \neq 0$, and the inequality $2 - 3a \neq 0$ excludes the existence of a triplet of invariant straight lines parallel to axis Ox. If $3a-1=0$, then the given system has two more invariant straight lines of the form: $l_7 = y+x+1$ and $l_8 = y-x-1$.

2) $b=(a-1)/2,\ c=a,\ d=(3a+1)/2,\ e=-a,\ A=B=1$
$(l_5 = y-x,\ l_6 = y-x-1,\ a(9a^2-1)(a^2-1)\neq 0)$;

3) $b=1-a,\ c=a,\ d=3a-1,\ e=2a-1,\ A=B=-1$
$(l_5=y+x,\ l_6=y+x+1,\ a(a-1)(2a-1)(3a-1)(3a-2)\neq 0)$;

4) $b=2a-1,\ c=a,\ d=3a-1,\ e=1-a,\ A=(1-2a)/(1-a),\ B=a/(a-1)$
$(l_5=y+(1-2a)x/(a-1),\ l_6=y+((1-2a)x-a)/(a-1),\ l_7=y-x)$.

If conditions **4)** hold, then (26) has seven invariant straight lines and, will be not considered. Moreover, it is sufficient to consider only the case **1)**, as the case **2)** (**3)**) can be reduced to the case **1)** via the change

$$a\to\frac{a}{2-3a},\ x\to y,\ y\to x,\ t\to(2-3a)t$$
$$(a\to 1-a,\ x\to x,\ y\to -y-1,\ t\to -t).$$

Inclusion of system (9) in the statement of Theorem in Section I is motivated. This system has the integrating factor

$$\mu(x,y)=\left[y(x+1)(y-x+1)\sqrt{x(y+1)(y-x)}\ \right]^{-1}.$$

3.7) Configuration $(2(2),2,2)$. Let cubic system (10) have distinct invariant straight lines l_j, $j=\overline{1,5}$, of which $l_1||l_2$, $l_3||l_4$ and l_5 has the degree of invariance equal to two. According to Proposition 1, the straight line l_5 must go through the points of intersection of straight lines l_1 and l_3, l_2 and l_4 (or l_1 and l_4, l_2 and l_3. This case is reduced to the previous one by changing the enumeration of straight lines). In our assumptions, via affine transformation and time rescaling the system (10) can be written in the form of (26). For (26) the straight lines $l_1=x$, $l_2=x+1$, $l_3=y$ and $l_4=y+1$ are invariant, and the equalities (13) look as:

$$\begin{aligned}
&B(c+e+2eB)=0,\ \ c+2(c+e)B+3eB^2=0,\\
&aA+2(c+e)AB+6eAB^2=0,\\
&a+b+(c+e)A+2dB+6eAB=0,\\
&d-a-b+(c+e-1)A=0,\\
&b+dA+2eA^2=0,\ \ A(2d-1+3eA)=0.
\end{aligned}\tag{29}$$

The straight line l_5 is given by the formula $x-y=0$. This line is invariant for (26) if $A=1$ together with $B=0$ are the solution of (29). Substituting in (29) these values of A and B, we obtain that $a=c=b+1=d+1=e-1=0$, which implies $GCD(P,Q)=y-x$.

3.8) Configuration $(2(2),2(2),2)$. Proposition 2 does not allow the realization of this configuration.

3.9) Configuration $(2(2),2(2),2(2))$. Taking into account Proposition 2, the invariant straight lines of this configuration should have a common point.

We consider the cubic system (15), where the straight lines $l_1 = x$ and $l_2 = y$ are invariant and have the degree of invariance equal to two. In this case the equalities (13) look as:

$$\begin{gathered} B(c+2eB)=0,\ B(2c+3eB)=0,\ B(2cA+dB+6eAB)=0,\\ a+cA+2dB+6eAB=0,\ a-cA=0,\\ b+dA+2eA^2=0,\ A(2d-1+3eA)=0. \end{gathered} \tag{30}$$

To determine the third invariant straight line $l_3 = Ax - y$, $A \neq 0$, with the same degree of invariance, we put in the equalities (30) $B = 0$ and resolve them for $A \neq 0$. The fourth and fifth equalities of ((30), $B = 0$) give $a = c = 0$. The condition $GCD(P, Q) = 1$ implies $e \neq 0$. From six and seven equalities of (30) we obtain $e = (2-d)(2d-1)/(9b)$ and $A = 3b/(d-2)$. Thus, we come to the system

$$\begin{cases} \dot{x} = x^2(bx+y),\ d(d+1)(2d-1)(d-2) \neq 0,\\ \dot{y} = y^2(dx+(2-d)(2d-1)y/(9b)), \end{cases}$$

which besides the invariant straight lines $x = 0$, $y = 0$, $3bx + (2-d)y = 0$ with the degree of invariance equal to two, also has the invariant straight line $3bx+(1-2d)y = 0$.

4 The phase portraits

We mention that the cubic system with at least four real invariant straight lines has no limit cycles [10]. Hence, the behaviour of trajectories in this system and, in particular, of system with six real invariant straight lines, is imposed by the type of singular points.

We denote by SP singular points; λ_1 and λ_2 the eigenvalues of SP; S − saddle ($\lambda_1\lambda_2 < 0$); N^s − stable node (λ_1, $\lambda_2 < 0$), N^u − unstable node (λ_1, $\lambda_2 > 0$); $S - N^{s(u)}$ − saddle-node with stable (unstable) parabolic sector; $P^{s(u)}$ − stable (unstable) parabolic sector; H − hyperbolic sector.

4.1. System (4). The coordinates of singular points of system (4) in the finite and infinite parts of the phase plane Oxy, also the eigenvalues λ_1, λ_2 of the characteristic equation, corresponding to each of these points, are shown in Tab.4.1. In this table the following notations: $\alpha = 1 + a$, $\delta = 1 - d$ were used.

Tab. 4.1

SP	$O_1(0,0)$	$O_2(-1,-1)$	$O_3(a,0)$	$O_4(0,-1)$	$O_5(0,a/\delta)$
$\lambda_1;\lambda_2$	$-a$; $-a$	α; α	$a\alpha$; $-a\delta$	$-a$; $a+\delta$	$-a$; $a(a+\delta)/\delta$
SP	$O_6(-1,0)$	$O_7(-1,(a+d)/\delta)$	$O_8(a,-1)$	$O_9(a,a)$	$I_1(1,0,0)$
$\lambda_1;\lambda_2$	α; $-a-d$	α; $\alpha(a+d)/\delta$	$a\alpha$; $\alpha\delta$	$a\alpha$; $a\alpha\delta$	-1; -1
SP	$I_2(0,1,0)$	$I_3(1,1,0)$	$I_4(1,-1/\delta,0)$		
$\lambda_1;\lambda_2$	$-\delta$; $-\delta$	-1; $2-d$	-1; $1+1/\delta$		

The singular point I_1 is a stable node. Taking into account that $a > 0$, at the point O_1 (O_2) the system (4) has a stable (unstable) node. Whatever are the parameters a, $a > 0$ and d, the types of points O_8 and O_9 coincide. In the case $a + \delta = 0$, i.e. $1 + a - d = 0$, $(a + d = 0;\ d = 2)$ the singular points O_4 and O_5 (respectively O_6 and O_7; I_3 and I_4) coincide.

By means of the straight lines $d = 0$, $d = 1$, $d = 2$, $a = 0$, $2+a-d = 0$, $1+a-d = 0$, $a+d-1 = 0$, $a+d = 0$ we divide the half-plane $a > 0$ of parameters space a and d in sectors (Fig. 4.1). In Fig. 4.1 by V we denote the semi-line $1+a-d = 0$, $d > 2$); by VI – the segment of straight line $(1 + a - d = 0,\ 1 < d < 2)$; by VII – the semi-line $(d = 2,\ a > 1)$; by $VIII$ – the segment $(d = 2,\ 0 < a < 1)$; by IX – the point $(2, 1)$; by XII – the semi-line $(a + d = 0,\ d < 0)$; by I – the open domain bounded by straight lines $a = 0$, $d = 2$, $1 + a - d = 0$ without the semi-line $(a - d + 2 = 0,\ 2 < d < +\infty)$ and so on.

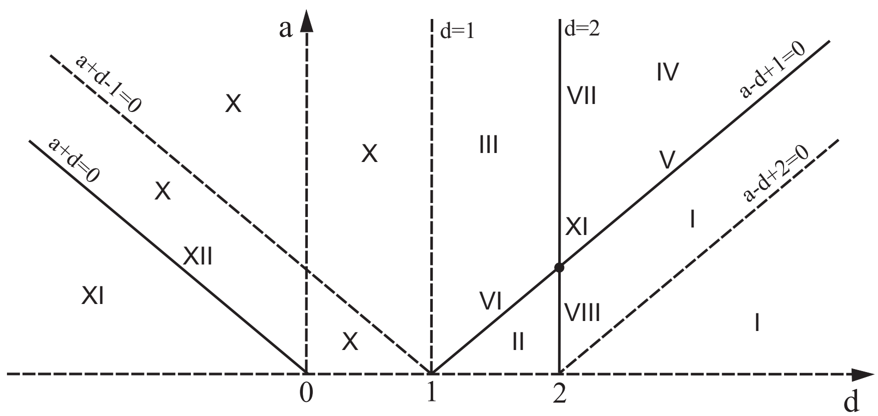


Fig. 4.1

For system (4) the results of qualitative investigation of singular points $O_3 - O_8$, $I_2 - I_4$ in each of the sectors $I - XII$ are given in Tab. 4.2.

Tab. 4.2

SP	I/II	III/IV	V/VI	$VII/VIII$	IX	X/XI	XII
O_3	N^u	N^u	N^u	N^u	N^u	S	S
O_4	N^s	S	$S{-}N^s$	S/N^s	$S{-}N^s$	S	S
O_5	S	N^s	–	N^s/S	–	S	S
O_6	S	S	S	S	S	S/N^u	$S{-}N^u$
O_7	S	S	S	S	S	N^u/S	–
O_8	S	S	S	S	S	N^u	N^u
I_2	N^u	N^u	N^u	N^u	N^u	N^s	N^s
I_3	N^s/S	S/N^s	N^s/S	$S{-}N^s$	$S{-}N^s$	S	S
I_4	S/N^s	N^s/S	S/N^s	–	–	S	S
$Fig. 1:$	1)/2)	3)/4)	5)/6)	7)/8)	9)	10)/11)	12)

4.2. System (5). For (5) the singular points and and the eigenvalues of the characteristic equation are shown in Tab. 4.3, where $\alpha = 1 + a$.

Tab. 4.3

SP	$O_1(-1,-1)$	$O_2(a,0)$	$O_3(0,-1)$	$O_4(0,0)$
$\lambda_1;\ \lambda_2$	$\alpha;\ \alpha$	$a\alpha;\ a\alpha$	$-a;\ \alpha-d$	$-a;\ a(\alpha-d)$
SP	$O_5(0,-a/\alpha)$	$O_6(-1,0)$	$O_7(-1,\frac{d-a}{\alpha+d})$	$O_8(a,-1)$
$\lambda_1;$ λ_2	$-a;$ $a(d-\alpha)/\alpha$	$\alpha;$ $\alpha(a-d)$	$\alpha;$ $\alpha(d-a)/(\alpha-d)$	$a\alpha;$ $\alpha(1-d)$
SP	$O_9(a,a/(d-\alpha))$	$I_1(1,0,0)$	$I_2(0,1,0)$	$I_3(1,1/\alpha,0)$
$\lambda_1;$ λ_2	$a\alpha;$ $a\alpha(d-1)/(\alpha-d))$	$-1;$ -1	$\alpha(d-\alpha);$ $\alpha(d-\alpha)$	$-1;$ $2-d/\alpha$
SP	$I_4(1,1/(d-\alpha),0)$			
$\lambda_1;\ \lambda_2$	$-1;\ 1+\alpha/(d-\alpha)$			

For the system (5) the singular points O_1 and O_2 are unstable nodes, but point I_1 is a stable node. At every point of the half-plane $a>0$ the points O_3 and O_4 are of the same type. If $a-d=0$ ($d=1$; $2a-d+2=0$), then the points O_6 and O_7 (respectively: O_8 and O_9; I_3 and I_4) coincide.

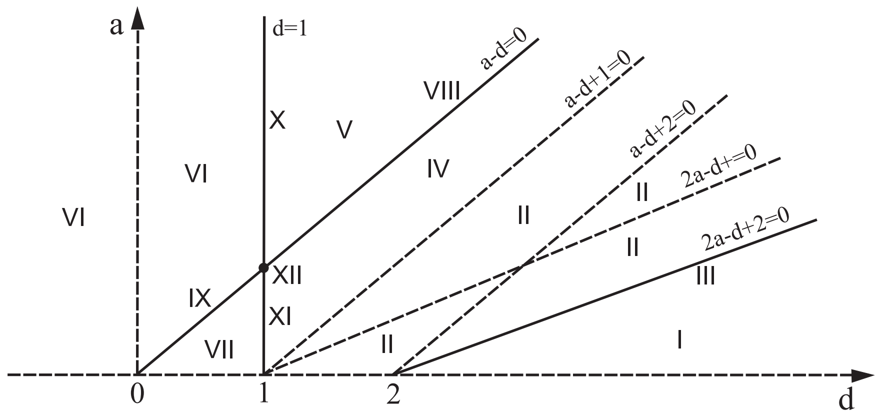


Fig. 4.2

The partition of the half-pane $a>0$ in sectors and the qualitative study of singular points O_4-O_9, I_2-I_4 are given in Fig. 4.2 and Tab. 4.4 respectively.

Tab. 4.4

SP	I/II	III	IV/V	VI/VII	$VIII/IX$	X/XI	XII
O_4	N^s	N^s	S	S	S	S	S
O_5	S	S	N^s	N^s	N^s	N^s	N^s
O_6	S	S	S/N^u	N^u/S	$S{-}N^u$	N^u/S	$S{-}N^u$
O_7	S	S	N^u/S	S	$-$	S/N^u	$-$
O_8	S	S	S	N^u	N^u/S	$S{-}N^u$	$S{-}N^u$
O_9	S	S	N^u	S/N^u	S/N^u	$-$	$-$
I_2	N^u	N^u	N^s	N^s	N^s	N^s	N^s
I_3	N^s/S	$S{-}N^s$	S	S	S	S	S
I_4	S/N^u	$-$	S	S	S	S	S
$Fig.\,2:$	1)/2)	3)	4)/5)	6)/5)	7)/8)	8)/7)	9)

4.3. System (6). This system has five singular points in the finite part of the phase plane: $O_1(0,0)$, $O_2(-1,0)$, $O_3(-1,-1)$, $O_4(0,-1)$, $O_5(-1,\ d/(1-d))$; and four singular points at the infinity: $I_1(1,0,0)$, $I_2(0,1,0)$, $I_3(1,1,0)$, $I_4(1,1/(d-1),0)$. Among these singular points only $O_1(0,0)$ has the both eigenvalues of the characteristic equation equal to zero (see Tab. 4.5). To determine the behavior of trajectories in the neighborhood of this point, we write the system (6) in polar coordinates $x=\rho\cos\theta$, $y=\rho\sin\theta$:

$$\begin{cases} \dfrac{d\rho}{d\tau} = \rho[\cos^3\theta(1+\rho\cos\theta)+\sin^2\theta(1+\rho\sin\theta)(d\cos\theta+\delta\sin\theta)], \\ \dfrac{d\theta}{d\tau} = \sin\theta\cos\theta(\sin\theta-\cos\theta)(\rho\cos\theta+\delta(1+\rho\sin\theta)), \end{cases} \tag{31}$$

where $\tau=\rho t$, $\delta=1-d$. The singular points of system (31) with the first coordinate $\rho=0$ and the second $\theta\in[0,2\pi)$, and their eigenvalues are $\{M_1(0,0), M_2(0,\pi):$ $\lambda_1=1, \lambda_2=d-1\}$; $\{M_3(0,\pi/2),\ M_4(0,3\pi/2):\ \lambda_1\cdot\lambda_2=-(1-d)^2\}$; $\{M_5(0,\pi/4):$ $\lambda_1=1/\sqrt{2}, \lambda_2=(1-d)/\sqrt{2}\}$; $\{M_6(0,5\pi/4):\ \lambda_1=-1/\sqrt{2}, \lambda_2=-(1-d)/\sqrt{2}\}$. The types of these points can differ only if d passes through value 1. If $d<1$, we have Fig. 4.3, and if $d>1$, we have Fig. 4.4.

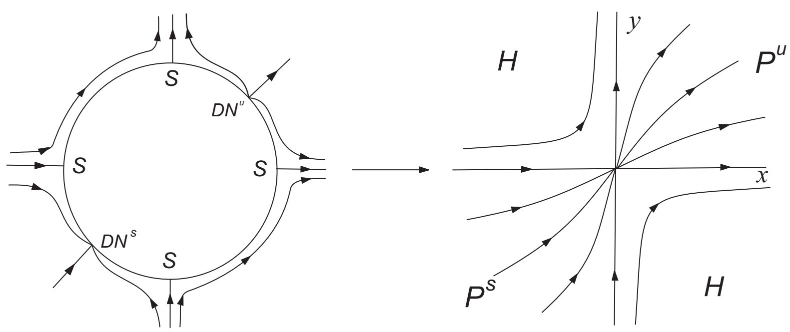


Fig. 4.3 ($d<1$).

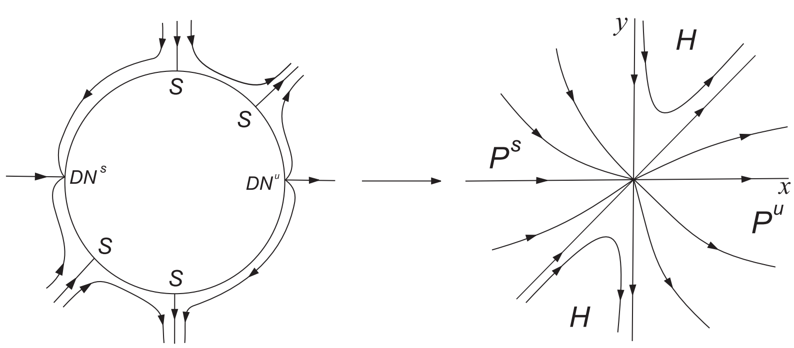


Fig. 4.4 ($d>1$).

In the case of system (6) we have Tab. 4.5.

Tab. 4.5

SP	$\lambda_1;\lambda_2$	$d<0$	$0<d<1$	$1<d<2$	$d=2$	$d>2$
O_1	$0;\,0$	P^uHP^sH	P^uHP^sH	P^uHP^sH	P^uHP^sH	P^uHP^sH
O_2	$1;\,-d$	N^u	S	S	S	S
O_3	$1;\,1$	N^u	N^u	N^u	N^u	N^u
O_4	$0;\,1-d$	$S-N^u$	$S-N^u$	$S-N^s$	$S-N^s$	$S-N^s$
O_5	$1;\,\frac{d}{1-d}$	S	N^u	S	S	S
I_1	$-1;\,-1$	N^s	N^s	N^s	N^s	N^s
I_2	$d-1;\,d-1$	N^s	N^s	N^u	N^u	N^u
I_3	$-1;\,2-d$	S	S	S	$S-N^s$	N^s
I_4	$-1;\,\frac{2-d}{1-d}$	S	S	N^s	$-$	S
$Fig. 3:$		1)	2)	3)	4)	5)

4.4. System (7). This system has the singular points: $O_1(0,0)$, $O_2(-1,0)$, $O_3(-1,-1)$, $O_4(0,\frac{1}{d-1})$, $I_1(1,0,0)$, $I_2(0,1,0)$, $I_3(1,1,0)$, $I_4(1,\frac{1}{d-1},0)$, whose characterizations are given in Tab. 4.6.

Tab. 4.6

SP	$\lambda_1;\lambda_2$	$d<1,\,d\neq 0$	$1<d<2$	$d=2$	$d>2$
O_1	$0;\,0$	P^uHP^sH	P^uHP^sH	P^uHP^sH	P^uHP^sH
O_2	$0;\,1$	$S-N^u$	$S-N^u$	$S-N^u$	$S-N^u$
O_3	$1;\,1-d$	N^u	S	S	S
O_4	$0;\,1/(1-d)$	$S-N^u$	$S-N^s$	$S-N^s$	$S-N^s$
I_1	$-1;\,-1$	N^s	N^s	N^s	N^s
I_2	$d-1;\,d-1$	N^s	N^u	N^u	N^u
I_3	$-1;\,2-d$	S	S	$S-N^s$	N^s
I_4	$-1;\,\frac{2-d}{1-d}$	S	N^s	$-$	S
$Fig. 4:$		1)	2)	3)	4)

As in the case of system (6), the behavior of the trajectories in the neighborhood of singular point $O_1(0,0)$ was established by using the blow-up method for (7) in polar coordinates:

$$\begin{cases} \dfrac{d\rho}{d\tau} = \rho[\cos^3\theta(1+\rho\cos\theta)+\sin^3\theta(1+d\rho\cos\theta+(1-d)\rho\sin\theta)], \\ \dfrac{d\theta}{d\tau} = \sin\theta\cos\theta(\sin\theta-\cos\theta)(1+\rho\cos\theta+(1-d)\rho\sin\theta)), \end{cases} \tag{32}$$

where $\tau=\rho t$. The singular points of (32) with $\rho=0$ and $\theta\in[0,2\pi)$ and their eigenvalues: $\{M_1(0,0),M_2(0,\pi),M_3(0,\pi/2),M_4(0,3\pi/2)\ :\ \lambda_1=-1,\lambda_2=1\}$; $\{M_5(0,\pi/4)\ :\ \lambda_1=\lambda_2=1/\sqrt{2}\}$; $\{M_6(0,5\pi/4)\ :\ \lambda_1=\lambda_2=-1/\sqrt{2}\}$, lead us to Fig. 4.3.

4.5. System (8). This system has in finite parts of the phase plane a singular point $O(0,0)$ with $\lambda_1=\lambda_2=0$ and at infinity singular points $I_1(1,0,0)$; $I_2(0,1,0)$;

$I_3(1,1,0)$ with $\lambda_1 = \lambda_2 = -1$; $\lambda_1 = \lambda_2 = 1$; $\lambda_1 = -1, \lambda_2 = 0$. We have that I_1 is N^s; $I_2 - N^u$; $I_3 - S$–N^s and $O - P^u HHP^u HH$ (see Fig. 5).

4.6. System (9). For (9) the singular points and the eigenvalues of the characteristic equation are shown in Tab. 4.7. In this table we used the notations: $\beta = a - 1$, $\gamma = 2a - 1$.

Tab. 4.7

SP	$O_1(0,0)$	$O_2(-1,0)$	$O_3(-1,-1)$	$O_4(1,0)$
$\lambda_1;\lambda_2$	$a;a$	$-2a;2\gamma$	$-\gamma;-\gamma$	$-2\beta;-2a$
SP	$O_5(-1,-2)$	$O_6(0,-1)$	$O_7(\beta/a,-1)$	$O_8(a/\beta,a/\beta)$
$\lambda_1;\lambda_2$	$-2\beta;2\gamma$	$\beta;\beta$	$-\beta\gamma/a;2\beta\gamma/a$	$-a\gamma/\beta;2a\gamma/\beta$
SP	$O_9(0,a/\gamma)$	$I_1(1,0,0)$	$I_2(0,1,0)$	$I_3(1,1/\beta,0)$
$\lambda_1;\lambda_2$	$-a\beta/\gamma;2a\beta/\gamma$	$a;a$	$\gamma;\gamma$	$\beta;\beta$
SP	$I_4(1,a/\gamma,0)$			
$\lambda_1;\lambda_2$	$-a\beta/\gamma;\ 2a\beta/\gamma$			

We divide the real axis in intervals $J_1 = (-\infty, 0)$, $J_2 = (0, 1/3)$, $J_3 = (1/3, 1/2)$, $J_4 = (1/2, 2/3)$, $J_5 = (2/3, 1)$, $J_6 = (1, +\infty)$; $J = J_1 \cup J_2 \cup \cdots \cup J_6$.

The eigenvalues λ_1 and λ_2 of the characteristic equation corresponding to each singular point, in intervals J_1 and J_6 differ only by sign. Therefore, from the qualitative point of view the phase portraits of system (9) in intervals J_1 and J_6, differ only by directions on trajectories.

Singular points O_7, O_8, O_9 and I_4 are saddles for every $a \in J$. The types of other singular points (i.e. $O_1 - O_6, I_1, I_2, I_3$) are shown in Tab. 4.8.

Tab. 4.8

SP	$J_1\ (J_6)$	$J_2,\ J_3$	$J_4,\ J_5$
O_1	$N^{s(u)}$	N^u	N^u
O_2	S	N^s	S
O_3	$N^{u(s)}$	N^u	N^s
O_4	$N^{u(s)}$	S	S
O_5	S	S	N^u
O_6	$N^{s(u)}$	N^s	N^s
I_1	$N^{s(u)}$	N^u	N^u
I_2	$N^{u(s)}$	N^u	N^s
I_3	$N^{s(u)}$	N^s	N^s
$Fig. 6:$	$1 \rightleftarrows)$	2)	3)

References

[1] Artes J., Llibre J. *On the number of slopes of invariant straight lines for polynomial differential systems*. J. of Nanjing University, 1996, **13**, 143–149.

[2] Artes J., Grünbaum B., Llibre J. *On the number of invariant straight lines for polynomial differential systems*. Pacific Journal of Mathematics, 1998, **184**, No. 2, 207–230.

[3] Cozma D., Şubă A. *The solution of the problem of center for cubic differential systems with four invariant straight lines*. Mathematical analysis and aplications (Iaşi, 1997). An. Ştiinţ. Univ. "Al. I. Cuza"(Iaşi), 1998, **44**, suppl., 517–530.

[4] Kooij R.E. *Cubic systems with four real line invariants*. Math. Proc. Camb. Phil. Soc., 1995, **118**, No. 1, 7–19.

[5] Llibre J., Vulpe N. *Planar cubic polynomial differential systems with the maximum number of invariant straight lines*. Rocky Mountain J. Math., 2006, **36**, No. 4, 1301–1373.

[6] Lyubimova R.A. *About one differential equation with invariant straight lines*. Differential and integral equations, Gorky Universitet, 1997, **1**, 19–22 (in Russian).

[7] Lyubimova R.A. *About one differential equation with invariant straight lines*. Differential and integral equations (Gorky Universitet), 1984, **8**, 66–69 (in Russian).

[8] Puţuntică V., Şubă A. *The cubic differential system with six real invariant straight lines along two directions*. Studia Universitatis. Seria Ştiinţe Exacte şi Economice, 2008, No. 8(13), 5–16.

[9] Rychkov G.S. *The limit cycles of the equation* $u(x+1)du = (-x + ax^2 + bxu + cu + du^2)dx$. Differentsial'nye Uravneniya, 1972, **8**, No. 12, 2257–2259 (in Russian).

[10] Suo Guangjian, Sun Jifang. *The* n*-degree differential system with* $(n-1)(n+1)/2$ *straight line solutions has no limit cycles*. Proc. of Ordinary Differential Equations and Control Theory, Wuhan, 1987, 216–220 (in Chinese).

[11] Şubă A., Cozma D. *Solution of the problem of the center for cubic differential system with three invariant straight lines in generic position*. Qualitative Theory of Dynamical Systems (Universitat de Lleida. Spaine), 2005, **6**, 45–58.

[12] Zhang Xiang, Ye Yanqian. *On the number of invariant lines for polynomial systems*. Proc. of the American Math. Soc., 1998, **126**, No. 8, 2249–2265.

V. Puţuntică
Department of Mathematics
Tiraspol State University
5 Gh. Iablocichin str., Chişinău, MD–2069
Moldova

E-mail: *vitputuntica@mail.ru*

Received February 5, 2009

A. Şubă
Department of Mathematics
State University of Moldova
60 A. Mateevici str., Chişinău, MD–2009
Moldova

E-mail: *suba@usm.md*

BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 2(60), 2009, Pages 131–134
ISSN 1024–7696

A note on a subclass of analytic functions defined by a differential operator

Alina Alb Lupaş, Adriana Cătaş

Abstract. By means of the Sălăgean differential operator we define a new class $\mathcal{BS}(m,\mu,\alpha)$ involving functions $f \in \mathcal{A}_n$. Parallel results for some related classes including the class of starlike and convex functions respectively are also obtained.

Mathematics subject classification: 30C45.
Keywords and phrases: Analytic function, starlike function, convex function, Sălăgean differential operator.

1 Introduction and definitions

Let $\mathcal{A}_n$ denote the class of functions of the form

$$f(z) = z + \sum_{j=n+1}^{\infty} a_j z^j \tag{1}$$

which are analytic in the open unit disc $U = \{z : |z| < 1\}$ and $\mathcal{H}(U)$ be the space of holomorphic functions in U, $n \in \mathbb{N} = \{1, 2, ...\}$.

Let $\mathcal{S}$ denote the subclass of functions that are univalent in U.

By $\mathcal{S}^*(\alpha)$ we denote a subclass of $\mathcal{A}_n$ consisting of starlike univalent functions of order α, $0 \leq \alpha < 1$, which satisfy

$$\mathrm{Re}\left(\frac{zf'(z)}{f(z)}\right) > \alpha, \quad z \in U. \tag{2}$$

Further, a function f belonging to $\mathcal{S}$ is said to be convex of order α in U, if and only if

$$\mathrm{Re}\left(\frac{zf''(z)}{f'(z)} + 1\right) > \alpha, \quad z \in U \tag{3}$$

for some α $(0 \leq \alpha < 1)$. We denote by $\mathcal{K}(\alpha)$ the class of functions in $\mathcal{S}$ which are convex of order α in U and denote by $\mathcal{R}(\alpha)$ the class of functions in $\mathcal{A}_n$ which satisfy

$$\mathrm{Re} f'(z) > \alpha, \quad z \in U. \tag{4}$$

It is well known that $\mathcal{K}(\alpha) \subset \mathcal{S}^*(\alpha) \subset \mathcal{S}$.

If f and g are analytic functions in U, we say that f is subordinate to g, written $f \prec g$, if there is a function w analytic in U, with $w(0) = 0$, $|w(z)| < 1$, for all $z \in U$ such that $f(z) = g(w(z))$ for all $z \in U$. If g is univalent, then $f \prec g$ if and only if $f(0) = g(0)$ and $f(U) \subseteq g(U)$.

Let D^m be the Sălăgean differential operator [3], $D^m : \mathcal{A}_n \to \mathcal{A}_n$, $n \in \mathbb{N}$, $m \in \mathbb{N} \cup \{0\}$, defined as

$$\begin{aligned} D^0 f(z) &= f(z), \\ D^1 f(z) &= Df(z) = zf'(z), \\ D^m f(z) &= D(D^{m-1} f(z)), \quad z \in U. \end{aligned}$$

We note that if $f \in \mathcal{A}_n$, then

$$D^m f(z) = z + \sum_{j=n+1}^{\infty} j^m a_j z^j, \quad z \in U.$$

To prove our main theorem we shall need the following lemma.

Lemma 1 (see [2])**.** *Let p be analytic in U with $p(0) = 1$ and suppose that*

$$Re\left(1 + \frac{zp'(z)}{p(z)}\right) > \frac{3\alpha - 1}{2\alpha}, \quad z \in U. \tag{5}$$

Then $Rep(z) > \alpha$ for $z \in U$ and $1/2 \leq \alpha < 1$.

2 Main results

Definition 1. We say that a function $f \in \mathcal{A}_n$ is in the class $\mathcal{BS}(m, \mu, \alpha)$, $n \in \mathbb{N}$, $m \in \mathbb{N} \cup \{0\}$, $\mu \geq 0$, $\alpha \in [0, 1)$ if

$$\left| \frac{D^{m+1} f(z)}{z} \left(\frac{z}{D^m f(z)} \right)^{\mu} - 1 \right| < 1 - \alpha, \qquad z \in U. \tag{6}$$

Remark 1. The family $\mathcal{BS}(m, \mu, \alpha)$ is a new comprehensive class of analytic functions which includes various new classes of analytic univalent functions as well as some very well-known ones. For example, $\mathcal{BS}(0, 1, \alpha) \equiv \mathcal{S}^*(\alpha)$, $\mathcal{BS}(1, 1, \alpha) \equiv \mathcal{K}(\alpha)$ and $\mathcal{BS}(0, 0, \alpha) \equiv \mathcal{R}(\alpha)$. Another interesting subclass is the special case $\mathcal{BS}(0, 2, \alpha) \equiv \mathcal{B}(\alpha)$ which has been introduced by Frasin and Darus [1] and also the class $\mathcal{BS}(0, \mu, \alpha) \equiv \mathcal{B}(\mu, \alpha)$ which has been introduced by Frasin and Jahangiri [2].

In this note we provide a sufficient condition for functions to be in the class $\mathcal{BS}(m, \mu, \alpha)$. Consequently, as a special case, we show that convex functions of order 1/2 are also members of the above defined family.

Theorem 1. *For the function $f \in \mathcal{A}_n$, $n \in \mathbb{N}$, $m \in \mathbb{N} \cup \{0\}$, $\mu \geq 0$, $1/2 \leq \alpha < 1$ if*

$$\frac{D^{m+2} f(z)}{D^{m+1} f(z)} - \mu \frac{D^{m+1} f(z)}{D^m f(z)} + \mu \prec 1 + \beta z, \quad z \in U, \tag{7}$$

where

$$\beta = \frac{3\alpha - 1}{2\alpha},$$

then $f \in \mathcal{BS}(m, \mu, \alpha)$.

Proof. If we consider

$$p(z) = \frac{D^{m+1} f(z)}{z} \left(\frac{z}{D^m f(z)} \right)^{\mu} \tag{8}$$

then $p(z)$ is analytic in U with $p(0) = 1$. A simple differentiation yields

$$\frac{zp'(z)}{p(z)} = \frac{D^{m+2} f(z)}{D^{m+1} f(z)} - \mu \frac{D^{m+1} f(z)}{D^m f(z)} + \mu - 1. \tag{9}$$

Using (7) we get

$$\mathrm{Re}\left(1 + \frac{zp'(z)}{p(z)}\right) > \frac{3\alpha - 1}{2\alpha}.$$

Thus, from Lemma 1 we deduce that

$$\mathrm{Re}\left\{ \frac{D^{m+1} f(z)}{z} \left(\frac{z}{D^m f(z)} \right)^{\mu} \right\} > \alpha.$$

Therefore, $f \in \mathcal{BS}(m, \mu, \alpha)$, by Definition 1. □

As a consequence of the above theorem we have the following interesting corollaries.

Corollary 1. *If* $f \in \mathcal{A}_n$ *and*

$$Re\left\{ \frac{2zf''(z) + z^2 f'''(z)}{f'(z) + zf''(z)} - \frac{zf''(z)}{f'(z)} \right\} > -\frac{1}{2}, \quad z \in U, \tag{10}$$

then

$$Re\left\{ 1 + \frac{zf''(z)}{f'(z)} \right\} > \frac{1}{2}, \quad z \in U. \tag{11}$$

That is, f *is convex of order* $\frac{1}{2}$.

Corollary 2. *If* $f \in \mathcal{A}_n$ *and*

$$Re\left\{ \frac{2z^2 f''(z) + z^3 f'''(z)}{zf'(z) + z^2 f''(z)} \right\} > -\frac{1}{2}, \quad z \in U, \tag{12}$$

then

$$Re\left\{ f'(z) + zf''(z) \right\} > \frac{1}{2}, \quad z \in U. \tag{13}$$

Corollary 3. *If $f \in \mathcal{A}_n$ and*

$$Re\left\{1+\frac{zf''(z)}{f'(z)}\right\} > \frac{1}{2}, \quad z \in U, \tag{14}$$

then

$$Ref'(z) > \frac{1}{2}, \quad z \in U. \tag{15}$$

In another words, if the function f is convex of order $\frac{1}{2}$ then $f \in \mathcal{BS}(0,0,\frac{1}{2}) \equiv \mathcal{R}\left(\frac{1}{2}\right)$.

Corollary 4. *If $f \in \mathcal{A}_n$ and*

$$Re\left\{\frac{zf''(z)}{f'(z)} - \frac{zf'(z)}{f(z)}\right\} > -\frac{3}{2}, \quad z \in U, \tag{16}$$

then f is starlike of order $\frac{1}{2}$.

References

[1] Frasin B.A., Darus M. *On certain analytic univalent functions*, Internat. J. Math. and Math. Sci., 2001, **25(5)**, 305–310.

[2] Frasin B.F., Jahangiri Jay M. *A new and comprehensive class of analytic functions*, Analele Universităţii din Oradea, Tom XV, 2008, 73–78.

[3] Sălăgean G.St. *Subclasses of univalent functions*, Lecture Notes in Math., Springer Verlag, Berlin, 1983, **1013**, 362–372.

Alina Alb Lupaş, Adriana Cătaş
Department of Mathematics and Computer Science
University of Oradea
1 Universitatii Street, 410087 Oradea
Romania
E-mail: *dalb@uoradea.ro, acatas@uoradea.ro*

Received August 22, 2009