Int. J. Appl. Math. Comput. Sci., 2000, Vol.10, No.3, 465–485

ON PARETO AND SALUKWADZE OPTIMIZATION PROBLEMS

WIESŁAW KOTARSKI*

In the paper, some problems of vector optimization are considered. Vector optimality is understood in the Pareto sense. Using the notion of Ponstein convexity, we formulate a 'scalarization' theorem. Two examples (vector optimization in $\mathbb{R}^2$ and an optimal-control problem for a parabolic equation with a vector performance index) are discussed. A Pareto boundary and a Salukwadze optimum are obtained for each of them. Additionally, for some vector optimization problems in $\mathbb{R}^2$, a criterion space is found. All calculations are performed with the use of Maple V. In the Appendix, a sketch of the proof of the main theorem on 'scalarization' is given.

Keywords: Pareto and Salukwadze optima, Pareto boundary, criterion space, scalarizarion

1. Pareto Optimality

Let Q be a closed, convex set with non-empty interior in a Banach space X, $I_i : X \mapsto \mathbb{R}$, $i = 1, \dots, s$ be given functionals and $I = [I_1, \dots, I_s]^T$, i.e. $I : X \mapsto \mathbb{R}^s$, be a vector performance index.

Consider the following vector optimization problem:

Problem (P): Find $x^0 \in Q$ such that

$$\text{Pareto} \min_{x \in Q \cap U(x^0)} I(x) = I(x^0),$$

where $U(x^0)$ is some neighbourhood of x^0.

Definition 1. A point $x^0 \in X$ is called *global (local) Pareto optimal* for Problem (P) if $x^0 \in Q$ and there is no $x^0 \neq x \in Q$ $(Q \cap U(x^0))$ with $I_i(x) \leq I(x^0)$ for $i = 1, \dots, s$ with strict inequality for at least one i, $1 \leq i \leq s$.

Following (Salukwadze 1974; 1979), we recall the following definitions:

Definition 2. A set $\Upsilon := \{y = I(x) \colon\ x \in Q\} \subset \mathbb{R}^s$ is called the *criterion space*.

This means that the criterion space is the image of the set Q through $I(x)$.

* Institute of Informatics, Silesian University, Będzińska 60, 41–200 Sosnowiec, Poland, e-mail: kotarski@gate.math.us.edu.pl

Definition 3. The product $\Upsilon_E := \Upsilon_1 \times \Upsilon_2 \times \cdots \times \Upsilon_s \subset \mathbb{R}^s$, where $\Upsilon_i = I_i(Q)$, $i = 1, \ldots, s$, is called the *extended criterion space.*

The components of $I(x^0)$ for a Pareto solution x^0 characterize a point in $\mathbb{R}^s$ lying on the boundary of Υ. In the rectangular coordinates with $I_1, \ldots, I_s$ axes, all the Pareto solutions to Problem (P) form the so-called Pareto boundary.

For special cases one can find a criterion space and a Pareto boundary which is some segment of the boundary of the criterion space. We give three examples to demonstrate that.

Example 1. Take

$$I_1 := x_1^2 + x_2^2, \quad I_2 := (1 - x_1)^2 + (1 - x_2)^2,$$

$$Q := \{(x_1, x_2) \in \mathbb{R}^2 : \ -1 \le x_1, x_2 \le 1\}. \qquad \blacklozenge$$

Example 2. Take

$$I_1 := x_1^2 - x_2^2, \quad I_2 := (1 - x_1)^2 + (1 - x_2)^2,$$

$$Q := \{(x_1, x_2) \in \mathbb{R}^2 : \ -1 \le x_1, x_2 \le 1\}. \qquad \blacklozenge$$

Example 3. Take

$$I_1 := \sin(x_1^2 + x_2^2), \quad I_2 := \sqrt{(1 - x_1)^2 + (1 - x_2)^2},$$

$$Q := \{(x_1, x_2) \in \mathbb{R}^2 : \ -1 \le x_1, x_2 \le 1\}. \qquad \blacklozenge$$

Usually it is not possible to obtain a criterion space analytically. But we need only few lines of a Maple V code to obtain an approximation of the criterion space numerically. Points from the square $-1 \le x_1, x_2 \le 1$ are generated randomly. Then, through $I(x)$, we obtain their images which are elements of the criterion space. Taking a sufficiently large number of points, we can obtain a satisfactory approximation of the criterion space. Knowing the form of the criterion space, we can have introductory information about the Pareto boundary.

In Figs. 1–3 the criterion space for Examples 1–3, is shown, respectively. The Maple programme below determines the underlying criterion space:

```
> n:=2000:
> die:=rand(-100..100):
> for i from 1 to n do
>  x:=[evalf(die()/100,evalf(die()/100)]:
>  I1:=x[1]^2+x[2]^2:
>  I2:=(1-x[1])^2+(1-x[2])^2:
>  l[i]:=[I1,I2]:
>  od:
> with(plots):
> pointplot([seq(l[i],i=1..n)],labels=['I1','I2'],
  title=`Criterion Space`);
```

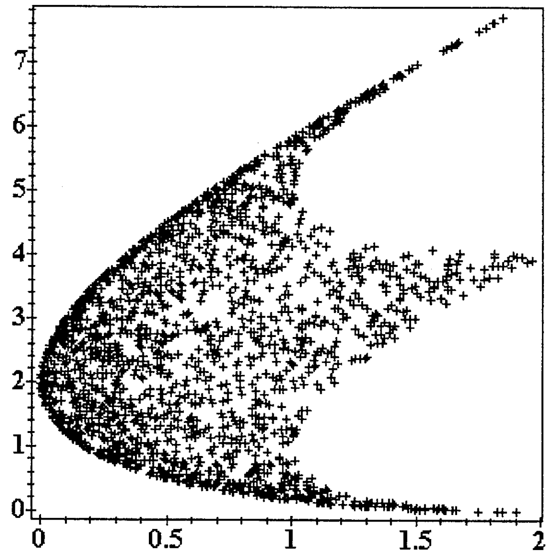

Fig. 1. Criterion space for Example 1.

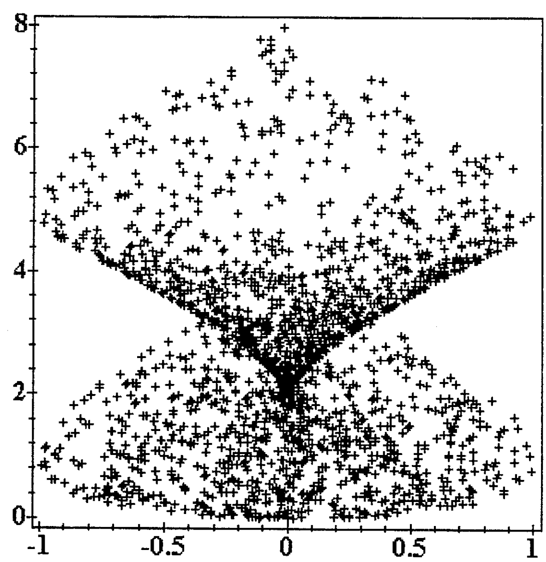

Fig. 2. Criterion space for Example 2.

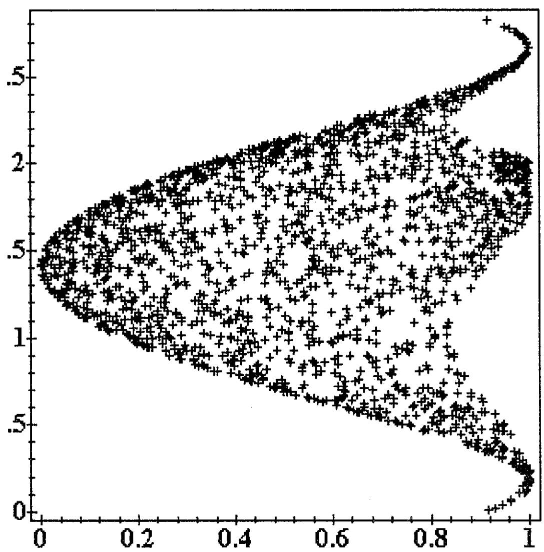

Fig. 3. Criterion space for Example 3.

The examples above show that the form of the criterion space is rather difficult to predict.

2. Problem 'Scalarization'

With Problem (P) the following scalar one can be associated:

Problem (S): For some $\lambda \in \mathbb{R}^s$ find $x_\lambda^0 \in Q$ such that

$$\sum_{i=1}^{s} \lambda_i I_i(x_\lambda^0) = \min_{x \in Q \cap U(x^0)} \sum_{i=1}^{s} \lambda_i I_i(x),$$

where $\lambda_i > 0$, $i = 1, \dots, s$, $\sum_{i=1}^{s} \lambda_i = 1$, Q being defined as in Problem (P).

Under additional assumptions, Problem (P) is equivalent to Problem (S) in the following sense: if x^0 is a Pareto optimal solution to Problem (P), then $x^0 = x_\lambda^0$ is an optimal solution to Problem (S) for some $\lambda_i > 0$, $i = 1, \dots, s$, $\sum_{i=1}^{s} \lambda_i = 1$, and vice versa.

The procedure of replacing Problem (P) by the equivalent Problem (S) is called the 'scalarization' of Problem (P).

A key role in the 'scalarization' problem is played by a special kind of convexity, namely the Ponstein convexity (Ponstein, 1967).

Definition 4. A functional $F\colon\ X \mapsto \mathbb{R}$ is called *Ponstein convex* if

$$F(x_2) \leq F(x_1) \Rightarrow F(\lambda x_1 + \mu x_2) < F(x_1), \quad \forall\, x_1 \neq x_2, \quad \lambda, \mu > 0, \quad \lambda + \mu = 1.$$

Strictly convex functionals are also Ponstein convex, but not every convex functional is Ponstein convex (Ponstein, 1967). The examples below show that the notions of convexity and Ponstein convexity generally are independent of each other.

In Figs. 4–7 we denote by SC, C and PC the strict convexity, convexity and Ponstein convexity, respectively. The notation 'SC: +' means that the function is strictly convex, while 'PC: −' means that the function is not Ponstein convex, etc.

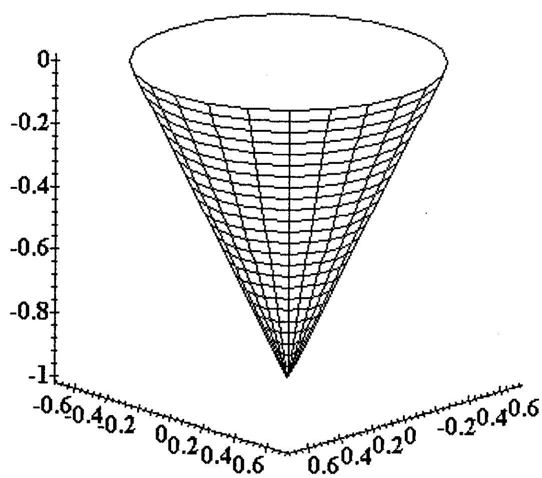

Fig. 4. SC: −, C: +, PC: +.

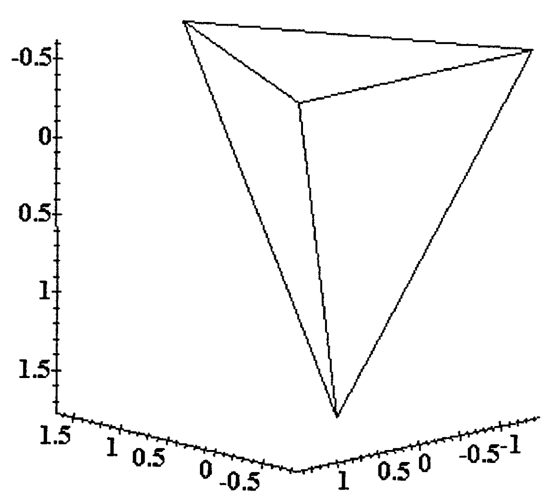

Fig. 5. SC: −, C: +, PC: −.

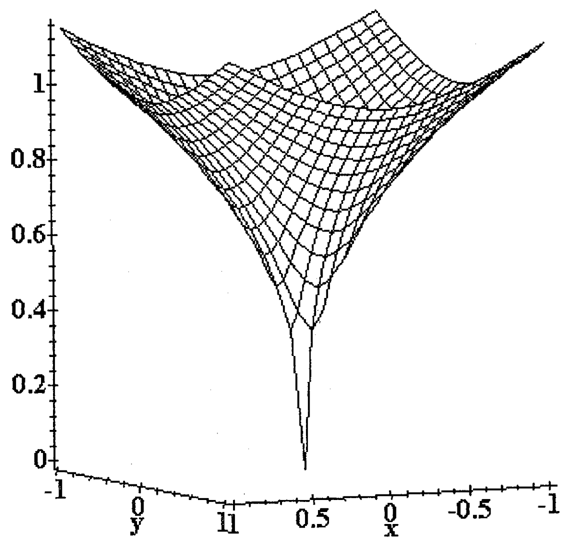

Fig. 6. SC: −, C: −, PC: +.

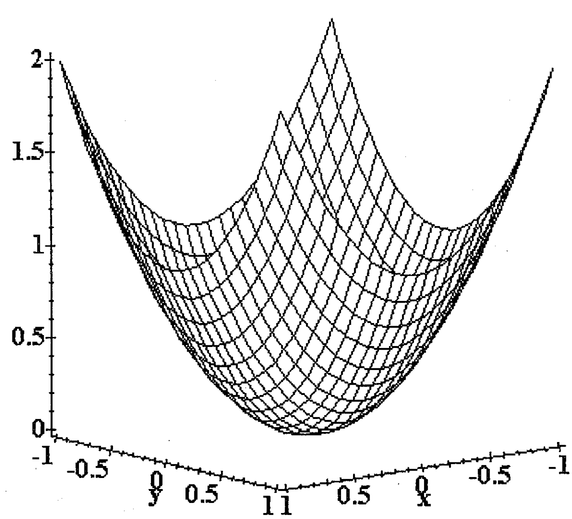

Fig. 7. SC: +, C: +, PC: +.

We have the following 'scalarization' theorem:

Theorem 1. *Assume that* $I_i : X \mapsto \mathbb{R}$, $i = 1, \ldots, s$ *are Fréchet differentiable, Ponstein convex and there exists a local Pareto optimum* x^0 *for Problem (P) such that* $x^0 \in \operatorname{dom} I_i$ *and* $\inf_x I_i(x) < I_i(x^0)$. *Then Problem (P) is equivalent to Problem (S) in the sense explained above.*

Remark 1. If, additionally, I_i, $i = 1, \ldots, s$ are convex, then every local Pareto optimum is also global Pareto one.

The proof of Theorem 1 is given in the Appendix. Accordingly, we can solve Problem (S) instead of Problem (P).

It is often observed that in the minimization of weighted combinations of objective functions an even spread of weights λ does not produce an even spread of points on the Pareto boundary. The points obtained using a uniformly spread set of values are actually concentrated in certain regions of the Pareto boundary. Roughly speaking, the weight λ is related to the 'slope' of the Pareto boundary in the criterion space. This problem is widely analysed in (Das and Dennis, 1997). To overcome this drawback of the 'scalarization' method, Das and Dennis (1998) proposed a new method called NBI (Normal-Boundary Intersection) which provides an even spread of points on the Pareto boundary. The NBI method is worked out for finite-dimensional spaces.

3. Salukwadze Optimality

Usually the set of Pareto solutions to Problem (P) has infinitely many elements. Therefore one has the dilemma which Pareto solution should be chosen as the best one. One of the methods to solve this problem was proposed by Salukwadze (1974; 1979). According to his suggestion, every component of the performance index, i.e. I_i, has to be minimized separately on the constrained set Q. A point with all minimal components is called the 'ideal point' (utopia or nadir). An 'ideal point' $\tilde{y}$ belongs to the extended criterion space Υ_E and usually it is not an element of Υ. A point on the Pareto boundary which is the nearest to the 'ideal point' in the sense of some metric is called the Salukwadze point. To that point in Q corresponds the Salukwadze optimum.

The Salukwadze point y^0 (if such a point exists at all) on the Pareto boundary nearest to the 'ideal point' $\tilde{y}$ in the sense of a metric ϱ (ϱ can be chosen arbitrarily, e.g. $\varrho(\tilde{y}, y) = \sqrt{\sum_{i=1}^{s} (\tilde{y}_i - y_i)^2}$) is an element of the criterion space Υ. This point determines a solution x^* such that $I(x^*) = y^0$. The point x^* is a Salukwadze optimum to Problem (P).

Assuming that $Q \subset X$ is a closed, convex and bounded (i.e. weakly compact) subset of a Hilbert space X for functions $I_i(x)$ which are continuous with respect to x, the existence of the Salukwadze optimum x^* results from the following considerations.

Define

$$I^{\lambda} := \min_{x \in Q} \sum_{i=1}^{s} \lambda_i I_i(x) = \sum_{i=1}^{s} \lambda_i I_i^{\lambda}.$$

We have the following estimate (see Appendix III in (Demyanov and Malozemtsev, 1975)):

$$\begin{aligned} \left| I^{\lambda} - I^{\tilde{\lambda}} \right| &= \left| \min_{x \in Q} \sum_{i=1}^{s} \lambda_i I_i(x) - \min_{x \in Q} \sum_{i=1}^{s} \tilde{\lambda}_i I_i(x) \right| \\ &\leq \max_{x \in Q} \left| \sum_{i=1}^{s} \lambda_i I_i(x) - \sum_{i=1}^{s} \tilde{\lambda}_i I_i(x) \right| \\ &= \max_{x \in Q} \left| \sum_{i=1}^{s} (\lambda_i - \tilde{\lambda}_i) I_i(x) \right| \to 0 \quad \text{as} \quad \lambda_i \to \tilde{\lambda}_i, \quad i = 1, \dots, s. \end{aligned}$$

This means that I^{λ} is continuous with respect to $\lambda = [\lambda_1, \dots, \lambda_s]^T$. The vector with components I_i^{λ}, $i = 1, \dots, s$ corresponds to that with I^{λ}.

The function $\sum_{i=1}^{s} (I_i^{\lambda} - \tilde{y}_i)^2$ representing the squared distance between the 'ideal point' $\tilde{y}$ and the points of the Pareto boundary is also continuous with respect to λ. Therefore, according to the Weierstrass Theorem, this function attains its minimum on the set $\lambda_i \geq 0$, $i = 1, \dots, s$, $\sum_{i=1}^{s} \lambda_i = 1$ for, say λ_i^0, $i = 1, \dots, s$.

If $\lambda_i^0 > 0$, $i = 1, \dots, s$, then I^{λ^0} (with $\lambda^0 = [\lambda_1^0, \lambda_2^0, \dots, \lambda_s^0]^T$) characterizes a point of the Pareto boundary and determines the Salukwadze point in the criterion space. A point x^* such that $I_i(x^*) = I_i^{\lambda^0}$, $i = 1, \dots, s$ is a Salukwadze optimum for Problem (P).

4. Examples

4.1. Vector Optimization in $\mathbb{R}^2$

We have to minimize a vector performance index in the Pareto sense as in Example 1. All the assumptions of Theorem 1 are satisfied, so we look for a minimum of the scalar function

$$\lambda\left(x_1^2 + x_2^2\right) + (1 - \lambda)\left[(1 - x_1)^2 + (1 - x_2)^2\right], \quad 0 \leq \lambda \leq 1.$$

Consequently, we get the global minimizers

$$x_{1\lambda} = x_{2\lambda} = \frac{1 - \lambda}{1 + \lambda}.$$

Their images through functionals $[I_1, I_2]^T$ for $\lambda \in (0, 1)$ characterize the Pareto boundary. The 'ideal' point is simply the origin, while the Salukwadze point obtained for $\lambda = 0.33$ has the coordinates $(0.5075, 0.4925)$ in the criterion space and $x_{1S} = x_{2S} = 0.5037$.

In Fig. 8 the criterion space, Pareto boundary and Salukwadze point are presented.

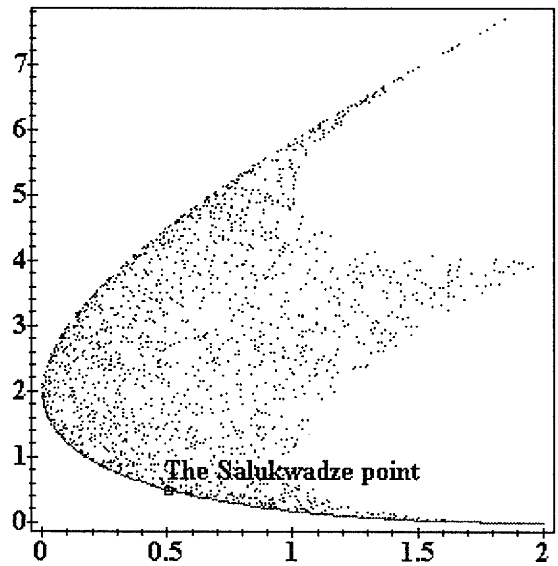

Fig. 8. Pareto boundary and Salukwadze point for Example 1.

4.2. Multi-Criterion Optimal-Control Problem

Consider the following Pareto optimization problem:

$$\frac{\partial y}{\partial t} - \frac{\partial^2 y}{\partial x^2} = u, \quad x \in (0,1), \quad t \in (0,T), \tag{1}$$

$$y(x,0) = y_p(x), \quad x \in (0,1), \tag{2}$$

$$y(0,t) = y(1,t) = 0, \quad t \in (0,T), \tag{3}$$

$$0 \le u(x,t) \le M, \quad x \in (0,1), \quad t \in (0,T), \tag{4}$$

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} \int_0^T \int_0^1 u^2(x,t)\,\mathrm{d}x\mathrm{d}t \\ \int_0^1 (y(x,T) - z_d(x))^2\,\mathrm{d}x \end{bmatrix} \to \min \quad \text{(in the Pareto sense).} \tag{5}$$

Equations (1)–(3) describe the process of heating a unit-length bar. While heating, the temperatures at both its ends are assumed to be constant and equal to zero. The initial distribution of the temperature is equal to $y_p(x)$. The time of heating T is assumed to be fixed and the maximal value of a non-negative control is less than a given number M.

The aim of the control is to achieve the required distribution of the temperature along the bar at the moment $t = T$ as close as possible to a given terminal state $z_d(x)$ caused by the smallest possible amount of the control energy. It is not possible to minimize both these quantities simultaneously. Therefore we look for a compromise, i.e. for some point on the Pareto boundary, namely the Salukwadze point.

Remark 2. The precise statement of the problem (1)–(5) requires a suitable choice of the state and control spaces (Kotarski, 1989; 1997a; 1997b; 1997c), but this problem will not be discussed here.

We have the following necessary and sufficient conditions of the Pareto optimality for the above problem (Kotarski, 1989; 1997a; 1997b; 1997c):

$$\frac{\partial y^0}{\partial t} - \frac{\partial^2 y^0}{\partial x^2} = u^0 \quad x \in (0,1), \quad t \in (0,T), \tag{6}$$

$$y^0(x,0) = y_p(x), \quad x \in (0,1), \tag{7}$$

$$y^0(0,t) = y(1,t) = 0, \quad t \in (0,T), \tag{8}$$

$$-\frac{\partial p}{\partial t} - \frac{\partial^2 p}{\partial x^2} = 0 \quad x \in (0,1), \quad t \in (0,T), \tag{9}$$

$$p(x,T) = (1-\lambda)\left[y^0(x,T) - z_d(x)\right] \quad x \in (0,1), \tag{10}$$

$$p(0,t) = p(1,t) = 0 \quad t \in (0,T), \tag{11}$$

$$\int_0^T \int_0^1 (p + \lambda u^0)(u - u^0)\,\mathrm{d}x\,\mathrm{d}t \geq 0, \quad 0 \leq u \leq M, \tag{12}$$

where u^0 and y^0 denote optimal control and optimal state, respectively, $\lambda \in (0,1)$.

The system of equations and inequalities (6)–(12) cannot generally be solved analytically because of mutual connections between the unknown functions u^0, y^0 and p. Therefore we have to resort to numerical methods.

In the region $\{(x,t) \in \mathbb{R}^2;\ 0 \leq x \leq 1,\ 0 \leq t \leq T\}$ we construct a grid of nodes (x_i, t_j). We have $x_i = ih = i/n+1$, $t_j = j/\tau = jT/m$, where h and τ stand for the step lengths along the x- and t-axes, respectively.

Writing $u_i^j = u(ih, j\tau)$ and $y_i^j = y(ih, j\tau)$, we can represent eqns. (6)–(11) in the difference form

$$\begin{cases} y^{j+1} = By^j + \tau u^j, & j = 0, 1, \dots, m-1, \\ y^0 = y_p(ih), & i = 1, \dots, n, \end{cases} \tag{13}$$

$$\begin{cases} p^{j-1} = Bp^j, & j = m, \quad m-1, \dots, 2, \\ p_i^m = \lambda_2\big(y^m - z_d(ih)\big), & i = 1, \dots, n, \end{cases} \tag{14}$$

where

$$B = \begin{bmatrix} 1-2\alpha, & \alpha & \dots & 0 \\ \alpha & 1-2\alpha & \alpha & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \alpha & 1-2\alpha & \alpha \\ 0 & \dots & \alpha & 1-2\alpha \end{bmatrix}$$

is an $n \times n$ tridiagonal matrix with $\alpha = \tau/h^2 \leq 1/2$ (this well-known condition guarantees the stability of the difference scheme).

The constraints on the controls can be presented in the form

$$0 \leq u_i^j \leq M, \quad i = 1, \dots, n, \quad j = 1, \dots, m. \tag{15}$$

The scalar performance index is approximated by the rectangle formula

$$I_h(u^j, y^j) := \lambda \sum_{i=1}^{n} \sum_{j=1}^{m} h\tau\big(u_i^j\big)^2 + (1-\lambda) \sum_{i=1}^{n} h\big(y_i^m - z_d(ih)\big)^2 \longrightarrow \min \tag{16}$$

The gradient is given by

$$I_h'\Big(\big(u_i^j\big), \big(y_i^j\big)\Big) = \big(p_i^j\big) + \lambda_1\big(u_i^j\big).$$

We used the gradient-projection method (Vasilev, 1981) for the approximated problem (13)–(16) described by the formula

$$\big(u_i^j\big)_{k+1} = \begin{cases} \big(u_i^j\big)_k - \beta_k\Big(\big(p_i^j\big)_k + \lambda\big(u_i^j\big)_k\Big) & \text{if} \quad 0 \leq \big(u_i^j\big)_k \leq M, \\ M & \text{if} \quad \big(u_i^j\big)_k > M, \\ 0 & \text{if} \quad \big(u_i^j\big)_k < 0, \end{cases} \tag{17}$$

where $i = 1, \dots, n, \quad j = 1, \dots, m.$

Remark 3. The values of the components of the performance index I_1 and I_2 usually have different orders of magnitude.Therefore taking into account their influence on the global performance index $\sum_{i=1}^{2} \lambda_i I_i(x)$ as well (according to the recommendations given in (Salukwadze, 1979)), one has to normalize them in the following way:

$$\tilde{I}_i = \frac{I_i - I_{i_{\min}}}{I_{i_{\max}} - I_{i_{\min}}}, \quad i = 1, 2,$$

where $I_{i_{\max}}$ and $I_{i_{\min}}$ denote the minimal and maximal values of I_i on the set Q, respectively. After normalization the 'ideal point' $\tilde{y}$ has two zero coordinates.

We applied normalization to the investigated problem. The convergence problems will not be discussed here. Numerical experiments showed that the computational process for the optimization problem under consideration was stable and convergent, but unfortunately the convergence speed was rather small and strongly dependent on the initial point u_0 and the step length β_0. They also confirmed the intuitive dependence of the parameter λ^* (which characterizes the Salukwadze point) on y_p, z_d and M. An in-depth analysis of that problem has not been made.

The problem (13)–(16) was solved for the following data:

$$n = 12, \quad m = 12, \quad \tau = \frac{1}{338}, \quad h = \frac{1}{13},$$
$$T = \frac{12}{338}, \quad \alpha = \frac{1}{2}, \quad M = 10, \quad y_p[i] = 0,$$
$$z_d[i] = \frac{(13 - i)i}{36}, \quad u_0[i, j] = \frac{1}{36}(13 - i)ij, \quad i, j = 1, \ldots, 12.$$

The initial coefficient of step β_0 was chosen experimentally as $8(1 - \lambda)^2 + 1$, the maximal number of iterations in the gradient-projection method was equal to 10 and the number of the calculated points on the Pareto boundary was equal to 21. The parameter ε in the stopping condition

$$\frac{|I_{k-1} - I_{k-2}|}{I_{k-2}} \leq \varepsilon \tag{18}$$

was equal to 0.001. We obtained $\lambda^* = 0.11$ which characterized the Salukwadze point on the Pareto boundary.

For that example we obtained only the Pareto boundary and Salukwadze point. Calculation of 21 points on the Pareto boundary took about 15 minutes on a PC with Pentium II processor. Finding the whole criterion space is rather a complicated and time-consuming problem. We observed that points on the Pareto boundary were not uniformly spread for even spreads of parameter λ (Fig. 9). We found experimentally the distribution of λ (Fig. 11) that ensured an almost even spread of points on the Pareto boundary (Fig. 10). In Fig. 12 an optimal control for the Salukwadze point is presented and in Fig. 13 the required and terminal states are given.

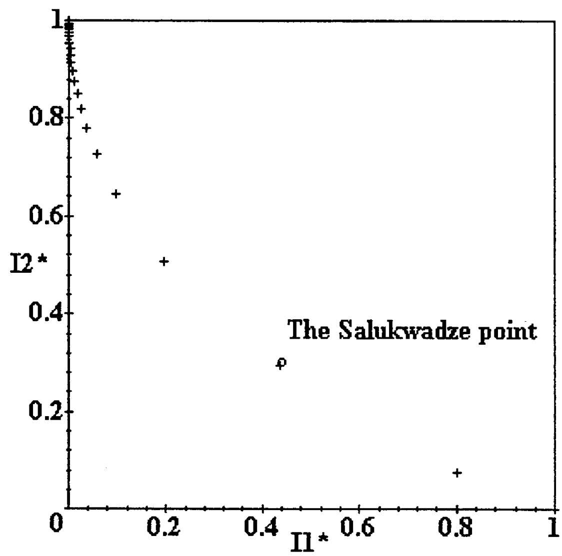

Fig. 9. Non-uniform spread of points.

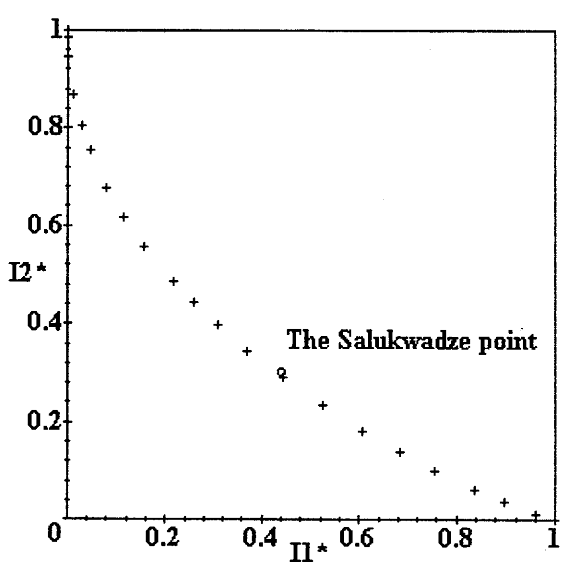

Fig. 10. Uniform spread of points.

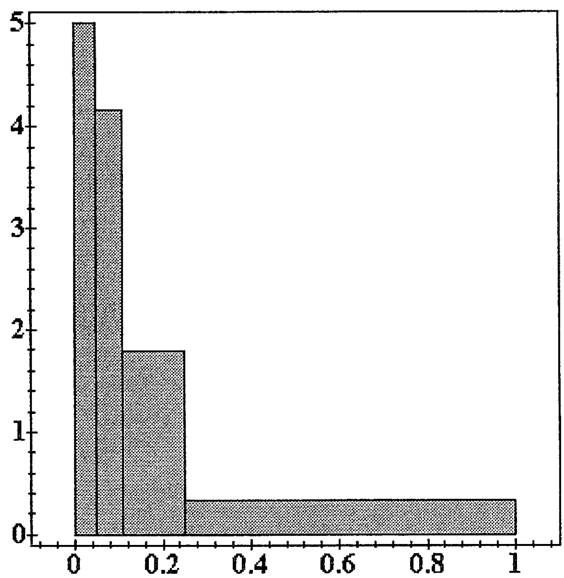

Fig. 11. Distribution of λ that ensures a uniform spread of points along the Pareto boundary.

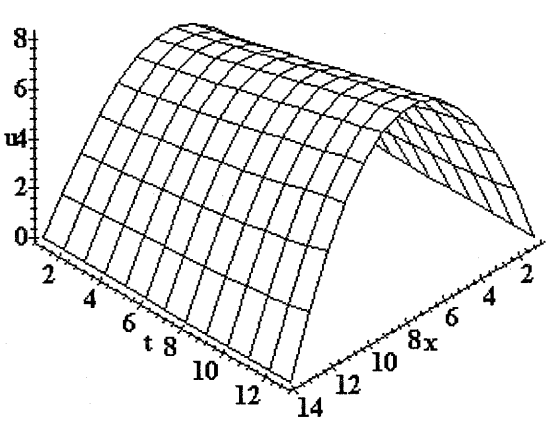

Fig. 12. Optimal distributed control.

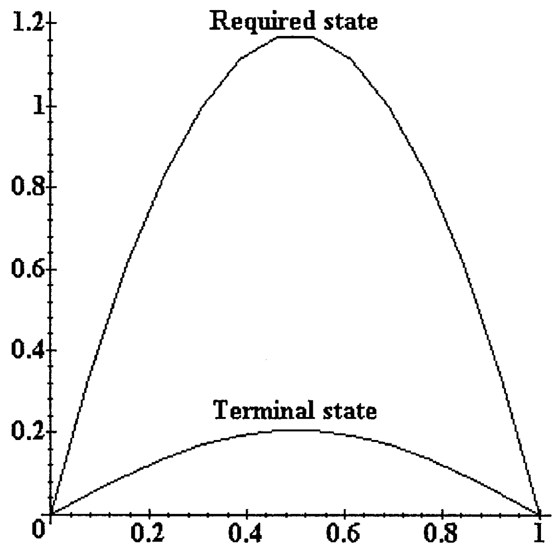

Fig. 13. Required and terminal states.

5. Final Remarks

The considerations of the paper are valid for vector optimization problems with a finite number of performance indices. Instead of a Pareto curve, we obtain a Pareto surface in higher dimensions. The 'scalarization' theorem is also valid for such problems.

Obtaining a good approximation of the Pareto surface is an important issue (Das and Dennis, 1998). For such problems the NBI method should be recommended instead of 'scalarization'.

The author of the paper works on a generalization of the NBI method to Hilbert spaces.

Appendix

In order to give a sketch of a proof for Theorem 1, we need some definitions of conical approximations (Censor, 1977; Girsanov, 1972). Let A be a set contained in a Banach space X and $F\colon\ X \mapsto \mathbb{R}$ be a given functional.

Definition A1. A set $\mathrm{TC}(A, x^0) := \{h \in X\colon\ \exists\, \varepsilon_0 > 0,\ \forall\, \varepsilon \in (0, \varepsilon_0),\ \exists\, r(\varepsilon) \in X;\ x^0 + \varepsilon h + r(\varepsilon) \in A\}$, where $r(\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0$ is called the tangent cone to the set A at the point $x^0 \in A$.

Definition A2. A set $\mathrm{AC}(A, x^0) := \{h \in X\colon\ \exists\, \varepsilon_0 > 0,\ \exists\, U(h),\ \forall\, \varepsilon \in (0, \varepsilon_0),\ \forall\, \bar{h} \in U(h);\ x^0 + \varepsilon\bar{h} \in A\}$, where $U(h)$ is a neighbourhood of h, is called the admissible cone to the set A at the point $x^0 \in A$.

Definition A3. A set $\mathrm{FC}(F,x^0) := \{h \in X: \ \exists\, \varepsilon_0 > 0, \ \exists\, U(h), \ \forall\, \bar{h} \in U(h), \ \forall\, \varepsilon \in (0,\varepsilon_0); \ F(x^0+\varepsilon\bar{h}) < F(x^0)\}$ is called the cone of descent directions of the functional F at the point $x^0 \in X$.

Definition A4. A set $\mathrm{NC}(F,x^0) := \{h \in X: \ \exists\, \varepsilon_0 > 0, \ \exists\, U(h), \ \forall\, \bar{h} \in U(h), \ \forall\, \varepsilon \in (0,\varepsilon_0); \ F(x^0+\varepsilon\bar{h}) \le F(x^0)\}$ is called the cone of non-descent directions of the functional F at the point $x^0 \in X$.

All the cones defined above are cones with vertices at the origin. The cones $\mathrm{AC}(A,x^0)$, $\mathrm{FC}(F,x^0)$ and $\mathrm{NC}(F,x^0)$ are open while the cone $\mathrm{TC}(A,x^0)$ is closed. If $\operatorname{int} A = \emptyset$, then $\mathrm{AC}(A,x^0)$ does not exist. Moreover, if $A_1,\dots,A_n, \in X$, $x^0 \in \bigcap_{i=1}^{n} A_i$, then

$$\bigcap_{i=1}^{n} \mathrm{TC}(A_i,x^0) \supset \mathrm{TC}\Big(\bigcap_{i=1}^{n} A_i, x^0\Big) \quad\text{and}\quad \bigcap_{i=1}^{n} \mathrm{AC}(A_i,x^0) = \mathrm{AC}\Big(\bigcap_{i=1}^{n} A_i, x^0\Big).$$

If the cones $\mathrm{TC}(A,x^0)$, $\mathrm{AC}(A,x^0)$, $\mathrm{FC}(F,x^0)$ and $\mathrm{NC}(F,x^0)$ are convex, then they are called regular and we denote them by $\mathrm{RTC}(A,x^0)$, $\mathrm{RAC}(A,x^0)$, $\mathrm{RFC}(F,x^0)$ and $\mathrm{RNC}(F,x^0)$, respectively.

Definition A5. Let K be a cone in X. The adjoint cone K^* of K is defined as

$$K^* := \{f \in X^*; \quad f(x) \ge 0 \quad \forall\, x \in K\},$$

where X^* denotes the dual space of X.

For Problem (P) we can formulate the following necessary condition for Pareto optimality:

Lemma A1. *For Problem (P) assume that*

(i) there exist cones $\mathrm{RAC}(Q,x^0)$, $\mathrm{RFC}(I_i,x^0)$, $\mathrm{RNC}(I_i,x^0)$, $i=1,\dots,s$, *and*

(ii) $x^0 \in Q$ *is a local Pareto optimum for Problem (P).*

Then

$$\mathrm{RFC}(I_i,x^0) \cap \Big(\bigcap_{j=1, j\ne i}^{s} \mathrm{RNC}(I_j,x^0)\Big) \cap \mathrm{RAC}(Q,x^0) = \emptyset, \quad i=1,\cdots,s. \tag{A1}$$

Proof. For the sake of contradiction, suppose that there is an i $(1 \le i \le s)$ such that (A1) is non-empty. Then there exists $h \in X$ such that:

$$\exists\, U^i(h) \ \ \exists\, \varepsilon^i > 0 \ \ \forall\, \bar{h} \in U^i(h) \ \ \forall\, \varepsilon \in (0,\varepsilon^i); \quad I_i(x^0+\varepsilon\bar{h}) < I_i(x^0),$$

$$\exists\, U^j(h) \ \ \exists\, \varepsilon^j > 0 \ \ \forall\, \bar{h} \in U^j(h) \ \ \forall\, \varepsilon \in (0,\varepsilon^j); \quad I_j(x^0+\varepsilon\bar{h}) \le I_j(x^0),$$
$$j=1,\cdots,s, \ \ j \ne i,$$

$$\exists\, U(h) \ \ \exists\, \varepsilon_0 > 0 \ \ \forall\, \bar{h} \in U(h) \ \ \forall\, \varepsilon \in (0,\varepsilon_0); \quad x^0+\varepsilon\bar{h} \in Q.$$

Let us define

$$\tilde{\varepsilon} := \min\left\{\varepsilon^i, \varepsilon^j, (j = 1, \ldots, s,\ j \neq i), \varepsilon_0\right\},$$

$$\tilde{U}(h) := U^i(h) \cap \Big(\bigcap_{j=1, j\neq i}^{s} U^j(h)\Big) \cap U(h).$$

Clearly, $x^0 \neq \hat{x} := x^0 + \tilde{\varepsilon}h(\tilde{\varepsilon}) \in Q$ and $I_i(\hat{x}) \leq I_i(x^0)$, $i = 1, \ldots, s$ with at least one strict inequality. This contradicts the Pareto optimality of x^0. ■

Condition (A1) in Lemma A1 can be reformulated in a more convenient form:

Theorem A1. *For Problem (P) assume that*

(i) there exist cones $\mathrm{RAC}(Q, x^0)$, $\mathrm{RFC}(I_i, x^0)$, $\mathrm{RNC}(I_i, x^0)$, $i = 1, \ldots, s$, *and*

(ii) $x^0 \in Q$ *is a local Pareto optimum for Problem (P).*

Then the following equations (the so-called Euler-Lagrange equations) must hold:

$$f_i + \sum_{j=1,\ j\neq i}^{s} f_j^{(i)} + \varphi = 0, \quad i = 1, \ldots, s,$$

where $f_i \in [\mathrm{RFC}(I_i, x^0)]^*$, $f_j^{(i)} \in [\mathrm{RNC}(I_j, x^0]^*$, $j = 1, \ldots, s$, $j \neq i$, $\varphi \in [\mathrm{RAC}(Q, x^0)]^*$, *and the functionals are not simultaneously equal to zero.*

Proof. If x^0 is a local Pareto optimum for Problem (P) and there exist suitable cones, then (A1) holds. Applying to (A1) the Dubovitskii-Milyutin Theorem (Lemma 5.11 in (Girsanov, 1972)), we obtain the required conclusion. ■

Remark A1. If, in addition to the assumptions of Theorem A1, I_i's are such that for every $i = 1, \ldots, s$, $[\mathrm{RFC}(I_i, x^0)]^* = [\mathrm{RNC}(I_i, x^0)]^*$, then s equations appearing there in the conclusion reduce to the single equation (the Euler-Lagrange Equation)

$$\sum_{i=1}^{s} f_j + \varphi = 0,$$

where $f_i \in [\mathrm{RFC}(I_i, x^0)]^*$, $i = 1, \ldots, s$, $\varphi \in [\mathrm{RAC}(Q, x^0)]^*$ and the functionals are not simultaneously equal to zero.

Remark A2. If a closed convex set Q is such that $\mathrm{int}\, Q = \emptyset$, then we can apply the cone $\mathrm{RTC}(Q, x^0)$ instead of $\mathrm{RAC}(Q, x^0)$.

To find conditions ensuring the equality $[\mathrm{RFC}(I_i, x^0)]^* = [\mathrm{RNC}(I_i, x^0)]^*$, we need the following results:

Lemma A2. *Let* $F\colon\ X \mapsto \mathbb{R}$ *be continuous and Ponstein convex. If* $x^0 \in \mathrm{dom}\, F$ *and* $\inf_x F(x) < F(x^0)$, *then* $\mathrm{int}\,\{x\colon\ F(x) \leq F(x^0)\} \neq \emptyset$ *and* $\{x\colon\ F(x) < F(x^0)\} = \mathrm{int}\,\{x\colon\ F(x) \leq F(x^0)\}$.

Proof. The proof of Lemma 5.4 given in (Censor, 1977) is valid without any changes for any arbitrary Banach space X. ■

Lemma A3. *If* $F: X \mapsto \mathbb{R}$ *is continuous and Ponstein convex,* $x^0 \in \operatorname{dom} F$, $\inf_x F(x) < F(x^0)$, *then*

$$[\mathrm{FC}(F, x^0)]^* = [\mathrm{NC}(F, x^0)]^* .$$

Proof. It is easy to see that

$$\mathrm{FC}(F, x^0) = \mathrm{AC}(A, x^0) \quad \text{and} \quad \mathrm{NC}(F, x^0) = \mathrm{AC}(B, x^0),$$

where $A := \{x;\ F(x) < F(x^0)\}$ and $B := \{x;\ F(x) \le F(x^0)\}$.

According to Proposition 1.2.5(i) of (Laurent, 1972) and Lemma A2, we have $\mathrm{AC}(A, x^0) = \mathrm{AC}(B, x^0)$, which establishes our claim. ■

The next lemma gives conditions under which every local Pareto optimum is also a global Pareto one for Problem (P).

Lemma A4. *Let* $I_i: X \mapsto \mathbb{R}$, $i = 1, \ldots, s$ *be convex functionals and* $Q \subset X$ *a convex set. Then every point being a local Pareto optimum for Problem (P) is also a point of a global Pareto optimum.*

Proof. For a proof see Theorem 5.1 in (Censor, 1977). ■

We formulate now necessary optimality conditions for Problem (P).

Theorem A2. *For Problem (P) let us assume that*

(i) $I_i: X \mapsto \mathbb{R}$ *is convex continuous and Ponstein convex,*

(ii) x^0 *is a local Pareto optimum for Problem (P) such that* $x^0 \in \operatorname{dom} I_i$ *and* $\inf_x I_i(x) < I_i(x^0)$, $i = 1, \ldots, s$,

(iii) Q *is closed and convex in* X.

Then x^0 *is a global Pareto optimum for Problem (P) if the Euler-Lagrange equation from Remark A1 is fulfilled.*

Proof. Since all the assumptions of Lemma A4 are fulfilled, any local Pareto optimum is also a global Pareto one.

The convexity of I_i ensures that of $A_i := \{x:\ I_i(x) < I_i(x^0)\}$ and $B_i := \{x:\ I_i(x) \le I_i(x^0)\}$, $i = 1, \ldots, s$. On the basis of Theorem 1.3.4 of (Laurent, 1972), the cones $\mathrm{FC}(I_i, x^0)$ and $\mathrm{NC}(I_i, x^0)$ are convex. With the same theorem, the convexity of Q implies that of $\mathrm{AC}(Q, x^0)$. All the assumptions of Theorem A1 are met and, additionally, the conditions of Remark A1 are satisfied, so the conclusion of our assertion follows. ■

Lemma A5. *Assuming that* $I_i\colon\ X \mapsto \mathbb{R},\ i = 1,\dots,s$ *are continuous, convex, and Ponstein convex with* $x^0 \in \operatorname{dom} I_i$ *and* $\inf_x I_i(x) < I_i(x^0)$, *we have*

$$\Big[\mathrm{RFC}\Big(\sum_{i=1}^{s}\lambda_i I_i, x^0\Big)\Big]^* \subset \sum_{i=1}^{s}\big[\mathrm{RFC}(I_i,x^0)\big]^*$$

for $\lambda_i > 0,\ i = 1,\dots,s,\ \sum_{i=1}^{s}\lambda_i = 1$.

Proof. The convexity of I_i ensures that of $A_i := \{x\colon\ I_i(x) < I_i(x^0)\}$ and $B_i := \{x\colon\ I_i(x) \le I_i(x^0)\},\ i = 1,\dots,s$. Using Lemma A3, we get

$$\big[\mathrm{RFC}(I_i,x^0)\big]^* = \big[\mathrm{RNC}(I_i,x^0)\big]^*,\quad i = 1,\dots,s.$$

Then, from Lemma 5.10 of (Girsanov, 1972), it follows that

$$\sum_{i=1}^{s}\big[\mathrm{RFC}(I_i,x^0)\big]^* = \Big[\bigcap_{i=1}^{s}\mathrm{RFC}(I_i,x^0)\Big]^*. \tag{A2}$$

Further, taking into account the following obvious facts:

(a) $\bigcap_{i=1}^{s} A_i \subset A$, where $A := \{x\colon\ \sum_{i=1}^{s}\lambda_i I_i(x) < \sum_{i=1}^{s}\lambda_i I_i(x^0)\}$,

(b) $A \subset B \Rightarrow \mathrm{RFC}(A,x^0) \subset \mathrm{RFC}(B,x^0)$ and

$$\mathrm{RFC}(A,x^0)\cap\mathrm{RFC}(B,x^0) = \mathrm{RFC}(A\cap B,x^0)$$

for arbitrary sets $A,\ B \subset X$ and $x^0 \in A\cap B$,

(c) $K_1 \subset K_2 \Rightarrow K_2^* \subset K_1^*$ for arbitrary cones $K_1,\ K_2$,

and the definitions of cones RFC and RAC, we obtain

$$\bigcap_{i=1}^{s}\mathrm{RFC}(I_i,x^0) = \bigcap_{i=1}^{s}\mathrm{RAC}(A_i,x^0) = \mathrm{RAC}\Big(\bigcap_{i=1}^{s}A_i,x^0\Big) \subset \mathrm{RAC}(A,x^0)$$

$$= \mathrm{RFC}\Big(\sum_{i=1}^{s}\lambda_i I_i,x^0\Big). \tag{A3}$$

From (A2) and (A3), for the adjoint cones we get

$$\Big[\mathrm{RFC}\Big(\sum_{i=1}^{s}\lambda_i I_i,x^0\Big)\Big]^* \subset \Big[\bigcap_{i=1}^{s}\mathrm{RFC}(I_i,x^0)\Big]^* = \sum_{i=1}^{s}\big[\mathrm{RFC}(I_i,x^0)\big]^*,$$

which completes the proof. ∎

Under additional assumptions on I_i, we get an equality in the conclusion of Lemma A5. This equality means that the functionals appearing in the Euler-Lagrange equation associated with the performance index have the same form both for the vector-valued Problem (P) and for scalar ones, i.e. for Problem (S). This implies the equivalence of Problems (P) and (S).

And now we are ready to give the proof of Theorem 1.

Proof of Theorem 1. It is sufficient to show that in the conclusion of Lemma A5 we have the opposite inclusion, namely '$\supset$'. From the inequality $\inf_x I_i(x) < I_i(x^0)$ it follows that $\mathrm{RFC}(I_i, x^0) \neq \emptyset$, $i = 1, \dots, s$. With Theorem 7.5 of (Girsanov, 1972) we have $\mathrm{RFC}(I_i, x^0) = \{\bar{x}\colon\ I_i'(x^0)\bar{x} < 0\}$. Hence $I_i'(x^0) \neq 0$, $i = 1, \dots, s$.

From Theorem 10.2 of (Girsanov, 1972) it follows that an arbitrary element of $\sum_{i=1}^{s} [\mathrm{RFC}(I_i, x^0)]^*$ has the form

$$f = \sum_{i=1}^{s} \alpha_i I_i'(x^0), \quad \alpha_i \leq 0, \quad i = 1, \dots, s.$$

It is easily seen that $f \equiv 0$ satisfies the inclusion '$\supset$'. Let $f \neq 0$. Setting $\beta := \sum_{i=1}^{s} \alpha_i$, we have $\beta < 0$. Further, write $\hat{\lambda}_i := \alpha_i / \beta$, $\hat{\lambda}_i \in [0, 1]$, $i = 1, \dots, s$.

We get

$$f = \sum_{i=1}^{s} \beta \frac{\alpha_i}{\beta} I_i'(x^0) = \sum_{i=1}^{s} \beta \hat{\lambda}_i I_i'(x^0).$$

One the other hand, from Theorems 7.5 and 10.2 of (Girsanov, 1972), we deduce that an arbitrary element of $[\mathrm{RFC}(\sum_{i=1}^{s} \lambda_i I_i(x^0))]^*$ has the form

$$\varphi = \gamma \sum_{i=1}^{s} \lambda_i I_i'(x^0) = \sum_{i=1}^{s} \gamma \lambda_i I_i'(x^0), \quad \gamma \leq 0.$$

The products $\beta \hat{\lambda}_i$ and $\gamma \lambda_i$ are non-positive. Therefore the functionals f and φ are the same up to some multiplicative constant. Taking into account Lemma A5, we get the required equality, which completes the proof. ■

References

Das I. and Dennis J. (1997): *A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems.* — Struct. Optim., Vol.14, No.1, pp.63–69.

Das I. and Dennis J. (1998): *Normal-Boundary Intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems.* — SIAM J. Optimiz., Vol.8, No.3, pp.631–657.

Censor Y. (1977): *Optimality in multi objective problems.* — Appl. Math. Optim., No.4, pp.41–59.

Demyanov W.F. and Malozemtsev W.M. (1975): *Introduction to Minimax Methods.* — Warsaw: Polish Scientific Publishers (in Polish).

Girsanov I.V. (1972): *Lectures on Mathematical Theory of Extremum Problems.* — New York: Springer–Verlag.

Kotarski W. (1989): *Characterization of Pareto optimal points in problems with multi-equality constraints.* — Optimization, Vol.20, No.1, pp.93–106.

Kotarski W. (1997a): *Quadratic Pareto-optimal control problem for a parabolic equation.* — Proc. *European Control Conference ECC'97*, Brussels, (CD–ROM Ref. No. 711).

Kotarski W. (1997b): *On the Salukwadze optimal control for a parabolic equation.* — Proc. 15-th IMACS World Congress, Berlin, Vol.V, pp.143–148, (CD–ROM Ref. No. 143).

Kotarski W. (1997c) *Some problems of optimal and Pareto optimal control for distributed parameter systems.* — Reports of Silesian University, No.1668, Katowice, Poland, pp.1–93.

Laurent P.J. (1972): *Approximation et Optimisation.* — Paris: Hermann.

Ponstein J. (1967): *Seven kinds of convexity.* — SIAM Rev., Vol.9, No.1, pp.115–119.

Salukwadze M.E. (1974): *On the existence of solutions in problems of optimization under vector-valued criteria.* — J. Optim. Th. Applicns., Vol.13, No.2, pp.203–217.

Salukwadze M.E. (1979): *Vector-Valued Optimization Problems in Control Theory.* — New York: Academic Press.

Vasilev F.P. (1981): *Methods for Solving Extremal Problems.* — Moscow: Nauka (in Russian).

Received: 14 October 1999
Revised: 14 July 2000