Working Paper

MULTIOBJECTIVE TRAJECTORY OPTIMIZATION
AND MODEL SEMIREGULARIZATION

Andrzej P. Wierzbicki

December 1980
WP-80-181

International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

MULTIOBJECTIVE TRAJECTORY OPTIMIZATION AND MODEL SEMIREGULARIZATION

Andrzej P. Wierzbicki

December 1980
WP-80-181

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

SUMMARY

The typical formulation of an optimal control or dynamic optimization problem is to optimize a scalar performance functional; less frequently, also vectors of performance functionals are considered in multiobjective optimization. However, there are practical problems -- mostly related to the use of dynamic control models in economic planning -- where the objectives are stated in terms of desirable trajectories. If the goal would be to approximate the desired trajectory from both sides, then the problem could be equivalently stated as a typical approximation problem. However, in many cases the desired trajectories have the meaning of aspiration levels: if possible, they should be exceeded.

The paper presents a mathematical formulation of a multiobjective trajectory optimization problem, various theoretical approaches to this problem -- including interpretations as a generalized Lagrangian functional approach and as a semi-regularization procedure for ill-posed problems, a review of possible computational approaches and examples of actual computations.

MULTIOBJECTIVE TRAJECTORY OPTIMIZATION AND MODEL SEMIREGULARIZATION

Andrzej P. Wierzbicki

1. MOTIVATION

Dynamic optimization problems are usually formulated in terms of minimization (or maximization) of a given objective functional, also called performance functional. Even if the performance of a dynamic system is specified in terms of closeness to a given trajectory, a performance functional corresponding to a distance from this trajectory is still being used. However, not all practical problems can be usefully formulated as optimization problems with given performance functionals.

Very often, particularly in economic applications, the purpose of optimization is not to propose 'the optimal solution', but rather to generate reasonable alternatives in response to users' requirements while eliminating clearly inferior alternatives. It is not likely that a user would specify his requirements in form of a performance functional. More likely, he would specify his aspirations in form of a reasonable or desirable trajectory of the dynamic system being investigated. Since the desirable trajectory reflects his judgment and experience, it might not be attainable for a particular model of the dynamic system being studied. However, if the desirable trajectory happens to be attainable, the user can often specify also what trajectories should be considered as naturally better than the desirable ones.

As an example, consider a dynamic economic model that specifies, for various monetary and fiscal policies, the resulting economic growth and inflation rates. An economist, while working with this model, is perfectly able to specify reasonable growth and inflation rates trajectories although these trajectories may not be attainable for the model. If they are attainable however, he would not be satisfied by them, particularly if he knew that he could obtain either higher growth rate or lower inflation rate or both. Thus, we cannot use the classical device of minimization of a performance functional corresponding to the distance from the desired trajectory; this device works well only when the desired trajectory is naturally better than the attainable ones. Another classical device is the formulation of a social welfare functional and its maximization; but the information needed for formulating the social welfare functional is much larger than the information contained in a desirable trajectory. Moreover, a social welfare functional implies 'the optimal solution' without allowing for the possibility of checking various alternatives by changing the desired trajectory.

Therefore, a concept of multiobjective trajectory optimization based on reference trajectories has been recently introduced (Wierzbicki 1979) and practically applied to some issues in economic modeling (Kallio et al. 1980). This concept, while being strongly related to some basic concepts in satisficing decision making (Wierzbicki 1980), deserves a separate study. The purpose of this paper is to present, in more detail, the theory, some computational approaches and applicational aspects of multiobjective trajectory optimization.

2. BASIC THEORY IN A NORMED SPACE

All the theory in this section could be introduced in referring to a more detailed dynamic model, for example, the classical control model described by an ordinary differential state equation and an output equation. However, the precise form of a dynamic model does not matter, and the theory is also applicable for models described by difference-differential equations (with delays), by partial differential equations, integral equations, etc.

To obtain a possible compact presentation of basic ideas, let us start with an abstract formulation in normed spaces. Let $u \in E_u$ be a *control trajectory*, shortly called *control*; E_u is a Banach space, say, the space of essentially bounded functions $L^\infty([t_0;t_1],R^m)$, or the space of square integrable functions $L^2([t_0,t_1],R^m)$, etc. Additionally, control constraints $u \in V \subset E_u$ might be given. Let $x \in E_x$ be a *state trajectory*, shortly *state*, defined by a mapping $X: E_u \to E_x$, $x = X(u)$. Conditions, under which the mapping X corresponds to a model of a *dynamic system* and can be expressed as a resolving operation for a *state equation* are given, for example, in Kalman et al. 1969, and will not be discussed here. A proper choice of a Banach space E_x might be the Sobolev space of absolutely continuous functions with essentially bounded derivatives $W^\infty([t_0;t_1],R^n)$ or with square integrable derivatives $W^2([t_0;t_1],R^n)$ -- see, e.g., Wierzbicki, 1977b. However, these properties are needed only for a more detailed development of the form of the dynamic model, and, at this stage of abstraction, E_u and E_x could be just any linear topological spaces.

More important are the assumptions concerning *output trajectory*, shortly *output* $y \in E_y$, defined as a result of a mapping $Y: E_x \times E_u \to E_y$, $y = Y(x,u)$. A properly chosen Banach space E_y should have the same character as the space E_u; thus, $E_y = L^\infty([t_0;t_1],R^p)$ or $E_y = L^2([t_0;t_1],R^p)$. Since the notion of an output is relative to the purpose of the model, we might consider only those output variables that are relevant for the purpose of multiobjective trajectory optimization, the number of those variables being p. Thus, a notion of a partial preordering (partial ordering of equivalence classes) is assumed to be given in the output space E_y. Although more general assumptions are possible, it is convenient to suppose that this partial preordering is transitive and, therefore, can be defined by specifying a positive cone $D \subset E_y$; the cone D is assumed to be closed, convex and proper, i.e. $D \neq E_y$. The partial preordering relation takes then the form

$$(1) \qquad y_1, y_2 \in E_y \quad , \qquad y_1 \leqslant y_2 \Leftrightarrow y_2 - y_1 \in D$$

with the corresponding equivalence relation

(2) $\quad y_1, y_2 \in E_y \quad , \quad y_1 \sim y_2 \Leftrightarrow y_2 - y_1 \in D \cap -D$

and the strong partial preordering relation

(3) $\quad y_1, y_2 \in E_y \quad , \quad y_1 \prec y_2 \Leftrightarrow y_2 - y_1 \in \tilde{D} \overset{df}{=} D \setminus (D \cap -D)$

as well as the strict partial preordering relation

(4) $\quad y_1, y_2 \in E_y \quad , \quad y_1 \prec\!\prec y_2 \Leftrightarrow y_2 - y_1 \in \mathring{D}$

where $\mathring{D}$ is the interior of the cone D. In some spaces, naturally defined positive cones might have empty interiors; however, we can define then the quasi-strict partial preordering through replacing $\mathring{D}$ in (4) by $\mathring{D}^q$, the quasi-interior of D

(5) $\quad \mathring{D}^q = \{y \in D : \langle y^*, y\rangle > 0 \ , \forall y^* \in \tilde{D}^* \overset{df}{=} D^* \setminus (D^* \cap -D^*)\}$

where

(6) $\quad D^* = \{y^* \in E_y^* : \langle y^*, y\rangle \geq 0, \forall y \in D\}$

is the dual cone to D, E_y^* being the dual space to E_y and $\langle \cdot, \cdot \rangle$ denoting the duality relation between E_y^* and E_y (the general form of a linear continuous functional from E_y^* over E_y).

For example, if $E_y = L^2([t_0;t_1], R^p)$, then a positive cone can be naturally defined by

(7) $\quad D = L_+^2([t_0;t_1], R^p) =$

$$\{y \in L^2([t_0;t_1], R^p) : y^i(t) \geq 0, \text{ a.e. for } t \in [t_0;t_1], \forall i = 1, \dots, p\} .$$

The equivalence classes (2) are then composed of functions that are equal to each other almost everywhere on $[t_0;t_1]$, which coincides with classical definitions of equivalence classes in L^2. The strong partial preordering (3) relates functions which have components $y_1^i(t) \leq y_2^i(t)$ a.e. on $[t_0;t_1]$, $\forall i = 1, \dots, p$, such that the inequality $y_1^i(t) < y_2^i(t)$ holds for at least one i and at least on a subset of $[t_0;t_1]$ of nonzero measure. Since the cone (7) has empty interior, there are no $y_1, y_2 \in E_Y$ that are strictly related. However, $D^* = D$ in this case (L^2 is a Hilbert space and its dual can be made identical with it). Moreover, D has a nonempty quasi-interior:

(8) $\overset{\circ}{D}^q = \{y \in L^2([t_0;t_1],R^p) : y^i(t) > 0 \text{ a.e. for } t \in [t_0;t_1], \forall i = 1,\ldots,p\}$

and the quasi-strict partial preordering relates functions with components $y_1^i(t) < y_2^i(t)$ a.e. on $[t_0;t_1]$, $\forall i = 1,\ldots,p$. For other examples of positive cones see Wierzbicki and Kurcyusz, 1977.

The set of admissible controls V and the mappings X,Y define together the *set of attainable outputs*

(9) $$Y_V = Y(X(V),V) \subset E_Y \quad .$$

Usually, we cannot describe the full set Y_V analytically because the mappings X,Y are too complicated; however, it is assumed that we can generate elements of this set, at least numerically, by solving the dynamic model for a given $u \in V$. On the other hand, suppose we are interested only in D-*maximal elements* $\hat{y} \in \hat{Y}_V$

(10) $$\hat{Y}_V = \{\hat{y} \in Y_V : Y_V \cap (\hat{y} + \tilde{D}) = \phi\}$$

which are natural generalizations of Pareto-maximal outputs for the case of trajectory optimization. If the cone $\overset{\circ}{D}$ is nonempty, it is sometimes convenient to consider also *weak* D-*maximal elements* $\hat{y} \in \hat{Y}_V^w$

(11) $$\hat{Y}_V^w = \{\hat{y} \in Y_V : Y_V \cap (\hat{y} + \overset{\circ}{D}) = \phi\}$$

or *quasi-weak* D-*maximal elements* $\hat{y} \in \hat{Y}_V^{wq}$ obtained as in (11) while replacing $\overset{\circ}{D}$ by $\overset{\circ}{D}^q$. Clearly, $\hat{Y}_V \subset \hat{Y}_V^{wq} \subset \hat{Y}_V^w \subset Y_V$. Sometimes it is also convenient to consider a smaller set $\hat{Y}_V^\varepsilon \subset \hat{Y}_V$ of D_ε-*maximal* elements of Y_V:

(12) $$\hat{Y}_V^\varepsilon = \{\hat{y} \in Y_V : Y_V \cap (\hat{y} + \tilde{D}_\varepsilon) = \phi\}$$

where D_ε is defined as a conical ε-neighborhood of D:

$$(13) \qquad D_\varepsilon = \{y \in E_y : \mathrm{dist}(y,D) < \varepsilon\|y\|\} \; ; \; \tilde{D}_\varepsilon = D_\varepsilon \setminus (D_\varepsilon \cap -D_\varepsilon) \quad .$$

Since dist(y,D) is a continuous functional of y, the cone D_ε is an open cone, that is, an open set augmented with the point 0 or the set $D_\varepsilon \cap -D_\varepsilon$. Thus, $\tilde{D}_\varepsilon$ is an open set, and D_ε-maximality is equivalent to weak D_ε-maximality.

For example, if $D = L^2_+([t_0;t_1],R^p)$ as in (7), then, using an argument via projections on cones in Hilbert spaces as in Wierzbicki and Kurcyusz (1977) it can be shown that:

$$(14) \quad D_\varepsilon = \{y \in L^2([t_0;t_1],R^p) : \|y_-\| < \varepsilon\|y\|\} ; y^i_-(t) = \min(0,y^i(t))$$

and D_ε has an interior: at any point $y \in D$ we can center a ball with radius $\delta < \varepsilon$, contained in D_ε.

A classical method of generating D-maximal elements of Y_V is that of maximizing a (quasi-) strictly positive linear functional $y^* \in \mathring{D}^{*q}$ over $y \in Y_V$:

$$(15) \qquad \hat{y} \in \mathrm{Arg} \max_{y \in Y_V} \langle y^*,y\rangle \quad , \qquad y^* \in \mathring{D}^{*q} \Rightarrow \hat{y} \in \hat{Y}_V \quad .$$

However, it is very difficult to express the experience and judgment of a user of the model in terms of a linear functional (called also weighting functional) $y^* \in \mathring{D}^{*q}$; in the case of dynamic trajectory optimization, it often becomes practically impossible. On the other hand, it is quite practical to express the experience and judgment in terms of a desirable output trajectory $\bar{y} \in E_y$, which should not be constrained to Y_V nor otherwise, called *reference trajectory* (also aspiration level trajectory, reference point).

Many authors -- see Wierzbicki (1979) for a review -- have considered the use of the norm $\|\bar{y} - y\|$ for generating D-maximal elements of Y_V. The most general results were obtained by Rolewicz (1975) for any Banach space E_y:

(16) $$\hat{y} \in \text{Arg} \max_{y \in Y_V} \|\bar{y} - y\| \quad , \qquad \bar{y} \in Y_{VD} \Rightarrow \hat{y} \in \hat{Y}_V$$

if $\bar{y}$ is D-dominating Y_V:

(17) $$Y_{VD} = \{\bar{y} \in E_y : \bar{y} - y \in D \text{ for all } y \in Y_V\} = \{\bar{y} \in E_y : Y_V \subset \bar{y} - D\}$$

and if the following condition is satisfied:

(18) $$D \cap (y - D) \subset B(0, \|y\|) \cup \{y\} \quad , \qquad \forall y \in E_y$$

where $B(0,\rho)$ denotes the open ball in the space E_y with radius ρ and center at 0. If E_y is Hilbert, then the condition (18) is satisfied iff

(19) $$D \subseteq D^* \quad .$$

However, the conditions (18) or (19), limiting the choice of the norm and the positive cone, are not very restrictive for applications; really restrictive is the requirement that $\bar{y}$ should be D-dominating all attainable outputs. To overcome this limitation, the notion of an *achievement scalarizing functional* has been introduced -- see, e.g., Wierzbicki (1980). An achievement scalarizing functional is a nonlinear continuous functional $s: E_y \to R^1$, with argument $y - \bar{y}$, where $y \in Y_V$ is an attainable output trajectory and $\bar{y} \in E_y$ is an arbitrary (not constrained to Y_V nor to Y_{VD}) desirable reference trajectory. An achievement scalarizing functional should, moreover, satisfy two axiomatic requirements:

(i) it should be *(quasi-) strictly order preserving*

(20) $$y_2 - y_1 \in \mathring{D} \;\; (\text{or } y_2 - y_1 \in \mathring{D}^q) \Rightarrow s(y_1 - \bar{y}) < s(y_2 - \bar{y})$$

or, if possible, *strongly order preserving*

(21) $$y_2 - y_1 \in \tilde{D} \Rightarrow s(y_1 - \bar{y}) < s(y_2 - \bar{y})$$

(ii) it should be *order representing*

(22) $$S_0 \overset{df}{=} \{y \in E_y : s(y-\bar{y}) \geq 0\} = \bar{y} + D \quad ;$$

$$s(y-\bar{y}) = 0 \quad \text{for all } y-\bar{y} \in D \backslash \mathring{D} \, (\text{or } y-\bar{y} \in D \backslash \mathring{D}^q)$$

or, at least, *order approximating* for some small $\varepsilon > 0$;

(23) $$\bar{y} + D \subset S_0 \overset{df}{=} \{y \in E_y : s(y-\bar{y}) \geq 0\} = \bar{y} + D_{\varepsilon 0} \subseteq \bar{y} + D_\varepsilon ; \; s(0) = 0$$

where the cone $D_{\varepsilon 0}$ is not necessarily of the form (13) and is a closed cone. However, in order to preserve similarity with $\tilde{D}_\varepsilon$, $\tilde{D}_{\varepsilon 0}$ is defined by $\tilde{S}_0 \overset{df}{=} \{y \in E_y : s(y-\bar{y}) > 0\} = \bar{q} + \tilde{D}_{\varepsilon 0}$. Therefore, $\tilde{D}_{\varepsilon 0}$ is an open set, and $D_{\varepsilon 0}$-maximality is equivalent to weak $D_{\varepsilon 0}$-maximality. The set $\hat{Y}_V^{\varepsilon 0} = \{\hat{y} \in Y_V : Y_V \cap (\hat{y} + \tilde{D}_{\varepsilon 0}) = \phi\}$ is understood in the above sense.

Thus, we can distinguish *strict achievement scalarizing functionals*, which satisfy the requirements (20) and (22), and *strong achievement scalarizing functionals*, which satisfy the requirements (21) and (23); the requirements (21) and (22) cannot be satisfied together. It is known that, if s is strongly order preserving, then, for any $\bar{y} \in E_y$:

(24) $$\hat{y} \in \operatorname{Arg} \max_{y \in Y_V} s(y-\bar{y}) \Rightarrow \hat{y} \in \hat{Y}_V$$

and if s is only (quasi-) strictly order preserving, then:

(25) $$\hat{y} \in \operatorname{Arg} \max_{y \in Y_V} s(y-\bar{y}) \Rightarrow \hat{y} \in \hat{Y}_V^w \; (\text{or } \hat{y} \in \hat{Y}_V^{wq}) \quad .$$

On the other hand, as shown in Wierzbicki (1980), if s is a strict achievement scalarizing functional, then

(26) $$\hat{y} \in \hat{Y}_V^w \; (\text{or } \hat{y} \in \hat{Y}^{wq}) \Rightarrow \hat{y} \in \operatorname{Arg} \max_{y \in Y_V} s(y-\hat{y}) \; , \; \max_{y \in Y_V} s(y-\hat{y}) = 0$$

and, if s is a strong achievment scalarizing functional, then

(27) $$\hat{y} \in \hat{Y}_V^{\varepsilon 0} \Rightarrow \hat{y} \in \operatorname{Arg} \max_{y \in Y_V} s(y-\hat{y}) \; , \; \max_{y \in Y_V} s(y-\hat{y}) = 0 \quad .$$

The conditions (26), (27) constitute not only necessary conditions for D-maximality even for nonconvex sets Y_V (corresponding to the separation of the sets Y_V and $\hat{y} + \mathring{D}^q$ or $\hat{y} + \tilde{D}_{\varepsilon 0}$ by the nonlinear functional s), but are also rather practical means for checking whether a given desirable $\bar{y}$ is attainable with surplus, attainable without surplus and D-maximal, or not attainable. In fact, for a strong achievement scalarizing functional s

$$
(28) \qquad \bar{y} \in (Y_V - D_{\varepsilon 0}) \backslash \hat{Y}_V^{\varepsilon 0} \Rightarrow \max_{y \in Y_V} s(y - \bar{y}) > 0
$$

$$
\bar{y} \in \hat{Y}_V^{\varepsilon 0} \subset Y_V - D_{\varepsilon 0} \Rightarrow \max_{y \in Y_V} s(y - \bar{y}) = 0
$$

$$
\bar{y} \notin Y_V - D_{\varepsilon 0} \Rightarrow \max_{y \in Y_V} s(y - \bar{y}) < 0
$$

where $(Y_V - D_{\varepsilon 0}) \backslash \hat{Y}_V^{\varepsilon 0}$ is the set of all output trajectories $D_{\varepsilon 0}$-dominated by an attainable trajectory, $Y_V - D_{\varepsilon 0} = \{\bar{y} \in E_y : \bar{y} = y - d, y \in Y_V, d \in D_{\varepsilon 0}\}$. The proof of relations (28) follows directly from the definition of $\tilde{D}_{\varepsilon 0}$ by $\tilde{S}_0 = \{y \in E_y : s(y - \bar{y}) > 0\} = \bar{q} + \tilde{D}_{\varepsilon 0}$. Similar conclusions hold for strict achievement scalarizing functionals.

Another important conclusion (see Wierzbicki 1980) from the conditions (26), (27) is *the controllability of modeling results by the user*: if, say, a strong achievement scalarizing functional is applied, then the user can obtain any $D_{\varepsilon 0}$-maximal output trajectory $\hat{y}$ as a result of maximization of $s(y - \bar{y})$ by suitably changing the reference trajectory $\bar{y}$, no matter what are other detailed properties of the functionals. Therefore, detailed properties of the functional s can be chosen in order to facilitate either computational optimization procedures, or the interaction between the user and the optimization model, or as a compromise between these two goals.

Various forms of achievement scalarizing functionals have been discussed in Wierzbicki (1980) in the case when $E_y = R^p$, together with some special forms when E_y is a Hilbert space. Here we consider in some more detail the construction of achievement scalarizing functionals in normed spaces.

A general construction of a strict achievement scalarizing functional in the case of $\mathring{D} \neq \phi$ can be obtained as follows. Suppose a value functional $v: D \quad R^1$ is given (that is, any strictly order preserving, nonnegative functional v defined for $y \in D$ -- similarly as in Debreu (1959)) and is equal zero for all $y \in D \backslash \mathring{D}$. Then:

$$
(29) \qquad s(y-\bar{y}) = \begin{cases} v(y-\bar{y}), & \text{if } y-\bar{y} \in D \\ -\rho \operatorname{dist}(y-\bar{y}, D), & \text{if } y-\bar{y} \notin D \quad ; \quad \rho > 0 \end{cases}
$$

is a strict achievement scalarizing functional. It is clearly order representing. If $y-\bar{y} \in D$, it is strictly order preserving. If $y_2 - y_1 \in \mathring{D}$, $y_2 - \bar{y} \in D$ and $y_1 - \bar{y} \notin D$, then $s(y_2 - \bar{y}) - s(y_1 - \bar{y}) > 0$ by the definition (29). If $y_2 - y_1 \in \mathring{D}$, $y_2 - \bar{y} \notin D$ and, thus, $y_1 - \bar{y} \notin D$, then denote $y_2 - y_1 = \tilde{y} \in \mathring{D}$ and observe that

$$
\operatorname{dist}(y_2-\bar{y}, D) = \min_{d \in D} \| y_1+\tilde{y}-\bar{y}-d \| = \min_{\tilde{d} \in D-\tilde{y}} \| y_1-\bar{y}-\tilde{d} \| = \operatorname{dist}(y_1-\bar{y}, D-\tilde{y}).
$$

On the other hand, since $\tilde{y} \in \mathring{D}$ and D is a convex cone, hence $D \subset \mathring{D} - \tilde{y}$. Any interior point of $D - \tilde{y}$ has a larger distance from the exterior point $y_1 - \bar{y}$ than $\operatorname{dist}(y_1 - \bar{y}, D-\tilde{y})$; hence $\operatorname{dist}(y_2-\bar{y}, D) <$ $\operatorname{dist}(y_1 - \bar{y}, D)$ and $s(y_2 - y) - s(y_1 - \bar{y}) > 0$ in all cases of $y_2 - y_1 \in \mathring{D}$, the functional (29) is strictly order preserving.

However, the functional (29) has several drawbacks. First, even if it would be possible to extend it for cases when $\mathring{D} = \phi$ and $\mathring{D}^q \neq \phi$, such an extension is not essential: in applications, weak or quasi-weak D-maximal elements of Y_V are not interesting, and much more important are D_ε-maximal elements. Moreover, the choice of a value functional with desired properties might be difficult in infinite-dimensional spaces, since the simplest value functional -- a positive linear functional -- cannot be continuously modified to zero for $y - \bar{y} \in D \backslash \mathring{D}$. Therefore, we shall relax the requirement of order representation to that of order approximation, while trying to obtain in return strong order preservation.

Choose any strongly positive linear functional $y^* \in \mathring{D}^{*q}$, of unit norm, $\| y^* \| = 1$. Then:

$$(30) \qquad s(y-\bar{y}) = \langle y^*, y-\bar{y}\rangle - \rho\,\mathrm{dist}(y-\bar{y}, D) \quad , \qquad \rho > 1$$

is a strong achievement scalarizing functional, with $\varepsilon > \frac{1}{\rho}$. In fact, $\langle y^*, y-\bar{y}\rangle$ is strongly order preserving, due to the definition $\overset{\circ}{D}{}^{*q} = \{y^* \in E_y^*: \langle y^*, y\rangle > 0 \ \forall y \in \tilde{D}\}$. The functional $-\mathrm{dist}(y-\bar{y}, D)$ is order preserving (neither strongly nor strictly), by an argument similar to the analysis of the functional (29). However, the sum of a strongly order preserving and an order preserving functional is, clearly, strongly order preserving. Moreover, by the definition of the norm in the dual space, $\langle y^*, y-\bar{y}\rangle \leq \|y-\bar{y}\|$ if $\|y^*\| = 1$. If, additionally, $y \in S_0 = \{y \in E_y : s(y-\bar{y}) \geq 0\}$, then $\rho\,\mathrm{dist}(y-\bar{y}, D) \leq \langle y^*, y-\bar{y}\rangle \leq \|y-\bar{y}\|$; hence $S_0 \subset \bar{y} + D_\varepsilon$ for $\varepsilon > \frac{1}{\rho}$. Clearly, $\bar{y} + D \subset S_0$ and $s(0) = 0$; thus the functional (30) is order approximating.

The functional (30) has also some drawbacks in applications. First, the choice of y^* is arbitrary; however, it does not much influence the applicability of the functional (30), particularly if $\rho >> 1$, since $\bar{y}$ is very often chosen as not attainable. Thus, any reasonable y^* -- for example, corresponding to equal weights for all components of output trajectories and all instants of time -- might be chosen; according to the controllability conclusion, this does not restrict the possibility of influencing the resulting D_ε-maximal output trajectories $\hat{y}$ by changing the reference trajectories $\bar{y}$. Second, the functional (30) is nondifferentiable. Although recent development of nondifferentiable optimization algorithms is remarkable, not all of these algorithms are directly applicable for dynamic optimization. Therefore, it might be useful to consider also achievement scalarizing functionals that are differentiable.

Observe that achievement scalarizing functionals are constructed by using a strictly or strongly order preserving functional of value functional type and supplementing it by a term expressing a distance from $y-\bar{y}$ to the cone D. While the first part can be chosen to be differentiable, it is the second part that introduces nondifferentiability. To facilitate computation and differentiation of functionals related to the distance, suppose E_y is a Hilbert space. Then, due to the Moreau theorem (1962; see Wierzbicki and Kurcyusz, 1977), the following holds:

$$(31) \qquad \mathrm{dist}(y-\bar{y},D) = \|(y-\bar{y})^{-D^*}\| = \|(\bar{y}-y)^{D^*}\|$$

when $(\cdot)^{-D^*}$ or $(\cdot)^{D^*}$ denotes the operation of projection on the cone $-D^*$ or D^*. Moreover, $\|(\bar{y}-y)^{D^*}\|^2$ is differentiable in y and its derivative is precisely $-(\bar{y}-y)^{D^*}$. Thus, if E_y is Hilbert, a differentiable modification of (30) is as follows:

$$(32) \qquad s(y-\bar{y}) = \langle y^*, y-\bar{y}\rangle - \frac{1}{2}\rho\|(\bar{y}-y)^{D^*}\|^2 \quad , \qquad \rho > 0 \quad .$$

This functional is strongly order preserving, by the same argument as in the analysis of (30), and its maximal points are D-maximal for any $\rho > 0$. However, the functional (32) is not order approximating and, if $\bar{y} = \hat{y}$ is D_ε-maximal, then the maximal points of (32) will generally not coincide with $\hat{y}$ for any $\rho > 0$. On the other hand, if ρ is sufficiently large, the maximal points of (32) usually approximate quite closely the maximal points of (30), and the requirement of order approximation does not play a decisive role. Thus, the functional (32) for sufficiently large ρ might have useful applications.

If E_y is Hilbert, then there is also a technically differentiable form of a strong achievement scalarizing functional, satisfying both (21) and (23):

$$(33) \qquad s(y-\bar{y}) = \frac{1}{2}\|y-\bar{y}\|^2 - \frac{1}{2}\rho\|(\bar{y}-y)^{D^*}\|^2 \quad ; \quad \rho > 1, \quad D \subseteq D^*$$

with $\varepsilon > \rho^{-\frac{1}{2}}$, see Wierzbicki (1977a). In (33), the role of a value functional for $y-\bar{y} \in D$ plays the (square) norm; hence the condition $D \subseteq D^*$, equivalent to the Rolewicz condition (18), is necessary for the strong order preservation property. If $y-\bar{y} \notin D$, the (square) norm is modified by the (square) distance term; if $\rho > 1$, this modification is sufficiently strong to imply strong order preservation. The property of order approximation results immediately from the form of (33).

Consider, however, a functional similar to (33):

$$(34) \qquad s(y-\bar{y}) = \|y-\bar{y}\| - \rho\|(\bar{y}-y)^{D^*}\| \quad ; \quad \rho > 1, \quad D \subseteq D^* \quad .$$

It is also a strong achievement scalarizing functional. It is clearly order approximating with $\varepsilon > \rho^{-1}$. Moreover, due to the Moreau theorem, $s(y-\bar{y}) = (a^2(y) + b^2(y))^{\frac{1}{2}} - \rho b(y)$, where $a(y) = \|(y-\bar{y})^D\|$, $b(y) = \|(y-\bar{y})^{-D^*}\|$. The operation of projection on cones, $(\cdot)^D$ or $(\cdot)^{-D^*}$, has the property (see Wierzbicki and Kurcyusz, 1977) that $\|(y-\bar{y}+\tilde{y})^{-D^*}\| \leq \|(y-\bar{y})^{-D^*}\|$ for all $\tilde{y} \in D$ and $\|(y-\bar{y}+\tilde{y})^D\| \geq \|(y-\bar{y})^D\|$ for all $\tilde{y} \in D^*$, hence also for $\tilde{y} \in D$ if $D \subseteq D^*$. Thus, if $y_2 - y_1 \in D$, then $a(y_2) \geq a(y_1)$ and $b(y_2) \leq b(y_1)$. Since $a(y_2) = a(y_1)$ and $b(y_2) = b(y_1)$ imply together $y_2 = y_1$, hence, if $y_2 - y_1 \in \tilde{D}$, we can have either $a(y_2) > a(y_1)$ and $b(y_2) \leq b(y_1)$ or $a(y_2) \geq a(y_1)$ and $b(y_2) < b(y_1)$. Now, consider the function $\psi(a,b) = (a^2 + b^2)^{\frac{1}{2}} - \rho b$. This function is clearly strictly increasing in respect to a. Since $\frac{\partial\psi}{\partial b}(a,b) = b(a^2 + b^2)^{-\frac{1}{2}} - \rho < 0$ for $\rho > 1$, the function ψ is strictly decreasing in respect to b. Therefore, if $y_2 - y_1 \in \tilde{D}$, then $s(y_2 - \bar{y}_1) - s(y_1 - \bar{y}) > 0$, and the functional (34) is strongly order preserving.

On the other hand, after a suitable choice of (different) values of ρ in (33) and (34), the level set $S_0 = \{y \in E_y : s(y-\bar{y} \geq 0\}$ can be made identical for these two functionals, and this level set has necessarily a corner point at $y = \bar{y}$. Thus, the differentiability of (33) has only technical character, and an essential nondifferentiability in terms of corner points of level sets is necessarily related to strong and strict achievement scalarizing functionals. Therefore, for computational purposes, it is useful to introduce another class of *approximate scalarizing functionals*. The approximate scalarizing functionals are supposed to have strong order preservation property (21), which implies that their maximal points are D-maximal. However, the requirement of order approximation (23) is further related by substituting D_ε, the conical ε-neighborhood of D, by another form of an ε-neighborhood:

$$(35) \qquad D_{\varepsilon\gamma} = \{y \in E_y : \text{dist}(y,D) < \varepsilon\gamma(\|y\|)\}$$

where $\gamma(\cdot)$ is any given strictly increasing function. For example, it is easy to check that (32) is an approximate scalarizing functional, with $\gamma(\|y\|) = \|y\|^{\frac{1}{2}}$. Approximate scalarizing functionals are not strictly applicable for checking D-maximality of a given

$\hat{y}$ via condition (27), nor attainability of a given $\bar{y}$ via conditions (28), since a maximum point of an approximate scalarizing functional might be different from a given D-maximal $\hat{y} = \bar{y}$. However, the set $D_{\varepsilon\gamma}$ approximates the cone D sufficiently closely for small ε, and the difference between its maximal point and a given D-maximal $\hat{y} = \bar{y}$ can be made very small. Thus, for practical purposes, approximate scalarizing functionals have all the advantages of strong scalarizing functionals.

To illustrate further the distinction between strong and approximate scalarizing functionals, consider still another variant of such functionals. Suppose we have, originally, a single-objective optimization problem with a performance functional:

(36) $$y^0 = Y^0(X(u),u) \in R^1 \quad .$$

Suppose that, after maximizing this functional and observing, for example, that there are many controls u and states x that result in nearly the same value of y^0 (a frequent case of practical nonuniqueness of solutions), we decided to supplement this performance functional with other objectives, stated in terms of a desirable shape of output trajectories:

(37) $$y^r = Y^r(X(u),u) \in E_y^r$$

where E_y^r is a normed space, with a positive cone D^r. After defining $y = (y^0,y^r)$, $E_y = R^1 \times E_y^r$ and $D = R_+^1 \times D^r$ we bring the problem back to the previous formulation, and any of the scalarizing functionals defined above can be used. However, this specific case suggests also a specific form of a strong scalarizing functional:

(38) $$s(y-\bar{y}) = y^0 - \bar{y}^0 - \rho\,\mathrm{dist}(y^r - \bar{y}, D^r) \quad ; \quad \rho > 0 \quad .$$

It is easy to check that this functional is order approximating with $\varepsilon > 1/\rho$. Moreover, it is strongly order preserving in a modified sense, with $\tilde{D} = (R_+^1 \times D^r)\setminus(\{0\} \times (D^r \cap -D^r))$ replaced by

$\tilde{\tilde{D}} = (R^1_+ \backslash \{0\}) \times (D^r \backslash (D^r \cap -D^r)) = \tilde{R}^1_+ \times \tilde{D}^r$. This modified sense of strong partial preordering results in modified D-maximal points that might be weakly D^r-maximal, in the second component y^r, but are always strongly maximal in the first component y^0. In fact, if $y_2 - y_1 \in \tilde{\tilde{D}}$, then $y_2^0 > y_1^0$ and $y_2^r - y_1^r \in \tilde{D}^r$. Since the functional $-\mathrm{dist}(y^r - \bar{y}^r, D^r)$ is (neither strictly nor strongly) order preserving, the first term in (38) guarantees that $s(y_2 - \bar{y}) > s(y_1 - \bar{y})$ for $y_2 - y_1 \in \tilde{\tilde{D}}$.

Suppose E^r_y is Hilbert and consider the following approximate scalarizing functional

$$(39) \qquad s(y-\bar{y}) = y^0 - \bar{y}^0 - \tfrac{1}{2}\rho\| (\bar{y}^r - y^r)^{D^{r*}} \|^2 \quad ; \quad \rho > 0$$

By a similar argument, this functional is strongly order preserving with $\tilde{\tilde{D}}$ replacing $\tilde{D}$. It is not order approximating, only γ-order approximating with $D_{\varepsilon\gamma}$ defined as in (35) and $\gamma(\|y\|) = \|y\|^{\frac{1}{2}}$.

Observe that the functionals (38), (39) correspond to one of the classical, widely used approaches to multiobjective optimization. In this approach, we choose one of the objectives -- say, y^0 -- to be maximized and represent other objectives -- say, y^r -- by parametrically changing constraints, $y^r - \bar{y}^r \in D^r$. The functionals (38), (39) represent, respectively, an exact and an exterior quadratic penalty functional for such a formulation. However, it is not widely known that, when using such penalty functionals, one does not have to increase ρ to infinity or otherwise iterate (e.g., introduce shifts) on penalty functionals. Since these functionals are (modified) strongly order preserving, each maximal point of them is (modified) D-maximal, no matter what $\rho > 0$ has been chosen and what are the actual violations $(\bar{y} - y^r)^{D^{r*}}$ of the constraints $y^r - \bar{y}^r \in D^r$, treated here as a type of soft constraints. This feature of the scalarizing functionals (38), (39) is particularly useful for dynamic optimization with trajectory constraints (taking a form, for example, of state constraints), since the iterations on penalty functions might be particularly cumbersome in such a case. While using functions (38), (39) for multiobjective trajectory optimization, it is sufficient to choose a reasonable value of $\rho > 0$ and to maximize (38) or (39) once in order to obtain a (modified) D-maximal alternative solution corresponding to a desirable shape $\bar{y}^r$ of output trajectory y^r.

Via penalty functions, functionals (38), (39) -- and, in fact, all other achievement scalarizing functionals -- are related to two other basic notions in mathematical optimization and modelling: those of generalized Lagrangian functionals and of regularization of solutions of ill-posed problems.

3. RELATIONS TO GENERALIZED LAGRANGIAN FUNCTIONALS

Consider the classical form of a mathematical programming problem with generalized inequalities:

$$(40) \qquad \underset{u \in U_0}{\text{minimize}}\ f^0(u) \quad ; \quad U_0 = \{u \in E_u : g(u) \in -D \subset E_g\}$$

where $f^0: E_u \to R'$, $g: E_u \to E_g$, D is a positive cone in E_g. Suppose E_x is a Banach space and E_g is a Hilbert space. Under various forms of regularity conditions -- see, e.g., Kurcyusz (1974) -- the necessary conditions for $\hat{u}$ being an optimal solution to this problem can be expressed via the well-known normal Lagrangian functional

$$(41) \qquad L(\eta,u) = f^0(u) + \langle \eta, g(u) \rangle$$

and take the known form

$$(42) \qquad L_u(\hat{\eta},\hat{u}) = f^0_u(\hat{u}) + g^*_u(\hat{u})\hat{\eta} = 0$$

where $g^*_u(\hat{u})$ is the adjoint operator to $g_u(\hat{u})$, and

$$(43) \qquad g(\hat{u}) \in -D \quad ; \quad \langle \hat{\eta}, g(\hat{u}) \rangle = 0 \quad ; \quad \hat{\eta} \in D^*$$

where $\hat{\eta} \in E^*_g$ is a normal Lagrange multiplier related to the solution $\hat{u}$. The triple condition (43) might be referred to as Kuhn-Tucher complementarity triple, widely known. However, it is not widely known that complementarity triple (43) is, in fact, equivalent to a single nonlinear equation for $\hat{\eta}$ (although this result has been, in fact, used in R^n by Rockafellar (1974), in a Hilbert space by Wierzbicki and Kurcyusz (1977) and independently proven in R^n by Mangasarian (1976)).

To show this in the case when E_g is a Hilbert space, we use the Moreau (1962) theorem: for any closed convex cone $D \subset E_g$ and any $p \in E_g$, $p_1 = (p)^{-D}$ and $p_2 = (p)^{D*}$ are the projections of p on the cones -D, D*, respectively, if and only if

$$(44) \qquad p_1 + p_2 = p, \quad p_1 \in -D, \quad \langle p_2, p_1 \rangle = 0, \quad p_2 \in D^* \quad .$$

Thus, denote $g(\hat{u}) + \hat{\eta} = p$; it is easy to check then that (43) holds if and only if

$$(45) \qquad (g(\hat{u}) + \hat{\eta})^{D*} = \hat{\eta}$$

or, equivalently, iff $(g(\hat{u}) + \hat{\eta})^{-D} = g(\hat{u})$ (one of these equations suffices and the other is redundant because of the definition $g(\hat{u}) + \hat{\eta} = p$.)

This basic fact has various consequences. For example, the sensitivity analysis of solutions of (40) might be based on appropriate implicit function theorems instead of analyzing the sensitivity of a system of inequalities, which is now the typical approach to this question -- see, e.g., Robinson (1976). Another important conclusion from equation (45) is that there are modified Lagrangian functionals that should possess an *unconstrained saddle point* in η, u at $\hat{\eta}, \hat{u}$. In fact, these are augmented Lagrangian functionals as introduced by Hesteness (1969) for problems with equality constraints in R^n, by Rockafellar (1974) for problems with inequality constraints in R^n, by Wierzbicki and Kurcyusz (1977) for problems with inequality constraints in a Hilbert space, and studied by many other authors. For problem (40), the augmented Lagrangian functional takes the form

$$(46) \qquad \Lambda(\eta,u,\rho) = f^0(u) + \tfrac{1}{2}\rho\|(g(u) + \tfrac{\eta}{\rho})^{D*}\|^2 - \tfrac{1}{2}\rho\|\tfrac{\eta}{\rho}\|^2 \quad , \ \rho > 0$$

and the first-order necessary conditions (42), (43) $\Longleftrightarrow$ (45) take the form

(47) $$\Lambda_u(\eta,u,\rho) = f_u^0(\hat{u}) + g_u^*(\hat{u})(\rho g(\hat{u}) + \hat{\eta})^{D^*} = 0$$

(48) $$\Lambda_\eta(\hat{\eta},\hat{u},\rho) = \frac{1}{\rho}((\rho,g(\hat{u}) + \hat{\eta})^{D^*} - \hat{\eta}) = 0$$

Other necessary and sufficient conditions for optimality of $\hat{u}$ in terms of saddle-points of (46) are given in Wierzbicki and Kurcyusz (1977).

Consider now the following specification of problem (40), taking into account (36), (37)

(49) $$f^0(u) = -Y^0(X(u),u) \quad ; \quad g(u) = \bar{y}^r - Y^r(X(u),u) \in -D^r$$

where u might be additionally constrained explicitly by $u \in V$. Consider the augmented Lagrangian functional (46) with $\eta = 0$:

(50) $$\Lambda(0,u,\rho) = -Y^0(X(u),u) + \tfrac{1}{2}\rho\|(\bar{y}^r - Y^r(X(u),u))^{D^{r*}}\|^2 =$$

$$= -s(Y(X(u),u)-\bar{y}) - \bar{y}^0$$

with $s(y-\bar{y})$ defined as in (39). The order-preservation properties of the approximate scalarizing functional (39) can be now interpreted as follows. Even if we fix $\eta = 0$ and admit violations of the constraint $\bar{y}^r - Y^r(X(u),u) \in -D^r$, and even under additional constraints $u \in V$, any minimal point of the augmented Lagrangian functional (50) is a D-maximal point of the set $Y_V = Y(X(V),V) = Y^0(X(V),V) \times Y^r(X(V),V)$ in the sense of the strong partial preordering induced by the cone $\tilde{\tilde{D}} = \tilde{R}_+^1 \times \tilde{D}^r$. Moreover, since:

(51) $$\Lambda(\eta,u,\rho) = -s(Y(X(u),u) - \bar{y} - \frac{\eta}{\rho}) - \bar{y}^0 - \tfrac{1}{2}\rho\|\frac{\eta}{\rho}\|^2$$

and the above conclusion holds independently of $\bar{y}$, hence it also holds for any fixed η. Thus, the conclusion can be considered as another generalization of Everett's theorem (196) and the reference trajectory $\bar{y}$ is, in a sense, related to the generalized Lagrange multiplier η.

However, the last analogy should not be taken too mechanistically. For example, the properties (28) of a strong scalarizing functional can be rewritten as

$$\min_{\bar{y}\in Y_V} \max_{u\in V} s(Y(X(u),u) - \bar{y}) = 0 \tag{52}$$

and the min-max points $(\hat{\bar{y}},\hat{u})$ correspond to D-maximal points of the set $Y_V = Y(X(V),V)$. On the other hand, (52) is not a saddle-point property, since $s(y-\bar{y})$ is not convex in $\bar{y}$, and it is easy to show examples such that $\max_{u\in V} \min_{\bar{y}\in Y_V} s(Y(X(u),u)-\bar{y}) < 0$. In order to obtain saddle-point properties, convexifying terms in η would have to be added to $s(Y(X(u),u) - \bar{y})$, as it was done in (51).

4. MULTIOBJECTIVE TRAJECTORY OPTIMIZATION AS SEMIREGULARIZATION OF MODEL SOLUTIONS

The monography of Tikkonov and Arsenin (1977) summarizes an extensive research on one of the basic problems of mathematical modeling -- that of regularization of solutions of ill-posed problems. Many results of this research relate to the usefulness of using distance functionals when solving problems with non-unique solutions or quasi-solutions (generalized solutions). The nonuniqueness of solutions of a mathematical model implies usually that the solutions would change discontinuously with small changes of parameters of the model. For example, if a dynamic linear programming model has practically nonunique solutions, that is, if there is one optimal basic solution but many other basic solutions result in almost the same value of the objective function, then a small change of parameters of the model results in large changes of the solution -- see Avenhaus (1980). The regularization of solutions of such a type of models consists then in choosing from experience a *reference solution* and considering the solution of the model that is closest to the reference solution in a chosen sense of distance; as proven by Tikkanov and Arsenin, this results not only in the selection of a solution, but also in continuous dependence of the selected solution on parameters of the model.

The regularization method can be illustrated as follows. Suppose a mathematical programming problem consists in minimizing the functional

$$f^0(u) = -Y^0(X(u),u) \tag{53}$$

for $u \in V$. Suppose the solutions of this problem are (possibly only practically) nonunique. Let a reference trajectory $\bar{y}^r$ be given in a normed space E_y^r of the outputs of the model, $y^r = Y^r(X(u),u)$. By a *normal solution* of the problem of minimizing $f^0(u)$ for $u \in V$ we define such a solution of this problem that minimizes, additionally, $\|\bar{y}^r - Y^r(X(u),u)\|$. This normalization is, clearly, relative to the output space E_y^r. However, it is easy to see that if, say, $f^0(u)$ and V are convex, Y and X are linear, and the unit ball in E_y^r is strongly convex, then the normal solution is unique relative to the output space -- that is, it determines uniquely the output trajectory y^r. Moreover, this output trajectory depends continuously on the reference trajectory $\bar{y}^r$. A stable computational method of determining the normal solution approximately consists in minimizing the functional:

$$\Phi(\bar{y}^r,u,\rho) = -Y^0(X(u),u) + \tfrac{1}{2}\rho\|\bar{y}^r - Y^r(X(u),u)\|^2 \tag{54}$$

for $\rho \to 0$. Again, under appropriate assumptions, it can be shown that output trajectories corresponding to minimal points of (54) converge to the output trajectory corresponding to the normal solution as $\rho \to 0$.

However, observe that (54) can be obtained from (50) if E_y^r is Hilbert and $D^r = \{0\}$, $D^{r*} = E_y^r$. Thus, the multiobjective trajectory optimization is strongly related to model regularization. Actually, the former can be considered as a generalization of the latter. In fact, define *semi-normal* solutions of the problem of minimizing $f^0(u)$ for $u \in V$ as such that minimize, additionally, $\text{dist}(Y^r(X(u),u),\bar{y}^r + D^r)$, where D^r is a positive cone in the space of output trajectories E_y^r. Now, even if f(u) and V were convex and Y and X linear, the output trajectory y^r corresponding to a semi-normal solution need not be unique -- since there might be many points in a convex set that are equidistant to a convex cone. However, the semi-normal solutions have good practical interpretation; the corresponding output trajectories are either close to or better than the desired reference trajectory $\bar{y}^r$, depending on its attainability. Moreover, when minimizing the functional (50), instead of (54), we obtain D-maximal points of the set $Y_V = Y(X(V),V) = Y^0(X(V),V) \times Y^r(X(V),V)$ for each $\rho > 0$. The same applies, clearly, to the functional (54), if we assume $D^r = \{0\}$, which gives another interpretation of regularization techniques. Thus, multiobjective trajectory optimization is a type of *model semiregularization technique:* for the selection of a solution of the model, a reference output trajectory is used together with a notion of a partial preordering of the output space.

5. COMPUTATIONAL ISSUES AND APPLICATIONS: A DIFFERENTIABLE TIME-CONTINUOUS CASE

If an achievement scalarizing functional is differentiable, then any method of dynamic optimization can be applied as a tool for obtaining an attainable, D-maximal trajectory $\hat{y}$ in response to a desirable trajectory $\bar{y}$. An efficient class of dynamic optimization techniques applicable in this case are gradient trajectory techniques, or *control space gradient techniques*, based on a reduction of the gradient of the minimized functional to control space. A general method for such a gradient reduction, independent on the particular type of the state equation, is described, for example, in Wierzbicki (1977b). Here we present only the simplest and well-known case of gradient reduction for problems with ordinary differential state equations.

As an example, consider the approximate scalarizing functional (39) and suppose y^0 is described by

$$(55) \qquad y^0 = \int_{t_0}^{t_1} F^0(x(t),u(t),t)\,dt + F^1(x(t_1)) \quad .$$

Moreover, assume the mapping X be given by solutions of the state equation

$$(56) \qquad x = X(u) \Longleftrightarrow \dot{x}(t) = F(x(t),u(t),t) \quad ; \quad x(t_0) = x_0 \in R^n$$

and the mapping Y^r -- by the output equation

$$(57) \qquad y^r = Y^r(X(u),u) \Longleftrightarrow y^r(t) = G^r(x(t),u(t),t) \in R^p .$$

Take $E^r_y = L^2([t_0;t_1],R^p)$ and $D^r = L^2_+([t_0;t_1],R^p)$; then

$$(58) \quad s(Y(X(u),u)-\bar{y}) = \int_{t_0}^{t_1} G^0(x(t),u(t),\bar{y}^r(t),t)dt + F^1(x(t_1)) - \bar{y}^0$$

where

$$(59) \quad G^0(x(t),u(t),\bar{y}^r(t),t) = F^0(x(t),u(t),t) - \tfrac{1}{2}\rho \sum_{i=1}^{p} (\bar{y}^{ri}(t) - G^{ri}(x(t),u(t),t))_+^2$$

and $(\tilde{y}^i)_+ = \max(0,\tilde{y}^i)$ for $\tilde{y}^i \in R^1$. By choosing $D^r = L^2_+([t_0;t_1],R^p)$ we assumed that all outputs improve as the corresponding values $y^{ri}(t)$ increase for (almost) all $t \in [t_0;t_1]$. Now, a reference output trajectory $\bar{y}^r(t) = (\bar{y}^{r1}(t),\ldots,\bar{y}^{ri}(t),\ldots,\bar{y}^{rp}(t))$ for $t \in [t_0;t_1]$ is assumed to be given by the model user. In fact, if p is not too large -- say, 3 or 4 -- the user can easily draw the number p of curves representing output trajectories desired by him. Moreover, experiments show that he is also able to evaluate easily the corresponding responses of the optimization model, $\hat{y}^0$ and $\hat{y}^r(t) = (\hat{y}^{r1}(t),\ldots,\hat{y}^{ri}(t),\ldots,\hat{y}^{rp}(t))$ for $t \in [t_0;t_1]$ and, if he does not like them, to change the reference trajectory in order to obtain new responses. Observe that the reference value $\bar{y}^0$ plays, in this case, a technical role and can be omitted. Thus, an interactive multiobjective dynamic optimization procedure can be organized, provided we could supply an efficient technique of maximizing the functional (57) subject to the state equation (56) and, possibly, other constraints. To simplify the presentation, suppose other constraints are already expressed as penalty terms in the functions F^0 or F^1.

Denote $S(u) = s(Y(X(u),u) - \bar{y})$. Then $S_u(u)$, the gradient of the functional (58) reduced to the control space, can be computed in the following way. The Hamiltonian function for the problem of maximizing (58) subject to (56) has the form

$$(60) \qquad H(\Psi(t),x(t),u(t),\bar{y}^r(t),t) = G^0(x(t),u(t),\bar{y}^r(t),t) + \Psi(t)F(x(t),u(t),t)$$

where $\Psi(t)F(x(t),u(t),t)$ is a short denotation for scalar product in R^n and $\Psi(t)$ is the costate (the adjoint variable for the state). To compute $S_u(u)(t)$ for $t \in [t_0;t_1]$, given $u(t)$ for $t \in [t_0;t_1]$, we first determine $x(t) = X(u)(t)$ by solving (56), written equivalently as

$$(61) \qquad \dot{x}(t) = H_\Psi(\Psi(t),x(t),u(t),\bar{y}^r(t),t) \quad ; \qquad x(t_0) = x_0 \quad .$$

Then the costate $\Psi(t)$ is determined for $t \in [t_0;t_1]$ by solving, in the reverse direction of time, the adjoint equation

$$(62) \qquad \dot{\Psi}(t) = -H_x(\Psi(t),x(t),u(t),\bar{y}^r(t),t) \quad ; \quad \Psi(t_1) = F_x^1(x(t_1))$$

and the reduced gradient in the control space is determined by

$$(63) \qquad S_u(u)(t) = H_u(\Psi(t),x(t),u(t),\bar{y}^r(t),t) \quad .$$

Typical conjugate directions algorithms of nonlinear programming can be adapted for making use of this reduced gradient. However, Fortuna (1974) has shown that, for dynamic optimization, conjugate directions perform much better if a modified reduced gradient is being used:

$$(64) \qquad \tilde{S}_u(u)(t) = -H_{uu}^{-1}(\Psi(t),x(t),u(t),\bar{y}^r(t),t)H_u(\Psi(t),x(t),u(t),\bar{y}^r(t),t) \quad .$$

This modification removes possible ill-conditioning of the algebraic part of the Hessian operator $S_{uu}(u)$, leaving only possible ill-conditioning of the compact part of this operator -- and the compact part has, in the limit, negligible influence on the convergence of conjugate direction algorithms in a Hilbert space. This abstract reasoning has been also confirmed by extensive computational tests.

Now, each continuous-time dynamic optimization problem, when solved on a digital computer, is ultimately discretized over time. While a discussion of results of recent world-wide extensive research on approximations of time-continuous optimization problems is beyond the scope of the paper, it is worthwhile to note some comments on this issue.

A conscientious approach to discretization of a time-continuous problem should start with the question whether time-continuity is really an essential aspect of the analyzed model. In many cases, time-continuity is assumed only for analytical convenience, and the actual model can be better built, parameter-fitted and validated in its time-discrete version. In such cases of a priori discretization, it is certainly better to abandon time-continuity at the very beginning and to develop the time-discrete versions, say, of the equations (55) ... (64). Some qualitative properties and conclusions from the time-continuous analysis might be still applied to time-discrete models; for example, the Fortuna modification of the reduced gradient, although motivated strictly for the time-continuous case only, gives good results also in the time-discrete case.

In rather special cases, time-continuity is essential. These cases are really hard, and great care should be devoted to the analysis of those qualitative properties of the optimization problem that make time-continuity essential (such as boundary-layer effects, appearance of relaxed controls, etc.). These qualitative properties should be taken into account when looking for alternative formulations of the problem, for an appropriate space of control functions, when choosing finite-dimensional bases for a sequence of subspaces approximating the control space, when determining what is the reduced gradient expressed in terms of a finite-dimensional basis. A naive discretization of equations (61) ... (64) can lead to serious errors, when, say, a naively discretized gradient equation (63) produces numbers that are in no correspondence to the gradient that would be consistent with a chosen discretization of the control space.

We close this section with a simple example, when the continuity of time is important only because it facilitates almost fully the analytical solution. Although it does not illustrate computational issues, the example illuminates some other important aspects of applications of multiobjective trajectory optimization.

Consider a simple model of relations between inflation and unemployment, as analyzed by Snower and Wierzbicki (1980) when comparing various economic policies. The inflation rate, $x(t)$, is influenced by monetary policies, that influence also the unemployment, $u(t)$. An adaptive price expectation mechanism and a linearized Phillips curve result in the following equation:

(65) $$\dot{x}(t) = rd(b - u(t)) \quad ; \quad x(0) = x_0$$

where unemployment $u(t)$ is taken as a dummy control variable, b is a parameter of the linearized Phillips curve, rd is a composite coefficient. The social welfare function related to inflation and unemployment is assumed in the form:

(66) $$U(x(t),u(t)) = 1 - \tfrac{1}{2}x^2(t) - \frac{q}{2}u^2(t)$$

where q is the weight attached to unemployment as compared to inflation. The intertemporal social welfare functional is assumed in the form

(67) $$W(x,u) = \int_0^{\infty} e^{-rt}\, U(x(t),u(t))dt \quad .$$

The problem of maximizing (67) subject to (65) can be easily solved analytically to obtain:

(68) $$\hat{u}(t) = \frac{\alpha_0}{d}\left(x_0 - \frac{bq}{d}\right)e^{-r\alpha_0 t} + b$$

(69) $$\hat{x}(t) = \left(x_0 - \frac{bq}{d}\right)e^{-r\alpha_0 t} + \frac{bq}{d}$$

where

(70) $$\alpha_0 = \tfrac{1}{2}\left(\left(1 + 4\frac{d^2}{d}\right)^{\frac{1}{2}} - 1\right) \quad .$$

However, if the initial inflation rate x_0 is high, the 'optimal' unemployment $\hat{u}(t)$ that results from this model for small t might be considered socially undesirable, too high. We could change the model by adding simply a constraint $u(t) \leq \bar{u}$. In this simple case, the constraining value $\bar{u}$ must be greater than b; otherwise, equation (65) would result in uncontrolled, increasing inflation. However, in more complicated models, it might be difficult to judge whether a control constraint is not too stringent. Therefore, it is reasonable to treat $\bar{u}$ as a desirable bound for trajectory rather than as a fixed constraint, and to formulate a multiobjective trajectory optimization problem: maximize the social welfare functional while, at the same time, trying to keep the unemployment smaller than $\bar{u}$.

Observe that, in this formulation, one of the outputs y^r of the model is just the input control u. However, such situations are quite frequent, when some important control variables appear directly as output variables in multiobjective trajectory optimization. Moreover, the unemployment u(t) is here only a dummy control variable; actually, the model should be controlled by a monetary policy that, after a transformation that was not included is the model for simplicity results in the unemployment u(t).

Suppose we apply the approximate achievement scalarizing functional (39) for this multiobjective trajectory optimization problem and choose the norm $\|u\|^2 = \int_0^\infty e^{-rt} u^2(t)dt$ for the control space. Then:

$$s(Y(x,u) - \bar{y}) = \int_0^\infty e^{-rt}(U(x(t),u(t)) - \tfrac{1}{2}\rho(u(t) - \bar{u})_t^2)dt - \bar{y}^0 \qquad (71)$$

Suppose $u(t) > \bar{u}$ for $t \in [0;t_1)$, $u(t_1) = \bar{u}$, $u(t) < \bar{u}$ for $t \in (t_1; +\infty)$. Then (71) transforms to

$$(72)\qquad s(Y(x,u)-\bar{y}) = \int_0^{t_1} e^{-rt}(U(x(t)) - \tfrac{1}{2}\rho(u(t)-\bar{u})^2)dt +$$

$$+ e^{-rt_1}\hat{W}(x(t_1)) - \bar{y}^0$$

where

$$(73)\quad \hat{W}(x(t_1)) = \frac{1}{r}\left(1 - \frac{1}{2(1+\alpha_0)}x^2(t_1) - \frac{bd}{(1+\alpha_0)^2}x(t_1) - \right.$$

$$\left. - \frac{(2+\alpha_0)\alpha_0\, qb^2}{2(1+\alpha_0)^2}\right)$$

is the minimal value of (67) depending on the initial state. The problem of minimizing (72) subject to (65) can be solved almost fully analytically to obtain:

$$(74)\quad \hat{u}(t) = \begin{cases} -\frac{rd}{q+\rho}\left(Ae^{-r\alpha_1 t} + Be^{-r\alpha_2 t}\right) + b, & t\in[0;t_1) \\ (\bar{u}-b)\, e^{-r\alpha_0(t-t_1)} + b, & t\in[t_1;+\infty) \end{cases}$$

$$(75)\quad \hat{x}(t) = \begin{cases} r\alpha_2 A e^{-r\alpha_1 t} + r\alpha_1 Be^{-r\alpha_2 t} - \frac{\rho}{d}(\bar{u}-b) + \frac{bq}{d}, & t\in[0;t_1) \\ \frac{d}{\alpha_0}(\bar{u}-b)e^{-r\alpha_0(t-t_1)} + \frac{bq}{d}, & t\in[t_1;+\infty) \end{cases}$$

where

$$(76)\qquad \alpha_1 = \tfrac{1}{2}((1+4\frac{d^2}{q+\rho})^{\frac{1}{2}} - 1);\ \alpha_2 = -\tfrac{1}{2}((1+4\frac{d^2}{q+\rho})^{\frac{1}{2}} + 1)$$

while the constants A,B and the time instant t, result from three conditions: the continuity of $\hat{u}(t)$ (implied by continuity of adjoint variable) and of $\hat{x}(t)$ at t_1 and the initial state $x_0=\hat{x}(0)$.

For example, the former two conditions determine A,B as functions of t_1:

(77) $$A = A(t_1) = e^{r\alpha_1 t_1} \cdot \frac{\bar{u}-b}{rd} \cdot \frac{q\alpha_0 - (q+\rho)\alpha_2}{\alpha_2 - \alpha_1} \; ;$$

$$B = B(t_1) = e^{r\alpha_2 t_1} \cdot \frac{\bar{u}-b}{rd} \cdot \frac{q\alpha_0 - (q+\rho)\alpha_1}{\alpha_1 - \alpha_2}$$

while the latter condition results in the following equation for t_1 that does not admit analytical solutions (must be solved numerically)

(78) $$r\alpha_2 A(t_1) + r\alpha_1 B(t_1) - \frac{\rho}{d}(\bar{u}-b) + \frac{bq}{d} = x_0 \; .$$

Nevertheless, (74),... (78) admit on easy interpretation of the influence of ρ and $\bar{u}$ on $\hat{u}(t)$ and $\hat{x}(t)$. The single-criterium solutions (68),(69) are compared with an example of solutions (74),(75) in Fig.1.

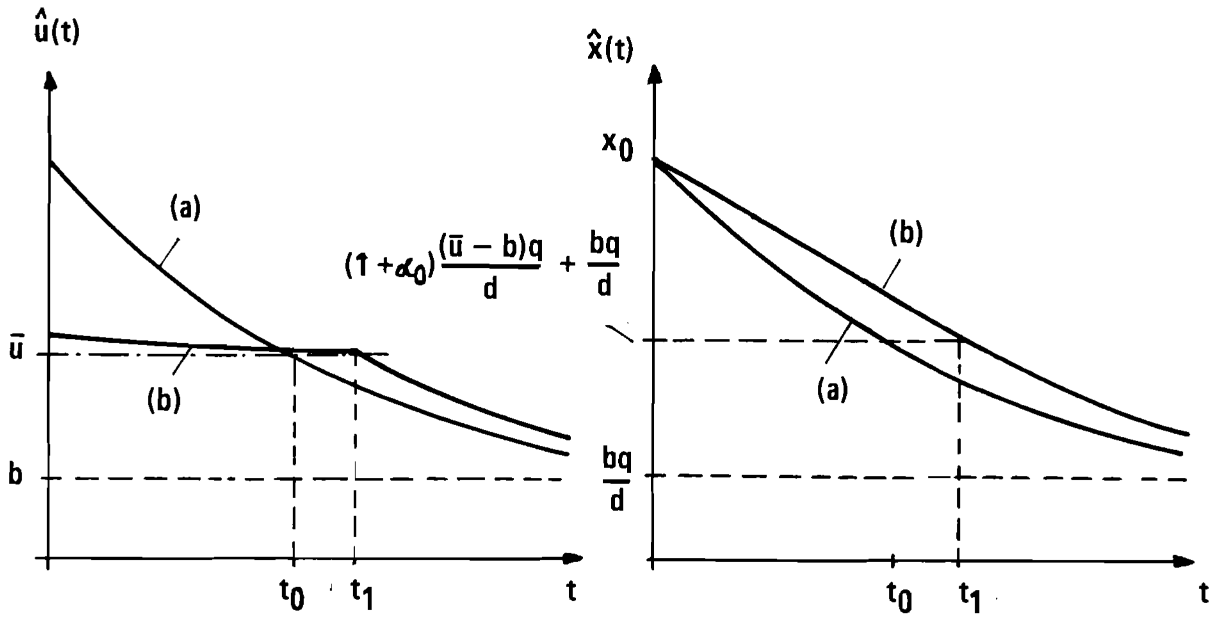

Figure 1. Examples of single-criterium 'optimal' solutions for unemployment $\hat{u}(t)$ and inflation $\hat{x}(t)$--case (a)-- compared with multicriteria D-maximal trajectories of these variables responding to a judgementably set reasonable level $\bar{u}$ of unemployment--case (b).

Observe that, if ρ is sufficiently large, the multicriteria D-maximal trajectory $\hat{u}$ has values $\hat{u}(t)$ only slightly greater than $\bar{u}$, and that the time t_1, at which $\hat{u}(t_1)=\bar{u}$ is also only slightly greater than the corresponding time t_0 for single-criterium case (the last observation follows from the fact that $\int_0^\infty (\hat{u}(t)-b)dt = \frac{1}{rd}(x_0 - \frac{bq}{d})$ for both cases). Thus, when applying multicriteria optimization, we can significantly reduce maximal unemployment while spreading the effects of this reduction over time. Clearly, in this simple example we could obtain similar results just by using an explicit constraint $u(t) \leq \bar{u}$. However, when using hard constraints, we must be careful not to specify $\bar{u} < b$; otherwise we would obtain $x(t) \to \infty$ as $t \to \infty$. When maximizing (71)--which is equivalent to a soft constraint on $u(t)$--we can assume $\bar{u} < b$ and still obtain well-defined results.

Observe also that one could interpret the achievement scalarizing functional (71) as just another form of welfare functional. This interpretation is correct; however, the modified welfare functional depends explicitly on judgementally set desirable bound $\bar{u}$ for unemployment, and in this aspect it differs basically from traditional welfare functionals. Moreover, it possesses the strong order preservation property. Thus, if $\hat{u}$ and $\hat{x}$ correspond to the maximum of this functional, then we cannot decrease the inflation $\hat{u}(t)$ at some t without increasing it at some other t or without decreasing the welfare functional $W(\hat{x},\hat{u})$.

6. COMPUTATIONAL ISSUES AND APPLICATIONS: A TIME-DISCRETE DYNAMIC LINEAR PROGRAMMING CASE

Many problems-- especially in economics (see, e.g. Kallio, Propoi, Seppälä 1980)--are formulated in terms of time-discrete dynamic programming models of the general form: maximize

$$(79) \qquad y^0 = Y^0(x,u) = \sum_{k=0}^{K=1} (c_k^* u_k + d_k^* x_k) + d_K^* x_K$$

subject to state equation constraints:

(80) $x_{k+1} = A_k x_k + B_k u_k$; x_0 - given

and to additional constraints

(81) $(x_k, u_k) \in V_k$

where V_k is a convex polyhedral set (described by linear inequalities), $u_k \in R^m$, $c_k^* \in R^{m*}$, $x_k \in R^n$, $d_k^* \in R^{n*}$, $A_k \in R^n x R^n$, $B_k \in R^n x R^m$. The trajectories x and u are, in this case, finite-dimensional, $u = \{u_0, \ldots u_k, \ldots u_{K-1}\} \in R^{mK}$, $x = \{x_0, \ldots x_k, \ldots x_K\} \in R^{n(K+1)}$, but can choose various norms in these trajectory spaces.

Various approaches have been devised to numerically solve this problem while taking advantage of its special structure (see, e.g. Kallio and Orchard-Hays 1980). For example, one of the efficient approaches is to solve this problem as a large scale static linear programming problem with the number of variables (m+n)K (excluding x_0, which is a given parameter) and generating an initial feasible basic solution by choosing admissible u and solving state equation (80) for x.

It often happens that the solutions of this problem are practically non-unique (many admissible solutions correspond to almost maximal values of y^0) and that we are interested, in fact, not only in y^0 but also in some output trajectories $y^r = \{y_0^r, \ldots y_k^r, \ldots y_{K-1}^r\} \in R^{pK}$ of the model (80)

(82) $y_k^r = C_k^r u_k + D_k^r x_k \in R^p$

where $C_k^r \in R^p \times R^m$, $D_k^r \in R^p \times R^n$. Suppose all output trajectories have to be maximized, thus the positive cone $D^r = R_+^{pK}$, $D = R_+^1 \times R_+^{pK}$.

A particularly convenient form of achievement scalarizing function for this class of problems has been developed by Wierzbicki (1978) and practically applied and further modified by Kallio, Lewandowski and Orchard-Hays (1980). The function corresponds to the choice of a maximum norm in the space

$E_y = R' \times E_y^r$ and has the form

$$(83)\quad s(y-\bar{y}) = \rho \min(y^0-\bar{y}^0, \min_{k,i}(y_k^{ri}-\bar{y}_k^{ri})) + y^{0*}(y^0-\bar{y}^0) + $$

$$+ \sum_{k,i} y_k^{ri*}(y_k^{ri}-\bar{y}_k^{ri})$$

or, if we introduce the surplus variable $w = y-\bar{y} \in R^{pK+1}$, $w = \{w^j\}_{j=0}^{j=pK}$

$$(84)\qquad s(w) = \rho \min_j w^j + y^* w$$

where $\rho > 0$ and y^* is a strictly positive linear function of unit norm in E_y^*. Because we have chosen maximum norm in E_y, E_y^* has the sum of absolute values norm, and y^* is simply a vector of positive weighting coefficients summing up to one, $y^* \in R^{(pK+1)*}$, $y^* = \{y^{j*}\}_{j=0}^{j=pK}$, $\sum_{j=0}^{pK} y^{j*} = 1$, $y^{j*} > 0$. Now, $\min_j w^j$ is strictly order preserving while $y^* w$ is strongly order preserving, thus $s(w)$ is strongly order preserving. Moreover, if $D = R_+^{pK+1}$, then $D \subset S_0 = \{w \in R^{pK+1} : s(w) \geq 0\} = D_{\varepsilon 0} \subset D_\varepsilon$, where D_ε has the form (13) with $\varepsilon > \frac{1}{\rho}$, since $s(w) \geq 0$ and $\|y^*\| = 1$ imply together $\rho \operatorname{dist}(w,D) = -\rho \min_j w^j \leq y^* w \leq \|w\|$. Thus, $s(w)$ is order-approximating and a strong achievement scalarizing function.

The problem of maximizing $s(w)$, however, can be written equivalently as another large scale linear programming problem, by introducing 2(pK+1) or even only (pK+1) additional linear constraints and pK+1 or even only 1 additional variables to the original problem. The modified problem is: maximize

$$(85)\qquad \tilde{s}(w,v) = y^* w + \rho v$$

with $v \in R^1$, subject to:

$$(86)\qquad v \leq w^j,\ j = 0,\ldots pK$$

(87) $$w^0 = \sum_{k=0}^{K-1} (c_k^* u_k + d_k^* x_k) + d_K^* x_K - \bar{y}^0$$

(88) $$\{w^j\}_{j=1}^{pK} = w^r = \{C_k^r u_k + D_k^r x_k - \bar{y}_k^r\}_{k=0}^{K-1}$$

and subject to (80),(81). Clearly we can set (87),(88) into (86), (85), thereby diminishing the number of additional constraints to (pK+1) and the number of additional variables to 1 (the variable v). An efficient algorithm for solving such problems has been developed by Orchard-Hays (see Kallio, Lewandowski, Orchard-Hays 1980).

According to the general theory from section 2, the choice of y^* and ρ does not affect principally the user of the model, who can obtain any desired D-maximal outputs of the model by changing the reference trajectory output $\bar{y}$. However, it might affect the easiness of interaction between the user and the model. This issue has been investigated in Kallio, Lewandowski and Orchard-Hays (1980) where $y^{j*} = \frac{1}{pK+1}$ and $\rho \geq 20$ resulted in good responses of the model. The particular model investigated was a Finnish forestry and forest industrial sector development model with maximized outputs representing the trajectory of the profit of the wood processing industries over time and the trajectory of income of the forestry from selling the wood to the industry over time (10 periods have been considered for each trajectory, hence the total number of objectives was 20; no intertemporal objective was included). Further improvements of the procedure have been also investigated, related to accumulating information about user's preferences revealed by the consecutive choice of reference trajectories $\bar{y}$ after a D-maximal trajectory $\hat{y}$ has been already proposed by the model. However, the main conclusions were the pragmatical and operational usefullness of the procedure; an example of trajectories $\bar{y}$ and $\hat{y}$ obtained in this model is shown in Fig.2.

It should be noted, finally, that achievement scalarizing function (83) is quite similar to functions used in goal programming techniques--see Charnes and Cooper (1961), Dyer (1972), Igmizio (1978), Kornbluth (1973). However, the use of function

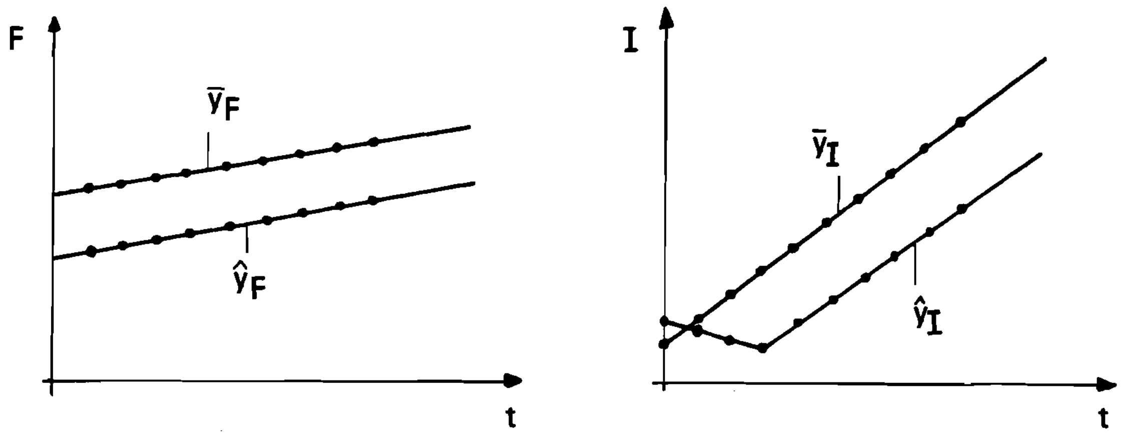

Figure 2. Forestry income trajectory (F) and forest industry profit trajectory (I) obtained in a multiobjective dynamic linear programming model: $\bar{y}$ - desired reference trajectories, $\hat{y}$ - corresponding D-maximal model outputs.

(83) is not related to some of the deficiencies known in applications of goal programming.

7. CONCLUSIONS

In many cases it is desirable and, as shown in this paper, both theoretically and practically possible to use multi-criteria trajectory optimization approaches to various dynamic system models. The approach is based on reference trajectories, when the user of the model specifies what are desirable output reference trajectories of a model and indicates what outputs would be even better than desirable ones, while the model responds with output trajectories that are not only attainable and nondominated in the sense of partial ordering in the output space as indicated by the user, but also correspond to the specified reference trajectories. On one hand, this approach is related to many interesting theoretical questions about the properties of achievement scalarizing functionals in normal spaces, their relations to augmented Lagrangian functionals, to regularization of solutions of ill-posed models; these questions have been investigated, to some extent, in the paper. On the other hand, this approach is also eminently pragmatical; the author hopes that the examples presented show the reasonability and pragmatical values of using the seemingly abstract and untractable notions of infinite-dimensional or high-dimensional multicriteria trajectory optimization.

REFERENCES

Avenhaus, R., R. Beedgen, and L. Schrattenholzer. Handling Uncertainties in Linear Programming Models. WP-80- (in preparation). Laxenburg, Austria: International Institute for Applied Systems Analysis.

Bell, D.E., R.L. Keeney, and H. Raiffa. 1977. Conflicting Objectives in Decisions. New York: Wiley/IIASA International Series on Applied Systems Analysis.

Charnes, A., and W.W. Cooper. 1961. Management Models and the Industrial Applications of Linear Programming, Vols. 1 and 2. New York: Wiley.

Cohn. J.L., and D.H. Marks. 1975. A review and evaluation of multiobjective programming techniques. Water Resources Research 11:208-220.

Da Cunha, N.O., and E. Polak. 1967. Constrained minimization under vector-valued criteria in finite dimensional space. Journal of Mathematical Analysis and Application 19: 103-124.

Debreu, G. 1959. Theory of Value. New York: Wiley.

Dyer, J.S. 1972. Interactive goal programming. Management Science 19(1):62-70.

Everett, H. 1963. Generalized Lagrange multiplier method for solving problems of optimum allocation of resources. Operations Research 11:399-417.

Fortuna, Z. 1974. Properties of Conjugate Direction Method in Hilbert Spaces. Ph. D. thesis, Institute of Automatic Control, Technical University of Warsaw, Warsaw (in Polish).

Hesteness, M.R. 1969. Multiplier and gradient methods. JOTA 4:303-320.

Ignizio, J.P. 1978. Goal programming: a tool for multiobjective analysis. Journal of Operational Research 29(II):1109-1119.

Intriligator, M.D. 1971. Mathematical Optimization and Economic Theory. Englewood Cliffs, N.J.: Prentice-Hall.

Kallio, M., A. Lewandowski, and W. Orchard-Hays. 1980. An Implementation of the Reference Point Approach for Multiobjective Optimization. WP-80-35. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kallio, M., and W. Orchard-Hays. 1979. Experiments with the Reduced Gradient Method for Linear Programming. WP-79-084. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kallio, M., A. Propoi and R. Seppälä. 1980. A Model for the Forest Sector. WP-80-34. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kalman, R., P.L. Falb, and M.A. Arbib. 1969. Topics on Mathematical System Theory. McGraw-Hill, New York.

Keeney, R.L., and H. Raiffa. 1976. Decisions with Multiple Objectives: Preferences and Value Trade-Offs. New York: Wiley.

Kornbluth, J.H.S. 1973. A survey of goal programming. Omega 1:193-205.

Kurcyusz, St. 1974. Necessary Conditions of Optimality for Problems with Constraints in a Function Space. Ph.D. thesis, Institute of Automatic Control, Technical University of Warsaw, Warsaw (in Polish).

Leitmann, G., and W. Schmittendorf. 1973. Some sufficient conditions for Pareto-optimal control Journal of Dynamical Systems, Measurement and Control 95(3).

Mangasarian, O.L. 1976. Equivalence of the complementarity problem to a system of nonlinear equations. SIAM J. Appl. Math. 31:89-92.

March, J.G., and H.A. Simon. 1958. Organizations. New York: Wiley.

Moreau, J.J. 1962. Decomposition orthogonale d'un espace hilbertien selon deux cones mutuellement polaires. C.R. Acad. Sci. Paris 225:238-240.

Robinson, S.M. 1976. Stability theory for systems of inequalities. SIAM J. Numeric. Anal. 23:497-513.

Rockefellar, R.T. 1974. Augmented Lagrangian multipliers functions and duality in nonconvex programming. SIAM J. Control 12:268-285.

Rolewicz, S. 1975. On Norm Scalarization in Infinite-Dimensional Banach Spaces. Working paper, Institute of Mathematics, Polish Academy of Sciences, Warsaw.

Sakluvadze, M.E. 1971. Optimization of vector functionals. Part I: Programming of optimal trajectories. Avtomatika i Telemekhanika 8:5-15. Part II: The analytic constructions of optimal controls. Ibidem 9:5-15 (in Russian).

Sakluvadze, M.E. 1974. On the existence of solutions in problems of optimization under vector-valued criteria. Journal of Optimization Theory and Applications 13(2).

Snower, D., and A.P. Wierzbicki. 1980. Notes on Optimal Policy Choice When Controlling Inflation and Unemployment. WP-80- (in preparation). Laxenburg, Austria: International Institute for Applied Systems Analysis.

Tikhonov, A.N., and W.J. Arsenin. 1975. Solutions of Ill-Posed Problems. Moscow: Nauka (in Russian); translated in 1977 by Fritz John and edited by V.H. Winston & Sons, Washington, D.C.

Wierzbicki, A.P. 1975. Penalty Methods in Solving Optimization Problems with Vector Performance Criteria. Proceedings of VI-th IFAC World Congress, Cambridge/Boston.

Wierzbicki, A.P. 1977a. Basic properties of scalarizing functionals for multiobjective optimization. Mathematische Operations-Forschung und Statistik, Ser. Optimization 8(1):55-60.

Wierzbicki, A.P. 1976. Models and Sensitivity of Control Systems. Warsaw, WNT (in Polish).

Wierzbicki, A.P., and St. Kurcyusz. 1977. Projection on a cone, penalty functionals and duality theory for problems with inequality constraints in Hilbert Space. SIAM J. Control and Optimization 15:25-26.

Wierzbicki, A.P. 1978. On the Use of Penalty Functions in Multiobjective Optimization. Proceedings of the International Symposium on Operations Research, Mannheim.

Wierzbicki, A.P. 1979. A Methodological Guide to Multiobjective Optimization. Proceedings of the 9-th IFIP Conference on Optimization Techniques, Warsaw. Also WP-79-122. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Wierzbicki, A.P. 1980. A Mathematical Basis for Satisficing Decision Making. WP-80-90. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Yu, P.L., and G. Leitmann. Compromise solutions, domination structures and Sakluvadze's solution. Journal of Optimization Theory and Applications 13:362-378.