Generalized Semi-Infinite Optimization and Anticipatory Systems

Tanil Ergenc
Institute of Applied Mathematics, Middle East Technical University (METU)
06531 Ankara, Turkey
Stefan Pickl
Department of Computer Science and DIMACS, University of New Mexico, U.S.A.
Nicole Radde
Department of Mathematics and Center of Applied Computer Science
University of Cologne, University of Cologne, Weyertal 80, 50931 Köln, Germany
Gerhard Wilhelm Weber
Institute of Applied Mathematics, Middle East Technical University (METU)
06531 Ankara, Turkey

Abstract This article is a small survey and pioneering as a starting point for a longer research project: to utilize generalized semi-infinite optimization for purposes of prediction. Firstly, it reflects the *analytical* and *inverse (intrinsic) behaviour* of *generalized semi-infinite optimization* problems $\mathcal{P}(f, h, g, u, v)$ and presents interpretations of them from the viewpoint of *anticipatory systems*. These differentiable problems admit an infinite set $Y(x)$ of inequality constraints y, which depends on the state x. Under suitable assumptions, we present global *stability* properties of the feasible set and corresponding *structural stability* properties of the entire optimization problem [89][90]. The achieved results are a basis of algorithm design.
In the course of explanation, the perturbational approach gives rise to *reconstructions*. By studying three applications of generalized semi-infinite optimization, secondly, we interpret these aspects of *inverse problems* in the sense of *prediction*. The three anticipatory systems are: *(i) Reverse Chebychev approximation*, where we describe a given system by a neighbouring easier one as long as possible under some error tolerance. We begin by a motivating problem from chemical engineering and turn then to time-dependent systems. *(ii) Time-minimal or -maximal optimization problems*, where we want to pull or push the time-horizon of some process to present time or into the future. We mention global warming and turn to further kinds of biosystems. *(iii) Computational biology*, where we are concerned with prediction and stability of DNA microarray gene-expression patterns.

AMS Subject Classification (1994): 05C20, 49J15, 90C34, 92D10, 92D20

Key words: Generalized semi-infinite optimization, structural stability, constraint qualification, prediction, anticipatory system, reverse Chebychev approximation, time-optimal control, thermo-couple, thermo-regulation, gene-expression patterns, inverse problems, statistical learning, experiment design.

1 Introduction.

This article is a first contribution to a general theory of anticipatory systems by means of semi-infinite optimization. In computational biology, medicine and technology research and applications, we have to deal with questions of a long-term understanding and forecasting, e.g., of genetic or metabolic processes or of temperature in atmosphere. *Modelling* of such processes by dynamical systems or optimization problems is followed by a *stability* and *anticipation* analysis and interpretations in biological or technical terms of how meaningful, possibly: optimal, the modelling has been. Based on this analysis and interpretation, maybe, the model must be improved. There is a tension, some tradeoff, between the bounded time-horizon in which the measurements and experiments were made, and the long term in which processes take place. Our topics in modelling, stability analysis and prediction are *inverse problems* of understanding and characterizing the inner nature and behaviour of systems, and they aim at *learning.* Just from these viewpoints we study and interpret our semi-infinite optimization problems.

The following introduction to generalized semi-infinite programming bases on [89][90]; for further foundations see also [73]. Concerning the results of them given in this papers, we do not work out the proofs given there, but figure out main inverse and dynamical-perturbational ideas underlying these results. Generalized semi-infinite ($\mathcal{GSI}$) optimization problems have the form

$$\mathcal{P}(f,h,g,u,v)\left\{\begin{array}{l}\text{Minimize } f(x) \text{ on } M[h,g], \text{ where}\\ M[h,g]:=\{\,x\in\mathbb{R}^n \mid h_i(x)=0 \ (i\in I),\ \ g(x,y)\geq 0 \ (y\in Y(x))\,\}.\end{array}\right.$$

The semi-*infinite* character lies in the typically infinite number of elements of $Y(=Y(x))$ [24][63], while the *generalized* character comes from the x-dependence of $Y(\cdot)$. We suppose these index sets to be finitely constrained ($\mathcal{F}$):

$$Y(x) \ = \ M_{\mathcal{F}}[u(x,\cdot),v(x,\cdot)] := \{\,y\in\mathbb{R}^q \mid u_k(x,y)=0 \ (k\in K),\ v_\ell(x,y)\geq 0 \ (\ell\in L)\},\quad (x\in\mathbb{R}^n).$$

Here we used the following **notation**: $h=(h_i)_{i\in I}$, $u=(u_k)_{k\in K}$, $v=(v_\ell)_{\ell\in L}$, where $h_i:\mathbb{R}^n\to\mathbb{R}$, $i\in I:=\{1,\ldots,m\}$, $u_k:\mathbb{R}^n\times\mathbb{R}^q\to\mathbb{R}$, $k\in K:=\{1,\ldots,r\}$, $v_\ell:\mathbb{R}^n\times\mathbb{R}^q\to\mathbb{R}$, $\ell\in L:=\{1,\ldots,s\}$ $(m<n;\, r<q)$. Let $f:\mathbb{R}^n\to\mathbb{R}$, $g:\mathbb{R}^n\times\mathbb{R}^q\to\mathbb{R}$, $h_i\,(i\in I)$, $u_k\,(k\in K)$, $v_\ell\,(\ell\in L)$ be continuously differentiable (C^1). By $Df(x)$, $D^Tf(x)$ we denote the row- (column) vector of the first-order partial derivatives $\frac{\partial}{\partial x_\kappa}f(x)$, and $D_xg(x,y)$, $D_yg(x,y)$ consist of $\frac{\partial}{\partial x_\kappa}g(x,y)$ and $\frac{\partial}{\partial y_\sigma}g(x,y)$. Let $\mathcal{U}\subset\mathbb{R}^n$, $M[h,g]\cap\mathcal{U}\neq\emptyset$, be some *bounded, open* set. For motivation and references see, e.g., [88][90]. Recent $\mathcal{GSI}$ applications appeared in *optimal design,* namely, of civil and aerospace structures, of experimental design, and in the inverse problems of discrete tomography of, e.g., VLSI chip design [10][15][65][89]. We make the following assumption in order to start our continuity and stability research:

Assumption $\mathbf{A}_{\mathcal{U}}$: $\cup_{x\in\overline{\mathcal{U}}} Y(x)$ is *bounded* (and hence, by continuity, compact).

In *generalized* semi-infinite optimization, $M[h,g]$ does not need not be closed [41]. However, the following assumption ensures closedness:

Assumption $\mathbf{B}_{\mathcal{U}}$: For all $x \in \overline{\mathcal{U}}$, the **linear independence constraint qualification (LICQ)** is fulfilled for $M_{\mathcal{F}}[u(x,\cdot), v(x,\cdot)]$, i.e., linear independence of

$$D_y u_k(\overline{x},\overline{y}),\ k \in K,\quad D_y v_\ell(\overline{x},\overline{y}),\ \ell \in L_0(\overline{x},\overline{y})$$

(considered as a family), where $L_0(\overline{x},\overline{y}) := \{\, \ell \in L \mid v_\ell(\overline{x},\overline{y}) = 0 \,\}$ consists of active indices. We shall realize strong Assumption $\mathrm{B}_{\mathcal{U}}$ to be a central condition of this article, but also a structural frontier overcome by recent research.

Under both assumptions we start our continuity and stability research. Using differential topology [34][37], they admit local linearization of $Y(x)$ $(x \in \overline{\mathcal{U}})$ by finitely many C^1-diffeomorphisms $\phi^j_x : \mathcal{V}^j \to S^j$ $(j \in J)$ in such a way that the image sets Z^j of indices are *x-independent* squares in a linear subspace. Herewith, $\mathcal{P}(f,h,g,u,v)$ becomes locally, namely in $\overline{\mathcal{U}}$, equivalently expressed as an **ordinary** *semi-infinite* optimization problem $\mathcal{P}_{\mathcal{OSI}}(f,h,g^0,u^0,v^0)$, where $M_{\mathcal{OSI}}[h,g^0] \cap \overline{\mathcal{U}} = M[h,g] \cap \overline{\mathcal{U}}$, f being unaffected [84][90].

On the upper stage of variable x, we shall use a constraint qualification, too. This *cq* geometrically means the existence of an at $M[h] = h^{-1}(\{0\})$ tangential, "inwardly" pointing direction at x:

Definition 1.1. We say that the **extended Mangasarian-Fromovitz constraint qualification (EMFCQ)** is fulfilled at an $\overline{x} \in M[h,g]$, if $EMF_{1,2}$ are satisfied:

$\mathbf{EMF}_1$. $Dh_i(\overline{x})$, $i \in I$, are linearly independent.

$\mathbf{EMF}_2$. There exists an "*EMF-vector*" $\zeta \in \mathbb{R}^n$ such that

$$\begin{aligned} Dh_i(\overline{x})\,\zeta &= 0 \quad \text{for all } i \in I, \\ D_x g^0_j(\overline{x},z)\,\zeta &> 0 \quad \text{for all } z \in \mathbb{R}^q,\ j \in J, \text{ with } (\phi^j_{\overline{x}})^{-1}(z) \in Y_0(\overline{x}), \end{aligned}$$

where $Y_0(\overline{x}) := \{\, y \in Y(\overline{x}) \mid g(\overline{x},y) = 0 \,\}$ consists of *active* indices.

EMFCQ is said to be fulfilled for $M[h,g]$ on $\overline{\mathcal{U}}$, if EMFCQ is fulfilled for all $x \in M[h,g] \cap \overline{\mathcal{U}}$.

For further information and versions of EMFCQ see [32][37][41][42][57][74], but also [14][36].

Let a local minimizer $\hat{x}$ of $\mathcal{P}(f,h,g,u,v)$ be given and EMFCQ be fulfilled at $\hat{x}$. Then, we can state the existence of *Lagrange multipliers* λ_i, μ_κ, such that the conditions KT_1:

$$Df(\hat{x}) = \sum_{i\in I} \lambda_i Dh_i(\hat{x}) + \sum_{\kappa\in\{1,\ldots,\hat{\kappa}\}} \mu_\kappa D_x g^0_{j^\kappa}(\hat{x}, z^\kappa), \qquad \mu_\kappa \geq 0 \ \ (\kappa \in \{1,\ldots,\hat{\kappa}\})$$

are satisfied, referring to *ordinary* semi-infinite ($\mathcal{OSI}$) data [32][84][90]. Now, we call $\hat{x}$ a <u>*$\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker point*</u>. Here, the points $z^\kappa \in Z^{j^\kappa}$ are suitable active indices. Below, $Z_0^j(x)$ stands for the set of $z \in Z^j$ being *active* for $g_j^0(x,\cdot)$. Referring to *all* the given $\mathcal{GSI}$ data, a further evaluation yields the following **Kuhn-Tucker conditions** with corresponding *Lagrange multipliers* λ_i, μ_κ, $\alpha_{\kappa,k}$, $\beta_{\kappa,\ell}$ [84][90]:

$$\mathbf{KT}_1.\ Df(\hat{x}) = \sum_{i\in I}\lambda_i Dh_i(\hat{x}) + \sum_{\kappa\in\{1,\ldots,\hat{\kappa}\}}\mu_\kappa D_x g(\hat{x},y^\kappa) - \sum_{k\in K}\alpha_{\kappa,k}D_x u_k(\hat{x},y^\kappa) - \sum_{\substack{\ell\in L_0(\hat{x},y^\kappa)\\ \kappa\in\{1,\ldots,\hat{\kappa}\}}}\beta_{\kappa,\ell}D_x v_\ell(\hat{x},y^\kappa)$$

$$\mathbf{KT}_2.\ \mu_\kappa,\ \beta_{\kappa,\ell} \geq 0 \quad (\ell \in L_0(\hat{x},y^\kappa),\ \kappa \in \{1,\ldots,\hat{\kappa}\}).$$

Again, the $y^\kappa \in Y_0(\hat{x})$ are active. Now, we call $\hat{x}$ a *$\mathcal{G}$ Kuhn-Tucker point*. Under general assumptions, the **necessary optimality conditions** $\mathrm{KT}_{1,2}$ were for the first time proved by Jongen, Rückmann and Stein [41]. Note, that the linear combination KT_1 contains the derivatives of *all* the defining functions. For further information see [49][74][84][90]. In fact, let LICQ be satisfied at a given point $\hat{x}$ as an element of $M[h]$, and $M[h] \cap \overline{\mathcal{U}}$ be *star-shaped* with *star point* $\hat{x}$. Moreover, $g_j^0(\cdot,z)$ $(z \in Z^j,\ j \in J)$ be *quasi-concave* and f be *pseudo-convex* on $M[h] \cap \overline{\mathcal{U}}$. This means the following implications for all $x \in M[h] \cap \overline{\mathcal{U}}$ [33][49]:

$$g_j^0(x,z) \geq g_j^0(\hat{x},z) \quad \Longrightarrow \quad D_x g_j^0(\hat{x},z)\,(x-\hat{x}) \geq 0 \text{ and } Df(\hat{x})\,(x-\hat{x}) \geq 0 \quad \Longrightarrow \quad f(x) \geq f(\hat{x})\,.$$

Then, $\hat{x}$ turns out to be a local minimizer of $\mathcal{P}(f,h,g,u,v)$ [45][84][90]. Concerning structural frontiers in ($\mathcal{F}$) *nonconvex* optimization see [44]. After this introduction of basic conditions, we make the following convention for the ease of presentation. In fact, as the theoretical treatment of the equality constraint functions is merely technical [26][66][80][90], we may delete them:

Convention: Until the end of Subsection 4.1 we assume: $I = \emptyset$, $K = \emptyset$.

Before we introduce the second-order condition for *strong stability* we state (under $\mathrm{A}_\mathcal{U}$, $\mathrm{B}_\mathcal{U}$):

<u>**Lemma 1.2**</u> [90]. Let $\hat{x} \in M[g] \cap \overline{\mathcal{U}}$ be given, and EMFCQ be fulfilled at $\hat{x}$. Then, $\hat{x}$ is a *$\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker point* for $\mathcal{P}(f,g,v)$, if and only if the extended Mangasarian-Fromovitz constraint qualification on $M[(g,-f+f(\hat{x}))]$, called $\widehat{EMFCQ}$, is violated at $\hat{x}$.

We prepare our introduction of strong stability of a stationary point by assuming that f,g,v are C^2 and putting for any bounded open neighbourhood $\mathcal{V} \subseteq \mathbb{R}^q$ of $\bigcup_{x\in\overline{\mathcal{U}}} Y(x)$ and any subset $\mathcal{M} \subseteq \mathbb{R}^n$:

$$\text{norm}[(f,g,v),\mathcal{M}] := \sup\left\{ \sup_{x\in\mathcal{M}} \left\{ |f(x)| + \sum_{i=1}^{n} |\frac{\partial f}{\partial x_i}(x)| + \sum_{\substack{i=1\\j=1}}^{n} |\frac{\partial^2 f}{\partial x_i \partial x_j}(x)| \right\},\right.$$

$$\sup_{\substack{x\in\mathcal{M}\\y\in\overline{\mathcal{V}}}} \max_{\substack{\eta\in\{g\}\cup\\\{v_\nu|\nu\in L\}}} \left\{ |\eta(x)| + \sum_{i=1}^{n} |\frac{\partial \eta}{\partial x_i}(x,y)| + \sum_{j=1}^{q} |\frac{\partial \eta}{\partial y_j}(x,y)| + \sum_{\substack{i=1\\j=1}}^{n} |\frac{\partial^2 \eta}{\partial x_i \partial x_j}(x)| + \right.$$

$$\left.\left.\sum_{i=1}^{n}\sum_{j=1}^{q} |\frac{\partial^2 \eta}{\partial x_i \partial y_j}(x)| + \sum_{\substack{i=1\\j=1}}^{q} |\frac{\partial^2 \eta}{\partial y_i \partial y_j}(x)| \right\}\right\}.$$

In $\mathcal{F}$ or $\mathcal{OSI}$ optimization we replace $\overline{\mathcal{V}}$ by J, Y or disregard v, using notation $\text{norm}_{\mathcal{F}}[\cdot,\cdot]$, $\text{norm}_{\mathcal{OSI}}[\cdot,\cdot]$ then. By continuity stated in Section 2, the next condition is well-defined.

Definition 1.3. Suppose a point $\hat{x}^u \in M[g] \cap \mathcal{U}$ for $\mathcal{P}(f,g,v)$ (of class C^2), $\mathcal{P}_{\mathcal{OSI}}(f,g^0,v^0)$ be locally (in $\overline{\mathcal{U}}$) representing $\mathcal{P}(f,g,v)$, and $\hat{x}^u$ be a $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker point of $\mathcal{P}(f,g,v)$. Then, we say that $\hat{x}^u$ is ($\mathcal{G}$-$\mathcal{O}$) **strongly stable**, if for some $\overline{\epsilon} > 0$ with $B(\hat{x}^u,\overline{\epsilon}) \subseteq \mathcal{U}$ and for each $\epsilon \in (0,\overline{\epsilon}]$ there is some $\delta > 0$ such that for each C^2-function $(\tilde{f},\widetilde{g^0})$ with $\text{norm}_{\mathcal{OSI}}[(f-\tilde{f},g^0-\widetilde{g^0}),B(\hat{x}^u,\epsilon)] \leq \delta$ the open ball $B(\hat{x}^u,\epsilon)$ contains an *ordinary Kuhn-Tucker point* $\hat{x}^d$ of $\mathcal{P}^*_{\mathcal{OSI}}(\tilde{f},\widetilde{g^0}) := \mathcal{P}_{\mathcal{OSI}}(\tilde{f},\widetilde{g^0},v^0)$, which is unique in $B(\hat{x}^u,\overline{\epsilon})$.

Referring to a $\mathcal{G}$ Kuhn-Tucker point $\hat{x}^u$ and to $\text{norm}[(f-\tilde{f},g-\tilde{g},v-\tilde{v}),B(\hat{x}^u,\epsilon)]$, we get the condition for ($\mathcal{G}$) **strong stability**.

Here, "*u, d*" stands for *(un)disturbed* (respectively). For our preferred ($\mathcal{G}$-$\mathcal{O}$) strong stability expressed by original $\mathcal{GSI}$ data, see [90]. In Section 3, we utilize an algebraical characterization of strong stability in the way of Kojima and Rückmann [67][47].

2 Stability of Feasible Sets and Its Characterization.

Results called *Manifold Theorem*, *Continuity Theorem*, *Genericity Theorem* and *Stability Theorem* [89][90] underline the importance of *EMFCQ* for concluding that $M[g,v] := M[g]$ is a topological manifold with boundary, behaving continuously and stable under perturbations of our defining C^1-functions. With these perturbations we remain inside of suitable C^1_S-open neighbourhoods of (g,v). Here, C^1_S stands for the **strong** or **Whitney topology**, which respects asymptotic effects (for topologies C^k_S, $k \in \mathbb{N}\cup\{\infty\}$ cf. [34][37]). We call a given $M \subseteq \mathbb{R}^n$ a **Lipschitzian manifold** (with boundary) of dimension κ, if for each $\overline{x} \in M$ there are open neighbourhoods $\mathcal{W}^1 \subseteq \mathbb{R}^n$ of $\overline{x}$, $\mathcal{W}^2 \subseteq \mathbb{R}^n$ of 0_n, and a bijective $\varphi : \mathcal{W}^1 \to \mathcal{W}^2$, $\varphi(\overline{x}) = 0_n$, with Lipschitzian continuity of φ, φ^{-1} such that φ carries $M\cap\mathcal{W}^1$ to the relatively open set $(\{0_{n-\kappa}\}\times\mathbb{R}^\kappa)\cap\mathcal{W}^2$ or to the set $(\{0_{n-\kappa}\}\times\{w\in\mathbb{R}\,|\,w\geq 0\}\times\mathbb{R}^{\kappa-1})\cap\mathcal{W}^2$ with (relative) boundary. So, Lipschitzian manifolds can locally be linearized, however, without preserving "angulars" in the boundary. According to our Convention, we shall focus on the case $\kappa=n$. In $\mathcal{F}$ optimization, that preservation is given by the stronger condition LICQ, using C^1-smooth linearizing "charts".

For topological *stability*, where the given and any arbitrarily slightly perturbed feasible set can be mapped onto each other by a global homeomorphism, EMFCQ is even *characterizing* in the sense of equivalence. In the next section, we embed this Stability Theorem in the model given by our entire problem $\mathcal{P}(f,g,v)$, where additionally the level parameter τ of the objective function arises.

3 Structural Stability of the Problem and Its Characterization.

3.1 Structural Stability of the Problem.

Under $\mathrm{A}_\mathcal{U}$, $\mathrm{B}_\mathcal{U}$, we still refer to the bounded set $M[g]$, but furthermore take f into consideration. The *structure* of the entire problem $\mathcal{P}(f,g,v)$ is established by all the lower level sets

$$L^\tau(f,g,v) \;:=\; \{x\in\mathbb{R}^n\,|\,x\in M[g,v],\;f(x)\leq\tau\}\quad(\tau\in\mathbb{R}).$$

We observe this structure under perturbation of *all* problem data now, and we define *structural stability.* Here, *descent* has to be preserved, if the level varies. We assume that all defining functions are C^2. Then, this global stability can essentially be *characterized* by EMFCQ of $M[g]$ and by strong stability of all considered stationary points.

Two problems $\mathcal{P}(f^1,g^1,v^1)$, $\mathcal{P}(f^2,g^2,v^2)$ (with defining C^2-functions) are called **structurally equivalent**:

$$\mathcal{P}(f^1,g^1,v^1)\;\sim_\mathcal{P}\;\mathcal{P}(f^2,g^2,v^2),$$

if there are continuous functions $\varphi_\mathcal{P}:\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$ and $\psi:\mathbb{R}\to\mathbb{R}$ with the properties $\mathcal{E}_{1,2,3}$:

E$_1$. $\varphi_{\mathcal{P},\tau}:\mathbb{R}^n\to\mathbb{R}^n$ is a homeomorphism, where $\varphi_{\mathcal{P},\tau}(x):=\varphi_\mathcal{P}(\tau,x)$, for every $\tau\in\mathbb{R}$.

E$_2$. $\psi:\mathbb{R}\to\mathbb{R}$ is a monotonically increasing homeomorphism.

E$_3$. $\varphi_{\mathcal{P},\tau}(L^\tau(f^1,g^1,v^1)) \;=\; L^{\psi(\tau)}(f^2,g^2,v^2)$ for all $\tau\in\mathbb{R}$.

Here, components have to go onto components continuously and from the global viewpoint of all level sets, and continuously connected along of them. To guarantee and prepare all these mapping tasks, the level shift function φ represents the pairwise *correspondence* of two level sets.

Now, let us consider the first problem as *undisturbed* and the second one as slightly *disturbed.* Then, we arrive at *structural stability* [25][40][43][80][90] (cf. also [2][9][37][69]):

$\mathcal{P}(f,g,v)$ (with defining C^2-functions) is called **structurally stable**, if there exists a C^2_S-neighbourhood $\mathcal{O}$ of (f,g,v), such that for each $(\tilde{f},\tilde{g},\tilde{v}) \in \mathcal{O}$

$$\mathcal{P}(f,g,v) \sim_{\mathcal{P}} \mathcal{P}(\tilde{f},\tilde{g},\tilde{v}).$$

3.2 Characterization Theorem.

Under $A_{\mathcal{U}}$, $B_{\mathcal{U}}$ we state:

Theorem 3.1 (**Characterization Theorem** or Structural Stability Theorem; [90]).
Let $M[g] \subset \mathcal{U}$ hold for problem $\mathcal{P}(f,g,v)$ (with defining C^2-functions).
Then, $\mathcal{P}(f,g,v)$ is *structurally stable, if and only if* the conditions $C_{1,2,3}$ are fulfilled:

C$_1$. *EMFCQ* holds for $M[g]$.

C$_2$. All the $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker points $\overline{x}$ of $\mathcal{P}(f,g,v)$ are ($\mathcal{G}$-$\mathcal{O}$) *strongly stable.*

C$_3$. For any two different $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker points $\overline{x}^1 \neq \overline{x}^2$ of $\mathcal{P}(f,g,v)$, the corresponding critical values are different, too: $f(\overline{x}^1) \neq f(\overline{x}^2)$.

In this main result, we could also make a further assumption, excluding certain inequality constraints z from the relative boundary ∂Z^j $(j \in J)$. Then, we could identify the $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker points by some $\mathcal{G}$ Kuhn-Tucker points. For the validity of our Characterization Theorem, however, we need not make such an assumption [90].

3.3 Proof of Characterization Theorem: Main Ideas.

Preparations.

We intensively apply *Implicit Function Theorem in Banach spaces* [37][55]; in particular, we state a *continuous dependence* of $(\tilde{g}^0,\tilde{v}^0)$ on $(\tilde{g},\tilde{v})$. Consequently, small perturbations on the data of $\mathcal{P}(f,g,v)$ cause slight perturbations on the data of $\mathcal{P}_{\mathcal{OSI}}(f,g^0,v^0)$. The *inverse problem* arises: *Can small perturbations of the $\mathcal{OSI}$ data be reconstructed under the problem representation from slight perturbations of the given $\mathcal{GSI}$ problem?* We give a conditionally positive answer. However, this answer will be fitting for the *perturbational* argumentations on Characterization Theorem:

Item 1: For representing $\mathcal{OSI}$ problem(s), $\tilde{v}^0$ is of special linearly affine form and, under sufficiently small perturbations of the $\mathcal{GSI}$ problem, we may treat them as *fixed.* Hence, besides the perturbations $f \to \tilde{f}$, for $\mathcal{P}_{\mathcal{OSI}}(f,g^0,v^0)$ we are concerned with $g^0 \to \widetilde{g^0}$ only. We therefore introduce the simplifying notation $\mathcal{P}^*_{\mathcal{OSI}}(f, g^0) := \mathcal{P}_{\mathcal{OSI}}(f, g^0, v^0)$.

Item 2: Subsequently, we mainly perform local perturbations for $\mathcal{P}^*_{\mathcal{OSI}}(f, g^0)$. Hereby, we treat the finitely many functions g^0_j $(j \in J)$ separately in small *disjoint* open sets $\mathcal{V}^*_j$ $(j \in J)$, such that their perturbations $g^0_j \to \widetilde{g^0_j}$ can be reconstructed by one single C^2-function $\tilde{g}$ (given below). Therefore, we would need the perturbationally stable

Assumption F*: For all $j^1, j^2 \in J$, $j^1 \neq j^2$, we have

$$\bigcup_{x \in M[g] \cap \overline{\mathcal{U}^0}} \left((\phi^{j^1}_x)^{-1}(Z^{j^1}_0(x)) \cap (\phi^{j^2}_x)^{-1}(Z^{j^2}_0(x)) \right) = \emptyset\,.$$

For the well-definedness (possibility) of this hardly controllable assumption we take into account that for any $x \in M[g] \cap \overline{\mathcal{U}^0}$ the sets $Z^{j^\kappa}_0(x)$ merely consist of *active* inequality constraints z. Herewith, they are *subsets* of Z^{j^κ} $(\kappa \in \{1,2\})$. While by definition for some preimages $(\phi^{j^1}_x)^{-1}(Z^{j^1})$ and $(\phi^{j^2}_x)^{-1}(Z^{j^2})$ an overlapping must exist, their subsets $(\phi^{j^1}_x)^{-1}(Z^{j^1}_0(x))$ and $(\phi^{j^2}_x)^{-1}(Z^{j^2}_0(x))$ need not intersect.
We are going to exploit Assumption F* *after* perturbations. However, if we may *suitably* choose our perturbed functions $\widetilde{g^0}$, then Assumption F* is naturally fulfilled (after perturbation), and we need not make it in the unperturbed situation. Now, under problem representation and joined by v, this function $\tilde{g}$ generates $\widetilde{g^0_j}$ locally in $\mathcal{V}^*_j$ $(j \in J)$. Then, for each $j \in J$, small perturbational (global) effects outside of $\mathcal{V}^*_j$ $(j \in J)$ have no influence and can be ignored. The function announced before is

$$\tilde{g}(x,y) := \begin{cases} \widetilde{g^0_j}(x, \phi^j_x(y)), & \text{if } y \in (\phi^j_x)^{-1}(Z^j) \text{ and } (x, \phi^j_x(y)) \in \mathcal{V}^*_j,\ j \in J \\ g(x,y), & \text{else.} \end{cases}$$

Item 3: Below we must consider a certain global perturbation of $\mathcal{P}^*_{\mathcal{OSI}}(f, g^0)$ to receive C^∞-data or, finally, some (global) "open and dense" property. Therefore, we apply on the one hand the perturbation technique from the proof of Genericity Theorem. On the other hand, whenever it is possible to turn from the $\mathcal{GSI}$ problem to an $\mathcal{OSI}$ (or $\mathcal{F}$) one, then we are back in the situation of Item 2 in order to perform local perturbations.

For our proof of Characterization Theorem, the *algebraical characterization of* ($\mathcal{G}$-$\mathcal{O}$) *strong stability* for a $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker point $\overline{x}$ is important [67][90]. Here, we assume EMFCQ at $\overline{x}$. That refined characterization refers to (restricted) Hessians of Lagrange functions, and it bases on a case study where we refer to the *reduction ansatz.* This **RA** demands strong stability in the sense of $\mathcal{F}$ optimization [47] for the local minimizers of the problem from the lower (y-) stage. Herewith, RA admits local representation of $\mathcal{P}(f, g, v)$ around $\hat{x}$ by Implicit Function Theorem [67][90]; see [31][91].

Case I: ELICQ and RA are fulfilled at $\hat{x}$.

Case II: EMFCQ – but *not* ELICQ – and *GRA* are fulfilled at $\hat{x}$.

Case III: EMFCQ – but *not* GRA – is fulfilled at $\hat{x}$.

In all these cases, we can also classify the *type* of the strongly stable stationary point $\overline{x}$: While in case I a saddle point, a local minimizer or local maximizer is detected by the "stationary index" of $\hat{x}$ (a topological invariant), in cases II, III we have a strict local minimizer throughout [90]; cf. [48][67][80].

Line of Proof.

Sufficiency Part:

Let $C_{1,2,3}$ be satisfied. We equivalently represent $\mathcal{P}(f,g,v)$ by $\mathcal{P}_{\mathcal{OSI}}(f,g^0,v^0)$ and interpret $\mathrm{C}_{1,2,3}$ as $\mathcal{OSI}$ conditions $\mathrm{C}_{\mathcal{OSI}\,1,2,3}$: ($\mathcal{OSI}$) constraint qualification EMFCQ, strong stability of all Kuhn-Tucker points in the sense of $\mathcal{OSI}$ optimization, and separateness of the $\mathcal{OSI}$ critical values. Under slight perturbations of the $\mathcal{GSI}$ data, v^0 does not (and need not) vary. Now, we are prepared for $\mathcal{OSI}$ explanations and, finally, $\mathcal{F}$ constructions from [40][43][80] in our $\mathcal{GSI}$ context.
An easy counterexample shows that C_3 can not generally be avoided for establishing structural stability [80]. Here, two connected sets, say: (arcwise) components, would have to be mapped onto one component, contradicting homeomorphy. A similar counterexample shows that the τ-dependence of the homeomorphisms is necessary, too. Moreover, each $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker point $\hat{x}^u$ has to be mapped to the corresponding stationary point $\hat{x}^d$ of the slightly perturbed problem $\mathcal{P}(\tilde{f},\tilde{g},\tilde{v})$. Finally, from the overall boundedness, EMFCQ and strong stability we conclude that the number of $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker points is *finite*: $\hat{x}^u_\sigma$ ($\sigma \in \{1,\ldots,\sigma^0\}$) [43][80][90].

We start by dynamically constructing the *level shift* ψ: We integrate a C^∞-vector field such that each critical value $f(\hat{x}^u_\sigma)$ becomes shifted in $\mathbb{R}$ to the corresponding critical value $\tilde{f}(\hat{x}^d_\sigma)$ ($\sigma \in \{1,\ldots,\sigma^0\}$).
Now, we may think $\psi = Id_{\mathbb{R}}$, otherwise referring to $f \circ \psi$. There are disjoint open ball neighbourhoods $B(\hat{x}^u_\sigma,\epsilon)$ around $\hat{x}^u_\sigma$, such that the smaller neighbourhoods $B(\hat{x}^u_\sigma,\frac{\epsilon}{2})$ contain $\hat{x}^d_\sigma$ ($\sigma \in \{1,\ldots,\sigma^0\}$). Without loss of generality we *assume* that the unperturbed and the perturbed lower level sets *coincide* in *all* the sets $B(\hat{x}^u_\sigma,\epsilon) \setminus \overline{B(\hat{x}^u_\sigma,\frac{\epsilon}{2})}$ ($\sigma \in \{1,\ldots,\sigma^0\}$).
Having performed this reduction of ψ and based on the previous assumption, we *local-globally* proceed by constructing $\varphi_{\mathcal{P},\tau}$ ($\tau \in \mathbb{R}^n$). At first, we realize which undisturbed sets have to be homeomorphically mapped onto which corresponding sets from the disturbed situation (*mapping task*). We distinguish three situations given by levels $\tau < \overline{\tau}$, $\tau = \overline{\tau}$, or $\tau > \overline{\tau}$. Some area from outside of the feasible set possibly must be "carried in". Apart from the stationary points, the level sets *transversally* intersect with the boundaries. On these *fundamental domains* our further construction will be raised.

Outside of $B(\hat{x}^u_\sigma, \epsilon)$ $(\sigma \in \{1, \ldots, \sigma^0\})$, we use "*EMF-technique*" based on Lemma 1.1 and applied on $L^\tau_{\mathcal{OSI}}(f, g^0)$ $(= L^\tau(f, g, v))$, $L^\tau_{\mathcal{OSI}}(\tilde{f}, \tilde{g}^0)$: Within of tubular neighbourhoods we transform undisturbed boundaries onto corresponding disturbed ones along trajectories of vector fields which are generated by EMF vectors. Using differential topology, this **global construction** is glued together in $\cup_{\sigma=1}^{\sigma^0}(B(\hat{x}^u_\sigma, \epsilon) \setminus \overline{B(\hat{x}^u_\sigma, \frac{\epsilon}{2})})$ with the **local construction** sketched next referring to one unperturbed stationary point $\hat{x}^u (= \hat{x}^u_\sigma) \in \{\hat{x}^u_1, \ldots, \hat{x}^u_{\sigma^0}\}$ and a corresponding perturbed point $\hat{x}^d$. Now, we are **inside** of $B(\hat{x}^u, \epsilon)$. We may restrict to $n \in \{2, 3\}$, doing dimensional *reduction* by successive hyperplane intersection otherwise.

Case 1. *$\hat{x}^u$ is lying in the interior $M_{\mathcal{OSI}}[g^0] (= M[g, v])$:*
Then, $\hat{x}^d$, being sufficiently slightly perturbed, lies in the interior of $M_{\mathcal{OSI}}[\tilde{g}^0]$. Both stationary points are *nondegenerate* [37], and for each τ we transform the τ-levels around $\hat{x}^u$ onto the local τ-levels at $\hat{x}^d$. In fact, this Morse theoretical local construction can be made by a C^1-diffeomorphism [43][80].

Case 2: *$\hat{x}^u$ is placed on the boundary of $M_{\mathcal{OSI}}[g^0]$:*
Then, $\hat{x}^d$ may lie on the boundary or in the interior of $M_{\mathcal{OSI}}[\tilde{g}^0]$. Without loss of generality we assume the boundary case. Actually, using an **implantation** of a suitable level structure, we turn from stationary points at the boundary to *fictive* stationary points in the interior. This level structure is locally given by *fictive* objective functions $\hat{f}^u$ and $\hat{f}^d$. For performing this implantation of $\hat{f}^u$, $\hat{f}^d$, we need precise knowledge of the configurations around the boundary points $\hat{x}^u, \hat{x}^d$, characterized by both position of cones or balls with respect to the boundaries and growth behaviours of f, $\tilde{f}$. We have two conical types and one radial type, governed by strong stability (under EMFCQ) [43][80][90]. By means of fictive interior problems, **extrapolating** the "characteristic" of $\hat{x}^u$, $\hat{x}^d$ and implanting fictive stationary points $\hat{x}^u_{fic}$, $\hat{x}^d_{fic}$, we arrive back in *case 1* (interior position). Hence, in case 2, the entire mapping task is also fulfilled.

Necessity Part:

Let $\mathcal{P}(f, g, v)$ be structurally stable; we prove $\mathcal{C}_{\mathcal{GSI}\,1,2,3}$ in *indirect* ways. Based on our assumptions, we carry over the proof of the $\mathcal{OSI}$ necessity part from [40] into our $\mathcal{GSI}$ setting.
Many details of argumentations are Morse theoretical [25][42][43][80][90]. To avoid loss of differentiability, we assume that all data are C^∞ [25]. This smoothness can be achieved by fine perturbations of all $\mathcal{OSI}$ data and, by tracing them back, of all $\mathcal{GSI}$ ones.

Here, we make the inequalities of different indices $\overline{z}^{\varrho^1} \neq \overline{z}^{\varrho^2}$ independent from each other (by small shifts).

$\underline{\mathcal{C}_1}$: Since $M[g]$ is compact, there exists the finite maximum $\tau^{\max} := \max\{f(x) | x \in M[g] (= L^\tau(f, g, v)\ \tau \in [\tau^{\max}, \infty))$. Under sufficiently slight perturbations, $M[\tilde{g}]$ remains compact. Let $\tilde{\tau}^{\max}$ for each slight perturbation $(\tilde{f}, \tilde{g}, \tilde{v})$ denote the maximal (feasible) value of $\tilde{f}$. Taking $\tau^* := \max\{\tau^{\max}, \psi^{-1}(\tilde{\tau}^{\max})\}$, the homeomorphism $\varphi_{\mathcal{P},\tau^*}$ gives topological equivalence between $M[g, v] = L^{\tau^*}(f, g, v)$ and

$M[\tilde{g},\tilde{v}] = L^{\psi(\tau^*)}(\tilde{f},\tilde{g},\tilde{v})$. By *Stability Theorem*, topological stability implies *EMFCQ*. In fact, by suitable perturbations any violation of EMFCQ at a feasible point leads to compact sets $M[\tilde{g}]$, $M[\tilde{\tilde{g}}]$, satisfying ELICQ but being *not* of the same homotopy type [26][42][80][90]. When, e.g., the two sets have a *different* finite number of connected components, this must contradict topological equivalence (cf. also [37]).

$\underline{\mathcal{C}_2}$: Suppose EMFCQ, but C_2 not fulfilled: some $\mathcal{G}$-$\mathcal{O}$ point $\hat{x}^u$ be *not* ($\mathcal{G}$-$\mathcal{O}$) strongly stable.

Lemma 3.2 (**Perturbation Lemma** [90]). Let a $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker point $\hat{x}^u$ of $\mathcal{P}(f,g,v)$ be given with EMFCQ being fulfilled, but ($\mathcal{G}$-$\mathcal{O}$) strong stability violated. Then, for each open C^2-neighbourhood $\mathcal{O}'$ of (f,g,v) there are $(\tilde{f},\tilde{g},\tilde{v}),(\tilde{\tilde{f}},\tilde{\tilde{g}},\tilde{\tilde{v}}) \in \mathcal{O}'$ and $k' \in \mathbb{N}$ such that:

(i) $\mathcal{P}(\tilde{f},\tilde{g},\tilde{v})$ has k' $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker points, all being ($\mathcal{G}$-$\mathcal{O}$) strongly stable, except one (namely, $\hat{x}$).

(ii) $\mathcal{P}(\tilde{\tilde{f}},\tilde{\tilde{g}},\tilde{\tilde{v}})$ has at least $k'+1$ $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker points, all being ($\mathcal{G}$-$\mathcal{O}$) strongly stable.

(iii) In both $\mathcal{P}(\tilde{f},\tilde{g},\tilde{v})$ and $\mathcal{P}(\tilde{\tilde{f}},\tilde{\tilde{g}},\tilde{\tilde{v}})$, EMFCQ is satisfied everywhere, and different $\mathcal{G}$-$\mathcal{O}$ Kuhn-Tucker points have different critical ($\tilde{f}$- or $\tilde{\tilde{f}}$-) values.

In $\mathcal{F}$ or $\mathcal{OSI}$ necessity parts of [25][40][80], these perturbations are realized by three steps. *Step 1* yields local isolation of $\hat{x}^u$ as a stationary point where (E)LICQ is guaranteed but unstability preserved. In *step 2*, outside of the local situation, (E)MFCQ and strong stability of all (other) stationary points are established. In *step 3*, finally, the unstable Kuhn-Tucker point $\hat{x}^u$ "splits": By this bi- (or tri-) furcation we locally get two new stationary points; they have *strong stability.* No, in this $\mathcal{GSI}$ situation, we use the algebraical characterization from our preparations. For $L^\tau(\tilde{f},\tilde{g},\tilde{v})$, $L^\tau(\tilde{\tilde{f}},\tilde{\tilde{g}},\tilde{\tilde{v}})$ we have to take into account each change of the homeomorphy type of a lower level set, when τ traverses $(-\infty,\infty)$. Based on the perturbations from above, we apply the following items on $\mathcal{P}(\tilde{f},\tilde{g},\tilde{v})$, and $\mathcal{P}(\tilde{\tilde{f}},\tilde{\tilde{g}},\tilde{\tilde{v}})$. We look at a C^2-problem $\mathcal{P}(\hat{f},\hat{g},\hat{v})$ with a compact feasible set fulfilling EMFCQ, and put $L_a^b(\hat{f},\hat{g},\hat{v}) := \{x \in M[\hat{g}] \,|\, a \leq \hat{f}(x) \leq b\}$ for some $a,b \in \mathbb{R}$, $a < b$ [67][89][90].

Item 1. If $L_a^b(\hat{f},\hat{g},\hat{v})$ does not contain a stationary point, then $L^a(\hat{f},,\hat{g},\hat{v})$ and $L^b(\hat{f},\hat{g},\hat{v})$ are homeomorphic.

Item 2. Let $L_a^b(\hat{f},\hat{g},\hat{v})$ contain exactly one stationary point $\hat{x}'$. Moreover, let $a < f(\hat{x}') < b$ and $\hat{x}'$ be ($\mathcal{G}$-$\mathcal{O}$) strongly stable. Then, $L^a(\hat{f},\hat{g},\hat{v})$ and $L^b(\hat{f},\hat{g},\hat{v})$ are *not* homeomorphic.

Here, *Item 2* can be expressed with attaching κ-cells (κ = stationary index at $\hat{x}'$). By Manifold Theorem and Lemma 1.1 we conclude for all noncritical levels τ: $L^\tau(\hat{f},\hat{g},\hat{v}) = M[(\hat{g},-\hat{f}+\tau)]$ is a *compact* topological manifold (with boundary).

So, their homology spaces (over $\mathbb{R}$) are of different *finite* dimensions [71]. Since these spaces are topological invariants, the two considered lower level sets cannot be homeomorphic [37].

Now, we are prepared to make a "discrete" statement on topological changes for the lower level sets: The homeomorphy type of $L^{\tau}(\tilde{\tilde{f}}, \tilde{\tilde{g}}, \tilde{\tilde{v}})$ changes (at least) at $k' + 1$ times, whereas the homeomorphy type of $L^{\tau}(\tilde{f}, \tilde{g}, \tilde{v})$ changes (at least) at $k' - 1$ times, but at most at k' times. This contradicts structural stability of $\mathcal{P}(f, g, v)$ [90].

$\underline{\mathcal{C}_3}$: Let C_3 be violated, but EMFCQ and strong stability be satisfied. By local addition of arbitrarily small constant functions on f, we get a problem $\mathcal{P}(f^*, g, v)$ satisfying C_3. Let k^* denote the number of critical points of $\mathcal{P}(f^*, g, v)$. Then the homeomorphy type of $L^{\tau}(f^*, g, v)$ changes k^* times, whereas the homeomorphy type of $L^{\tau}(f, g, v)$ changes less than k^* times. Hence, we are faced again with a situation which is incompatible with structural stability of $\mathcal{P}(f, g, v)$. ∎

Our optimality conditions, topological results and techniques together prepare *iteration procedures* for treating $\mathcal{P}(f, g, v)$. For detailed explanation of the design see [60][85][88][89][90]. Further new approaches and numerical methods are presented in [23][54][65][75][76][77][78][79].

4 Generalizations.

We generalize our inverse and perturbational results along the following two directions:

(I) $M[g]$ is *unbounded* (noncompactness),

(II) f is of the *nondifferentiable $\mathcal{GSI}$ maximum-minimum-type*, i.e., the composition $f = f_p \circ f_{p-1} \circ \ldots \circ f_1$ of finitely many functions which are of max-type $f_j(x) = \max_{\varsigma \in \Upsilon^j(x)} w_j(x, \varsigma)$ or of min-type $f_j(x) = \min_{\varsigma \in \Upsilon^j(x)} w_j(x, \varsigma)$.

On *(I)*: We overcome noncompactness by turning to the family of *excised* subsets of $\overline{M[g]}$. The effect of intersection is generated by subtracting lower semi-continuous functions from g [68][80][90], yielding cuts, e.g., by cylinders or balls, by $\mathbb{R}^n$ itself or by bizarre sets. Referring to *all* excised sets, we get the condition of **excisional topological stability** which can actually be characterized by the overall validity of EMFCQ in the unbounded set $M[g]$. The *(Excisional) Stability Theorem* is given in [90].

On *(II)*: In the case where f is of *max-type*, nonsmoothness can be overcome by representing $\mathcal{P}(f, g, v)$ as minimization of x_{n+1} over the *epigraph* $E(f) := \{(x, x_{n+1}) | \, x \in M[g], \; f(x) \leq x_{n+1}\}$. ¿From this problem in $\mathbb{R}^{n+1}$ we obtain our stationary points of $\mathcal{P}(f, g, v)$ and the appropriate condition of strong stability [80][81][90]. Now, (max-) structural stability of our nondifferentiable problem can again be characterized by EMFCQ, strong stability and the technical separateness

condition. This *Characterization Theorem* and the one for the *case combination* with *(I)* are presented in [90]. A classical example of minimization of max-type functions is given by *Chebychev approximation*; we refer to this in Section 5.
In case of a *min-type* f, we turn to $E(-f)$ and use geometrical insights from the max-type case. Now, in our general case of finite *max-min* composition, we *unfold* nondifferentiability step by step, finally getting our *max-min structural stability* and its characterizing conditions [87].

Remark: In *(II)*, we treated the discrete-combinatorial nondifferentiability structure underlying f by *unfolding* or *lifting* along continuous parameters. For further examples of tracing back structures in the way *"discrete → continuous"*, or *"continuous → continuous"*, *"continuous → discrete"* and *"discrete → discrete"*, cf. [52][90].

5 Anticipatory Systems.

5.1 Prediction and Reverse Chebychev Approximation.

(a) Approximation of a Thermo-Couple Characteristic (Chemical Engineering).

The following motivation of reverse Chebychev approximation from chemical engineering was formulated by Hoffmann and Reinhard [35] and also modelled in [90]. A *thermo-couple* f is some spline of polynomials with different degrees between 3 and 13. It is defined on an interval $[a, b]$ $(a < b)$. ¿From the engineer's point of view, the practical use of a thermo-couple characteristic is very sophisticated. There are several reasons, which call for an *approximation* of the characteristic by a simpler function. For instance, the characteristic cannot be presented in a closed form, the polynomials' degree is too large, and often only a small region of temperature is of practical interest. (Applications in chemical engineering can be found in [93].) Hence, the engineer may look for an approximation by means of only one polynomial $p(y) = \sum_{k=0}^{n^\circ} x_{k+1} y^k$ $(y \in \mathbb{R})$ of some order n° such that the domain of approximation is as large as possible, certain *interpolation* properties are fulfilled and lower and upper *error bounds* are not violated. This optimization problem is naturally called a **reverse Chebychev approximation problem**. Therein we put $n = n^\circ + 2$, $x_\circ = (x_1, \ldots, x_{n^\circ+1})^T$, $x^T = (x_\circ^T, x_n)$ and

$$\Psi(x_\circ, y) := \sum_{k=0}^{n^\circ} x_{k+1} y^k, \quad \delta(x_\circ, y) := \Psi(x_\circ, y) - f(y) \quad (x_\circ \in \mathbb{R}^{n-1},\ y \in \mathbb{R}),$$

referring to some $\alpha \in [a, b)$. Then we may model our problem in $\mathbb{R}^n$ in the following way, which can easily be seen to be of generalized semi-infinite character:

$$\mathcal{P}_{\mathcal{GSI}} \quad \left\{ \begin{array}{l} \text{Minimize} \quad -x_n \text{ on } M_{\mathcal{GSI}}, \\ \text{where } M_{\mathcal{GSI}} \ := \ \{ x \in \mathbb{R}^n \mid \Psi(x_\circ, \hat{y}^i) - \hat{f}^i \ = \ 0 \ \ (i \in \mathsf{I}(x_n)), \\ \delta(x_\circ, y) - \delta^I(y) \geq 0 \ \ (y \in [\alpha, x_n]), \\ -\delta(x_\circ, y) + \delta^{II}(y) \geq 0 \ \ (y \in [\alpha, x_n]), \\ x_n - \alpha \geq 0, \ \ -x_n + b \geq 0 (x^T = (x_\circ^T, x_n), \ x_\circ \in \mathbb{R}^{n-1}, \ x_n \in \mathbb{R}) \}. \end{array} \right\}$$

Namely, taking two different numbers $c, d \in \mathbb{R} \setminus [a,b]$ we would define $Y^v(x) = [\alpha, x_n]$ $(v \in \{1,2\})$, $Y^3(x) \equiv \{c, d\}$, $x \in \mathbb{R}^n$ $(q^v = 1,\ v \in \{1,2,3\})$. However, with the appearance of the x-dependent set $\mathsf{I}(x_n) \subseteq I = \{1, \ldots, m\}$, there is an additional generalization of discrete character in $\mathcal{P}_{\mathcal{GSI}}$. Here, the most important practical situation is given by $\mathsf{I}(x_n) = \{i \in I \mid \hat{y}^i \in [\alpha, x_n]\}$. (A motivation of general index sets can be given in optimal control theory [89][90].) Originally, the points $(\hat{y}^i, \hat{f}^i)$ can be interpreted as interpolation points. Moreover, the continuous functions f, δ^I, δ^{II} need not always be continuously differentiable. In this paper, we stick to the case of $\mathsf{I} \equiv I$ (x-independence) and of C^1-functional data.

Of course, there are other *Chebychev bases*, for instance trigonometric ones, to which one may suitably refer instead of $\{\, x \mapsto x^j \ (x \in [a,b]) \mid j \in \{0, \ldots, n^\circ\} \,\}$.

Based on our mathematical modelling by a generalized semi-infinite optimization problem, all our structure and stability considerations from the previous sections can be applied on our example from chemical engineering here.

(b) Optimization of Anticipatory Systems.

In the previous part *(a)*, we extended the region of approximation, consisting of values of the variable y. This variable, with respect to which we want to extend the horizon optimally, can be of some, e.g., physical dimension, for instance, of a variable or transformation of time, or just time t itself. This interpretation of our foregoing chemical process, or of another biochemical, physical, technological, economical or social process, can be given in terms of *prediction*: We are looking for a *maximal* time-interval along of which the process can be well described or controlled. This wide (in time) understanding, or *anticipation*, of the considered process is meant in an approximative sense where, additionally, interpolation requests can be integrated. Here, by that interval maximization, we want to optimize the entire anticipatory system. By those interpolating conditions, we can put emphasis on some very important data, or highly accurate measurements or experiments. Finally, this understanding can be expressed as a solution of two kinds of problems: *(i) inverse problems*, dealing with the "inner" properties of a system or process based on "outer", "selected" or "sample wise" experimental or measurement data [3][10], and of the problems from recently rising *(ii) statistical learning*, dealing with estimation of parameters based on those data [28].

We underline that all our structural and stability considerations from the previous sections about generalized semi-infinite optimization can by this problem representation be applied here. This is very important for validating our data and the model based on these data or, in terms of statistical learning theory, for testing or *goodness-of-fit measuring* of the model quality. In fact, by stability (in time t) of the process dynamics or, in the topological sense of generalized semi-infinite programming problems representing the system to be optimized (cf. previous sections), we express the robustness and well-posedness of a system.

In Characterization Theorem on structural stability, we stated that strong stability of the stationary points to be one of the central features. This is a second-order condition (in terms of derivatives) which can also be measured by "topological invariants" such as eigenvalues of Hessian matrices (on extended tangential spaces) or by so-called *Morse indices* (for a basic introduction cf. [37][48]). In inverse problems and statistical learning, second-order conditions can be found with the help of covariance matrices. There, significant information is given by confidence regions (ellipsoids), i.e., by the lenghts of its principal axes. These lengths and the projections of the ellipsoids on coordinate axes measure "ellipticity", where also the correlation coefficients of the unknown parameters are indices for [3].

In the following subsection, we continue with our interpretation $y = t$, i.e., of a widely (in time) reliable prediction. Then, anticipation will be understood by right now, at present time, studying the time-horizon of observation, by pushing it into future, which means: *maximization.* Before we maximize, we briefly introduce a corresponding *minimization* problem which is classical in generalized semi-infinite optimization.

5.2 Prediction and Heating Processes: Time-Optimal Control.

(a) Time-Minimal Control in Heating and Cooling Processes.

Let us think that a given ball B consists of a homogeneous material. We study the following problem of heating or cooling B from an initial to a terminal temperature [50][51][90]:

$$\mathcal{P}^{\mathsf{tm}} \quad \left\{ \begin{array}{l} \text{Min} \quad \mathsf{I}(T,u) := T \quad \text{such that there is a bounded function} \\ \theta : [0,R] \times [0,\infty) \to \mathbb{R}, \ \text{ where } \ \theta|(0,R] \times (0,\infty) \\ \text{is partially differentiable,} \quad u = \theta(R,\cdot)|[0,T] \text{ is continuous,} \\ \text{and } \theta_t(r,t) = a\Delta\theta(r,t) = \frac{a}{r^2}\frac{\partial}{\partial r}(r^2\theta_r(r,t)) \\ ((r,t) \in (0,R] \times (0,\infty)), \theta(r,0) = \theta_0 \\ (r \in [0,R]), \theta(R,T) = \theta_E, T \geq 0, |\sigma_u(R,t)| \leq \sigma^* \quad (t \in [0,T]). \end{array} \right. \tag{1}$$

Here, $\Delta\theta$ represents the Laplacian of θ and R denotes the radius of B. The temperature $\theta(r,t)$ is a function of the radial variable r, where r measures the distance from the center point 0_3 of B, and of the time t. Moreover, we start with an **initial** temperature θ_0 and finish with an intended **target** (end) temperature $\theta_E > \theta_0$ (or $\theta_E < \theta_0$, respectively). Because of this inequality, each time $\widehat{T}$ which is optimal for $\mathcal{P}^{\mathsf{tm}}$, can not be zero ($\widehat{T} > 0$). The temperature is essentially governed by the implied heat equation, where $a > 0$ stands for the heat conductivity. This can effectively be realized by the substitution $v(r,t) := r\theta(r,t)$. We interpret $u_T(\cdot) := u(\cdot) = \theta(R,\cdot)|[0,T]$ as a *control variable* ($T \geq 0$). Now, we are focussing on *partial differential equations* with the following unique solution of the (boundary-value) problem:

$$\theta(r,t) = 2r \sum_{k=1}^{\infty} \tfrac{(-1)^{k+1}}{k\pi} \exp(-a(\tfrac{k\pi}{R})^2 t)\theta_0 \tfrac{1}{r} \sin(\tfrac{k\pi}{R} r) + \tfrac{2a}{R} \sum_{k=1}^{\infty} (-1)^{k+1} k\pi \cdot$$
$$\int_0^t \exp(-a(\tfrac{k\pi}{R})^2 (t-s))\, u(s) ds \cdot \tfrac{1}{r} \sin(\tfrac{k\pi}{R} r).$$

Furthermore, $\sigma_u(r,t)$ denotes thermal stress tangential to the boundary $\partial\mathsf{B}$ of B ($r = R$); σ^* is a given upper bound of the stress.

Under suitable physical assumptions, at the boundary $\sigma_u = \sigma_u^{\theta_0}$ has the form

$$(\sigma_u(R,t) =)\ \sigma_u^{\theta_0}(R,t) \;=\; \tfrac{E\alpha}{1-\mu} \big(\tfrac{3}{R^3} \int_0^R \theta(r,t) r^2\, dr \;-\; u(t)\big).$$

Here, E is the modulus of elasticity, μ and α are the coefficients of cross-extension and linear heat extension, respectively. For more detailed explanations and references we refer to [50][51], where also an interpretation of $\mathcal{P}^{\mathsf{tm}}$ as a problem from **two-stage optimization** is given, based on the achieved representation of temperature.

On the lower stage, for each $T \geq 0$ we consider the one-parameter family $(\mathcal{P}_T^{\mathsf{nm}})_{T\in[0,\infty)}$ of *norm-minimal control problems* on the thermal stress at the boundary, given by the *approximation problem*

$$\mathcal{P}_T^{\mathsf{nm}} \qquad \{\ \mathrm{Min}_{u_T}\ ||\sigma_{u_T}^{\theta_0}(R,\cdot)||_{\infty,T},\ \text{ where } u_T \in C([0,T],\mathbb{R}) \text{ fulfills } u_T(T) = \theta_E.\ \}$$

Item [50]: For each $T \geq 0$ the problem $\mathcal{P}_T^{\mathsf{nm}}$ has *precisely one* solution $\hat{u}_T$. ▪ This (unique) solution $\hat{u}_T$ of $\mathcal{P}_T^{\mathsf{nm}}$ ($T \geq 0$) is $\hat{u}_T(t) := \frac{\theta_E - \overline{y}_T(T)}{\overline{u}_T(T)} \cdot \overline{u}_T(t) + \overline{y}_T(t)$ ($t \in [0,T]$), where $(\overline{u}_T, \overline{y}_T)$ is the unique solution of some system of integral equations (cf. [50][51][90] for details):

$$\overline{u}_T(t) - \int_0^t k(t-s)\overline{u}_T(s)\, ds \;=\; 1\overline{y}_T(t) - \int_0^t k(t-s)\overline{y}_T(s)\, ds \;=\; \overline{\theta}_0(t)\ (t \in [0,T])$$

The mapping $u_\vee^0(t,T) := \hat{u}_T(t)$ is called a *core of a Kuhn-Tucker function* or, more precisely: a *global minimizer function.* Inserting the optimal control variables ($u =$) $\hat{u}_T$ into the given problem $\mathcal{P}^{\mathsf{tm}}$ leads to the upper stage, given by the following *generalized semi-infinite ($\mathcal{GSI}$) optimization problem* of class C^0, with $x := T$ and $y := t$:

$$\mathcal{P}_{\mathcal{GSI}}(f,g,v) \qquad \left\{ \begin{array}{l} \text{Min } \ f(x) := x \ \text{ such that} \\ \pm\sigma_{u_T}^{\sigma_0}(R,y) + \sigma^* \geq 0 \ \ (y \in Y(x)), \\ x \geq 0, \text{ where } Y(x) := [0,x] \ \ (x \in \mathbb{R}). \end{array} \right\}$$

Here, g, u comprise the three or two continuous inequality constraints on x and y, respectively. The problem $\mathcal{P}^{\mathrm{tm}}$ is an example for a *terminal problem.* For first numerical treatments including convergence results see [46][51][90]. Because of the form of this generalized semi-infinite problem with its special implication of time, i.e., t and T, such that a differential equation is becoming reflected, the corresponding stability condition is right between the stability conditions on differential equations [2][19] and structural stability of generalized semi-infinite optimization (Subsection 3.1). This special form of optimization problems and stability conditions is important not only for technical applications but also for *biosystems.* In the following part *(b)* and in Subsection 5.3, we mention three of them, located in control of global warming, of temperature control for premature infants and in genetics, respectively.

(b) Control of Global Warming, Optimization of Anticipatory Systems.

When maximizing the time-horizon, we usually do not refer to a terminal state where the system should be controlled to, but to a set where the state trajectories are requested to lie as long as possible. As we did in Section 5.1, a maximization problem can directly be translated to a minimization problem. The problem of global warming [61] can on the one hand be considered as a controllability problem of keeping the temperature in atmosphere (or stratosphere) within certain bounds, or to achieve emission reduction (assumed to contribute to global warming). For this reduction, *Kyoto Protocol* requests a collaboration between the countries, called *joint implementation* [59]; [56] offers optimization and dynamical systems theory. On the other hand, it can be interpreted as the maximization problem of respecting the temperature bounds in time as long as possible. Besides the basic characters and targets of the considered problems, there are some further important differences between the time-minimal control (part *(a)*) and the global warming problem. In fact, the latter one refers to atmosphere as a thin and gaseous boundary layer of the earth rather than to a solid (ball) such as the earth itself. This surrounding layer is not homogeneous. In particular, the temperature also depends on the degree of latitude [61]. A more refined model needs to incorporate also topographical aspects, the underlying distributions of the continents or oceans and, finally, corresponding carbondioxid cycles. All these reasons require that our *global* problem of earth warming is formulated with the temperature $\theta(r,x,t)$, depending also on the 2- (or 3-) dimensional *locally* interpreted variable x. Our second application of thermoregulation comes from medicine: It deals with keeping the heads of premature infants in an appropriate, i.e., not too warm surrounding temperature. Such a care is very important for those babies [7]. Herewith, we have turned to applications from computational biology and medicine. Our third practical field is located in genetics and it consists in the modelling and prediction of DNA microarray patterns. Here, in the sense of our Subsection 5.2 *(b)*, optimized anticipation means a maximized time-horizon. In following Subsection 5.3, we introduce into this field of research.

reverse version of *Chebychev approximation* in the sense of Subsection 5.1 *(b)*. In Subsection 5.2 *(b)*, we made a very related approach by directly addressing the time-horizon T and maximizing it. There, however, the controllability aspect of reaching a terminal state is included. In computational biology, such a required end state may be imposed as a medical threshold or intended health state of the patient. For both of these approaches *(b)* from Subsections 5.1-2 with their formulations by *generalized semi-infinite optimization*, all structural and stability results and reflections from Sections 1-4 can be utilized for the prediction of gene-expression patterns.

Another approach by optimization theory to computational biology exists in the investigation of protein structure. This research is very important for, e.g., the discovery and design of drugs. One distinguishes between four kinds of structure: *primary structure*, which denotes the amino acid sequence, *secondary structure*, which refers to common substructures into which the amino acid chain forms, *tertiary structure*, related to the three dimensional structure of a single protein, and *quarternary structure* for a complex of several proteins. All these can be characterized differently in our topological terms of the previous sections. While in a protein string the primary structure gives more *pointwise* information, the secondary structures connects pointwise to *local* information; tertiary structure incorporates *global* shape information in space, whereas quarternary structure allows *disconnected* configurations given by the appearance of different protein chains. For more information and, in particular, secondary structure prediction, see [92].

In computational biology, DNA microarrays are used. We began to study them in a larger class of *chips*, to which also microchips belong. With the structure of their atom clusters in boundary layers and with further topics *discrete tomography* is concerned [10][17]. We look at these chips from the unifying perspectives of *inverse problems* [3], of *experimental design* [15][16] (cf. also [6]) and of *statistical learning* [28] (cf. [22]).

6 Conclusion.

This paper contains a first approach to interpret and optimize anticipatory systems as generalized semi-infinite optimization problems. We introduced into this wide, well motivated field of mathematical programming problems, and we studied their structure and stability. Finally, we represented problems from chemical engineering and of heating or cooling in the form of generalized semi-infinite problems. These problems are reverse Chebychev or time-optimal problems; we interpreted both problem classes as optimization of anticipatory systems. The stability results which we provided before, serve for a validation of these systems from the viewpoint of mathematical modelling, and for a testing of them from the viewpoint of statistical learning theory.

The present contribution may be a first contribution and encouragement for a new view and treatment for important prediction problems from various fields of

Here, we ask for *stability* or *instability* of the system of differential equations depending on certain (control) parameters in the matrix $M(E)$. We call them *expression-metabolic (em-) parameters.* Answering this question also means *anticipation* of the future, and *(statistical) learning.* Based on (in-) stability, which we detect by our algorithm for the given parameter constellations (having done the time-discretization) [1][21][19], we optimize the process and experimental design such that our mathematical model is permanently under improvement (optimization). Here, we need an intensive interdisciplinary exchange between mathematicians and biologists. This all means a joint process of *learning* in a wide sense.

That approximation problem in its simplest version is unconstrained. In principle it can be solved with well-known mathematical methods. But there may occur problems in a biological sense, like, e.g., having much more genes than time-points, what leads us to an underdetermined system of equations. We face this problem by introducing some biological meaningful constraints, so that we have an approximation problem with inequality constrains. We assume, e.g., that between two time-steps the decrease of the transcript concentration is restricted by a constant vector.
The matrix-valued function M may, for simplicity, be regarded to be constant. Then, by an Euler discretization, our time-continuous dynamics turns into a time-discrete one, and we can interpret the entries of M, i.e., our em-parameters m_{ij}, as the coefficients or rates of how gene i influences gene j. These interrelations become represented by a *gene regulatory network.* We would like to obtain a network with a biologically comparable and reasonable interpretation. Therefore, as a further constrain, it is useful to limit the maximum outdegree and indegree of a node. Since bounding of the maximal outdegree leads to a loss of a valuable decomposition property of our minimization, we bind the indegrees. This is done by means of binary variables. They make our modelling and inverse problem become a *mixed integer programming problem* [20] which can computationally be treated by a *branch and cut algorithm.* If we look at the bounds in a time-depending way, the programming problem becomes *semi-infinite.* In the following, we present a further approach to semi-infiniteness.

In fact, sometimes there is an infinite set of data given which may be countable or uncountable. For example, if we expect periodic behaviour of a process, we could repeat a finite data set or the hypothesis of a continuous nature law periodically. These sets or laws may, however, depend on randomness or they may be numerically or by communication systems hard to evaluate. Therefore, we try to approximate them in a very convenient way so that noise becomes ruled out. An optimal approximation would again mean *discrete* or *Chebychev approximation.* In the case of an infinite number of data represented, e.g., by a continuous variable, this approximation problem can be represented by a semi-infinite optimization problem [32][37]. If, however, our aim is to maximize the time-domain where the approximation under some error bounds, possibly, under interpolation conditions, takes place, we do the

5.3 Computational Biology and Medicine: Prediction of Gene-Expression Patterns.

In this biological and medical context, we point out a few relations between generalized semi-infinite optimization and forecasting into an open time-horizon. Here, we refer to optimized anticipatory systems in the sense indicated in the parts *(b)* of Subsections 5.1 and 5.2, and we give some closer explanations.

DNA-chip or microarray experiments offer the possibility to observe several thousands of expressed genes within a cell simultaneously. This can even be the whole genome of an organism. It is therefore a very convenient technique, e.g., to examine the differences in concentrations of gene products, i.e., messenger RNA, between treated and non-treated cells or between different states during development of a cell, which can be detected by several measurements at different time-points. This is especially interesting if we ask for the reaction of a cell when treated with a certain medicine. In the first step of such a chip-experiment the concerning genes are labelled with fluorescent markers. Then they are verified via hybridization with specific probes called oligonucleotides, which are localized on the chip in a matrix-like scheme. A laser scanner reads out the light signal. In order to compare between different genomes one can label them with different markers, so that they can be detected via various colours. It is even possible to get a quantitative result if you distinguish between different nuances of the intensity of the genes' signals, which can be encoded with different integers. Our aim is to model the development of a cell's expression states over time and to make a prediction. Therefore, we canonically represent the matrices of integers which we get from time-series data by column (expression-) vectors $E(t)$. We represent this process with a system of ordinary differential equations $\dot{E} = M(E)E$ and find the matrix-valued function $M(E)$ by means of *Gaussian* or *least-squares approximation* based on the finitely many DNA-chip measurements [18][19]. This approximation means *nonlinear optimization*; as there are only finitely many data given, we also call it *discrete approximation* or *data fitting*. To be more precise, we ask for the least sum of squares of differences (errors) between difference-like quotients, based on the measurements, and the corresponding values of $M(E)E$. We use a parametrical ansatz for $M(E)$, whose selection bases on the biologists' experiences and expectations in view of the future, e.g., polynomial, piecewise polynomial (spline), exponential or periodic (e.g., trigonometric) developments. Especially, Hill-curves are often used. The hypothesis for using that kind of function to model the functional dependence of $\dot{E}_i$ on E_j is simply that the concentration of the product of the regulating gene must first reach a certain threshold value before it has a meaningful effect on the gene it regulates. By this ansatz and the least-squares approximation we approach the open time-horizon.

We may first refer to the deterministic optimization and control model. (This can later be extended by switching into a stochastic framework, where we take into account normally distributed and uncorrelated DNA-chip data errors and where we may extract the model parameters according to maximum-likelihood estimation.)

science, technology and economy by means of modern optimization theory. As a recent and more and more important application field, we mention modern *life sciences* which encompass research and education about biosystems and human sciences from micro and macro perspectives under criteria like learning and improvement of the quality of live.

Acknowledgement: The authors would like to express their gratitude to Professor Daniel Dubois for his encouragement and support, and to the anonymous referees for their valuable criticism.

References

[1] Akhmet, M.U., Gebert, J., Öktem, H., Pickl, S.W., and Weber, G.W., An improved algorithm for analytical modeling and anticipation of gene expression patterns, preprint no. 6, Institute of Applied Mathematics, Middle East Technical University, Ankara, submitted for publication to special issue of European Journal of Operational Research (EJOR) at the occasion of *XXI EURO Summer Institute 2003 - Stochastic and Heuristic Methods in Optimization*, Neringa, Lithuania (July-August 2003).

[2] Amann, H., *Gewöhnliche Differentialgleichungen*, Walter de Gruyter, Berlin, New York, 1983.

[3] Aster, A., Borchers, B., Thurber, C., *Parameter Estimation and Inverse Problems*, Academic Press, to appear in 2004.

[4] Barthold, F.-J., *Zur Kontinuumsmechanik inverser Geometrieprobleme*, habilitation thesis, Braunschweig University of Technology, Germany, 2000.

[5] Berge, C., *Espaces topologique, fonctions multivoque*, Dunod, Paris, 1966.

[6] Bertram, A., Böhlke, T., Gaffke, N., Heiligers, B. and Offinger, R., On the generation of discrete isotropic orientation distribution for linear elastic crystals, preprint, University of Magdeburg, Germany, Faculty of Engineering, Faculty of Mathematics (2000).

[7] Breuss, M., Fischer, B., and Meister, A., The unsteady thermoregulation of premature infants - a model and its application, in: *Discrete Modelling and Discrete Algorithms in Continuum Mechanics*, eds.: Sonar, Th., and Thomas, I., Logos Verlag, Berlin (2001) 47-56.

[8] Censor, Y., Binary steering in discrete tomography reconstruction with sequential and simultaneous iterative algorithms, preprint, University of Haifa, Israel, 2001.

[9] Chen, L.-Y., Goldenfeld, N., Oono, Y., and Paquette, C., Selection, stability and renormalization, *Physica A* 204 (1994) 111-133.

[10] Diner, C., Dogan, A., Özbudak, F., Weber, G.-W., and Yasar, Ö., On the applied mathematics of discrete tomography, preprint, Institute of Applied Mathematics, Middle East Technical University, Ankara (2004).

[11] Diestel, R., *Graphentheorie*, Springer, 1996.

[12] Dontchev, A.L., Hager, W.W., Poore, A.B., and Yang, B., Optimality, stability, and convergence in nonlinear control, Appl. Math. Optim. 31 (1995) 297-326.

[13] Dubois, D., Sabatier, Ph. For a naturalist approach to anticipation: from catastrophe theory to hyperincursive modelling, *International Journal of Computing Anticipatory Systems, CHAOS* 4 (1999) 35-59.

[14] Fragnelli, V., Patrone, F., Sideri, E., and Tijs, S., Balanced games arising from infinite linear models, preprint, University of Tilburg, Netherland, 1999.

[15] Gaffke, N., and Heiligers, B., Algorithms for optimal design with applications to multiple polynomial regression, *Metrika* 42 (1995) 173-190.

[16] Gaffke, N., and Heiligers, B., Approximative designs for polynomial regression: invariance, admissibility, and optimality, in: *Handbook of Statistics* 13, eds.: Ghosh, S., and Rao, C.R., Elsevier Science (1996) 1149-1199.

[17] Gardner, R. J., and Gritzmann, P., Uniqueness and complexity in discrete tomography, in: *Discrete Tomography - Foundations, Algorithms and Applications*, eds.: Herman, G.T., and Kuba, A., Birkhäuser Verlag, Boston, Basel, Berlin (1999) 85-113.

[18] Gebert, J., Lätsch, M., Ming Poh Quek, E., and Weber, G.-W. Analyzing and optimizing genetic network structure via path-finding, preprint, IAM of METU, Ankara, submitted for publication to Journal of Computational Technologies (2003).

[19] Gebert, J., Lätsch, M., Pickl, S.W., Weber, G.-W., and Wünschiers, R., An algorithm to analyze stability of gene-expression pattern, preprint no. 8, Institute of Applied Mathematics, Middle East Technical University, Ankara, revised version, submitted for spezial issue *Operations Research and Data Analysis* of Annals of Operations Research, guest eds.: Boros, E., Hammer, P.L., and Kogan, A. (2003).

[20] Gebert, J., Lätsch, M., Radde, N., Pickl, S.W., Weber, G.-W., and Wünschiers, R., Genetic networks and anticipation of gene expression patterns, preprint no. 13, Institute of Applied Mathematics, Middle East Technical University, Ankara,

revised version, to appear in *International Journal of Computing Anticipatory Systems*, Liege, Belgium, 2004.

[21] Gebert, J., Pickl, S.W., Shokina, N.,Weber, G.-W., and Wünschiers, R., Algorithmic analysis of gene expression data with polyhedral structures, in: *Similarity Methods - 5th International Workshop*, at the occasion of workshop of same name, Stuttgart (November 2002), by Institute of Statics and Dynamics of Aerospace Structures, University of Stuttgart, eds.: B. Kröplin, S. Rudolph und J. Häcker, ISBN 3-930683-47-4 (2002) 79-87.

[22] Gebert, J., Lätsch, M., Pickl, S.W., Weber, G.-W., and Wünschiers, R., Towards DNA chip analysis and discrete tomography by reverse problems and optimization, proceedings of *4th MATHMOD Vienna - Fourth International Symposium on Mathematical Modelling*, Vienna, February 2003 (2003) 1298-1302.

[23] Geletu, A., and Hoffmann, A., Solving generalized semi-infinite programming problems via global optimization, to appear in special issue of *European J. Operational Research* at the occasion of Workshop *Smooth and Nonsmooth Optimization - Theory and Applications*, Rotterdam, 2001.

[24] Goberna, M. A., and López, M. A., *Linear Semi-Infinite Optimization*, John Wiley, New York, 1998.

[25] Guddat, J., and Jongen, H. Th., Structural stability and nonlinear optimization, *Optimization* 18 (1987) 617–631.

[26] Guddat, J., Jongen, H. Th., and Rückmann, J.-J., On stability and stationary points in nonlinear optimization, *J. Austral. Math. Soc., Ser. B* 28 (1986) 36–56.

[27] Haase, M., An exploration of self-similar structures with wavelets, in: *Similarity Methods*, eds.: Kröplin, B., Rudolph, S., and Brückner, S., Institut für Statik und Dynamik der Luft- und Raumfahrtkonstruktionen, University of Stuttgart (2000) 183-192.

[28] Hastie, T., Tibshirani, R., and Friedman, J., *The Elements of Statistical Learning - Data Mining, Inference and Prediction*, Springer Series in Statistics, 2001.

[29] Herman, G.T., and Kuba, A., eds., *Discrete Tomography - Foundations, Algorithms and Applications*, Birkhäuser Verlag, Boston, Basel, Berlin (1999).

[30] Hestenes, M. R., *Calculus of Variations and Optimal Control Theory*, John Wiley, New York, 1966.

[31] Hettich, R., and Jongen, H.Th., Semi-infinite programming: conditions of optimality and applications, in: *Optimization Techniques 2*, ed.: Stoer, J., Lecture Notes in Control and Information Sci., Springer (1978) 1-11.

[32] Hettich, R., and Zencke, R., *Numerische Methoden der Approximation und semi-infiniten Optimierung*, Teubner, Stuttgart, 1982.

[33] Higgins, J. E., and Polak, E., Minimizing pseudoconvex functions on convex compact sets, *Journal of Optimization Theory and Applications* 65 (1990) 1-27.

[34] Hirsch, M. W., *Differential Topology*, Springer, 1976.

[35] Hoffmann, A., and Reinhard, R., On reverse Chebychev approximation problems, preprint, Ilmenau University of Technology, Department of Mathematics (1994).

[36] Jongen, H.Th., personal communication, Alicante, 1999.

[37] Jongen, H.Th., Jonker, P., and Twilt, F., *Nonlinear Optimization in Finite Dimensions.* Nonconvex Optimization and Its Applications 47, Kluwer (2000).

[38] Jongen, H.Th., Jonker, P., and Twilt, F., Critical sets in parametric optimization, *Mathematical Programmming* 34 (1986) 333-353.

[39] Jongen, H.Th., and Pallaschke, D., On linearization and continuous selections of functions, *Optimization* 19 (1988) 343-353.

[40] Jongen, H.Th., and Rückmann, J.-J., On stability and deformation in semi-infinite optimization, in: *Semi-Infinite Programming*, eds.: Reemtsen, R., and Rückmann, J.-J., Kluwer Academic Publishers, Boston, Dordrecht, London (1998) 29-67.

[41] Jongen, H.Th., Rückmann, J.-J., and Stein, O., Generalized semi-infinite optimization: a first order optimality condition and examples, *Mathematical Programming* 83 (1998) 145-158.

[42] Jongen, H.Th., Twilt, F., and Weber, G.-W., Semi-infinite optimization: structure and stability of the feasible set, *J. Optim. Theory Appl.* 72 (1992) 529-552.

[43] Jongen, H.Th., and Weber, G.-W., Nonlinear optimization: characterization of structural stability, *J. Global Optimization* 1 (1991) 47-64.

[44] Jongen, H.Th., and Weber, G.-W. Nonconvex optimization and its structural frontiers, in: *Modern Methods in Optimization*, eds.: Krabs, W., and Zowe, J., Lecture Notes in Econom. and Math. Systems 378 (1992) 151-203.

[45] Kaiser, C., and Krabs, W., Ein Problem der semi-infiniten Optimierung im Maschinenbau und seine Verallgemeinerung, working paper, Darmstadt University of Technology, Department of Mathematics, 1986.

[46] Kaplan, A., and Tichatschke, R., On the numerical treatment of a class of semi-infinite terminal problems, *Optimization* 41 (1997) 1-36.

[47] Kojima, M., Stable stationary solutions in nonlinear programs, in: *Analysis and Computation of Fixed Points*, ed.: Robinson, S.M., Academic Press, New York (1980) 93-138.

[48] Kojima, M., and Hirabayashi, R., Continuous deformations of nonlinear programs, *Math. Programming Study* 21 (1984) 150-198.

[49] Krabs, W., *Optimization and Approximation*, John Wiley, New York, 1979.

[50] Krabs, W., On time-minimal heating or cooling of a ball, *International Series of Numerical Mathematics* 81, Birkhäuser, Basel (1987)121-131.

[51] Kropat, E., Pickl, S.W., Rößler, A., and Weber, G.-W., A new algorithm from semi-infinite optimization for a problem from time-minimal optimal control, *J. Computational Technologies* 5, 4 (2000) 67-81.

[52] Kropat, E., Pickl, S.W., Rößler, A., and Weber, G.-W., On theoretical and practical relations between discrete optimization and nonlinear optimization, in special issue *Colloquy Optimization - Structure and Stability of Dynamical Systems* of Journal of Computational Technologies 7, Special Issue (2002) 27-62.

[53] Lee, E. B., and Markus, L., *Foundations of Optimal Control Theory*, John Wiley, New York, 1969.

[54] Levitin, E., and Tichatschke, R., A branch and bound approach for solving a class of generalized semi-infinite programming problems, preprint, University of Trier (1996).

[55] Ljusternik, L.A., and Sobolew, W. I., *Elemente der Funktionalanalysis*, Akademie Publ. House, Berlin, 1960.

[56] Lozuvanu, D., Pickl, S.W., and Weber, G.-W. Optimization, monotonicity and the determination of Nash equilibria, preprint no. 14, Institute of Applied Mathematics, Middle East Technical University, Ankara, to appear in *International Journal of Computing Anticipatory Systems*, Liege, Belgium, 2004.

[57] Mangasarian, O.L., and Fromovitz, S., The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, *J. Math. Anal. Appl.* 17 (1967) 37-47.

[58] Maurer, H., and Pickenhain, S., Second-order sufficient conditions for control problems with control-state constraints, *J. Optim. Theory Appl.* 86 (1995) 649-667.

[59] Pickl, S.W., *Der τ-value als Kontrollparameter. Modellierung und Analyse eines Joint-Implementation Programmes mithilfe der kooperativen dynamischen Spieltheorie und der diskreten Optimierung*, Shaker-Verlag, Aachen, 1999.

[60] Pickl, S.W., and Weber, G.-W., An algorithmic approach by linear programming problems in generalized semi-infinite optimization, *J. Computational Technologies* 5, 3 (2000) 62-82.

[61] Pickl, S.W., and Weber, G.-W., Optimal control of heating processes - with special emphasis on earth warming, in: *Operations Research Proceedings 2001*, eds.: Chamoni, P., Leisten, R., Martin, A., Minnemann, J., and Stadler, H. (2002) 247-254.

[62] Pontryagin, L.S., Boltryanski, V.G., Gamkrelidze, R.V., and Mishchenko, E.F., *The Mathematical Theory of Optimal Processes*, Interscience Publishers, John Wiley, New York, 1962.

[63] Reemtsen, R., and Rückmann, J.-J., *Semi-Infinite Programming*, Kluwer Academic Publishers, Boston, Dordrecht, London, 1998.

[64] Rockafellar, R. T., *Convex Analysis*, Princeton University Press, 1970.

[65] Royset, J.O., Polak, E., and Der Kiureghian, A., Adaptive approximations and exact penalization for the solution of generalized semi-infinite min-max problems, SIAM J. Optimization 14, 1 (2003) 1- 34.

[66] Rückmann, J.-J., Stability of noncompact feasible sets in nonlinear optimization, in: *Parametric Optimization and Related Topics III*, eds.: Guddat, J., Jongen, H.Th., Kummer, B., and Nožička, F., Peter Lang Publ. House, Frankfurt a.M., Bern, New York (1993) 163-214.

[67] Rückmann, J.-J., On the existence and uniqueness of stationary points in semi-infinite optimization, Mathematical Programming, Ser. A 86 (1999) 387-415.

[68] Rückmann, J.-J., and Weber, G.-W., Semi-infinite optimization: excisional stability of noncompact feasible sets, Sibirskij Mathematicheskij Zhurnal 39 (1998) 129-145 (in Russian). An English version of this article appeared in Siberian Mathematical Journal 39 (1998) 113-125.

[69] Shub, M., *Global Stability of Dynamical Systems*, Springer, Berlin, Heidelberg, New York, 1987.

[70] Sonar, Th., and Thomas, I. (eds.), *Discrete Modelling and Discrete Algorithms in Continuum Mechanics*, proceedings of GAMM workshop (Braunschweig, 2000), Logos Verlag, Berlin, 2001.

[71] Spanier, E. H., *Algebraic Topology*, McGraw-Hill, New York, 1966.

[72] Stein, O., On level sets of marginal functions, Optimization 48 (2000) 43-67.

[73] Stein, O., *Bi-level Strategies in Semi-infinite Programming*, Nonconvex Optimization and Its Applications, Kluwer Academic Publishers, Boston, Dordrecht, London, 2003.

[74] Stein, O., and Still, G., On optimality conditions for generalized semi-infinite programming problems, Journal of Optimization Theory and Applications 104 (2000) 443-458.

[75] Stein, O., and Still, G., On generalized semi-infinite optimization and bilevel optimization, European J. Operational Research 142, 3 (2002) 444-462.

[76] Stein, O., and Still, G., Solving semi-infinite optimization problems with interior point techniques, *SIAM Journal on Control and Optimization* 42, 3 (2003) 769-788.

[77] Still, G., Generalized semi-infinite programming: Theory, methods, *European J. Operational Research* 119 (1999) 301-313.

[78] Still, G., Generalized semi-infinite programming: Numerical aspects, *Optimization* 49, 3 (2001) 223-242.

[79] Still, G., Discretization in semi-infinite programming: Rate of convergence, *Math. Program., Ser. A* 91, 53-69, (2001).

[80] Weber, G.-W., *Charakterisierung struktureller Stabilität in der nichtlinearen Optimierung*, doctoral thesis, Aachen University of Technology, Department of Mathematics, Aachener Beiträge zur Mathematik 5, eds.: Bock, H.H., Jongen, H.Th., and Plesken, W., Augustinus Publ. House, Aachen, 1992.

[81] Weber, G.-W., Minimization of a max-type function: characterization of structural stability, in: *Parametric Optimization and Related Topics III*, eds.: Guddat, J., Jongen, H. Th., Kummer, B., and Nožička, F., Peter Lang Publ. House, Frankfurt a.M., Bern, New York (1993) 519-538.

[82] Weber, G.-W., On the topology of parametric optimal control, J. Austral. Math. Soc., Ser. B 39 (1998) 463-497.

[83] Weber, G.-W., Optimal control theory: on the global structure and connections with optimization. Part 1, J. Computational Technologies 4, 2 (1999) 3-25.

[84] Weber, G.-W., Generalized semi-infinite optimization: on some foundations, J. Computational Technologies 4, 3 (1999) 41-61.

[85] Weber, G.-W., Generalized semi-infinite optimization: on iteration procedures and topological aspects, in: *Similarity Methods*, eds.: Kröplin, B., Rudolph, S., and Brückner, S., Institute of Statics and Dynamics of Aviation and Space-Travel Constructions (1998) 281-309.

[86] Weber, G.-W., Approaches to mechanics from discrete and nonlinear optimization and control, lecture held at Workshop *Discrete Modelling and Discrete Algorithms in Continuum Mechanics*, org.: Sonar, Th., and Thomas, I., Braunschweig (november 2000).

[87] Weber, G.-W., Impulses from results of inverse problems to discrete tomography, lecture held at Workshop *The Mathematics of Discrete Tomography*, org.: Gardner, R. J., and Gritzmann, P., Oberwolfach (december 2000).

[88] Weber, G.-W., Generalized semi-infinite optimization: theory and applications in optimal control and discrete optimization, in: special issue *Optimality Conditions, Generalized Convexity and Duality in Vector Optimization*, eds.: Cambini, A., and Martein, L., of Journal of Statistics and Management Systems 5 (2002) 359-388.

[89] Weber, G.-W. On the topology of generalized semi-infinite optimization, in: special issue *Optimization* (at the occasion of *French-German-Italian Conference on Optimization* in Montpellier, September 2000) of Journal of Convex Analysis 9, 2 (2002) 665-691.

[90] Weber, G.-W., *Generalized Semi-Infinite Optimization and Related Topics*, Research and Exposition in Mathematics 29, Heldermann Verlag, Lemgo, 2003.

[91] Wetterling, W.W.E., Definitheitsbedingungen für relative Extrema bei Optimierungs- und Approximationsaufgaben, Numer. Math. 15 (1970) 122-136.

[92] Yilmaz, Özlem *A Two-Stage Mathematical Programming Algorithm for Predicting Secondary Structures of Proteins*, master thesis, Graduate School of Engineering, Koc University, Istanbul, 2003.

[93] Zlokarnik, M., *Dimensional Analysis and Scale-up in Chemical Engineering*, Springer, Berin, Heidelberg, New York, 1991.