Proc. 12th Conf. on Variational Calculus, Optimal Control and Applications, Trassenheide, 23-27.9.1996
(Eds. W.H.Schmidt, K.Heier, L.Bittner, R.Bulirsch) ISNM 124 (1998), Birkhuser, Basel, pp.87-96.

Existence results for some nonconvex optimization problems governed by nonlinear processes.

Tomáš Roubíček

Mathematical Institute, Charles University
and
Institute of Information Theory and Automation, Academy of Sciences

Abstract. Optimal control problems with nonlinear equations usually do not possesses optimal solutions. Nevertheless, if the cost functional is uniformly concave with respect to the state, the solution may exist. Using the Balder's technique based on a Young-measure relaxation, Bauer's extremal principle and investigation of extreme Young measures, the existence is demonstrated here for the case of nonlinear ordinary and partial differential equations.

Key words. optimal control, nonlinear differential equations, relaxation, Bauer's extremal principle, extreme Young measures, existence theory.

1991 Mathematics Subject Classification. 49J15, 49J20.

1. Introduction.

The optimal control problems usually fail to have solutions unless the cost functional to be minimized is convex and the controlled system is linear with respect to the controls. This failure is typically due to oscillations effects: minimizing sequences tend to oscillate faster and faster which eventually prevent them to be convergent in a norm topology so that the limit passage through the involved Nemytskiĭ mappings is impossible. A typical example of this sort is the following optimal control problem:

$$(1.1)\quad \begin{cases} \text{Minimize} & \displaystyle\int_0^T (u(t)^2-1)^2 + y(t)^2 \,\mathrm{d}t & & \text{(cost functional)} \\ \text{subject to} & \mathrm{d}y/\mathrm{d}t = u(t) \quad \text{for a.a. } t\in(0,T), & y(0)=0, & \text{(state equation)} \\ & -1 \le u(t) \le 1 \quad \text{for a.a. } t\in(0,T). & & \text{(control constraints)} \end{cases}$$

Minimizing sequences of controls inevitably oscillate faster and faster around -1 and 1, converging weakly* (but not strongly) to $u=0$. Yet, $u=0$ is not an optimal control.

Sometimes, even problems with nonconvex cost functionals or nonlinear state equations may have a solution. E.g., if changing the sign in the term y^2 in (1.1) so that we deal with the cost functional

$$\int_0^T (u(t)^2-1)^2 - y(t)^2 \mathrm{d}t,$$

the problem (1.1) does have a solution! Using constructive methods, this phenomenon has been observed for cost functionals and systems linear with respect to the state y already by Gabasov and Kirillova [8; Section 5.3], Macki and Strauss [10; Section 4.5], Neustadt [12] or Olech [13]. For more complicated problems (typically convex cost functionals and linear systems or, in a multidimensional case, the variational constraint $\nabla y = u$) see also Cellina and Colombo [5], Cesari [6] and [7; Chapter 16], Mariconda [11] or Raymond [14–16].

Recently, Balder [1] proposed another scheme based on the following three steps: first to relax the original problem to ensure an existence of so-called optimal relaxed controls, then to use a Bauer extremal principle [2] to show that at least one optimal relaxed control is an extreme point of the set of all admissible relaxed controls, and then to show that such extreme points are essentially the original controls, which eventually yields an optimal control for the original problem. However, Bauer's principle requires the relaxed problem to be concave. Therefore, the existence investigations for the original problem are reduced basically to two questions:

- Which data qualification does guarantee the concave structure of the relaxed problem?
- When every extreme point of the set of admissible relaxed controls is the original control?

In Section 2 of this contribution, we want to illustrate the basic situation on the simplest optimal control problem for the ordinary differential equations with additively separated cost functional as well as the state equation as in (1.1) with $S(t)$ uniformly bounded. Considering problems having a uniformly concave cost functional but a "slightly" nonlinear state equation, both above questions can be affimatively answered; cf. respectively Lemmas 1-2 and 3. The methods used for the first question are intimately related with a sufficiency of the Pontryagin maximum principle, which requires basically a convexity of the relaxed problem (at least "at the optimal point"); see Gabasov and Kirillova [8; Section VII.2] or, for the case of general integral processes, also Schmidt [19]. By this technique one can show the concave structure of the relaxed problem even if the controlled system is "slightly" nonlinear with respect to the state on the assumption that the cost functional is "enough" uniformly concave with respect to the state. As to the second question, we will use (and modify) the results by Berliocchi and Lasry [3] and Castaing and Valadier [4].

Then, in Sections 3 and 4, we want to expose briefly the generalizations to problems with unbounded controls and governed by partial differential equations.

2. Main results.

Avoiding problems with state-space constraints as well as problems with non-additively coupled states and controls (which would cause considerable complications), we consider the optimal control problem for a (system of) ordinary differential equations in the form:

$$
(\mathsf{P}) \quad \begin{cases} \text{Minimize} & \displaystyle\int_0^T g(t,y(t))+h(t,u(t))\,\mathrm{d}t & \text{(cost functional)}\\ \text{subject to} & \dfrac{\mathrm{d}y}{\mathrm{d}t}=G(t,y(t))+H(t,u(t)) \ \text{ for a.a. } t\in(0,T) & \text{(state equation)}\\ & y(0)=0, & \text{(initial condition)}\\ & u(t)\in S(t) \qquad \text{for a.a. } t\in(0,T), & \text{(control constraints)}\\ & y\in W^{1,q}(0,T;\mathbb{R}^n), \quad u\in L^\infty(0,T;\mathbb{R}^m)\ . \end{cases}
$$

where $g:(0,T)\times\mathbb{R}^n\to\mathbb{R}$, $G:(0,T)\times\mathbb{R}^n\to\mathbb{R}^n$, $h:(0,T)\times\mathbb{R}^m\to\mathbb{R}$, $H:(0,T)\times\mathbb{R}^m\to\mathbb{R}^n$ are Carathéodory functions satisfying the growth conditions

$$
\text{(2.1a)} \qquad \exists a\in L^q(0,T)\ \ \exists b\in\mathbb{R}: \quad |G(t,r)|\le a(t)+b|r|, \qquad |H(t,s)|\le a(t),
$$

$$
\text{(2.1b)} \qquad \exists a\in L^1(0,T): \quad |g(t,r)|\le a(t), \quad |h(t,s)|\le a(t)
$$

with some $q\in(1,+\infty)$. Moreover, $G(t,\cdot)$ is Lipschitz continuous in the sense

$$
\text{(2.1c)} \qquad \exists a\in L^1(0,T): \quad |G(t,r_1)-G(t,r_2)|\le a(t)|r_1-r_2|,
$$

As to the multivalued mapping $S:(0,T)\rightrightarrows\mathbb{R}^m$, we suppose it in this section as bounded, measurable, and in the form $S(t)=M(t,S_0)$ for some $S_0\subset\mathbb{R}^m$ compact and $M:(0,T)\times\mathbb{R}^m\to\mathbb{R}^m$ a Carathéodory mapping such that both $M(t,\cdot)$ and $M(t,\cdot)^{-1}$ are Lipschitz continuous uniformly with respect to $t\in(0,T)$. Note that (P) covers also (1.1) for S qualified appropriately.

Following ideas by Young [21], we extend the set of admissible controls $U_{\mathrm{ad}}:=\{u\in L^\infty(0,T;\mathbb{R}^m);\ u(t)\in S(t)$ for a.a. $t\in(0,T)\}$ to the set of admissible relaxed controls $\bar U_{\mathrm{ad}}:=\{\nu\in\mathcal{Y}(0,T;\mathbb{R}^m);\ \mathrm{supp}(\nu_t)\subset S(t)$ for a.a. $t\in(0,T)\}$, where $\mathcal{Y}(0,T;\mathbb{R}^m):=\{\nu: t\mapsto\nu_t:(0,T)\to\mathrm{rca}(\mathbb{R}^m)$ weakly measurable; ν_t is a probability measure for a.a. $t\in(0,T)\}$ denotes the set of the so-called Young measures and $\mathrm{rca}(\mathbb{R}^m)\cong C_0(\mathbb{R}^m)^*$ stands for Radon measures on $\mathbb{R}^m$. It is known that U_{ad} is weakly* dense in $\bar U_{\mathrm{ad}}$ if imbedded via the mapping $i:u\mapsto\nu$ with $\nu_t=\delta_{u(t)}$ where $\delta_s\in\mathrm{rca}(\mathbb{R}^m)$ denotes the Dirac measure supported at $s\in\mathbb{R}^m$; cf. [17] or, for S constant, also Cesari [7] or Warga [20]. The relaxed problem is then created by the continuous extension of the original problem (P) from U_{ad} to $\bar U_{\mathrm{ad}}$, which gives:

$$
(\mathsf{RP}) \quad \begin{cases} \text{Minimize} & \displaystyle\int_0^T\Big(g(t,y(t))+\int_{\mathbb{R}^m}h(t,s)\nu_t(\mathrm{d}s)\Big)\,\mathrm{d}t\\ \text{subject to} & \dfrac{\mathrm{d}y}{\mathrm{d}t}=G(t,y(t))+\displaystyle\int_{\mathbb{R}^m}H(t,s)\nu_t(\mathrm{d}s) \ \text{ for a.a. } t\in(0,T),\\ & y(0)=0,\\ & \mathrm{supp}(\nu_t)\subset S(t) \qquad \text{for a.a. } t\in(0,T),\\ & y\in W^{1,q}(0,T;\mathbb{R}^n), \quad \nu\in\mathcal{Y}(0,T;\mathbb{R}^m)\ . \end{cases}
$$

It is known (see e.g. Warga [20] or also [17]) that (RP) is actually a correct relaxation of (P) in the sense that (RP) always posseses a solution and min(RP) = inf(P). Moreover, if ν solves (RP) and $\nu=i(u)$ for some $u\in U_{\mathrm{ad}}$, then u solves (P).

Obviously, (RP) just represents minimization over $\bar{U}_{\rm ad}$ of the extended cost functional Φ defined by $\Phi(\nu) := \int_0^T \big(g(t,[y(\nu)](t)) + \int_{\mathbb{R}^m} h(t,s)\nu_t(\mathrm{d}s)\big)\,\mathrm{d}t$ with $y = y(\nu) \in W^{1,q}(0,T;\mathbb{R}^n)$ being the unique solution to the initial-value problem in (RP). To investigate the geometrical properties of Φ, we will calculate its Gâteaux differential with respect to the geometry comming from $L^1(0,T;C_0(\mathbb{R}^m))^* \supset U_{\rm ad}$. This is, in fact, a standard task undertaken within derivation of the Pontryagin maximum principle for the relaxed controls. This principle usually needs Fréchet differentiability with respect to y which can be guaranteed by the following assumptions on the partial derivative of g and G with respect to the variable r, denoted respectively by $g'(t,r)$ and $G'(t,r)$:

$$\text{(2.2a)} \qquad \exists a \in L^1(0,T) \quad \exists b : \mathbb{R} \to \mathbb{R} \text{ continuous} : \quad |g'(t,r)| \le a(t) + b(|r|),$$
$$|g'(t,r_1) - g'(t,r_2)| \le (a(t) + b(|r_1|) + b(|r_2|))|r_1 - r_2|,$$
$$\text{(2.2b)} \qquad \exists a \in L^{1+\varepsilon}(0,T) \quad \exists b : \mathbb{R} \to \mathbb{R} \text{ continuous} : \quad |G'(t,r)| \le a(t) + b(|r|),$$
$$|G'(t,r_1) - G'(t,r_2)| \le (a(t) + b(|r_1|) + b(|r_2|))|r_1 - r_2|,$$

The maximum principle involves the so-called adjoint equation

$$\text{(2.3)} \qquad \frac{\mathrm{d}\lambda}{\mathrm{d}t} \;=\; -\lambda(t)G'(t,y(t)) - g'(t,y(t)) \;, \qquad \lambda(T) = 0 \;.$$

The assumption (2.2) ensure that the terminal-value problem (2.3) possesses precisely one solution $\lambda \in W^{1,1}(0,T;\mathbb{R}^n)$. Likewise in a procedure by Gabasov and Kirillova [8; Section VII.2] or (for the general integral processes) by Schmidt [19] developed to prove sufficiency of the maximum principle for optimal control problems, we can establish the following incrementation formula.

Lemma 1. *Let (2.1) and (2.2) be satisfied, let $\nu, \tilde{\nu} \in \bar{U}_{\rm ad}$, $y, tildey \in W^{1,q}(0,T;\mathbb{R}^n)$ solve the initial-value problem in* (RP) *with ν and $\tilde{\nu}$ respectively, and let $\lambda \in W^{1,1}(0,T;\mathbb{R}^n)$ solve (2.3). Then*

$$\text{(2.4)} \qquad \Phi(\tilde{\nu}) - \Phi(\nu) \;=\; \int_0^T \int_{\mathbb{R}^m} [\lambda(t)H(t,s) + h(t,s)]\,[\tilde{\nu}_t - \nu_t](\mathrm{d}s)\mathrm{d}t$$
$$+ \int_0^T \Delta_g(t) + \lambda(t)\Delta_G(t)\mathrm{d}t$$

with the second-order correcting terms $\Delta_g(t)$ and $\Delta_G(t)$ defined by

$$\text{(2.5a)} \qquad \Delta_g(t) := g(t,\tilde{y}(t)) - g(t,y(t)) - g'(t,y(t))(\tilde{y}(t) - y(t)),$$
$$\text{(2.5b)} \qquad \Delta_G(t) := G(t,\tilde{y}(t)) - G(t,y(t)) - G'(t,y(t))(\tilde{y}(t) - y(t)).$$

Proof. Using the extended state equation both for y and $\tilde{y}$, the per-partes integration and the adjoint equation (2.3), we can calculate:

$$\Phi(\tilde{\nu}) - \Phi(\nu) \;-\; \int_0^T \int_{\mathbb{R}^m} [\lambda(t)H(t,s) + h(t,s)]\,[\tilde{\nu}_t - \nu_t](\mathrm{d}s)\mathrm{d}t$$

$$
\begin{aligned}
&= \int_0^T \Big(g(t,\tilde{y}(t)) - g(t,y(t)) - \int_{\mathbb{R}^m} \lambda(t) H(t,s)[\tilde{\nu}_t - \nu_t](\mathrm{d}s) \Big)\,\mathrm{d}t \\
&= \int_0^T g(t,\tilde{y}(t)) - g(t,y(t)) + \lambda(t) \left(G(t,\tilde{y}(t)) - G(t,y(t)) - \frac{\mathrm{d}(\tilde{y}(t)-y(t))}{\mathrm{d}t} \right) \mathrm{d}t \\
&= \int_0^T g(t,\tilde{y}(t)) - g(t,y(t)) + \lambda(t)(G(t,\tilde{y}(t)) - G(t,y(t))) + \frac{\mathrm{d}\lambda}{\mathrm{d}t}(\tilde{y}(t)-y(t))\mathrm{d}t \\
&= \int_0^T \Big(g(t,\tilde{y}(t)) - g(t,y(t)) - g'(t,y(t))(\tilde{y}(t)-y(t)) \\
&\qquad\qquad + \lambda(t)(G(t,\tilde{y}(t)) - G(t,y(t)) - G'(t,y(t))(\tilde{y}(t)-y(t))) \Big)\,\mathrm{d}t \\
&=: \int_0^T \Delta_g(t) + \lambda(t)\Delta_G(t)\mathrm{d}t.
\end{aligned}
$$

□

The formula (2.4) enables us to investigate concavity of the extended cost functional Φ. Let us take a sufficiently large radius R so that $|[y(u)](t)| \leq R$ for any $u \in U_{\mathrm{ad}}$ and any $t \in (0,T)$, where $y(u) \in W^{1,q}(0,T;\mathbb{R}^n)$ denotes the unique solution to the initial-value problem in (P). Concretely, let us take

$$
R = \sup_{u \in U_{\mathrm{ad}}} \sup_{t \in (0,T)} \Big| [y(u)](t) \Big| \, . \tag{2.6}
$$

Furthermore, let $a(t) := \sup_{|r|\leq R} |g'(t,r)|$ and $A(t) := \sup_{|r|\leq R} |G'(t,r)|$. As (2.2) ensures $a, A \in L^1(0,T)$, we can define

$$
b(t) := \int_t^T a(\tau)\mathrm{d}\tau \ , \quad B(t) := \int_t^T A(\tau)\mathrm{d}\tau \ . \tag{2.7}
$$

Lemma 2. *Let us assume (2.1), (2.2) and $G(t,\cdot)$ twice continuously differentiable, and let $g(t,\cdot)$ be uniformly concave on the ball of the radius R from (2.6) in the sense*

$$
\begin{aligned}
\forall r, \tilde{r} \in \mathbb{R}^n : \quad & \max(|r|,|\tilde{r}|) \leq R \implies \\
& g(t,\tilde{r}) - g(t,r) - g'(t,r)(\tilde{r}-r) \leq -\alpha(t)|\tilde{r}-r|^2
\end{aligned} \tag{2.8}
$$

with the modulus $\alpha \geq 0$ satisfying, for b and B defined by (2.7), the condition

$$
\alpha(t) \geq \frac{b(t)}{2} \mathrm{e}^{B(t)} \sup_{|r|\leq R} |G''(t,r)| \ . \tag{2.9}
$$

Then Φ is concave on $\bar{U}_{\mathrm{ad}}$.

Proof. From the adjoint equation (2.3) we can estimate $\mathrm{d}|\lambda|/\mathrm{d}t \leq A(t)|\lambda(t)| + a(t)$ so that by the Gronwall inequality one gets

$$
|\lambda(t)| \ \leq \ \left(\int_t^T a(\tau) \mathrm{e}^{-\int_t^T A(\theta)\mathrm{d}\theta} \mathrm{d}\tau \right) \mathrm{e}^{\int_t^T A(\tau)\mathrm{d}\tau} \ . \tag{2.10}
$$

To simplify the notation, we can also (a bit more pessimistically) estimate

$$|\lambda(t)| \;\leq\; b(t)\mathrm{e}^{B(t)}\;. \tag{2.11}$$

By the Taylor expansion, we can estimate $|G(t,\tilde{y}(t))-G(t,y(t))-G'(t,y(t))(\tilde{y}(t)-y(t))| \leq \sup_{|r|\leq R}\frac{1}{2}|G''(t,r)|\,|\tilde{y}(t)-y(t)|^2$. Then (2.8) with (2.9) and (2.11) ensure the following inequality

$$\begin{aligned}\Delta_g(t)+\lambda(t)\Delta_G(t) &\leq -\alpha(t)|\tilde{y}(t)-y(t)|^2+\frac{1}{2}|\lambda(t)|\;|G''(t,y(t))|\;|\tilde{y}(t)-y(t)|^2\\ &\leq \left(-\alpha(t)+\frac{b(t)}{2}\mathrm{e}^{B(t)}\sup_{|r|\leq R}\frac{1}{2}|G''(t,r)|\right)|\tilde{y}(t)-y(t)|^2\leq 0.\end{aligned} \tag{2.12}$$

so that the second right-hand term in (2.4) is non-positive. Since the first right-hand term in (2.4) represents just the Gâteaux differential of Φ, i.e. $[\nabla\Phi(\nu)](\tilde{\nu}-\nu)=\int_0^T\int_{\mathbb{R}^m}[\lambda(t)H(t,s)+h(t,s)]\,[\tilde{\nu}_t-\nu_t](\mathrm{d}s)\mathrm{d}t$, we obtained

$$\forall \nu,\tilde{\nu}\in\bar{U}_{\mathrm{ad}}:\qquad \Phi(\tilde{\nu})-\Phi(\nu)-[\nabla\Phi(\nu)](\tilde{\nu}-\nu)\leq 0\;, \tag{2.13}$$

which just says that Φ is concave on $\bar{U}_{\mathrm{ad}}$. □

Lemma 3. *If $\nu\in\bar{U}_{\mathrm{ad}}$ is an extreme point of $\bar{U}_{\mathrm{ad}}$ (i.e. $\nu=\frac{1}{2}\nu^1+\frac{1}{2}\nu^2$ for some $\nu^1,\nu^2\in\bar{U}_{\mathrm{ad}}$ implies $\nu_1=\nu_2$), then $\nu=i(u)$ for some $u\in U_{\mathrm{ad}}$.*

Sketch of the proof. If ν_t were not a Dirac mass for a.a. $t\in(0,T)$, then ν_t would not be an extreme point in the set of probability measures on $S(t)$, and thus the multivalued mapping $C:(0,T)\rightrightarrows \mathrm{rca}(\mathbb{R}^m)^2$ defined by

$$\begin{aligned}C(t):=\Big\{\;&(\mu^1,\mu^2)\in\mathrm{rca}(\mathbb{R}^m)^2;\\ &\mu^1,\mu^2\geq 0,\;\mu^1(S(t))=1=\mu^2(S(t)),\;\frac{1}{2}\mu^1+\frac{1}{2}\mu^2=\nu_t\;\Big\}\end{aligned} \tag{2.14}$$

is not a sigleton for a.a. $t\in(0,T)$. Since S is measurable, ν is weakly measurable, and $C_0(\mathbb{R}^m)$ is separable, the multivalued mapping C is measurable. Then this mapping admits a measurable selection $t\to(\nu_t^1,\nu_t^2)$ which is not equal to (ν_t,ν_t) for a.a. $t\in(0,T)$. This shows that $\nu^1\neq\nu^2$ but $\nu=\frac{1}{2}\nu^1+\frac{1}{2}\nu^2$ so that ν is not an extreme point in $\bar{U}_{\mathrm{ad}}$, a contradiction. □

Proposition 1. *Let (2.1), (2.2), and (2.8)–(2.9) be satisfied. Then* (P) *has an optimal solution.*

Proof. The relaxed problem (RP) consists in minimization of the weakly* continuous functional Φ on the convex weakly* compact set $\bar{U}_{\mathrm{ad}}$ so that it certainly has a solution by the standard compactness arguments. By Lemma 2, Φ is concave, so that by

Bauer's extremal principle [2] at least one solution to (RP) is an extreme point of $\bar{U}_{\rm ad}$. By Lemma 3, this extreme optimal control for (RP) can be represented by the ordinary control from $U_{\rm ad}$ which is obviously optimal also for (P). □

Remark. If the controlled system is linear with respect to the state, i.e. $G(t,\cdot)$ is affine, then obviously $G'' \equiv 0$ and one can take $\alpha \equiv 0$ in (2.8) which then just requires $g(t,\cdot)$ to be concave.

3. A generalization: unbounded controls.

If the multivalued mapping $S : (0,T) \rightrightarrows \mathbb{R}^m$ acting as control constraints in (P) is not bounded, several sophisticated approaches must be still incorporated. First, we must suppose certain (for simplicity polynomial) coercivity of the problem, say:

$$c_0|s|^p \le g(t,r) + h(t,s) \le a_0(t) + C_1|s|^p \ , \tag{3.1}$$

$$|G(t,r) + H(t,s)| \le a_1(t) + b_1|r| + c_1|s|^{p-\varepsilon} \tag{3.2}$$

for some $a_0 \in L^1(0,T)$, $a_1 \in L^q(0,T)$, $c_0, C_0, b_1, c_1 \in \mathbb{R}$, $p \in [1,+\infty)$, $q \in (1,+\infty)$, and $\varepsilon > 0$. Then the modified formulation of the original problem (P) involves $L^p(0,T;\mathbb{R}^m)$ in place of $L^\infty(0,T;\mathbb{R}^m)$. As to the measurable multivalued mapping S, it is now natural to suppose it again in the form $S(t) := M(t,S_0)$ with some $S_0 \in \mathbb{R}^m$ closed and the Carathéodory mapping M satisfying $\max(|M(t,s)|, |M(t,s)^{-1}|) \le a(t) + b|s|$ for some $a \in L^p(0,T)$ and $b \in \mathbb{R}$. Note that this growth condition makes the Nemytskiĭ mapping $\mathcal{N}_M$ generated by M a homeomorphism on $L^p(\Omega;\mathbb{R}^m)$ whose inverse transforms $U_{\rm ad}$ onto the set $\mathcal{N}_M^{-1}(U_{\rm ad}) = \{u \in L^p(0,T;\mathbb{R}^m);\ u(t) \in S_0 \text{ for a.a. } t \in (0,T)\}$ which uses a fixed constraint S_0.

Then the correct relaxed problem looks as (RP) but with $\mathcal{Y}(0,T;\mathbb{R}^m)$ replaced by $\mathcal{Y}^p(0,T;\mathbb{R}^m)$ defined as

$$\mathcal{Y}^p(\Omega;\mathbb{R}^m) \ = \ \Big\{\nu \in \mathcal{Y}(0,T;\mathbb{R}^m);\ \ \textstyle\int_0^T \int_{\mathbb{R}^m} |s|^p \nu_t(\mathrm{d}s)\mathrm{d}t < +\infty\Big\}.$$

This set contains just those Young measures, called L^p-Young measures, that can be attained by sequences bounded in $L^p(0,T;\mathbb{R}^m)$; cf. [17]. The quite nontrivial fact that such (RP) has a solution and $\inf(\mathsf{P}) = \min(\mathsf{RP})$ relies on a nonconcetration of energy of any minimizing sequence $\{u_k\}_{k\in\mathbb{N}}$ for (P), i.e. the relative weak compactness of the set $\{|u_k|^p;\ k \in \mathbb{N}\}$ in $L^1(0,T)$; cf. [17].

By the assumed coercivity (3.1), all solutions to (RP) must belong to the set

$$\mathcal{Y}^p_{\varrho_0}(0,T;\mathbb{R}^m) = \Big\{\nu \in \mathcal{Y}(0,T;\mathbb{R}^m);\ \ \Big(\textstyle\int_0^T \int_{\mathbb{R}^m} |s|^p \nu_t(\mathrm{d}s)\mathrm{d}t\Big)^{1/p} \le \varrho_0\Big\}$$

for some $\varrho_0 \in \mathbb{R}$ sufficiently large.

Proposition 2. *Let (2.1c), (2.2), (3.1), (3.2), and (2.8)–(2.9) be satisfied for the radius*

$$R = \sup_{u\in U_{\rm ad},\ \|u\|_{L^p(0,T;\mathbb{R}^m)} \le 2^{1/p}\varrho_0} \|y(u)\|_{C(0,T;\mathbb{R}^n)}\ . \tag{3.3}$$

Then (P) *has an optimal solution.*

Proof. Let us consider a problem $(\mathsf{RP})_\varrho$ with $\mathcal{Y}^p_\varrho(0,T;\mathbb{R}^m)$ in place of $\mathcal{Y}^p(0,T;\mathbb{R}^m)$. By the coercivity (3.1), $(\mathsf{RP})_\varrho$ has the same set of solutions as (RP) provided $\varrho \geq \varrho_0$. By (2.8) valid for R satisfying (3.3), the problem $(\mathsf{RP})_{2^{1/p}\varrho_0}$ has the concave cost functional Φ so that, by Bauer's extremal principle, it must admit at least one optimal control ν which is an extreme point in the convex weakly* compact set $\mathcal{Y}^p_{2^{1/p}\varrho_0}(0,T;\mathbb{R}^m)$. By the coercivity, we also know that even $\nu \in \mathcal{Y}^p_{\varrho_0}(0,T;\mathbb{R}^m)$.

However, then $\nu = i(u)$ for some $U_{\rm ad}$. Indeed, in the opposite case the Young measures $\nu^1, \nu^2 \in \mathcal{Y}(0,T;\mathbb{R}^m)$ obtained in the proof of Lemma 3 would also satisfy the estimate

$$\int_0^T\!\!\int_{\mathbb{R}^m} |s|^p \nu^1_t({\rm d}s){\rm d}t = \int_0^T\!\!\int_{\mathbb{R}^m} |s|^p (2\nu_t - \nu^2_t)({\rm d}s){\rm d}t \leq 2\int_0^T\!\!\int_{\mathbb{R}^m} |s|^p \nu_t({\rm d}s){\rm d}t \leq 2\varrho_0^p$$

because $\frac{1}{2}\nu^1_t + \frac{1}{2}\nu^2_t = \nu_t$. In other words, $\nu^1 \in \mathcal{Y}^p_{2^{1/p}\varrho_0}(0,T;\mathbb{R}^m)$. Replacing the role of ν^1 and ν^2, we get $\nu^2 \in \mathcal{Y}^p_{2^{1/p}\varrho_0}(0,T;\mathbb{R}^m)$, as well. This would show that ν is not an extreme point in $\mathcal{Y}^p_{2^{1/p}\varrho_0}(0,T;\mathbb{R}^m)$, a contradiction.

Then obviously u is the sought optimal control for (P). □

4. A generalization: distributed-parameter problems.

The presented method readily extends for distributed-parameter controlled systems which we want to illustrate here briefly for systems governed by elliptic partial differential equations; for Fredholm integral equations see [18]. In fact, the only peculiarity is that both the state y and the adjoint state λ are required to be bounded in L^∞-norm, which may require certain additional regularity.

We will consider a bounded convex domain $\Omega \subset \mathbb{R}^n$, $n \leq 3$, with a Lipschitz boundary $\partial\Omega$ and the following optimal cotrol problem for one elliptic equation with homogeneous Dirichlet boundary conditions:

$$(\mathsf{P}')\quad \begin{cases} \text{Minimize} & \displaystyle\int_\Omega g(x,y(x)) + h(x,u(x))\ {\rm d}x & \text{(cost functional)} \\ \text{subject to} & {\rm div}(\nabla y(x)) = G(x,y(x)) + H(x,u(x)) & \text{(state equation)} \\ & y|_{\partial\Omega} = 0, & \text{(boundary condition)} \\ & u(x) \in S(x) \quad \text{for } x \in \Omega, & \text{(control constraints)} \\ & y \in W^{1,2}(\Omega), \quad u \in L^p(\Omega;\mathbb{R}^m)\ , & \end{cases}$$

where $g, G : \Omega \times \mathbb{R} \to \mathbb{R}$, $h, H : \Omega \times \mathbb{R}^m \to \mathbb{R}$ are Carathéodory functions satisfying (3.1)–(3.2) and additionally

$$G(x,\cdot) \ \text{ nondecreasing for a.a. } x \in \Omega \tag{4.1}$$

$$|G(x,r)| \leq a(x) + b|r|^{c/q}\ , \quad |H(x,s)| \leq a(x) + b|s|^{p/q} \tag{4.2}$$

for some $a \in L^q(\Omega)$, $b \in \mathbb{R}$, and $q \geq 1$ and $c < +\infty$ such that $q > n/2$ and, if $n = 3$, also $c \leq 2n/(n-2)$. Note that the involved boundary-value problem has, for any control $u \in L^p(\Omega;\mathbb{R}^m)$, a unique weak solution $y = y(u) \in W_0^{1,2}(\Omega)$.

Besides, the multivalued mapping $S : \Omega \rightrightarrows \mathbb{R}^m$ is subjected to the same assumptions as in Section 2 with Ω in place of $(0,T)$.

Let us note that, for u ranging a bounded set in $L^p(\Omega;\mathbb{R}^m)$, $y(u)$ ranges a bounded set in $W^{1,2}(\Omega) \subset L^c(\Omega)$. By (4.2), $G(x,y(u)) + H(x,u)$ then ranges a bounded set in $L^q(\Omega) \subset W^{-1+\varepsilon,2}(\Omega)$ for $\varepsilon \in [0,1]$ such that $q > 2n/(n+2-2\varepsilon)$; the last imbedding is just adjoint to $L^q(\Omega)^* \equiv L^{q/(q-1)}(\Omega) \supset W_0^{1-\varepsilon,2}(\Omega) \equiv W^{-1+\varepsilon,2}(\Omega)^*$. Moreover, the standard elliptic regularity (see Grisward [9]) shows that $y(u)$ is bounded $W^{2,2}(\Omega)$ provided $G(x,y(u)) + H(x,u)$ ranges a bounded set in $L^2(\Omega)$. By interpolation for the linear operator $f \mapsto y$ with y solving the problem $\mathrm{div}(\nabla u) = f$ and $y|_{\partial\Omega} = 0$, the solution $y(u)$ ranges a bounded set in $W^{1+\varepsilon,2}(\Omega)$ so that, using a well-known imbedding theorem, $W^{1+\varepsilon,2}(\Omega) \subset C^0(\Omega)$ provided $\varepsilon > (n-2)/2$, and thus the solution $y(u)$ lives in $C^0(\Omega)$, which is essential for our theory. Note that $1 \geq \varepsilon > (n-2)/2$ and $q > 2n/(n+2-2\varepsilon)$ yield respectively the mentioned restrictions $n \leq 3$ and $q > n/2$.

Again, we make relaxation by a continuous extension of (P') from U_{ad} on $\bar{U}_{\mathrm{ad}}$ (defined as previously but with Ω in place of $(0,T)$):

$$(\mathsf{RP}') \qquad \begin{cases} \text{Minimize} & \displaystyle\int_\Omega \Big(g(x,y(x)) + \int_{\mathbb{R}^m} h(x,s)\nu_x(\mathrm{d}s) \Big)\,\mathrm{d}x \\ \text{subject to} & \mathrm{div}(\nabla y(x)) = G(x,y(x)) + \displaystyle\int_{\mathbb{R}^m} H(x,s)\nu_x(\mathrm{d}s)\ , \\ & \mathrm{supp}(\nu_x) \subset S(x) \qquad \text{for a.a. } x \in \Omega, \\ & y \in W_0^{1,2}(\Omega), \quad \nu \in \mathcal{Y}^p(\Omega;\mathbb{R}^m)\ . \end{cases}$$

Using Green's formula instead of per-partes integration, as in Lemma 1 one can derive the incrementation formula for $\Phi(\nu) := \int_\Omega (g(x,[y(\nu)](x)) + \int_{\mathbb{R}^m} h(x,s)\nu_x(\mathrm{d}s))\,\mathrm{d}x$ with $y = y(\nu) \in W_0^{1,2}(\Omega)$ satisfying $\mathrm{div}(\nabla y(x)) = G(x,y(x)) + \int_{\mathbb{R}^m} H(x,s)\nu_x(\mathrm{d}s)$ in the weak sense, which now looks as

$$\begin{aligned} \Phi(\tilde{\nu}) - \Phi(\nu) \ = \ & \int_\Omega \int_{\mathbb{R}^m} [\lambda(x)H(x,s) + h(x,s)]\,[\tilde{\nu}_x - \nu_x](\mathrm{d}s)\mathrm{d}x \\ & + \int_\Omega \Delta_g(x) + \lambda(x)\Delta_G(x)\mathrm{d}x \end{aligned}$$

with the adjoint state $\lambda \in W_0^{1,2}(\Omega)$ satisfying (in the weak sense)

$$\mathrm{div}(\nabla\lambda(x)) = G'(x,y(x))\lambda + g'(x,y(x)) \tag{4.3}$$

We will assume, beside (2.2), the following growth conditions on G' and g':

$$|G'(x,r)| \leq b(|r|) \qquad |g'(x,r)| \leq a(x) + b(|r|) \tag{4.4}$$

with some $a \in L^q(\Omega)$ and $b : \mathbb{R} \to \mathbb{R}$ continuous. Note that, by similar regularity arguments as used for y, the adjoint state λ lives in $W^{1+\varepsilon,2}(\Omega) \subset C^0(\Omega)$, which allows us to establish the criterion (4.5) below.

Proposition 3. *Let (2.2), (2.8), (3.1), (3.2), (4.1)–(4.4) be fulfilled. Then* (P') *has an optimal solution provided the uniform-concavity coefficient* α *of* g *satisfies*

$$\alpha(x) \;\geq\; \frac{1}{2} \sup_{\substack{u\in U_{\rm ad}\\ \|u\|_{L^p(\Omega;\mathbb{R}^m)}\leq 2^{1/p}\varrho_0}} \|\lambda(u)\|_{C^0(\Omega)} \sup_{|r|\leq R} |G''(x,r)| \tag{4.5}$$

with $R = \sup_{u\in U_{\rm ad},\ \|u\|_{L^p(\Omega;\mathbb{R}^m)}\leq 2^{1/p}\varrho_0} \|y(u)\|_{C^0(\Omega)}$, *where* ϱ_0 *is a sufficiently large radius of the ball in* $L^p(\Omega;\mathbb{R}^m)$ *where every minimizing sequence for* (P') *eventually lives.*

Acknowledgements. This research was partly covered by the grant No. 201/96/0228 of the Grant Agency of the Czech Republic. The author also appreciates his visit at Ernst-Moritz-Arndt-Universität in Greifswald and useful discussions with Professor Werner H. Schmidt.

References.

1. E.J. Balder, *New existence results for optimal controls in the absence of convexity: the importance of extremality.* SIAM J. Control Optim. **32** (1994), 890–916.
2. H. Bauer, *Minimalstellen von Funktionen und Extremalpunkte.* Archiv d. Math. **9** (1958), 389–393, **11** (1960), 200–205.
3. H. Berliocchi, J.-M. Lasry, *Intégrandes normales et mesures paramétrées en calcul des variations.* Bull. Soc. Math. France **101** (1973), 129–184.
4. C. Castaing, M. Valadier, *Convex Analysis and Measurable Multifunctions.* Lecture Notes in Math. **580**, Springer, Berlin, 1977.
5. A. Cellina, G. Colombo, *On a classical problem of the calculus of variations without convexity assumptions.* Annales Inst. H. Poincaré, Anal. Nonlin. **7** (1990), 97–106.
6. L. Cesari, *An existence theorem without convexity conditions.* SIAM J. Control **12** (1974), 319–331.
7. L. Cesari, *Optimization: Theory and Applications.* Springer, New York, 1983.
8. R. Gabasov, F. Kirillova, *Qualitative Theory of Optimal Processes.* Nauka, Moscow, 1971.
9. P. Grisvard, *Elliptic problems in nonsmooth domains.* Pitman, Boston, 1985.
10. J. Macki, A. Strauss, *Introduction to Optimal Control Theory.* Springer, New York, 1982.
11. C. Mariconda, *On a parametric problem of the calculus of variations without convexity assumptions.* J. Math. Anal. Appl. **170** (1992), 291–297.
12. L.W. Neustadt, *The existence of optimal controls in the absence of convexity conditions.* J. Math. Anal. Appl. **7** (1963), 110–117.
13. C. Olech, *Integrals of set-valued functions and linear optimal control problems.* Clloque sur la Théorie Math. du Contrôle Optimal, C.B.R.M., Vander Louvain, 1970, pp.109–125.
14. J.-P. Raymond, *Existence theorems in optimal control theory without convexity assumptions.* J. Optim. Theory Appl. **67** (1990), 109–132.
15. J.-P. Raymond, *Existence theorems without convexity assumptions for optimal control problems governed by parabolic and elliptic systems.* Appl. Math. Optim. **26** (1992), 39–62.
16. J.-P. Raymond, *Existence and uniqueness results for minimization problems with non-convex functionals.* J. Optim. Theory Appl. **82** (1994), 571–592.
17. T. Roubíček, *Relaxation in Optimization Theory and Variational Calculus.* W. de Gruyter, Berlin, 1997.
18. T. Roubíček, W.H. Schmidt, *Existence of solutions to nonconvex optimal control problems governed by nonlinear Fredholm integral equations.* (submitted)
19. W.H. Schmidt, *Maximum principles for processes governed by integral equations in Banach spaces as sufficient optimality conditions.* Beiträge zur Analysis **17** (1981), 85–93.
20. J. Warga, *Optimal Control of Differential and Functional Equations.* Academic Press, New York, 1972.
21. L.C. Young, *Generalized curves and the existence of an attained absolute minimum in the calculus of variations.* Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, Classe III **30** (1937), 212–234.

Tomáš Roubíček
Mathematical Institute, Charles University, Sokolovská 83, CZ-186 00 Praha 8
e-mail: `roubicek@karlin.mff.cuni.cz`