EQUADIFF 4

Rolf Klötzler

On a general conception of duality in optimal control

In: Jiří Fábera (ed.): Equadiff IV, Czechoslovak Conference on Differential Equations and Their Applications. Proceedings, Prague, August 22-26, 1977. Springer-Verlag, Berlin, 1979. Lecture Notes in Mathematics, 703. pp. [189]--196.

Persistent URL: http://dml.cz/dmlcz/702219

ON A GENERAL CONCEPTION OF DUALITY IN OPTIMAL CONTROL

R. Klötzler, Leipzig

Many problems in the theory of differential equations and its applications can be formulated as problems of optimal control. For these problems again several conceptions of duality have been developed which are very useful from theoretical and numerical point of view. For example we all know in the theory of elasticity the important duality between the principle of Dirichlet and its dual problem as the principle of Castigliano.

In general, if we denote the original problem by

(1) $F(x) \longrightarrow \text{Min}$

subject to all $x \in X$,

then a dual problem is defined in general sense by any problem

(2) $L(y) \longrightarrow \text{Max}$

subject to all $y \in Y$,

with the property $F(x) \geqq L(y) \quad \forall\, x \in X,\ y \in Y$.

As a rule one aspires to construct such dual problems which satisfy the strong duality condition

$$\inf_X F \text{ (or } \min_X F \text{) } = \sup_Y L \text{ (or } \max_Y L \text{) } .$$

It is easily seen that such a conception of duality leads to both-side estimates of $\inf_X F$ and often also to corresponding error estimates with respect to an optimal solution x_0 .

For regular variational problems already K.O. Friedrichs [3] introduced dual variational problems in 1928. His theory requires besides assumptions of differentiability mainly convexity properties of the integrand. In the last decade by M.M. Cvetanov [9] , R.T. Rockafellar [8] and Ekeland/Temam [2] several investigations were stated, which may be viewed as an extension of the original conception of Friedrichs with respect to general problems of optimal control. In these papers the former assumptions of differentiability are essentially weakened, however convexity properties are again supposed and instead of Legendre transformation by Friedrichs now Fenchel's theory of conjugate functions is applied.

In the present paper we shall delineate a new conception of duality, which avoids any requirements on the convexity of the original problem. Simultaneously this treatment carries on relevant investigations on Bellman's differential equation and extensions of the classical theory of Hamilton and Jacobi by the author [5] - [7].

We consider problems of optimal control of the type

$$(3)\qquad J(x,u) := \int_{\Omega} f(t,x,u)\,dt + \int_{\partial\Omega} l(t,x)\,do \longrightarrow \mathrm{Min}$$

subject to all vector-valued state functions $x \in X$, control functions $u \in U(x)$, and constraints

$$(4)\qquad x^i_{t_\alpha} = g^i_\alpha(t,x,u) \qquad (i = 1,\dots,n;\ \alpha = 1,\dots,m)\ .$$

Here Ω is a strongly Lipschitz domain of R^m,

$$X = \left\{ x \in W_p^{1,n}(\Omega) \;\middle|\; (t,x(t)) \in \bar{G} \text{ on } \bar{\Omega}\ ,\ b(t,x(t)) = \sigma \text{ on } \partial\Omega \right\}$$

with $p > m$,

$$U(x) = \left\{ u \in L_p^r(\Omega) \;\middle|\; u(t) \in V(t,x(t)) \subset R^r \text{ a.e. on } \Omega \right\}$$

for every $x \in X$,

G is an open set of R^{n+m}, and $V(.,.)$ is assumed to be a normal map from $\bar{G}$ into R^r in the sense of Joffe/Tichomirov [4] p.338. Further we suppose l and b are real continuous functions on $\partial\Omega \times R^n$ and f as well as g^i_α are real functions on $\bar{G} \times R^r$ satisfying the Carathéodory condition in the following meaning: they are (Lebesgue-) measurable functions with respect to the first argument t and continuous functions for almost every fixed $t \in \Omega$. Therefore $f(.,x(.),u(.))$ and $g^i_\alpha(.,x(.),u(.))$ are measurable functions on Ω for every <u>process</u> $\langle x,u \rangle$, that means for every admissible pair $\langle x,u \rangle$ of problem (3). We denote the set of all processes by $\mathcal{P}$ and require the following additional assumptions:

$$(5a)\qquad \mathcal{P} \neq \emptyset$$

(5b) $f(.,x(.),u(.))$ is minorized by a function $\psi \in L_1^1(\Omega)$ $\forall \langle x,u \rangle \in \mathcal{P}$.

In consequence of (5b) $f(.,x(.),u(.))$ is summable on Ω (in the broad sense) for every $\langle x,u\rangle \in \mathcal{P}$ and accordingly $J(x,u)$ is well-defined on $\mathcal{P}$.

Now we prepare the formulation of a corresponding duality principle. For this purpose we introduce the denotations

H for the <u>Pontryagin function</u> defined by

(6) $$H(t,\xi,v,y) := -f(t,\xi,v) + y_i^\alpha g_\alpha^i(t,\xi,v) \quad ,$$

$\mathcal{H}$ for the <u>Hamiltonian function</u> defined by

(7) $$\mathcal{H}(t,\xi,y) := \sup_{v\in V(t,\xi)} H(t,\xi,v,y) \qquad \text{on } \bar{G}\times R^{nm} \quad ,$$

and $Q(t)$ for the following cuts of $\bar{G}$

(8) $$Q(t) := \begin{cases} \{\xi\in R^n \mid (t,\xi)\in\bar{G}\} & \forall\ t\in\Omega \\ \{\xi\in R^n \mid (t,\xi)\in\bar{G}\ ,\ b(t,\xi)=\sigma\} & \forall\ t\in\partial\Omega\ . \end{cases}$$

Moreover we select a subset $\mathcal{Y}$ from $W^{1,m}(G)$ consisting of all functions $S\in W^{1,m}_\infty(G)$ having the following properties:

(9a) each class of distribution derivatives of S^α ($\alpha=1,\dots,m$) contains a bounded representative S_j^α ($j=1,\dots,n+m$) ;

(9b) there are uniformly bounded sequences of functions $z_k^\alpha\in C^1(R^{n+m})$ and their derivatives satisfying pointwise the conditions

$$\lim_{k\to\infty} z_k^\alpha = S^\alpha \text{ and } \lim_{k\to\infty} (\partial z_k^\alpha/\partial\xi^j) = S_j^\alpha \quad \text{on } \bar{G}\ .$$

Obviously $\mathcal{Y}$ contains the set $C^{1,m}(\bar{G})$ of continuous differentiable (vector-valued) functions. By means of mollified functions of S^α we can easily see that also the set $D^{1,m}(\bar{G})$ of piecewise continuous differentiable functions belongs to $\mathcal{Y}$ .In the following text we denote generally $S_i^\alpha = S_{t^i}^\alpha$ for $i=1,\dots,m$ and $S_{i+m}^\alpha = S_{\xi^i}^\alpha$ for $i=1,\dots,n$.

With these definitions we are starting from theorems on set-valued functions by Joffe/Tichomirov [4] ch. 8 and conclude the following lemmas.

<u>Lemma</u> 1 . $\mathcal{H}(.,.,.)$ is a measurable function on $\bar{G}\times R^{nm}$. If dom $\mathcal{H}(t,\xi,.)\neq\emptyset$,then $\mathcal{H}(t,\xi,.)$ is a convex function .

Lemma 2 . Setting $y_i^\alpha = S^\alpha_{\xi^i}(t,\xi)$ in (7) for arbitrary functions $S \in \mathcal{Y}$ we obtain $\mathcal{H}(.,.,S_\xi(.,.))$ as a measurable function on G .

Lemma 3 . For each $S \in \mathcal{Y}$ and

$$(10) \qquad \delta_S(t,\xi) := S^\alpha_{t^\alpha}(t,\xi) + \mathcal{H}(t,\xi,S_\xi(t,\xi))$$

the function

$$(11) \qquad \Lambda_S(.) := \sup_{\xi \in Q(.)} \delta_S(.,\xi)$$

is measurable on Ω .

In connection with the general equation of Hamilton-Jacobi

$$\phi^\alpha_{t^\alpha} + \mathcal{H}(t,\xi,\phi_\xi) = 0$$

it seems adequate to define this function δ_S as the "defect" of the Hamilton-Jacobi equation with respect to S .

Now we fix a process $\langle x,u \rangle$ and a function $S \in \mathcal{Y}$. We obtain by using the expression (6) the equation

$$J(x,u) = \int_{\partial\Omega} l(t,x)\, do + \int_{\Omega} \left\{ -H(t,x,u,S_\xi(t,x)) + S^\alpha_{\xi^i}(t,x)\, g^i_\alpha(t,x,u) \right\} dt \quad .$$

Furthermore, we insert in its second integrand $S^\alpha_{t^\alpha}(t,x) - S^\alpha_{t^\alpha}(t,x)$. Because of the fact that in consequence of (9) and (4)

$$\int_\Omega \left[S^\alpha_{t^\alpha}(t,x) + S^\alpha_{\xi^i}(t,x)\, g^i_\alpha(t,x) \right] dt$$

$$= \int_\Omega \left[S^\alpha_{t^\alpha}(t,x) + S^\alpha_{\xi^i}(t,x)\, x^i_{t^\alpha}(t) \right] dt$$

$$= \lim_{k\to\infty} \int_\Omega \left[z^\alpha_{kt^\alpha}(t,x) + z^\alpha_{k\xi^i}(t,x)\, x^i_{t^\alpha}(t) \right] dt$$

$$= \lim_{k\to\infty} \int_\Omega \frac{d\, z^\alpha_k(t,x(t))}{dt^\alpha}\, dt$$

$$= \lim_{k\to\infty} \int_{\partial\Omega} z^\alpha_k(t,x)\, n_\alpha(t)\, do = \int_{\partial\Omega} S^\alpha(t,x)\, n_\alpha(t)\, do$$

holds, the following equality results:

$$J(x,u) = \int_{\Omega} \{-H(t,x,u,S_{\xi}(t,x)) + S^{\alpha}_{t^{\alpha}}(t,x)\} \, dt + \int_{\partial\Omega} [S^{\alpha}(t,x)\, n_{\alpha}(t) + l(t,x)] \, do \; .$$

Here the symbols $n_{\alpha}(t)$ $(\alpha = 1,\dots,m)$ denote the components of the unit vector of the exterior normal on $\partial\Omega$ at the point t. If we observe that in consequence of (5),(7),Lemma 2 and $\mathcal{H}(t,x,S_{\xi}(t,x)) \geqq H(t,x,u,S_{\xi}(t,x)) \in L^1_1(\Omega)$ the function $\mathcal{H}(t,x,S_{\xi}(t,x))$ is summable too on Ω ,the last equation leads to the estimate

$$J(x,u) \geqq \int_{\Omega} \{-\mathcal{H}(t,x,u,S_{\xi}(t,x)) + S^{\alpha}_{t^{\alpha}}(t,x)\} \, dt + \int_{\partial\Omega} [S^{\alpha}(t,x)\, n_{\alpha}(t) + l(t,x)] \, do$$

and through Lemma 3 to

$$\text{(12)} \qquad J(x,u) \geqq L(S) := -\int_{\Omega} \Lambda_S(t)\, dt + \int_{\partial\Omega} \inf_{\xi \in Q(t)} [S^{\alpha}(t,\xi) n_{\alpha}(t) + l(t,\xi)] \, do \; .$$

From this development of formula (12) it is easy to notice in which cases there the equality occurs.We summarize our results in the following duality theorem.

Theorem 1 . Let $\langle x,u\rangle$ be a process and $S \in \mathcal{Y}$. Then $J(x,u) \geqq L(S)$ in the sense of the detailed formulation of (12). Here the equality holds if and only if the following conditions are fulfilled:

$$\text{(13a)} \quad H(t,x,u,S_{\xi}(t,x)) = \mathcal{H}(t,x,S_{\xi}(t,x)) \text{ a.e. on } \Omega \; ,$$

$$\text{(13b)} \quad \delta_S(t,x(t)) = \Lambda_S(t) \quad \text{a.e. on } \Omega \; ,$$

$$\text{(13c)} \quad S^{\alpha}(t,x)n_{\alpha}(t) + l(t,x) = \inf_{\xi \in Q(t)} [S^{\alpha}(t,\xi)n_{\alpha}(t) + l(t,\xi)] \quad \text{a.e. on } \partial\Omega \; .$$

In virtue of ths Theorem 1 a duality is defined between the original problem (3) and its dual problem $L(S) \to \text{Max}$ on $\mathcal{Y}$. This duality is a far-reaching generalization of several conceptions of duality which we cited above in the introduction.We can easily demonstrate that through a reduction of problem (3) to a Bolza problem the dual functional of Friedrichs and Rockafellar is generated by L(S) under the special statement

$$(14) \qquad S^{\alpha}(t,\xi) = y_0^{\alpha}(t) + y_i^{\alpha}(t)\xi^i \qquad (\alpha = 1,\dots,m) .$$

Hence the duality of Friedrichs, Cvetanov, Rockafellar and Ekeland/Temam is formally included in our conception (12) by specialization on linear-affine functions S with respect to ξ. From this fact it is obvious that in general the dual problem, restricted on the class $\mathcal{Y}_0 \subset \mathcal{Y}$ of functions (14), does not generate so good lower bounds of $\inf_{\mathcal{P}} J$ as $\sup L(S)$ on the whole $\mathcal{Y}$. An instructive comparison is supplied by the following example.

Example 1. It is to find in Euclidean metric the shortest way in the domain $\overline{G}_0 = \{ \xi \in R^2 \mid r_1 \leqq |\xi| \leqq r_2\}$, $r_1 < r_2$, from an initial point $\xi_1 = (0,-r_1)$ to the endpoint $\xi_2 = (0,r_1)$. - Here we obtain $\inf_{\mathcal{P}} J = \pi r_1 = \sup_{\mathcal{Y}} L(S)$, attained by $S(\xi) = r_1 \arctan(\xi^2/|\xi^1|)$; but on the other hand $\sup_{\mathcal{Y}_0} L(S) = 2r_1$.

A further difference between these duality conceptions is the following. The duality of Rockafellar has for convex problems the advantage of being symmetric, as the double dual problem coincides with the original one. On the other hand, our duality in the sense of (12) leads to fundamental differences between the analytical structure of the functionals J and L so that this new duality is not symmetric.

As an application of Theorem 1 let us discuss the case in which for a given process $\langle x,u\rangle$ and $S \in \mathcal{Y}$ the equality $J(x,u) = L(S)$ is valid. Then the pair $(\langle x,u\rangle, S)$ is said to be a saddle point of the duality condition (12) and $\langle x,u\rangle$, S are optimal solutions of (3) and of its dual problem respectively. Thus we can interpret the condition (13) equivalent to the saddle point property as a generalized form of Pontryagin's maximum principle. In this form it is especially a sufficient criterion for optimality of the process $\langle x,u\rangle$. In a recent paper [5] we proved that for problems (3) without state restrictions (disregarding boundary conditions) the condition (13) includes Pontryagin's maximum principle in the original form (for m = 1) and in the generalized form by L.Cesari [1] (for $m > 1$). The converse question is in general still unsolved: to what extent the condition (13) and the existence of the corresponding $S \in \mathcal{Y}$ is necessary for an optimal process $\langle x,u \rangle$. Only for special classes of (3) with $m=1$ it is known that the Bellman function realizes this condition. For convex problems the stability theory of Rockafellar [8] answers this question.

Finally we mention two further results without giving their proofs, which are similar to the proof of Theorem 1.

Theorem 2 . The result of Theorem 1 holds even if we replace the set $\mathcal{Y}$ by $\mathcal{Y}_p := \{ S = S_1 + S_2 \mid S_1 \in \mathcal{Y},\ S_2 \in W_p^{1,m}(\Omega) \}$.

Theorem 3 . Let $\langle x,u \rangle$ be a process and $S \in \mathcal{Y}_p$,restricted by the condition

$$(15)\qquad \delta_S(t,\xi) = S^{\alpha}_{t^\alpha}(t,\xi) + \mathcal{H}(t,\xi,S_{\xi}(t,\xi)) \leqq \sigma$$

for a.e. $t \in \Omega$ and every $\xi \in Q(t)$.

Then the inequality

$$(16)\qquad J(x,u) \geqq L_0(S) := \int_{\partial\Omega} \inf_{\xi \in Q(t)} \left[S^{\alpha}(t,\xi) n_{\alpha}(t) + l(t,\xi) \right] do$$

is valid.Here the equality holds if and only if

$$(17a)\qquad H(t,x,u,S_{\xi}(t,x)) = \mathcal{H}(t,x,S_{\xi}(t,x)) \quad \text{a.e. on } \Omega ,$$

$$(17b)\qquad \delta_S(t,x(t)) = \sigma \quad \text{a.e. on } \Omega ,$$

$$(17c)\qquad S^{\alpha}(t,x)n_{\alpha}(t) + l(t,x) = \inf_{\xi \in Q(t)} \left[S^{\alpha}(t,\xi) n_{\alpha}(t) + l(t,\xi) \right] \quad \text{a.e. on } \partial\Omega .$$

The estimate (16) induces a modified dual problem stated by the object

$$(18)\qquad L_0(S) \longrightarrow \text{Max on } \mathcal{Y}_p$$

under the constraint $\delta_S(t,.) \leqq \sigma$ for a.e. $t \in \Omega$.

In consequence of Lemma 1 this modified dual problem is a convex optimal problem on an infinite dimensional function space with a linear objective functional.If we denote the feasible set of (18) by $\mathcal{O}\!\mathcal{L}$ and regard it as a subset of $W_p^{1,m}(G)$, then formula (16) is true also on the closure $\overline{\mathcal{O}\!\mathcal{L}}$ so that $\sup_{\overline{\mathcal{O}\!\mathcal{L}}} L_0 \leqq \inf_{\mathcal{P}} J$.

Example 2 (parametric variational problems) .

We consider simple integrals (m = 1)

$$J(x) = \int_{\sigma}^{T} f(x,\dot{x})dt \longrightarrow \mathrm{Min} \quad \text{on} \quad W_p^{1,n}(\sigma,T)$$

under boundary conditions $x(\sigma) = x_o$, $x(T) = x_T$ and state restrictions $x(t) \in \bar{G}_o \subset R^n \ \forall t \in [0,T]$, where G_o is a domain satisfying $\partial G_o \in C_1^o$. Besides (5) we assume $f \geqq \sigma$ and $f(x,.)$ is a positive homogeneous function of the degree 1 . - Now we obtain by some here omitted computations under the additional assumption $S_t \equiv \sigma$ the result $L_o(S) = S(x_T) - S(x_o)$ and $\overline{\mathfrak{A}} = \{S \in W_p^{1,1}(G_o) \mid S_{\xi}(\xi) \in \mathcal{F}(\xi)\}$ a.e. on G_o , where $\mathcal{F}(\xi)$ is the convex figuratrix set at the point ξ in the sense of Carathéodory defined by

$$\mathcal{F}(\xi) = \{z \in R^n \mid z_i v^i \leqq f(\xi,v) \ \forall v \in R^n\} .$$

References

[1] L.Cesari, Optimization with partial differential equations in Dieudonné-Rashevsky form and conjugate problems, Arch.Rat. Mech.Anal. 33 (1969), 339-357

[2] I.Ekeland,R.Temam, Analyse convexe et problèmes variationnels, Gauthier-Villars, Paris 1974

[3] K.Friedrichs, Ein Verfahren der Variationsrechnung das Minimum eines Integrals als das Maximum eines anderen Ausdrucks darzustellen, Göttinger Nachr. 1929, 13-20

[4] A.D.Joffe,V.M.Tichomirov, Teoria ekstremalnych zadač, Nauka Moskva 1974

[5] R.Klötzler, On Pontryagin's maximum principle for multiple integrals, Beiträge z. Analysis 8 (1976), 67-75

[6] R.Klötzler, Einige neue Aspekte zur Bellmanschen Differentialgleichung, Materialy vsesoyuznogo simpoziuma po optimalnomu upravleniu i differencialnym igram, Tbilisi 1976, 146-154

[7] R.Klötzler, Weiterentwicklungen der Hamilton-Jacobischen Theorie, Sitzungsberichte der AdW der DDR (to appear)

[8] R.T.Rockafellar, Conjugate convex functions in optimal control and the calculus of variations, Journ.Math.Anal.Appl. 32 (1970), 174-222

[9] M.M.Cvetanov, O dvoystvennosti v zadačach variacionnogo isčislenia, Dokl. Bolgarskoj Akad. Nauk 21 (1968), 733-736

Author's address: Karl-Marx-Universität, Sektion Mathematik,
Karl-Marx-Platz, 701 Leipzig, DDR