E. A. Galperin
Department of Mathematics Universite du Quebec a Montreal
(C.P. 8888, Succ. Centre Ville, Montreal, Quebec H3C 3P8, Canada
E-mail:galperin.efim@uqam.ca)

Contiguity and Dynamic Programming

(Recommended by Prof. A. Martynyuk)

Some aspects of the principle of optimality [1, p.83] are considered, and a modification is proposed for the derivation of the main functional equations of dynamic programming to demonstrate that those equations are valid also in the case of non-optimal remaining trajectories under certain contiguity condition that is defined and analyzed in the paper. Control systems with incomplete information or structural limitations on controls do not, in general, satisfy the contiguity condition. Control problems for such systems may have optimal solutions which, however, cannot be obtained by dynamic programming. This fact is shown on example of a widely used engineering system for which optimal trajectories have all its remaining parts non optimal and non contiguous to the optimal trajectories. The paper presents theoretical justification of dynamic programming for contiguous systems without the principle of optimality, and discussion of some problems important for applications.

Рассмотрены некоторые аспекты принципа оптимальности [1, с. 83] и предложена его модификация для вывода основных функциональных уравнений динамического программирования, чтобы показать, что эти уравнения справедливы также в случае неоптимальных остаточных траекторий при некотором условии смежности. Приведены определение и анализ этого условия. Системы управления с неполной информацией или структурными ограничениями управления не удовлетворяют условию смежности в общем случае. Задачи управления для таких систем могут иметь оптимальные решения, но их нельзя получить методом динамического программирования. Этот факт продемонстрирован на примере широко применяемой технической системы, для которой все остаточные части оптимальных траекторий не оптимальны и не смежны с оптимальными траекториями. Дано теоретическое обоснование динамического программирования для смежных систем без принципа оптимальности и рассмотрены некоторые важные прикладные задачи.

K e y   w o r d s: dynamic programming, total optimality, contiguity.

1. Introduction. In the ten line section from [1, § 3, p. 83] we read: «In each process, the functional equation governing the process was obtained by an application of the following intuitive:

Principle of optimality. An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

The mathematical transliteration of this simple principle will yield all the functional equations we shall encounter throughout the remainder of the book. A proof by contradiction is immediate».

In control literature, the mathematical transliteration of the above principle for deterministic processes is related to certain types of functional equations, see, e.g., [1—8]. Now we reproduce some of them in authors' notations from [1, Ch. IX, § 13] where a new formalization of the calculus of variations is proposed for non-autonomous problems

$$f(a,c) = \max_u J(u) = \max_u \int_a^T F(x,u,t)\, dt, \quad \frac{dx}{dt} = G(x,u,t), \ 0 \le a \le t \le T, \ x(a) = c, \tag{1}$$

parameterized by a, c with variable $a \in [0, T)$. Splitting the integral over two segments $[a, a+S]$ and $[a+S, T]$, we have, according to the principle of optimality,

$$f(a,c) = \max_{u[a,\,a+S]} \left[\int_a^{a+S} F(x,u,t)\, dt + f(a+S,c(a+S)) \right], \tag{2}$$

for all a and all S such that $0 \le a < a + S \le T$. For small S this yields

$$f(a,c) = \max_{u[a,\,a+S]} [F(c,u(a),a)\, S + f(a,c) + S f_a + S f_c dc(S)/dS + o(S)]. \tag{3}$$

In (3) the terms $f(a,c)$ cancel out and, dividing by S we get, to the first order as $S \to 0$

$$0 = \max_{u(a)} [F(c,u(a),a) + f_a + f_c G(c,u(a),a)], \tag{4}$$

$$-f_a = \max_v [F(c,v,a) + f_c G(c,v,a)], \ v = u(a), \tag{5}$$

and for optimal $v = v^o$ in the interior of the region, we get, cf. [1, Ch. IX, § 13, (8)]:

$$-f_a = F(c,v^o,a) + f_c G(c,v^o,a), \tag{6}$$

$$0 = F_v(c,v^o,a) + f_c G_v(c,v^o,a). \tag{7}$$

Solving this system for f_a, f_c and equating $f_{ac} = f_{ca}$ yields a partial differential equation for $v^o(a,c)$ which is the extremal control $u^o(t,x(t))$.

Remark 1. If the extremal (supposed to be optimal) v^o of (5), is not in the interior of the admissible region, then PDEs (6), (7), must be replaced by extremality (stationarity) conditions of the Karush — Kuhn — Tucker type. Moreover, in this case the classical theorem on mixed partial derivatives to find u^o, v^o may be inapplicable. In engineering and economics, optimal policies (controls) usually belong to boundaries of admissible regions for controls and states (trajectories), which necessitates the replacement of those PDEs (not the Bellman equation (5)). This extension is beyond the scope of the paper.

There is vast literature on the subject, see, e.g., [1—8] and references therein, mainly concerned with the application of the optimality principle to various kinds of problems, derivation of the corresponding functional equations, their solution by different methods under certain assumptions and conditions, investigation of the existence of solutions and stability properties, numerical studies and computational experiments, etc.

However, the principle of optimality is an important statement in its own right and it is interesting to study it as such, without an accent on its numerous applications, particular functional equations, technical details or connections with other theories.

The statement of the principle of optimality is very general. It does not specify a notion of optimality to which the principle is referred. In [1], it is applied to max (min) optimality for certain classes of functionals, and to optimality in games with a saddle point. The statement does not specify a set of possible decisions (admissible controls), nor a class of problems to which it is applicable, nor allowable restrictions, etc., leaving one to believe that it is a universally valid principle, i.e., «a comprehensive and fundamental law, doctrine, or assumption» (Webster's dictionary), under the sole hypothesis that an optimal policy exists. It also does not require convexity, monotonicity, differentiability, etc., though such properties may be imposed on a case-to-case basis in order to obtain or solve functional equations resulting from the principle and to assure that an extremal (stationary) solution is, in fact, optimal.

2. Observations. The formulation of the principle as presented in [1, p. 83], see Section 1 above, contains some explicit and implicit assumptions, such as the existence of an optimal policy which includes optimal remaining decisions as its parts, the existence of «the state resulting from the first decision», and the existence of non-optimal policies and corresponding trajectories emanating from the same initial or intermediate state along an optimal trajectory. Among the most important suppositions, we can list the following properties.

2.1. *The congruence property.* For simplicity, suppose first that an optimal policy is unique and defined by a sequence of decisions $\{u_1, ..., u_n, ..., u_N\}$ or by a continuous function $u(t)$, $0 \le t \le T$, $T \le \infty$.

Then, «the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision». Denote this subsequent optimal policy by the sequence $\{v_n, ..., v_N\}$, $1 < n < N$, or, respectively, by the function $v(t)$, $0 \le t_1 \le t \le T$, with the understanding that n or t_1 correspond to the state resulting from decisions $\{u_1, ..., u_{n-1}\}$, or $u(t)$, $0 \le t \le t_1$. The congruence property is the requirement of the principle of optimality that $v_i = u_i$, $i = n, ..., N$, or, respectively, $v(t) \equiv u(t)$, $t_1 \le t < T$. Indeed, otherwise, $\{u_1, ..., u_N\}$, or $u(t)$, $0 \le t < T$, would not constitute an optimal policy as specified in the principle, if

an optimal policy is unique. This can be taken as an immediate proof by contradiction which, indeed, appeared in the literature, see Section 4.

If an optimal policy is not unique, then u_i ($i = 1, ..., N$) and $u(t)$ above may be set-valued. In this case, the congruence property means that $v_i \subseteq u_i$, $i = n, ..., N$, or $v(t) \subseteq u(t)$, $t_1 \leq t < T$. For this and other reasons, see Section 5, we prefer to consider congruence as a separate assumption meaning the inclusion of «remaining decisions» in the remaining part of the original policy, without requirement of their optimality. There are different terms expressing this notion in the literature, e.g., «joining controls» in [6, p. 83], or «concatenated» pieces of admissible trajectory in [7, p. 87].

2.2 *Additivity of costs in the optimality criterion (cost function).* This assumption can be seen from the functional equation (2) and other similar equations derived from the principle of optimality. In some sources, e. g., in [6, pp. 84, 85], it is postulated separately in order to prove the principle, see Section 3. In integral criteria, it follows from the additivity of the integral with respect to intervals of integration. This assumption represents a restriction on the class of problems to which the principle of optimality is applied.

2.3 *Feedback structure of optimal policies.* This property is implied by the statement that «the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision». It can be seen from various equations in [1 — 4] that define the optimal policy, from (6), (7) for $v^o(a, c) \Rightarrow u^o(t, x(t))$, and from other similar equations in the literature.

2.4. *Independence on the preceding data and (or) decisions.* The optimality of remaining decisions «with regard to the state resulting from the first decision» must not explicitly depend on the initial state nor the initial decision. This requirement is manifested by the word «whatever» in the formulation of the principle. In practice, it means that feedback controls obtained from equations implied by the principle explicitly depend on the current state, and not on the initial state nor preceding states, nor preceding decisions.

2.5. *Invariance of the Bellman function $f(\cdot)$.* This hypothesis is not mentioned in the principle itself as formulated in [1, p. 83], see Section 1, and it does not follow from the principle nor from any of the preceding observations. However, it is used in derivation of the functional equations and PDEs of dynamic programming, see (2), (3) at left and in the bracket, and it is tacitly assumed in applications and research works that followed in subsequent publications. This hypothesis is supported by a number of solved examples, though without any theoretical justification. The invariance of the function $f(\cdot)$ may be violated with modification of parameters or arguments at later steps (subsequent parts) of controlled trajectory.

3. Proof by contradiction. Geometrically, the principle of optimality can be illustrated as follows. Consider a period of time $[t_0, T]$, $T \leq \infty$, an admissible

region X (state space), and denote by $x(t) \in X$ a point on a trajectory which may depend on a policy (control) $u(x, t)$ applied to the process to assure the optimality of some criterion $J(u(\cdot), \cdot)$, say, $\max J$, see [1, 2]. A trajectory $x[t_0, T]$ can be discrete or continuous. Admissible controls are piecewise continuous functions, and the cost function (optimality criterion) $J(u(\cdot), \cdot)$ is assumed to be additive [1—8], that is

$$J[t_1, t_3] = J[t_1, t_2] + J[t_2, t_3], \quad t_0 \le t_1 < t_2 < t_3 \le T. \tag{8}$$

The policy (control)

$$u[t_1, t_3] = u^o(x, t) = \arg\max J(u, \cdot), \quad t \in [t_1, t_3] \tag{9}$$

and the corresponding piece of trajectory $x[t_1, t_3]$ are called optimal. The statement of the principle is simple: if $u[t_0, T]$ is optimal for the whole trajectory $x[t_0, T]$, then remaining parts $u[t, T]$, $x[t, T]$ are also optimal for any $t \in (t_0, T)$, provided that $x(t) \in x[t_0, T]$. Shortly: remaining parts of an optimal policy and trajectory are themselves optimal with respect to the same criterion.

In some sources, see, e.g., [8, footnote to (4.2.16)], this condition is qualified as sufficient condition. Other sources present a proof of the principle as a necessary condition of optimality. In [6, pp. 86, 87], or [7, p. 87], a simple proof of the principle is given which can be summarized as follows.

P r o o f. Setting $t_1 = t_0$, $t_3 = T$ in (8) and assuming the existence of optimal policies (trajectories) starting at any point $x(t) \in x[t_0, T]$, one can see that if $J[t_0, T] = J^o$ is optimal, say, $J^o = \max J[t_0, T]$ for some $u^o(t)$, $t \in (t_0, T)$, then it implies

$$\max J[t_0, T] = J[t_0, t_2] + \max J[t_2, T]. \tag{10}$$

Indeed, if $J[t_2, T]$ in (10) is not maximal, then there is a better policy $u[t_2, T]$, thus, $J[t_0, T]$ at the left in (10) is not optimal, contradicting the assumption of its optimality.

4. Total optimality. Theorem 1. If for an optimal trajectory, all remaining decisions are optimal, then every part of that trajectory is itself optimal.

P r o o f. Consider relation (8) again with $t_3 = T$, not fixing t_1 nor t_2. By assumption, $J[t_1, T]$, $J[t_2, T]$ are optimal, say, maximal, as corresponding to optimal remaining decisions, thus, instead of (8), we can write

$$\max J[t_1, T] = \max J[t_1, t_2] + \max J[t_2, T], \quad t_0 \le t_1 < t_2 \le T. \tag{11}$$

Now, if $J[t_1, t_2]$ in (11) is not maximal, there is a better policy $u[t_1, t_2]$, thus, $J[t_1, T]$ at the left in (11) is not optimal, contradicting the assumption of its optimality.

C o r o l l a r y. If two remaining parts $x\,[t_1,T]$, $x\,[t_2,T]$ of a trajectory (not necessarily optimal) are optimal, then the piece $x\,[t_1,t_2]$ of that trajectory is also optimal.

D e f i n i t i o n. If all parts of a trajectory including the whole trajectory itself are optimal, then such a trajectory (process) is called totally optimal.

We see that, according to the proof by contradiction, the principle of optimality under assumptions 2.1—2.5 implies total optimality. It is well known that certain optimal autonomous systems are totally optimal. Certain optimal non autonomous systems may also be totally optimal. Clearly, every totally optimal system satisfies the principle of optimality.

5. Contiguity and dynamic programming. If an optimal trajectory (policy) is not totally optimal, then the principle of optimality is not valid for such a trajectory (policy). However, a modification of the derivation of the functional equations for the optimal cost and optimal policy allows one to conclude that the same resulting equations (5) — (7) hold under certain supplementary condition, irrespective of the principle of optimality.

As in [1], we make a blanket assumption that all functions have continuous partial derivatives up to the order we may need, and that optimal trajectories we may consider do exist. In the course of derivation, we shall indicate all additional conditions that are required and must be verified when solving practical problems.

If the principle of optimality does not apply to the problem, then the transition from (1) to (2) is not valid. Moreover, an optimal cost $f(a, c)$ in (1) for the remaining policies on $[a, T)$, $a > 0$, need not be preserved as the same function for all a because the optimal policy $u_t^o(\cdot)$ in the integrand may not be congruent with $u_a^o(\cdot)$, that is, $u_t^o(t) \ne u_a^o(t)$ for $t > a$, where $u_t^o(\cdot), u_a^o(\cdot)$ are curves, not partial derivatives. With this in mind, denote

$$f^{\,a}(a,c) = \max_u J(u) = \max_u \int_a^T F(x,u,t)\,dt, \tag{12}$$

where $c = x\,(a)$ is an intermediate point for $t = a \in [0, T)$. Now, we have

$$f^{\,a}(a,c) = \max_{u[a,\,T]}\left[\int_a^{a+S} F(x,u,t)\,dt + \int_{a+S}^{T} F(x,u,t)\,dt\right]; \tag{13}$$

$$f^{\,a}(a,c) = \max_{u[a,\,a+S]}\left[\int_a^{a+S} F(x,u,t)\,dt + f^{\,a+S}(a+S,x\,(a+S)) + g^{\,a+S}(\cdot)\right], \tag{14}$$

where $g^{\,a+S}(\cdot)$ is the difference between the value of the second integral in (13) with the optimal control $u^o[a,T]$, and the optimal value $f^{\,a+S}(\cdot)$ of the same in-

tegral maximized separately by $u^*[a + S, T)$ which may not be restriction of $u^o[a, T)$ onto $[a + S, T)$, that is, $u^*(t) \ne u^o(t)$ for $t \in [a + S, T)$ (non-congruence), if the trajectory is not totally optimal.

The weak invariance hypothesis: assume that $\lim f^{a+S}(\cdot) = f^a(\cdot)$ as $S \to 0$ for every $a \in [0, T)$, and same for its partial derivatives of the first order.

Now, expanding the first two terms of (14) into Taylor series with respect to S and denoting $v = u(a)$, we obtain

$$f^a(a, c) = \max_{u[a,\, a+S]} [F(c, v, a)\, S + f^a(a, c) + S f_a^a(\cdot) + S f_c^a(\cdot)\, dc(S)/dS +$$
$$+ o(S) + g^{a+S}(\cdot)]. \tag{15}$$

In (15), we cancel out the term $f^a(a, c)$, then divide the equation by $S > 0$, take $f_a^a(a, c)$ to the left hand side, and consider $S \to 0$. Suppose that the following assumption holds

$$\lim_{S \to 0} g^{a+S}(\cdot)/S = 0 \text{ for all } a \in (0, T), \tag{16}$$

which we call the contiguity condition. Assume also that the Taylor series in (15) is convergent, and all terms in the bracket (14) with their respective partial derivatives are continuous in all arguments. Then noting that $u(a) = v$, $x(a) = c$, we get the equation

$$-f_a^a(a, c) = \max_v [F(c, v, a) + f_c^a(a, c)\, G(c, v, a)], a \in [0.\ T) \tag{17}$$

and superscript a is redundant because of (16), thus, (17) is identical to (5). Since by notation $a = t$, $c = x(t)$, $v = u(a) = u(t)$, so relation (17) can be written in the usual form

$$-f_t(t, x) = \max_{u(t)} [F(x, u, t) + f_x(t, x)\, G(x, u, t)], t \in [0, T). \tag{18}$$

If optimal $u^o(t)$ is in the interior of the admissible region, then we have the equations

$$-f_t(t, x) = F(x, u^o, t) + f_x(t, x)\, G(x, u^o, t), t \in [0, T), \tag{19}$$

$$0 = F_u(x, u^o, t) + f_x(t, x)\, G_u(x, u^o, t), t \in [0, T) \tag{20}$$

which are identical to (6), (7). This means that the contiguity condition (16) is sufficient for derivation of those equations without the optimality of remaining trajectories. Vice versa, it is easy to see that, if the limit in (16) is nonzero or does not exist, the equation (17) does not follow from (15), thus, equations (19), (20) are invalid despite sufficient smoothness and convergence of the Taylor series in (15). Hence, under those regularity conditions, the following statement is proved:

Theorem 2. Under the Weak Invariance Hypothesis and the assumptions cited in observations 2.2, 2.4, the derivation of the functional equations and PDEs of dynamic programming is valid if and only if the contiguity condition (16) is satisfied, irrespective of the principle of optimality.

Remark 2. It is worth noting that for totally optimal systems (e.g., certain autonomous systems), we have $g^{a+S}(\cdot) \equiv 0$ for all a, S which means precisely the optimality of all remaining trajectories. Also, the validity of derivation may not imply the existence of solutions (same as indicated in [1] on many occasions).

Contiguity condition (16) can be simplified. To make it independent of the concrete problem above and special notations in (12) — (16), we return to considerations of Sections 3, 4, see (8), (10), (11). In our notations of (11), the limit as $S \to 0$ [[1], Ch. IX, § 13], corresponds to the limit as $t_2 \to t_1$, $t_1 \in [t_0, T)$ in (11). Without optimality of remaining decisions, we have to consider two different trajectories and, instead of (11), the following relations:

$$x[t_1, T] = \lim_{t_2 \to t_1}(x[t_1, t_2] \cup x^*[t_2, T]), \; x(t_2) = x^*(t_2), \; t_0 \le t_1 < t_2 < T, \quad (21)$$

$$J(x[t_1, T]) = \lim_{t_2 \to t_1}\{J(x[t_1, t_2]) + \max J(x^*[t_2, T])\}. \quad (22)$$

(The set-wise limit in (21) means simple coincidence as $t_2 = t_1$, $t_1 \in [t_0, T)$, without consideration of various definitions of such limits.)

Here $x[t_1, T]$ designates a remaining (not necessarily optimal on $[t_1, T]$) part of an entire optimal trajectory $x^o[t_0, T]$, on which part the value $J(x[t_1, T])$ is attained under original optimal control $u^o[t_0, T]$. If we remove the equality signs at left and the limits in (21), (22), then $x[t_1, t_2]$ designates the initial part of semi-trajectory $x[t_1, T]$ with the value $J(x[t_1, t_2])$ upon $x[t_1, t_2]$. Notation $x^*[t_2, T]$ means certain optimal semi-trajectory starting at $x(t_2)$ with the value $\max J(x^*[t_2, T])$ attained on it. This semi-trajectory $x^*[t_2, T]$ may not coincide with the remaining part $x[t_2, T]$ of the trajectory $x[t_1, T]$.

In cases when the principle of optimality is valid for the system under consideration, we have identically $x^*[t_2, T] \equiv x[t_2, T]$, equalities in (21), (22) are valid without the limit operation, relation (22) coincides with (11), and relation (21) becomes trivial tautology omitted in the proof of Theorem 1.

If the principle of optimality is invalid, then $x^*(t) \ne x(t)$, $t \in (t_2, T]$, but relations (21), (22) may still hold in the limit for any $t_1 \in [t_0, T)$. A careful reader can see that it is the equalities of the type (21), (22) that are actually necessary and sufficient for the derivation of the functional equations and PDEs of dynamic programming in [1, Ch.9, § 3, 6, 13, 22], as it is presented above in (12) to (20).

Denote $J(x^o[t_2, T])$ the (non optimal) value of $J(\cdot)$ over remaining part $x^o[t_2, T]$ of the semi-trajectory $x[t_1, T]$, $t_2 > t_1 = a \ge t_0$, under original optimal

control $u^o(t) \in u^o[t_0, T]$. Let extr $J = J(x^*[t_2, T])$ be the extremal (max or min) value of $J(\cdot)$ over a different semi-trajectory $x^*[t_2, T]$ starting at the same point $x^*(t_2) = x^o(t_2)$ but optimized separately by some control $u^*(t) \neq u^o(t)$, $t \in [t_2, T)$.

Instead of $g^{a+S}(\cdot)$ in (14) — (16), consider the functional

$$C[t_1, t_2] := C(u^o[t_1, T], u^*[t_2, T]) = J(t_1, x^o[t_2, T]) - J(x^*[t_2, T]),$$

$$u^o[t_1, T] \subset u^o[t_0, T], \tag{23}$$

where t_1, t_2 are variable markers indicating two starting points of two non congruent remaining trajectories. The quantity $C[t_1, t_2]$ of (23) is identical to $g^{a+S}(\cdot) > 0$ for min-problem, and to $g^{a+S}(\cdot) < 0$ for max-problem as in (14), (15). The value $J(t_1, x^o[t_2, T])$ is equal to the second integral in (13), and $J(x^*[t_2, T])$ equals the term $f^{a+S}(\cdot)$ in (14).

The functional (23) is a measure of *J*-proximity of semi-trajectories $x^o[t_2, T]$, $x^*[t_2, T]$ of which $x^*[t_2, T]$ is optimal with respect to the same criterion $J(\cdot)$ but does not coincide with (non optimal) remaining part $x^o[t_2, T]$. For totally optimal systems (where the principle of optimality is valid), we have $C[\cdot] \equiv 0$. Otherwise, $C[\cdot]$ should be of higher order than the time difference $t_2 - t_1 \to 0$ for all $t_1 \in [t_0, T)$ in accordance with (16) to assure the applicability of dynamic programming methods.

Denote $S = t_2 - t_1$ (same S as in [1, Ch. 9, § 3, 6, 13, 22]), then (23) can be rewritten as follows:

$$C[t_1, t_1 + S] := C(u^o[t_1, T], u^*[t_1 + S, T] =$$

$$= J(t_1, x^o[t_1 + S, T]) - J(x^*[t_1 + S, T]), \tag{24}$$

where $t_1 = a$ in (12) to (16), and we get an equivalent representation for (16):

$$\lim_{S \to 0} \{C[t_1, t_1 + S] / S\} = 0 \text{ for all } t_1 \in [t_0, T). \tag{25}$$

Under assumptions of continuous partial derivatives adopted in [1], the functional $C[t_1, t_1 + S]$ is continuously differentiable with respect to *S*. Consider the limit

$$C(t_1) = \lim_{S \to 0} C[t_1, t_1 + S] = \lim_{S \to 0} \{J(t_1, x^o[t_1 + S, T]) - J(x^*[t_1 + S, T])\}. \tag{26}$$

If $C(t_1) \neq 0$, then in (25) we have $\pm \infty$, and non-contiguity follows. If $C(t_1) = 0$, the l'Hopital rule applies, and (25) is equivalent to the condition

$$\lim_{S \to 0} \{dC[t_1, t_1 + S] / dS\} = 0 \text{ for all } t_1 \in [t_0, T), \tag{27}$$

which is the contiguity condition for general systems and functionals, not necessarily of the type considered in [1, 2], as in (12).

6. Non contiguous optimal dynamical systems. It is tempting to think that dynamic programming, applicable even without the principle of optimality, might be universally applicable to almost all systems which most likely satisfy the Invariance Hypothesis and the contiguity condition, with the exception of, maybe, some pathological examples. Unfortunately, this is not the case.

Non-contiguity (i.e., non satisfaction of the contiguity condition) in optimal dynamical systems is a rather fine property that invalidates the extension (application) of an optimal control over some part of a trajectory for use over another part of the same trajectory. Strangely as it may seem, non-contiguous systems are common in practice. In such cases, dynamic programming is inapplicable, and other methods can be employed, see [5—11].

Systems with incomplete information and/or structure constraints are generally non-contiguous, even if they are autonomous. To illustrate the existence of such systems, we consider an engineering problem, see [10; 11, p. 4—7], of optimal stabilization of a stationary linear oscillator (engines, turbines, vibratory machines, etc.) which we study in two distinct settings to demonstrate the fundamental difference in regard to optimality that may occur with modifications in the structure of a controller.

Example 1 [10]. In all textbooks on undergraduate mathematics, mechanics, or control theory, the following linear oscillator equation is considered

$$d^2x / dt^2 + x = u, \ t \geq 0, \ x(0) = x_0, \ dx(0)/dt = v_0. \tag{28}$$

Denoting $dx/dt = v$ (velocity), we convert (28) into the normal form:

$$dx/dt = v, \ dv/dt = -x + u, t \geq 0, \ x(0) = x_0, \ v(0) = v_0. \tag{29}$$

The oscillator (28) or (29) should be stabilized to minimize the functional

$$f(x_0, v_0) = \min J(u) = \int_0^\infty (ax^2 + bv^2 + cu^2)\, dt, a, b, c = \text{const} > 0. \tag{30}$$

The system is autonomous, its optimal trajectories are totally optimal, the principle of optimality applies, thus, denoting $f(x, v) = \min J(u)$ for $t > 0$, the Bellman equation follows with $f_T \equiv 0$ for $T = \infty$ (fixed), cf. [1, Ch. IX, § 3, (12)]:

$$f_T(x, v) = \min [ax^2 + bv^2 + cu^2 + v\partial f / \partial x + (-x + u)\partial f / \partial v] = 0. \tag{31}$$

With the control space open, it implies for optimal u^o,

$$2cu^o + \partial f / \partial v = 0. \tag{32}$$

Substituting optimal control u^o from (32) into (31) (which removes the min operation) and solving the nonlinear PDE, we obtain the minimal cost ($J(u)$ of (30) is strictly convex) and the unique optimal feedback :

$$f(x,v) = (c+a)^{1/2}(b-2cp)^{1/2}\, x^2 - 2cpxv + (bc - 2c^2p)^{1/2}v^2, \qquad (33)$$

$$u^o = px - (\,b/c - 2p)^{1/2}v, \ \ p = 1 - (1 + a/c)^{1/2}. \qquad (34)$$

We see that feedback (34) does not depend on initial conditions in (29). It means that every remaining part of an optimal trajectory is itself optimal. Also, for fixed T, the equation (31) does not depend on boundary conditions $x(t_0)$, $x(T)$, $0 \le t_0 < T < \infty$, which means that every intermediate part $x[t_0, T]$ of the optimal trajectory is itself optimal (total optimality) with the same control (34) (congruence), cf. Theorem 1.

Example 2 [10]. Suppose that the space coordinate $x(t)$ cannot be measured (fluid friction which is common in engineering), thus, the control must be of the form $u = u(v)$. If $a = 0$ in (30), then $p = 0$, and u^o of (34) is feasible. However, in (30) we have $a > 0$, thus, in (34) $p < 0$ and u^o is infeasible. In the class of controls $u(v)$ depending on velocity, equation (31) does not have a solution, though «approximate» solutions with small $|p|$ do exist, all non optimal since they actually correspond to small $a > 0$ in (30).

In the spirit of dynamic programming and noting the result (34), let us look for the optimal feedback in the class of controls

$$u = -2qv, \ \ q > 0, \qquad (35)$$

which represents a nonsingular Lagrange problem of the calculus of variations with two constraints (29) if we eliminate u from (30) by the substitution $u = x + v'$ from (29), and then substitute (35) into (29).

For this class of controls, one can also write the Bellman equation by simply substituting (35) into (31), and then considering $q(\cdot)$ as a control. Denoting $u = -2q(\cdot)v = u^o(\cdot)$, one gets the same equations (31), (32) with the solution (34) — nothing new.

However, the optimal regulator of the form (35) with $q = $ const (optimal dampening) does exist. For system (29), (30), (35), the characteristic equation is $r^2 + 2qr + 1 = 0$, so that Re $r_{1,2} < 0$, the integral (30) is convergent and, for $q \ne 1$, it has the value

$$J(x_0, v_0, q) = [(a+c)\, x_0^2 + cv_0^2]\, q + \tfrac{1}{4}(a+b)(x_0^2 + v_0^2)\, q^{-1} + ax_0v_0. \qquad (36)$$

From the equation $\partial J/\partial q = 0$, we find the extremal value

$$q_0^2 = \tfrac{1}{4}(a+b)(x_0^2 + v_0^2)/[(a+c)\, x_0^2 + cv_0^2], \ \ q_0 > 0, \qquad (37)$$

for which $\partial^2 J/\partial q^2 > 0$, yielding the minimum in (36), see [10, formulae (5.7)—(5.12)]. Since $J(\cdot)$ of (36) is continuously differentiable at $q = 1$, for all values of parameters in (36), formulae (37), (36) are valid also for $q = 1$, covering the whole range of possible friction fluid densities.

Comparing (34) and (35), (37), one can see that:

1) u^o of (34) does not depend on initial data meaning that u^o is optimal not only for any remaining part of an optimal trajectory (Bellman's principle of optimality) but also for all trajectories in the feasible space X. It means that equations (33), (34) establish a property which may be called field optimality in the sense that for the system (28) — (30) there exists one single optimal control (34) such that every trajectory starting from any point $x \in X = R^2$ is optimal . Moreover, there is a unique surface (33) which renders the optimal value $f(x)$ of the functional (30) on the trajectory starting from that point x.

2) $u = - q_0 v$ of (35), (37) does depend on initial data x_0, v_0, meaning that, being optimal for the whole trajectory corresponding to those data, it is not optimal for any remaining part thereof, hence, Bellman's principle of optimality is invalid in the class of controls (35), if $a > 0$ in (30).

3) If $a = 0$ in (30), we have $u \equiv u^o$, see (34), (35), (37), thus, for this class of functionals in (30), Bellman's principle of optimality is valid in the class of controls (35) for the particular system (29). Moreover, field optimality is preserved.

Nonexistence of solutions of the Bellman equation (31) for linear system (29), (30), (35), which equation can be formally written as proposed in [1], means that this autonomous system with incomplete information is non-contiguous. Its non-contiguity invalidates the derivation of Bellman's equations presented in [1]. In contrast, this derivation is valid for contiguous systems even without optimality of remaining semi-trajectories because the limit operation as $S \to 0$ removes the necessity of exact coincidence (principle of optimality) of two (maybe, different) semi-trajectories starting at the same point $x(t)$, $t > t_0 = 0$. Those two semi-trajectories need not coincide. They must be touching at $x(t + S)$ as $S \to 0$ to the second or higher order, i.e., lie in a weak neighborhood of one another (cf. the curvature in the Weierstrass form of the Euler—Lagrange equations).

7. Concluding remarks and open problems. In this paper, the notions of total optimality and contiguity are defined, and a study of contiguity of dynamical systems is presented in application to dynamic programming. It is demonstrated that, according to the currently available proof [6, 7], the principle of optimality is a partial statement about total optimality. With free variations of controls, optimal autonomous systems are usually totally optimal. There are optimal non autonomous systems that are also totally optimal. Whether or not every such system can be transformed (by a non autonomous transformation) into a totally optimal system, or reduced to such a system is an open question of much practical importance.

Dynamic programming approach does not require optimality of remaining semi-trajectories: a much weaker contiguity condition would suffice. Many practical problems with structural constraints, incomplete information, and other specific requirements, do not satisfy the contiguity condition, even if such systems are autonomous. In some cases, optimal problems for such systems can be reformulated to allow for formal setting of dynamic programming equations. Those equations may have solutions which may not respect the constraints. It is interesting to find out how much the costs on real trajectories (with constraints) would exceed the costs of the formal dynamic programming solution without constraints, or with constraints partially satisfied.

Among other open questions, we can list the following problems:

1. Find an alternative proof of Bellman's principle of optimality, different from that presented in Section 3, which would not imply total optimality, Section 4.

2. Prove or disprove the equivalence of the Weak Invariance Hypothesis (Section 5) and Bellman's Invariance Hypothesis (Section 2, observation 2.5).

3. Find an example of a contiguous but not totally optimal dynamical system, or prove that contiguity implies total optimality.

4. Extend dynamic programming to the case when optimal trajectories are on the boundary, replacing gradient conditions by the Karush—Kuhn—Tucker type conditions.

5. Example 1 presents an instance where Bellman's equations imply field optimality. Study the conditions for this phenomenon.

Розглянуто деякі аспекти принципу оптимальності [1, с.83] і запропоновано його модифікацію для виведення основних функціональних рівнянь динамічного програмування, щоб показати, що ці рівняння є справедливими також у випадку неоптимальних залишкових траєкторій за певною умовою суміжності. Наведено визначення та аналіз цієї умови. Системи управління з неповною інформацією або структурними обмеженнями управління не задовольняють умові суміжності у загальному випадку. Задачі управління для таких систем можуть мати оптимальні розв'язки, але їх не можна отримати методом динамічного програмування. Цей факт показано на прикладі технічної системи, що широко використовується, для якої всі залишкові частини оптимальних траєкторій не оптимальні та не суміжні з оптимальними траєкторіями. Дано теоретичне обґрунтування динамічного програмування для суміжних систем без принципу оптимальності та розглянуто деякі важливі прикладні задачі.

1. *Bellman R.* Dynamic Programming. — Princeton, N. J. : Princeton University Press, 1957.
2. *Bellman R. E., Glicksberg I., Gross O. A.* Some Aspects of the Mathematical Theory of Control Processes. — Santa Monica, California: RAND Corporation, 1958.
3. *Bellman R.* Adaptive Control Processes: A Guided Tour. — Princeton, N. J. : Princeton University Press, 1961.
4. *Bellman R. E., Dreyfus S. E.* Applied Dynamic Programming. — Princeton, N. J. : Princeton University Press, 1962.

5. *Lee E., Marcus L.* Foundations of Optimal Control Theory. — N.Y. : Wiley, 1967.

6. *Leitmann G.* The Calculus of Variations and Optimal Control. — N. Y. : Plenum Press, 1981.

7. *Knowles G.* An Introduction to Applied Optimal Control. — Academic Press, 1981.

8. *Bryson A. E., Jr., and Yu-Chi Ho* Applied Optimal Control. — Waltham, Massachusets: Blaisdell Publ. Co., a Division of Ginn & Co., 1969.

9. *Galperin E. A.* Translational symmetry of certain integrals with application to the calculus of variations and optimal control// Mathematical and Computer Modeling. — 2002. — **36**. — P. 717—728.

10. *Galperin E. A., Dergacheva E. I.* Optimum stabilization of a linear system with incomplete information// Automation and Remote Control. — 1968. — **8**. — P. 1193—1202.

11. *Galperin E. A., Zheng Q.* Global Solutions in Optimal Control and Games. — Montreal: NP Research Publ., 1991.