ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria II: WIADOMOŚCI MATEMATYCZNE XII (1969)

Ministerstwo Nauki i Szkolnictwa Wyższego

A. Turowicz i H. Górecki (Kraków)

Sterowanie optymalne

1. Wstęp

1.1. Zagadnienia sterowania optymalnego zostały postawione w latach pięćdziesiątych obecnego stulecia przez Feldbauma [10], La Salle'a [26], Bellmana [1], Pontriagina [4] i Bushawa [6]. Najdonioślejsze wyniki uzyskali Pontriagin, Bellman i Krasowski. Od Pontriagina [30] pochodzi głośna zasada maksimum, Bellman stosował do problemów sterowania optymalnego swoją metodę programowania dynamicznego [2], a Krasowski [16] zapoczątkował stosowanie metod analizy funkcjonalnej. Te metody rozwinął w Polsce Kulikowski w pracach ogłoszonych w latach 1959 i 1960 [20], [21], [22], [23] i [24]. Uogólnienie zasady maksimum na układy opisywane równaniami różniczkowymi cząstkowymi uzyskał Butkowskij [7]. Tego rodzaju równania opisują układy, które w technice noszą nazwę *układów o parametrach rozłożonych* (w przestrzeni), w przeciwieństwie do układów opisywanych równaniami różniczkowymi zwyczajnymi, które noszą nazwę *układów o parametrach skupionych*.

Problematyka sterowania optymalnego wywodzi się z techniki. Podamy kilka przykładów.

1.2. Przeprowadzenie obiektu z jednego stanu w drugi w najkrótszym czasie. Niech x oznacza kąt obrotu wału silnika, t czas, u napięcie przyłożone do zacisków silnika. Należy przeprowadzić wał silnika z położenia x_0 do położenia x_1 w najkrótszym czasie, przy czym napięcie nie może przekroczyć z góry ustalonej wartości $|u| \leqslant a$. Załóżmy, że w przybliżeniu proces można opisać równaniami:

$$x'' = u(t), \quad x(0) = 0, \quad x'(0) = 0, \quad x(T) = x_1, \quad x'(T) = 0. \tag{1.2}$$

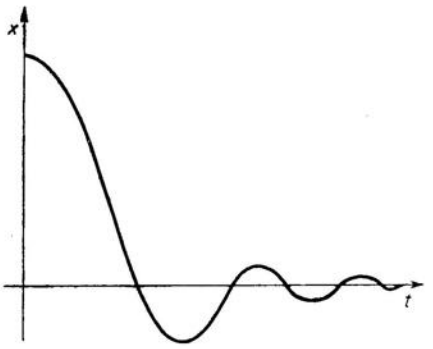
Zagadnienie polega na wyznaczeniu takiej funkcji $u(t)$, spełniającej warunek $|u(t)| \leqslant a$, aby T było najmniejsze. Jak to udowodnił Feldbaum, funkcja $u(t)$ ma dwa przedziały stałości: $u(t) = a$ dla $0 \leqslant t \leqslant \frac{1}{2}T$ i $u(t) = -a$ dla $\frac{1}{2}T < t \leqslant T$. Odpowiada to metodzie sterowania dwuwartościowego „cała naprzód, cała wstecz" (w literaturze angielskiej "bang-bang").

1.3. Ekonomiczne zużycie paliwa. Samolot zabiera określoną ilość paliwa. Zagadnienie optymalnego sterowania będzie polegało na tym, aby tak regulować prędkość, jako funkcję ilości paliwa, z uwzględnieniem warunków atmosferycznych, aby samolot miał największy zasięg. Matematyczne sformułowanie tego zagadnienia jest znacznie bardziej skomplikowane niż w poprzednim przykładzie. Widać jednak analogie: minimum czasowe zastąpione jest przez maksimum zasięgu, ograniczenie napięcia przez ograniczoną ilość paliwa. Jest jednak i zasadnicza różnica: w drugim przykładzie występują zakłócenia przypadkowe w postaci nie znanych dokładnie warunków atmosferycznych, zmieniających się w nieprzewidziany sposób.

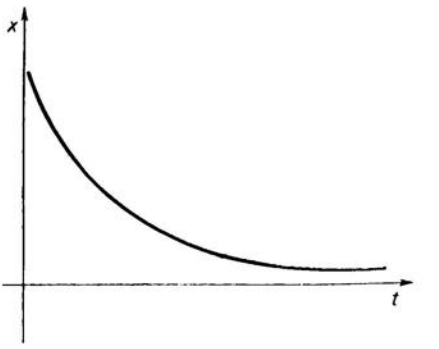
1.4. Optymalne tłumienie zmienne w układach pomiarowych. Niech x oznacza odchylenie napięcia w potencjometrze od wartości żądanej. Problem polega na sprowadzeniu tego odchylenia w najkrótszym czasie do zera. Tego rodzaju problemy spotykamy we wszystkich przyrządach pomiarowych typu galwanometru lub autokompensatora. Równanie różniczkowe opisujące proces z wystarczającą dla praktyki dokładnością może być napisane w postaci

$$x''+2ux'+bx = 0, \tag{1.4}$$

gdzie $u > 0$, $b > 0$; stałe u i b zależą od konstrukcji przyrządu. Jeżeli współczynnik tłumienia u jest bardzo mały, tak że u^2/b jest znacznie mniejsze od 1, to proces jest oscylacyjny, x szybko zbliża się do 0 w początkowej fazie, lecz wolno się ustala, gdyż oscylacje powoli zanikają (rys. 1).



Rys. 1



Rys. 2

Jeżeli współczynnik tłumienia u jest duży, tak że $u^2/b > 1$, to proces ma charakter aperiodyczny: w początkowej fazie wolno zbliża się do 0 (rys. 2). Czas ustalenia się procesu (osiągnięcia wartości $x = 0$) jest w obu przypadkach teoretycznie nieskończenie długi. W pierwszym przypadku

proces tłumienia przebiega korzystnie w początkowej fazie, w drugim we fazie końcowej.

Optymalne rozwiązanie otrzymuje się, jeżeli zmieni się konstrukcję przyrządu w ten sposób, aby współczynnik tłumienia był początkowo mały, a potem duży. Stawiamy problem: znaleźć funkcję $u(x)$ taką, aby proces opisany równaniem

$$x'' + 2u(x) \cdot x' + bx = 0$$

dobiegał końca w skończonym, minimalnym czasie. Problem ten został praktycznie rozwiązany przed powstaniem ogólnej teorii sterowania optymalnego.

1.5. Przykład sterowania dwuparametrowego. Rozważmy problem opisany w przykładzie pierwszym, którego realizację przeprowadza się za pomocą silnika prądu stałego, obcowzbudnego, sterowanego równocześnie prądem wzbudzenia i napięciem przyłożonym do zacisków wirnika. Proces można opisać w przybliżeniu równaniem różniczkowym nieliniowym

$$f(x'', x', x, u_1, u_2) = 0, \tag{1.5}$$

gdzie $u_1(t)$ oznacza napięcie na zaciskach wzbudzenia, $u_2(t)$ napięcie na zaciskach obwodu wirnika, x prędkość kątową silnika. Problem polega na przeprowadzeniu rozruchu silnika w minimalnym czasie przez dobór optymalny funkcji sterujących $u_1(t)$ i $u_2(t)$. Można wykazać, że minimalny czas rozruchu osiągalny przy równoczesnym optymalnym sterowaniu funkcjami $u_1(t)$ i $u_2(t)$ jest krótszy niż minimalny czas osiągalny za pomocą sterowania optymalnego jedną z funkcyj $u_1(t)$ i $u_2(t)$ przy założeniu, że druga jest stała.

1.6. Sterowanie optymalne w ekonomii. Teoria sterowania optymalnego znalazła szerokie i ważne zastosowanie w optymalnym planowaniu ekonomicznym, jak na przykład planowanie inwestycji, optymalny rozdział zasobów, zagadnienie optymalnej sieci transportu, optymalnego rozkładu obciążeń w systemie energetycznym i w wielu innych. Zagadnienia te nie mieszczą się w ramach niniejszego artykułu.

2. Postawienie problemu i pojęcia podstawowe. Modele matematyczne procesów sterowania

2.1. Równanie różniczkowe zwyczajne. Niech $\boldsymbol{x}$ oznacza wektor n-wymiarowy, zwany wektorem fazowym lub wektorem stanu, $\boldsymbol{u}$ wektor r-wymiarowy, zwany wektorem sterowania, t czas, $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t)$ funkcję wektorową n-wymiarową określoną dla $\boldsymbol{x} \in R^n$, $\boldsymbol{u} \in U \subset R^r$, $t \geqslant 0$ (U zbiór ogra-

niczony, domknięty). Zakładamy, że funkcja $f(x, u, t)$ ma pochodne ciągłe $\partial f/\partial x^i$, gdzie x^i oznaczają współrzędne wektora x, i że jest ciągła ze względu na u i t. Proces sterowania jest opisany równaniem

$$\frac{dx}{dt} = f\big(x(t), u(t), t\big) \tag{2.1}$$

dla $t_0 \leqslant t \leqslant t_1$, $t_0 \geqslant 0$; t_1 może równać się $+\infty$.

Sterowaniem dopuszczalnym nazywamy funkcję t określoną w $[t_0, t_1]$, kawałkami ciągłą, taką że $u(t) \in U$.

Dany jest funkcjonał $F(x, u)$, określony dla funkcji $x(t)$, absolutnie ciągłych w $[t_0, t_1]$ i dopuszczalnych sterowań. Problem polega na znalezieniu dopuszczalnego sterowania $u^*(t)$ takiego, że funkcjonał $F(x, u)$ osiąga minimalną (ewentualnie maksymalną) wartość dla $x = x^*(t)$, $u = u^*(t)$, gdzie $x^*(t)$ oznacza rozwiązanie równania (2.1) przy podstawieniu $u(t) = u^*(t)$, spełniające warunki $x^*(t_0) = x_0$, $x^*(t_1) = x_1$, gdzie x_0, x_1 są zadanymi punktami przestrzeni R^n.

Sterowanie $u^*(t)$ nazywamy *sterowaniem optymalnym*, a $x^*(t)$ *trajektorią optymalną* lub *torem optymalnym*.

Uwagi. Sformułowany problem różni się istotnie od zagadnienia rozwiązania równania różniczkowego, gdyż sprowadza się do wyznaczenia prawej strony równania (2.1) tak, aby jego rozwiązanie spełniało nałożone warunki. Problem jest zbliżony do zagadnień rozważanych w rachunku wariacyjnym, lecz zazwyczaj nie może być traktowany metodami klasycznymi z powodu założenia, że U jest zbiorem domkniętym i ograniczonym.

Istnieje także inny wariant tego problemu, gdy żądamy, aby tor optymalny leżał w ograniczonym i domkniętym zbiorze $X \subset R^n$.

2.2. Równanie różnicowe. Rozważamy ciąg fazowych wektorów n-wymiarowych $\{x_n\}$ i ciąg r-wymiarowych wektorów sterowania $\{u_n\}$. Zakładamy, że $u_k \in U$ dla wszelkich k naturalnych, gdzie $U \subset R^r$ jest zbiorem domkniętym i ograniczonym. Proces sterowania jest opisany równaniem

$$x_{n+1} = f(x_n, u_n, n), \qquad n = 0, 1, 2, \ldots \tag{2.2}$$

Oznaczmy przez $\tilde{x}_N$ ciąg skończony $(x_0, x_1, \ldots, x_N)$, przez $\tilde{u}_N$ ciąg skończony $(u_0, u_1, \ldots, u_{N-1})$. Dany jest funkcjonał $F(\tilde{x}_N, \tilde{u}_N)$. Problem polega na wyznaczeniu takiego ciągu $\{u_k^*\}$, aby funkcjonał F osiągnął minimalną (lub maksymalną) wartość dla $\{u_k\} = \{u_k^*\}$ i $\{x_k\} = \{x_k^*\}$, gdzie $\{x_k^*\}$ jest rozwiązaniem równania (2.2) przy podstawieniu $u_n = u_n^*$, dla zadanej początkowej wartości x_0 i końcowej x_N.

Uwaga. Możliwe są tu różne warianty: 1° liczba N może być zadana z góry; 2° liczba N może być dowolną liczbą naturalną; 3° N może wynosić $+\infty$, wtedy zadana jest graniczna wartość $\lim\limits_{N\to\infty} x_N$.

2.3. Równanie różniczkowo-różnicowe. Zachowujemy oznaczenia przypadku 2.1. Proces sterowania jest opisany równaniem

$$x'(t) = f[x(t-\tau), x(t), u(t), t], \tag{2.3}$$

gdzie $\tau > 0$. Równanie (2.3) przedstawia model procesów z opóźnieniem. Możliwy jest bardziej skomplikowany wariant, w którym prawa strona równania (2.3) zawiera wektory $x(t-\tau_1), \ldots, x(t-\tau_k)$, $k \leqslant n$, $\tau_j \geqslant 0$, $j = 1, 2, \ldots, k$, jeżeli trzeba uwzględnić występowanie w procesie sterowanym różnych opóźnień, których ilość nie przekracza jednak wymiaru przestrzeni fazowej.

Problem sterowania optymalnego formułuje się analogicznie, jak w przypadku 2.1. Trzeba jednak uwzględnić, że dla równania (2.3) jako warunek początkowy trzeba określić funkcję wektorową $x(t)$ w przedziale $[t_0-\tau, t_0]$. Model określony równaniem (2.3) reprezentuje procesy fizyczne, w których występuje tzw. opóźnienie transportowe, polegające na tym, że efekt sterowania ujawnia się w wektorze fazowym nie od razu, lecz po pewnym czasie τ.

2.4. Równanie różniczkowo-całkowe. Proces sterowania opisujemy równaniem

$$x' = f\Big[x(t), u(t), \int_{-\infty}^{t} K_1(t, s, x(s), u(s))ds, \ldots, \int_{-\infty}^{t} K_p(t, s, x(s), u(s))\,ds\Big], \tag{2.4}$$

$$p \leqslant n.$$

Całki występujące w (2.4) odpowiadają faktowi, że w niektórych procesach występuje zjawisko „pamięci", czy „dziedziczności", tzn. zależności wektora fazowego od przebiegu procesu w przeszłości. Jeżeli zaczynamy sterować procesem w chwili t_0, to należy przyjąć $u(t) = 0$ dla $t < t_0$.

Problem sterowania formułuje się tak, jak w przypadku 2.1.

2.5. Równanie różniczkowe cząstkowe. Jako przykład podamy model sterowania procesem nagrzewania

$$\frac{\partial x}{\partial t} = \frac{\partial^2 x}{\partial t^2} + f\Big[x, \frac{\partial x}{\partial t}, l, t, u(t, l)\Big]. \tag{2.5}$$

W tym równaniu $x(t, l)$ i $u(t, l)$ są funkcjami skalarnymi, zależnymi od czasu t i długości l.

Problem sterowania jest, w przypadku zależności wektora fazowego od wielu zmiennych, znacznie trudniejszy.

2.6. Modele procesów stochastycznych i adaptacyjnych. We wszystkich pokazanych dotąd modelach mieliśmy do czynienia z procesami

deterministycznymi, w których nie występują zakłócenia przypadkowe. W praktycznych zastosowaniach teorii sterowania zdarzają się procesy, których przebieg zależy nie tylko od wektora sterującego, lecz także od nie dającego się z góry dokładnie określić wektora zakłóceń przypadkowych $z(t)$. Wyróżnimy trzy ważne klasy modeli procesów sterowania, zależnych od wektora $z(t)$:

(A) *Modele stochastyczne.* Znany jest rozkład prawdopodobieństwa (najczęściej normalny) wektora zakłóceń $z(t)$. Wprowadzając formalnie wektor $z(t)$ do funkcji f w równaniach (2.1) i (2.5), otrzymujemy modele stochastyczne procesów sterowania. Wektor fazowy będzie zależał stochastycznie od wektora zakłóceń, będzie więc reprezentował proces stochastyczny. Ze względu na trudności w posługiwaniu się rozkładami prawdopodobieństwa, zwykle stosuje się metodę zastąpienia wektorów $x(t)$ i $z(t)$ przez ich wartości średnie. W języku polskim ukazały się dwie książki omawiające szczegółowiej procesy tego typu, a mianowicie Bellmana [3] i Feldbauma [11].

(B) *Modele rozgrywające.* Załóżmy, że rozkład prawdopodobieństwa wektora $z(t)$ nie jest znany. Zakładamy wobec tego, że $z(t)$ zmienia się w sposób najbardziej niekorzystny z punktu widzenia celu sterowania. Przez sterowanie optymalne rozumie się teraz sterowanie ekstremalizujące funkcjonał F, gdy spełnione jest powyższe założenie o wektorze $z(t)$. Sterowań, które byłyby optymalne gdyby wektor zakłóceń zmieniał się w inny sposób, nie bierze się pod uwagę. Rozwiązanie realizuje się przez wyznaczenie tzw. punktu siodłowego funkcjonału F, w którym on osiąga minimum ze względu na $u(t)$ i maksimum ze względu na $z(t)$ (lub na odwrót). Rozwiązania uzyskuje się metodami teorii gier. Bliższe szczegóły można znaleźć w książkach Krasowskiego-Pospiełowa [17] i Isaacsa [15].

(C) *Modele adaptacyjne.* Załóżmy, że znany jest typ rozkładu prawdopodobieństwa wektora $z(t)$, lecz nie znamy charakteryzujących go parametrów, jak np. średnia wartość, czy wariancja itp. Poprawne rozwiązanie problemu wyznaczenia optymalnego sterowania wymagałoby znajomości tych parametrów. Do adaptacyjnych układów sterowania stosuje się następującą metodę: przyjmuje się prowizorycznie pewne wartości parametrów uważane za prawdopodobne i obserwuje się efekt sterowania, wyznaczonego przy tym założeniu. Na podstawie tej obserwacji koryguje się stopniowo, co pewien czas, wartość parametrów. Ponieważ szukamy rozkładu prawdopodobieństwa *a posteriori*, stosuje się zwykle wzór Bayesa (por. [2]). Główne problemy adaptacyjne teorii sterowania, wraz z bogatym wykazem literatury, można znaleźć w artykule Jemieljanowa i Dubowskiego [9]. W układach adaptacyjnych korektura odbywa się automatycznie dzięki sprzężeniu zwrotnemu.

2.7. Ograniczenia wektora sterującego. Warunki techniczne narzucają na wektor sterowania pewne ograniczenia. Ograniczenia te można podzielić na dwie zasadnicze klasy (por. [28]):

(A) Ograniczenia lokalne w czasie polegają na tym, że wektor sterowania $\boldsymbol{u}(t)$ ma spełniać warunek: $\boldsymbol{u}(t) \in U$, gdzie U jest zbiorem ograniczonym, domkniętym, $U \subset R^r$. Najczęściej spotykamy cztery rodzaje ograniczeń tego typu:

(Aa) $U = \{\boldsymbol{u}: \max_j |u^j(t)| \leqslant M,\ j = 1, \ldots, r\}$, gdzie u^j jest współrzędną wektora sterowania. Jest to ograniczenie amplitudy każdej współrzędnej wektora $\boldsymbol{u}$.

(Ab) $U = \{\boldsymbol{u}: \sum_{j=1}^{r} |u^j(t)| \leqslant M\}$; tu suma amplitud jest ograniczona.

(Ac) $U = \{\boldsymbol{u}: \left[\sum_{j=1}^{r} |u^j(t)|^2\right]^{1/2} \leqslant M\}$; mamy tu ograniczenie całkowitej mocy źródła sterującego.

(Ad) U jest zbiorem skończonym punktów przestrzeni R^r, $U = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_N\}$.

(B) Ograniczenia globalne w czasie polegają na tym, że pewien funkcjonał zależny od $\boldsymbol{u}(t)$ w czasowym przedziale $[t_0, t_1]$ ma ograniczoną wartość. Najczęściej spotykane w zagadnieniach technicznych ograniczenia tego typu są:

(Ba) $\int_{t_0}^{t_1} \sum_{j=1}^{r} |u^j(t)|\, dt \leqslant M$; jest to ograniczenie wydajności źródła sterującego.

(Bb) $\left[\int_{t_0}^{t_1} \sum_{j=1}^{r} |u^j(t)|^2 dt\right]^{1/2} \leqslant M$; jest to ograniczenie całkowitej energii źródła sterującego.

Przy rozwiązywaniu problemu sterowania odgrywają rolę pewne własności zbioru U, jak np. jego wypukłość. W przypadkach (Aa), (Ab) i (Ac) U jest zbiorem wypukłym.

2.8. Ograniczenia wektora fazowego. Przy pewnych zagadnieniach technicznych żądamy, aby wektor stanu spełniał pewne warunki ograniczające, co zwykle bardzo komplikuje sprawę. Najczęściej spotykamy: (a) ograniczenia wszystkich, lub niektórych składowych wektora stanu: $|x^j(t)| \leqslant m_j$, $1 \leqslant j \leqslant n$, przy czym j przyjmuje wszystkie lub niektóre wartości dopuszczalne; (b) ograniczenia nałożone na pewien funkcjonał zależny od wektora $\boldsymbol{x}(t)$, np. $\left|\sum_{i=1}^{n} a_i x^i(t)\right| \leqslant M$, gdzie a_i stałe.

2.9. Kryteria optymalności sterowania. Formułując problem sterowania optymalnego mieliśmy zawsze do czynienia z zagadnieniem ekstremum

pewnego funkcjonału. Funkcjonał ten nosi nazwę *funkcji celu*. Sterowanie jest optymalne, gdy funkcja celu osiąga maksymalną, lub minimalną wartość, zależnie od tego, czy pożądane są jej duże, czy małe wartości. Obok nazwy funkcja celu używana jest również nazwa *wskaźnik jakości*, lub *kryterium optymalności*. Rolę kryterium optymalności sterowania można ująć w ten sposób: celem sterowania jest osiągnięcie maksimum zysków lub minimum strat przy przejściu układu ze stanu początkowego do stanu pożądanego. W zagadnieniach ekonomicznych oblicza się zyski i straty w dobrach ekonomicznych. W zagadnieniach technicznych jako koszt może występować np.: zużycie energii, paliwa, materiału, długość czasu potrzebnego do realizacji procesu technologicznego itp.

Do najczęściej stosowanych wskaźników jakości należą w przypadku modeli procesów ciągłych następujące funkcjonały:

1° Zadana jest funkcja $\boldsymbol{g}(t)$, przedstawiająca pożądany przebieg procesu. Szukamy sterowania optymalnego minimalizującego funkcjonał:

$$F(\boldsymbol{x}, \boldsymbol{u}) = \max_{t_0 \leqslant t \leqslant t_1} \|\boldsymbol{g}(t) - \boldsymbol{x}(t)\|$$

(może być też $t_1 = \infty$), to znaczy wymagamy, aby maksymalne odchylenie wektora fazowego od jego pożądanej wartości, było najmniejsze. Można przy tym posługiwać się rozmaitymi definicjami normy w przestrzeni R^n.

2° Często spotyka się funkcje celu określone przez całkę $\int_{t_0}^{T} F(\boldsymbol{x}(t), \boldsymbol{u}(t))\, dt$, która ma mieć minimalną wartość. Wymieniamy następujące przypadki:

(2a) $F(\boldsymbol{x}, \boldsymbol{u}) \equiv 1$. Zagadnienie sprowadza się do znalezienia minimalnego czasu $T - t_0$ dla trwania przebiegu optymalnego. Jest to tzw. problem czasowo minimalny.

(2b) $F(\boldsymbol{x}, \boldsymbol{u}) = \|\boldsymbol{f}(t) - \boldsymbol{x}(t)\|$ przy różnych definicjach normy.

(2c) $F(\boldsymbol{x}, \boldsymbol{u}) = \|\boldsymbol{f}(t) - \boldsymbol{x}(t)\|^2 + \lambda \|\boldsymbol{u}\|^2, \quad \lambda > 0.$

3° Spotyka się również wskaźnik jakości postaci $\lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{T} \|\boldsymbol{f}(t) - \boldsymbol{x}(t)\|^2 dt$, szuka się sterowania minimalizującego.

2.10. Dla modeli procesów dyskretnych spotykamy często następujące kryteria optymalności: Oznaczmy przez $\{\boldsymbol{g}_n\}$ pożądany przebieg procesu sterowanego.

1° Szukamy sterowania optymalnego, przy którym $\max_{0 \leqslant n \leqslant N} \|\boldsymbol{g}_n - \boldsymbol{x}_n\|$ osiąga minimalną wartość (może być też $N = \infty$).

2° Szukamy sterowania optymalnego, przy którym równość $\boldsymbol{x}_n = \boldsymbol{g}_n$ zachodzi przy najmniejszym n. Mamy tu problem czasowo optymalny dla procesu dyskretnego. Pewnym wariantem tego problemu jest, gdy

szukamy sterowania, przy którym nierówność $\|\boldsymbol{g}_n - \boldsymbol{x}_n\| < \varepsilon$, gdzie $\varepsilon > 0$ jest zadana, jest spełniona przy możliwie małym n. Rozważanie tego problemu okazuje się konieczne, gdy osiągnięcie równości $\boldsymbol{x}_n = \boldsymbol{g}_n$ nie jest możliwe. Z drugiej strony rozwiązanie tego problemu przy dostatecznie małym ε zupełnie wystarcza dla praktyki technologicznej.

3° Szukamy sterowania optymalnego, przy którym funkcjonał $\sum_{n=0}^{N} F(\boldsymbol{x}_n, \boldsymbol{u}_n)$ osiąga najmniejszą wartość. W szczególności ważny jest przypadek $F(\boldsymbol{x}_n, \boldsymbol{u}_n) = \|\boldsymbol{g}_n - \boldsymbol{x}_n\|$.

4° Szukamy sterowania optymalnego, przy którym funkcjonał $F(\boldsymbol{x}_N)$ osiąga ekstremalną wartość. Funkcja celu jest tu zależna tylko od końcowej wartości wektora fazowego $\boldsymbol{x}_n$ dla $n = N$.

5° Funkcja celu ma postać $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \|\boldsymbol{g}_n - \boldsymbol{x}_n\|^2$.

3. Metody teorii sterowania optymalnego

3.1. Ze względu na obszerność materiału ograniczymy się do modeli procesów deterministycznych: (a) ciągłych opisanych równaniem różniczkowym zwyczajnym lub zastępującym je równaniem całkowym; (b) dyskretnych opisanych równaniem różnicowym.

Jeżeli idzie o procesy ciągłe, to klasyczne metody rachunku wariacyjnego nie mogą być zazwyczaj stosowane, gdyż w najważniejszych, a przy tym najprostszych zagadnieniach, ekstremum wskaźnika jakości bywa osiągane, gdy sterowanie optymalne znajduje się na brzegu zbioru U sterowań dopuszczalnych, a nie w jego wnętrzu.

Ogromne usługi mogą oddać, i oddają, metody oparte na najnowszych wynikach dotyczących ekstremów warunkowych (por. pracę Dubowickiego i Milutina [8]) oraz na twierdzeniu Kuhna-Tuckera. Dla tych ostatnich metod podstawowymi pracami są Kuhna-Tuckera [19] i Hurwicza [14]. Z polskich prac poświęconych zagadnieniom sterowania, w których metody te są używane, lub częściowo referowane wymienimy prace [25], [27] i [28]. Lista ta nie jest kompletna. Metody Dubowickiego-Milutina oraz Kuhna-Tuckera są tak doniosłe, że powinny być udostępnione inżynierom pracującym naukowo w specjalnym opracowaniu. Dlatego też nie będą one referowane w niniejszym artykule.

3.2. Zmiana kryteriów optymalności. W praktyce inżynierskiej bywają stosowane metody, pozwalające przez zmianę wskaźnika jakości uprościć zagadnienie wyznaczenia sterowania optymalnego (por. [31]). Wtedy zamiast rozwiązania problemu postawionego otrzymuje się pewną jego aproksymację. Podamy przykłady:

(A) Przypuśćmy, że przy zadanym modelu procesu sterowanego chcemy zminimalizować funkcjonał $F(\boldsymbol{x}, \boldsymbol{u})$ przy przejściu w przestrzeni fazowej z punktu $\boldsymbol{x}_0$ do punktu $\boldsymbol{x}_1$ o współrzędnych x_1^k, $k = 1, \ldots, n$. Wprowadźmy zamiast funkcjonału F funkcjonał

$$F_1(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}, \boldsymbol{u}) + \alpha \sum_{i=1}^{n} |x^i(t_1) - x_1^i|^p, \tag{3.2.1}$$

gdzie $\alpha > 0$ jest bardzo dużą liczbą, zaś dla p przyjmuje się najczęściej jedną z wartości $p = 2, 1, \frac{1}{2}$. Rezygnujemy przy tym z żądania, aby w chwili t_1 obiekt znalazł się w punkcie $\boldsymbol{x}_1$ lecz żądamy, aby funkcjonał F_1 miał minimalną wartość. Jest widoczne, że przy dużych wartościach współczynnika α punkt $\boldsymbol{x}(t_1)$ musi leżeć blisko punktu $\boldsymbol{x}_1$, jeżeli F_1 ma mieć minimalną wartość.

(B) W podobny sposób można pozbyć się ograniczeń nałożonych na wektor fazowy lub na wektor sterujący. Przypuśćmy, że ograniczenia te mają postać

$$G_j(\boldsymbol{x}, \boldsymbol{u}, t) \leqslant 0, \quad j = 1, 2, \ldots, m, \tag{3.2.2}$$

gdzie G_j są pewnymi funkcjonałami. Wprowadźmy funkcjonał

$$F_2(\boldsymbol{x}, \boldsymbol{u}, t) = F_1(\boldsymbol{x}, \boldsymbol{u}) + \beta \sum_{j=1}^{m} H_j(\boldsymbol{x}, \boldsymbol{u}, t), \tag{3.2.3}$$

gdzie $\beta > 0$ jest bardzo dużą liczbą, zaś

$$H_j \begin{cases} \equiv 0 & \text{dla } G_j \leqslant 0, \\ = G_j \quad \text{lub} \quad = G_j^2 & \text{dla } G_j > 0, \ j = 1, \ldots, m. \end{cases} \tag{3.2.4}$$

Zamiast żądać, aby były spełnione nierówności (3.2.2), wymagamy, aby funkcjonał F_2 miał minimalną wartość. Jest widoczne, że gdy β jest bardzo duże, znalezione optymalne sterowania i optymalny tor tylko bardzo nieznacznie mogą wyjść poza obszar określony warunkami (3.2.2).

Tego rodzaju zmiana kryterium optymalności upraszcza zagadnienie matematyczne. Natomiast odpowiedź na pytanie, o ile otrzymane rozwiązanie odchyla się od optymalnego rozwiązania dla pierwotnego problemu minimalizacji funkcjonału F, może być trudna.

3.3. Redukcja problemu sterowania optymalnego do problemu czasowo optymalnego. Jeśli szukamy minimum funkcjonału

$$F(\boldsymbol{x}, \boldsymbol{u}) = \int_{t_0}^{t_1} f^0\big(\boldsymbol{x}(t), \boldsymbol{u}(t)\big)\, dt,$$

gdzie $f^0(\boldsymbol{x}, \boldsymbol{u}) > 0$ i f^0 jest ciągłą funkcją zmiennych $\boldsymbol{x}$ i $\boldsymbol{u}$, to można postąpić w sposób następujący. Oznaczmy

$$\tau(t) = \int_{t_0}^{t} f^0\big(\boldsymbol{x}(s), \boldsymbol{u}(s)\big)\, ds.$$

Wówczas $\frac{d\tau}{dt} = f^0\big(\boldsymbol{x}(t), \boldsymbol{u}(t)\big) > 0$. Wobec tego funkcja $\tau(t)$ ma funkcję odwrotną $t(\tau)$ i $\frac{dt}{d\tau} = \big[f^0\big(\boldsymbol{x}(t), \boldsymbol{u}(t)\big)\big]^{-1}$.

Oznaczmy

$$\boldsymbol{y}(\tau) = \boldsymbol{x}\big(t(\tau)\big),\ \boldsymbol{v}(\tau) = \boldsymbol{u}\big(t(\tau)\big),\ \boldsymbol{y}(0) = \boldsymbol{x}(t_0),\ \boldsymbol{y}(\tau_1) = \boldsymbol{x}\big(t(\tau_1)\big) = \boldsymbol{x}(t_1).$$

Jeśli równanie różniczkowe opisujące proces sterowany ma postać

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t),$$

to otrzymamy

$$\frac{d\boldsymbol{y}}{d\tau} = \boldsymbol{f}\big(\boldsymbol{y}, \boldsymbol{v}, t(\tau)\big) \cdot [f^0(\boldsymbol{y}, \boldsymbol{v})]^{-1}. \tag{3.3}$$

Funkcjonał F będzie równy $F(\boldsymbol{x}, \boldsymbol{u}) = \tau(t_1) = \tau_1$.

Zagadnienie może być sformułowane obecnie następująco: przeprowadzić obiekt opisany równaniem (3.3) z położenia $\boldsymbol{y}(0)$ w położenie $\boldsymbol{y}(\tau_1)$ w minimalnym czasie τ_1. W ten sposób bardzo ogólna klasa problemów może być sprowadzona do problemu czasowo optymalnego.

4. Metoda pola orientorowego

Autorem metody pola orientorowego jest T. Ważewski (por. [29], [32], [33], [34], [35] i [36]). Rozważmy zagadnienie sterowania procesem opisanym przez równanie

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}), \tag{4.1}$$

przy czym spełniony jest warunek

$$\boldsymbol{u}(t) \in U\big(t, \boldsymbol{x}(t)\big). \tag{4.2}$$

Problem jest tu postawiony ogólniej niż się to zwykle robi, gdyż przyjmujemy, że zbiór U (zwany *dziedziną sterowania*) zmienia się w czasie. Wobec tego odpowiedź na pytanie, które sterowanie jest dopuszczalne, zależy od czasu t i od pozycji w przestrzeni fazowej $\boldsymbol{x}(t)$. Przypadek zwykle rozpatrywany, gdy dziedzina sterowania U jest niezmienna, stanowi szczególny przypadek objęty obecnym rozumowaniem. Geometryczna interpretacja równań (4.1) i (4.2) jest następująca: równanie

(4.1) przy spełnieniu warunku (4.2) nie wyznacza jednej wartości pochodnej $\dot{x}(t)$ w chwili t. Wiadomo tylko, że $\dot{x}(t)$ należy do pewnego zbioru wektorów, który oznaczymy przez $N(t, x)$:

$$N(t, x) = \{v: v = f(t, x, u),\ u \in U(t, x(t))\}, \tag{4.3}$$

czyli słowami: $N(t, x)$ jest zbiorem wszystkich wektorów v spełniających równość $v = f(t, x, u)$, gdzie u jest dowolnym elementem zbioru $U(t, x(t))$. Definicję (4.3) można zapisać krócej w postaci

$$N(t, x) = f(t, x, U). \tag{4.4}$$

Zbiór $N(t, x)$ będziemy nazywali (za Ważewskim) *orientorem*; ponieważ $N(t, x)$ określone jest dla wszelkich par (t, x), mamy określone *pole orientorowe*. Równości (4.1), (4.2) i (4.4) można zanotować krótko

$$\dot{x}(t) \in N(t, x). \tag{4.5}$$

Warunek (4.5) stanowi uogólnienie równania różniczkowego. Gdyby zbiór $N(t, x)$ składał się z jednego tylko wektora, warunek (4.5) byłby zwykłym równaniem różniczkowym. Ogólnie warunek (4.5) określa pewną „miotełkę" kierunków prędkości wektora fazowego $x(t)$.

Warunek (4.5) był badany przez S. K. Zarembę [37] pod nazwą *równania paratyngensowego*.

Trajektorią pola orientorowego $N(t, x)$ nazywamy każdą krzywą $\tilde{x}(t) = (t, x(t))$, absolutnie ciągłą, która prawie wszędzie spełnia warunek (4.5). Zwróćmy uwagę, że $\tilde{x} \in R^{n+1}$, podczas gdy $x \in R^n$. Trzeba więc odróżniać trajektorie pola orientorowego od torów leżących w przestrzeni fazowej, określonych w 2.1.

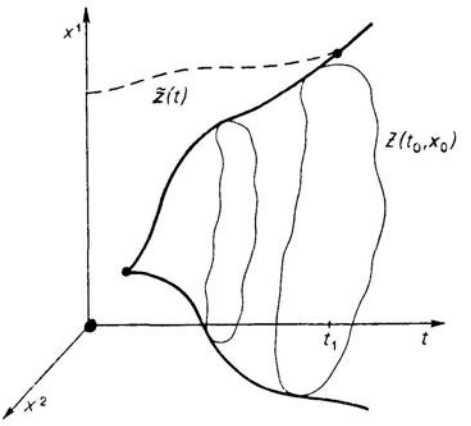
Strefą emisji wychodzącą z punktu x_0 w chwili t_0 nazywamy zbiór punktów $\tilde{x}(t) \in R^{n+1}$, spełniających warunki: (a) $t \geqslant t_0$; (b) $\tilde{x}(t)$ jest trajektorią pola orientorowego; (c) $x(t_0) = x_0$. Strefę emisji oznaczamy przez $Z(t_0, x_0)$. Jej brzeg (pobocznicę) oznaczamy przez $B(t_0, x_0)$.

Rozważmy zagadnienie czasowo optymalne: chcemy, wychodząc z punktu $\tilde{x}(t_0) = (t_0, x_0)$, doścignąć w najkrótszym czasie ruchomy punkt $\tilde{z}(t) \in R^{n+1}$. Nastąpi to tylko wtedy, gdy tor $\tilde{z}(t)$ ma punkty wspólne ze strefą $Z(t_0, x_0)$. Niech minimalny czas, dla którego $\tilde{z}(t) \in Z(t_0, x_0)$ wynosi t_1. Jest wtedy $\tilde{z}(t_1) \in B(t_0, x_0)$ (rys. 3). Zagadnienie czasowo optymalne będzie rozwiązane, gdy 1° wyznaczymy czas t_1; 2° wyznaczymy sterowanie $u(t)$ takie, że całka równania (4.1) $x(t)$ spełnia równość

$$\tilde{x}(t_1) = (t_1, x(t_1)) = \tilde{z}(t_1).$$

Jak widać, jeśli będziemy umieli wyznaczyć brzeg $B(t_0, x_0)$, to wyznaczenie czasu t_1 sprowadzi się do wyznaczenia pierwszego punktu wspólnego $z(t)$ z tym brzegiem.

Nasuwa się przypuszczenie, że trajektoria optymalna powinna w całości przebiegać po brzegu $B(t_0, x_0)$. Zostało dowiedzione, że istnieje trajektoria pola orientorowego $N(t, x)$, położona w całości na brzegu strefy emisji, jeżeli spełnione są następujące warunki: (a) zbiór $N(t, x)$



Rys. 3

jest określony dla każdego punktu (t, x), gdy $t \geqslant t_0$ i jest domknięty; (b) zbiory $N(t, x)$ są wszystkie zawarte w pewnej n-wymiarowej kuli o promieniu nie zależnym od t i x; znaczy to, że istnieje skończony kres górny długości wektorów należących do wszystkich zbiorów $N(t, x)$; (c) $N(t, x)$ zmienia się w sposób półciągły w zależności od (t, x) (rozumiemy przez to, że gdy $(t, x) \to (t_0, x_0)$, to maksimum odległości punktów zbioru $N(t, x)$ od zbioru $N(t_0, x_0)$ dąży do 0); (d) zbiory $N(t, x)$ są wypukłe (por. [37]).

Istotną trudność przy stosowaniu twierdzenia o trajektorii pola orientorowego sprawia warunek (d), gdyż w najważniejszych zagadnieniach optymalnego sterowania, spotykanych w praktyce, warunek (d) może nie być spełniony. Trudność tę można pokonać w następujący sposób:

Wprowadza się pojęcie quasi-trajektorii pola orientorowego: *quasi-trajektorią pola orientorowego* $N(t, x)$ jest funkcja $\tilde{x}(t)$, absolutnie ciągła, o tej własności, że istnieje ciąg funkcji absolutnie ciągłych $\{x_n\}$ taki, że dla $n \to \infty$, $\tilde{x}_n(t) \to \tilde{x}(t)$, a przy tym odległość punktu $\tilde{x}_n'(t)$ od zbioru $N\big(t, x_n(t)\big) \to 0$.

Funkcje ciągu $\{\tilde{x}_n\}$ nie muszą być trajektoriami, gdyż $x_n'(t)$ nie musi należeć do $N\big(t, x_n(t)\big)$.

Jeżeli przez $E(t, x)$ oznaczymy osłonę wypukłą zbioru $N(t, x)$, to znaczy najmniejszy zbiór wypukły domknięty zawierający zbiór $N(t, x)$, to przy założeniach (a), (b), (c) można dowieść (por. [36]), że zbiór quasi-trajektorii pola orientorowego $N(t, x)$ jest identyczny ze zbiorem trajektorii pola orientorowego $E(t, x)$.

Dla zastosowań praktycznych zastąpienie trajektoryj przez quasi-trajektorie jest dopuszczalne, gdyż dotarcie do dostatecznie małego otoczenia celu jest praktycznie zupełnie wystarczające.

Dla zbioru wypukłego $E(t, x)$ istnieje ściśle określony zbiór $Q(t, x)$, będący najmniejszym zbiorem, dla którego osłona wypukła jest identyczna ze zbiorem $E(t, x)$. Np. jeżeli $E(t, x)$ jest wielościanem wypukłym, to $Q(t, x)$ będzie zbiorem jego naroży. Zbiór $Q(t, x)$ nazywamy *tendorem zbioru* $N(t, x)$.

Można udowodnić (por. [36]), że zbiór wszystkich trajektorii pola orientorowego $E(t, x)$ jest identyczny ze zbiorem wszystkich trajektorii pola orientorowego $Q(t, x)$.

Znaczenie tego twierdzenia jest takie: jeżeli ograniczymy się przy wyborze sterowania do sterowań spełniających warunek $u(t) \in Q(t, x)$, to otrzymamy wszystkie trajektorie i quasi-trajektorie odpowiadające sterowaniom spełniającym warunek $u(t) \in N(t, x)$. Metoda sterowania „bang-bang" sprowadza się właśnie do tego, że wartość funkcji sterującej należy do tendora zbioru $N(t, x)$.

Od quasi-trajektorii należy odróżnić pojęcie toru poślizgowego, wprowadzone przez Filippowa. *Torem poślizgowym pola orientorowego* $N(t, x)$ nazywamy funkcję $\tilde{x}(t)$ absolutnie ciągłą, dla której istnieje ciąg trajektorii pola orientorowego $N(t, x)\{\tilde{x}_n(t)\}$ taki, że $\tilde{x}_n(t) \to \tilde{x}(t)$ dla $n \to \infty$. Można dowieść (por. [29]), że istnieją quasi-trajektorie nie będące torami poślizgowymi. Znane są również warunki dostateczne na to, aby każda quasi-trajektoria była jednocześnie torem poślizgowym (por. [32]).

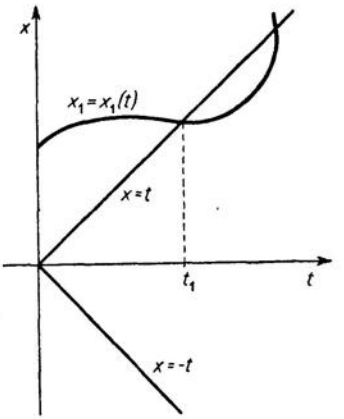
Metoda pól orientorowych eliminuje funkcję sterującą i wysuwa na pierwszy plan wyznaczenie trajektorii optymalnych. Gdy trajektoria optymalna jest wyznaczona, zagadnienie wyznaczenia optymalnego sterowania jest znacznie łatwiejsze. Trajektorii optymalnej poszukuje się na brzegu strefy emisji, który należy naprzód wyznaczyć.

4.2. Przykład. Rozważmy równanie różniczkowe nieautonomiczne

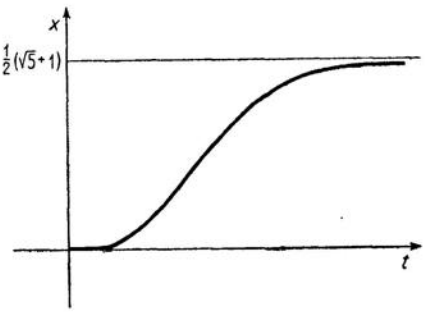
$$\dot{x} = -x^2+t^2+u(t), \qquad x(0) = 0 \ (x \text{ skalar}). \tag{4.2.1.}$$

Załóżmy, że funkcja sterująca $u(t)$ może przyjmować tylko wartości $+1$ i -1. Zatem orientor $N(t, x)$ składa się z dwu punktów $(-x^2+t^2+1)$ oraz $(-x^2+t^2-1)$. Osłona wypukła $E(t, x)$ jest odcinkiem na płaszczyźnie (t, x) łączącym oba punkty orientora. Tendor $Q(t, x)$ jest identyczny z $N(t, x)$. Dla każdego problemu sterowania optymalnego procesem opisanym przez (4.2.1) jest sterowanie „bang-bang", co jest oczywiste, skoro istnieją tylko dwie dopuszczalne wartości funkcji sterującej. Brzeg strefy emisji otrzymamy kładąc stale albo $u = 1$, albo $u = -1$. W pierwszym przypadku jedynym rozwiązaniem równania (4.2.1) jest $x(t) = t$, a drugim będzie $x(t) = -t$. Zatem strefą emisji jest domknięty obszar

określony warunkami $-t \leqslant x \leqslant t$, $t \geqslant 0$. Przestrzeń fazowa jest jednowymiarowa. Do punktu przestrzeni fazowej $x = x_1$ docieramy w minimalnym czasie, który graficznie można wyznaczyć przez przecięcie brzegu strefy emisji z prostą $x - x_1$. Czas minimalny wynosi $t_1 = |x_1|$. Jeżeli



Rys. 4



Rys. 5

cel, który ma być osiągnięty jest ruchomy, $x_1 = x_1(t)$, to będzie osiągnięty, jeśli krzywa $x_1(t)$ przecina brzeg strefy emisji. Czas minimalny wyznaczony jest przez pierwszy punkt przecięcia (rys. 4).

4.3. Przykład. Weźmy teraz równanie różniczkowe autonomiczne

$$\dot{x} = x - x^2 + u(t), \qquad x(0) = 0 \quad (x \text{ skalar}). \tag{4.3.1}$$

Przyjmiemy, że $0 \leqslant u(t) \leqslant 1$. Orientor jest odcinkiem łączącym punkty $x - x^2$, $x - x^2 + 1$ leżące na osi x. Tendor składa się z tych dwu punktów. Brzeg strefy emisji otrzymujemy kładąc stale $u(t) = 0$ lub $u(t) = 1$. W pierwszym przypadku otrzymujemy rozwiązanie $x(t) \equiv 0$, w drugim

$$x(t) = \frac{1-\sqrt{5}}{2} + \frac{3\sqrt{5}-5}{(3-\sqrt{5}) + 2e^{-t\sqrt{5}}}. \tag{4.3.2}$$

Krzywa (4.3.2) ma asymptotę poziomą $x = \dfrac{1+\sqrt{5}}{2}$, $x(t)$ jest stale rosnące i ma jeden punkt przegięcia (rys. 5).

Strefa emisji leży między dodatnią osią t a krzywą (4.3.2). Cel stały $x = x_1$ będzie osiągnięty, jeżeli $0 < x_1 < \dfrac{\sqrt{5}+1}{2}$, cel ruchomy, jeżeli krzywa $x_1 = x_1(t)$ ma punkty wspólne z brzegiem strefy emisji. Odcięta pierwszego z tych punktów przecięcia równa jest czasowi minimalnemu. Trajektoria optymalna leży na brzegu.

5. Metoda obszaru osiągalnego

5.1. Niech proces sterowania będzie opisany równaniem

$$x(t) = x_0(t) + \int_{t_0}^{t} h(t, \tau) u(\tau)\, d\tau, \tag{5.1.1}$$

gdzie x, x_0, h są wektorami n-wymiarowymi, zaś u jest funkcją sterującą skalarną. Funkcja $x_0(t)$ jest zadana i ciągła dla $t \geqslant 0$. (5.1.1) można uważać za rozwiązanie równania różniczkowego

$$\dot{x}(t) = A(t)x(t) + b(t) \cdot u(t), \tag{5.1.2.}$$

gdzie A jest macierzą $n \times n$, zaś b jest n-wymiarowym wektorem; zatem problem rozważany należy do klasy zagadnień sterowania procesami opisanymi przez równanie różniczkowe zwyczajne liniowe. W dalszym ciągu będziemy posługiwali się wyłącznie równaniem (5.1.1), nie odwołując się do równania (5.1.2).

Zakładamy, że składowe wektora h, funkcje $h^i(t, \tau)$, są ograniczone w każdym skończonym przedziale zbioru $t_0 \leqslant \tau \leqslant t < \infty$, są ciągłe ze względu na t i kawałkami ciągłe ze względu na τ. Następnie zakładamy, że funkcje $h^i(t, \tau)$ są liniowo niezależne w dowolnym przedziale $t_0 \leqslant \tau \leqslant t$, czyli że dla dowolnego niezerowego wektora λ mamy:

$$\lambda^T \cdot h(t, \tau) = \sum_{i=1}^{n} \lambda^i h^i(t, \tau) \not\equiv 0 \quad \text{dla} \quad t_0 \leqslant \tau \leqslant t \tag{5.1.3}$$

(symbol λ^T oznacza transpozycję wektora λ, który jest macierzą jednokolumnową; λ^T jest macierzą jednowierszową).

Jeżeli iloczyn skalarny (5.1.3) zeruje się tylko w zbiorze punktów izolowanych, to proces sterowany nazywamy *normalnym*. Sterowaniami dopuszczalnymi będą sterowania czyniące zadość warunkowi:

$$\left(\int_{t_0}^{T} |u(t)|^p\, dt \right)^{1/p} \leqslant C_p, \quad 1 \leqslant p \leqslant \infty, \tag{5.1.4}$$

gdzie liczba p jest zadana.

Szczególnie ważne przypadki otrzymujemy dla $p = 1$, $p = 2$ i $p = \infty$, w tym ostatnim przypadku warunek (5.1.4) sprowadza się przy sterowaniach $u(t)$ kawałkami ciągłych do warunku $\max_{t_0 \leqslant t \leqslant T} u(t) \leqslant C_\infty$. W zastosowaniach mamy do czynienia zwykle ze sterowaniami kawałkami ciągłymi. Przejmiemy jednak ogólniejsze założenie, że $u(t)$ jest funkcją mierzalną w sensie Lebesgue'a.

W przypadku $p = 1$ będziemy zakładali, że $h(t, \tau)$ jest ciągłe ze względu na obie zmienne.

5.2. Cel sterowania. Celem tym jest znalezienie sterowania dopuszczalnego takiego, aby dla danej funkcji ciągłej $\boldsymbol{y}(t)$ równość

$$\int_{t_0}^{t} \boldsymbol{h}(t, \tau)\, u(\tau)\, d\tau = \boldsymbol{y}(t) \tag{5.2}$$

zachodziła dla minimalnego czasu $t = T^0$.

Inaczej mówiąc idzie o to, aby w minimalnym czasie przeprowadzić wektor fazowy z pozycji $\boldsymbol{x}_0(t_0)$ do pozycji $\boldsymbol{x}_0(t) + \boldsymbol{y}(t)$.

5.3. Obszar osiągalny. Załóżmy, że $\boldsymbol{x}_0(t_0) = 0$. Zbiór punktów $\boldsymbol{x}(T)$ otrzymanych przy sterowaniu wszelkich dopuszczalnych sterowań nazywamy *obszarem osiągalnym* i oznaczamy przez $R(t_0, T)$. Zwróćmy uwagę na różnicę definicji obszaru osiągalnego i strefy emisji. Strefa emisji leżała w przestrzeni $(n+1)$-wymiarowej o punktach $(t, \boldsymbol{x})$, natomiast obszar osiągalny R leży w przestrzeni fazowej n-wymiarowej. $R(t_0, T)$ jest przekrojem strefy emisji płaszczyzną $t = T$.

Obszar $R(t_0, T)$ zależy od p i od C_p. Można dowieść, że dla $T > t_0$ jest zbiorem domkniętym, ograniczonym, symetrycznym względem punktu $\boldsymbol{x} = 0$, ciągłym względem T, i że jest ciałem wypukłym, tzn. że jest nie tylko zbiorem wypukłym, lecz że nadto zawiera punkty, których całe n-wymiarowe otoczenie do $R(t_0, T)$ należy.

Przez ciągłość względem T rozumiemy następującą własność: jeżeli punkt $\boldsymbol{x}_1$ leży zewnątrz $R(t_0, T)$, to dla dostatecznie małego $\delta > 0$ będzie leżał zewnątrz $R(t_0, T+\delta)$.

Proces sterowania możemy przedstawić w następujący sposób: $\boldsymbol{y}(T)$ jest punktem osiągalnym, jeżeli dla pewnego skończonego T, $\boldsymbol{y}(T) \in R(t_0, T)$.

Z ciągłości wynika, że pierwsze zetknięcie się $\boldsymbol{y}(T)$ z $R(t_0, T)$ może nastąpić tylko na brzegu obszaru $R(t_0, T)$, który oznaczymy przez $\partial R(t_0, T)$ i ono wyznacza minimalny czas T^0.

Mamy więc twierdzenie: *Jeśli przy pewnym skończonym $T = T_1$ punkt $\boldsymbol{y}(T_1) \in R(t_0, T_1)$, to istnieje optymalny czas T^0, $T^0 \leqslant T_1$, taki, że $\boldsymbol{y}(T^0) \in \partial R(t_0, T^0)$, zaś dla $T < T^0$ jest $\boldsymbol{y}(T) \notin R(t_0, T)$. Warunkiem koniecznym, aby sterowanie było optymalne, jest by* (5.1.4) *było spełnione ze znakiem równości.*

Oznaczmy przez $S(\boldsymbol{\lambda})$ hiperpłaszczyznę oporową hiperpowierzchni ∂R, daną równaniem $\boldsymbol{\lambda}^T \cdot \boldsymbol{x} = c$.

Można dowieść, że warunkiem koniecznym i dostatecznym na to, aby za pomocą sterowania $u(t)$ można było dotrzeć do dowolnego punktu zbioru $S(\boldsymbol{\lambda}) \cap \partial R(t_0, T)$ jest, aby iloczyn skalarny

$$\boldsymbol{\lambda}^T \cdot \boldsymbol{x}(T) = \int_{t_0}^{T} \boldsymbol{\lambda}^T \cdot \boldsymbol{h}(T, \tau)\, u(\tau)\, d\tau \tag{5.3.1}$$

osiągał dla tego sterowania maksymalną wartość (przy zadanym $\boldsymbol{\lambda}$).

Maksymalizacja wyrażenia (5.3.1) daje sterowanie optymalne, jako funkcję czasu zależną od parametrów: $\boldsymbol{\lambda}, t_0, T^0, C_p$: $u^0 = u^0(t; \boldsymbol{\lambda}, t_0, T_0 C_p)$. Trzeba więc znaleźć T^0 (o ile istnieje) oraz $\boldsymbol{\lambda}$ odpowiadające punktowi $\boldsymbol{y}(T_0) \in \partial R(t_0, T^0)$. Do tego celu służy twierdzenie:

Warunkiem koniecznym i dostatecznym na to, aby punkt $\boldsymbol{y}(t)$ *leżał na* $\partial R(t_0, t)$ *w momencie* $t = T$ *jest, aby*

$$\max_{\boldsymbol{\lambda}} \left[\frac{\boldsymbol{\lambda}^T \boldsymbol{y}(t)}{\int_{t_0}^{t} \boldsymbol{\lambda}^T \boldsymbol{h}(t, \tau) u^0(\tau; \boldsymbol{\lambda}, t_0, t, C_p) d\tau} \right]_{t=T} = 1. \tag{5.3.2}$$

Punkt $\boldsymbol{y}(t)$ jest wewnętrznym lub zewnętrznym punktem zbioru $R(t_0, t)$, zależnie od tego, czy lewa strona (5.3.2) jest mniejsza czy większa niż 1. Przy wzrastaniu t, pierwsza wartość T, spełniająca (5.3.2) dla $\boldsymbol{\lambda} = \boldsymbol{\lambda}^0$, jest optymalnym czasem T^0, zaś $\boldsymbol{\lambda}^0$ spełnia warunek $\boldsymbol{y}(T^0) \in S(\boldsymbol{\lambda}^0) \cap \partial R(t_0, T^0)$.

Widzimy więc, że rozwiązanie zagadnienia sterowania optymalnego istnieje, gdy równość (5.3.2) jest spełniona przy pewnym skończonym T.

5.4. Na brzegu ∂R odróżniamy następujące rodzaje punktów: (a) punkty, w których istnieje jedyna zewnętrzna normalna do ∂R, $\boldsymbol{\lambda}$, i jedna hiperpłaszczyzna $S(\boldsymbol{\lambda})$ styczna do ∂R; te punkty nazywamy *punktami gładkości* brzegu ∂R; (b) punkty kątowe rzędu k; są to *punkty przecięcia hiperpłaszczyzny oporowej* $S(\boldsymbol{\lambda})$ z ∂R takie, że wektor $\boldsymbol{\lambda}$ może się zmieniać wewnątrz k-wymiarowego stożka; jest więc $2 \leqslant k \leqslant n$ (dla $n = 1$ mamy punkt gładkości).

Jeżeli zbiór $S(\boldsymbol{\lambda}) \cap \partial R$ zawiera więcej niż jeden punkt, to (wobec wypukłości zbioru R) musi zawierać albo odcinek, albo pewien zbiór płaski, albo ogólnie kawałek l-wymiarowej hiperpłaszczyzny, gdzie $1 \leqslant l \leqslant n-1$. Przy tym przez l rozumiemy najwyższy wymiar hiperpłaszczyzny zawierającej $S(\boldsymbol{\lambda}) \cap \partial R$. Zbiór $S(\boldsymbol{\lambda}) \cap \partial R$ nazywamy wtedy *l-wymiarową krawędzią brzegu* ∂R. Jeśli ∂R ma krawędź, to sterowanie $u(t)$ maksymalizujące wyrażenie (5.3.1) nie jest wyznaczone jednoznacznie.

5.5. Można dowieść twierdzenia: *Sterowanie dopuszczalne jest tylko wtedy optymalne, gdy*

$$u^0(t) = \frac{C_p |\boldsymbol{\lambda}^T \cdot \boldsymbol{h}(T, t)|^{q-1} \cdot \operatorname{sgn}(\boldsymbol{\lambda}^T \cdot \boldsymbol{h}(T, t))}{\left(\int_{t_0}^{T} |\boldsymbol{\lambda}^T \boldsymbol{h}(T, \tau)|^q d\tau \right)^{1/p}}, \tag{5.5.1}$$

gdzie

$$\frac{1}{p} + \frac{1}{q} = 1, \quad 1 < p \leqslant \infty.$$

Dla $p = \infty$, $q = 1$, równość (5.5.1) ma postać

$$u^0(t) = C_\infty \operatorname{sgn}\big(\boldsymbol{\lambda}^T \boldsymbol{h}(T, t)\big).$$

W przypadku $p = 1$ ($q = \infty$) wzór (5.5.1) pozostaje prawdziwy, jeśli prawą jego stronę rozumiemy jako granicę otrzymaną dla $q \to \infty$. Okazuje się, że wtedy sterowanie optymalne $u^0(t)$ jest dystrybucją, jeśli proces sterowania jest normalny. Wzór dla sterowania $u^0(t)$ może być zapisany za pomocą dystrybucji δ Diraca. Wobec (5.5.1) wzór (5.3.2) można zapisać w postaci

$$\left[\max_{\boldsymbol{\lambda}} \frac{\boldsymbol{\lambda}^T \boldsymbol{y}(t)}{\left(\int_{t_0}^{T} |\boldsymbol{\lambda}^T \boldsymbol{h}(t, \tau)|^q d\tau\right)^{1/q}}\right]_{t=T} = C_p. \tag{5.5.2}$$

Długość wektora $\boldsymbol{\lambda}$ nie odgrywa roli, więc można przyjąć, że $\boldsymbol{\lambda}^T \boldsymbol{y}(t) = 1$ i wtedy mamy

$$\left[\min_{\boldsymbol{\lambda}} \left(\int_{t_0}^{t} |\boldsymbol{\lambda}^T \boldsymbol{h}(t, \tau)|^q d\tau\right)^{1/q}\right]_{t=T} = C_p^{-1}. \tag{5.5.3}$$

Wzory (5.5.1), (5.5.2) i (5.5.3) zostały wyprowadzone przez R. Kulikowskiego.

Można udowodnić, że dla $1 < p < \infty$ na $\partial R(t_0, T)$ nie istnieją punkty kątowe.

5.6. Sterowanie wielowymiarowe. Zakładamy obecnie, że funkcją sterującą jest r-wymiarowy wektor $\boldsymbol{u}^j(t)$ o składowych $u^j(t)$, $j = 1, \ldots, r$. Proces sterowania jest odpisany równaniem

$$\boldsymbol{x}(t) = \boldsymbol{x}_0(t) + \int_{t_0}^{t} H(t, \tau)\boldsymbol{u}(\tau)\, d\tau = \boldsymbol{x}_0(t) + \sum_{j=1}^{r} \int_{t_0}^{t} \boldsymbol{h}_j^T(t, \tau) u^j(\tau)\, d\tau, \tag{5.6.1}$$

gdzie $\boldsymbol{h}_j(t, \tau)$ jest j-tą kolumną (wektorem) $n \times r$-wymiarowej macierzy $H(t, \tau)$. Dla elementów macierzy H oznaczonych przez $h_{ij}(t, \tau)$ zakłada się, że spełniają warunki analogiczne do przyjętych w przypadku sterowania jednowymiarowego.

Wymienimy dwa rodzaje ograniczeń dla sterowania $\boldsymbol{u}(t)$:

(a) ograniczenie indywidualne typu

$$\left(\int_{t_0}^{T} |u^j(t)|^{p_j} dt\right)^{1/p_j} \leqslant C_j, \quad j = 1, \ldots, r, \ 1 \leqslant p_j \leqslant \infty;$$

(b) ograniczenie wspólne

$$\left(\int_{t_0}^{T} \sum_{j=1}^{r} |u^j(t)|^p dt\right)^{1/p} \leqslant C_p, \quad 1 \leqslant p \leqslant \infty.$$

W przypadku ograniczeń indywidualnych dla każdej składowej $u^j(t)$ możemy, wobec liniowości równania (5.6.1), zastosować zasadę superpozycji, to znaczy rozwiązać zagadnienie sterowania optymalnego dla każdej składowej $u^j(t)$ z osobna, kładąc 0 za pozostałe składowe, przy czym z równania (5.6.1) otrzymamy wektor $x_j(t)$. Zachodzi wzór $x(t) = \sum_{j=1}^{r} x_j(t)$.

W przypadku wspólnego ograniczenia można udowodnić wzory analogiczne do wzorów (5.5.1) i (5.5.2) przy założeniu $1 < p < \infty$:

$$u^j_{\mathrm{opt}}(t) = \frac{C_p |\lambda^T h_j(T, t)|^{q-1} \operatorname{sgn}\left(\lambda^T h_j(T, t)\right)}{\left(\int_{t_0}^{T} \sum_{j=1}^{r} |\lambda^T h(T, \tau)|^q d\tau\right)^{1/p}}, \tag{5.6.2}$$

$$\left[\max_{\lambda} \frac{\lambda^T y(t)}{\left(\int_{t_0}^{t} \sum_{j=1}^{r} |\lambda^T h_j(t, \tau)|^q d\tau\right)^{1/q}}\right]_{t=T} = C_p, \qquad \frac{1}{p} + \frac{1}{q} = 1. \tag{5.6.3}$$

Przypadki $p = 1$ i $p = \infty$, więcej skomplikowane, pomijamy.

6. Zasada maksimum

6.1. Sformułujemy zasadę maksimum pochodzącą od L. S. Pontriagina i jego uczniów. Podane sformułowanie nie pokrywa się z oryginalnym ujęciem Pontriagina, które można znaleźć w [30]. Ograniczymy się do ujęcia dogodniejszego w zastosowaniach (por. [5]). Zasada będzie sformułowana dla problemu czasowo optymalnego. Proces sterowania jest opisany równaniem

$$\dot{x} = f(x, u), \tag{6.6.1}$$

gdzie $x(t) \in R^n$, $u(t) \in U \subset R^r$. U nazywamy *dziedziną sterowania*. Sterowaniem dopuszczalnym będzie każda funkcja $u(t)$ kawałkami ciągła o wartościach należących do U, ciągła na końcach odcinka $[t_0, t_1]$, w którym jest określona.

Dane są w R^n dwa punkty x_0, x_1, początkowa i końcowa pozycja wektora fazowego. Szukamy sterowania przeprowadzającego wektor fazowy z x_0 do x_1 w minimalnym czasie, a więc $x(t_0) = x_0$, $x(t_1) = x_1$. Mamy znaleźć sterowanie, dla którego $t_1 - t_0$ ma minimalną wartość (t_0 ustalone).

Wprowadźmy pomocniczą funkcję H, zależną od x, u oraz n-wymiarowego wektora $\psi(t)$:

$$H(\psi, x, u) = \psi^T f = \sum_{k=1}^{n} \psi^k f^k(x, u). \tag{6.1.2}$$

Wektor $\psi(t)$ ma spełniać równanie różniczkowe

$$(6.1.3)\qquad \dot{\psi}(t) = -\frac{\partial H}{\partial x}, \quad \text{czyli} \quad \dot{\psi}^k(t) = -\frac{\partial H}{\partial x^k}, \quad k = 1, \ldots, n.$$

Przy tych oznaczeniach i założeniach mamy twierdzenie:

Warunkiem koniecznym optymalności procesu sterowania jest istnienie takiego niezerowego rozwiązania równania (6.1.3) *w przedziale* $[t_0, t_1]$, *by w każdej chwili* τ, *w której* $u(t)$ *jest ciągłe, zachodził związek*

$$(6.1.4)\qquad H\big(\psi(\tau), x(\tau), u(\tau)\big) = \max_{v \in U} H\big(\psi(\tau), x(\tau), v\big),$$

a przy tym zachodziła nierówność

$$(6.1.5)\qquad H\big(\psi(t_1), x(t_1), u(t_1)\big) \geqslant 0.$$

6.2. Przykład zastosowania zasady maksimum. Masa m znajduje się w chwili $t = 0$ w punkcie $x = 0$ w spoczynku. Chcemy ją przesunąć po linii prostej poziomej (bez tarcia) w minimalnym czasie t_1 do punktu $a > 0$, przy czym prędkość końcowa ma wynosić 0. Zadanie polega na znalezieniu siły $f(t)$ jako funkcji czasu, przy założeniu $|f(t)| \leqslant K$. Mamy więc warunki:

$$m\ddot{x} = f(t), \quad x(0) = 0, \quad \dot{x}(0) = 0, \quad x(t_1) = a, \quad \dot{x}(t_1) = 0.$$

Oznaczmy $x = x_1$, $\dot{x} = x_2$, $\frac{1}{m} f(t) = u(t)$, $\frac{K}{m} = M$. Mamy więc układ równań:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u(t) \end{cases} \quad \text{przy warunkach} \quad \begin{cases} x_1(0) = 0, & x_1(t_1) = a, \\ x_2(0) = 0, & x_2(t_1) = 0, \\ |u(t)| \leqslant M. \end{cases}$$

Tworzymy funkcję $H = x_2\psi_1 + u\psi_2$. Jest $\dot{\psi}_1 = 0$, $\dot{\psi}_2 = -\psi_1$. Stąd $\psi_1(t) = c_1$, $\psi_2(t) = c_2 - c_1 t$, $H = x_2 c_1 + u(c_2 - c_1 t)$.

Maksimum funkcji H otrzymamy, przyjmując $u(t) = M \cdot \operatorname{sgn}(c_2 - c_1 t)$. Łatwo wyliczyć, że

$$x_1(t) = \int_0^t (t-s)\,u(s)\,ds, \qquad x_2(t) = \int_0^t u(s)\,ds.$$

Z danych warunków otrzymujemy $\int_{t_0}^{t_1} \operatorname{sgn}(c_2 - c_1 s)\,ds = 0$. Stąd wynika, że $c_1 \neq 0$ i że $c_2 - c_1 s$ musi zmienić znak w przedziale $[0, t_1]$, co może nastąpić tylko dla $s = c_2/c_1$. Jest więc

$$\operatorname{sgn}(c_2 - c_1 s) = \operatorname{sgn} c_2 \qquad \text{dla} \quad 0 \leqslant s < \frac{c_2}{c_1},$$

$$\operatorname{sgn}(c_2 - c_1 s) = -\operatorname{sgn} c_2 \qquad \text{dla} \quad \frac{c_2}{c_1} < s \leqslant t_1.$$

Wobec tego

$$\int_0^{\frac{c_2}{c_1}} \operatorname{sgn} c_2\,ds - \int_{\frac{c_2}{c_1}}^{t_1} \operatorname{sgn} c_2\,ds = 0.$$

Stąd wynika, że $c_2/c_1 = \frac{1}{2}t_1$. Sterowanie optymalne będzie

$$u(t) = M \quad \text{dla} \quad 0 \leqslant t < \tfrac{1}{2}t_1,$$
$$u(t) = -M \quad \text{dla} \quad \tfrac{1}{2}t_1 < t \leqslant t_1.$$

Podstawiając znalezione $u(t)$ we wzór na $x_1(t)$ otrzymujemy

$$\int_0^{\frac{t_1}{2}} (t_1 - s)M\,ds - \int_{\frac{t_1}{2}}^{t_1} (t_1 - s)M\,ds = a,$$

a stąd $t_1 = 2\sqrt{a/M}$.

6.3. Podany przykład był bardzo łatwy i tym samym maskował wielkie trudności spotykane przy stosowaniu zasady maksimum do rozwiązywania konkretnych problemów. Występują one zwłaszcza, gdy równanie (6.1.1) nie jest liniowe. Musimy nie tylko znaleźć ogólne rozwiązanie równania (6.1.3), ale nadto wyznaczyć stałe całkowania tak, aby sterowanie doprowadziło wektor fazowy do pozycji $\boldsymbol{x}_1$. W przypadku nieliniowego równania (6.1.1), znalezienie analitycznych wzorów dla $\boldsymbol{x}(t)$ i $\psi(t)$ oraz dla optymalnego sterowania $\boldsymbol{u}(t)$ jest zazwyczaj niemożliwe. Przybliżone rozwiązanie za pomocą komputerów wymaga przy dużych wartościach wymiaru n bardzo wielkiej pamięci komputera. Trzeba zwykle drogą prób dobierać wartości początkowe dla wektora $\psi(t)$, które by dały rozwiązanie problemu odpowiadające zadanym warunkom.

6.4. Rozważmy szczególny przypadek, gdy problem sterowania jest liniowy, tzn gdy równanie (6.1.1) ma postać

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + B\boldsymbol{u}, \tag{6.4.1}$$

gdzie A jest macierzą $n \times n$, zaś B jest macierzą $n \times p$, przy czym elementy tych macierzy mogą być stałe, bądź też funkcjami czasu.

W tym przypadku zasada maksimum stanowi warunek nie tylko konieczny, lecz i dostateczny optymalności rozwiązania, byleby w ogóle istniało sterowanie przeprowadzające wektor fazowy z położenia $\boldsymbol{x}_0$ w położenie $\boldsymbol{x}_1$.

Jeśli macierz A jest stabilna, to znaczy, że wszystkie wartości własne mają ujemne części rzeczywiste, to wtedy dla dowolnego $\boldsymbol{x}_0$ istnieje ste-

rowanie optymalne przeprowadzające wektor fazowy z położenia $\boldsymbol{x}_0$ w położenie $\boldsymbol{x}_1 = 0$, pod warunkiem, że dziedzina sterowania U zawiera punkt $\boldsymbol{u} = 0$ jako punkt wewnętrzny.

6.5. Problem sterowania czasowo optymalnego można uogólnić, zakładając, że położenie początkowe $\boldsymbol{x}_0$ i końcowe $\boldsymbol{x}_1$ nie są ustalone, lecz że należą do hiperpowierzchni M_0 i M_1, określonych następująco:

$\boldsymbol{x} \in M_0$, gdy spełnione są równania $g_k(\boldsymbol{x}) = 0$, $k = 1,\ldots,\alpha$, $\alpha < n$;

$\boldsymbol{x} \in M_1$, gdy spełnione są równania $h_k(\boldsymbol{x}) = 0$, $k = 1,\ldots,\beta$, $\beta < n$.

Zakładamy, że każda z funkcji g_k, h_k ma w każdym punkcie gradient różny od 0.

Wówczas sterowanie optymalne musi czynić zadość nie tylko zasadzie maksimum, lecz musi nadto spełniać tzw. warunki transwersalności:

$$\psi^T \cdot \delta\boldsymbol{x}_0 = 0, \qquad \psi^T \cdot \delta\boldsymbol{x}_1 = 0$$

dla wszelkich wektorów $\delta\boldsymbol{x}_0$ i $\delta\boldsymbol{x}_1$, spełniających warunki:

$$\delta\boldsymbol{x}_0^T \cdot \operatorname{grad} g_k(\boldsymbol{x}_0) = 0, \qquad k = 1,\ldots,\alpha,$$
$$\delta\boldsymbol{x}_1^T \cdot \operatorname{grad} h_k(\boldsymbol{x}_1) = 0, \qquad k = 1,\ldots,\beta.$$

Warunki transwersalności można więc wypowiedzieć tak: wektory $\psi(t_0)$ i $\psi(t_1)$ mają być odpowiednio ortogonalne do hiperpowierzchni M_0 i M_1.

6.6. Zasadę maksimum można sformułować ogólniej, gdy zamiast zagadnienia czasowo optymalnego rozważamy problem osiągnięcia ekstremum funkcjonału $\int_{t_0}^{t_1} f^0(\boldsymbol{x}, \boldsymbol{u})\,dt$. Załóżmy, że $\boldsymbol{u}(t)$ jest sterowaniem dopuszczalnym, przeprowadzającym wektor fazowy z punktu $\boldsymbol{x}_0$ do punktu $\boldsymbol{x}_1$, zaś $\boldsymbol{x}(t)$ jest torem odpowiadającym temu sterowaniu. Zmieniamy definicję funkcji H, przyjmując obecnie, że

$$H(\psi, \boldsymbol{x}, \boldsymbol{u}) = \sum_{i=0}^{n} \psi^i f^i(\boldsymbol{x}, \boldsymbol{u}), \tag{6.6.1}$$

gdzie ψ^0 jest pewną stałą, zaś $\psi^k(t)$, $k = 1,\ldots,n$, spełniają równania różniczkowe

$$\dot{\psi}^i = -\frac{\partial H}{\partial x^i}, \qquad i = 1,\ldots,n. \tag{6.6.2}$$

Wektor ψ jest obecnie wektorem $(n+1)$-wymiarowym. Warunkiem koniecznym na to, by $\boldsymbol{u}(t)$ było sterowaniem optymalnym jest, aby istniała stała $\psi^0 \leqslant 0$ i takie rozwiązanie układu równań (6.6.2), aby w każdej chwili τ, w której sterowanie $\boldsymbol{u}(t)$ jest ciągłe, zachodził związek

$$H\big(\psi(\tau), \boldsymbol{x}(\tau), \boldsymbol{u}(\tau)\big) = \max_{v \in U} H\big(\psi(\tau), \boldsymbol{x}(\tau), \boldsymbol{v}\big) = 0 \tag{6.6.3}$$

dla $t_0 \leqslant \tau \leqslant t_1$.

6.7. Zasada maksimum może być stosowana do procesów opisanych nieautonomicznych równaniem różniczkowym

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \quad \boldsymbol{x}(t_1) = \boldsymbol{x}_1. \tag{6.7.1}$$

Niech wskaźnikiem jakości będzie $\int_{t_0}^{t_1} f^0\big(t, \boldsymbol{x}(t), \boldsymbol{u}(t)\big)\, dt$.

Zagadnienie możemy sprowadzić do rozwiązania równania różniczkowego autonomicznego przez podwyższenie wymiaru przestrzeni o 1. Wprowadzamy nową współrzędną x^{n+1}, spełniającą równania

$$\dot{x}^{n+1} = 1, \quad x^{n+1}(0) = 0. \tag{6.7.2}$$

Jest oczywiście $x^{n+1} = t$.

Oznaczmy wektor $(n+1)$-wymiarowy o współrzędnych $(x^1, x^2, \ldots, x^{n+1})$ przez $\tilde{\boldsymbol{x}}$, a wektor o współrzędnych $(f^1, f^2, \ldots, f^n, 1)$ przez $\tilde{\boldsymbol{f}}$. Wówczas równanie (6.7.1) i (6.7.2) można zapisać w postaci:

$$\dot{\tilde{\boldsymbol{x}}} = \tilde{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{u}). \tag{6.7.3}$$

W ten sposób otrzymaliśmy równanie autonomiczne z warunkami $\tilde{\boldsymbol{x}}(t_0) = (\boldsymbol{x}_0, t_0)$, $\tilde{\boldsymbol{x}}(t_1) = (\boldsymbol{x}_1, t_1)$.

1. Programowanie dynamiczne

7.1. Rozważmy jeszcze raz zagadnienie sterowania optymalnego procesu opisanego równaniami

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \quad \boldsymbol{u} \in U. \tag{7.1.1}$$

Szukamy minimum funkcjonału

$$F(\boldsymbol{x}, \boldsymbol{u}) = \int_0^T G[\boldsymbol{x}(t), \boldsymbol{u}(t)]\, dt, \tag{7.1.2}$$

gdzie T jest stałe.

Punktem wyjścia dla metody programowania dynamicznego jest tzw. *zasada optymalności.*

Niech $\boldsymbol{u}^*(t)$ oznacza sterowanie optymalne, a $\boldsymbol{x}^*(t)$ tor optymalny. Obieramy t_1, $0 < t_1 < T$. Tor został podzielony na dwie części, odpowiadające przedziałom czasowym $[0, t_1]$ i $[t_1, T]$, przy czym $\boldsymbol{x}^*(t_1) = \boldsymbol{x}_1$. Zasada optymalności mówi:

Drugi odcinek toru jest torem optymalnym.

Jest to oczywiste, gdyż gdyby można było znaleźć inne sterowanie $v(t)$ takie, że przy odpowiadającym mu torze $\tilde{x}(t)$

$$\int_{t_1}^{T} G(\tilde{x}(t), v(t))\,dt < \int_{t_1}^{T} G(x^*(t), u^*(t))\,dt,$$

przy czym $\tilde{x}(t_1) = x_1$, to wtedy sterowanie $u^*(t)$ nie byłoby optymalne w przedziale $[0, T]$, gdyż nie minimalizowałoby całki (7.1.2): mniejszą wartość tej całki otrzymalibyśmy przy sterowaniu $w(t)$, określonym przez równości

$$w(t) = u^*(t) \quad \text{dla} \quad 0 \leqslant t \leqslant t_1,$$

$$w(t) = v(t) \quad \text{dla} \quad t_1 < t \leqslant T.$$

Możemy ująć zasadę optymalności również w ten sposób:

Sterowanie optymalne dla $t_1 \leqslant t \leqslant T$ *zależy od* $x(t_1)$, *natomiast nie zależy ani od* $x(t)$, *ani od* $u(t)$ *dla* $0 \leqslant t < t_1$.

Sformułowana wyżej zasada optymalności nie odnosi się do pierwszego odcinka toru. Jest tak z następującego powodu: Jeżeli sterowanie optymalne $u^*(t)$ doprowadziło w chwili t_1 do punktu x_1, a więc do punktu czasoprzestrzeni (t_1, x_1), to może się zdarzyć, że pewne sterowanie $\hat{u}(t)$, które wyznacza tor $\hat{x}(t)$ doprowadzi w chwili t_1 do punktu $\hat{x}(t_1) = x_1$ i że przy tym

$$\int_{t_0}^{t_1} G(\hat{x}(t), \hat{u}(t))\,dt < \int_{t_0}^{t_1} G(x^*(t), u^*(t))\,dt,$$

ale przy tym optymalne prowadzenie wektora fazowego z punktu $\hat{x}(t_1)$ w przedziale czasowym $[t_1, T]$ daje tak wielką wartość całki $\int_{t_1}^{T} G(x(t), u(t))\,dt$, że zastąpiwszy w przedziale $[0, t_1]$ sterowanie $\hat{u}^*(t)$ przez $u(t)$ nie możemy osiągnąć minimum dla całki (7.1.2). Widzimy więc, że sterowanie $u^*(t)$, które jest optymalne w $[0, T]$, nie musi być optymalne w $[0, t_1]$. Podany przykład ilustruje ten fakt dla problemu czasowo optymalnego.

7.2. Przykład. Niech przestrzeń fazowa będzie dwuwymiarowa, $x \in R^2$. Proces sterowania jest opisany równaniem

$$\dot{x} = u(t). \tag{7.2.1}$$

Niech sterowanie dopuszczalne u spełnia warunki określające dziedzinę sterowania U:

$$1 \leqslant u^1(t) \leqslant 2,$$

$$u^2(t) = 0 \quad \text{gdy} \quad x^1 < 1,$$

$$u^2(t) = (x^1 - 1)(x^1 - t) \quad \text{gdy} \quad x^1 \geqslant 1,$$

gdzie przez x^1, x^2, u^1, u^2 oznaczyliśmy składowe wektorów fazowego i sterującego. Chcemy przeprowadzić wektor fazowy z punktu $(0,0)$ do punktu $(2, 0)$ w najkrótszym czasie. Łatwo sprawdzić, że jedynym sterowaniem, za pomocą którego punkt $(2,0)$ można osiągnąć, jest sterowanie $u^1(t) \equiv 1$, co pociąga za sobą $u^2(t) \equiv 0$, gdyż tylko przy tym sterowaniu tor pozostaje w płaszczyźnie $x^2 = 0$. Jeśli ten tor, dany równaniami $x^1 = t$, $x^2 = 0$, osiągający punkt $(2, 0)$ po czasie $T = 2$, podzielimy na dwie części odpowiadające przedziałom czasowym $0 \leqslant t \leqslant 1$ i $1 \leqslant t \leqslant 2$, to koniec pierwszego odcinka jest w punkcie $(1, 0)$ przestrzeni fazowej. Za pomocą sterowania $u^1(t) = 2$ dla $0 \leqslant t \leqslant \frac{1}{2}$ ten punkt można osiągnąć w krótszym czasie $t_1 = \frac{1}{2}$, lecz z punktu czasoprzestrzeni $(\frac{1}{2}, 1, 0)$ za pomocą żadnego sterowania nie można dotrzeć do punktu $(2,0)$ przestrzeni fazowej (ściśle mówiąc do punktu $(T, 2, 0)$ czasoprzestrzeni, jakiekolwiek obierzemy T), gdyż wszystkie sterowania dla dróg wychodzących z $(\frac{1}{2}, 1, 0)$ spełniają warunek $u^2(t) > 0$ dla $t > \frac{1}{2}$, tak, że trajektorie spełniają warunek $x^2(t) > 0$ dla $t > \frac{1}{2}$.

7.3. Metoda programowania dynamicznego jest przystosowana do obliczania na maszynach. Wobec tego zastąpimy równanie różniczkowe (7.1.1) przez równanie różnicowe

$$\frac{\boldsymbol{x}(t_{k+1}) - \boldsymbol{x}(t_k)}{t_{k+1} - t_k} = \boldsymbol{f}\big(\boldsymbol{x}(t_k), \boldsymbol{u}(t_k)\big), \tag{7.3.1}$$

gdzie $t_k = k \cdot h$, $h > 0$, $k = 0, 1, \ldots, (N-1)$.

Oznaczając $\boldsymbol{x}(t_k) = \boldsymbol{y}(k)$, $\boldsymbol{u}(t_k) = \boldsymbol{v}(k)$, $h\boldsymbol{f}\big(\boldsymbol{x}(t_k), \boldsymbol{u}(t_k)\big) = \boldsymbol{g}\big(\boldsymbol{y}(k), \boldsymbol{v}(k)\big)$ możemy przepisać (7.3.1) w postaci

$$\boldsymbol{y}(k+1) - \boldsymbol{y}(k) = \boldsymbol{g}\big(\boldsymbol{y}(k), \boldsymbol{v}(k)\big). \tag{7.3.2}$$

Analogicznie całkę (7.1.2) zastępujemy przez sumę

$$F_1(\boldsymbol{y}, \boldsymbol{v}) = h \sum_{k=0}^{N-1} G\big(\boldsymbol{y}(k), \boldsymbol{v}(k)\big) = \sum_{k=0}^{N-1} H\big(\boldsymbol{y}(k), \boldsymbol{v}(k)\big) \tag{7.3.3}$$

gdzie $H = h \cdot G$.

Widzimy, że problem sprowadza się do wyznaczenia minimum funkcjonału (7.3.3) zależnego od zmiennych $\boldsymbol{v}(0)$, $\boldsymbol{v}(1), \ldots, \boldsymbol{v}(N-1)$.

Metoda programowania dynamicznego pozwala sprowadzić zagadnienie wyznaczenia minimum funkcji N zmiennych do kolejnego wyznaczania minimów funkcyj jednej zmiennej. Rozumujemy tak: chcemy wyznaczyć sterowanie optymalne na ostatnim etapie, to znaczy $\boldsymbol{v}(N-1)$. Dla uproszczenia zapisu przy przedstawieniu metody programowania dynamicznego przyjmiemy najpierw dla N pewną szczególną wartość, np. $N = 4$. Łatwo będzie zauważyć, że metoda nie zależy od ilości etapów N.

Będzie więc

$$F_1 = \sum_{k=0}^{3} H\big(\boldsymbol{y}(k), \boldsymbol{v}(k)\big).$$

Szukamy

$$(7.3.4)\quad \min_{\substack{\boldsymbol{v}(i)\in U\\ i=0,1,2,3}} \sum_{k=0}^{3} H\big(\boldsymbol{y}(k), \boldsymbol{v}(k)\big) =$$

$$= \min_{\substack{\boldsymbol{v}(i)\in U\\ i=0,1,2}} \Big[\sum_{k=0}^{2} H\big(\boldsymbol{y}(k), \boldsymbol{v}(k)\big) + \min_{\boldsymbol{v}(3)\in U} H\big(\boldsymbol{y}(3), \boldsymbol{v}(3)\big)\Big] =$$

$$= \min_{\boldsymbol{v}(0)\in U} \Big\{H\big(\boldsymbol{y}(0), \boldsymbol{v}(0)\big) + \min_{\boldsymbol{v}(1)\in U} \Big[H\big(\boldsymbol{y}(1), \boldsymbol{v}(1)\big) +$$

$$+ \min_{\boldsymbol{v}(2)\in U} \big\{H\big(\boldsymbol{y}(2), \boldsymbol{v}(2)\big) + \min_{\boldsymbol{v}(3)\in U} H\big(\boldsymbol{y}(3), \boldsymbol{v}(3)\big)\big\}\Big]\Big\}.$$

Zaczynamy od wyznaczenia $\min_{\boldsymbol{v}(3)\in U} H\big(\boldsymbol{y}(3), \boldsymbol{v}(3)\big)$, a więc minimum funkcji jednej zmiennej $\boldsymbol{v}(3)$ (co prawda, ta zmienna jest wektorem, a więc reprezentuje wiele zmiennych rzeczywistych). Jeżeli rozwiążemy to zadanie, otrzymamy, że minimum jest osiągnięte dla $\boldsymbol{v}(3) = \boldsymbol{v}^*(3)$, które jest zależne od $\boldsymbol{y}(3)$, $\boldsymbol{v}^*(3) = \boldsymbol{\varphi}_3\big(\boldsymbol{y}(3)\big)$. Z kolei obliczamy

$$(7.3.5)\quad \min_{\boldsymbol{v}(2)\in U}\Big[H\big(\boldsymbol{y}(2), \boldsymbol{v}(2)\big) + H\Big(\boldsymbol{y}(3), \boldsymbol{\varphi}_3\big(\boldsymbol{y}(3)\big)\Big)\Big].$$

Uwzględniając (7.3.2), możemy (7.3.5) zapisać w postaci

$$(7.3.6)\quad \min_{\boldsymbol{v}(2)\in U}\Big\{H\big(\boldsymbol{y}(2), \boldsymbol{v}(2)\big) + H\Big[\boldsymbol{y}(2) +$$

$$+ \boldsymbol{g}\big(\boldsymbol{y}(2), \boldsymbol{v}(2)\big), \boldsymbol{\varphi}_3\big(\boldsymbol{y}(2) + \boldsymbol{g}\big(\boldsymbol{y}(2), \boldsymbol{v}(2)\big)\big)\Big]\Big\}.$$

Mamy znów znaleźć minimum funkcji jednej zmiennej $\boldsymbol{v}(2)$. Jeśli to minimum jest osiągnięte dla $\boldsymbol{v}(2) = \boldsymbol{v}^*(2) = \boldsymbol{\varphi}_2\big(\boldsymbol{y}(2)\big)$, to możemy wyrażenie (7.3.6) zapisać jako pewną funkcję zmiennej $\boldsymbol{y}(2)$, którą oznaczymy przez $B_2\big(\boldsymbol{y}(2)\big)$.

Wtedy z kolei trzeba obliczyć

$$(7.3.7)\quad \min_{\boldsymbol{v}(1)\in U}\big[H\big(\boldsymbol{y}(1), \boldsymbol{v}(1)\big) + B_2\big(\boldsymbol{y}(2)\big)\big] =$$

$$= \min_{\boldsymbol{v}(1)\in U}\Big[H\big(\boldsymbol{y}(1), \boldsymbol{v}(1)\big) + B_2\Big(\boldsymbol{y}(1) + \boldsymbol{g}\big(\boldsymbol{y}(1), \boldsymbol{v}(1)\big)\Big)\Big].$$

Analogicznie jak poprzednio wyrażenie (7.3.7) po rozwiązaniu zadania wyznaczenia minimum możemy zapisać w postaci $B_1\big(\boldsymbol{y}(1)\big)$ i ostatecznie mamy do wyznaczenia

$$(7.3.8)\quad \min_{\boldsymbol{v}(0)\in U}\Big[H\big(\boldsymbol{y}(0), \boldsymbol{v}(0)\big) + B_1\Big(\boldsymbol{y}(0) + \boldsymbol{g}\big(\boldsymbol{y}(0), \boldsymbol{v}(0)\big)\Big)\Big].$$

W ten sposób został zrealizowany plan zastąpienia poszukiwania minimum funkcji wielu zmiennych wektorialnych przez kolejne wyznaczanie minimów funkcji jednej zmiennej wektorialnej. Zadanie jednak pozostaje bardzo trudne, gdyż idzie o znalezienie zależności $\boldsymbol{v}^*(k)$ od wektora $\boldsymbol{y}(k)$. Tylko wyjątkowo udaje się wyrazić tę zależność wzorem analitycznym. Na ogół zadanie może być rozwiązane tylko za pomocą komputera i wymaga umieszczenia w jego pamięci stabelaryzowanej funkcji $\boldsymbol{v}^*(k) = \boldsymbol{\varphi}_k(\boldsymbol{y}(k))$. Również wartości $H(\boldsymbol{y}(k), \boldsymbol{v}^*(k))$ muszą być umieszczone w pamięci. Jeżeli ilość wymiarów przestrzeni fazowej lub przestrzeni wektorów sterowania jest duża, to zadanie jest praktycznie nierozwiązalne, gdyż z jednej strony pojemność pamięci komputera okazuje się za mała, z drugiej zaś czas obliczania wypada niezmiernie wielki. Wobec tego stosowalność metody programowania dynamicznego Bellmana jest uzależniona od wielkości wymiarów przestrzeni i od pojemności pamięci komputera oraz jego sprawności.

7.4. Podkreślmy charakterystyczną cechę metody Bellmana: rachunki zaczynają się od obliczania sterowania optymalnego na ostatnim etapie, a kończą się obliczeniem sterowania w momencie początkowym. Obecnie ujmiemy tę metodę za pomocą wzoru rekurencyjnego. Wzór (7.3.8) (brany dla dowolnego N) służy do obliczania minimum funkcjonału F_1 ze wzoru (7.3.3) odpowiadającego procesowi o N etapach. Oznaczmy

$$\min_{\substack{\boldsymbol{v}(k)\in U \\ k=0,\ldots,(N-1)}} F_1 = S_N(\boldsymbol{y}(0)). \tag{7.4.1}$$

Przy tym oznaczeniu otrzymujemy wzór rekurencyjny

$$S_N(\boldsymbol{y}(0)) = \min_{\boldsymbol{v}(0)\in U} \{H(\boldsymbol{y}(0), \boldsymbol{v}(0)) + S_{N-1}(\boldsymbol{y}(1))\}. \tag{7.4.2}$$

Zwróćmy uwagę na to, że drugi składnik w nawiasie we wzorze (7.4.2) jest, zgodnie z zasadą optymalności, równy minimum funkcjonału F_1 w $(N-1)$ etapowym procesie z warunkiem początkowym $\boldsymbol{y} = \boldsymbol{y}(1)$.

Opierając się na (7.3.2) otrzymujemy

$$S_N(\boldsymbol{y}(0)) = \min_{\boldsymbol{v}(0)\in U} \left\{H(\boldsymbol{y}(0), \boldsymbol{v}(0)) + S_{N-1}\left(\boldsymbol{y}(0) + \boldsymbol{g}(\boldsymbol{y}(0), \boldsymbol{v}(0))\right)\right\}. \tag{7.4.3}$$

W ten sposób otrzymaliśmy algorytm Bellmana dla procesów dyskretnych.

7.5. Metoda Bellmana jest zasadniczo pomyślana dla procesów dyskretnych, lecz bywa stosowana również do procesów ciągłych. Wróćmy do równań (7.1.1) i (7.1.2), przy czym zapiszemy je w ogólniejszej postaci:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \tag{7.5.1}$$

$$F(\dot{\boldsymbol{x}}, \boldsymbol{u}) = \int_{t_0}^{T} \boldsymbol{g}[\boldsymbol{x}(t), \boldsymbol{u}(t), t]dt; \tag{7.5.2}$$

T jest ustalone. Niech $t_1 = t_0+h < T$, $h > 0$. Jest

$$F(\boldsymbol{x}, \boldsymbol{u}) = \int_{t_0}^{t_0+h} G[\boldsymbol{x}(t), \boldsymbol{u}(t), t]dt + \int_{t_0+h}^{T} G[\boldsymbol{x}(t), \boldsymbol{u}(t), t]dt. \tag{7.5.3}$$

Oznaczmy minimum funkcjonału F, to jest jego wartość przy przejściu z $(t_0, \boldsymbol{x}_0)$ do $(T, \boldsymbol{x}^*(T))$ po torze optymalnym $\boldsymbol{x}^*(t)$ przez $S(t_0, \boldsymbol{x}_0)$. Zakładamy, że S ma pochodne cząstkowe $\partial S/\partial x^k$, $\partial S/\partial t$. Jest wówczas

$$S(t_0, \boldsymbol{x}_0) = \min_{\boldsymbol{u}(t) \in U} \left\{ \int_{t_0}^{t_0+h} G[\boldsymbol{x}(t), \boldsymbol{u}(t), t]dt + \int_{t_0+h}^{T} G[\boldsymbol{x}(t), \boldsymbol{u}(t), t]dt \right\}. \tag{7.5.4}$$

Zgodnie z zasadą optymalności

$$S(t_0, \boldsymbol{x}_0) = \min_{\boldsymbol{u}(t) \in U} \left\{ \int_{t_0}^{t_0+h} G[\boldsymbol{x}(t), \boldsymbol{u}(t), t]dt + S(t_0+h, \boldsymbol{x}(t_0+h)) \right\} = \tag{7.5.5}$$

$$= \min_{\boldsymbol{u}(t) \in U} \left\{ hG[\boldsymbol{x}(t_0), \boldsymbol{u}(t_0), t_0] + h \sum_{k=1}^{n} \frac{\partial S}{\partial x^k} \dot{x}^k \Big|_{(t_0, \boldsymbol{x}_0)} + h \frac{\partial S}{\partial t}\Big|_{(t_0, \boldsymbol{x}_0)} + \right.$$

$$\left. + o(h) + S(t_0)\boldsymbol{x}(t_0) \right\},$$

gdzie symbol Landaua $o(h)$ oznacza funkcję spełniającą warunek $o(h)/h \to 0$ dla $h \to 0$. Odejmując w (7.5.5) po obu stronach $S(t_0, \boldsymbol{x}_0)$ i dzieląc przez h otrzymujemy po przejściu do granicy $h \to 0$

$$\min_{\boldsymbol{v} \in U} \left\{ G[\boldsymbol{x}(t_0), \boldsymbol{u}(t_0)] + \sum_{k=1}^{n} \frac{\partial S}{\partial x^k} \dot{x}^k \Big|_{(t_0, \boldsymbol{x}_0)} + \frac{\partial S}{\partial t}\Big|_{(t_0, \boldsymbol{x}_0)} \right\} = 0. \tag{7.5.6}$$

Uwzględniając (7.5.1) i fakt, że $\partial S/\partial t$ od $\boldsymbol{u}(t)$ nie zależy, a pozostałe składniki w nawiasie w (7.5.6) zależą od $\boldsymbol{u}(t)$ tylko w punkcie $t = t_0$, otrzymujemy

$$-\frac{\partial S}{\partial t}\Big|_{(t_0, \boldsymbol{x}_0)} = \min_{\boldsymbol{u}(t_0) \in U} \left\{ G[\boldsymbol{x}(t_0), \boldsymbol{u}(t_0), t_0] + \sum_{k=1}^{n} \frac{\partial S}{\partial x^k} f^k \Big|_{(t_0, \boldsymbol{x}_0)} \right\}. \tag{7.5.7}$$

Punkt $(t_0, \boldsymbol{x}_0)$ można było obrać dowolnie, więc otrzymujemy ostatecznie tzw. równanie Bellmana:

$$-\frac{\partial S}{\partial t} = \min_{\boldsymbol{u}(t) \in U} \left\{ G[\boldsymbol{x}(t), \boldsymbol{u}(t), t] + \mathbf{grad}_{\boldsymbol{x}}^T S \cdot \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t) \right\}. \tag{7.5.8}$$

Wektor $\mathbf{grad}_{\boldsymbol{x}} S$ nie zależy od $\boldsymbol{u}(t)$.

Jeżeli na $\boldsymbol{u}(t)$ nie są nałożone żadne ograniczenia, to w celu wyznacze-

nia optymalnego sterowania przyrównuje się do 0 pochodną względem wektora u z wyrażenia w nawiasie we wzorze (7.5.8):

$$\text{grad}_u\,[G(x, u, t)+\text{grad}_x^T\, Sf(x, u, t)] = 0. \tag{7.5.9}$$

7.6. Metoda rozwiązywania problemu sterowania optymalnego za pomocą równania Bellmana była atakowana z tego powodu, że założenie o istnieniu pochodnych cząstkowych $\partial S/\partial x^k$ i $\partial S/\partial t$ nie jest spełnione, nieraz nawet w przypadku bardzo prostych, a ważnych problemów. Dowiedziono jednak (por. [5]), że do poprawności metody Bellmana wystarczy, aby pochodne istniały poza pewnym zbiorem wyjątkowym, który ma mniejszy wymiar niż rozpatrywana czasoprzestrzeń. Dzięki temu metoda Bellmana może być stosowana w bardzo wielu przypadkach.

Rzecz jasna, że zarzuty krytyczne nie odnosiły się do metody programowania dynamicznego, stosowanego do procesów dyskretnych.

7.7. Przykład. Proces sterowania jest opisany układem równań

$$\dot{x}^1 = x^2, \quad \dot{x}^2 = u(t), \quad |u(t)| \leqslant 1. \tag{7.7.1}$$

Zadanie: przejść w minimalnym czasie z punktu (x_0^1, x_0^2) do punktu $(0, 0)$. W naszym przykładzie $G(x, u, t) \equiv 1$. Równania (7.7.1) zależą tylko od x i $u(t)$; będzie więc $S = T(x)$, $S'_t = 0$. Równanie (7.5.8) przyjmie postać

$$\min_{-1\leqslant u\leqslant 1}\left\{1+\frac{\partial T}{\partial x^1}\,x^2+\frac{\partial T}{\partial x^2}\,u\right\} = 0.$$

Stąd sterowanie optymalne $u^* = -\text{sgn}\dfrac{\partial T}{\partial x^2}$, a więc

$$1+\frac{\partial T}{\partial x^1}\,x_2-\frac{\partial T}{\partial x^2}\,\text{sgn}\frac{\partial T}{\partial x^2} = -\left|\frac{\partial T}{\partial x^2}\right|+1+\frac{\partial T}{\partial x^1}\,x_2 = 0. \tag{7.7.2}$$

Trzeba znaleźć rozwiązania dla dwu równań:

$$1+\frac{\partial T}{\partial x^1}\,x^2-\frac{\partial T}{\partial x^2} = 0 \quad \text{gdy} \quad \text{sgn}\frac{\partial T}{\partial x^2} = +1, \tag{7.7.3.a}$$

$$1+\frac{\partial T}{\partial x^1}\,x^2+\frac{\partial T}{\partial x^2} = 0 \quad \text{gdy} \quad \text{sgn}\frac{\partial T}{\partial x^1} = -1. \tag{7.7.3.b}$$

Dla równania (7.7.3.a) istnieją dwie całki pierwsze $T-x^2$ oraz $2x^1+(x^2)^2$, przeto

$$T = x^2+\varphi\left(2x^1+(x^2)^2\right), \tag{7.7.4}$$

gdzie $\varphi(v)$ jest pewną funkcją różniczkowalną. Aby ją wyznaczyć, trzeba znaleźć wartość T wzdłuż jakiejkolwiek krzywej nie będącej charaktery-

styką równania (7.7.3.a). Najdogodniej wziąć w tym celu tor optymalny odpowiadający sterowaniu $u^*(t) \equiv 1$ (bo dla $u^*(t) = -1$ otrzymalibyśmy właśnie charakterystykę (7.7.3.a), docierający do $(0, 0)$ w zadanym czasie, np. $T = 2$. Całkując równania (7.7.1) przy $u(t) = 1$ i warunku $x^1(2) = x^2(2) = 0$ znajdziemy

$$x^1 = \tfrac{1}{2}(t-2)^2, \quad x^2 = t-2. \tag{7.7.5}$$

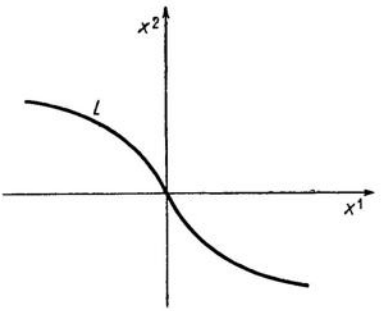
Z optymalności toru (7.7.5) wynika, że $T = 2-t$. Uwzględniając te wyniki otrzymamy z (7.7.4):

$$2-t \equiv t-2+\varphi[2(t-2)^2], \quad \varphi[2(t-2)^2] = -2(t-2).$$

Kładąc $2(t-2)^2 = v$, $t-2 = \mp\sqrt{v/2}$, znajdziemy $\varphi(v) = \mp 2\sqrt{v/2} = \mp\sqrt{2v}$, a stąd

$$T = x^2 \mp \sqrt{4x^1+2(x^2)^2}. \tag{7.7.6}$$

Szczegółowa dyskusja, którą pomijamy, prowadzi do wniosku, że należy przyjąć $\operatorname{sgn}\dfrac{\partial T_{xx}}{\partial x^2} = 1$ w obszarze bieżącym nad krzywą L (rys. 6),



Rys. 6

która dla $x^1 < 0$ ma równanie $2x^1+(x^2)^2 = 0$, $x^2 > 0$, zaś dla $x^1 > 0$ ma równanie $-2x^1+(x^2)^2 = 0$, $x^2 < 0$ oraz że w równaniu (7.7.6) należy przyjąć znak dolny tak, że ostatecznie

$$T = x^2 + \sqrt{4x^1+2(x^2)^2}. \tag{7.7.7}$$

Analogiczne rozważanie dla równania (7.7.3.b) prowadzą do wyniku, że poniżej krzywej L jest

$$T = -x^2 + \sqrt{2(x^2)^2-4x^1}, \qquad u^*(t) = 1 \tag{7.7.8}$$

Ogólny wzór na optymalne sterowania jest więc

$$u^*(t) = \operatorname{sgn}(x^1+\tfrac{1}{2}x^2\cdot|x^2|). \tag{7.7.9}$$

8. Zasada maksimum dla procesu dyskretnego

8.1. Procesy dyskretne liniowe. Rozważmy problem optymalnego sterowania dla procesu opisanego równaniem (por. [12])

$$x(i+1)-x(i) = A(i)x(i)+g(i,u(i)), \quad i = 0, 1,\ldots,k. \tag{8.1.1}$$

Tutaj i jest numerem momentu czasu, $x(i) \in R^n$, $u(i) \in U \subset R^r$, $A(i)$ jest macierzą $n \times n$, $g \in R^n$, wektor-funkcja g jest określona dla $i = 0,1, \ldots,(k-1)$ oraz $u(i) \in U$.

Warunki: początkowy $x(0) = x_0$, końcowy $x(k) = (x_k^1,\ldots, x_k^{n-1}, x^n)$, gdzie $n-1$ składowych wektora $x(k)$ jest ustalone, natomiast n-ta składowa jest dowolna. Wobec tego tor optymalny wychodzi z punktu x_0 i ma dotrzeć do pewnej prostej równoległej do osi x^n, która będzie prostą celową. Tor oczywiście nie jest krzywą ciągłą, lecz ciągiem punktów przestrzeni R^n. Zadanie polega na tym, aby wyznaczyć takie sterowanie $u(i)$, aby n-ta współrzędna punktu $x(k)$ była największa. Ta współrzędna jest więc wskaźnikiem jakości. Ilość rozważanych momentów czasowych k jest ustalona.

8.2. Założenie zasadnicze. Oznaczmy przez $\|A(i)\|$ normę macierzy

$$\|A(i)\| = \max_{|x|\leqslant 1} |A(i)x|, \tag{8.2.1}$$

gdzie przez $|x|$ rozumiemy euklidesową długość wektora x. Oznaczamy przez $N(i)$ zbiór wszystkich wektorów $g(i,u)$, gdzie u jest dowolnym wektorem $u \in U$.

Zakładamy, że $N(i)$ jest zbiorem domkniętym, ograniczonym i wypukłym dla $i = 0,1,\ldots,(k-1)$.

Zbiorem osiągalnym, oznaczonym przez W, nazywamy zbiór wszystkich $x(i)$, $i = 0,1,\ldots,k$, które można osiągnąć z punktu x_0 przyjmując dowolne sterowanie dopuszczalne, to znaczy takie, że $u(i) \in U$. Przez W_i oznaczamy zbiór wszystkich punktów zbioru W dla ustalonego i. W jest więc połączeniem wszystkich W_i dla $i = 0,1,\ldots,k$.

8.3. Zasada optymalnego toru. Przy przyjętych założeniach można dowieść indukcyjnie, że jeżeli $u(i)$ jest sterowaniem optymalnym, to $x(i)$ należy do brzegu zbioru W_i. Inaczej mówiąc, gdy sterowanie jest optymalne, to odpowiadający mu tor $x(i)$ przebiega stale po brzegu zbioru osiągalnego W.

Oznaczmy przez $w_i(x(i))$ zbiór wszystkich $x(i+1)$ osiągalnych przy dowolnym wyborze $u(i)$ z danego punktu $x(i)$. Inaczej mówiąc, $w_i(x(i))$ jest zbiorem osiągalnym po jednym kroku z danego punktu $x(i)$. Z zasady optymalnego toru wynika twierdzenie: *jeżeli $u(i)$ jest sterowaniem optymalnym, to $x(i+1)-x(i)$ należy do brzegu zbioru $w_i(x(i))$.*

8.4. Oznaczmy przez $f(x, u, i)$ prawą stronę równania (8.1.1). Niech p będzie dowolnym wektorem w przestrzeni R^n. Oznaczamy przez $H(i, x, u, p)$ iloczyn skalarny

$$H(i, x, u, p) = p^T f(x, u, i). \tag{8.4.1}$$

Można dowieść następujące twierdzenia (por. [12]):

1. *Jeżeli $u(i)$ jest sterowaniem optymalnym i jeżeli brzeg zbioru W_i ma w punkcie $x(i)$ hiperpłaszczyznę styczną z niezerową normalną zewnętrzną $p(i)$, a brzeg zbioru W_{i+1} ma w punkcie $x(i+1)$ hiperpłaszczyznę styczną z niezerową normalną zewnętrzną $p(i+1)$, to wtedy dla wszelkich $v \in U$ jest*

$$H\big(i, x(i), u(i), p(i+1)\big) \geqslant H\big(i, x(i), v, p(i+1)\big).$$

2. *Przy tych samych założeniach co w twierdzeniu* 1 *można tak dobrać długości wektorów $p(i)$, $p(i+1)$, aby zachodził związek*

$$p(i) - p(i+1) = A^T(i)p(i+1).$$

Twierdzenia te wyrażają zasadę maksimum dla procesu opisanego równaniem (8.1.1).

Na zakończenie zauważmy, że jeżeli istnieje w ogóle sterowanie prowadzące z punktu x_0 do prostej celowej, to musi istnieć sterowanie optymalne.

8.5. Zajmijmy się bardziej skomplikowanym przypadkiem (por. [13]), w którym dziedzina sterowania zmienia się z czasem, wektor fazowy podlega ograniczeniom zależnym od czasu, a zamiast żądania, aby tor dotarł do prostej celowej przy największej n-tej współrzędnej, postawimy żądanie, by otrzymać minimum pewnego funkcjonału wypukłego ze względu na x i u. Zamiast $x(i)$, $u(i)$ będziemy pisali x_i, u_i. Niech $x_i \in X_i \subset R^n$, $i = 0, \ldots, k$; $u_i \in U_i \subset R^r$; $x_{i+1} - x_i = f_i(x_i, u_i)$. Zakładamy, że $f_i(x, u)$ są funkcjami liniowymi na $X_i \times U_i$. Niech funkcje $\sigma_i(x, u)$ o wartościach rzeczywistych będą określone odpowiednio na zbiorach $X_i \times U_i$ i wypukłe na nich.

Szukamy optymalnego sterowania $\{\hat{u}_i\}$ i odpowiadającego mu optymalnego toru $\{\hat{x}_i\}$, takich, aby funkcjonał $\sum_{i=0}^{k-1} \sigma_i(x_i, u)$ osiągnął minimum dla $x_i = \hat{x}_i$, $u_i = \hat{u}_i$. Mamy twierdzenie:

Warunkiem koniecznym i dostatecznym, aby $\{\hat{x}_i\}$ i $\{\hat{u}_i\}$ były optymalne, jest, aby istniał ciąg wektorów $\{p_i\}$, $p_i \in R^n$, $i = 0, \ldots, k$, $p_0 = 0$, taki, aby $p^T(\hat{x}_k - x_k) \geqslant 0$ dla wszystkich $x_k \in X_k$ i aby

$$H_i(\hat{x}_i, \hat{u}_i, p_{i+1}, p_i) \geqslant H_i(x_i, u_i, p_{i+1}, p_i)$$

dla wszystkich $x_i \in X_i$, $u_i \in U_i$, $i = 0, 1, \ldots, k-1$, *gdzie*

$$H_i(x, u, p, q) = p^T f_i(x, u) + (p-q)^T x - \sigma_i(x, u).$$

Praca [13] zawiera jeszcze dwa inne twierdzenia, dotyczące pochodnych problemów, które pomijamy, gdyż wymagałyby dłuższych objaśnień.

Prace cytowane

[1] R. Bellman, J. Glicksberg and O. Gross, *On the bang - bang control problem*, Quart. Appl. Math. 14 (1956).

[2] R. Bellman, *Dynamic programming*, Princeton 1957.

[3] — *Adaptacyjne procesy sterowania*, Warszawa 1965.

[4] В. Г. Болтянский, Р. В. Гамкрелидзе и Л. С. Понтрягин, *К теории оптимальных процессов*, Докл. Акад. Наук СССР 110 (1956).

[5] В. Г. Болтянский, *Математические методы оптимального управления*, Москва 1966.

[6] D. W. Bushaw, *Optimal discontinous forcing terms*, Contrib. to the theory of non-linear oscillations, Vol. 4, Princeton 1958.

[7] А. Г. Бутковский, *Принцип максимума для оптимальных систем с распределенными параметрами*, Автом. и Телемех. (1961), т. 22. № 10,

[8] А. И. Дубовицкий и А. А. Милютин, *Задачи на экстремум при наличии ограничений*, Журн. Вычислит. Матем. и Матем. Физ. (1965), № 3.

[9] С. В. Емелянов и Е. Н. Дубровский, *Теория адаптивных систем упрявления*, Итоги Науки и Техн., Москва 1967.

[10] А. А. Фелдбаум, *Оптимальные процесси в системах автоматического регулирования*, Автом. и Телемех. № 14 (1963).

[11] A. A. Feldbaum, *Podstawy teorii optymalnych układów sterowania*, Warszawa 1967.

[12] H. Halkin, *Optimal control for systems described by difference equations*, W: *Advances in control systems*, Vol. I, New York—London 1964.

[13] — B. W. Jordan, E. Polak and J. B. Rosen, *Theory of optimum discrete time systems*, Proc. III Congr. IFAC, Paper 28 B, London 1966.

[14] L. Hurwicz, *Programming in linear spaces. Studies on linear and non-linear programming*, Stanford 1958.

[15] R. Isaacs, *Differential games*, New York 1965.

[16] Н. Н. Красовский, *К теории оптимального управления*, Автом. и Телемех. т. 18, № 11 (1957).

[17] N. N. Krasowski i G. G. Pospiełow, *Podstawy automatyki i cybernetyki technicznej*, Warszawa 1965.

[18] E. Kreindler, *Contributions to the theory of time optimal control*, Journ. Franklin Instit. 275, No. 4 (1963), str. 314-344.

[19] H. Kuhn and A. W. Tucker, *Non-linear programming*, Proc. 2-nd Berkeley Symposium on Mathem. Statistics and Probability (1950), str. 481-492.

[20] R. Kulikowski, *On optimum control with constants*, Bull. Pol. Acad. Sci. Sér. Techn. 7 (1959), str. 285-294.

[21] — *Concerning the synthesis of the optimum non-linear control*, ibidem 7 (1959), str. 339-391.

[22] — *Synthesis of a class of optimum control systems*, ibidem 7 (1959), str. 663-671.

[23] — *Synthesis of optimum control systems with area-bounded control*, ibidem 8 (1960), str. 178-186.

[24] — *Concerning a class of optimum control systems*, ibidem 8 (1960), str. 595-600.
[25] — *Optimizations of aggregated dynamic systems*, Arch. Autom. i Telemech. 11 (1966), str. 227-255.
[26] J. P. La Salle, *Basic principle of the bang-bang servo*, Bull. Amer. Math. Soc. 60 (1954).
[27] J. Majerczyk-Gómułka i K. Makowski, *Wyznaczanie optymalnego sterowania procesami dynamicznymi metodą funkcjonałów Lagrange'a*, Prace IPPT PAN 9 (1967).
[28] K. Malanowski i S. Rolewicz, *Zastosowanie płaszczyzn podpierających do wyznaczania sterowania czasowo optymalnego*, Arch. Autom. i Telemech. 10 (1965), str. 149-170.
[29] A. Pliś, *Trajectories and quasi-trajectories of an orientor field*, Bull. Pol. Acad. Sci. Sér. Math. 11 (1963), str. 369-370.
[30] Л. С. Понтрягин, В. Г. Болтянский, Р. В. Гамкрелидзе, и Е. Ф. Мищенко, *Математическая теория оптимальных процессов*, Москва 1961.
[31] Р. И. Стаховский, *Метод изохрон в решении задач оптималного управления*, Москва 1967.
[32] A. Turowicz, *Sur les trajectoires et quasitrajectoires des systèmes de commande non-linéaires*, Bull. Pol. Acad. Sci. Sér. Math. 10 (1962), str. 529-531.
[33] — *Sur les zones d'émission de trajectoires et de quasitrajectoires des systèmes de commande non-linéaires*, ibidem 11 (1963), str. 47-50.
[34] — *Remarque sur les zones d'émission de trajectoires et de quasitrajectoires des systèmes de commande non-linéaires*, ibidem 11 (1963), str. 241-243.
[35] T. Ważewski, *O problemie optymalnego sterowania w przypadku nieliniowym*, Arch. Autom. i Telemech. 7 (1962), str. 19-32.
[36] — *On an optimal control problem. Differential equations and their applications*, Proc. Confer. in Prague (1962), str. 229-242.
[37] S. K. Zaremba, *O równaniach paratyngensowych*. Dodatek do Roczników PTM 9 (1935).