

OPT

OŚRODEK POSTĘPU TECHNICZNEGO

**Katedra Automatyki i Telemechaniki
Politechniki Warszawskiej**

Władysław Findeisen

WSTĘP DO TECHNIKI SYSTEMÓW

**Do użytku
służbowego**

Kurs: Identyfikacja i Optymalizacja Procesów

MATERIAŁY SZKOLENIOWE

**WARSZAWA
1969-1970**

1. METODYKA OGÓLNA

1.1. Postępowanie od zadania do struktury i parametrów układu

Tok postępowania prowadzący do projektu układu sterowania określonym obiektem /procesem/ można przedstawić jako złożony z wielu etapów:

- /1/ Sformułowanie zadania
- /2/ Ustalenie matematycznego modelu obiektu
- /3/ Wyznaczenie sterowania optymalnego
- /4/ Wybór struktury układu sterowania
- /5/ Optymalizacja układu sterowania.

W pierwszym przybliżeniu można te etapy zilustrować następująco. Załóżmy, że obiekt jest scharakteryzowany przez zbiór wielkości wejściowych \underline{u} i zbiór współrzędnych stanu \underline{x} , a przedziałem czasu, w którym rozpatrujemy działanie obiektu jest $t_1 \leq t \leq t_2$. Wówczas:

Zadanie sterowania może brzmieć

$$\max Q [\underline{x}, \underline{u}, t, t_2, \underline{x}(t_2)] \quad /1/$$

gdzie Q jest wynikającym z zadania ekonomicznego funkcjonałem albo funkcją. Przy nieustalonym z góry czasie końcowym t_2 lub stanie końcowym $\underline{x}(t_2)$ wielkości te mogą wchodzić do wskaźnika jakości obok przebiegów wejściowych \underline{u} oraz stanu \underline{x} , co zaznaczono w zapisie /1/. Zadaniu ekstremalizacji wskaźnika Q towarzyszą z reguły ograniczenia

$$\text{na wielkości wejściowe, } \underline{u} \in U \quad /2/$$

$$\text{na współrzędne stanu, } \underline{x} \in X \quad /3/$$

Model matematyczny obiektu jest to na przykład opis jego własności w postaci równania stanu

$$\dot{\underline{x}} = \underline{f}(\underline{x}, \underline{u}, t) \quad /4/$$

wraz z warunkami początkowymi $\underline{x}(t_1) = \underline{x}_1$. Wśród warunków początkowych część może być narzucona z góry /np. skład surowca podlegający przeróbce w reaktorze/, a część może odgrywać rolę sterowań /np. dozowane ilości domieszek dodawanych na początku reakcji/. Przy opisie obiektu określone być mogą ograniczenia na wielkości wejściowe i współrzędne stanu, które wzbogacą zbiór /2/, /3/.

Wyznaczenie sterowania optymalnego polega na rozwiązaniu zadania:

$$\max_{\underline{u}'} \tilde{Q}$$

/5/

z uwzględnieniem /2/, /3/, /4/. W zapisie przyjęliśmy, że wśród wielkości wejściowych \underline{u} tylko część \underline{u}' , może wchodzić w grę jako wielkości sterujące. Pozostała część \underline{u}'' , stanowi zakłócenia bądź na wartości narzucone z góry. Ponieważ zakłócenia będą przypadkowe, ekstremalizować względem \underline{u}' można tylko wybraną ocenę statystyczną \tilde{Q} , co zaznaczono wężykiem. Rozwiązanie zadania /5/ dostarcza nam istotnej informacji: obliczone $\max_{\underline{u}'} \tilde{Q}$ stanowi największą wartość wskaźnika jakości, jaką można otrzymać w danym obiekcie. Jest to optymalizacja sterowania - nie bierzemy tu pod uwagę ograniczeń w realizacji sterowania optymalnego. Ograniczenia takie, wynikające np. z okresowego działania optymalizatora, jego skończonych możliwości obliczeniowych itd. obniżają osiąganą wartość \tilde{Q} .

Wybór struktury układu sterowania polega na podjęciu decyzji, jakie wielkości mają być mierzone i według jakich powiązań wyniki tych pomiarów mają wpływać na obiekt. Jest tu bardzo dużo możliwości wyboru. Na przykład, sterowanie w układzie otwartym mogłoby polegać na pomiarze zakłóceń \underline{u}'' i oddziaływaniu na wejścia \underline{u}' . Ale możnaby również mierzyć nie wszystkie elementy \underline{u}'' , lub mierzyć je okresowo - co sprowadza się do wyboru reprezentacji zakłóceń. Można też zrezygnować z części wejść \underline{u}' , obniżając w ten sposób koszty realizacji sterowania, z jednocześnie być może pogorszeniem wskaźnika jakości. Można również proponować układy zamknięte sterowania, układy mieszane itd. Na przykład, możemy dla naszego obiektu przyjąć strukturę układu sterowania jak na rys.1. Przyjęto tu rezygnację z części potencjalnych wielkości sterujących, zachowując \underline{u}''^* i \underline{u}' . Na rys.1 wybrane wielkości sterujące służą, w układzie zamkniętym, do otrzymania zadanej trajektorii stanu \underline{x} : stan obiektu jest mierzony i porównywany z wartościąadaną, jak zwykle w układach regulacji. Natomiast wartość zadana stanu jest wyznaczona jako wartość optymalna $\hat{\underline{x}}$, w oparciu o zmierzone bądź wyliczone cechy zakłóceń \underline{u}'' ; nie wszystkie zakłócenia są obserwowane bądź nie są obserwowane w pełni, co zaznaczono wprowadzając symbol \underline{u}''^* . Układ z rys.1 jest przykładem układu dwuwarstwowego, często stosowanego w praktyce.

Optymalizacja układu sterowania sprowadza się, po wyborze struktury, do odpowiedniego doboru algorytmów i współczynników, według których wykonywane będą działania wyznaczające wszelkie sterowania. Problem ten jest zbliżony do znanego z praktyki konwencjonalnej doboru typu regulatora /P, PI, PID/ i doboru nastaw. Zaznaczając symbolicznie na rys.1, że chodzi o dobór parametrów $\underline{\alpha}$ oraz parametrów $\underline{\beta}$, zadanie optymalizacji formuluje się jako

$$\max_{\underline{\alpha}, \underline{\beta}} \tilde{Q} \quad /6/$$

gdzie $\underline{\alpha}$, $\underline{\beta}$ podlegają zwykle pewnym ograniczeniom fizycznym, $\underline{\alpha} \in A$, $\underline{\beta} \in B$. Zadanie /6/ jest zadaniem optymalizacji układu sterowania, w odróżnieniu od poprzedniego zadania /5/, które było zadaniem optymalizacji sterowania.

Łatwo zauważyć, że zawsze będzie

$$\max_{\underline{\alpha}, \underline{\beta}} \tilde{Q} \leq \max_{\underline{u}'} \tilde{Q} \quad /7/$$

a różnica między tymi liczbami jest stratą jakości, wywołaną ograniczonością struktury i algorytmów sterowania. Zadaniem projektanta jest oczywiście znalezienie właściwego kompromisu między stratą jakości działania obiektu, a kosztem realizacji układu sterowania.

1.2. Dekompozycja zadań złożonych

Rzeczywisty obiekt automatyzacji jest zazwyczaj zbyt złożony na to, by można było bezpośrednio rozwiązywać zadanie typu /1/. Wyobraźmy sobie ciąg technologiczny jak pokazano schematycznie na rys.2. Zadanie optymalizacji może tu polegać, na przykład, na zapewnieniu minimum kosztów przy zadanej wielkości produkcji, albo na maximum wydajności przy zadanej ilości surowców itd. Jeśli oznaczyć taki wskaźnik przez Q oraz przyjmując, że \underline{u} jest zbiorem wszystkich wielkości sterujących /nie zwracając uwagi na to, na który z obiektów z rys.2 mają one wpływ bezpośredni/, to zadanie optymalnego sterowania można zapisać jako

$$\max_{\underline{u} \in U} Q(\underline{u}) \quad /8/$$

gdzie zaznaczono, że wybór \underline{u} jest ograniczony /przez równania

obiektów i ograniczenia nierównościowe/. Gdy liczba wielkości sterujących jest duża oraz duża jest liczba równań ograniczających, co związane jest z liczbą i rzędem obiektów w układzie z rys.1, to rozwiązanie zadania /8/ jest niemożliwe ze względów obliczeniowych. Skorzystać trzeba z możliwości dekompozycji zadania /8/. Główne zasady dekompozycji są następujące.

Założmy, że globalny wskaźnik jakości Q udało się przedstawić w postaci sumy

$$Q(\underline{u}) = Q_1(\underline{u}^1, \underline{v}) + Q_2(\underline{u}^2, \underline{v}) + \dots + Q_N(\underline{u}^N, \underline{v}) \quad /9/$$

gdzie każdy wskaźnik Q_i zależy od zmiennych koordynacyjnych \underline{v} oraz od takiego zbioru \underline{u}^i wielkości sterujących, który nie ma elementów wspólnych z żadnym innym zbiorem \underline{u}^j , $j \neq i$. Q_i nazywamy wskaźnikami lokalnymi. Jeżeli suma zbiorów $\underline{u}^1, \underline{u}^2, \dots, \underline{u}^N$, \underline{v} zawiera wszystkie elementy \underline{u} , to można napisać równoważność

$$\max_{\underline{u}} Q(\underline{u}) = \max_{\substack{\underline{u}^1 \\ \underline{u}^2 \\ \dots \\ \underline{u}^N \\ \underline{v}}} Q(\underline{u}, \underline{v}) \quad /10/$$

a z kolei, przy rozłączności zbiorów \underline{u}^i postulowanej wyżej, mamy możliwość dwustopniowej /dwupoziomowej/ optymalizacji

$$\max_{\substack{\underline{u}^1 \\ \underline{u}^2 \\ \dots \\ \underline{u}^N \\ \underline{v}}} Q(\underline{u}, \underline{v}) = \max_{\underline{v}} \left[\max_{\underline{u}^1} Q_1(\underline{u}^1, \underline{v}) + \max_{\underline{u}^2} Q_2(\underline{u}^2, \underline{v}) + \dots + \max_{\underline{u}^N} Q_N(\underline{u}^N, \underline{v}) \right] \quad /11/$$

W optymalizacjach "lokalnych" lub "pierwszego poziomu", tj. w zadaniach typu

$$\max_{\underline{u}^i} Q_i(\underline{u}^i, \underline{v}) \quad /12/$$

wielkość \underline{v} odgrywa rolę parametru, zatem optymalizacje te odbywać się mają przy pewnych wartościach narzuconych z góry. Ze sformułowań nie wynika, czy będą to z góry narzucone wartości

wielkości wejściowych czy też współrzędnych stanu obiektów składowych procesu złożonego, przedstawionego na rys.2. Możliwe jest jedno i drugie.

Istotą trudności wielopoziomowej optymalizacji jest nie podział wskaźnika jakości na części Q_1, Q_2, \dots, Q_N , lecz odpowiednie odseparowanie związków ograniczających, by można było możliwie wiele elementów \underline{u} pomieścić w zbiorach \underline{u}^1 . Problemem tym zajmiemy się później szczegółowo; teraz zwrócimy tylko uwagę, że zaznaczone we wzorze /8/ ograniczenie globalne $\underline{u} \in U$ musi zostać rozdzielone tak, by powstały ograniczenia lokalne zależne co najwyżej od zmiennych koordynacyjnych:

$$\underline{u}^1 \in U^1(\underline{v}), \underline{u}^2 \in U^2(\underline{v}), \dots, \underline{u}^N \in U^N(\underline{v}), \underline{v} \in V. \quad /13/$$

W pełniejszy zatem sposób zapiszemy zależności /10/, /11/ jako

$$\max_{\underline{u}} Q(\underline{u}) = \max_{\underline{u}, \underline{v}} Q(\underline{u}, \underline{v}) \quad /10'/$$

$$\underline{u} \quad \underline{u}^1 \in U^1(\underline{v})$$

$$\underline{u}^2 \in U^2(\underline{v})$$

...

$$\underline{u}^N \in U^N(\underline{v})$$

$$\underline{v} \in V$$

$$\max_{\underline{v} \in V} Q(\underline{u}, \underline{v}) = \max_{\underline{v} \in V} \left[\max_{\underline{u}^1 \in U^1(\underline{v})} Q_1(\underline{u}^1, \underline{v}) + \max_{\underline{u}^2 \in U^2(\underline{v})} Q_2(\underline{u}^2, \underline{v}) + \dots + \max_{\underline{u}^N \in U^N(\underline{v})} Q_N(\underline{u}^N, \underline{v}) \right] \quad /11'/$$

W odniesieniu do układu z rys.2 możemy w pierwszym przybliżeniu przyjąć, że zmiennymi koordynacyjnymi będą te wszystkie wielkości, za pośrednictwem których obiekty składowe układu wpływają na siebie wzajemnie. Będą to więc, na przykład, natężenia przepływu materiałów, wielkości charakteryzujące ich skład chemiczny, ich temperatury i tak dalej. Nie uściślając tego bardziej w obecnej chwili stwierdzamy, że po dokonaniu dekompozycji zadania globalnego mamy przed sobą N zadań sterowania poszczególnymi obiektami. Zadania te będziemy następnie rozpatrywać tak jak to pokazano w punkcie poprzednim. Pojawia się oczywiście zadanie dodatkowe, mianowicie koordynacja zadań lokalnych przez wybór \underline{v} zapewniający ekstremum globalne, patrz /11/. Jest to jednak, w zasadzie, ten sam typ zadania - optymalizacja funkcjonału w obecności ograniczeń.

2. FORMUŁOWANIE ZADANIA OPTYMALIZACJI

2.1. Uwagi ogólne

Ustalony dla obiektu wskaźnik jakości, który ma podlegać ekstremalizacji, musi mieć określony sens ekonomiczny i techniczny. Wynikać on zatem powinien z rzeczywistej oceny znaczenia ekonomicznego poszczególnych cech i efektów procesu. Przykładem wskaźników mogą być: zysk z prowadzenia procesu, koszt, wydajność itd. Można przyjąć jako regułę, że wyboru tego wskaźnika nie może dokonać samodzielnie projektant układu sterowania. Zwrócimy tu tylko uwagę na to, że w większości przypadków sformułowaniu wskaźnika jakości towarzyszyć musi sprecyzowanie dodatkowych związków ograniczających. Na przykład, minimalizacja kosztów produkcji, przeprowadzana bez dodatkowego ograniczenia:

$$\text{wielkość produkcji} \geq \text{wartości zadanej}$$

doprowadziłyby oczywiście do rozwiązania, przy którym koszty będą minimalne, ale produkcja równa zero. Podobnie, maksymalizacja wielkości produkcji bez ograniczenia zużycia surowców czy energii może dać w wyniku matematycznym złe wykorzystanie surowców, itd.

Z punktu widzenia sformułowania zadania optymalizacji bardzo istotne jest, czy rozpatrujemy sytuację deterministyczną czy też stochastyczną, to znaczy czy uwzględniamy fakt istnienia przypadkowych zakłóceń. W związku z tym możemy rozróżnić:

a/ Zadanie sterowania optymalnego, gdy żądamy

$$\max Q \quad /14/$$

bądź też maximum pewnej cechy statystycznej, na przykład

$$\max E [Q] \quad /15/$$

albo, co stanowi przypadek ważny dla niektórych zastosowań

$$\max \min [Q] \quad /16/$$

Zapis /16/ oznacza, że chodzi o sterowanie zapewniające maximum minimalnej wartości Q /maximum względem sterowania, minimum względem zakłóceń/.

b/ Zadanie sterowania optymalno-wystarczającego, gdy żądamy

$$\max Q, \quad \text{lub też} \quad \max E [Q] \quad /17/$$

przy jednoczesnym warunku, by

$$\min [Q] \geq Q_d \quad /18/$$

Zapis /17/, /18/ oznacza, że mamy maksymalizować Q lub $E [Q]$, ale że przy najmniej sprzyjającym zakłóceniu wartość Q nie ma spaść poniżej pewnej granicy.

Czytelnik zauważy, że np. sformułowanie /15/ wymagać będzie przy rozwiązywaniu zadania znajomości rozkładów prawdopodobieństwa zakłóceń, a sformułowanie /16/ oraz /18/ - wymagać będzie znajomości przedziału zmienności zakłóceń, natomiast nie konieczności ich rozkładów.

Zadania optymalizacji można dalej klasyfikować na statyczne i dynamiczne. Rozróżnienie to jest nader ważne z punktu widzenia metod rozwiązywania. Nie będziemy tego rozwijać szerzej w tym miejscu. Zwróćmy tylko uwagę, że w zadaniach statycznych poszukuje się rozwiązań w postaci pewnych liczb /np. najlepszych parametrów prowadzenia procesu ustalonego - ciśnień, temperatur itp./, a w zadaniach dynamicznych rozwiązaniami są funkcje czasu - na przykład przebieg w czasie prądu łuku w elektrycznym piecu stalowniczym. Zadania dynamiczne są znacznie bardziej pracochłonne w rozwiązywaniu, to znaczy absorbują znacznie większe środki obliczeniowe.

2.2. Przykład formułowania zadania dla obiektu

Rozpatrzmy jako obiekt sterowania reaktor chemiczny o idealnym zmieszaniu, przedstawiony na rys.3. Do reaktora dopływa materiał A z natężeniem F_A kg/s o temperaturze T_A °K, materiał B z natężeniem F_B kg/s o temperaturze T_B °K, odpływa zaś strumień materiału o natężeniu F_D kg/s. W reaktorze odbywa się reakcja $A \rightarrow B$ z prędkością $r = C_A k(T) s^{-1}$, gdzie T temperatura w reaktorze, $k(T)$ stała prędkości reakcji, C_A stężenie składnika A. Reakcja jest endotermiczna, ciepło reakcji wynosi h J/kg, do reaktora doprowadzany jest strumień energii cieplnej o natężeniu H J/s.

Jeżeli reaktor z rys.3 pracuje sposobem ciągłym, to zadanie optymalizacji będzie sformułowane w odniesieniu do stanu ustalonego odbywającego się w reaktorze procesu. Może to być np. zadanie następujące:

osiągnąć w stanie ustalonym $\max (1 - C_A) F_D$ kg/s /19/
przy ograniczeniach

$$F_B \leq F_{Bm} \quad /20/$$

$$C_A \leq C_{Am} \quad /21/$$

Wielkość $(1 - C_A)F_D$ jest natężeniem wypływu składnika B w kg/s, czyli ma charakter natężenia produkcji. Ograniczenie $F_B \leq F_{Bm}$ ma sens oczywisty, a ograniczenie $C_A \leq C_{Am}$ jest warunkiem na jakość, czy "czystość" strumienia wypływającego /nie ma on zawierać zbyt dużo nieprzereagowanego składnika A/.

Przy istnieniu przypadkowych zakłóceń, wskaźnik jakości należałoby sformułować w postaci wartości średniej

$$\max E \left[(1 - C_A)F_D \right] \quad \text{kg/s} \quad /19'/$$

To oczywiste sprecyzowanie można często, w fazie pierwszego formułowania zadania, pomijać. Mniej oczywiste byłoby natomiast uwzględnienie przypadkowego charakteru zjawiska w ograniczeniach. Na przykład ograniczenie nierównościowe /20/ może być w istocie warunkiem

$$E [F_B] \leq F_{Bm} \quad /20'/$$

albo też

$$\max [F_B] \leq F_{Bm} \quad /20''/$$

Czytelnik łatwo zinterpretuje techniczny sens obu tych sformułowań oraz zauważy, że to które z nich będzie właściwe zależy np. od istnienia zbiornika buforowego materiału B przed reaktorem.

Dla tego samego reaktora, gdyby pracował on sposobem cyklicznym, aktualne byłoby postawienie zadania optymalizacji dynamicznej. Na przykład, możnaby żądać takiego wyboru początkowego załadowania reaktora oraz przebiegu czasowego dopływu ciepła H , aby osiągnąć zadany końcowy produkt w reaktorze, przy jednoczesnej ekstremalizacji wybranego wskaźnika jakości. Na przykład, mogłoby to być zadanie, ażeby w całkowicie załadowanym reaktorze

$$\text{minimalizować czas trwania procesu } t_2 - t_1 = \int_{t_1}^{t_2} dt \quad /22/$$

$$\text{z ograniczeniem na skład końcowy } C_A(t_2) \leq C_{A2} \quad /23/$$

$$\text{oraz na skład początkowy } C_A(t_1) \geq C_{A1} > C_{A2} \quad /24/$$

W ujęciu statystycznym, możemy zadanie /22/ sformułować jako

$$\min E \left[\int_{t_1}^{t_2} dt \right] \quad /22'/$$

$$H(t)$$

Minimalizacja wartości średniej, wskazana przez /22'/, ma sens oczywisty. Gdy jednak rozpatrywany reaktor stanowi jedno z ogniw złożonego procesu technologicznego, może być na przykład niedopuszczalne by czas $t_2 - t_1$ przekroczył kiedykolwiek pewną wartość graniczną t_g , ustaloną z warunków koordynacji całości procesu. Wówczas wymaganie /22'/ uzupełnić trzeba dodatkowym ograniczeniem nierównościowym

$$\max \left[\int_{t_1}^{t_2} dt \right] \leq t_g \quad /25/$$

gdzie symbol max odnosi się oczywiście do wartości maksymalnej ze względu na zakłócenia.

Czytelnik zwróci uwagę, że w ustaleniach typu /19/ czy /22/ nie zastanawiamy się nad tym, czy wchodzące w te wzory wielkości są wejściami (\underline{u}), współrzędnymi stanu (\underline{x}) czy innymi jeszcze zmiennymi charakteryzującymi obiekt. Są to po prostu te zmienne, które mają określony charakter techniczno-ekonomiczny.

2.3. Przykład dekompozycji zadania w układzie złożonym

W pracy [3] przedstawiono m.inn. sformułowanie zadania optymalizacji stanu ustalonego dla surowni cukrowni. Układ składa się z trzech obiektów: stacja dyfuzji, stacja oczyszczania soku i stacja wyparna. Przy zadanej wielkości przerobu buraków, optymalizacja ekonomiczna sprowadza się do minimalizacji kosztów /włączając w koszty również straty cukru/. Zadanie sformułowano łącznie jako minimalizację natężenia strat w zł/h:

$$\min [Q = \gamma_1 F_w c_{kw} + \gamma_2 (F_1 c_{k1} - F_2 c_{k2}) + (\gamma_2 - \gamma_3) F_2 \sum_{i=1}^{19} c'_{M1} K_{M1} + \gamma_4 F_p + F_2 \sum_{i=1}^4 \gamma_5 T_i^2] \quad /26/$$

W wyrażeniu tym ujęto najważniejsze składniki kosztów i strat. $F_w c_{kw}$ oznacza stratę cukru przez przejście do wysłodków /patrz rys.4/, $F_1 c_{k1} - F_2 c_{k2}$ stanowi stratę cukru w stacji oczyszczania $\gamma_4 F_p$ wyraża koszt pary grzejnej. Składnik zawierający sumę $c'_{M1} K_{M1}$ oznacza stratę cukru przez przejście do melasu - przejście to jest związane z zawartościami dziewiętnastu różnych niecukrów melasotwórczych c'_{M1} w soku za stacją

oczyszczania oraz współczynnikami absorpcji cukru przez i-ty niecukier K_{M1} . Wreszcie, ostatni składnik wyraża stratę wartości cukru /stratę jakości/ przez przebywanie w zbyt wysokiej temperaturze T_1 w i-tym dziale wyparki. Zwróćmy uwagę na niektóre współczynniki γ_i , które są cenami: γ_1 jest ceną cukru liczoną za stacją dyfuzji /zatem bez kosztów oczyszczania/, γ_2 - cenę cukru w soku za stacją oczyszczania, γ_3 - cenę melasu.

Dla rozwiązywania zadań optymalizacji i sterowania, wskaźnik /26/ rozłożono tak, by otrzymać wskaźnik oddzielny dla każdej ze stacji, ale by ich suma stanowiła nadal wartość Q :

$$Q_1 + Q_2 + Q_3 = Q \quad /27/$$

Wskaźniki Q_1, Q_2, Q_3 ustalono jak następuje:

$$Q_1 = \gamma_1 F_w c_{kw} + (\gamma_2 - \gamma_3) F_1 \sum_{i=1}^{19} c_{\mu i} K_{\mu i} \quad /28/$$

$$Q_2 = \gamma_2 (F_1 c_{k1} - F_2 c_{k2}) - (\gamma_2 - \gamma_3) \left[F_1 \sum_{i=1}^{19} c_{\mu i} K_{\mu i} - F_2 \sum_{i=1}^{19} c'_{\mu i} K_{\mu i} \right] \quad /29/$$

$$Q_3 = \gamma_4 F_p + F_2 \sum_{i=1}^4 \gamma_5 T_i^2 \quad /30/$$

Zwróćmy specjalną uwagę na budowę wskaźnika dla stacji dyfuzji, wzór /28/. Straty cukru przez przejście do melasu, jak ujęto we wzorze /26/, zależą od zawartości niecukrów w soku oczyszczonym. Pojawienie się jednak niecukrów zależy od pracy stacji dyfuzji /głównie od rozkładu temperatury/. W związku z tym konieczne było uwzględnienie niecukrów w ocenie procesu dyfuzji, stąd odpowiedni składnik we wzorze /28/, zawierający udziały niecukrów melasotwórczych $c_{\mu i}$. Straty stąd pochodzące ulegają jednak zmniejszeniu przez akcję oczyszczania soku, co z kolei uwzględniono w ocenie pracy stacji oczyszczania - we wzorze /29/ występuje składnik ze znakiem przeciwnym do znaku strat. Dzięki takiej budowie wskaźników /28/, /29/ udaje się uzyskać zależność każdego z nich od tych wszystkich istotnych wielkości sterujących każdej ze stacji, które mają wpływ na

wskaźnik globalny /26/. Czytelnik łatwo sprawdzi, że spełniony jest również związek /27/. Zdajemy sobie oczywiście sprawę z tego, że niektóre wielkości sterujące wpływać mogą sprzecznie na wskaźniki częściowe /28/ - /30/. Na przykład, zwiększenie ilości wody dodawanej do krajanki w stacji dyfuzji poprawi pracę tej stacji /zmniejszy straty F_{wckw} /, ale równocześnie zwiększy ilość pary potrzebnej w stacji wyparnej do zagęszczenia soku, czyli zwiększy składnik $\gamma_4 F_p$. W schemacie optymalizacji dwupoziomowej, jaką zastosowalibyśmy do rozpatrywanego przykładu, ilość wody dodawanej do krajanki będzie zatem musiała być zmienną koordynacyjną, której wartość zostanie zdecydowana na wyższym poziomie.

3. MODEL MATEMATYCZNY OBIEKTU

3.1. Metodyka postępowania

Skoncentrujemy uwagę na opisie właściwości obiektu dynamicznego przez równania stanu

$$\dot{\underline{x}} = \underline{f}(\underline{x}, \underline{u}, t) \quad /31/$$

wraz z równaniami wyjść

$$\underline{y} = \underline{g}(\underline{x}, \underline{u}, t) \quad /32/$$

Dodatkowo do równań /31/ podać należy warunki początkowe $\underline{x}(t_1)$, a do obu równań - ewentualne ograniczenia nierównościowe na \underline{x} oraz \underline{u} .

Opis typu /31/, /32/ wydaje się na pierwszy rzut oka nieogodny o tyle, że wymaga znajomości bądź ustalenia rzędu obiektu, czyli liczby równań w układzie /31/, to jest liczby współrzędnych stanu. Okazuje się jednak, że przy ustalaniu opisu obiektu z danych fizycznych, czyli przy modelu "analitycznym", operowanie pojęciem współrzędnych stanu jest właśnie bardzo dogodne i przejrzyste od strony zjawiskowej. Natomiast przy identyfikacji modelu obiektu całkowicie doświadczalnej, tylko pozornie można ominąć sprawę ustalenia rzędu obiektu. Mianowicie, gdy opisywać obiekt przy pomocy operatora "wejście - wyjście"

$$\underline{y}(t) = A[\underline{u}(t)] \quad /33/$$

pamiętać należy, że przebieg na wyjściu $\underline{y}(t)$ zależy również

od warunków początkowych, które nie są niczym innym jak stanem początkowym obiektu. Napisać zatem należy

$$\underline{y}(t) = A [\underline{u}(t), \underline{x}(t_1)] \quad /34/$$

a tu już identyfikacja wektora liczbowego $\underline{x}(t_1)$ wymaga stwierdzenia, ile on zawiera elementów. Dodatkowo, fakt że w serii doświadczeń identyfikacyjnych niektóre elementy $\underline{x}(t_1)$ okazały się zerowe nie przesądza sprawy na sytuacje przyszłe i włączyć się tu należy w zagadnienie obserwowalności obiektu - a to znowu sprowadza nas na grunt zdefiniowania współrzędnych stanu.

Łatwo zauważyć, że pochodne w stosunku do równań /31/, /32/ formy opisu własności obiektu otrzymamy przez odpowiednie przeróbki. I tak:

Charakterystyki statyczne otrzymamy zakładając stan ustalony, to jest /mowa tu o obiekcie niezależnym od czasu/

$$\underline{\dot{q}} = \underline{f}(\underline{x}, \underline{u}) \quad /35/$$

$$\underline{y} = \underline{g}(\underline{x}, \underline{u}) \quad /36/$$

oraz, jeśli zachodzi potrzeba, rugując zmienną \underline{x} ze - związków /35/, /36/.

Transmitancje macierzowa obiektu otrzymamy po linearyzacji związków /31/, /32/, to jest wychodząc z równań /dla przyrostów/

$$\dot{\underline{x}} = A \underline{x} + B \underline{u} \quad /37/$$

$$\underline{y} = C \underline{x} + D \underline{u} \quad /38/$$

gdzie A, B, C, D macierze niezależne od czasu. Dokonując transformacji równań /37/, /38/ przy zerowych warunkach początkowych, otrzymamy rugując $\underline{x}(s)$ wzór na transmitancję macierzową

$$\underline{G}(s) = C [s I - A]^{-1} B + D \quad /39/$$

Postępowanie zmierzające do ustalenia opisu /31/, /32/ na drodze analitycznej, to jest na podstawie danych o obiekcie, można przedstawić jako złożone z kilku etapów:

- a/ Określenie współrzędnych stanu /oraz ewent. ograniczeń na te współrzędne/
- b/ ułożenie równań stanu
- c/ określenie wielkości wejściowych /oraz ewent. ograniczeń na te wielkości/
- d/ określenie wielkości wyjściowych

e/ ułożenie równań wyjścia.

Podkreślić należy, że etapy a/, b/, c/ wyczerpują opis dynamiki obiektu. Można je wykonywać bez znajomości zadania optymalizacji, a tym bardziej przed określeniem struktury układu sterowania. Natomiast wybór wielkości wyjściowych, etap d/, zależy od tego, czy zmienne takie występują w zadaniu, to znaczy w sformułowaniu wskaźnika jakości i towarzyszących temu sformułowaniu ograniczeniach. Wybór wielkości wyjściowych zależy również od struktury układu sterowania czyli od decyzji, jakie zmienne będą w obiekcie mierzone. To samo dotyczy oczywiście etapu e/.

Określenie współrzędnych stanu rozumieć należy, zgodnie z ich definicją, jako znalezienie takiego zestawu zmiennych, których wartości w dowolnej chwili t "podsumowałyby" całą przeszłość procesu; interesuje nas przy tym zestaw zawierający minimalną liczbę zmiennych. W praktyce może się okazać, że niektóre ze współrzędnych stanu nie są istotne dla rozpatrywanych aspektów procesu w obiekcie i można zrezygnować z ich uwzględnienia w opisie. Zależy to od zadania sterowania; nie będziemy tego analizować obecnie, lecz dopiero w punkcie 4.

Ułożenie równań stanu polega na sformułowaniu, dla n współrzędnych stanu x_1, x_2, \dots, x_n układu n równań pierwszego rzędu o postaci

$$\frac{dx_1}{dt} = f_1(x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_r, t), x_1(t_1) = x_{10}$$

/40/

$$\dots \dots \dots$$

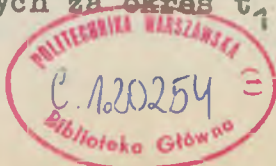
$$\frac{dx_n}{dt} = f_n(x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_r, t), x_n(t_1) = x_{n0}$$

Wypisany przykład równań odnosi się do obiektu o stałych skupionych. Jeśli nie zaznaczono inaczej, symbole $x_1, x_2, \dots, u_1, u_2, \dots$ rozumiane są jako wartości odpowiednich zmiennych w tej samej chwili t.

W odróżnieniu od tego, obiekt z opóźnieniami stanu będzie opisany równaniami typu

$$\frac{dx_1}{dt} = f_1(x_1(t-h_1), x_2(t-h_2) \dots, u_1(t), u_2(t), \dots, t) \quad /41/$$

gdzie h_1 odpowiednie opóźnienia. Warunkami początkowymi dla takiego obiektu nie są już tylko wartości $x_1(t_1)$, jak w przypadku /40/, lecz "stan początkowy zupełny", to znaczy przebiegi każdej ze współrzędnych za okres $t_1 - h_1$:



$$x_1(t), \quad t_1 - h_1 \leq t \leq t_1$$

Obiekt o stałych rozłożonych, jakim będzie np. reaktor rowy, opisany będzie często równaniami cząstkowymi typu

$$\frac{\partial x_1}{\partial t} = f_1(x_1, x_2, \dots, \frac{\partial x_1}{\partial t}, \frac{\partial x_2}{\partial t}, \dots, u_1, u_2, \dots, u_r, t) \quad /43/$$

We wszystkich tych opisach jest rzeczą istotną, że po prawej stronie równań stanu nie występują pochodne współrzędnych stanu względem czasu, pochodne wielkości wejściowych względem czasu oraz wielkości wyjściowe. Jeżeli zatem w równaniu stanu wystąpi np. strumień produktu wyjściowego z reaktora i nie da się go wyeliminować zależnością funkcyjną od współrzędnych stanu oraz od innych wielkości wejściowych, to ten strumień jest z punktu widzenia matematycznego modelu obiektu wielkością wejściową. Ma to np. miejsce gdy strumień wypływający jest wymuszany pompą, gdyż nie zależy on wówczas od poziomu cieczy w zbiorniku.

Określenie wielkości wejściowych polega na przeglądzie prawych stron równań stanu. Te wielkości, które tam występują a nie są współrzędnymi stanu, mogą być zmiennymi wejściowymi, lub parametrami. Decyzja co do tego leży przynajmniej częściowo w rękach projektanta układu, który może na przykład zdecydować, że temperatura czynnika dopływającego do reaktora będzie stała /stanie się wtedy parametrem w równaniach/, lub też będzie sterowana /stanie się wielkością wejściową sterującą/, albo pozostanie zmienna w sposób przypadkowy /będzie wielkością wejściową zakłócającą/. Decyzje tego rodzaju należą do wyboru struktury układu sterowania - w stadium ustalania modelu obiektu należy sobie tylko zdać sprawę z tego, jak wpływają na proces poszczególne wielkości wejściowe, czyli jak one występują w równaniach stanu. Należy również ustalić, jakim ograniczeniom /technologicznym/ podlega każda z wielkości wejściowych.

Określenie wielkości wyjściowych jest stosunkowo proste; wspomnieliśmy już, że będą to w pierwszym rzędzie te zmienne występujące we wskaźniku jakości i ograniczeniach zewnętrznych, które nie są współrzędnymi stanu ani wielkościami wejściowymi, czyli nie wystąpiły w równaniach stanu. W całości modelu są bowiem tylko trzy rodzaje zmiennych: wielkości wejściowe, współrzędne stanu i wielkości wyjściowe, zgodnie ze schematem rys.5.

Ułożenie równań wyjść polega na sformułowaniu zależności

wielkości wyjściowych $y_1, y_2 \dots y_m$ od współrzędnych stanu i wielkości wejściowych. Zależności te nie są już równaniami różniczkowymi, gdyż całość dynamiki obiektu zawarta jest w równaniach stanu. Ponieważ równania wyjść mogą uzależniać y_1 wprost od wielkości wejściowych u /patrz /32/ może się okazać, że trzeba na tym etapie uzupełnić listę wielkości wejściowych /może istnieć zmienna nie wchodząca w równania stanu, a występująca np. wprost we wskaźniku jakości/.

3.2. Przykład obiektu o stałych skupionych

Rozpatrzmy ułożenie modelu matematycznego dla reaktora przedstawionego już poprzednio na rys.3. Postępować będziemy według metodyki omówionej w punkcie 3.1., to znaczy wykonamy:

- określenie współrzędnych stanu
- ułożenie równań stanu
- określenie wielkości wejściowych

Współrzędne stanu wybieramy wychodząc ze stwierdzenia, że przeszłość procesu w reaktorze zostanie podsumowana jednoznacznie trzema liczbami:

- W_A - ilość nagromadzonego składnika A, kg
- W_B - ilość nagromadzonego składnika B, kg /44/
- E - ilość nagromadzonego ciepła, J

Dogodniej jednak będzie ułożyć równania stanu, gdy jako trzy współrzędne stanu przyjmiemy wielkości nieco inne:

- W - ilość materiału w reaktorze, kg
- C_A - stężenie wagowe składnika A, bez wymiaru /45/
- T - temperatura mieszaniny, °K

Wielkości te są oczywiście w prosty sposób związane z poprzednimi. Podkreślamy tu, że wybór współrzędnych stanu nie jest jednoznaczny, natomiast dla danego obiektu liczba współrzędnych stanu jest określona jednoznacznie. Czynniki decydujące o wyborze tego czy innego zestawu zmiennych jako współrzędnych stanu, mogą być różne. W naszym przykładzie zwróćmy uwagę na to, że w zestawie /45/ występuje zmienna C_A , która wchodzi również do wskaźnika jakości procesu /19/ i ograniczenia /21/, oraz zmienna T , na którą spodziewać się możemy ograniczenia, która wchodzi w stałą prędkość reakcji $k(T)$ i która jest zarazem istotną charakterystyką strumienia materiału opuszczającego reaktor.

Wystąpią następujące ograniczenia współrzędnych stanu:

$$0 \leq W \leq W_m, \text{ gdzie } W_m \text{ pojemność reaktora} \quad /46/$$

$$T \leq T_m, \text{ gdzie } T_m \text{ maksymalna dopuszczalna temperatura} \quad /47/$$

Ułożenie równań stanu opiera się z reguły na bilansie masy oraz bilansie energii. W naszym przypadku, z ogólnego bilansu masy mamy

$$\frac{dW}{dt} = F_A + F_B - F_D \quad \text{kg/s} \quad /48/$$

Z bilansu masy dla składnika A

$$\frac{d(W C_A)}{dt} = F_A - C_A F_D - W C_A k(T) \quad \text{kg/s} \quad /49/$$

Z bilansu energii cieplnej

$$\frac{d(c W T)}{dt} = c F_A T_A + c F_B T_B - c F_D T + H - h W C_A k(T) \quad /50/$$

W równaniu /50/ c oznacza ciepło właściwe J/kg-deg, dla uproszczenia przyjęte jako jednakowe dla wszystkich substancji. Równania /49/, /50/ nie są jeszcze równaniami stanu, gdyż jako współrzędne stanu przyjęliśmy C_A , T , a nie ilość składnika A oraz ilość energii. Ponieważ jednak

$$\frac{d(W C_A)}{dt} = W \frac{d C_A}{dt} + C_A \frac{dW}{dt} \quad /51/$$

a wielkość $\frac{dW}{dt}$ jest określona przez /48/, znając /49/ oraz /51/ łatwo napiszemy równanie na $\frac{dC_A}{dt}$. Podobnie postąpimy z zależnością /50/ i ostatecznie otrzymamy zestaw równań stanu w końcowej postaci:

$$\frac{dW}{dt} = F_A + F_B - F_D \quad \text{kg/s} \quad /52/$$

$$\frac{dC_A}{dt} = \frac{1 - C_A}{W} F_A - \frac{C_A}{W} F_B - C_A k(T) \quad 1/s \quad /53/$$

$$\frac{dT}{dt} = \frac{T_A - T}{W} F_A + \frac{T_B - T}{W} F_B + \frac{H}{cW} - \frac{h C_A k(T)}{c} \quad \text{deg/s} \quad /54/$$

Warunkami początkowymi w chwili $t = t_1$ dla tego układu równań będą oczywiście wartości załadowania $W(t_1)$, udziału składnika A, $C_A(t_1)$, oraz temperatury $T(x_1)$.

Określenie wielkości wejściowych odbywa się przez przegląd równań stanu /52/, /53/, /54/. Oprócz współrzędnych stanu występują tam zmienne

$F_A, F_B, F_D, T_A, T_B, H$

/55/

Jest to zestaw wielkości wejściowych, gdyż żadnej z nich nie można wyrugować przy pomocy zmiennych pozostałych i współrzędnych stanu. W szczególności, natężenie odpływu produktu F_D nie zależy od zasobu W w schemacie procesu jak na rys.3.

Nie decydujemy na razie, które z sześciu wielkości wejściowych będą wielkościami sterującymi, a które pozostawimy jako dopuszczalne zakłócenia.

Zwróćmy tylko uwagę na niewątpliwe ograniczenia wielkości wejściowych:

$$F_A \geq 0, F_B \geq 0, F_D \geq 0, H \geq 0 \quad /56/$$

W szczególności, $H \geq 0$ oznacza, że reaktor nie może być chłodzony, a zatem jego temperatura T jest pośrednio ograniczona od dołu, a nie tylko od góry jak podaje /47/.

3.3. Przykład obiektu o stałych rozłożonych

Rozpatrzmy teraz, w podobny sposób, obiekt przedstawiony schematycznie na rys.6. W podłużnym reaktorze przesuwany jest materiał B z prędkością v_2 m/s, a w przeciwnym kierunku porusza się materiał A z prędkością v_1 m/s. Między tymi składnikami zachodzi reakcja $A \rightarrow B$ z prędkością $r = C_A k(T)$ 1/s. Reakcja jest endotermiczna, reaktor jest ogrzewany z zewnątrz w sposób wymuszający profil temperatury $T(l, t)$.

Jako współrzędne stanu elementu długości reaktora, zaznaczonego na rys.6, można przyjąć

ilość materiału A: $dW_A = x_1 dl \quad /57/$

gdzie $x_1(l, t)$ gęstość zapełnienia substancją A w kg/m

ilość materiału B: $dW_B = x_2 dl \quad /58/$

gdzie $x_2(l, t)$ gęstość zapełnienia substancją B w kg/m

ilość nagromadzonego ciepła: $(c_1 dW_A + c_2 dW_B) T \quad /59/$

Przy zrobionym wyżej założeniu, że temperatura $T(l, t)$ będzie narzucona z zewnątrz, współrzędna stanu /59/ jest nieistotna. Jako pozostałe dwie współrzędne stanu dogodnie będzie przyjąć $x_1(l, t)$ oraz $x_2(l, t)$. Jako niewątpliwe ograniczenie stanu przyjmujemy

$$x_1 + x_2 \leq x_m$$

/60/

2233

gdzie x_m maksymalna gęstość zapełnienia reaktora w kg/m.

Równania stanu dla elementu reaktora wypiszemy w zależności bilansowych.

Dla składnika A otrzymamy

$$\frac{d}{dt} (dW_A) = v_1 x_1 (1 + dl, t) - v_1 x_1 (1, t) - k(T) x_1 (1, t) dl \quad \text{kg/s} \quad /61/$$

gdzie ostatni wyraz oznacza ubytek A skutkiem reakcji, zachodzącej w elemencie reaktora o długości dl.

Po wprowadzeniu $dW_A = x_1 dl$ oraz wykorzystaniu rozwinięcia

$$x_1 (1 + dl, t) = x_1 (1, t) + \frac{\partial x_1}{\partial l} dl \quad /62/$$

uzyskujemy równanie stanu

$$\frac{\partial x_1}{\partial t} = v_1 \frac{\partial x_1}{\partial l} - k(T) x_1 (1, t) \quad \text{kg/m.s} \quad /63/$$

Dla składnika B napiszemy

$$\frac{d}{dt} (dW_B) = v_2 x_2 (1, t) - v_2 x_2 (1 + dl, t) + k(T) x_1 (1, t) dl \quad \text{kg/s} \quad /64/$$

i w konsekwencji podobnego przekształcenia jak wyżej otrzymamy równanie stanu

$$\frac{\partial x_2}{\partial t} = -v_2 \frac{\partial x_2}{\partial l} + k(T) x_1 (1, t) \quad \text{kg/m.s} \quad /65/$$

Równaniom /63/, /65/ towarzyszyć muszą cztery warunki brzegowe, na przykład dotyczące:

$$\text{stanu początkowego reaktora: } x_1 (1, 0), x_2 (1, 0) \quad /66/$$

$$\text{wartości } x_1, x_2 \text{ na krańcach: } x_1 (L, t), x_2 (0, t) \quad /67/$$

Łatwo zauważyć, że warunki /67/ są dane przez natężenia dopływu i odpływu materiału:

$$x_1 (L, t) = \frac{F_A (L, t)}{v_1} \quad /68/$$

$$x_2 (0, t) = \frac{F_B (0, t)}{v_2} \quad /69/$$

Wielkości wejściowe wynikną z przeglądu równań stanu. Rozpatrując prawe strony /63/, /65/ widzimy, że na proces możemy wpływać przez:

prędkości przesuwu materiału $v_1(t), v_2(t)$ /70/

temperaturę $T(l, t)$ /71/

Temperatura $T(l, t)$ jest w istocie współrzędną stanu, ale ponieważ założyliśmy, że będzie ona wymuszona, spełnia ona dla równań stanu /63/, /65/ rolę wielkości wejściowej /nie jest zależna od stanu x_1, x_2 oraz pozostałych wejść v_1, v_2 /.

Zwróćmy uwagę, że wielkości wejściowe /70/, /71/ nie wyczerpują możliwości sterowania obiektem - jak zawsze, można na proces wpływać przez warunki brzegowe. W rozpatrywanym przypadku, dość oczywiste byłoby użycie jako wielkości sterujących warunków /68/, /69/.

Ograniczenia na wielkości wejściowe /70/, /71/ będą zapewne

$$v_1 \geq 0, \quad v_2 \geq 0, \quad T \leq T_m \quad /72/$$

4. ANALIZA ROZWIĄZANIA ZADANIA OPTYMALIZACJI

4.1. Uwagi ogólne

Mając sformułowany wskaźnik jakości Q , dane ograniczenia oraz model matematyczny obiektu możemy myśleć o rozwiązaniu zadania optymalizacji, czyli w zasadzie o rozwiązaniu zadania

$$\max_Q \quad \underline{u} \quad /73/$$

Ponieważ zadanie /73/ rozwiązywane jest z uwzględnieniem równań obiektu

$$\dot{\underline{x}} = \underline{f}(\underline{x}, \underline{u}, t) \quad /74/$$

otrzymamy wynik w postaci sterowania optymalnego $\hat{u}(t)$ określa zarazem przebieg optymalny stanu obiektu, czyli w interpretacji przestrzennej - tak zwaną optymalną trajektorię stanu $\hat{x}(t)$. W sumie otrzymamy wyniki w postaci:

$$\hat{u}(t), \quad \hat{x}(t), \quad \hat{Q} = Q(\hat{u}, \hat{x}) \quad /75/$$

Metody rozwiązywania zadań typu /73/ są przedmiotem teorii optymalizacji. W tym miejscu zastanowimy się nad tym, jakie wnioski z dostarczanych przez teorię rozwiązań można wyciągnąć dla struktury układu sterowania obiektem. Najważniejsze problemy jakie wypada tu poruszyć są następujące:

Istotne i nieistotne wielkości wejściowe

Rozwiązanie \hat{u} (t) opiewa w zasadzie na wszystkie wielkości wejściowe występujące w równaniach stanu obiektu. Zaznaczyliśmy, co prawda, już w punkcie 1.1, że część u " spośród \underline{u} trzeba niekiedy z góry przyjąć za zakłócenia. Wówczas oczywiście rozwiązanie optymalne określa tylko \hat{u} . Ale w każdym przypadku może się okazać, że część u^2 wielkości wejściowych \underline{u} jest dla sterowania optymalnego nieistotna - inaczej mówiąc, rozwiązanie optymalne dopuszcza dowolność wartości niektórych wielkości wejściowych. Ma to oczywiście duże znaczenie praktyczne, gdyż pozwala - bez utraty optymalności - zredukować liczbę wielkości wejściowych podlegających sterowaniu. Fakt nieistotności wielkości wejściowych \underline{u}^2 oraz istotności pozostałych \underline{u}^1 możemy zapisać jako równość

$$\max Q = \max Q \quad /76/$$

czyli
$$Q(\hat{u}^1, \underline{u}^2) = \hat{Q} \quad /76'/$$

gdzie \underline{u}^2 może przyjmować wartości dowolne.

Istotne i nieistotne współrzędne stanu

Podobnie jak to ma miejsce z wielkościami wejściowymi może się zdarzyć, że rozwiązanie określające optymalną trajektorię $\hat{x}(t)$ pozostawia dowolność przebiegu niektórych współrzędnych stanu. Będą to "nieistotne" współrzędne stanu /nieistotne nie dla obiektu w ogóle, lecz dla danego zadania optymalizacji w tym obiekcie/. Prosty przykład takiej sytuacji jest, że temperatura cieczy w zbiorniku jest nieistotna dla procesów regulacji poziomu; jest ona natomiast istotną współrzędną stanu, gdy np. zadaniem zbiornika jest mieszanie cieczy ciepłej z zimną.

Fakt nieistotności części współrzędnych stanu \underline{x}^2 oraz istotności współrzędnych pozostałych \underline{x}^1 można zapisać jako równość

$$Q(\hat{x}^1, \underline{x}^2, \hat{u}^1, \underline{u}^2) = \hat{Q} \quad /77/$$

gdzie \underline{x}^2 może przyjmować wartości dowolne, podobnie jak \underline{u}^2 .

Znaczenie praktyczne istotności współrzędnych stanu występuje w tych licznych przypadkach, gdy struktura układu sterowania zawiera sprzężenia zwrotne. Jest bowiem oczywiste, że

nieistotnych współrzędnych stanu nie trzeba mierzyć. Jeżeli sprzężenia zwrotne nie opierają się na współrzędnych stanu, lecz na wielkościach wyjściowych \underline{y} , to można analogicznie mówić o istotnych i nieistotnych wielkościach wyjściowych \underline{y}^1 oraz \underline{y}^2 , jeśli zachodzi równość

$$Q(\hat{\underline{y}}^1, \underline{y}^2, \hat{\underline{u}}^1, \underline{u}^2) = \hat{Q} \quad /78/$$

Nieistotność części współrzędnych stanu stanowi też ważne uproszczenie obliczeniowe - można dane zadanie optymalizacji rozwiązywać w przyszłości przyjmując rząd obiektu odpowiednio niższy.

Wrażliwość sterowania optymalnego

Obok "nieistotnych" wielkości wejściowych, można również mówić o "mało istotnych" elementach zbioru \underline{u} - będą to te elementy, których pominięcie w sterowaniu nieznacznie tylko pogarsza osiąganą wartość Q . Chodzi tu o wrażliwość sterowania optymalnego na pominięcie elementu u^1 wśród wielkości wejściowych sterujących. Efekt ilościowy wyrazi się różnicą

$$\Delta Q_1 = \max_{\underline{u}} Q - \max_{\underline{u}^1} Q \quad /79/$$

gdzie \underline{u}^1 jest zbiorem \underline{u} bez elementu u_1 . Wartość /79/ można konfrontować z obniżką kosztu sterowania, wywołaną pominięciem elementu u_1 .

Badać można wiele innych rodzajów wrażliwości. Między innymi rozpatrywać możemy wrażliwość sterowania optymalnego na pomijanie współrzędnych stanu w wyznaczonej trajektorii optymalnej $\hat{\underline{x}}(t)$. Odniesiemy to do układu sterowania podanego na rys.1, gdzie wyznacza się trajektorię stanu $\hat{\underline{x}}$, za którą nadąża stan obiektu. Gdy \underline{x} jest pełnym wektorem stanu, badamy czy można opuścić element x_1 w tym wektorze. Interesuje nas zatem różnica

$$\Delta Q_1 = \max_{\underline{x}} Q - \max_{\underline{x}^1} Q \quad /80/$$

gdzie \underline{x}^1 - zbiór \underline{x} bez elementu x_1 , a Q jest wartością wskaźnika jakości osiąganego w układzie z rys.1. Strata /80/ może być konfrontowana z odpowiednią obniżką kosztu realizacji sterowania.

4.2. Przykład reaktora o idealnym zmieszaniu

Rozpatrzmy zadanie optymalizacji stanu ustalonego, sformułowane dla opisywanego już reaktora zależnościami /19/, /20/, /21/:

$$\max (1 - C_A) F_D \quad \text{kg/s} \quad /19/$$

$$F_B \leq F_{Bm} \quad /20/$$

$$C_A \leq C_{Am} \quad /21/$$

dla obiektu o modelu matematycznym, opisanym przez /52/, /53/, /54/

$$\frac{dW}{dt} = F_A + F_B - F_D \quad /52/$$

$$\frac{dC_A}{dt} = \frac{1 - C_A}{W} F_A - \frac{C_A}{W} F_B - C_A k(T) \quad /53/$$

$$\frac{dT}{dt} = \frac{T_A - T}{W} F_A + \frac{T_B - T}{W} F_B + \frac{H}{cW} - \frac{hC_A k(T)}{c} \quad /54/$$

z ograniczeniami $W \leq W_m$, $T \leq T_m$, $H \geq 0$.

Zadanie stanu ustalonego sprowadza się do wymagania

$$\frac{dW}{dt} = 0, \quad \frac{dC_A}{dt} = 0, \quad \frac{dT}{dt} = 0$$

i w rezultacie możemy nasze zadanie optymalizacji przepisać jako zadanie programowania nieliniowego:

$$\max (1 - C_A) F_D \quad /81/$$

$$F_A, F_B, F_D, H, T_A, T_B$$

$$\text{przy } F_B \leq F_{Bm}, \quad C_A \leq C_{Am}, \quad W \leq W_m, \quad T \leq T_m, \quad H \geq 0 \quad /82/$$

oraz

$$F_A + F_B - F_D = 0 \quad /83/$$

$$\frac{1 - C_A}{W} F_A - \frac{C_A}{W} F_B - C_A k(T) = 0 \quad /84/$$

$$\frac{T_A - T}{W} F_A + \frac{T_B - T}{W} F_B + \frac{H}{cW} - \frac{hC_A k(T)}{c} = 0 \quad /85/$$

Rozwiązanie tego problemu jest łatwe i brzmi

$$\hat{F}_A = C_{Am} (1 - C_{Am})^{-1} [W_m k(T_m) + F_{Bm}] \quad /86/$$

$$\hat{F}_B = F_{Bm} \quad /87/$$

$$\hat{F}_D = (1 - C_{Am})^{-1} [W_m C_{Am} k(T_m) + F_{Bm}] \quad /88/$$

$$cF_A \hat{T}_A + cF_{Bm} \hat{T}_B + \hat{H} = hW_m C_{Am} k(T_m) + c \hat{F}_D T_m \quad /89/$$

Zwróćmy uwagę na następujące cechy rozwiązań /86/-/89/.

a/ Wartości \hat{F}_A , \hat{F}_D wymagają do ich wyznaczenia znajomości stałej prędkości reakcji $k(T_m)$, co oznacza potrzebę wiernego opisu procesu.

b/ Wartości \hat{T}_A , \hat{T}_B , \hat{H} nie są wyznaczane oddzielnie, lecz w postaci jednego warunku /89/; oznacza to, że dla danych \hat{T}_A , \hat{T}_B można podać takie \hat{H} , które zapewni pełną optymalność procesu. Spośród zatem trzech wielkości wejściowych T_A , T_B , H jedna tylko jest istotna, to znaczy wystarczy jedną sterować - w zasadzie dowolną, praktycznie oczywiście dopływem ciepła H . Zwróćmy uwagę, że ograniczenie $H \geq 0$ wprowadza pewne ograniczenie zakresu dowolności T_A , T_B , nie zmieniając istoty sprawy.

4.3. Przykład reaktora rurowego

Rozpatrzmy zadanie optymalizacji stanu ustalonego, sformułowane dla hipotetycznego reaktora opisanego w punkcie 3.3. Przyjmijmy, że zadanie optymalizacji brzmi /patrz rys.6/

$$\max F_B(L) \quad /90/$$

przy ograniczeniu

$$F_B(0) \leq F_{Bm} \quad /91/$$

$$\frac{F_A(0)}{F_A(L)} \leq a_m \quad /92/$$

gdzie F_A , F_B są natężeniami przepływu w kg/s.

Model matematyczny reaktora był już określony - mamy równania stanu

$$\frac{\partial x_1}{\partial t} = v_1 \frac{\partial x_1}{\partial l} - k(T) x_1(1, t) \quad /63/$$

$$\frac{\partial x_2}{\partial t} = -v_2 \frac{\partial x_2}{\partial l} + k(T) x_1(1, t) \quad /65/$$

oraz warunki brzegowe

$$x_1(1, 0), \quad x_2(1, 0) \quad - \text{stan początkowy}$$

$$x_1(L, t) = \frac{F_A(L, t)}{v_1}, \quad x_2(0, t) = \frac{F_B(0, t)}{v_2}$$

- warunki zależne od wielkości wejściowych

Rozwiązanie zadania optymalizacji stanu ustalonego przebiega jak następuje. Dla stanu ustalonego mamy z równań stanu równania różniczkowe zwyczajne:

$$v_1 \frac{dx_1}{dl} - k(T) x_1(l) = 0 \quad /93/$$

$$-v_2 \frac{dx_2}{dl} + k(T) x_1(l) = 0 \quad /94/$$

które należy rozwiązać przy ograniczeniach

$$x_1 + x_2 \leq x_m \quad /95/$$

$$T \leq T_m \quad /96/$$

oraz warunkach brzegowych danych przez sterowanie

$$x_1(L) = \frac{F_A(L)}{v_1} \quad /97/$$

$$x_2(0) = \frac{F_B(0)}{v_2} \quad /98/$$

Łatwo przewidzieć, że rozwiązanie zadania $\max F_B(L)$ przy ograniczonym dopływie $F_B(0) \leq F_{Bm}$ wypadnie na brzegu ograniczeń

$$T = T_m$$

$$F_B(0) = F_{Bm}$$

Przy $T = T_m = \text{const}$ równania stanu ustalonego /93 /94/ dają się łatwo rozwiązać, dając

$$x_1(l) = \frac{F_A(L)}{v_1} \exp \left[- \frac{k(T_m)}{v_1} (L - l) \right] \quad /99/$$

oraz

$$x_2(l) = \frac{F_A(L)}{v_2} \exp \left[- \frac{k(T_m)}{v_1} (L - l) \right] - \frac{F_A(L)}{v_2} \exp \frac{-k(T_m)}{v_1} L + \frac{F_B(0)}{v_2} \quad /100/$$

Przy tych rozwiązaniach, wielkość $F_B(L) = v_2 x_2(L)$, którą maksymalizujemy, zależy od wielkości wejściowych według wzoru

$$F_B(L) = F_A(L) \left[1 - \exp\left(-\frac{k(T_m)}{v_1} L\right) \right] + F_B(0) \quad /101/$$

Maksymalizacji /101/ mają towarzyszyć ograniczenia /91/ /92/, które przy uwzględnieniu rozwiązań /99/ /100/ przyjmują postać

$$\exp\left(-\frac{k(T_m)}{v_1} L\right) \leq a_m \quad /102/$$

$$x_1(L) + x_2(L) = \frac{F_A(L)}{v_1} + \frac{F_B(L)}{v_2} \leq x_m \quad /103/$$

bowiem łatwo stwierdzić, że maximum $x_1(l)$ oraz maximum $x_2(l)$ przypada na końcu reaktora, $l = L$. Ponieważ celem jest $\max F_B(L)$, ograniczenie /103/ będzie w rozwiązaniu spełnione równościowo, dając wzór

$$F_B(L) = v_2 \left[x_m - \frac{F_A(L)}{v_1} \right] \quad /104/$$

Zależność ta pokazuje, że $F_B(L)$ rośnie gdy rośnie prędkość v_1 . Ale prędkość v_1 jest ograniczona przez nierówność /102/. Rozwiązując tę nierówność względem v_1 , otrzymamy

$$v_1 \leq \frac{k(T_m) L}{-\ln a_m} \quad /105/$$

czyli rozwiązanie na prędkość optymalną

$$\hat{v}_1 = \frac{k(T_m) L}{-\ln a_m} \quad /106/$$

Używając łącznie /101/ oraz /104/ otrzymamy rozwiązanie na $\hat{F}_A(L)$:

$$\hat{F}_A(L) = \frac{\hat{v}_1 (v_2 x_m - F_{Bm})}{v_2 + \hat{v}_1 (1 - a_m)} \quad /107/$$

Przy $v_1 = \hat{v}_1$ oraz $F_A(L) = \hat{F}_A(L)$ otrzymamy rezultat optymalizacji

$$\hat{F}_B(L) = \hat{F}_A(L) (1 - a_m) + F_{Bm}$$

czyli

$$\hat{F}_B(L) = \frac{F_{Bm} + x_m \hat{v}_1 (1 - a_m)}{1 + \frac{\hat{v}_1}{v_2} (1 - a_m)} \quad /108/$$

Wielkość ta roslaby z wartością v_2 . Przyjmujemy, że prędkość v_2 nie podlega sterowaniu /jest narzucona z góry/.

5. STRUKTURA OGÓLNA UKŁADU STEROWANIA

5.1. Kontynuacja przykładów

Rozpatrujemy dalej przykład z reaktorem, omawiany ostatnio w punkcie 4.2. Zadanie optymalizacji stanu ustalonego w obiekcie opisanym równaniami /52/ /53/ /54/ możemy przedstawić jako złożone z dwóch części:

- a/ należy zapewnić stałość współrzędnych stanu, tj. $W = \text{const}$,
 $C_A = \text{const}$, $T = \text{const}$,
- b/ ustalone wartości W , C_A , T oraz wartości wielkości sterujących powinny być optymalne.

Punkt a/ można określić jako stabilizację procesu w reaktorze, a punkt b/ - jako optymalizację.

Układ sterowania, który zapewniłby stałość np. współrzędnej stanu W musi zapewnić $\frac{dW}{dt} = 0$. Można to zrobić, w oparciu o równanie /83/, przez takie sterowanie wielkością np. F_D aby otrzymać

$$F_A + F_B - F_D = 0, \quad \text{czyli } F_D = F_A + F_B$$

Sterowanie to wymagałoby ustalenia F_B w oparciu o pomiar $F_A + F_B$; byłoby to sterowanie w układzie otwartym. Błąd w pomiarze $F_A + F_B$ spowoduje $F_A + F_B - F_D \neq 0$, zatem wywoła $\frac{dW}{dt} \neq 0$ i w konsekwencji zmianę stanu zapełnienia reaktora.

Stałość współrzędnej stanu W znacznie skuteczniej zapewnić można w zamkniętym układzie sterowania, spełniającym bezpośrednio wymaganie stabilizacji $W = \text{const}$. Pokazano to na rys. 7, podając zarazem dwa pozostałe układy stabilizacji, dla współrzędnych stanu C_A oraz T . Do ustalenia struktury układu sterowania należy również wybór, jaką z wielkości wejściowych powiązać z regulatorem danej współrzędnej stanu. Przyjęliśmy na rys.7 przykładowo, że regulator zapełnienia W wpływa na dopływ F_D , regulator stężenia C_A wpływa na dopływ F_A , regulator temperatury - na dopływ ciepła H . Nie była to jedyna możliwość wyboru. Kształtowanie układów stabilizacji omawiamy bliżej w podrozdziale 6; tu zwróćmy tylko uwagę, że dla realizacji stabilizacji stanu potrzeba przynajmniej tyle wielkości sterujących, ile jest współrzędnych stanu. Muszą to ponadto być wielkości, spełniające warunki sterowalności obiektu - w tym przypadku, sterowalności odniesionej do utrzymania zadanego stanu. Problem sterowalności omawiamy nieco szerzej w punkcie 6.1.

Z punktu widzenia optymalizacji stanu ustalonego, mamy - po wprowadzeniu trzech regulatorów stabilizujących - następujące zadanie:

$$\max_{W_o, T_o, C_{Ao}, F_B, T_A, T_B} (1 - C_A) F_D \quad /109/$$

to znaczy to samo zadanie co /81/, lecz względem innych zmiennych decyzyjnych: $W_o, T_o, C_{Ao}, F_B, T_A, T_B$. Zadanie /109/ będzie rozwiązywane z uwzględnieniem tych samych ograniczeń co poprzednio, tj. /82/, /83/, /84/, /85/. Rozwiązania zadania /109/ brzmią /przy założeniu, że $k(T)$ rośnie z temperaturą/:

$$\hat{W}_o = W_m \quad /110/$$

$$\hat{C}_{Ao} = C_{Am} \quad /111/$$

$$\hat{T}_o = T_m \quad /112/$$

$$\hat{F}_B = F_{Bm} \quad /113/$$

$$\hat{T}_A, \hat{T}_B = \text{dowolne} \quad /114/$$

Pierwsze trzy z tych liczb określają optymalne wartości zadane, jakie należy ustawić w regulatorach stabilizujących. Czwarta liczba podaje optymalną wartość dopływu F_B /dopływ ten byłby w praktyce stabilizowany regulatorem przepływu, co uwzględniono w rys.7/. Wartości T_A, T_B są dowolne, gdyż $T = \text{const}$ gwarantuje stałość bilansu cieplnego, a $T = \text{const} = T_m$ zapewnia optymalność tego stanu ustalonego.

Gdy sobie wyobrazimy, że wartości zadane obliczone jako /110/ /111/ /112/ są wyznaczane bieżąco w czasie trwania procesu, układ sterowania reaktorem staje się układem dwuwarstwowym /rys.7/. Warstwa pierwsza zawiera regulatory stabilizujące, wpływające bezpośrednio na proces, warstwa druga zawiera urządzenie wyznaczające optymalne wartości zadane dla regulatorów.

Wróćmy teraz na chwilę do rozwiązań tego samego zadania, danych przez /86/ - /89/. Bezpośrednia realizacja tych rozwiązań dałaby układ otwarty sterowania optymalnego. Układ z rys.7 ma w porównaniu z nim szereg istotnych zalet:

- stabilizacja stanu reaktora jest osiągnięta niezależnie od dokładnej znajomości jego równań;
- stan optymalny wyznaczony jest niezależnie od stabilizacji, zatem błędy optymalizatora nie wywołują wytrącenia reaktora z równowagi;

- do wyznaczenia stanu optymalnego, w tym przykładzie, można nie znać wartości stałej prędkości reakcji $k(T_m)$ - wystarczająca informacja, że rośnie ona z temperaturą;
- w przypadku niewielkich, a szybkozmiennych zakłóceń, można zrezygnować z interwencji optymalizatora: układy stabilizacji zapewnią dostrojenie wielkości sterujących do stanu stałości współrzędnych stanu, przy niewielkiej utracie optymalności. Optymalizator może zatem działać okresowo.

Wymienione wyżej i inne jeszcze cechy powodują, że struktury wielowarstwowe znajdują powszechne zastosowanie w układach sterowania złożonymi procesami.

Rozpatrzmy jeszcze przykład wyboru struktury układu sterowania dla reaktora o stałych rozłożonych z rys.6, którego model był omówiony w punkcie 3.3, a rozwiązanie zadania optymalizacji - w punkcie 4.3.

Jeśli, podobnie jak w poprzednim przykładzie, przedstawić zadanie optymalizacji stanu ustalonego jako:

a/ zadanie stabilizacji stanu,

b/ zadanie optymalizacji

to rozpatrując punkt a/ zauważymy od razu pewną różnicę. Stan ustalony reaktora nie jest teraz określony przez liczby x_1, x_2, \dots , lecz przez funkcje zmiennej przestrzennej, $x_1(l)$ oraz $x_2(l)$. Ścisła stabilizacja tego stanu przez układy regulacji nie jest możliwa - trzeba by np. mierzyć x_1 w każdym punkcie l i użyć wielkiej liczby regulatorów. Jest to tym bardziej abstrakcyjne, że nie mamy dla reaktora z rys.6 odpowiedniej liczby wielkości wejściowych: nie mamy możliwości wpływać na współrzędną $x_1(l)$ lub $x_2(l)$ inaczej, jak tylko przez dopływy na krańcach reaktora, bądź przez prędkość transportu v_1 - również scharakteryzowaną jedną liczbą, a nie rozkładem $v_1(l)$.

Musimy więc zrezygnować ze ścisłej stabilizacji stanu wzdłuż długości reaktora.

Rozpatrzmy zatem sprawę wyboru struktury układu stabilizacji od innego końca: dla reaktora z rys.6 mamy trzy wielkości sterujące, mianowicie dopływ materiału $F_A(L)$, dopływ $F_B(L)$ oraz prędkość v_1 . Pominęliśmy prędkość v_2 , którą traktujemy jako narzuconą z góry, oraz dopływ ciepła, o którym uważamy, że jest taki by zapewnić na całej długości optymalną temperaturę $T(l) = T_m$. Przy użyciu tych trzech wielkości sterujących możemy stabilizować trzy wartości współrzędnych stanu reaktora-

ra, na przykład $x_1(L)$, $x_2(L)$, $x_2(0)$. Wytyczne wyboru tych czy innych wartości weźmiemy z rozwiązania zadania optymalizacji, punkt 4.3. Mianowicie dogodnie jest przyjąć za wielkości regulowane te zmienne, które w rozwiązaniu zadania optymalizacji leżą na ograniczeniach. Wówczas bowiem wartości zadane dla odpowiednich regulatorów wynikają wprost z ustaleń ograniczających. Przypomnijmy, że w rozwiązaniu opisanym w punkcie 4.3. aktywne okazały się ograniczenia:

$$F_B(0) \leq F_{Bm}$$

$$x_1(L) + x_2(L) \leq x_m$$

$$\frac{F_A(0)}{F_A(L)} \leq a_m$$

Sugeruje to by przyjąć za wielkości regulowane, czyli za wielkości wyjściowe obiektu:

$$y_1 = x_1(L) + x_2(L) \quad /115/$$

$$y_2 = \frac{F_A(0)}{F_A(L)} = \frac{v_1 x_1(0)}{F_A(L)} \quad /116/$$

oraz za trzecią wielkość regulowaną natężenie dopływu $F_B(0)$.

Pozostaje ustalenie wielkości sterujących, które byłyby powiązane z regulatorami wielkości y_1 , y_2 . Przyjmiemy:

$$u_1 = F_A(L) \quad \text{dla regulatora wielkości wyjściowej } y_1$$

$$u_2 = v_1 \quad \text{dla regulatora wielkości wyjściowej } y_2$$

Całość struktury układu stabilizacji przedstawia schematycznie rys.8.

Drugą warstwą układu sterowania byłaby warstwa optymalizacji. Jej zadanie polega na ustaleniu optymalnych wartości zadanych dla regulatorów. Zgodnie z rozwiązaniem zadania optymalizacji /punkt 4.3./, wartości optymalne są równe:

$$\hat{y}_{1d} = x_m$$

$$\hat{y}_{2d} = a_m$$

$$[F_B(0)]_d = F_{Bm}$$

5.2. Wielowarstwowe układy sterowania

Rysunek 7 sugerował bardzo praktyczny i mający wiele zalet "dwuwarstwowy" układ sterowania reaktorem, którego istotą było rozdzielenie funkcji stabilizacji procesu od funkcji je-

go optymalizacji. Rozdzielenie funkcji można prowadzić i dalej - jeśli przyjąć, że czynności optymalizacji wymagają znajomości modelu procesu, oraz że model ten trzeba aktualizować, to nazywając tę aktualizację modelu adaptacją możemy jej wykonywanie przypisać następnej "warstwie" układu sterowania, jak to pokazuje rys.9. Istotne jest teraz zwrócić uwagę, że o ile układ stabilizacji musi wpływać na obiekt /na przykład rozpatrywany reaktor/ w sposób ciągły w czasie, lub bardzo często, to już wyznaczenie wartości zadanych /funkcje warstwy optymalizacji/ może się odbywać stosunkowo rzadziej - np. co okres $T_2 > T_1$ /patrz rys.9/. Jeszcze rzadziej mogą się okazać potrzebne interwencje adaptacji, czyli zmiany modelu użytego w optymalizacji: otrzymamy uporządkowanie częstości wykonywania swych funkcji przez poszczególne warstwy sterowania

$$T_3 > T_2 > T_1, \quad \text{lub} \quad \frac{1}{T_3} < \frac{1}{T_2} < \frac{1}{T_1} \quad /117/$$

Częstości te mają bezpośrednie znaczenie ekonomiczne jeśli przyjąć, że wszystkie funkcje, lub też co najmniej funkcje optymalizacji i adaptacji, będą wykonywane przez cyfrową maszynę sterującą: koszt realizacji każdej z funkcji będzie maleł z częstotliwością powtarzania danego obliczenia.

Rzecz jasna, zwiększanie każdego z okresów T_1, T_2, T_3 w układzie z rys.9 musi wywołać obniżenie osiąganego wartości wskaźnika jakości, w stosunku do wartości jaka byłaby osiągnięta przy całkowicie optymalnym, ciągłym sterowaniu danym obiektem:

$$\max_{\underline{u}} Q \geq \max_{\underline{\gamma}} Q \quad /118/$$

gdzie $\underline{\gamma}$ jest wektorem parametrów, zadanych warstwie adaptacji - zwróćmy uwagę, że na cały układ z rys.9 można patrzeć jak na obiekt sterowania, dla którego wielkością sterującą zewnętrzną jest właśnie wektor parametrów $\underline{\gamma}$.

Oczywiście, różnica wskazana nierównością /118/ zależy nie tylko od okresów T_1, T_2, T_3 , ale także od nieidealności czy uproszczeń w algorytmach stabilizacji, optymalizacji i adaptacji. Nazwijmy wybór wartości okresów T_1, T_2, T_3 oraz wybór struktur algorytmów stabilizacji, algorytmów optymalizacji i algorytmów adaptacji - decyzjami strukturalnymi. Możemy wówczas sformułować następujące zadanie optymalizacji układu ste-

rowania, w odróżnieniu od optymalizacji sterowania:

max [Q - C]
decyzje
strukturalne

/119/

We wzorze /119/ Q oznacza wskaźnik jakości obiektu sterowania, osiągany w układzie z rys.9, natomiast C oznacza sprovedzony do tych samych jednostek /np. zł/godz./ koszt realizacji układu sterowania - na przykład maszyny cyfrowej wraz z urządzeniami powiązania z obiektem. Maksymalizacja różnicy Q - C odbywa się względem "decyzji strukturalnych" - to znaczy trzeba tak wybrać poszczególne algorytmy i okresy ich powtarzania, aby Q - C osiągnęło maximum.

Problem postawiony w postaci /119/ jest trudny do bezpośredniego rozwiązywania, gdyż trudno znaleźć potrzebne do jego rozwiązania zależności Q = f /decyzje strukturalne oraz C = f /decyzje strukturalne/. Istnieją natomiast możliwości ustalenia pewnej metodyki postępowania, którą przedstawiamy w punkcie 5.3.

Istnieje wiele różnych możliwości i przykładów tworzenia wielowarstwowego układu sterowania. W tym miejscu omówimy jeszcze przypadek charakterystyczny przez to, że w warstwie najwyższej pojawi się problem optymalizacji dynamicznej - w schematach z rys.7 i rys.9 sugerowano, że w warstwach powyżej pierwszej występuje optymalizacja statyczna.

Rozpatrzmy następujący hipotetyczny przykład automatyzacji kompleksowej procesu produkcyjnego. Dana jest elektrownia cieplna, sprzężona z urządzeniem odsalania wody morskiej. Należy osiągnąć minimum kosztu produkcji energii elektrycznej i wody za okres $t_2 - t_1$, przy spełnieniu ograniczeń określonych przez moc obciążenia W_L (MW) oraz zapotrzebowania na wodę $F_D(t/h)$. Jeśli natężenie strumienia kosztów oznaczyć przez Q /składają się na to koszty paliwa, energii elektrycznej, potrzeb własnych itd./, zadanie brzmi

$$\min \int_{t_1}^{t_2} Q dt \quad /120$$

przy ograniczeniach:

moc produkowana równa obciążeniu $W = W_L$ /121/

ilość wody równa zapotrzebowaniu mniej dopuszczalna zmiana zasobu

$$\int_{t_1}^{t_2} F dt = \int_{t_1}^{t_2} F_D dt - h(t_1) + h(t_2) \quad /122/$$

Zmiennymi decyzyjnymi są w zasadzie położenia wszystkich zaworów regulacyjnych w obu częściach instalacji, \underline{u} .

Zmienne te, wraz z zakłóceniami \underline{z} , wpływają na współczesne stanu obiektu \underline{x} , co możemy wyrazić zapisem zadania w postaci:

$$\min_{\underline{u}} \int_{t_1}^{t_2} Q(\underline{u}, \underline{x}, \underline{z}, t) dt \quad /120'/$$

przy

$$W(\underline{u}, \underline{x}, \underline{z}) = W_L(t) \quad /121'/$$

$$\int_{t_1}^{t_2} F(\underline{u}, \underline{x}, \underline{z}) = \int_{t_1}^{t_2} F_D(t) dt - h(t_1) + h(t_2) \quad /122'/$$

Pojawienie się argumentu t w natężeniu kosztów Q związane jest z tym, że niektóre ceny /np. na energię elektryczną/ mogą być zmienne w czasie. Zwróćmy uwagę, że

- postać funkcji $Q(\underline{u}, \underline{x}, \underline{z}, t)$ wynika z danych "ekonomicznych" i jest z reguły jawna

- postaci funkcji $W(\underline{u}, \underline{x}, \underline{z})$ oraz $F(\underline{u}, \underline{x}, \underline{z})$ wynikają z równań procesu; uzyskanie tych zależności w postaci jawnej zwykle nie jest możliwe.

Teoretycznie rzecz biorąc, można wykonać /120/, znajdując $\hat{\underline{u}}(t)$. Przy dużej wymiarowości \underline{u} , \underline{x} oraz złożonych równaniach obiektu rozwiązania takiego uzyskać jednak nie można. Jedną z koncepcji uproszczenia zadania, zarazem koncepcją dogodną w realizacji technicznej sterowania, jest zastosowanie struktury wielowarstwowej.

W wielowarstwowym układzie sterowania elektrownia oraz instalacja odsalania wody miałyby szereg układów regulacji stałowartościowej, wpływających na odpowiednie zawory i utrzymujących stan obu procesów $\underline{x}^1, \underline{x}^2$ zadanych wartościach $\underline{x}_d^1, \underline{x}_d^2$ /rys.10/. Z kolei, urządzenia następnej "warstwy" określałyby optymalne wartości $\underline{x}_d^1, \underline{x}_d^2$. W warstwie tej optymalizator elektrowni minimalizuje odpowiednie natężenie obiektu Q_1 uwzględniając zadanie W_L , a optymalizator odsalania wody minimalizuje Q_2 , spełniając zadane natężenie produkcji F . Natężenie prze-

plywu ciepła H od elektrowni do instalacji odsalania nie może być ustalone optymalnie przez którykolwiek z tych optymalizatorów. Potrzebna jest dodatkowa optymalizacja, $\min (Q_1 + Q_2)$ względem H ; daje to łącznie dwupoziomowy układ optymalizacji. Dla układu tego zadane są W_L oraz F , rozwiązaniami są x_d^1 oraz x_d^2 . Przyjmujemy, że zmienność w czasie tych wszystkich zmiennych jest na tyle powolna, by całość omówionego zadania optymalizacji miała charakter optymalizacji statycznej.

Pozostaje ustalenie wartości natężenia produkcji wody F . Ponieważ w systemie istnieje zbiornik wody, możliwe są rozwiązania $F(t) \neq F_D(t)$; powstaje tu zadanie optymalizacji dynamicznej typu planowania produkcji.

$$\min_{F(t)} \int_{t_1}^{t_2} (Q_1 + Q_2) dt \quad /123/$$

przy danych $W_L(t)$, $F_D(t)$ oraz początkowym i końcowym stanie zasobu $h(t_1)$, $h(t_2)$.

Otrzymuje się rozwiązanie $\hat{F}(t)$ /wymaga ono w ogólnym przypadku predykcji $F_D(t)$ oraz $W_L(t)$ na okres $(t_2 - t_1)$; chwilowa wartość \hat{F} stanowi zadanie /typu ograniczenia równościowego/ dla omówionej wyżej optymalizacji statycznej.

Wszystkie użyte przy optymalizacji modele obiektu musiałyby zapewne być adaptacyjne, co nie zostało zaznaczone na rysunku. Zauważmy, że problem dynamiczny przypisany najwyższej warstwie układu sterowania może wymagać innych jeszcze rodzajów adaptacji; na przykład, można zmieniać stan docelowy zasobu $h(t_2)$ czy też różnicę $h(t_2) - h(t_1)$, w kolejnych okresach, według pewnej taktyki długofalowej. Gdyby, na przykład $t_2 - t_1$ wynosiło 24 godziny, możnaby stopniowo obniżać wymagane $h(t_2)$ od poniedziałku do piątku, a następnie uzupełnić bilans wody zwiększoną produkcją w sobotę i niedzielę.

W przedstawionym przykładzie wystąpiło kilka zagadnień bardzo charakterystycznych dla automatyzacji złożonych procesów:

- niemożność rozwiązania zadania optymalizacji "wprost",
- koncepcja uproszczenia zadania optymalizacji drogą wprowadzenia struktury warstwowej,
- konieczność używania uproszczonych modeli, np. modeli statycznych dla obiektów w istocie dynamicznych,
- możliwość wielopoziomowego rozwiązania zadań optymalizacji,

- potrzeba opanowania praktycznych algorytmów wyznaczania rozwiązań zadań optymalizacji.

5.3. Metodyka projektowania układu wielowarstwowego

Rozpatrujemy przypadek, gdy sterowanie obiektu o wektorze stanu \underline{x} , wejściu zakłócającym \underline{z} i wejściu nastawiającym \underline{u} jest dokonywane w układzie wielowarstwowym, jak to przedstawia schematycznie rys.9.

W warstwie pierwszej /warstwa "sterowania bezpośredniego" lub "stabilizacji"/ odbywa się ustalanie wejść nastawiających \underline{u} tak, by stan \underline{x} był zgodny z trajektorią zadaną \underline{x}_d . Na rysunku 9 przedstawiono to jako prosty układ sprzężenia zwrotnego; możliwe są oczywiście struktury bardziej złożone. Układ sterowania pierwszej warstwy ma stałą strukturę oraz nastawialny wektor parametrów $\underline{\alpha}$. Wektor ten może być zmieniany w trakcie działania adaptacyjnego /w zasadzie dokonywanego w warstwie trzeciej/, w pewnych regularnych lub nieregularnych odstępach T_3^α .

Warstwa druga /warstwa "optymalizacji" dokonuje wyznaczenia trajektorii zadanej \underline{x}_d . Przyjmuje się, że jest to dokonywane w oparciu o pewną reprezentację z^2 zakłócenia \underline{z} , oraz przy użyciu algorytmu o stałej strukturze i wektorze nastawialnych parametrów $\underline{\beta}$. Wyznaczanie zadanej trajektorii \underline{x}_d odbywa się w regularnych lub nieregularnych odstępach T_2 .

Warstwa trzecia /warstwa "adaptacji"/ zmienia wektor parametrów $\underline{\beta}$ w regularnych lub nieregularnych odstępach T_3^β oraz wektor $\underline{\alpha}$ w odstępach T_3^α . Adaptacje te są oparte na reprezentacjach z^3 względnie z^1 zakłócenia \underline{z} . Działania adaptacyjne odbywają się według algorytmów o stałej strukturze, z nastawialnym wektorem parametrów $\underline{\gamma}$.

Zakłada się, że wektor $\underline{\gamma}$ pozostaje stały przez cały czas działania układu. Zauważyliśmy już, że jeśli potraktować obiekt i układ sterowania jako jedną całość, to jest tu tylko jedna wielkość sterująca; wektor $\underline{\gamma}$. Osiągnięta w wyniku optymalizacji względem $\underline{\gamma}$ wartość wskaźnika jakości zależy jednak bardzo znacznie od podjętych decyzji strukturalnych: struktur algorytmów sterowania, budowy reprezentacji zakłóceń z^1 , z^2 , z^3 , odstępów optymalizacji i adaptacji T_2 , T_3^α , T_3^β . W ten sposób, "decyzje strukturalne" stanowią również wielkość wejściową układu, jak pokazano symbolicznie na rys.9.

Uwaga nasza będzie skupiona na projektowaniu i optymalizacji tego układu sterowania /w odróżnieniu od optymalizacji sterowania u/. Istotną cechą naszych rozważań będzie rozpatrywanie kompromisu pomiędzy kosztem wykonywania sterowania a stratą na optymalności. Ponieważ bezpośrednia optymalizacja wskaźnika jakości

$$\max E [Q(\underline{x}, \underline{u}, \underline{z}, \underline{x}_2, t_2)] \quad /124/$$

względem decyzji strukturalnych nie jest, jak już podkreślaliśmy, możliwa, musimy dla celów praktycznych znaleźć metodę określania tych decyzji. Nawiązując do opisu układu z rys.9 decyzje strukturalne obejmują:

Budowę reprezentacji zakłóceń

$$z^1, z^2, z^3$$

Strukturę algorytmu sterowania bezpośredniego

$$\underline{u} = \underline{u}(\underline{\alpha}, \underline{x}_d, \underline{x})$$

Strukturę algorytmu optymalizacji

$$\underline{x}_d = \underline{x}_d(\underline{\beta}, z^2)$$

Struktury algorytmów adaptacji

$$\underline{\alpha} = \underline{\alpha}(\underline{\gamma}, z^1)$$

$$\underline{\beta} = \underline{\beta}(\underline{\gamma}, z^3)$$

Odstępy próbkowania

$$T_2, T_3^\alpha, T_3^\beta$$

Warto w tym miejscu zrobić dwie uwagi. Po pierwsze, decyzje odnośnie wartości $T_2, T_3^\alpha, T_3^\beta$ można zastąpić specyfikacją liczby przedziałów optymalizacji względnie adaptacji w całkowitym przedziale sterowania (t_1, t_2) , tym samym pozwalając na zmienność $T_2, T_3^\alpha, T_3^\beta$. Zmienne $T_2, T_3^\alpha, T_3^\beta$ byłyby praktycznie dopuszczalne np. w sytuacji, gdy maszyna sterująca jest używana "part-time", a obliczenia sterowania mają w niej priorytet. Po drugie, całość decyzji "strukturalnych" możnaby uważać za jeszcze jedno wejście do układu, nazwać je "organizacją", a ewentualną automatyczną realizację przypisać następnej warstwie: warstwie "samoorganizacji".

Jednym z możliwych sposobów kolejnego podjęcia decyzji strukturalnych jest rozpoczęcie projektowania układu od decyzji dotyczących najbardziej bezpośrednio tylko warstwy pierwszej:

struktura algorytmu sterowania bezpośredniego

$$\underline{u} = \underline{u}(\underline{\alpha}, \underline{x}_d, \underline{x}) \quad /125/$$

budowa reprezentacji zakłóceń z^1, z^2

odstęp próbkowania adaptacji T_3^α

Dążymy zatem do określenia optymalnego $\underline{\alpha}$ /jako funkcji czasu kT_3^α oraz zakłócenia z^1 / oraz optymalnego \underline{x}_d /jako funkcji czasu t oraz zakłócenia z^2 /, nie ograniczając jednak struktur tych zależności.

Mamy zatem rozwiązać problem /124/ względem $\underline{\alpha}, \underline{x}_d$:

$$\max_{\underline{\alpha}, \underline{x}_d} E [Q(\underline{x}, \underline{u}, \underline{z}, \underline{x}_2, t_2)] \quad /126/$$

z uwzględnieniem równania obiektu, ograniczeń obiektu, algorytmu sterowania bezpośredniego /125/ oraz decyzji ograniczających strukturę z^1, z^2 i odstęp adaptacji T_3^α .

Rozwiązanie zadania /126/ daje

$$\hat{\underline{\alpha}} = \hat{\underline{\alpha}}(kT_3) = \hat{\underline{\alpha}}(z^1) \quad /127/$$

$$\hat{\underline{x}}_d = \hat{\underline{x}}_d(t) = \hat{\underline{x}}_d(z^2) \quad /128/$$

$$E[Q]_{opt} = E[Q] \quad /129/$$

$$\begin{aligned} \underline{x}_d &= \hat{\underline{x}}_d \\ \underline{\alpha} &= \hat{\underline{\alpha}} \end{aligned}$$

W tym stadium można sprawdzić wrażliwość rezultatu /129/ na decyzje /125/, z^2, T_3^α . Zauważyć przy tym można, że /125/ z^1, T_3^α tworzą jedną grupę, wpływającą głównie na uchyb nadążania $\|\underline{x}_d - \underline{x}\|$, podczas gdy wybór z^2 wpływa bardziej bezpośrednio na jakość optymalizacji. Nasuwa się tu alternatywa postępowania: wyznaczać \underline{x}_d tak, jakby sterowanie bezpośrednie działało idealnie, $\underline{x} = \underline{x}_d$, i wybrać najwłaściwsze z^2 . Następnie optymalizować wybór /125/, z^1, T_3^α pod kątem małego uchybu nadążania \underline{x} za \underline{x}_d . Podkreślić warto, że wybór z^2 jest bardzo istotny, gdyż zbyt bogata reprezentacja zakłóceń z^2 może być kosztowna, a zbyt uboga może prowadzić do dużej straty na optymalności. Zauważmy również, że dyskusja wyboru z^2 może wykazać, jakie składniki lub cechy zakłóceń \underline{z} nie mają dużego znaczenia dla optymalizacji.

Jak widać, zasadniczą cechą metody kolejnego podejmowania decyzji jest to, że w pierwszym stadium ignorujemy fakt, że zarówno /127/ jak /128/ nie będzie dokładnie zrealizowane przez warstwę optymalizacji względnie adaptacji. Pozwala to jednak

dyskutować słuszność pierwszych decyzji, bez określenia struktur algorytmów sterowania w warstwach wyższych. W rezultacie wartość oceny jakości ostatecznie osiągnięta w całym układzie będzie niższa niż /129/, gdyż będzie $\underline{x}_d \neq \hat{\underline{x}}_d$ oraz $\underline{\alpha} \neq \hat{\alpha}$, skutkiem przybliżeń wprowadzonych w warstwie drugiej i trzeciej. Jeżeli zatem obliczona wartość /129/, np. przy pewnym wybranym z^2 , jest już za mała, należy korygować decyzje strukturalne związane z warstwą pierwszą.

Drugim stadium projektowania będzie określenie tych spośród pozostałych decyzji strukturalnych, które dotyczą najbardziej bezpośrednio warstwy drugiej:

struktura algorytmu optymalizacji $\underline{x}_d = \underline{x}_d(\beta, z^2)$ /130/
 budowa reprezentacji zakłóceń z^3
 odstępy próbkowania T_2, T_3^β

Po podjęciu tych decyzji można określić optymalne β jako funkcję czasu mT_3^β oraz reprezentacji zakłóceń z^3 , nie nakładając na razie ograniczeń na strukturę tej zależności.

Punktem wyjścia jest $\hat{\underline{x}}_d$ otrzymane w stadium poprzednim. Przyjmując, że $\underline{\alpha} = \hat{\alpha}$ wg /127/, możemy uważać za znaną wrażliwość /129/ na $\underline{x}_d \neq \hat{\underline{x}}_d$, gdyż w zasadzie możemy obliczać

$$E [Q] = Q_E(\underline{x}_d) \quad /131/$$

$$\begin{aligned} \underline{x}_d &\neq \hat{\underline{x}}_d \\ \underline{\alpha} &= \hat{\alpha} \end{aligned}$$

a zatem możemy dyskutować wpływ wyboru struktury /130/, z^3 oraz odstępów T_2, T_3^β . Przeciwwagą w stosunku do zwiększenia złożoności /130/ i skracania odstępów T_2, T_3^β jest zmniejszenie straty jakości

$$E [Q] - E [Q] \quad /132/$$

$$\begin{aligned} \underline{x}_d = \hat{\underline{x}}_d & \quad \underline{x}_d \neq \hat{\underline{x}}_d \\ \underline{\alpha} = \hat{\alpha} & \quad \underline{\alpha} = \hat{\alpha} \end{aligned}$$

Dla wybranego zbioru decyzji strukturalnych, wartość /132/ należy minimalizować, lub wartość /131/ należy maksymalizować^{x/},

x/ Należy zwrócić uwagę, że jeśli minimalizacja /132/ celu znalezienia $\hat{\beta}$ nie jest wykonalna, można pójść drogą pośrednią polegającą na określeniu takiego β , które minimalizowałoby pewną normę $\|\hat{\underline{x}}_d - \underline{x}_d\|$, gdzie $\underline{x}_d = \underline{x}_d(\beta, z^2)$. W tym jednak przypadku osiągnięta wartość $\|\hat{\underline{x}}_d - \underline{x}_d\|$ nie informuje o stracie jakości /za wyjątkiem przypadku gdy norma wyniesie zero/

otrzymując optymalną wartość:

$$\underline{\beta} = \underline{\beta} (mT_3^\beta) = \underline{\beta} (z^3) \quad /133/$$

Zauważmy, że decyzje drugiego stadium projektowania rozkładają się na dwie grupy. Odstęp próbkowania T_2 można uważać za parametr bezpośrednio wpływający na optymalizację w tym sensie, że dyskretyzacja $\hat{x}_d(t)$ obniża jakość nawet w przypadku, gdyby /130/ było idealne pod względem budowy zależności \hat{x}_d od z^2 . Z kolei ograniczenie struktury /130/ oraz z^3 , T_3^β powoduje dalsze obniżenie jakości oraz ma wpływ na złożoność adaptacji.

Trzecie stadium projektowania polega na określeniu ostatnich danych strukturalnych:

$$\text{struktura algorytmu adaptacji } \underline{\beta} = \underline{\beta} (\underline{\gamma}, z^3) \quad /134/$$

$$\text{struktura algorytmu adaptacji } \underline{\alpha} = \underline{\alpha} (\underline{\gamma}, z^1) \quad /135/$$

oraz na określeniu optymalnej wartości $\underline{\gamma}$.

Punktem wyjścia są $\hat{\underline{\alpha}}$ oraz $\hat{\underline{\beta}}$ otrzymane w stadiach poprzednich. Strata jakości spowodowana przez $\underline{\beta} \neq \hat{\underline{\beta}}$ oraz $\underline{\alpha} \neq \hat{\underline{\alpha}}$ określona będzie przez różnicę

$$E [Q]_{\substack{\underline{\alpha} = \hat{\underline{\alpha}} \\ \underline{\beta} = \hat{\underline{\beta}}}} - E [Q]_{\substack{\underline{\alpha} \neq \hat{\underline{\alpha}} \\ \underline{\beta} \neq \hat{\underline{\beta}}}} \quad /136/$$

a wybór złożoności /134/ i /135/ powinien być dokonany w kompromisie z wartością /136/. Wartość /136/ należy minimalizować względem $\underline{\gamma}$, otrzymując wartość $\underline{\gamma} = \hat{\underline{\gamma}}$.

Należy zauważyć, że procedura prowadząca do otrzymania $\hat{\underline{\gamma}}$ jest równoważna maksymalizacji wskaźnika jakości obiektu

$$\max_{\underline{\gamma}} E [Q (\underline{x}, \underline{u}, \underline{z}, \underline{x}_2, t_2)] \quad /137/$$

z uwzględnieniem równania obiektu, ograniczeń na stan \underline{x} i sterowanie \underline{u} oraz z uwzględnieniem ograniczeń struktury układu sterowania:

$$\underline{u} = \underline{u} (\underline{\alpha}, \underline{x}_d, \underline{x}) \quad \text{jak założono w /125/}$$

$$\underline{x}_d = \underline{x}_d (nT_2) = \underline{x}_d (\underline{\beta}, z^2) \quad \text{jak założono w /130/}$$

$$\underline{\beta} = \underline{\beta} (mT_3^\beta) = \underline{\beta} (\underline{\gamma}, z^3) \quad \text{jak założono w /134/}$$

$$\underline{\alpha} = \underline{\alpha} (kT_3^\alpha) = \underline{\alpha} (\underline{\gamma}, z^1) \quad \text{jak założono w /135/}$$

Zaletą podejmowania decyzji strukturalnych kolejno polega na tym, że ich wpływ na osiąganą jakość ostateczną /137/

jest rozpatrywany niejako oddzielnie. Jeśli osiągnięty wynik /137/ nie jest zadowalający, projektant układu musi rozpocząć procedurę iteracyjną, przechodząc ponownie przez trzy grupy decyzji:

$$T_3^\alpha, z^1, z^2, \underline{u} = \underline{u}(\alpha, \underline{x}_d, \underline{x}) \quad \text{/pierwsza warstwa/}$$

$$T_2, T_3^\beta, z^3, \underline{x}_d = \underline{x}_d(\beta, z^2) \quad \text{/druga warstwa/}$$

$$\beta = \beta(\gamma, z^3), \alpha = \alpha(\gamma, z^1) \quad \text{/trzecia warstwa/}$$

Wartość wskaźnika jakości osiągnana bez ograniczeń na układ sterowania:

$$\max_{\underline{u}} E \left[Q(\underline{x}, \underline{u}, \underline{z}, \underline{x}_2, t_2) \right] = \hat{Q}_E$$

powinna przy tym stale stanowić "poziom odniesienia".

Przykład

Dla ilustracji metody projektowania rozpatrzmy następujący prosty przykład.

Zadanie sterowania. Dany jest obiekt o równaniu $\dot{x} = u + z$. Wskaźnik jakości /maksymalizowany/ ma postać

$$Q = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \left(-\frac{x^2}{2} - zx \right) dt \quad \text{/138/}$$

Założenie $z(t)$ jest stacjonarne, $z(t) = \theta = 0$, $z^2(t) = \theta^2$, $R(\tau)$ znane.

W rozwiązaniu tego zadania optymalna jakość /otrzymywana przy $\hat{x} = -z$, $\hat{u} = 2z$ / wynosi

$$\max Q = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \frac{z^2}{2} dt \quad \text{/139/}$$

Wartość oczekiwana optymalnej jakości $E[\hat{Q}] = \frac{1}{2} \theta^2$. Ograniczenia strukturalne układu sterowania /decyzje projektowe/ będą wartość tę obniżać.

Pierwsza warstwa sterowania. Przyjmiemy strukturę układu zamkniętego, z regulatorem ciągłym proporcjonalnym

$$u = \alpha (x_d - x)$$

Po włączeniu regulatora wielkość wyjściowa x wynosi

$$x = \frac{\alpha}{1 + \alpha} x_d + \frac{1}{1 + \alpha} z$$

a wskaźnik jakości

$$Q = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \left[-\frac{\alpha^2}{2(1+\alpha)^2} x_d^2 - \frac{2\alpha + \alpha^2}{(1+\alpha)^2} z x_d - \frac{3 + 2\alpha}{2(1+\alpha)^2} z^2 \right] dt$$

Parametr α uznamy za nie podlegający adaptacji. Wartość zadana x_d ma być dobierana przez układ drugiej warstwy / układ optymalizacji/, w oparciu o reprezentację zakłóceń z^2 . Przyjmujemy następującą reprezentację zakłóceń z^2 : wartości $z(t)$ próbkowane z odstępem T , oraz cechy statystyczne zakłócenia:

$$\theta, \sigma^2, R(\tau).$$

Rozpatrzmy wskaźnik jakości za jeden okres T :

$$Q_k = \frac{1}{T} \int_{kT}^{(k+1)T} \left[-\frac{\alpha^2}{2(1+\alpha)^2} x_d^2 - \frac{2\alpha + \alpha^2}{(1+\alpha)^2} z x_d - \frac{3 + 2\alpha}{2(1+\alpha)^2} z^2 \right] dt$$

Sterowanie optymalne maksymalizujące tę jakość

$$\hat{x}_d = -\frac{2 + \alpha}{\alpha} z$$

Ponieważ tylko $z(kT)$ jest mierzone, \hat{x}_d trzeba zbudować w oparciu o estymatę zakłócenia z na okres T w przód. Załóżmy, że jako estymatę przyjmujemy wartość oczekiwaną warunkową

$$z^* = E[z(kT + \tau) | z(kT)] = \frac{z(kT)}{\sigma^2} R(\tau)$$

Podstawiając

$$\hat{x}_d = -\frac{2 + \alpha}{\alpha} z^*$$

do wskaźnika jakości, otrzymujemy dla okresu T

$$\hat{Q}_k = \frac{1}{T} \int_{kT}^{(k+1)T} \left[-\frac{(2 + \alpha)^2}{2(1 + \alpha)^2} z^*(z^*)^2 + \frac{(2 + \alpha)^2}{(1 + \alpha)^2} z z^* - \frac{3 + 2\alpha}{2(1 + \alpha)^2} z^2 \right] dt$$

Wartość oczekiwaną \hat{Q} wyniesie

$$E[\hat{Q}] = \frac{(2 + \alpha)^2}{2(1 + \alpha)^2} \cdot \frac{1}{\sigma^2 T} \int_0^T R^2 d\tau - \frac{3 + 2\alpha}{2(1 + \alpha)^2} \sigma^2$$

Zauważmy, że gdy $T \rightarrow 0$ /co oznacza kosztowną reprezentację $z(t)$, wskaźnik jakości dąży do

$$E[\hat{Q}] = \frac{1}{2} \sigma^2.$$

Żeby zbadać zależność od T , założmy $R(\tau) = \sigma^2 e^{-|\tau|}$, wówczas

$$\frac{1}{\sigma^2 T} \int_0^T R^2 d\tau = \frac{\sigma^2}{2T} (1 - e^{-2T})$$

oraz

$$E[\hat{Q}] = \left[\frac{(2+\alpha)^2}{2(1+\alpha)^2} \cdot \frac{1}{2T} (1 - e^{-2T}) - \frac{3+2\alpha}{2(1+\alpha)^2} \right] \sigma^2.$$

Wykres wartości $E[\hat{Q}]$ w funkcji T przedstawiono na rys.11.

Wykres wykonano przyjmując $\alpha = 9$. Przedstawia on stratę jakości powodowaną przez decyzje strukturalne pierwszej warstwy sterowania: przyjęto strukturę regulacyjną z parametrem $\alpha = 9$, oraz reprezentację $z(t)$ w postaci próbek użytą do wyznaczenia optymalnej wartości zadanej x_d . Linie "sterowanie idealne" oraz "bez sterowania" obrazują graniczne wartości, otrzymane odpowiednio przy pełnej optymalizacji i przy braku sterowania w ogóle.

Druga warstwa sterowania. Przyjmijmy następujące decyzje strukturalne:

- wartość x_d stała w przedziałach o długości T_2 , ustalona w zależności od $z(kT_2)$ wg algorytmu

$$x_d = \beta_1 + \beta_2 z(kT_2)$$

- parametry β_i nastawiane w oparciu o reprezentację zakłóceń z^3 , w postaci θ , σ^2 , $R(\tau)$, tj. w oparciu o cechy statystyczne $z(t)$,

- okres adaptacji $T_3^\beta > T_2$

Łatwo pokazać, że w tym przypadku najlepsze x_d jest określone przez

$$x_d(kT_2 + \tau) = -\frac{2+\alpha}{\alpha} \cdot \frac{z(kT_2)}{\sigma^2 T_2} \int_0^{T_2} R d\tau = \hat{\beta}_2 z(kT_2)$$

oraz $\hat{\beta}_1 = 0$.

Wartość oczekiwana wskaźnika jakości przy $\beta_i = \hat{\beta}_i$ wyniesie w jednym okresie T_2 :

$$E[\hat{Q}_k] = \frac{(2+\alpha)^2}{2(1+\alpha)^2} \left(\frac{1}{T_2} \int_0^{T_2} R d\tau \right)^2 \frac{z^2(kT_2)}{\sigma^4} - \frac{3+2\alpha}{2(1+\alpha)^2} \left[\sigma^2 - \frac{\sigma^2 - z^2(kT_2)}{\sigma^4} \cdot \frac{1}{T_2} \int_0^{T_2} R^2 d\tau \right]$$

a uśrednienie po wszystkich T_2 daje

$$E [\hat{Q}] = \left[\frac{(2 + \alpha)^2}{2(1 + \alpha)^2} \cdot \frac{1}{\sigma^4} \left(-\frac{1}{T_2} \int_0^{T_2} R d\tau \right)^2 - \frac{3 + 2\alpha}{2(1 + \alpha)^2} \right] \sigma^2.$$

Aby zdać sobie sprawę z wpływu wartości T_2 , przyjmijmy jak poprzednio $R(\tau) = \sigma^2 e^{-\tau}$. Wówczas

$$\frac{1}{\sigma^2} \left(-\frac{1}{T_2} \int_0^{T_2} R d\tau \right)^2 = \sigma^2 \left[\frac{1}{T_2} (1 - e^{-T_2}) \right]^2$$

oraz

$$E [\hat{Q}] = \left[\frac{(2 + \alpha)^2}{2(1 + \alpha)^2} \cdot \frac{1}{T_2^2} (1 - e^{-T_2})^2 - \frac{3 + 2\alpha}{2(1 + \alpha)^2} \right] \sigma^2.$$

Wykres $E[\hat{Q}]$ w funkcji T_2 przedstawiono na rys.12. Dla parametru α przyjęto wartość 9. Wykres ten podaje stratę jakości spowodowaną uproszczeniem optymalizacji $x_d = \text{const}$ w przedziałach T_2 , tj. spowodowaną decyzjami o budowie drugiej warstwy układu sterowania. Jeśli byśmy przyjęli koszt jednokrotnego wyznaczenia $x_d = \hat{\beta}_2$ z (kT_2) za c_2 , to natężenie kosztu w czasie wyniesie $\frac{c_2}{T_2}$. Wprowadzając to do rys.12 możemy zestawić wskaźnik jakości z kosztem realizacji sterowania i znaleźć ogólne maximum.

Trzecia warstwa sterowania. Zadaniem układu trzeciej warstwy byłoby nastawianie parametrów β_1, β_2 . W rozpatrywanym przykładzie będzie to:

- wyznaczanie wartości cech statystycznych $\theta, \sigma^2, R(\tau)$ na okres T_3
- adaptacja β_2 .

Przyjmijmy, że nastawianie β_2 odbywa się wg algorytmu

$$\beta_2 = -\frac{2 + \alpha}{\alpha} \gamma$$

gdzie powinno być

$$\gamma = \hat{\gamma} = \frac{1}{\sigma^2 T_2} \int_0^{T_2} R(\tau) d\tau.$$

Założmy, że skutkiem niedostatecznej znajomości statystyki zakłócenia wartość γ została ustalona błędnie jako $\gamma = k\hat{\gamma}$. Wtedy $\beta \neq \hat{\beta}$ i wskaźnik jakości wyniesie

$$E[\hat{Q}]_{\beta \neq \hat{\beta}} = \left[\frac{(2 + \alpha)^2}{(1 + \alpha)^2} \left(k - \frac{k^2}{2} \right) \hat{\gamma}^2 - \frac{3 + 2\alpha}{2(1 + \alpha)^2} \right] \sigma^2.$$

Przyjmijmy teraz, że γ jest wyznaczane z błędem przypadkowym, takim że $E[\gamma] = \hat{\gamma}$, $E[\gamma^2] = \hat{\gamma}^2(1 + \mu^2)$, czyli innymi słowami $E[k] = 1$, $E[k^2] = 1 + \mu^2$.

Jeśli zmienna przypadkowa k jest nieskorelowana z zakłóceniem z , to wartość oczekiwana Q może być obliczona jako

$$E[\hat{Q}]_{E[\beta] = \hat{\beta}} = \left[\frac{(2 + \alpha)^2}{2(1 + \alpha)^2} (1 - \mu^2) \hat{\gamma}^2 - \frac{3 + 2\alpha}{2(1 + \alpha)^2} \right] \sigma^2.$$

Wykres wskaźnika jakości w zależności od wariancji μ^2 przedstawiono na rys.13. Przyjęto tu $\alpha = 9$ oraz $R(\tau) = \sigma^2 e^{-(\tau)}$. Okres T_2 wynosi $T_2 = 0,5$. Jeśli przyjąć, że wyznaczanie γ z mniejszym błędem /małe μ^2 / jest bardziej kosztowne można przedstawić hipotetyczną krzywą natężenia kosztów jak przedstawiono na rys.13. Można też znaleźć wspólne maximum.

6. UKŁADY STEROWANIA BEZPOŚREDNIEGO

6.1. Sterowalność obiektu i wybór wielkości sterujących

Przyjmijmy, zgodnie z rozważaniami z punktu 5.1., że zadaniem układu sterowania bezpośredniego jest utrzymać zadany stan obiektu $\underline{x} = \underline{x}_d$, albo ogólnej - zapewnić nadążanie trajektorii stanu obiektu $\underline{x}(t)$ za przebiegiem zadany trajektorii $\underline{x}_d(t)$. Problemem naszym jest w pierwszym rzędzie znalezienie wielkości sterujących \underline{u} /wśród wielkości wejściowych obiektu \underline{u}' /, które spełniałyby odpowiednie warunki sterowalności.

Zakładając, że w chwili t_1 obiekt znajduje się w stanie $\underline{x}(t_1)$, rozpatrzyć należy dwa etapy sterowania: dotarcie od stanu początkowego do punktu leżącego na trajektorii zadanej $\underline{x}_d(t)$, które nastąpi w pewnej chwili t_1' oraz następnie poruszanie się wzdłuż tej trajektorii. Pierwszy z tych etapów można nazwać przebiegiem przejściowym; założymy, że obiekt jest sterowalny w tym sensie, tzn. że czas $t_1' - t_1$ jest skończony. Etap drugi, nadążanie za trajektorią, rozpatrzyć można jak następuje.

Dane jest równanie obiektu:

$$\dot{\underline{x}} = \underline{f}(\underline{x}, \underline{u}, t) \quad /138/$$

gdzie \underline{u} - wielkości sterujące.

Oznaczmy odchylenie od trajektorii

$$\underline{\xi} = \underline{x}_d - \underline{x} \quad /139/$$

Wprowadźmy $\underline{\xi}$ do /138/

$$\dot{\underline{\xi}} = \dot{\underline{x}}_d - \underline{f}(\underline{x}_d - \underline{\xi}, \underline{u}, t) \quad /140/$$

Poruszanie się ściśle po trajektorii zadanej wymaga $\underline{\xi} = \underline{0}$ oraz $\dot{\underline{\xi}} = \underline{0}$. Stawiamy zatem warunek, by układ równań /otrzymany z /140//

$$\dot{\underline{x}}_d - \underline{f}(\underline{x}_d - \underline{\xi}, \underline{u}, t) = \underline{0} \quad /141/$$

miał rozwiązanie zerowe, $\underline{\xi} = \underline{0}$. Ma to być osiągalne przez dobór wartości \underline{u} , które w /141/ są parametrami.

Zbadajmy prosty przypadek liniowy; dany jest obiekt o równaniu

$$\dot{\underline{x}} = A \underline{x} + B \underline{u}$$

który ma nadażać za trajektorią $\underline{x}_d(t)$. Wprowadzamy $\underline{\xi} = \underline{x}_d - \underline{x}$ i otrzymujemy

$$\dot{\underline{\xi}} = A \underline{\xi} - B \underline{u} - A \underline{x}_d + \dot{\underline{x}}_d \quad /142/$$

Stan ustalony nadażania za trajektorią, $\dot{\underline{\xi}} = 0$, daje

$$A \underline{\xi} = B \underline{u} + A \underline{x}_d - \dot{\underline{x}}_d$$

Jeśli istnieje A^{-1} , to równanie to ma jedyne rozwiązanie $\underline{\xi} = \underline{0}$, przy warunku

$$B \underline{u} = \dot{\underline{x}}_d - A \underline{x}_d \quad /143/$$

Warunek ten wyznacza jednoznacznie \underline{u} gdy istnieje B^{-1} :

$$\underline{u} = B^{-1} (\dot{\underline{x}}_d - A \underline{x}_d) \quad /144/$$

Sterowalność w nadażaniu za trajektorią wymaga zatem, by B było macierzą kwadratową nieosobliwą, jeżeli \underline{u} ma n elementów. Jeśli \underline{u} ma $m > n$ elementów, to musimy mieć wpływ na n spośród nich w sposób taki, by można było kształtować niezależnie wartości wszystkich n elementów wektora $B\underline{u}$.

Ogólniejszy przypadek nieliniowy rozpatrzyć można w sposób następujący.

Załóżmy, że pewna trajektoria \underline{x}_d jest wywołana przez sterowanie \underline{u}_d , spełniając równanie obiektu

$$\dot{\underline{x}}_d = \underline{f}(\underline{x}_d, \underline{u}_d, t)$$

Zażądajmy odchylenia $\underline{\xi}$ od tej trajektorii; niech dla wywołania $\underline{x}_d + \underline{\xi}$ potrzebne będzie sterowanie $\underline{u}_d + \underline{u}$:

$$\dot{\underline{x}}_d + \dot{\underline{\xi}} = \underline{f}(\underline{x}_d + \underline{\xi}, \underline{u}_d + \underline{u}, t)$$

Zapišmy prawą stronę tego równania w postaci rozwinięcia, słusznego dla małych $\underline{\xi}$ i małych \underline{u} ; otrzymamy wówczas związek

$$\dot{\underline{\xi}} = \frac{\partial \underline{f}}{\partial \underline{x}} \underline{\xi} + \frac{\partial \underline{f}}{\partial \underline{u}} \underline{u} \quad /145/$$

wyznaczający przyrost sterowania \underline{u}

$$\frac{\partial \underline{f}}{\partial \underline{u}} \underline{u} = \dot{\underline{\xi}} - \frac{\partial \underline{f}}{\partial \underline{x}} \underline{\xi} \quad /146/$$

Podobnie jak w interpretacji warunku /143/, jeśli odchylenie $\underline{\xi}$ ma mieć kierunek dowolny, to \underline{u} musi mieć n elementów, a macierz $\frac{\partial \underline{f}}{\partial \underline{u}}$ musi być macierzą kwadratową nieosobliwą. Warunek ten ma być spełniony dla każdej trajektorii \underline{x}_d .

6.2. Wielkości wyjściowe i obserwowalność

Przy rozpatrywaniu sterowalności obiektu nie trzeba zastanawiać się nad tym, czy obiekt będzie sterowany w układzie otwartym czy zamkniętym i co będzie w nim mierzone. Gdy natomiast rozważamy dalej układ sterowania bezpośredniego i decydujemy ukształtować go jako układ zamknięty, pojawia się potrzeba analizy obserwowalności obiektu. Z wyjątkiem przypadku, gdy dla pomiaru dostępne są bezpośrednio wszystkie współrzędne stanu, układ zamknięty sterowania opiera się na pomiarze wybranych wielkości wyjściowych \underline{y} , gdzie

$$\underline{y} = \underline{g}(\underline{x}, \underline{u}, t) \quad /147/$$

Powstaje pytanie, czy mierząc \underline{y} określać możemy jednoznacznie wektor stanu \underline{x} .

Rozpatrzmy to zagadnienie w sposób następujący. Zadana jest trajektoria \underline{x}_d i jesteśmy zainteresowani w spełnieniu równania:

$$\underline{\xi} = \underline{x}_d - \underline{x} = \underline{0} \quad /148/$$

Przyjmijmy, że wyjście \underline{y} zależy jawnie tylko od stanu \underline{x}

$$\underline{y} = \underline{g}(\underline{x}) \quad /149/$$

oraz założmy, że istnieje funkcja odwrotna do /149/, czyli $\underline{x} = \underline{g}^{-1}(\underline{y})$. Wówczas odchylenie od trajektorii /148/ zapiszemy

$$\underline{\xi} = \underline{x}_d - \underline{g}^{-1}(\underline{y}) \quad /150/$$

Aby móc sprawdzić spełnienie wymagania $\underline{\xi} = \underline{0}$ w oparciu o znajomość /pomiar/ \underline{y} , trzeba, by $\underline{g}^{-1}(\underline{y})$ było jednoznaczne ^{*}/.
Możliwość sprawdzenia czy $\underline{\xi} = \underline{0}$ nazwiemy obserwowalnością obiektu.

W przypadku liniowym zapiszemy

$$\underline{y} = C \underline{x} \quad /151/$$

zatem $\underline{x} = C^{-1} \underline{y}$, czyli dla spełnienia wymagania obserwowalności trzeba by jednoznacznie istniało C^{-1} .

Wynika stąd dla praktyki wnioski, że liczba wielkości wyjściowych /mierzonych/ musi być równa liczbie współrzędnych stanu.

Podkreślmy, że wprowadziliśmy tu obserwowalność stanu w sposób nieco odmienny i stawiający ostrzejsze wymagania, niż to spotyka się w literaturze dotyczącej optymalnego sterowania dynamicznego. Wiąże się to z postawieniem przez nas warunku, że mamy mieć możliwość sprawdzenia w każdej chwili, czy spełnione jest wymaganie /148/. Nie uciekamy się przy tym do żadnej obróbki dynamicznej sygnału $\underline{y}(t)$, gdyż zakładać musimy, że obiekt znajdować się może w stanie ustalonym. Zwróćmy też uwagę, że funkcja /147/ wzgl. /149/ nie może zależeć od dodatkowych parametrów, gdyż wówczas nie wyznaczymy dokładnie stanu procesu. Wiąże się to z dokładnością pomiarów.

6.3. Charakterystyki statyczne i transmitancje

Mając wybrane wielkości wejściowe i wyjściowe obiektu możemy przedstawić go w postaci schematu blokowego jak na rys.14. Dla konkretnego projektowania układów sterowania bezpośrednio /układów stabilizacji/ musimy znać charakterystyki statyczne i transmitancje obiektu regulacji przedstawionego na rys.14.

Jak to już wspomniano w punkcie 3.1, charakterystyki statyczne określone są równaniami

^{*}/ $\underline{g}^{-1}(\underline{y})$ istnieje, gdy macierz $\frac{\partial \underline{g}}{\partial \underline{x}}$ jest kwadratowa i nieosobliwa

$$\underline{f}(\underline{x}, \underline{u}) = \underline{0} \quad /35/$$

$$\underline{y} = \underline{g}(\underline{x}, \underline{u}) \quad /36/$$

z których wyrugować trzeba \underline{x} , aby otrzymać żądane związki

$$\underline{y} = \underline{\psi}(\underline{u}) \quad /152/$$

Związki te można w zasadzie otrzymać z /35/, /36/ zawsze - jeśli nie na drodze analitycznej, to numerycznej.

Jako przykład postępowania analitycznego rozpatrzmy obiekt o stałych rozłożonych przedstawiony na rys.6 i opisany w punkcie 3.3. Równania stanu mają tu postać:

$$\frac{\partial x_1}{\partial t} = v_1 \frac{\partial x_1}{\partial l} - k(T) x_1$$

$$\frac{\partial x_2}{\partial t} = -v_2 \frac{\partial x_2}{\partial l} + k(T) x_1$$

Przyjmijmy, zgodnie z punktem 5.1, że jako wielkości wejściowe sterujące przyjęto

natężenie dopływu $F_A(L, t)$ /wpływ poprzez warunek brzegowy/

prędkość transportu $v_1(t)$

to znaczy, że pozostałe wielkości wejściowe $v_2, T, F_B(0)$ są parametrami.

Wielkościami wyjściowymi będą

$$\text{zapełnienie na końcu reaktora } y_1 = x_1(L, t) + x_2(L, t)$$

$$\text{stosunek przepływów } y_2 = \frac{F_A(0, t)}{F_A(L, t)}$$

Równania wyjść mają postać

$$y_1 = x_1(L, t) + x_2(L, t) \quad /153/$$

$$y_2 = \frac{v_1 x_1(0, t)}{F_A(L, t)} \quad /154/$$

Równania stanu dają dla stanu ustalonego układ równań

$$v_1 \frac{dx_1}{dl} - k(T) x_1 = 0 \quad /155/$$

$$-v_2 \frac{dx_2}{dl} + k(T) x_1 = 0 \quad /156/$$

Rozwiązanie równania /155/ ma postać

$$x_1 (1) = \frac{F_A (L)}{v_1} \exp \left[- \frac{k (T)}{v_1} (L - 1) \right] \quad /157/$$

a rozwiązanie na $x_2 (1)$ brzmi

$$x_2 (1) = \frac{F_A (L)}{v_2} \exp \left[- \frac{k (T)}{v_1} (L - 1) \right] - \frac{F_A (L)}{v_2} \exp \left[- \frac{k (T)}{v_1} L \right] + \frac{F_B (0)}{v_2} \quad /158/$$

Wzory /157/, /158/ pozwalają wyznaczyć zależność y_1, y_2 od $F_A (L)$ oraz od v_1 , czyli od wielkości wejściowych. Będą to związki

$$y_1 = \frac{F_A (L)}{v_1} + \frac{F_A (L)}{v_2} \left[1 - \exp - \frac{k (T)}{v_1} L \right] + \frac{F_B (0)}{v_2} = \Psi_1 \left[F_A (L), v_1 \right] \quad /159/$$

$$y_2 = \exp - \frac{k (T)}{v_1} L = \Psi_2 (v_1) \quad /160/$$

Transmitancje obiektu z rys.14 otrzymamy w zasadzie z linearyzowanych równań stanu i równań wyjść

$$\dot{\underline{x}} = A \underline{x} + B \underline{u} \quad /137/$$

$$\underline{y} = C \underline{x} + D \underline{u} \quad /138/$$

jako macierz

$$\underline{G} (s) = C [sI - A]^{-1} B + D \quad /139/$$

tak jak to przedstawiono w punkcie 3.1, wzory /37/, /38/, /39/.

Rozpatrzmy, na przykład, reaktor z rys.3, traktowany jako obiekt o trzech wielkościach regulowanych W, C_A, T i trzech wielkościach sterujących F_D, F_A, H . Równania stanu obiektu /52/, /53/, /54/ trzeba zlinearyzować /symbole zmiennych oznaczają teraz przyrosty/:

$$\frac{dW}{dt} = - F_D + F_A \quad /161/$$

$$\frac{dC_A}{dt} = - \frac{(1 - C_{Ao}) F_{Ao} - C_{Ao} F_{Bo}}{W_o^2} W - \left[\frac{F_{Ao} + F_{Bo}}{W_o} + k (T_o) \right] C_A - C_{Ao} \frac{\partial k (T_o)}{\partial T} T + \frac{1 - C_{Ao}}{W_o} F_A \quad /162/$$

$$\frac{dT}{dt} = - \frac{(T_{Ao} - T_o) F_{Ao} + (T_{Bo} - T_o) F_{Bo} + H_o)c}{W_o^2} W - \frac{h k (T_o)}{c} C_A -$$

$$- \left[\frac{F_{Ao} + F_{Bo}}{W_o} - \frac{h}{c} C_{Ao} \frac{\partial k (T_o)}{\partial T} \right] T + \frac{T_{Ao} - T_o}{W_o} F_A + \frac{1}{c W_o} H \quad /163/$$

W równaniach tych W_o , C_{Ao} , T_o , F_{Ao} , F_{Bo} , T_{Ao} , T_{Bo} , H_o oznaczają wartości ustalone w punkcie równowagi, w otoczeniu którego dokonano linearyzacji. Wprowadzając oznaczenia współczynników w równaniach /161/ /161/ /163/, zapiszemy je w skrócie

$$\frac{dW}{dt} = - F_D + F_A \quad /164/$$

$$\frac{dC_A}{dt} = a_{21} W + a_{22} C_A + a_{23} T + b_{22} F_A \quad /165/$$

$$\frac{dT}{dt} = a_{31} W + a_{32} C_A + a_{33} T + b_{32} F_A + b_{33} H \quad /166/$$

oraz w postaci macierzowej

$$\begin{bmatrix} \dot{W} \\ \dot{C}_A \\ \dot{T} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} W \\ C_A \\ T \end{bmatrix} + \begin{bmatrix} -1 & 1 & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} F_D \\ F_A \\ H \end{bmatrix} \quad /167/$$

Transformując /167/ przy zerowych warunkach początkowych, otrzymamy wynik typu /39/:

$$\begin{bmatrix} W(s) \\ C_A(s) \\ T(s) \end{bmatrix} = \begin{bmatrix} s & 0 & 0 \\ -a_{21} & s-a_{22} & -a_{23} \\ -a_{31} & -a_{32} & s-a_{33} \end{bmatrix}^{-1} \begin{bmatrix} -1 & 1 & 0 \\ 0 & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} F_D(s) \\ F_A(s) \\ H(s) \end{bmatrix} \quad /168/$$

W rozpatrywanym przykładzie nie wchodziły w grę równania wyjść, gdyż wielkościami regulowanymi były wprost współrzędne stanu.

Macierz transmitancji występująca w równaniu /168/ zawiera ogółem 9 elementów. Wypiszemy je po kolei

$$\frac{W(s)}{F_D(s)} = \frac{\begin{vmatrix} -1 & 0 & 0 \\ 0 & s-a_{22} & -a_{23} \\ 0 & -a_{32} & s-a_{33} \end{vmatrix}}{\begin{vmatrix} s & 0 & 0 \\ -a_{21} & s-a_{22} & -a_{23} \\ -a_{31} & -a_{32} & s-a_{33} \end{vmatrix}} = \frac{-(s-a_{22})(s-a_{33}) + a_{23}a_{32}}{s(s-a_{22})(s-a_{33}) - a_{23}a_{32}s} = -\frac{1}{s} \quad /169/$$

$$\frac{W(s)}{F_A(s)} = \frac{1}{s} \quad /170/$$

$$\frac{W(s)}{H(s)} = 0 \quad /171/$$

$$\frac{C_A(s)}{F_D(s)} = \frac{-a_{21}(s-a_{33}) - a_{23}a_{31}}{s[(s-a_{22})(s-a_{33}) - a_{23}a_{32}]} \quad /172/$$

$$\frac{C_A(s)}{F_A(s)} = \frac{(s-a_{33})(b_{22}s+a_{21}) + a_{23}b_{32}s + a_{23}a_{31}}{s[(s-a_{22})(s-a_{33}) - a_{23}a_{32}]} \quad /173/$$

$$\frac{C_A(s)}{H(s)} = \frac{a_{23}b_{33}}{(s-a_{22})(s-a_{33}) - a_{23}a_{32}} \quad /174/$$

$$\frac{T(s)}{F_D(s)} = \frac{-a_{31}(s-a_{22}) - a_{21}a_{31}}{s[(s-a_{22})(s-a_{33}) - a_{23}a_{32}]} \quad /175/$$

$$\frac{T(s)}{F_A(s)} = \frac{(s-a_{22})(b_{32}s + a_{31}) + b_{22}a_{32}s + a_{21}a_{32}}{s[(s-a_{22})(s-a_{33}) - a_{23}a_{32}]} \quad /176/$$

$$\frac{T(s)}{H(s)} = \frac{b_{33}(s-a_{22})}{(s-a_{22})(s-a_{33}) - a_{23}a_{32}} \quad /177/$$

Tak jak można się było spodziewać po rządzie obiektu, mianowniki w transmitancjach /169/ - /177/ są co najwyżej stopnia trzeciego.

Rozpatrzony przykład dotyczył obiektu o stałych skupionych; obliczenie transmitancji układu o stałych rozłożonych, opisywanego równaniami cząstkowymi, jest zwykle znacznie trudniejsze.

Rozpatrzmy obiekt z rys.6. Jego równania stanu /por.punkt 3.3/ mają postać:

$$\frac{\partial x_1}{\partial t} = v_1 \frac{\partial x_1}{\partial l} - k(T) x_1 \quad /63/$$

$$\frac{\partial x_2}{\partial t} = -v_2 \frac{\partial x_2}{\partial l} + k(T) x_1 \quad /65/$$

Jeśli wejściami sterującymi będą $F_A(L, t) = \frac{x_1(L, t)}{v_1}$ oraz $v_1(t)$, a wielkościami wyjściowymi /patrz punkt 5.1/

$$y_1(t) = x_1(L, t) + x_2(L, t) \quad /115/$$

$$y_2(t) = \frac{v_1 x_1(0, t)}{F_A(L, t)} \quad /116/$$

to poszukiwane transmitancje otrzymamy poprzez odpowiednią linearyzację równań /63/, /65/, /115/, /116/. Linearyzacja ma dotyczyć zależności od zmiennej t , przy czym interesują nas przyrosty x_1, x_2, F_A, v_1 .

Linearyzacja równań stanu daje

$$\frac{\partial x_1}{\partial t} = \left(\frac{\partial x_1}{\partial l} \right)_0 v_1 + v_{10} \frac{\partial x_1}{\partial l} - k(T_0) x_1 \quad /178/$$

$$\frac{\partial x_2}{\partial t} = -v_{20} \frac{\partial x_2}{\partial l} + k(T_0) x_1 \quad /179/$$

Jako warunki brzegowe i początkowe należy teraz przyjąć, jako dla zmiennych przyrostowych, wartości

$$x_1(l, 0) = 0; \quad x_2(l, 0) = 0 \quad /180/$$

$$x_1(L, t) = \frac{1}{v_{10}} F_A(L, t) - \frac{F_{A0}(L)}{v_{10}^2} v_1; \quad x_2(0, t) = 0 \quad /181/$$

Linearyzacja równań wyjścia daje

$$y_1(t) = x_1(L, t) + x_2(L, t) \quad /182/$$

$$y_2(t) = \frac{v_{10}}{F_{A0}(L)} x_1(0, t) + \frac{x_{10}(0)}{F_{A0}(L)} v_1(t) - \frac{v_{10} x_{10}(0)}{F_{A0}^2(L)} F_A(L, t) \quad /183/$$

W równaniach /178/, /179/ $\frac{\partial x_1}{\partial l}$, $\frac{\partial x_2}{\partial l}$ oznaczają funkcje o wartościach przyrostowych, natomiast $\left(\frac{\partial x_1}{\partial l} \right)_0$ jest współ-

czynnikiem wynikającym ze stanu ustalonego w punkcie równowagi, czyli z ustalonej wartości $x_1(1)$. Jego wartość obliczymy przez rozwiązanie równania

$$\left(\frac{dx_1}{dl}\right)_0 v_{10} - k(T_0) x_{10} = 0 \quad /184/$$

co daje

$$x_{10}(1) = \frac{F_{Ao}(L)}{v_{10}} \exp\left[-\frac{k(T_0)}{v_{10}}(L-1)\right] \quad /185/$$

oraz

$$\left(\frac{\partial x_1}{\partial l}\right)_0 = \frac{k(T_0)F_{Ao}(L)}{v_{10}^2} \exp\left[-\frac{k(T_0)}{v_{10}}(L-1)\right] = h(1) \quad /186/$$

Równania /178/ /179/ i warunki /181/ można transformować wg Laplace'a:

$$sx_1(1, s) - x_1(1, 0) = h(1)v_1(s) + v_{10} \frac{dx_1(1, s)}{dl} - k(T_0)x_1(1, s) \quad /187/$$

$$sx_2(1, s) - x_2(1, 0) = -v_{20} \frac{dx_2(1, s)}{dl} + k(T_0)x_1(1, s) \quad /188/$$

$$x_1(L, s) = \frac{1}{v_{10}} F_A(L, s) - \frac{F_{Ao}(L)}{v_{10}^2} v_1(s) \quad /189/$$

$$x_2(0, s) = 0 \quad /190/$$

Uwzględniając $x_1(1, 0) = 0$, $x_2(1, 0) = 0$ przepiszemy

$$\frac{dx_1(1, s)}{dl} = \frac{1}{v_{10}} [s + k(T_0)] x_1(1, s) - \frac{v_1(s)}{v_{10}} h(1) \quad /191/$$

$$\frac{dx_2(1, s)}{dl} = -\frac{s}{v_{20}} x_2(1, s) + \frac{k(T_0)}{v_{20}} x_1(1, s) \quad /192/$$

Równania /191/, /192/ z warunkami /189//190/ dają się rozwiązać

$$x_1(1, s) = \left\{ \frac{k(T_0)F_{Ao}(L)}{v_{10}^2 s} \left[\exp(-\beta(L-1)) - \exp(-\alpha(L-1)) \right] \right\} v_1(s) + \frac{1}{v_{10}} \exp(-\alpha(L-1)) F_A(L, s) \quad /193/$$

$$x_2(l, s) = \left[\frac{1}{s + v_2 \beta} \cdot \frac{k^2(T_0) F_{A0}(L)}{v_{10}^2 \cdot s} \exp(-\beta(L-l)) - \frac{1}{s + v_{20} \alpha} \cdot \frac{k(T_0) F_{A0}(L)}{v_{10}} \cdot \left(1 + \frac{k(T_0)}{s}\right) \exp(-\alpha(L-l)) \right] \cdot (1 - \exp(-\gamma l)) v_1(s) + \left[\frac{1}{s + v_{20} \alpha} \frac{k(T_0)}{v_{10}} \exp(-\alpha(L-l)) - \frac{1}{s + v_2 \beta} \frac{k(T_0)}{v_{10}} \exp(-\beta(L-l)) \right] \cdot (1 - \exp(-\gamma l)) F_A(L, s) \quad /194/$$

przy czym

$$\beta = \frac{k(T_0)}{v_{10}}, \quad \alpha = \frac{s + k(T_0)}{v_{10}}, \quad \gamma = \frac{s}{v_{20}}.$$

Mając rozwiązania analityczne /193/, /194/ oraz równania wyjść /115/, /116/ określimy wielkości wyjściowe. Potrzebne są nam do tego $x_1(L, s)$ i $x_2(L, s)$ z rozwiązań /193/, /194/ oraz $x_{10}(0)$, które otrzymamy ze /185/. W rezultacie otrzymamy szukane zależności

$$y_1(s) = G_{11}(s) v_1(s) + G_{12}(s) F_A(L, s) \quad /195/$$

$$y_2(s) = G_{21}(s) v_1(s) + G_{22}(s) F_A(L, s) \quad /196/$$

gdzie $G_{11}(s)$, $G_{12}(s)$, $G_{21}(s)$, $G_{22}(s)$ dadzą się wyznaczyć z rozwiązań /193/, /194/, /185/ i równań /115/, /116/. Transmittancje te będą mieć postać skomplikowaną.

W rozpatrzonym przykładzie możliwe było uzyskanie rozwiązań analitycznych, patrz /193/, /194/. W związku z tym transmittancje $G_{ij}(s)$ otrzymać można było również w postaci wzorów.

Sytuacja taka nie zawsze ma miejsce. Mianowicie układ równań linearyzowanych względem czasu po transformowaniu względem Laplace'a, który ma postać ogólną /porównaj /191/, /192/:

$$\frac{dx(l, s)}{dl} = \underline{f} [x(l, s), u(l, s), l, s] \quad /197/$$

może przy danych warunkach brzegowych $x(L, s)$ lub $x(0, s)$ nie mieć rozwiązania analitycznego. Różnania /197/ mogą być np. nieliniowe /mówimy o zmiennej l /, a w warunkach brzegowych x część może być dana na początek ($l = 0$), a część na koniec ($l = L$) przedziału zmiennej l . W takim przypadku można tylko

poszukiwać numerycznego wyznaczenia charakterystyk częstotliwościowych. Podstawiając mianowicie w /197/ $s = j\omega$ napiszemy

$$\frac{d\underline{x}(l, j\omega)}{dl} = \underline{f} [\underline{x}(l, j\omega), \underline{u}(l, j\omega), l, j\omega] \quad /198/$$

Zmienna $\underline{x}(l, j\omega)$, ściśle każdy jej element $x_1(l, j\omega)$, $x_2(l, j\omega), \dots$ będzie mieć część rzeczywistą i część urojona. Podobnie zmienne znane w zadaniu, to znaczy $\underline{u}(l, j\omega)$ oraz warunki brzegowe $\underline{x}(L, j\omega)$ względnie $\underline{x}(0, j\omega)$. Zapisując

$$\underline{x}(l, j\omega) = \underline{x}_R(l, \omega) + j \underline{x}_u(l, \omega)$$

możemy z układu równań /198/ otrzymać, przez rozdzielenie części rzeczywistej i części urojonej, dwa układy równań zawierających parametr ω :

$$\frac{d\underline{x}_R(l, \omega)}{dl} = \text{Re } \underline{f} [\underline{x}(l, j\omega), \underline{u}(l, j\omega), l, j\omega] \quad /199/$$

z warunkami brzegowymi $\underline{x}_R(L, \omega)$ bądź $\underline{x}_R(0, \omega)$, oraz

$$\frac{d\underline{x}_u(l, \omega)}{dl} = \text{Im } \underline{f} [\underline{x}(l, j\omega), \underline{u}(l, j\omega), l, j\omega] \quad /200/$$

z warunkami brzegowymi $\underline{x}_u(L, \omega)$ bądź $\underline{x}_u(0, \omega)$.

Równania /199/, /200/ nadają się do rozwiązania numerycznego, jeśli zadać wartość parametru ω oraz wszystkich warunków i sterowań.

Przyjmijmy jeszcze, że dane jest linearyzowane z punktu widzenia zmiennej t i transformowane równanie wyjść /por. /115/, /116/

$$\underline{y}(s) = \underline{g} [\underline{x}(l, s), \underline{u}(l, s), l, s] \quad /201/$$

gdzie \underline{g} jest w zasadzie funkcjonalem, wiążącym wartość zmiennej wyjściowej \underline{y} z przebiegiem $\underline{x}(l)$ czy $\underline{u}(l)$. Przy $s = j\omega$ równanie /201/ daje

$$\underline{y}_R(\omega) = \text{Re } \underline{g} [\underline{x}(l, j\omega), \underline{u}(l, j\omega), l, j\omega] \quad /202/$$

$$\underline{y}_u(\omega) = \text{Im } \underline{g} [\underline{x}(l, j\omega), \underline{u}(l, j\omega), l, j\omega] \quad /203/$$

Przypomnijmy, że jeśli na wejście obiektu dynamicznego liniowego doprowadzić sterowanie w postaci impulsu jednostkowego /impulsu Diraca/ $\delta(t)$, to transformata Fourier'a przebiegu wyjściowego $y(j\omega)$ jest transmitancją widmową, a wartości $y(j\omega)$ dla różnych ω są punktami charakterystyki amplitudowo-fazowej.

Rozpatrzmy dla naszego obiektu o stałych rozłożonych jedno jedyne sterowanie, na przykład w postaci warunku brzegowego $x_1(L, t)$ albo przebiegu czasowego jednej wielkości sterującej w określonym punkcie $l = l_1$, czyli $u_1(l_1, t)$. Rozpatrzmy również jedną wielkość wyjściową, $y_1(t)$. Sytuacja taka opisana będzie równaniami, będącymi konkretnym przypadkiem /199/, /200/ oraz /202/ /203/, a mianowicie /przyjmujemy sterowanie przez $x_1(L, t)$ o postaci przebiegu $\delta(t)$ /.

Z równań stanu:

$$\frac{d \underline{x}_r(l, \omega)}{dl} = \text{Re } \underline{f} [\underline{x}(l, j\omega), l, j\omega] \quad /204/$$

$$\frac{d \underline{x}_u(l, \omega)}{dl} = \text{Im } \underline{f} [\underline{x}(l, j\omega), l, j\omega] \quad /205/$$

- warunek brzegowy na x_1 :

$$x_1(L, j\omega) = 1, \text{ czyli } x_{1r}(L, \omega) = 1, x_{1u}(L, \omega) = 0$$

- warunki brzegowe na pozostałe zmienne-zerowe, zadane bądź na koniec jako $x_1(L, j\omega)$, bądź na początek jako $x_1(0, j\omega)$;

Z równania wyjścia:

$$y_{1r}(\omega) = \text{Re } g_1 [\underline{x}(l, j\omega), l, j\omega] \quad /206/$$

$$y_{1u}(\omega) = \text{Im } g_1 [\underline{x}(l, j\omega), l, j\omega] \quad /207/$$

Przy znanej postaci równań, jeśli rozwiążemy numerycznie układy /204/, /205/, otrzymamy wartości współrzędnych $x_{1r}(l, \omega)$, $x_{1u}(l, \omega)$ itd. Podstawiając je do /206/, /207/ wyliczamy współrzędne punktu charakterystyki amplitudowo-fazowej $y_{1r}(\omega)$, $y_{1u}(\omega)$.

Przedstawimy jeszcze prosty przykład. Dane są linearyzowane i unormowane równania przeciwprądowego wymiennika ciepła /patrz [1], rozdz.2/:

$$\frac{\partial x_1}{\partial t} = -\frac{\partial x_1}{\partial l} - N_1 (x_1 - x_2)$$

$$\frac{\partial x_2}{\partial t} = \frac{1}{\beta} \frac{\partial x_2}{\partial l} - \frac{N_2}{\beta} (x_2 - x_1)$$

gdzie $x_1(l, t)$, $x_2(l, t)$ - temperatury, β , N_1 , N_2 - współczynniki. Długość $0 \leq l \leq 1$. Niech warunki brzegowe dla tych równań mają postać /dla odchyień/:

$$x_1(l, 0) = 0, \quad x_2(l, 0) = 0$$

$$x_1(0, t) = u(t), \quad x_2(1, t) = 0$$

Interesującą nas wielkością wyjściową jest temperatura $x_2(0, t)$, czyli równanie wyjścia ma postać

$$y(t) = x_2(0, t).$$

Linearyzowane równania stanu transformujemy wg Fouriera:

$$j\omega x_1(1, j\omega) - x_1(1, 0) = -\frac{dx_1(1, j\omega)}{dl} - N_1 [x_1(1, j\omega) - x_2(1, j\omega)]$$

$$j\omega x_2(1, j\omega) - x_2(1, 0) = \frac{1}{\beta} \frac{dx_2(1, j\omega)}{dl} - \frac{N_2}{\beta} [x_2(1, j\omega) - x_1(1, j\omega)]$$

Uwzględniając zerowe warunki $x_1(1, 0)$, $x_2(1, 0)$ oraz kładąc

$$x_1(1, j\omega) = x_{1r}(1, \omega) + j x_{1u}(1, \omega)$$

$$x_2(1, j\omega) = x_{2r}(1, \omega) + j x_{2u}(1, \omega)$$

otrzymamy cztery równania, powstałe przez rozdzielenie części rzeczywistych i urojonych

$$\frac{d x_{1r}}{dl} = -N_1 x_{1r} + \omega x_{1u} + N_1 x_{2r}$$

$$\frac{d x_{1u}}{dl} = -\omega x_{1r} - N_1 x_{1u} + N_1 x_{2u}$$

$$\frac{d x_{2r}}{dl} = -N_2 x_{1r} + N_2 x_{2r} - \beta \omega x_{2u}$$

$$\frac{d x_{2u}}{dl} = -N_2 x_{1u} + \beta \omega x_{2r} + N_2 x_{2u}$$

dla których warunkami brzegowymi są

$$x_{1r}(0, \omega) = 1 \text{ /przyjęto } u(t) = \delta(t)$$

$$x_{1u}(0, \omega) = 0$$

$$x_{2r}(1, \omega) = 0$$

$$x_{2u}(1, \omega) = 0$$

Równania powyższe należy rozwiązać, poszukując wartości $x_2(0, j\omega)$ czyli liczb $x_{2r}(0, \omega)$ oraz $x_{2u}(0, \omega)$, dla różnych wartości ω . Zadanie to jest wykonalne numerycznie.

Rezultatem będą punkty charakterystyki amplitudowo-fazowej jak pokazano na rysunkach w [1] .

6.4. Niektóre aspekty projektowania i realizacji

Mając określone wielkości regulowane, wybrane wielkości sterujące oraz dane charakterystyki statyczne i transmitancje przystąpić możemy do projektowania układu sterowania bezpośredniego. Tok tego projektowania jest w zasadzie konwencjonalny i nie będziemy go omawiać w szczegółach; zwrócimy tylko uwagę na niektóre aspekty, wiążące projektowanie układu sterowania bezpośredniego z resztą systemu.

Wybór wskaźników jakości regulacji

W rozważaniach optymalizacji w układzie wielowarstwowym wyznacza się dla układu sterowania bezpośredniego funkcję nadążania za trajektorią zadaną $x_d(t)$ wzgl. $y_d(t)$. Układ sterowania bezpośredniego projektować potrafimy wg żądań typu

$$\min \|y_{d1} - y_1\|, \quad \min \|y_{d2} - y_2\|, \quad \dots \quad \min \|y_{dn} - y_n\|$$

/208/

przy czym wybór wskaźników jakości, to jest norm występujących w wymaganiach /208/, nie jest obojętny.

Byłoby najlepiej, gdyby wybór ten był powiązany ze wskaźnikiem jakości Q , który występuje w ogólnym zadaniu optymalizacji obiektu sterowania. Przedstawimy to na przykładzie.

Założmy, że wskaźnik jakości Q /minimalizowany/ zależy od rozpatrywanej wielkości regulowanej w sposób następujący

$$Q = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} [2y + (4 - y)^2] dt \quad /209/$$

W układzie regulacji będzie występował uchyb $e = y_d - y$. Należy znaleźć najwłaściwszą normę dla minimalizacji tego uchybu. Wprowadźmy $y = y_d - e$ do /209/

$$Q = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} [2y_d - 2e + (4 - y_d + e)^2] dt \quad /210/$$

W wyrażeniu /210/ wielkość e będzie przypadkowo zmienna /pod wpływem zakłóceń działających na obiekt/, natomiast y_d jest stałą liczbą. Porządkując odpowiednio /210/ otrzymamy

$$Q = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} [2y_d + (4 - y_d)^2 + (8 - 2y_d - 2)e + e^2] dt \quad /211/$$

Przyjmijmy, że wartość średnia $e(t)$ w przedziale $t_2 - t_1$ wyniesie zero, co odpowiada istnieniu regulatora z całkowaniem /PI, PID/. Wówczas minimalizacja /211/ sprowadza się do wymagania

$$\min \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} e^2 dt \quad /212/$$

czyli do wskaźnika jakości regulacji w postaci wartości średniokwadratowej, oraz zarazem do takiego ustalenia wartości zadanej y_d , by było

$$\min_{y_d} [2y_d + (4 - y_d)^2] \quad /213/$$

co daje $\hat{y}_d = 3$. Wartość ta jest, w tym przypadku, identyczna z wartością optymalną \hat{y} jaką wyliczylibyśmy wprost z minimalizacji /209/. Wobec istnienia uchybu $e \neq 0$ zgodność taka nie zawsze wystąpi.

Wymagania odsprzężenia /autonomizacji/

Specyficzne dla układu o wielu wielkościach regulowanych jest wymaganie odsprzężenia polegające na tym, by zmiana każdej z wartości zadanych y_d wpływała na jedną i tylko jedną wielkość regulowaną y . Spełnienie tego wymagania sprowadza się, jak wiadomo /patrz np. [1] rozdz.11/ do pewnych warunków strukturalnych na macierzową transmitancję regulatora $R(s)$. Wiadomo również, że zadośćuczynienie warunkom autonomizacji spowodować może obniżenie osiągalnej jakości regulacji w układzie, nie należy zatem tego wymagania stawiać zbyt pochopnie.

Kompensacja zakłóceń

W układzie o jednej wielkości regulowanej pełną kompensację zakłóceń otrzymać można tylko na drodze sprzężenia dodatkowego z pomiarem wielkości zakłócającej. W układzie o wielu wielkościach regulowanych istnieją możliwości uzyskania analogicznego efektu bez pomiaru zakłócenia, przez wykorzystanie istniejących sprzężeń skrośnych w obiekcie. Prowadzić to może do swego rodzaju układów kaskadowych /również patrz [1], rozdz. 11/.

Bezpośrednie sterowanie cyfrowe /DDC/

Niezawodność cyfrowych maszyn sterujących pozwala już na to, by powierzyć im nie tylko funkcje warstw wyższych /optymalizacji, adaptacji/, ale również funkcje sterowanie bezpośredniego, por. rys.7,9. Mamy wówczas do czynienia z tzw. bezpośrednim sterowaniem cyfrowym /DDC/, gdzie żądane zależności między wielkością regulowaną i jej uchybem a wielkością sterującą obiektem realizowane są przez algorytm cyfrowy w maszynie. Zazwyczaj jest to pewna wersja dyskretna algorytmu PI lub PID, patrz np. [8]. Specyficzne tu będzie impulsowe działanie każdego z obwodów regulacji, patrz okres T_1 na rys.9, gdyż maszyna cyfrowa wykonuje tylko jeden algorytm jednocześnie. Z punktu widzenia maszyny, istotne będzie oszczędne zaprojektowanie tych często powtarzanych działań oraz właściwe ich umieszczenie w całości systemu operacyjnego maszyny sterującej.

7. UKŁADY OPTYMALIZACJI

7.1. Struktury układów optymalizacji statycznej

Przyjmijmy, dla ustalenia uwagi, że przez układ optymalizacji rozumiemy drugą warstwę sterowania z rys.9, to jest wyznaczenie optymalnej trajektorii stanu \hat{x}_d . Celem jest uzyskanie maximum wskaźnika jakości obiektu Q , a podstawą wyznaczania trajektorii optymalnej mogą być pomiary zakłóceń lub inne dane, zależne od wybranej struktury układu optymalizacji.

Skoncentrujmy uwagę najpierw na zadaniu optymalizacji statycznej; w zadaniu takim wyznaczany jest optymalny stan ustalony, to znaczy liczby $x_{d1}, x_{d2}, \dots, x_{dn}$. Możliwe struktury układu optymalizacji są następujące:

- a/ układ otwarty
- b/ układ zamknięty
- c/ układ mieszany
- d/ układ otwarty z adaptacją modelu.

Układ otwarty przedstawiony jest schematycznie na rys.15. Układem takim jest również struktura warstwy optymalizacji na rys.9. Istotą jest oparcie wyznaczenia \hat{x}_d na pomiarze i ewent. obróbce statystycznej zakłóceń, cechą charakterystyczną - potrzeba posiadania dokładnego modelu obiektu /wraz z warstwą sterowania bezpośredniego/. Na rys.15 zaznaczono, że optymalizacja

zator wyznacza \hat{x}_d w zależności od reprezentacji zakłóceń z' . W przeprowadzeniu tego wyznaczania możliwe są dwa przypadki. Pierwszy przypadek zachodzi, gdy zadanie optymalizacji było rozwiązane analitycznie względem zmiennych niezależnych z' i odpowiednie wzory wynikowe $\hat{x}_d = f(z')$ są sprowadzone w urządzenie realizujące sterowanie. Drugi przypadek występuje wtedy, gdy zadanie nie ma rozwiązania analitycznego i wobec tego, dla każdej wartości z' , trzeba je numerycznie rozwiązywać - zazwyczaj przy użyciu procedury szukania ekstremum. Szukanie odbywa się na modelu matematycznym obiektu, "wewnątrz" optymalizatora, jak to ilustruje rys.16. Łatwo zauważyć, że w przypadku, jak na rys.16 realizacja optymalizatora praktycznie musi się opierać na maszynie matematycznej. Można w pewnym sensie sprowadzić przypadek drugi do przypadku pierwszego; wyobraźmy sobie mianowicie, że obliczenie numeryczne optymalnych decyzji \hat{x}_d przeprowadzono dla dużego i należytego wybranego zbioru wartości z' . Powstały wówczas zbiór danych liczbowych można aproksymować pewnymi hiperpowierzchniami, wyznaczającymi - już w postaci analitycznej - potrzebne zależności.

$$x_{d1} = f_1(z')$$

$$x_{d2} = f_2(z')$$

.....

$$x_{dn} = f_n(z')$$

/214/

Te właśnie zależności trzeba wówczas "wbudować" w optymalizator, działający ściśle wg rys.15. Trzeba zwrócić uwagę, że wyznaczenie zależności /214/ może być niezwykle pracochłonne, ze względu na konieczność przygotowania danych liczbowych w postaci wielokrotnego numerycznego rozwiązywania zadania optymalizacji.

Zamknięty układ optymalizacji przedstawiony jest na rys. 17. Mamy to do czynienia z takim samym wyznaczaniem \hat{x}_d na drodze szukania ekstremum Q, jak na rys.16, z tą różnicą, że szukanie odbywa się wprost na obiekcie, a nie na jego modelu. Odpada zatem kwestia wierności modelu, pojawia się natomiast bardzo istotny sposób sprawa czasu znalezienia ekstremum. Na rys.16 model obiektu "poddany" jest tylko tym zakłóceniom z' , które chcemy uwzględnić w optymalizacji. W czasie procedury szukania w układzie z rys.17 działają wszystkie zakłócenia, co

sytuację utrudnia. Ponadto w grę wchodzi dynamika obiektu, to jest konieczność odczekiwania na zakończenie przebiegu przejściowego po każdej zmianie decyzji \underline{x}_d - pamiętajmy, że optymalizować mamy stan ustalony. Przy szukaniu na modelu problem ten nie występuje, gdyż "model obiektu" na rys.16 jest oczywiście modelem statycznym. Omówione cechy powodują, że praktyczne zastosowanie układów typu rys.17 w czystej formie jest bardzo małe.

Układ mieszany powstaje przez połączenie sterowania otwartego rys.15 lub rys.16 ze sterowaniem w układzie zamkniętym, jak to pokazuje rys.18. Mamy teraz połączenie zalet obu układów: szybkość reakcji właściwą układowi otwartemu oraz niewrażliwość na niedokładność modelu, która jest cechą układu zamkniętego. Podkreślić może trzeba, że zarówno w układzie z rys.17 jak rys.18 występuje pomiar wskaźnika jakości Q . Osiągana zatem dokładność nie może przekroczyć granicy określonej przez uchyb tego pomiaru. Biorąc pod uwagę, że Q jest zwykle wyznaczone pośrednio, jako funkcja kilku wielkości mierzonych w obiekcie, możemy patrzeć sceptycznie na uzyskiwane dokładności lepszych niż procenty. Tymczasem spodziewane zwiększenie Q dzięki zastosowaniu optymalizacji może być rzędu podobnego lub mniejszego - co oczywiście stawia rezultaty działania układu zamkniętego pod znakiem zapytania i zwiększa zatem nasz nacisk na posiadanie dokładnego modelu obiektu.

Układ otwarty z adaptacją modelu ma praktycznie największe znaczenie spośród dotąd omówionych. Polega to na tym, że skoro wyznaczanie $\hat{\underline{x}}_d$ typu sterowania otwartego jest zawsze szybkie oraz skuteczne wtedy, gdy model obiektu jest prawidłowy, należy szukać możliwości poprawiania tego modelu na bieżąco /"on line"/. Jest tu wiele wariantów strukturalnych. Pierwszy z nich przedstawiono na rys.19: parametry modelu β /nie jest istotne, czy są to parametry w zależnościach /214/ czy też parametry modelu obiektu pokazanego na rys.16/ są dobiierane drogą bezpośredniego szukania ekstremum wskaźnika Q . Łatwo zauważyć, że struktura rys.19 jest ewolucją układu z rys.18 - po prostu szukanie odbywa się względem parametrów β , a nie względem sterowań \underline{x}_d . Inny rodzaj układu adaptacji przedstawiono na rys.20. Tutaj, w oparciu o odpowiednie obserwacje wejść \underline{x}_d i wyjścia Q budowany jest model zależności $Q(\underline{x}_d)$, z którego wynikają właściwe parametry β dla relacji typu /214/

lub dla modelu obiektu, użytego dla optymalizacji wg rys.16. Jeszcze inną możliwość przedstawia w istocie rzeczy rys. 9: zaznaczono tam, że parametry β dobierane są w układzie otwartym, w oparciu o pomiar i obróbkę zakłóceń.

7.2. Struktury układów optymalizacji dynamicznej

Układy optymalizacji dynamicznej mają szereg cech specyficznych, dzięki którym dogodnie jest ich struktury omawiać oddzielnie od struktur układów optymalizacji statycznej. Rozpatrzmy struktury następujące:

- a/ układ otwarty,
- b/ układ zamknięty,
- c/ układ zamknięty dwuwarstwowy,
- d/ układ repetycyjny,
- e/ układ ze sprzężeniem optymalizującym.

Układ otwarty dynamicznego sterowania optymalnego przedstawiono na rys.21. Zwróćmy uwagę, że na rysunku tym optymalizator wyznacza sterowania \underline{u} , a nie trajektorię optymalną \hat{x}_d . Są i inne różnice w stosunku do rys.15: ponieważ do wyznaczenia \underline{u} w chwili t potrzebna jest m.in. znajomość zakłóceń w całym przyszłym przedziale $t_1 - t$, niezbędna się staje predykcja przebiegu $\underline{z}(t)$; posiadany przez optymalizator model matematyczny obiektu musi być modelem dynamicznym, przy czym nadal od jego dokładności zależy powodzenie sterowania; pojawił się nowy czynnik sterujący w postaci warunków początkowych $\underline{x}(t_1)$, itd.

Układ zamknięty optymalizacji dynamicznej w postaci najbardziej klasycznej pokazuje rys.22. Algorytm optymalizatora jest tu wyznaczony jako $\hat{u}(\underline{x})$, a nie w postaci generacji przebiegu czasowego $\hat{u}(t)$. Wiadomo z teorii optymalizacji dynamicznej, że istniejąca w układzie rys.22 kontrola stanu obiektu daje na ogół mniejszą wrażliwość optymalizacji na niedokładność modelu obiektu oraz na zakłócenia, niż to ma miejsce w układzie otwartym z rys.21. Jeżeli model jest całkowicie dokładny, a zakłóceń nie ma, oba układy są równoważne. Jeżeli model jest niedokładny lub istnieją zakłócenia, układ zamknięty z rys.22 przestaje być układem ściśle optymalnego sterowania.

Odmianą w pewnym sensie układu zamkniętego z rys.22 jest układ dwuwarstwowy, przedstawiony na rys.23. Istnieje tu kontrola stanu, lecz sterowanie \underline{u} wyznaczane jest nie z rozwiąza-

nia zadania dynamicznego, lecz z warunku zgodności \underline{x} z trajektorią zadaną $\hat{\underline{x}}_d$. Natomiast $\hat{\underline{x}}_d$ wyznaczone jest z zadania dynamicznego, i to stanowi o różnicy między układem z rys.23 a układami na rys.15, 16. Zwróćmy uwagę, że $\hat{\underline{x}}_d(t)$ wyznaczone byłoby w oparciu o dynamikę samego obiektu, a nie wypadkową dynamikę obiektu z regulatorami. Na przykład, gdyby prowadzić cykliczny proces w reaktorze z rys.3, wyznaczenie $\hat{\underline{x}}_d(t)$ polegałoby na ustaleniu optymalnego narastania temperatury i malenia stężenia C_A , oczywiście z użyciem równań stanu obiektu. Regulatory dbałyby następnie o utrzymanie tych optymalnych przebiegów w obiekcie, eliminując tym samym wiele spośród zakłóceń.

Działanie układu repetycyjnego /rys.24/ jest zbliżone do układu zamkniętego z rys.22. Sterowanie $\hat{\underline{u}}(t)$ wyznaczone jest tu według zasad układu otwartego, na cały przedział $t_2 - t_1$, lecz wynik sterowania /w postaci stanu obiektu/ jest po upływie pewnego czasu T kontrolowany. Jeżeli stan obiektu nie zgadza się z przewidywanym przez model istniejący w optymalizatorze, wówczas stan rzeczywisty w danej chwili $t_1 + T$ uznawany jest za stan początkowy dla nowego, skróconego przedziału optymalizacji. Na nowo wyznacza się sterowanie $\hat{\underline{u}}(t)$, tym razem na przedział $t_2 - t_1 - T$, po czym po upływie nowego okresu T konfrontuje się wynik ze stanem obiektu itd. Realizacja układu repetycyjnego musi się operować na maszynie matematycznej, gdyż istotą działania jest wielokrotne powtórzenie obliczenia optymalizacyjnego. Zwróćmy uwagę, że wyznacza się tu $\hat{\underline{u}}(t)$, co jest łatwiejsze niż znalezienie zależności $\hat{\underline{u}}(\underline{x})$.

Układ ze sprzężeniem optymalizującym ma w prostym przypadku postać jak na rys.25. Istotą rzeczy jest tutaj to, że urządzenie sterujące wykonuje bieżąco funkcję ekstremalizacji pewnego wskaźnika liczbowego, jest więc w zasadzie urządzeniem /lub programem/ optymalizacji statycznej. Natomiast ten wskaźnik liczbowy jest tak wybrany, że otrzymuje się rezultat w postaci dynamicznego sterowania optymalnego danym obiektem. Rozpatrzmy to na prostym przykładzie.

Dany jest obiekt pierwszego rzędu, nieliniowy, o równaniu

$$\dot{\underline{x}} = f_1(\underline{x}, u, t) \quad /215/$$

oraz wskaźnik jakości /minimalizowany/

$$Q = \int_{t_1}^{t_2} f_0(x, u, t) dt \quad /216/$$

Warunek konieczny optymalności /zasada maximum/ brzmi w tym przypadku

$$\max [H = -f_0 + \Psi_1 f_1 + \Psi_2] = 0 \quad /217/$$

Zróbmy założenie, że w rozwiązaniu optymalnym $f_0 > 0$ /oznacza to np. że koszt nie spada do zera i nie zmienia się w zysk/. Wówczas maksymalizacja hamiltonianu /217/ jest równoważna maksymalizacji stosunku jego części dodatniej do ujemnej

$$\max \left[N_1 = \frac{\Psi_1 f_1 + \Psi_2}{f_0} \right] \quad /218/$$

Założmy dalej, że $\Psi_1 f_1$ dominuje nad składnikiem Ψ_2 , który wyraża koszt upływu czasu. Musi być wówczas $\Psi_1 f_1 > 0$, niezależnie od znaku Ψ_2 . Jeśli ponadto na całej trajektorii optymalnej $\dot{x} > 0$, czyli $f_1 > 0$ /można to stwierdzić z pewnością badając przybliżone rozwiązania analityczne/, to musi być $\Psi_1 > 0$. Można wówczas zastąpić /218/ przez wymaganie

$$\max \left[N = \frac{f_1 + \Psi_2 / \Psi_1}{f_0} \right] \quad /219/$$

co oznacza bieżącą /to jest w każdej chwili czasu/ maksymalizację wskaźnika N. Wartości f_1 oraz f_0 mogą być mierzone w obiekcie. Składnik Ψ_2 / Ψ_1 trzeba oszacować z danych analitycznych. W zadaniu niezależnym od czasu i o swobodnej chwili końcowej t_2 czynnik Ψ_2 / Ψ_1 będzie równy zeru.

Wykazuje się [7], że układ o strukturze rys.25 może mieć znacznie mniejszą wrażliwość na niedokładność modelu obiektu /model wchodzi tylko w oszacowanie Ψ_2 / Ψ_1 / oraz na zakłócenia, niż inne struktury układów optymalizacji dynamicznej.

7.3. Kryteria adaptacji modeli

Z rozpatrzenia struktur układów optymalizacji wynika, że duże znaczenie mają układy otwarte. W układach otwartych musi być użyty model obiektu; wystąpić on może w różnej formie. Jeśli zatrzymać uwagę na układach optymalizacji statycznej, to model może mieć np. postać "modelu sterowania" wyrażonego wzorami /214/, bądź też postać "modelu całościowego", który wyra-

załby zależność

$$Q = f(\underline{x}_d, \underline{z}) \quad /220/$$

jak tego wymaga działanie układu z rys.16.

Model /220/ nie daje się na ogół wyznaczyć w postaci jawnej z równań obiektu i wzoru na wskaźnik jakości Q . Tym niemniej, mając równania obiektu, wskaźnik jakości i ograniczenia możemy wyznaczać \underline{x}_d zapewniające $\max Q$ /metodą Kubna-Tuckera lub metodą szukania ekstremum/, zatem nie jest rzeczą istotną czy zależność /220/ dana jest jawnie czy niejawnie.

W otwartym układzie optymalizacji na plan pierwszy wysuwa się sprawa adaptacji modelu, gdyż trudno przypuszczać by model był kiedykolwiek wyznaczony z uwzględnieniem wszystkich wpływów zakłócających.

Przypomnijmy teraz, dla koncentracji rozważań, że rozpatrujemy obiekt rzeczywisty, który jest opisany pewną zależnością

$$Q = Q(\underline{x}_d, \underline{z}) \quad /221/$$

oraz że obiekt ten jest sterowany w oparciu o użycie modelu, z konieczności przybliżonego. Model ten ma stałą strukturę i zbiór nastawialnych parametrów β . Nie jest istotne, czy jest to "model sterowania" typu /214/ czy "model całościowy" o postaci typu /220/. W danym przypadku istnieje sterowanie optymalne oparte na modelu \underline{x}_d^* oraz powstaje pytanie, według jakiego kryterium dobierać należy parametry modelu β . Odpowiedzi można udzielić stosunkowo prosto - należy osiągnąć

$$\min_{\beta, \underline{z}} E \left[Q(\hat{\underline{x}}_d, \underline{z}) - Q(\underline{x}_d^*, \underline{z}) \right] \quad /222/$$

gdzie $Q(\underline{x}_d^*, \underline{z})$ oznacza wskaźnik osiągnięty w obiekcie rzeczywistym przy sterowaniu opartym na modelu (\underline{x}_d^*) , natomiast $Q(\hat{\underline{x}}_d, \underline{z})$ oznacza wskaźnik, jaki byłby osiągnięty w obiekcie przy sterowaniu ściśle optymalnym $(\hat{\underline{x}}_d)$.

Operowanie celem /222/ jest trudne, bowiem wymaga znajomości ścisłego opisu procesu $Q(\underline{x}_d, \underline{z})$ oraz wyznaczenia sterowania ściśle optymalnego $\hat{\underline{x}}_d$. Ostatnią trudność można ominąć stawiając zamiast /222/ wymaganie

$$\max_{\beta, \underline{z}} E \left[Q(\underline{x}_d^*, \underline{z}) \right] \quad /223/$$

Wynik ekstremalizacji /223/ będzie, z punktu widzenia doboru, taki sam jak ekstremalizacja /222/, bowiem tylko \underline{x}_d^* zależy od β . Różnica polega na tym, że warzosc liczbowa /222/ wyraża stratę optymalności spowodowaną użyciem modelu przybliżonego, podczas gdy wartość /223/ pozwala jedynie badać wrażliwość sterowania opartego na modelu na zmiany parametrów β .

Użycie /223/ wymaga znajomości dokładnej równań obiektu, bądź też adaptacji doświadczalnej polegającej na optymalizacji β drogą szukania wprost na obiekcie połączonym z optymalizatorem. Odpowiadałoby to schematowi z rys.19.

Jeżeli użycie ocen bezpośrednich /222/ lub /223/ nie jest możliwe, szukać należy dróg pośrednich. Najważniejszą z nich jest minimalizacja różnicy między modelem całościowym

$$Q^* = Q^*(\underline{x}_d, \underline{z}, \beta) \quad /224/$$

a zależnością rzeczywistą /221/, w określonym zakresie zmian \underline{x}_d oraz \underline{z} , na przykład minimalizacja wartości średniokwadratowej

$$\min_{\beta, \underline{z}} E \left[Q^*(\underline{x}_d, \underline{z}, \beta) - Q(\underline{x}_d, \underline{z}) \right]^2 \quad /225/$$

Należy dobrze zauważyć różnicę między /225/ a /222/. Obliczenie /222/ wymaga ścisłej znajomości opisu obiektu $Q(\underline{x}_d, \underline{z})$, gdyż trzeba do opisu tego podstawić wartości $\hat{\underline{x}}_d, \underline{x}_d^*$. Ocena /225/ może być stosowana przy $Q(\underline{x}_d, \underline{z})$ znanym tylko w postaci wyników doświadczalnych. Stosując /225/ dążymy do zgodności modelu $Q^*(\underline{x}_d, \underline{z}, \beta)$ z rzeczywistością $Q(\underline{x}_d, \underline{z})$.

Zwróćmy uwagę, że nie jest obojętne na jakim zbiorze danych $\underline{x}_d, \underline{z}$ oparta zostanie budowa modelu /225/. Ponieważ model /224/ służyć ma do wyznaczania sterowania optymalnego \underline{x}_d przy danych zakłóceniach \underline{z} , istotne jest by model odpowiadał rzeczywistości pod względem zależności od \underline{x}_d ; mówiąc ściślej, zależność modelowa:

$$Q^*(\underline{x}_d) \quad \underline{z} = \text{const}$$

musi wyznaczać tę samą wartość optymalną \underline{x}_d^* , co wartość optymalna $\hat{\underline{x}}_d$ wyznaczana z zależności rzeczywistej $Q(\underline{x}_d)_{\underline{z}=\text{const}}$. A więc, na przykład, zapewnione być musi prawidłowe wyznaczenie pochodnych $\frac{\partial Q}{\partial \underline{x}_d}$. Celu tego możemy nie osiągnąć, jeśli sprawdzimy zgodność modelu z rzeczywistością tylko dla wartoś-

ci optymalnych \hat{x}_d . Sprawa ta jest bliżej zilustrowana w pracy [4].

Osiągnięta wartość oceny /225/ nie wskazuje, ściśle rzecz biorąc, na ile jakość sterowania opartego o model $Q^*(x_d, z, \beta)$ będzie się różnić od jakości optymalnej $Q(\hat{x}_d, z)$.

8. OPTIMALIZACJA WIELOPOZIOMOWA

8.1. Sformułowanie zadania

Rozwiązywanie zadań optymalizacji dla obiektów dynamicznych o dużej wymiarowości lub zadań statycznych o dużej liczbie zmiennych decyzyjnych oraz związków ograniczających wymaga dużych nakładów obliczeniowych. Nakłady te mogą być w wielu przypadkach znacznie zmniejszone jeśli zadanie pierwotne poddamy dekompozycji, to jest podziałowi na kilka zadań częściowych, które byłyby z kolei koordynowane przez wyższy poziom sterowania.

Sposób podziału zadania jest oczywisty jeśli zadania częściowe mogą być sformułowane w taki sposób, że nie mają one zmiennych wspólnych. Oznaczałoby to jednak, że zadanie pierwotne jest w istocie zbiorem zadań nie sprzężonych ze sobą bezpośrednio, poza tym że składają się one na wspólny wskaźnik jakości. Sytuacja taka nie zachodzi na ogół w praktyce; rozpatrywać będziemy zatem przypadek, gdy zadania częściowe nie są rozdzielne z natury, a ich separacja jest uzyskiwana przez wprowadzenie zmiennych koordynacyjnych. Zmiennymi koordynacyjnymi zajmuje się drugi, tj. wyższy poziom sterowania. Jego zadaniem jest zarówno uzyskanie ekstremum globalnego jakości, jak też zapewnienie dopasowania do siebie wzajemnie wszystkich zadań częściowych, zgodnie z ich istniejącymi w rzeczywistości powiązaniem. Jest oczywiste, że praktyczny wynik w postaci zmniejszenia nakładu obliczeń otrzymamy tylko wówczas, gdy zmiennych koordynacyjnych będzie mniej niż zmiennych decyzyjnych w zadaniu pierwotnym. Przypadek ten zachodzi wówczas, gdy w zadaniu pierwotnym nie wszystkie zmienne są powiązane ze sobą przez każdą z zależności ograniczających. W przypadku, gdy np. ograniczenia liniowe zadania optymalizacji statycznej wyrażone są macierzą /por. 236/ oznacza to, że część elementów w tej macierzy jest zerami.

Podstawową strukturą jest, jak wynika z powyższego omówienia, struktura dwupoziomowa. Łącząc ze sobą czyli koordynując kilka układów dwupoziomowych dodajemy trzeci poziom sterowania i tak dalej. Sens tego postępowania polegać musi jednak zawsze na oszczędności nakładu obliczeniowego.

Rozpatrujemy, dla koncentracji uwagi, zadanie optymalizacji statycznej

$$\max Q(\underline{u}) \quad /226/$$

przy ograniczeniu

$$\underline{u} \in U. \quad /227/$$

Aczkolwiek duże praktyczne znaczenie mają suboptymalne rozwiązania zadania /226/, /227/ zajmować się będziemy dalej tylko wielopoziomowymi rozwiązaniami ściśle optymalnymi, gdyż tylko takie rozważanie można przeprowadzić całkowicie jednoznacznie.

Wprowadzamy zbiór zmiennych koordynacyjnych \underline{v} i dokonujemy podziału zadania /226/, /227/ tak, że ma ono postać

$$\max_{\underline{u} \in U} Q(\underline{u}) = \max_{\substack{\underline{u} \in U \\ \underline{v} \in V}} Q(\underline{u}, \underline{v}) = \max_{\substack{\underline{u}^1 \in U^1(\underline{v}) \\ \underline{u}^2 \in U^2(\underline{v}) \\ \dots \\ \underline{u}^N \in U^N(\underline{v}) \\ \underline{v} \in V}} Q[Q_1(\underline{u}^1, \underline{v}), Q_2(\underline{u}^2, \underline{v}), \dots, Q_N(\underline{u}^N, \underline{v})] \quad /228/$$

gdzie $\underline{u}^i \subset \underline{u}$ są pewnymi częściami zbioru \underline{u} . Związek /228/ wyraża, że przedstawiamy Q jako pewną funkcję N czynników Q_1, Q_2, \dots, Q_N , gdzie Q_1 zależy od \underline{u}^1 oraz /być może/ \underline{v} , a także że zmieniamy sformułowanie zależności ograniczających w taki sposób, by powstało N oddzielnych grup ograniczeń na $\underline{u}^1, \underline{u}^2, \dots$, zależnych od wektora \underline{v} , oraz pewne ograniczenie na sam wektor \underline{v} .

Rozwiązanie dwupoziomowe zadania /226/ z użyciem dekompozycji wyrażonej w /228/ jest możliwe, jeżeli a/ \underline{u}^i , $i=1, \dots, N$ są rozłącznymi podzbiorami \underline{u} , oraz jeżeli b/ postać /228/ pozwala na wykonanie rozłącznej ekstremalizacji względem każdego z \underline{u}^i :

$$\max_{\underline{u} \in U} Q(\underline{u}) = \max_{\underline{v} \in V} Q \left[\max_{\underline{u}^1 \in U^1(\underline{v})} Q_1(\underline{u}^1, \underline{v}), \max_{\underline{u}^2 \in U^2(\underline{v})} Q_2(\underline{u}^2, \underline{v}), \dots, \max_{\underline{u}^N \in U^N(\underline{v})} Q_N(\underline{u}^N, \underline{v}) \right] \quad /229/$$

Po oznaczeniu

$$\max_{\underline{u}^i \in U^i(\underline{v})} Q_i(\underline{u}^i, \underline{v}) = \hat{Q}_1(\underline{v}) \quad /230/$$

związek /229/ przyjmie postać

$$\max_{\underline{u} \in U} Q(\underline{u}) = \max_{\underline{v} \in V} Q[\hat{Q}_1(\underline{v}), \hat{Q}_2(\underline{v}), \dots, \hat{Q}_N(\underline{v})] \quad /231/$$

Wzór /230/ wyraża zadanie pierwszego poziomu, lub zadanie lokalne optymalizacji. Zadanie to jest parametryczne względem \underline{v} . Wzór /231/ wyraża zadanie drugiego poziomu, lub zadanie optymalizacji globalnej. Jest ono wykonywane względem zmiennych \underline{v} .

8.2. Sposoby dekompozycji

Związki /228/ i /229/ pokazują, że dla danego zadania /226/, /227/ może istnieć wiele sposobów dekompozycji, wszystkie dopuszczalne z punktu widzenia optymalizacji globalnej. Może być różna liczba utworzonych zadań częściowych, różna liczba wprowadzonych zmiennych koordynacyjnych /wymiar \underline{v} /, różny sposób przeformułowania związków ograniczających. W zasadzie, najlepszy będzie ten sposób dekompozycji, przy którym nakład obliczeniowy łączny dla całej optymalizacji będzie najmniejszy. Ten jednak punkt widzenia jest trudny do ujęcia analitycznego, omówimy zatem tylko pokrótce niektóre cechy i zasady dokonywania dekompozycji. Rozpatrzmy oddzielnie dekompozycję wskaźnika jakości $Q(\underline{u})$ oraz zależności ograniczających $\underline{u} \in U$.

Wskaźnik jakości $Q(\underline{u})$ pozwala na rozłączną ekstremalizację względem $\underline{u}^i / \underline{u}^i$ są rozłącznymi podzbiorami \underline{u} jeżeli jest on jednego z następujących typów:

a/ wyłącznie addytywny

$$Q(\underline{u}) = \sum_{i=1}^N Q_i(\underline{u}^i, \underline{v}) \quad /232/$$

b/ wyłącznie multiplikatywny

$$Q(\underline{u}) = \prod_{i=1}^N Q_i(\underline{u}^i, \underline{v}) \quad /233/$$

z warunkiem, że $Q_i(\underline{u}^i, \underline{v}) \geq 0$, $i = 1, 2, \dots, N$,

c/ mieszany, o rozłącznej części addytywnej i multiplikatywnej

$$Q(\underline{u}) = \prod_{i=1}^1 Q_1(\underline{u}^i, \underline{v}) + \sum_{i=1+1}^N Q_1(\underline{u}^i, \underline{v}) \quad /234/$$

z warunkiem, że $Q_1(\underline{u}^i, \underline{v}) \geq 0$, $i = 1, 2, \dots, 1$,

d/ mieszany, z nieantagonistycznym wpływem czynników wspólnych części addytywnej i multiplikatywnej $\ast/$:

$$Q(\underline{u}) = \prod_{i=1}^1 Q_1(\underline{u}^i, \underline{v}) + \sum_{i=1}^{i \leq 1} Q_1(\underline{u}^i, \underline{v}) + \sum_{i=1+1}^N Q_1(\underline{u}^i, \underline{v}) \quad /235/$$

z warunkiem, że $Q_1(\underline{u}^i, \underline{v}) \geq 0$, $i = 1, 2, \dots, 1$

Globalne ograniczenie $\underline{u} \in U$ jest, zgodnie z /228/, przekształcone w zbiory ograniczeń $\underline{u}^i \in U^i(\underline{v})$, $\underline{v} \in V$. Dokonuje się tego przez wprowadzenie zmiennych koordynacyjnych \underline{v} . Trudno o generalne reguły; zazwyczaj mamy do czynienia z kompromisem między liczbą wprowadzonych zmiennych, a prostotą rozwiązywania utworzonych zadań. Rozpatrzmy pewne przypadki dla ilustracji.

Przyjmijmy, że nałożone na \underline{u} ograniczenia są liniowymi lub nieliniowymi nierównościami

$$A\underline{u} \geq \underline{b} \quad /236/$$

przy czym macierz A nie jest blokowo-diagonalna, lecz ma szeregi elementów zerowych. Wyobraźmy sobie przeniebrowanie zmiennych takie, by otrzymana macierz A składała się z bloków, powiązanych ze sobą przez niektóre kolumny:

$\ast/$ Zauważmy, że np. wskaźnik $Q(\underline{u}) = Q_1(\underline{u}^1, \underline{v}) Q_2(\underline{u}^2, \underline{v}) Q_3(\underline{u}^3, \underline{v}) - Q_2(\underline{u}^2, \underline{v})$ nie może być ekstremalizowany rozłącznie ze względu na \underline{u}^2 , gdyż sposób w jaki $Q_2(\underline{u}^2, \underline{v})$ wpływa na całość poprzez część addytywną i multiplikatywną jest antagonistyczny.

$$\left[\begin{array}{cccc}
 a_{11} & \dots & a_{11} & \\
 \dots & & \dots & \\
 a_{s1} & \dots & a_{s1} & \\
 & & a_{(s+1)1} & \dots a_{(s+1)k} \\
 & & \dots & \dots \\
 & & a_{t1} & \dots a_{tk} \\
 & & & a_{(t+1)k} \dots a_{(t+1)n} \\
 & & & \dots \dots \\
 & & a_{rk} & \dots a_{rn}
 \end{array} \right] \left[\begin{array}{c}
 u_1 \\
 \dots \\
 \dots \\
 u_1 \\
 \dots \\
 \dots \\
 u_k \\
 \dots \\
 \dots \\
 u_n
 \end{array} \right] \geq \left[\begin{array}{c}
 b_1 \\
 \dots \\
 \dots \\
 b_s \\
 \dots \\
 \dots \\
 b_{s+1} \\
 \dots \\
 \dots \\
 b_t \\
 \dots \\
 \dots \\
 b_{t+1} \\
 \dots \\
 \dots \\
 b_r
 \end{array} \right] \quad /237/$$

Macierz w równaniu /237/ można przekształcić w macierz blokowo-diagonalną przez usunięcie zmiennych pierwotnych u_1, u_k i wprowadzenie na ich miejsce zmiennych koordynacyjnych v'_1, v''_1, v'_k, v''_k oraz wprowadzenie nowych dwóch ograniczeń równościowych

$$\begin{aligned}
 v'_1 - v''_1 &= 0, \\
 v'_k - v''_k &= 0
 \end{aligned} \quad /238/$$

Wynikiem jest nowy układ związków ograniczających, który zastąpi /237/.

Macierz opisująca te ograniczenia jest teraz blokowo-diagonalna, kosztem wzrostu jej wymiaru o dwie kolumny i dwa wiersze:

$$\left[\begin{array}{cccc}
 a_{11} & \dots & a_{11} & \\
 \dots & & \dots & \\
 a_{s1} & \dots & a_{s1} & \\
 & & a_{(s+1)1} & \dots a_{(s+1)k} \\
 & & \dots & \dots \\
 & & a_{t1} & a_{tk} \\
 & & & a_{(t+1)k} \dots a_{(t+1)n} \\
 & & & \dots \dots \\
 & & & a_{rk} \quad a_{rn} \\
 & & 1 & -1 \\
 & & & 1 \quad -1
 \end{array} \right] \left[\begin{array}{c}
 u_1 \\
 \dots \\
 \dots \\
 u_{1-1} \\
 v'_1 \\
 v''_1 \\
 \dots \\
 \dots \\
 u_{k-1} \\
 v'_k \\
 v''_k \\
 \dots \\
 \dots \\
 u_n
 \end{array} \right] \geq \left[\begin{array}{c}
 b_1 \\
 \dots \\
 \dots \\
 b_s \\
 \dots \\
 \dots \\
 b_{s+1} \\
 \dots \\
 \dots \\
 b_t \\
 \dots \\
 \dots \\
 b_{t+1} \\
 \dots \\
 \dots \\
 b_r \\
 0 \\
 0
 \end{array} \right] \quad /239/$$

Górna część macierzy występującej w /239/ opisuje oddzielne zadania częściowe; co więcej, w tym szczególnym przypadku są one również rozdzielnie ograniczone. Na przykład, pierwsze zadanie częściowe ma ograniczenia

$$\begin{aligned} a_{11}u_1 + a_{12}u_2 + \dots + a_{1(l-1)}u_{l-1} &\stackrel{\geq}{=} b_1 - a_{1l}v_1' \\ \dots & \\ a_{s1}u_1 + a_{s2}u_2 + \dots + a_{s(l-1)}u_{l-1} &\stackrel{\geq}{=} b_s - a_{sl}v_1' \end{aligned} \quad /240/$$

lub w skróconym zapisie

$$\underline{u}^1 \in U^1(v_1') \quad /240'/$$

Zadanie drugie ma ograniczenia

$$\begin{aligned} a^{(s+1)}(l+1)u_{l+1} + a^{(s+1)}(l+2)u_{l+2} + \dots + a^{(s+1)}(k-1)u_{k-1} &\stackrel{\geq}{=} \\ &= b_{s+1} - a^{(s+1)}l v_1'' - a^{(s+1)}k v_k' \end{aligned} \quad /241/$$

$$a_t(l+1)u_{l+1} + a_t(l+2)u_{l+2} + \dots + a_t(k-1)u_{k-1} \stackrel{\geq}{=} b_t - a_{tl}v_1'' - a_{tk}v_k'$$

lub w skróconym zapisie

$$\underline{u}^2 \in U^2(v_1'', v_k') \quad /241'/$$

Prawe strony /240/, /241/ zawierają różne zmienne koordynacyjne, co stanowi właśnie wspomnianą wyżej rozdzielność tych ograniczeń. Fakt ten znacznie ułatwia rozwiązywanie zadań tego rodzaju przy użyciu metod Lagrange'owskich, jak to omawiamy nieco dalej.

W przypadku nieliniowym ograniczenie $\underline{u} \in U$ nie może być przedstawione zależnością typu /236/. Tym niemniej można uzyskać dekompozycję zadania przez wprowadzenie zmiennych koordynacyjnych na miejsce niektórych zmiennych zadania pierwotnego. Rozpatrzmy na przykład zadanie o ograniczeniach

$$\begin{aligned} a_{11}u_1u_2 + a_{12}u_2 + \sum_{i=3}^k a_{1i}u_i &\geq b_1 \\ a_{21}u_1u_2 + \sum_{i=k+1}^n a_{2i}u_i &\geq b_2 \end{aligned} \quad /242/$$

Usuwamy zmienną u_1 w obu ograniczeniach wprowadzając na to miejsce odpowiednio v_1' , v_1'' oraz dodatkowe wymaganie

$$v_1' - v_1'' = 0 \quad /243/$$

Rezultatem jest

$$(a_{11}v_1' + a_{12})u_2 + \sum_{i=3}^k a_{1i}u_i \geq b_1$$

$$a_{21}v_1'' u_2 + \sum_{i=k+1}^n a_{2i} u_i \geq b_2 \quad /244/$$

Ograniczenia są teraz liniowe względem \underline{u} , lecz nadal sprzężone poprzez zmienną u_2 . Połączenie to usuniemy wprowadzając v_2', v_2'' w miejsce u_2 . W rezultacie otrzymujemy układ

$$\sum_{i=3}^k a_{1i}u_i \geq b_1 - (a_{11}v_1' + a_{12}) v_2'$$

$$\sum_{i=k+1}^n a_{2i}u_i \geq b_2 - a_{21}v_1'' v_2'' \quad /245/$$

oraz ograniczenia dla drugiego poziomu

$$v_1' - v_1'' = 0,$$

$$v_2' - v_2'' = 0 \quad /246/$$

Zauważmy, że ograniczenia /245/ są rozdzielone i liniowe, ponadto ich prawe strony zawierają różne zmienne koordynacyjne.

Separację ograniczeń /242/ można również uzyskać wprowadzając mniejszą liczbę zmiennych, na przykład tylko jedną zmienną v dla zastąpienia iloczynu $u_1 u_2$. Otrzymamy

$$\sum_{i=2}^k a_{1i}u_i \geq b_1 - a_{12}v$$

$$\sum_{i=k+1}^n a_{2i}u_i \geq b_2 - a_{21}v \quad /247/$$

Prawe strony /247/ zawierają teraz tę samą zmienną koordynacyjną, a ponadto ograniczenia na v , które muszą być uwzględnione przy rozwiązywaniu zadania drugiego poziomu, wymagają znajomości rozwiązań parametrycznych z poziomu pierwszego $\hat{u}_1(v)$. Ograniczenia te będą brzmieć

$$a_{12}v \geq b_1 - \sum_{i=2}^k a_{1i} \hat{u}_i(v)$$

$$a_{21}v \geq b_2 - \sum_{i=k+1}^n a_{2i} \hat{u}_i(v) \quad /248/$$

Rozwiązywanie zadań z ograniczeniami /247/, /248/ może być bardziej złożone, niż zadań ograniczonych przez /245/, /246/. Może być zatem racjonalne wprowadzać więcej zmiennych koordynacyjnych niż niezbędne minimum.

Dekompozycja ograniczeń $\underline{u} \in U$ w obu rozpatrywanych przypadkach, liniowym i nieliniowym, opierała się na częściowym odseparowaniu zmiennych już w zadaniu pierwotnym, przez co droga podziału była częściowo wskazana. Sytuacja taka nie zawsze ma miejsce - na przykład może istnieć jedno ograniczenie o postaci

$$h(u_1, u_2, \dots, u_n) \geq b \quad /249/$$

Dogodny sposób dekompozycji powstaje wówczas przez wprowadzenie zmiennych "agregowanych"

$$v_1 = g_1(u_1, u_2, \dots, u_1)$$

$$v_2 = g_2(u_{1+1}, u_{1+2}, \dots, u_k) \quad /250/$$

...

$$v_N = g_N(u_{k+\alpha}, \dots, u_n)$$

tak żeby ograniczenie /249/ przybrało postać

$$h(v_1, v_2, \dots, v_N) \geq b \quad /249/$$

Ograniczenia równościowe /250/ będą stosowane do zadań częściowych, ograniczenie /249/ - do zadania drugiego poziomu.

Zawsze należy pamiętać o tym, że zgodnie z /228/ dekompozycja $Q(\underline{u})$ oraz podział ograniczeń $\underline{u} \in U$ musi się opierać na tych samych rozłącznych podzbiorach $\underline{u}^i \subset \underline{u}$. Należy zazwyczaj rozpoczynać próbę podziału od dekompozycji ograniczeń, a następnie sprawdzić czy proponowany podział \underline{u} na \underline{u}^i prowadzi do takiej postaci wskaźnika jakości

$$Q = Q [Q_1(\underline{u}^1, \underline{v}), Q_2(\underline{u}^2, \underline{v}), \dots, Q_N(\underline{u}^N, \underline{v})]$$

która pozwala na rozłączną ekstremalizację.

8.3. Podstawowe metody rozwiązywania

Wzór /230/ wyrażał zadania optymalizacji pierwszego poziomu, z parametrycznym wynikiem $Q_1(\underline{v})$. Optymalne decyzje dla pierwszego poziomu byłyby również parametryczne względem \underline{v} :

$$\hat{u}^1(\underline{v}), \hat{u}^2(\underline{v}), \dots, \hat{u}^N(\underline{v}) \quad /251/$$

Wzór /231/ wyrażał zadanie optymalizacji drugiego poziomu. Rozwiązaniem tego zadania będzie $\underline{v} = \hat{\underline{v}}$, które z kolei określi wartości $\hat{u}^1, \hat{u}^2, \dots$, zgodnie z /251/. Zadania zarówno pierwszego jak drugiego poziomu mogą być rozwiązywane jakakolwiek odpowiednią dla nich metodą. Jeżeli zadanie pierwszego poziomu /230/ oraz zadanie drugiego poziomu /231/ są powiązane ze sobą wprost przez zmienną \underline{v} , mówimy że rozwiązanie wielopoziomowe jest "bezpośrednie". Przykłady takich rozwiązań podajemy dalej.

W dziedzinie programowania nieliniowego, Rosen i inni [6] opracowali metody iteracyjne pod nazwą "partition programming", przy pomocy których można sposobem bezpośrednim rozwiązywać zadania częściowo nieliniowe wypukłe o postaci

$$\min \left[Q(\underline{u}, \underline{v}) = \sum_{i=1}^N c_i^T(\underline{v}) \underline{u}^i + \varphi(\underline{v}) \right] \quad /252/$$

przy ograniczeniach

$$A_1(\underline{v}) \underline{u}^1 \geq \underline{b}_1(\underline{v}), \quad i = 1, 2, \dots, N \quad /253/$$

Metody te mogą być przydatne w odpowiednich przypadkach. Zauważmy, że wskaźnik jakości /252/ jest liniowy względem \underline{u} oraz nieliniowy względem \underline{v} . Ograniczenia /253/ są liniowe względem \underline{u} , lecz zarówno macierz A jak wektor B mogą być funkcjami \underline{v} /porównaj na przykład, zależności /247/, /248//. Przyjmuje się, że macierz A jest blokowo diagonalna, tak że ograniczenia zadania są opisane przez układ nierówności /253/, w które wchodzi macierze zadań częściowych A_1 . Jak już wskazywaliśmy, wiele zadań można podzielić i wprowadzić do postaci /253/ przez wprowadzenie dodatkowych zmiennych i dodatkowych ograniczeń. Możemy jednak nie mieć wpływu na postać $Q(\underline{u})$ tak, by sprowadzić ten wskaźnik do postaci $Q(\underline{u}, \underline{v})$ wskazanej przez /252/.

Metodę "Lagrange'owską" zamiast bezpośredniej możemy ze skutkiem zastosować w przypadku następującym. Przyjmujemy, że ograniczenie drugiego poziomu $\underline{v} \in V$ jest zbiorem ograniczeń równościowych liniowych względem elementów $v_1', v_1'', v_2', \dots, v_{N-1}''$ wektora koordynacyjnego \underline{v} :

$$\varepsilon_1(\underline{v}) = \alpha'_1 v'_1 + \alpha''_1 v''_1 + \alpha'_2 v'_2 + \dots + \alpha''_{N-1} v''_{N-1} = 0$$

/254/

$$\varepsilon_{N-1}(\underline{v}) = \gamma'_1 v'_1 + \gamma''_1 v''_1 + \gamma'_2 v'_2 + \dots + \gamma''_{N-1} v''_{N-1} = 0$$

W szczególnym przypadku, ograniczenia drugiego poziomu mogą mieć po prostu postać /por. 238/

$$\varepsilon_1(\underline{v}) = v'_1 - v''_1 = 0$$

$$\varepsilon_2(\underline{v}) = v'_2 - v''_2 = 0$$

...

$$\varepsilon_{N-1}(\underline{v}) = v'_{N-1} - v''_{N-1} = 0$$

/255/

Dla zadania drugiego poziomu, patrz /231/, możemy napisać funkcję Lagrange'a i określić rozwiązanie \underline{v} jako jej punkt siodłowy

$$\min_{\lambda} \max_{\underline{v}} L(\underline{v}, \lambda) = \min_{\lambda} \max_{\underline{v}} \left\{ Q[\hat{Q}_1(\underline{v}), \hat{Q}_2(\underline{v}), \dots, \hat{Q}_N(\underline{v})] + \lambda^T \underline{g}(\underline{v}) \right\} \quad /256/$$

Założmy teraz, że wskaźnik jakości jest addytywny /por. 232/ i zależy od rozłącznych części \underline{v} /jeśli w ogóle zależy jawnie od \underline{v} /:

$$Q(\underline{u}) = Q_1(\underline{u}^1, v'_1) + Q_2(\underline{u}^2, v''_1, v'_2) + \dots + Q_N(\underline{u}^N, v''_{N-1}) \quad /257/$$

oraz że wszystkie ograniczenia w zadaniach częściowych są rozdzielne względem zmiennych koordynacyjnych /por. 240/

$$h_1(\underline{u}^1, v'_1) \geq 0$$

$$h_2(\underline{u}^2, v''_1, v'_2) \geq 0$$

...

$$h_N(\underline{u}^N, v''_{N-1}) \geq 0$$

/258/

Przyjeliśmy tu, dla uproszczenia rozważań, że każde z zadań częściowych ma jedno /skalarne/ ograniczenie. Wobec /258/ $\hat{Q}_1(\underline{v})$ staje się $\hat{Q}_1(v'_1)$, $\hat{Q}_2(\underline{v})$ staje się $\hat{Q}_2(v''_1, v'_2)$ i tak dalej, zatem uwzględniając /255/, /257/ i /258/ wyrażenie /256/ przybiera postać szczególną

$$\min_{\lambda} \max_{\underline{v}} L(\underline{v}, \lambda) = \min_{\lambda} \left\{ \max_{v'_1} [\hat{Q}_1(v'_1) + \lambda_1 v'_1] + \max_{v''_1, v'_2} [\hat{Q}_2(v''_1, v'_2) - \lambda_1 v''_1 + \lambda_2 v'_2] + \dots + \max_{v''_{N-1}} [\hat{Q}_N(v''_{N-1}) - \lambda_{N-1} v''_{N-1}] \right\} \quad /259/$$

Wykorzystaliśmy tu zarówno addytywność i rozłączność /257/ względem elementów \underline{v} , jak charakter zależności / /254/ wzgl. /255/. Postać /259/ wskazuje, że maksymalizacja wyrazów

w nawiasach kwadratowych względem elementów \underline{v} może być przesunięta do zadań częściowych, które będą miały na przykład postać

$$\max_{\underline{v}'_1} [\hat{Q}_1(\underline{v}'_1) + \lambda_1 \underline{v}'_1] = \max_{\underline{v}'_1} \left[\max_{\underline{u}'_1 \in U'_1(\underline{v}'_1)} Q_1(\underline{u}'_1, \underline{v}'_1) + \lambda_1 \underline{v}'_1 \right] \quad /260/$$

Rozwiązania zadań częściowych będą parametryczne względem $\underline{\lambda}$, a nie względem \underline{v} :

$$\hat{\underline{u}}^1(\lambda_1), \hat{\underline{v}}^1(\lambda_1), \hat{\underline{u}}^2(\lambda_1, \lambda_2) \quad \text{itd} \quad /261/$$

Zauważmy, że $\underline{\lambda}$ są to zmienne dodatkowe w zadaniu, zmienne koordynacyjne \underline{v} muszą być użyte również, aby zapewnić potrzebną separację zadań częściowych. Korzyść z użycia $\underline{\lambda}$ polega na tym, że w pewnych przypadkach rozwiązania /161/ zawierają mniej parametrów niż rozwiązania /251/. Istnieją efektywne metody iteracyjne dla rozwiązywania zadań typu /256/, przy założeniach takich jak /254/, /257/ i /258/.

Przykłady

Rozpatrzmy zadanie programowania wypukłego, które można nawiązać do układu trzech zbiorników wodnych na rzece, rys. 26. Każdy ze zbiorników zasila trzech odbiorców. Wpływy wody oznaczone są u_1, u_2, \dots, u_9 . Zapas wody wynosi odpowiednio z_1, z_2, z_3 . Zbiornik 1 może oddać ilość wody u_{10} do zbiornika 2, zbiornik 2 ilość wody u_{11} do zbiornika 3. Wskaźnik jakości, który należy zminimalizować ma postać

$$Q(\underline{u}) = \sum_{i=1}^9 (a_i - u_i)^2 \quad /262/$$

a decyzje u_i są ograniczone przez nierówności, wynikające z bilansu wody

$$h_1(\underline{u}) = u_1 + u_2 + u_3 + u_{10} - z_1 \leq 0$$

$$h_2(\underline{u}) = u_4 + u_5 + u_6 - u_{10} + u_{11} - z_2 \leq 0 \quad /263/$$

$$h_3(\underline{u}) = u_7 + u_8 + u_9 - u_{11} - z_3 \leq 0$$

oraz przez wymaganie, by wszystkie przepływy były dodatnie

$$u_i \geq 0, \quad i = 1, 2, \dots, 11 \quad /264/$$

Dekompozycja zadania /262/, /263/ będzie podyktowana raczej przez układ ograniczeń /263/, bowiem /262/ ma postać czysto addytywną i może być podzielone dowolnie. Jedną z możliwości jest wprowadzenie dwóch zmiennych koordynacyjnych, $\underline{v}_1 = u_{10}$

oraz $v_2 = u_{11}$, tak by utworzyć trzy zadania częściowe, z ograniczeniami następującymi

$$h_1(\underline{u}^1, v_1) = u_1 + u_2 + u_3 + v_1 - z_1 \leq 0, \quad u_1 \geq 0$$

$$h_2(\underline{u}^2, v_1, v_2) = u_4 + u_5 + u_6 - v_1 + v_2 - z_2 \leq 0, \quad u_1 \geq 0 \quad /265/$$

$$h_3(\underline{u}^3, v_2) = u_7 + u_8 + u_9 - v_2 - z_3 \leq 0, \quad u_1 \geq 0$$

Wskaźnik jakości /262/ będzie wówczas podzielony jak następuje

$$Q(\underline{u}) = Q_1(\underline{u}^1) + Q_2(\underline{u}^2) + Q_3(\underline{u}^3) \quad /266/$$

gdzie

$$Q_1(\underline{u}^1) = \sum_{i=1}^3 (a_i - u_i)^2$$

$$Q_2(\underline{u}^2) = \sum_{i=4}^6 (a_i - u_i)^2 \quad /267/$$

$$Q_3(\underline{u}^3) = \sum_{i=7}^9 (a_i - u_i)^2$$

Po rozwiązaniu zadań częściowych, zadanie drugiego poziomu będzie

$$\min_{v_1, v_2} [\hat{Q}_1(v_1) + \hat{Q}_2(v_1, v_2) + Q_3(v_2)] \quad /268/$$

Ażeby otrzymać $\hat{Q}_1(v_1)$, trzeba rozwiązać zadanie częściowe pierwsze. Weźmy w tym celu wskaźnik jakości $Q_1(\underline{u}^1, v_1)$ oraz pierwszą nierówność z /265/ i utwórzmy funkcję Lagrange'a

$$L(u_1, u_2, u_3, \lambda) = (a_1 - u_1)^2 + (a_2 - u_2)^2 + (a_3 - u_3)^2 + \lambda (u_1 + u_2 + u_3 + v_1 - z_1) \quad /269/$$

Poszukiwanie punktu siodłowego /269/ daje wyrażenia

$$\begin{aligned} \hat{u}_1 &= a_1 - \Delta a = f_1(v_1) \\ \hat{u}_2 &= a_2 - \Delta a = f_2(v_1) \\ \hat{u}_3 &= a_3 - \Delta a = f_3(v_1) \\ \hat{\lambda} &= 2 \Delta a \end{aligned} \quad /270/$$

gdzie

$$\Delta a = \frac{1}{3} (a_1 + a_2 + a_3 + v_1 - z_1) \quad /271/$$

przy czym rozwiązanie /270/ są punktem siodłowym funkcji Lagrange'a /269/, czyli spełniają warunki Kuhna-Tuckera, w obszarze

$$0 \leq \hat{u}_1 \leq a_1$$

$$0 \leq \hat{u}_2 \leq a_2$$

$$0 \leq \hat{u}_3 \leq a_3$$

/272/

$$z_1 - v_1 \leq a_1 + a_2 + a_3 = \alpha_1$$

lub, inaczej to określając, w obszarze

$$z_1 - \alpha_1 \leq v_1 \leq z_1 - \alpha_1 + 3 a_1$$

/273/

gdzie a_1 jest najmniejszym spośród a_1 tego zadania częściowego.

Wykres rozwiązań /270/ podaje rys. 27. Zgodnie z /272/ rozwiązania /270/ są ważne w obszarze $2 \leq v_1 \leq 5$ dla przykładowych wartości a_1, z_1 przyjętych dla rys.27. dla $v_1 < 2$ optymalne wartości u_1 minimalizujące Q_1 są to oczywiście $u_1 = a_1$, a dla $v_1 > 5$ trzeba rozwiązać nowe zadanie optymalizacji:

$$\min [Q_1(\underline{u}^1) = (a_2 - u_2)^2 + (a_3 - u_3)^2]$$

/274/

z ograniczeniem $h_1(\underline{u}^1, v_1)$ tym co poprzednio /265/ oraz z wartością $\hat{u}_1 = 0$. Rozwiązanie jest łatwe i będzie obowiązywać w przedziale $5 \leq v_1 \leq 7$. Przy $v_1 = 7$ osiągamy $\hat{u}_2 = 0$, zatem dla $v_1 > 7$ zadanie optymalizacji brzmi

$$\min [Q_1(\underline{u}^1) = (a_3 - u_3)^2]$$

/275/

przy ciągle tym samym ograniczeniu $h_1(\underline{u}^1, v_1)$ według /265/ oraz z wartościami $\hat{u}_1 = 0, \hat{u}_2 = 0$. Wszystkie wyniki $\hat{u}_1(v_1)$ są wykreślone na rys.27.

Musimy teraz obliczyć $\hat{Q}_1(v_1)$. Wykorzystując /270/ dojdziemy po paru przekształceniach do

$$\hat{Q}_1(v_1) = \frac{1}{3} (\alpha_1 + v_1 - z_1)^2$$

/276/

przy czym wzór ten obowiązuje w obszarze /273/.

Biorąc poza tym obszarem rozwiązania \hat{u}_1 wskazane na rys. 27, możemy przedstawić $\hat{Q}_1(v_1)$, dla rozpatrywanego przykładu, jak pokazano na rys.28. Zauważmy że Q_1 jest zerem dla $v_1 < z_1 - \alpha_1$ oraz jest wyrażeniem kwadratowym /276/ w przedziale określonym przez /273/. Gdybyśmy chcieli uprościć nieco zadanie drugiego poziomu, możnaby przyjąć /276/ jako przybliżenie dla wszyst-

kich $v_1 \geq z_1 - \alpha_1$, otrzymując w ten sposób

$$\hat{Q}_1(v_1) = \begin{cases} \frac{1}{3} (d_1 + v_1 - z_1)^2 & \text{jeśli } v_1 \geq z_1 - \alpha_1 \\ 0 & \text{w obszarze pozostałym} \end{cases} \quad /277/$$

oraz podobnie dla pozostałych dwóch zadań częściowych:

$$\hat{Q}_2(v_1, v_2) = \begin{cases} \frac{1}{3} (\alpha_1 - v_1 + v_2 - z_2)^2 & \text{jeśli } v_1 - v_2 \leq \alpha_2 - z_2 \\ 0 & \text{w obszarze pozostałym} \end{cases} \quad /278/$$

$$\hat{Q}_3(v_2) = \begin{cases} \frac{1}{3} (\alpha_3 - v_2 - z_3)^2 & \text{jeśli } v_2 \leq \alpha_3 - z_3 \\ 0 & \text{w obszarze pozostałym} \end{cases} \quad /279/$$

Odpowiednie wykresy dla wybranych danych liczbowych przedstawiono na rys.29 i rys.30.

Funkcja drugiego poziomu polega na wykonaniu

$$\min_{v_1, v_2} [\hat{Q}_1(v_1) + \hat{Q}_2(v_1, v_2) + \hat{Q}_3(v_2)] \quad /280/$$

z ograniczeniami pochodzącymi z bilansów

$$\begin{aligned} v_1 - z_1 &\leq 0 \\ v_2 - v_1 - z_2 &\leq 0 \end{aligned} \quad /281/$$

oraz warunkami na nieujemność

$$v_1 \geq 0, \quad v_2 \geq 0 \quad /282/$$

Zauważmy, że ograniczenia /281/ można wyprowadzić wprost z problemu fizycznego lub też otrzymać z /265/ przez przyjęcie dla u_1, u_2, \dots, u_9 ich wartości krańcowych $u_1 = 0$.

Zadanie /280/ z warunkami /281/, /282/ może być rozwiązane jakąkolwiek odpowiednią metodą. Rozpatrzmy, na przykład, wprowadzenie trzeciego poziomu sterowania, tworząc na poziomie drugim zadanie parametryczne

$$\min_{v_1} [\hat{Q}_1(v_1) + \hat{Q}_2(v_1, v_2)] = \hat{Q}_{12}(v_2) \quad /283/$$

z ograniczeniem $0 \leq v_1 \leq z_1$, oraz pozostawiając dla poziomu trzeciego ostateczną optymalizację

$$\min_{\underline{u}} Q(\underline{u}) = \min_{v_2} [\hat{Q}_{12}(v_2) + \hat{Q}_3(v_2)] \quad /284/$$

z ograniczeniem $0 \leq v_2 \leq z_2 + z_1$.

Obliczenia wskazane przez /283/ i /284/ mogą być, w rozpatrywanym przykładzie, łatwo przeprowadzone. Wyniki są podane na rys.31. Dla decyzji na trzecim poziomie trzeba dodać $Q_{12}(v_2)$ do $\hat{Q}_3(v_2)$ wziętego z rys.30 i stwierdzić położenie minimum tej sumy. Wynik brzmi $\hat{v}_2 = \frac{4}{3}$, a wskaźnik jakości $\hat{Q}(u) = 2\frac{7}{9}$.

Dalej należy wziąć wartość $\hat{v}_2 = \frac{4}{3}$, użyć wykresu $\hat{v}_1(v_2)$ z rys.31 aby stwierdzić, że $\hat{v}_1 = \frac{3\frac{2}{3}}{2}$, zastosować wykres z rys. 27 dla określenia $\hat{u}_1 = \frac{4}{9}$, $\hat{u}_2 = 1\frac{4}{9}$, $\hat{u}_3 = 2\frac{4}{9}$, oraz postąpić podobnie dla rozwiązań $\hat{u}_4 \div \hat{u}_9$ /odnośnych wykresów nie pokazujemy/. Szkic struktury rozwiązania podano na rys.32.

Jako drugi przykład rozpatrzmy zadanie, w którym wskaźnik jakości ma postać iloczynową.

Dane jest zadanie maksymalizacji

$$\max [Q(u) = Q_1(u_1, u_2, u_3) \cdot Q_2(u_3, u_4, u_5)] \quad /285/$$

gdzie

$$Q_1(u_1, u_2, u_3) = \alpha^2 - (a_1 - u_1)^2 - (a_2 - u_2)^2 - (a_3 - u_3)^2 \geq 0 \quad /286/$$

$$Q_2(u_3, u_4, u_5) = \beta^2 - (a_3 - u_3)^2 - (a_4 - u_4)^2 - (a_5 - u_5)^2 \geq 0$$

z ograniczeniami $u_1 \geq 0$ oraz

$$h_1(u_1, u_2, u_3) = z_1 - u_1 - u_2 - u_3 \geq 0$$

$$h_2(u_3, u_4, u_5) = z_2 - u_3 - u_4 - u_5 \geq 0 \quad /287/$$

$$h_3(u_1, u_2, u_3, u_4, u_5) = z_3 - u_1 - u_2 - u_3 - u_4 - u_5 \geq 0$$

Zadanie to może być ilustrowane schematem z rys. 33. Wprowadzając nową zmienną $v' = u_3$ w funkcje Q_1, h_1 oraz zmienną $v'' = u_3$ w funkcje Q_2, h_2 wraz z nowym ograniczeniem

$$v' - v'' = 0 \quad /288/$$

możemy utworzyć dwa zadania częściowe.

Zadanie częściowe 1

$$\max_{u_1, u_2} Q_1(u_1, u_2, v') = \max_{u_1, u_2} [\alpha^2 - (a_1 - u_1)^2 - (a_2 - u_2)^2 - (a_3 - v')^2] \quad /289/$$

przy ograniczeniu

$$h_1(u_1, u_2, v') = z_1 - u_1 - u_2 - v' \geq 0 \quad /290/$$

z rozwiązaniami parametrycznymi

$$\hat{Q}_1(v'), \hat{u}_1(v'), \hat{u}_2(v')$$

Zadanie częściowe 2

$$\max_{u_4, u_5} Q_2(u_4, u_5, v'') = \max_{u_4, u_5} [\beta^2 - (a_4 - u_4)^2 - (a_5 - u_5)^2 - (a_3 - v'')^2] \quad /291/$$

przy ograniczeniu

$$h_2(u_4, u_5, v'') = z_2 - u_4 - u_5 - v'' \geq 0 \quad /292/$$

z rozwiązaniami parametrycznymi

$$\hat{Q}_2(v''), \hat{u}_4(v''), \hat{u}_5(v'')$$

Rozwiązania zadań częściowych wchodzą w zadanie drugiego poziomu

$$\max_{v', v''} \hat{Q}_1(v') \cdot \hat{Q}_2(v'') \quad /293/$$

z ograniczeniem

$$v' - v'' = 0 \quad /294/$$

oraz

$$h_3(u_1, u_2, v', u_4, u_5) = z_3 - u_1 - u_2 - v' - u_4 - u_5 \geq 0 \quad /295/$$

lub też, co da ten sam rezultat, przy ograniczeniu otrzymanym z h_3 w założeniu $u_1 = 0$:

$$z_3 - v' \geq 0 \quad /296/$$

Rozwiązaniem zadania drugiego poziomu będą \hat{v}' , \hat{v}'' .

Zadanie częściowe 1 rozwiążemy szukając punktu siodłowego funkcji Lagrange'a

$$L(u_1, u_2, \lambda) = \alpha^2 - (a_1 - u_1)^2 - (a_2 - u_2)^2 - (a_3 - v')^2 + \lambda (z_1 - u_1 - u_2 - v') \quad /279/$$

Sprawdzenie warunków Kuhna-Tuckera daje jako współrzędne punktu siodłowego

$$\hat{u}_1 = \frac{1}{2} (z_1 - v' + a_1 - a_2)$$

$$\hat{u}_2 = \frac{1}{2} (z_1 - v' - a_1 + a_2) \quad /298/$$

$$\hat{\lambda} = a_1 + a_2 - z_1 + v'$$

przy czym rozwiązanie /298/ obowiązuje w obszarze

$$0 \leq u_1 \leq a_1$$

$$0 \leq u_2 \leq a_2 \quad /299/$$

albo, gdy przyjmą że $a_2 \geq a_1$, w obszarze

$$z - a_1 - a_2 \leq v' \leq z_1 + a_1 - a_2 \quad /300/$$

W tak określonym obszarze $\hat{Q}_1(v')$ można obliczyć jako

$$\hat{Q}_1(v') = \alpha^2 - (a_1 - \hat{u}_1)^2 - (a_2 - \hat{u}_2)^2 - (a_3 - v')^2 \quad /301/$$

co daje

$$\hat{Q}_1(v') = \alpha^2 - \frac{1}{2} (a_1 + a_2 - z_1 + v')^2 - (a_3 - v')^2 \quad /302/$$

Nie rozpatrzyliśmy, dla oszczędności miejsca, rozwiązań właściwych dla obszaru poza określonym przez /300/.

Należałoby rozwiązania te znaleźć w sposób podobny jak w przykładzie poprzednim.

Dla drugiego zadania częściowego utworzymy funkcję Lagrange'a

$$L(u_4, u_5, \mu) = \beta^2 - (a_4 - u_4)^2 - (a_5 - u_5)^2 - (a_3 - v'')^2 + \mu(z_2 - u_4 - u_5 - v'') \quad /303/$$

która ma punkt siodłowy

$$\hat{u}_4 = \frac{1}{2} (z_2 - v'' + a_4 - a_5)$$

$$\hat{u}_5 = \frac{1}{2} (z_2 - v'' - a_4 + a_5) \quad /304/$$

$$\hat{\mu} = a_4 + a_5 - z_2 + v''$$

w obszarze określonym przez

$$0 \leq u_4 \leq a_4$$

$$0 \leq u_5 \leq a_5$$

/305/

albo inaczej, gdy $a_5 \geq a_4$, w obszarze

$$z_2 - a_4 - a_5 \leq v'' \leq z_2 + a_4 - a_5$$

/306/

Wynik optymalizacji brzmi w tym obszarze

$$\hat{Q}_2(v'') = \beta^2 - \frac{1}{2} (a_4 + a_5 - z_2 + v'')^2 - (a_3 - v'')^2 \quad /307/$$

Ażeby z kolei rozwiązać zadanie drugiego poziomu

$$\max_{v', v''} \hat{Q}_1(v') \cdot \hat{Q}_2(v'')$$

z ograniczeniami /294/, /296/ utworzymy funkcję Lagrange'a

$$L(v', v'', \gamma_1, \gamma_2) = \left[\alpha^2 - \frac{1}{2} (a_1 + a_2 - z_1 + v')^2 - (a_3 - v')^2 \right] \quad /308/$$

$$\left[\beta^2 - \frac{1}{2} (a_4 + a_5 - z_2 + v'')^2 - (a_3 - v'')^2 \right] + \gamma_1 (z_3 - v') + \gamma_2 (v' - v'')$$

Stosując warunki Kuhna-Tuckera znajdujemy punkt siodłowy tej funkcji; przy danych liczbowych następujących

$$\begin{array}{lll} a_1 = a_4 = 2 & z_1 = 8 & \alpha^2 = 100 \\ a_2 = a_5 = 4 & z_2 = 8 & \beta^2 = 100 \\ a_3 = 3 & z_3 = 12 & \end{array}$$

rozwiązanie zadania brzmi

$$\hat{v}' = \hat{v}'' = \hat{u}_3 = \frac{8}{3}$$

oraz w ślad za tym, użytkując /298/, /304/ otrzymujemy

$$\hat{u}_1 = \frac{5}{3}, \hat{u}_2 = \frac{11}{3}, \hat{u}_4 = \frac{5}{3}, \hat{u}_5 = \frac{11}{3}$$

Dwupoziomowa struktura rozwiązywania zadania, jaką tu zastosowaliśmy, przedstawiona jest schematycznie na rys.34.

8.4. Metody obliczeniowe

Przedstawione w poprzednim punkcie przykłady rozwiązane były na drodze analitycznej. W zadaniach o wymiarowości i stopniu skomplikowania odpowiadających praktyce, rozwiązywanie analityczne nie jest możliwe i sięgać trzeba do metod numerycznych. Stosowane tu metody iteracyjne będą zbliżone do metod szukania ekstremum, stosowanych w optymalizacji statycznej. Nie będziemy rozwijać w niniejszym tekście zagadnień dynamicznych, aczkolwiek rozwiązywanie ich numerycznie w układach wielopoziomowych jest również znane.

Rozpatrzmy zadanie, rozwiązywane dwupoziomowo metodą Lagrange'owską. Przypomnijmy /patrz równania /254/ do /261/, że gdy ograniczenia drugiego poziomu są liniowe, wskaźnik jakości jest addytywny i ponadto każda jego część zależy od innej części wektora zmiennych koordynacyjnych \underline{v} /jeśli w ogóle zależy od \underline{v} /, to zadania pierwszego poziomu i drugiego poziomu są ze sobą skojarzone poprzez mnożniki Lagrange'a.

Weźmy dla ilustracji przykład zadania, sformułowanego przez związki /262/, /263/, /264/ i rys.26. Dla zastosowania metody mnożników Lagrange'a wprowadzimy zmienne koordynacyjne

$$v'_1, v''_1 \text{ zamiast } u_{10} \quad /309/$$

$$v'_2, v''_2 \text{ zamiast } u_{11} \quad /310/$$

otrzymując zadania częściowe /por. /265/, /266/, /267//

$$1/ \min [Q_1(\underline{u}^1) = \sum_{i=1}^3 (a_i - u_i)^2] \quad /311/$$

z ograniczeniem

$$h_1(\underline{u}^1, v_1') = u_1 + u_2 + u_3 + v_1' - z_1 \leq 0, \quad u_i \geq 0 \quad /312/$$

$$2/ \min [Q_2(\underline{u}^2) = \sum_{i=4}^6 (a_i - u_i)^2] \quad /313/$$

z ograniczeniem

$$h_2(\underline{u}^2, v_1'', v_2') = u_4 + u_5 + u_6 - v_1'' + v_2' - z_2 \leq 0, \quad u_i \geq 0 \quad /314/$$

$$3/ \min [Q_3(\underline{u}^3) = \sum_{i=7}^9 (a_i - u_i)^2] \quad /315/$$

z ograniczeniem

$$h_3(\underline{u}^3, v_2'') = u_7 + u_8 + u_9 - v_2'' - z_3 \leq 0, \quad u_i \geq 0 \quad /316/$$

Założmy, że zadania częściowe rozwiązaliśmy, otrzymując parametryczne wyniki

$$\hat{Q}_1(v_1'), \quad \hat{Q}_2(v_1'', v_2'), \quad \hat{Q}_3(v_2'') \quad /317/$$

Zadanie drugiego poziomu polega na minimalizacji sumy

$$\min [\hat{Q}_1(v_1') + \hat{Q}_2(v_1'', v_2') + \hat{Q}_3(v_2'')] \quad /318/$$

z uwzględnieniem ograniczeń następujących

$$v_1' - v_1'' = 0$$

$$v_2' - v_2'' = 0$$

/319/

Napiszemy funkcję Lagrange'a dla zadania /318/, /319/

$$L(v_1', v_1'', v_2', v_2'', \lambda_1, \lambda_2) = \hat{Q}_1(v_1') + \hat{Q}_2(v_1'', v_2') + \hat{Q}_3(v_2'') + \lambda_1(v_1' - v_1'') + \lambda_2(v_2' - v_2'') \quad /320/$$

oraz przegrupujemy jej wyrazy, zapisując zarazem że poszukujemy punktu siodłowego

$$\max_{\underline{\lambda}} \min_{\underline{v}} [\hat{Q}_1(v_1') + \lambda_1 v_1' + \hat{Q}_2(v_1'', v_2') - \lambda_1 v_1'' + \lambda_2 v_2' + \hat{Q}_3(v_2'') - \lambda_2 v_2''] \quad /321/$$

Wykorzystamy możliwość rozłącznej ekstremalizacji

$$\max_{\lambda_1, \lambda_2} \left\{ \min_{v_1'} [\hat{Q}_1(v_1') + \lambda_1 v_1'] + \min_{v_1'', v_2'} [\hat{Q}_2(v_1'', v_2') - (\lambda_1 v_1'' + \lambda_2 v_2')] + \right.$$

$$+ \min_{v_2''} \left[\hat{Q}_3(v_2'') - \lambda_2 v_2'' \right] \quad /322/$$

Teraz przypomnijmy, że np. $\hat{Q}_1(v_1')$ oznaczało rezultat minimalizacji /311/ z ograniczeniem /312/ - możemy zatem pierwszy wyraz z /322/ wraz z odpowiednim zadaniem pierwszego poziomu zapisać jako jedną ekstremalizację:

$$\min_{v_1'} \left[\hat{Q}_1(v_1') + \lambda_1 v_1' \right] = \min_{v_1', u_1, u_2, u_3} \left[Q_1(u_1, u_2, u_3) + \lambda_1 v_1' \right] \quad /323/$$

z ograniczeniem

$$h_1(u_1, u_2, u_3, v_1') = u_1 + u_2 + u_3 + v_1' - z_1 \leq 0 \quad /324/$$

Nie jest istotne, czy ograniczenie /324/ uwzględnimy w operacji /323/ metodą utworzenia funkcji Lagrange'a czy w inny sposób. Istotne jest natomiast, że w nowopowstałym zadaniu częściowym ekstremalizujemy zmodyfikowany wskaźnik jakości

$$Q_1(u_1, u_2, u_3) + \lambda_1 v_1' \quad /325/$$

w porównaniu z samym tylko $Q_1(u_1, u_2, u_3)$ powstałym z dekompozycji wskaźnika globalnego. Składnik $\lambda_1 v_1'$ stanowi "wartość zmiennej koordynacyjnej". Łatwo zauważyć, że gdy w zadaniu /323/ $\lambda_1 v_1'$ jest dodatnie /koszt/, to w zadaniu drugim wystąpi $(-\lambda_1 v_1'')$ co oznacza wartość przeciwną /zysk/.

W ostatecznym rozwiązaniu zadania globalnego muszą być spełnione ograniczenia /319/ oraz spełnione wymaganie /322/, co oznacza właściwy dobór zarówno mnożników λ_1, λ_2 jak zmiennych v_1', v_2', v_1'', v_2'' .

Możliwe są dwa warianty iteracyjnego rozwiązania tego zagadnienia.

Pierwsza z nich, metoda doboru mnożników /zwana też metodą cen/, polega na zakładaniu wartości mnożników $\lambda_1, \lambda_2 \dots$ i na ulepszaniu ich w kolejnych iteracjach. Miarą niewłaściwości doboru mnożników są niezgodności

$$v_1' \neq v_1'', \quad v_2' \neq v_2'' \quad /326/$$

Nasuwa się zatem reguła postępowania iteracyjnego

$$\lambda_1^{i+1} = \lambda_1^i + \varrho (v_1' - v_1'')^i \quad /327/$$

$$\lambda_2^{i+1} = \lambda_2^i + \varrho (v_2' - v_2'')^i$$

...

Można łatwo ustalić znak współczynnika ϱ ; w zadaniu na minimalizację współczynnik ϱ winien być dodatni. Od wartości tego współczynnika zależy bieg procesu iteracyjnego, czego nie będziemy rozpatrywać.

W bardziej złożonych lub innych odpowiednich przypadkach można oczywiście niezgodność /326/ redukować przy pomocy jednej ze standardowych procedur poszukiwania ekstremum, stawiając zadanie

$$\min_{\lambda_1, \lambda_2, \dots} \left[k_1(v'_1 - v''_1)^2 + k_2(v'_2 - v''_2)^2 + \dots \right] \quad /328/$$

Wybór współczynników wagi k_1, k_2, \dots zależy musi od konkretnego zadania; szybkość zbieżności i nakład obliczeń dla rozwiązania zadania /328/ może być większy niż w przypadku posłużenia się metodą opartą na /327/.

Metoda doboru współrzędnych /zwana też metodą realizowalną/ rozwiązuje Lagrange'owskie zadanie dwupoziomowe w sposób również iteracyjny, lecz wychodząc z zakładania wartości zmiennych koordynacyjnych $v'_1 = v''_1, v'_2 = v''_2$. Nie ma tu, w żadnym etapie procedury, niezgodności /326/. Przyjmuje się natomiast, że współczynnik np. λ_1 ma początkowo inną wartość w zadaniu częściowym pierwszym, niż w zadaniu częściowym drugim. Zadanie częściowe np. pierwsze brzmi teraz /patrz różnicę względem /323//

$$\max_{\lambda'_1} \min_{u_1, u_2, u_3} \left[Q_1(u_1, u_2, u_3) + \lambda'_1 v_1 \right] \quad /329/$$

to znaczy, że przy narzuconym z drugiego poziomu v_1 otrzymujemy w rozwiązaniu wartość λ'_1 .

Postępowanie iteracyjne polega na redukcji różnic $\lambda'_1 \neq \lambda''_1$, gdzie λ''_1 pochodzi z rozwiązania drugiego zadania częściowego

$$\max_{\lambda''_1, \lambda'_2} \min_{u_4, u_5, u_6} \left[Q_2(u_4, u_5, u_6) - \lambda''_1 v_1 + \lambda'_2 v_2 \right] \quad /330/$$

Reguła postępowania może być następująca

$$v_1^{i+1} = v_1^i + \varrho (\lambda'_1 - \lambda''_1)$$

$$v_2^{i+1} = v_2^i + \varrho (\lambda'_2 - \lambda''_2) \quad /331/$$

...

Podobnie jak poprzednio można stwierdzić, że w zadaniu na minimum współczynnik ϱ powinien być dodatni.

Również w metodzie doboru współrzędnych można zastosować, zamiast /331/, standardową procedurę iteracyjną. Zadanie dla niej brzmi wówczas

$$\min_{v_1, v_2, \dots} [k_1 (\lambda_1' - \lambda_1'')^2 + k_2 (\lambda_2' - \lambda_2'')^2 + \dots] \quad /332/$$

Rozpatrywaliśmy, na gruncie przykładu, dwie iteracyjne metody właściwe dla metody Lagrange'owskiej rozwiązywania zadania dwupoziomowego. Cechą charakterystyczną było to, że dobór /mnożników λ_1 lub współrzędnych v_1 / odbywał się w oparciu o niezgodności /typu $v_1' \neq v_1''$ lub $\lambda_1' \neq \lambda_1''$ /.

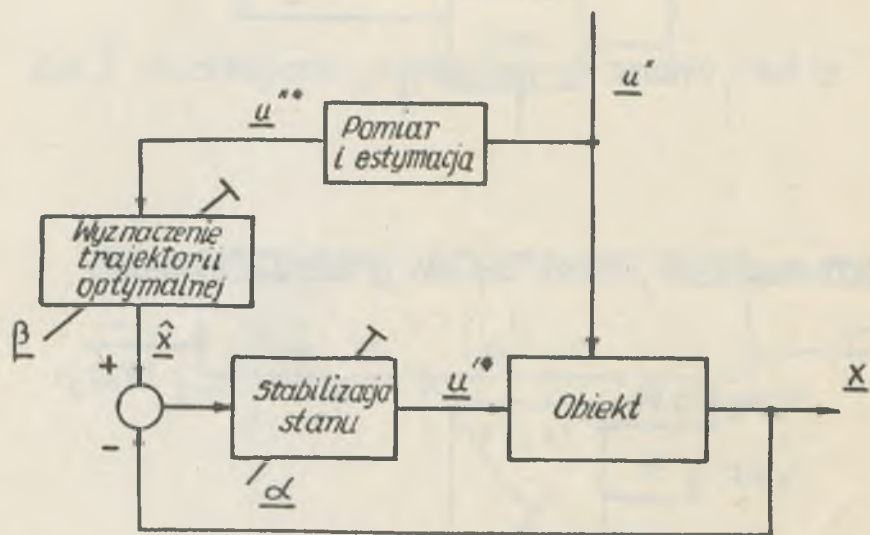
Procedury iteracyjne można również stosować do rozwiązywania zadania dwupoziomowego metodą bezpośrednią. Wróćmy do tego samego przykładu w rozwiązaniu takim, jak przedstawione w punkcie 8.3. Dla drugiego poziomu mamy tam zadanie /por.268/

$$\min_{v_1, v_2} [\hat{Q}_1(v_1) + \hat{Q}_2(v_1, v_2) + \hat{Q}_3(v_2)] \quad /333/$$

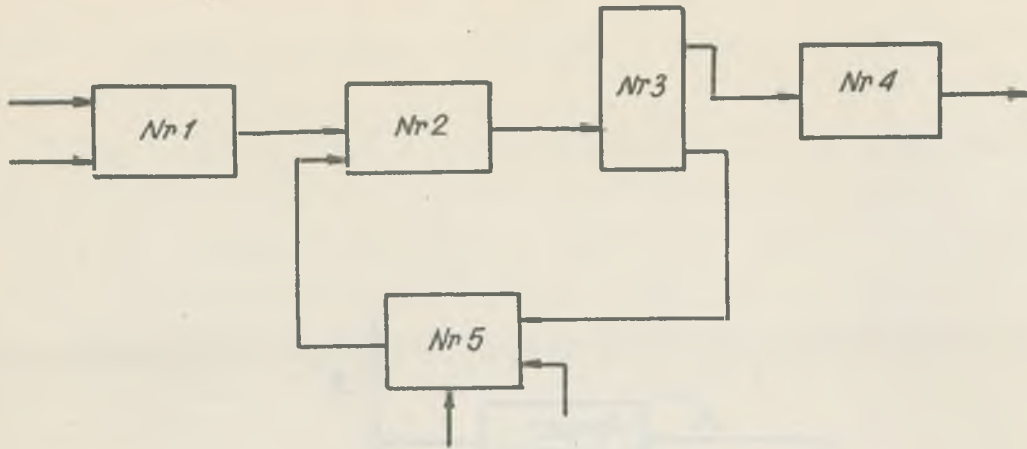
Do zadania /333/ można zastosować wprost odpowiednią procedurę szukania minimum. Zauważmy, że w każdym kroku procedury trzeba będzie obliczać nowe wartości $\hat{Q}_1(v_1)$, $\hat{Q}_2(v_1, v_2)$, ... czyli rozwiązać wszystkie zadania częściowe dla na nowo zadanych wartości v_1, v_2 . Uniknąć tego rozwiązywania nie można również w poprzednich metodach.

Literatura

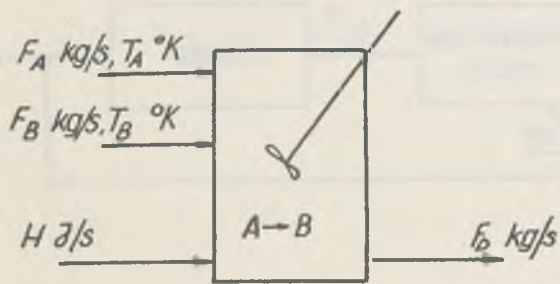
- 1 Findeisen W. Technika regulacji automatycznej. Wyd. 2, Warszawa 1969, PWN
- 2 Findeisen W., Lefkowitz I. Design and Applications of Multilayer Control, Proceedings the 4-th Congress of IFAC, June 1969, Warszawa, paper 42.1.
- 3 Findeisen W., Manitius A., Pułaczewski J. Multilevel Optimization and Dynamic Coordination of Mass Flows in a Beet Sugar Plant, 4-th Congress of IFAC, June 1969, Warszawa, paper 66.5.
- 4 Findeisen W. Modele przybliżone dla sterowania optymalnego i kryteria ich adaptacji. Archiwum Automatyki i Telemekhaniki, tom 13, Nr 1, 1968.
- 5 Lasdon L. Duality and Decomposition in Mathematical Programming. IEEE Trans. on Systems Science and Cybernetics, Vol. SSC-4, July 1968.
- 6 Rosen J.B., in "Recent Advances in Mathematical Programming", R.L. Graven and P. Wolfe, Eds. New York 1963, Mc.Graw Hill.
- 7 Wierzbicki A. Zasada maximum a synteza regulatorów optymalnych. Cz.III. Wrażliwość struktur. Archiwum Automatyki i Telemekhaniki, tom 14, Nr 1, 1969.
- 8 Zbiorowe, Poradnik inżyniera automatyka, Warszawa 1959. PWN.



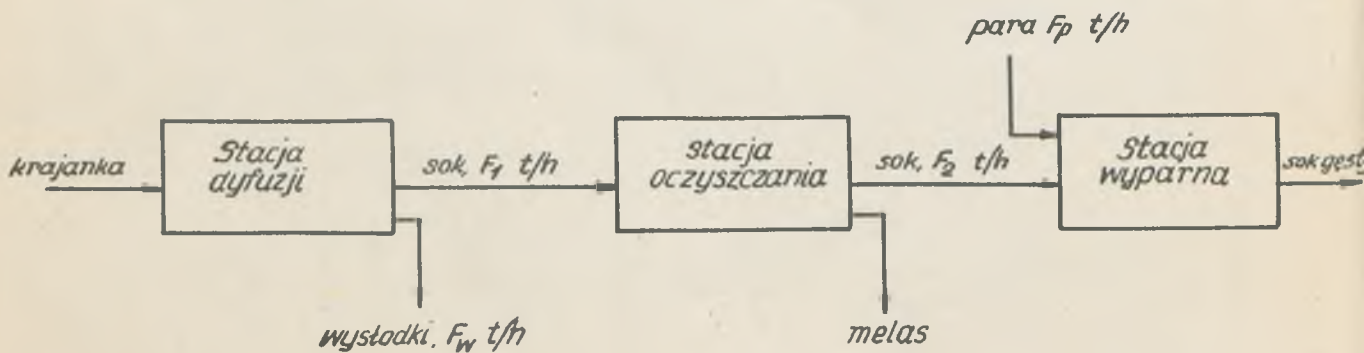
Rys.1 Przykład struktury układu



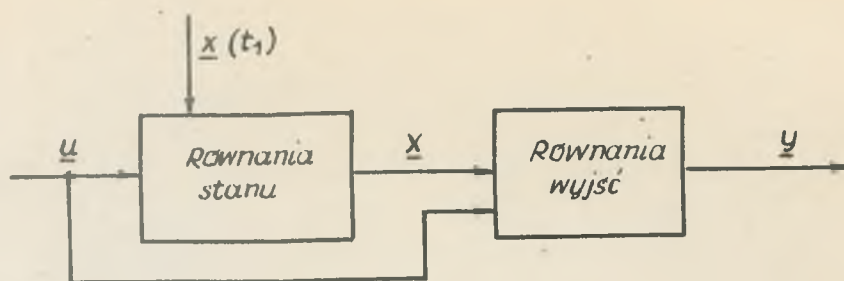
Rys. 2 Schematyczne przedstawienie procesu złożonego



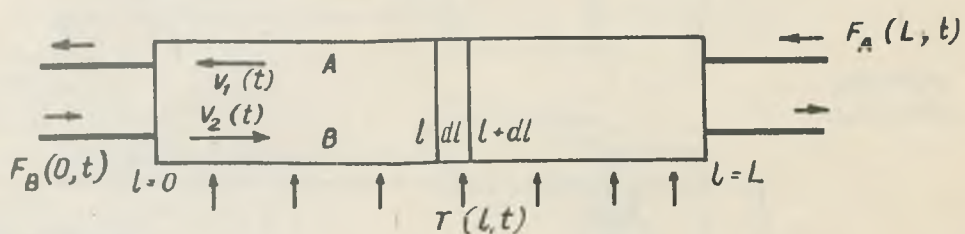
Rys. 3 Schemat reaktora



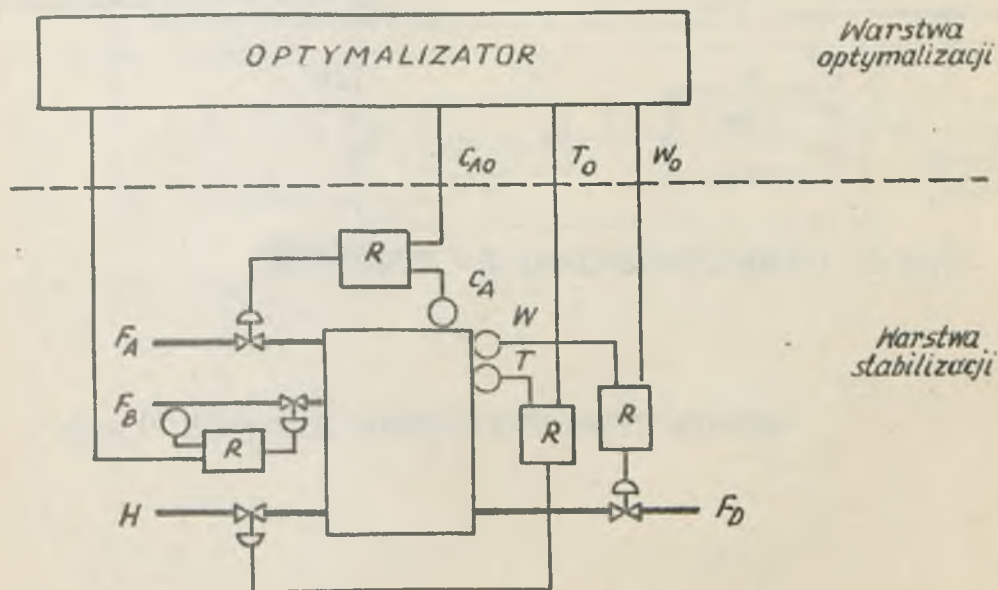
Rys. 4 Schemat procesu w surowni cukrowni



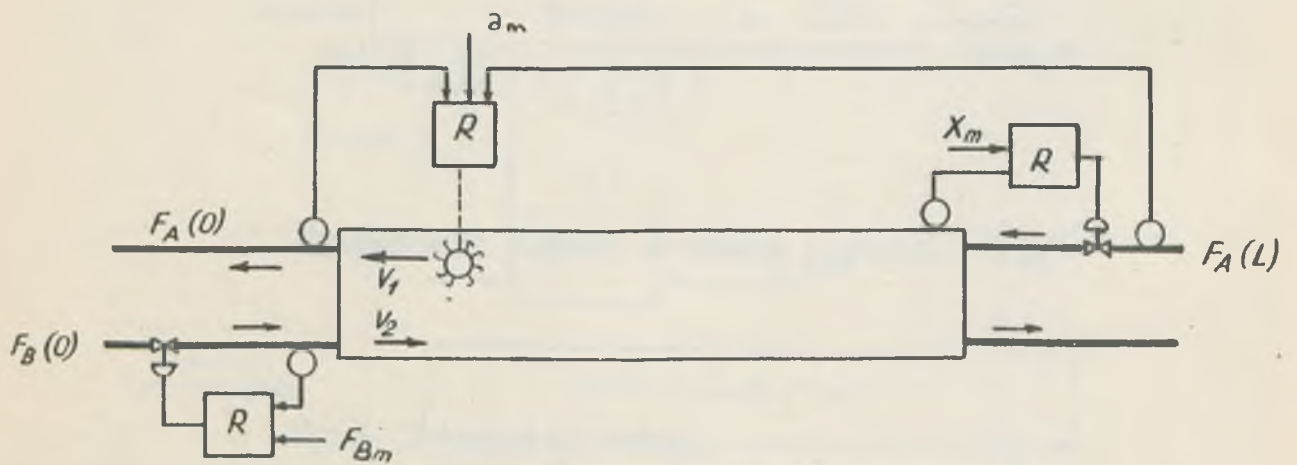
Rys. 5 Schematyczne przedstawienie modelu obiektu



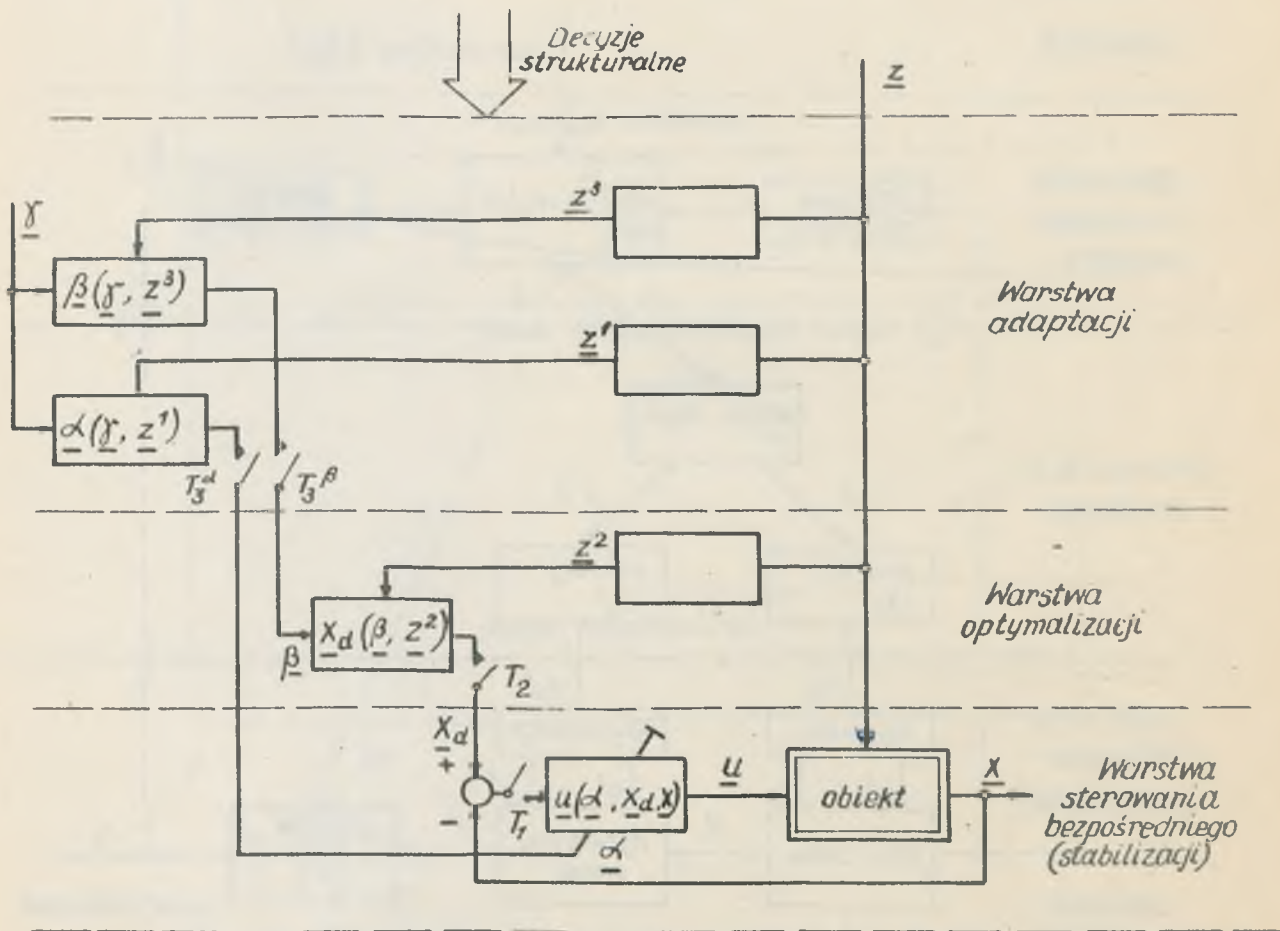
Rys. 6 Przykład obiektu o statych rozłożonych



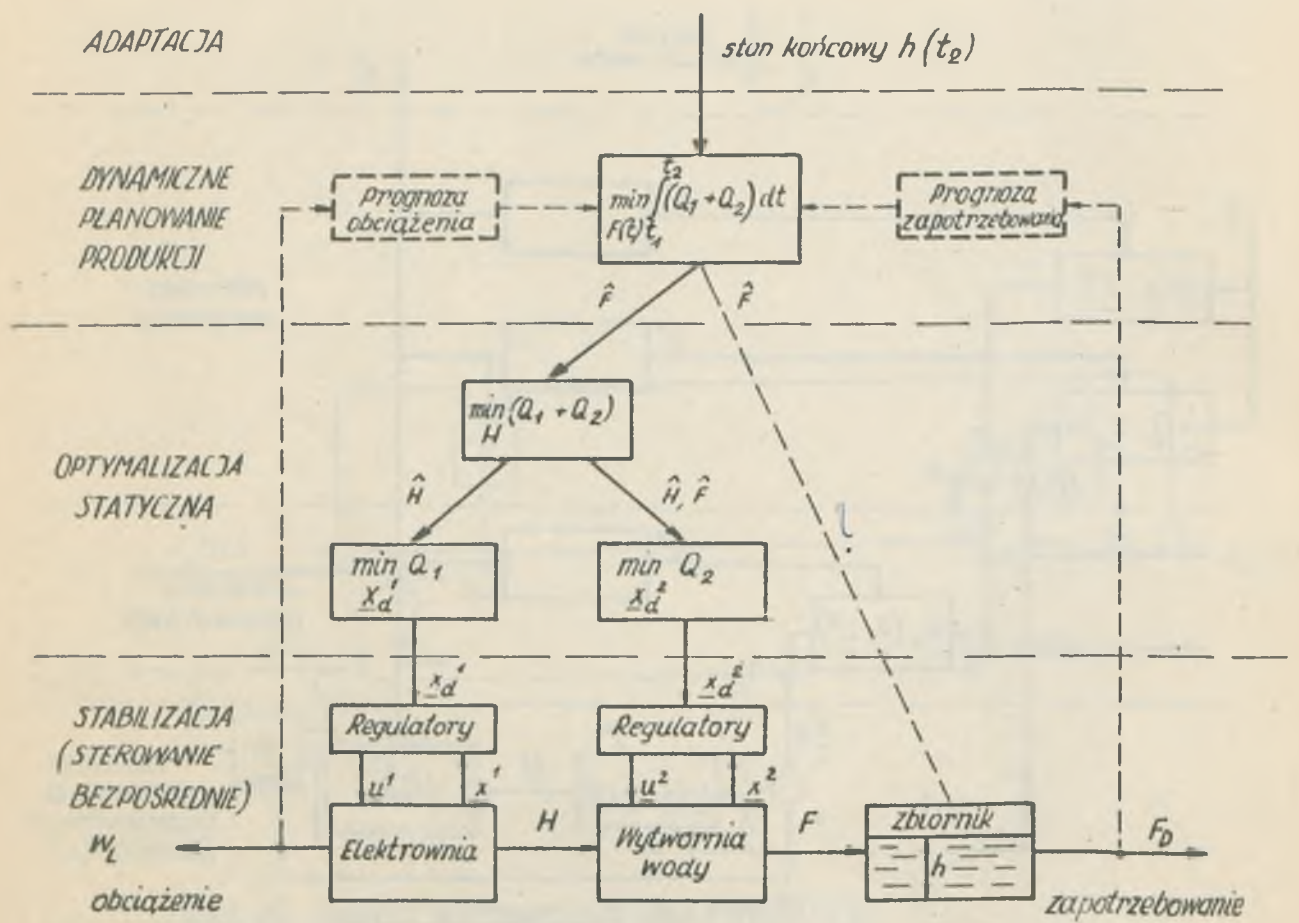
Rys. 7 Dwuwarstwowy układ sterowania reaktora R-regulatory



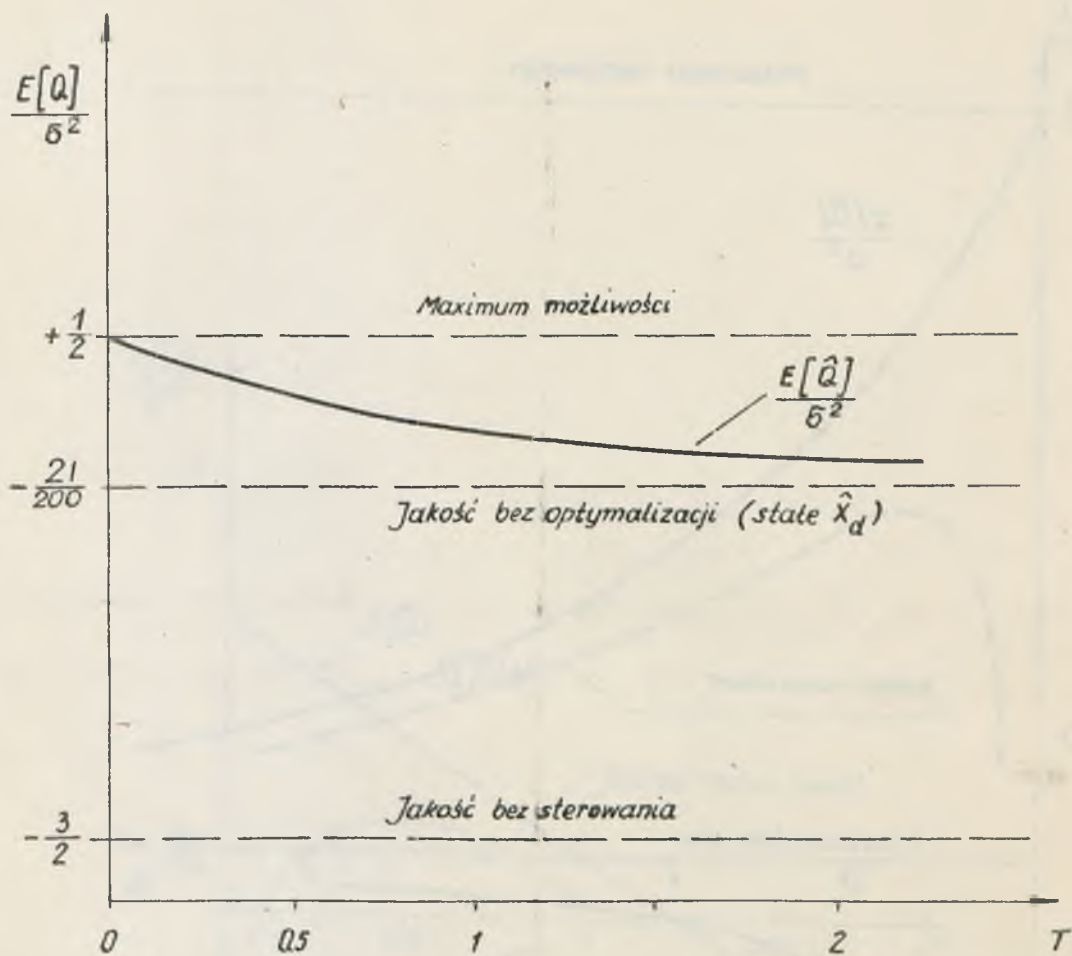
Rys. 8 Układ sterowania dla reaktora



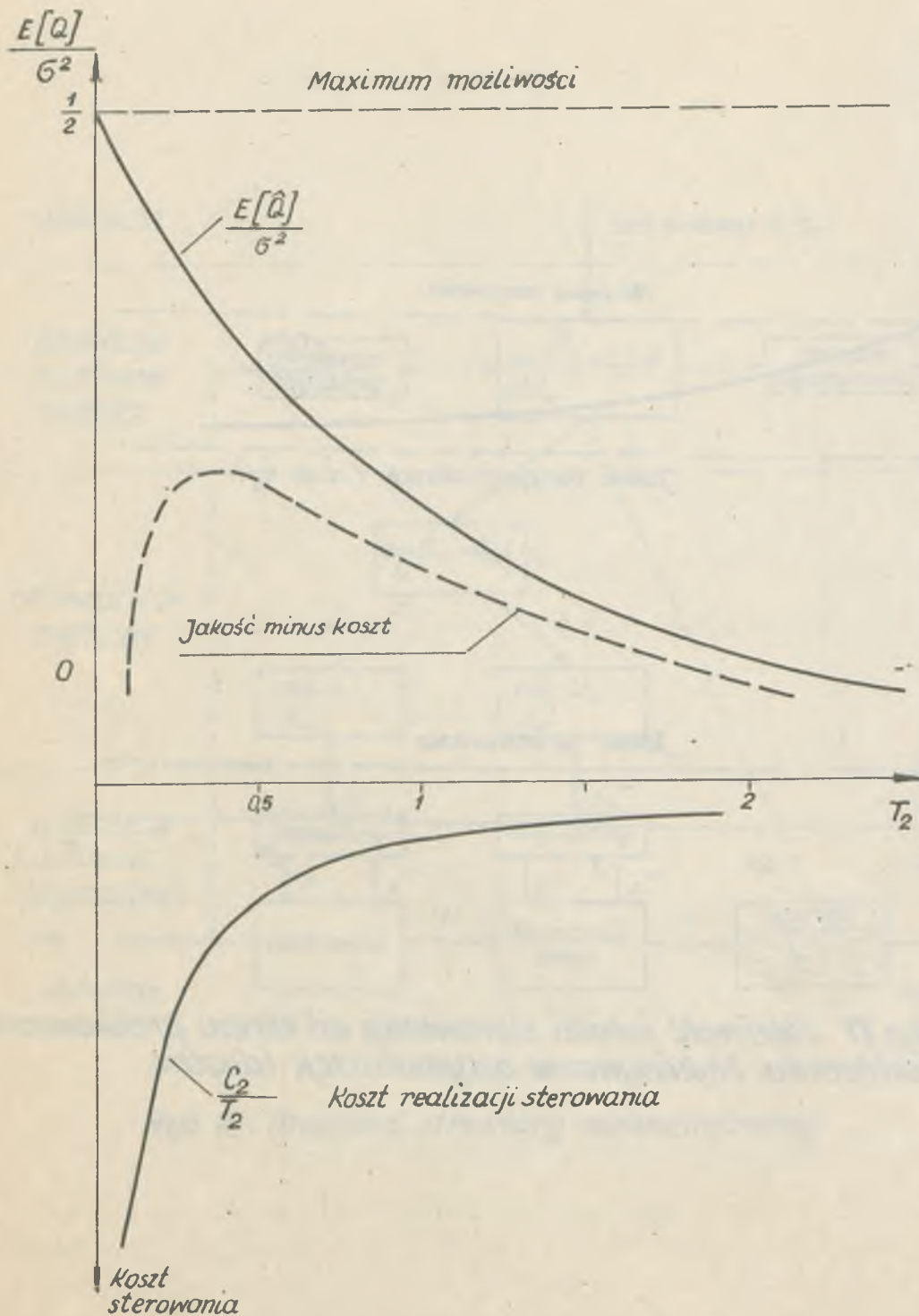
Rys.9 Schemat układu wielowarstwowego



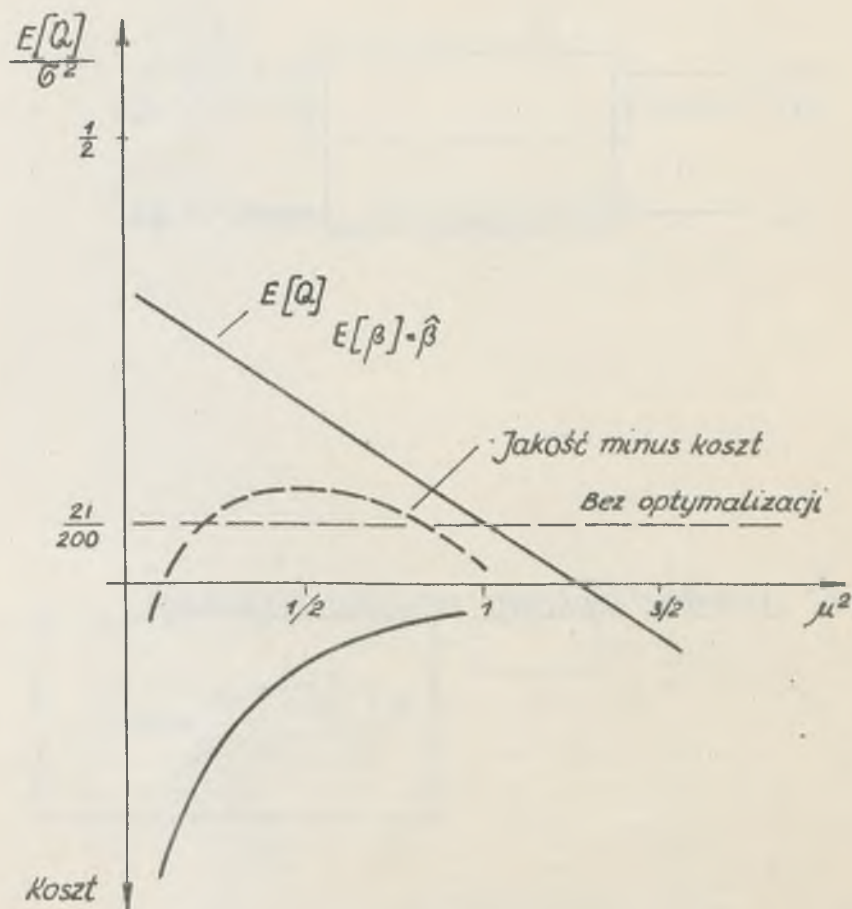
Rys. 10 Przykład struktury wielowarstwowej



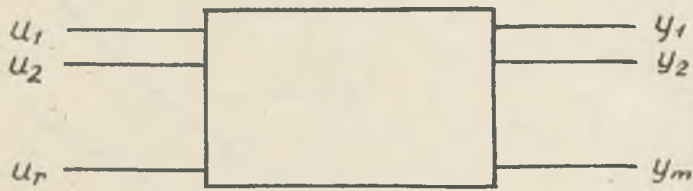
Rys.11 Zależność jakości sterowania od okresu próbkowania zakłócenia. Wykonywanie optymalizacji idealne.



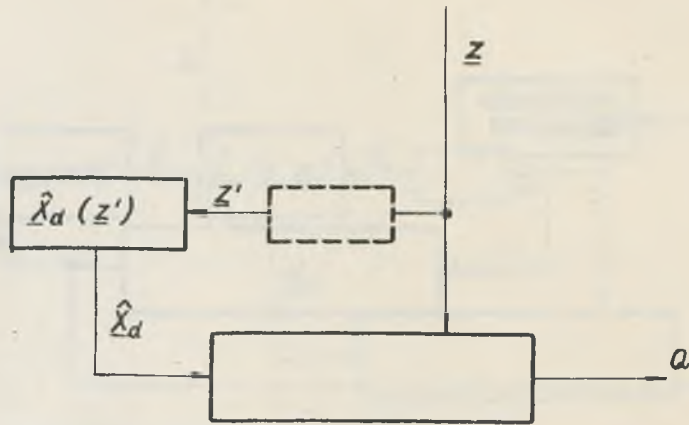
Rys. 12 Zależność jakości sterowania od okresu działania drugiej warstwy sterowania



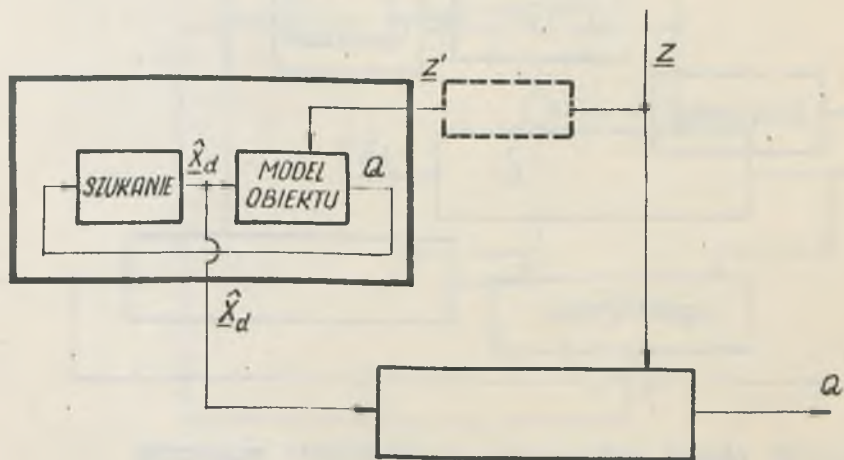
Rys. 13 Pogorszenie się jakości sterowania wywołane niepewnością określenia cech statystycznych zakłócenia. Krzywa kosztu podaje przybliżony koszt określenia tych cech z daną wariancją.



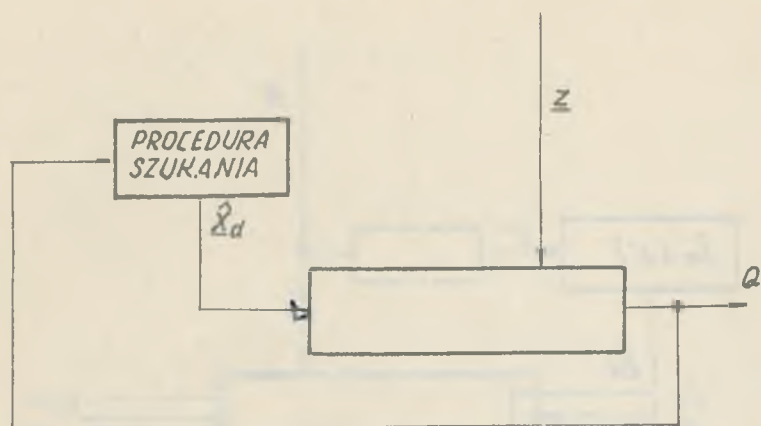
Rys.14 Schemat blokowy obiektu regulacji



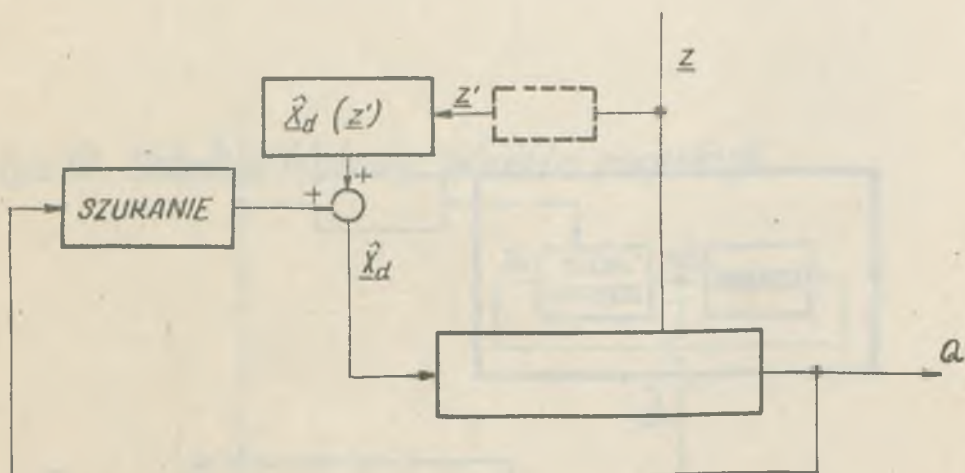
Rys.15 Otwarty układ optymalizacji



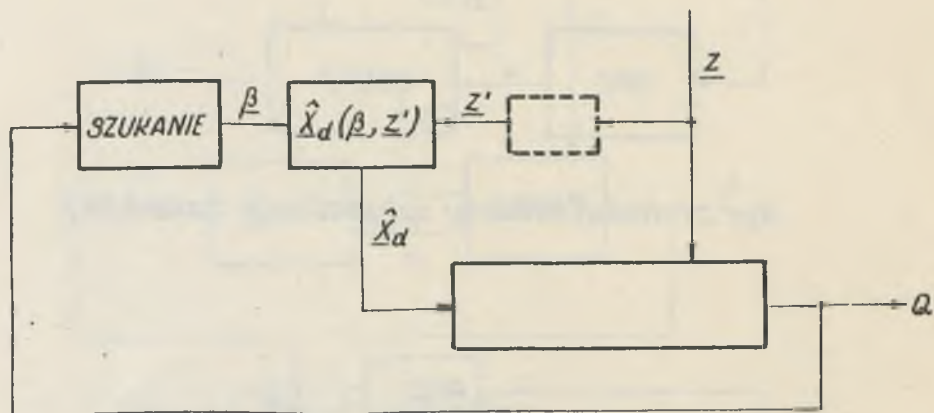
Rys.16 Otwarty układ optymalizacji, z zaznaczeniem szukania ekstremium na modelu obiektu



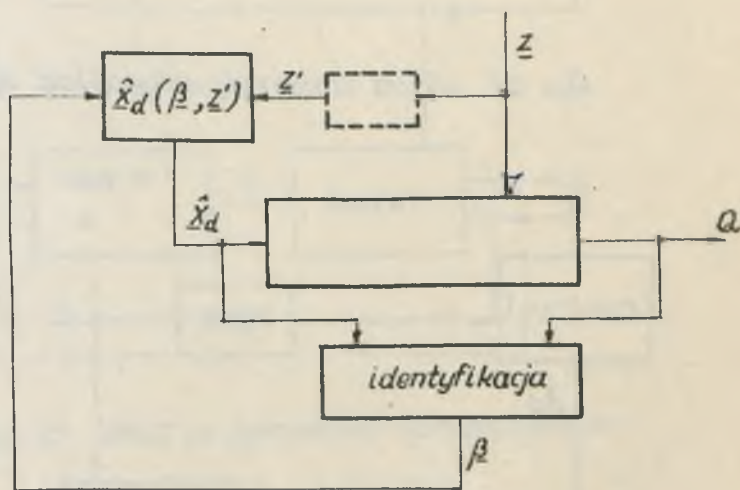
Rys. 17 Układ optymalizacji o strukturze zamkniętej



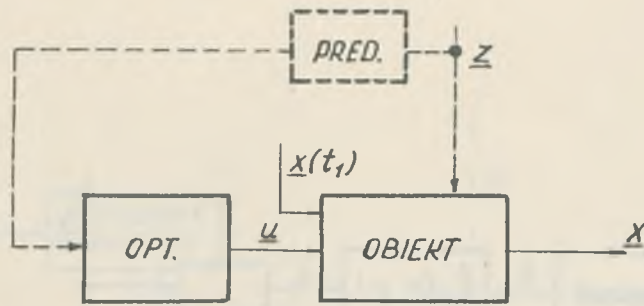
Rys 18 Układ optymalizacji o strukturze mieszanej



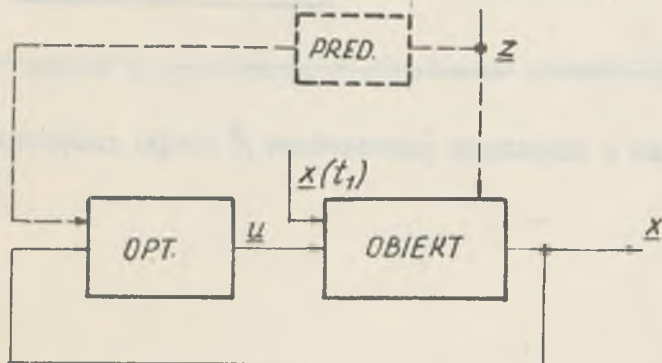
Rys. 19 Układ z adaptacją parametrów β drogą szukania



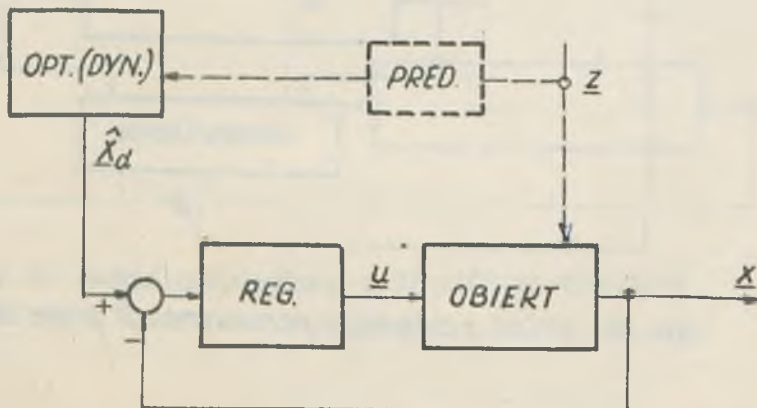
Rys. 20 Układ z adaptacją parametrów β przez identyfikację



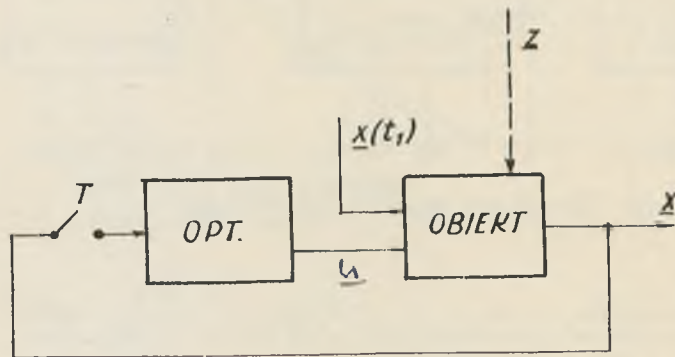
Rys. 21 Układ otwarty optymalizacji dynamicznej



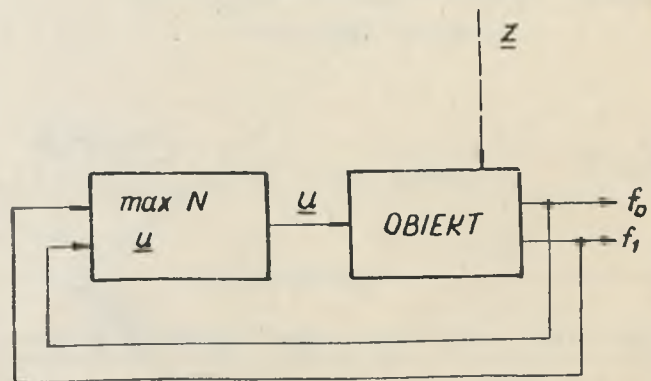
Rys. 22 Układ zamknięty optymalizacji dynamicznej



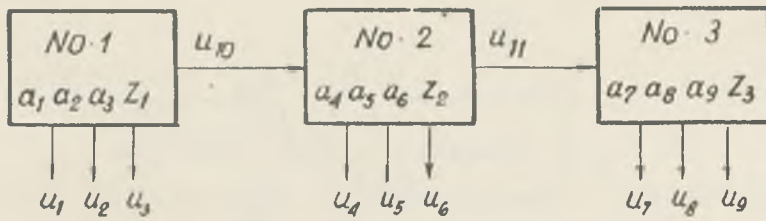
Rys. 23 Układ dwuwarstwowy optymalizacji dynamicznej



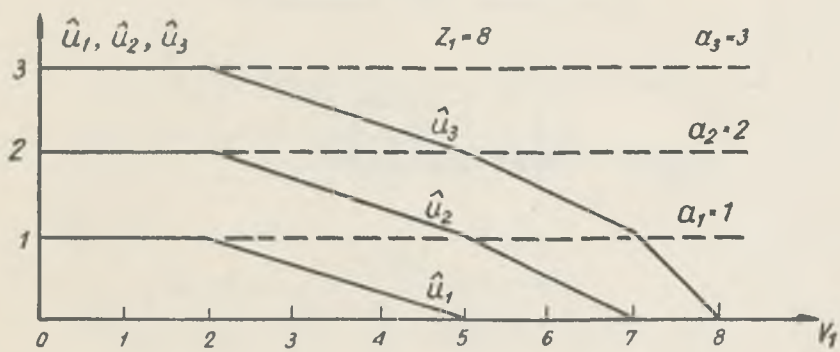
Rys. 24 Układ repetycyjny



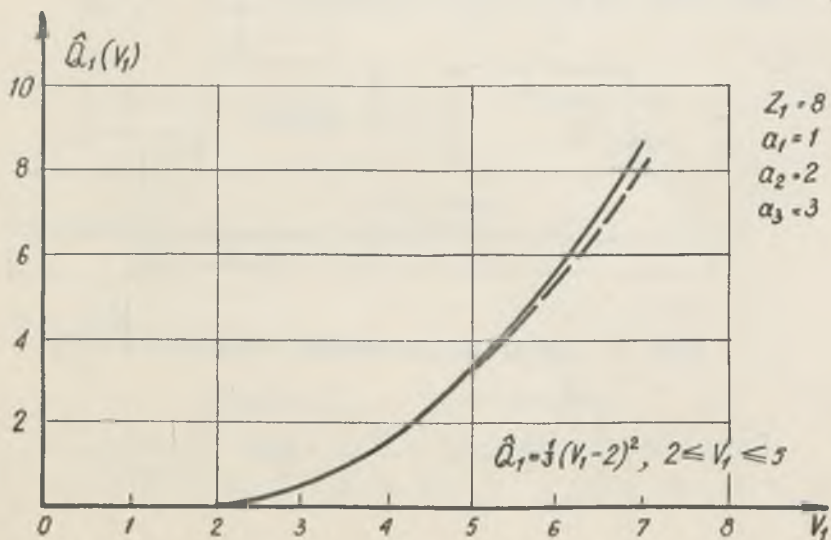
Rys. 25 Układ ze sprzężeniem optymalizującym



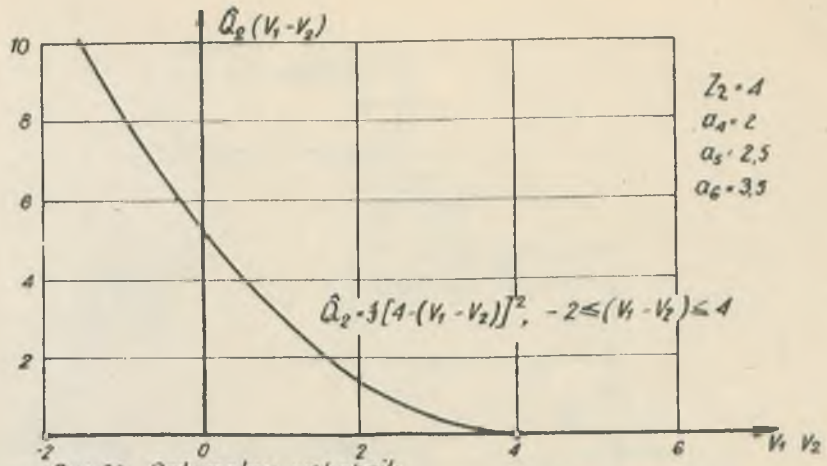
Rys. 26 Przykład systemu wodnego



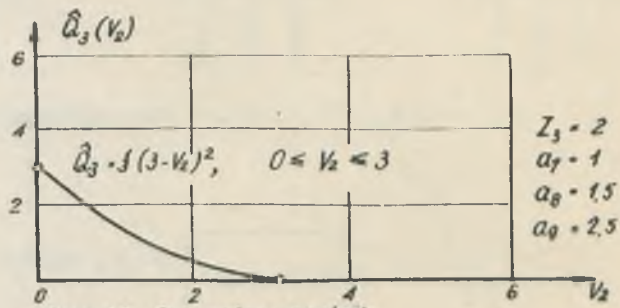
Rys. 27 Wykres rozwiązań optymalnych
zadanie częściowe 1



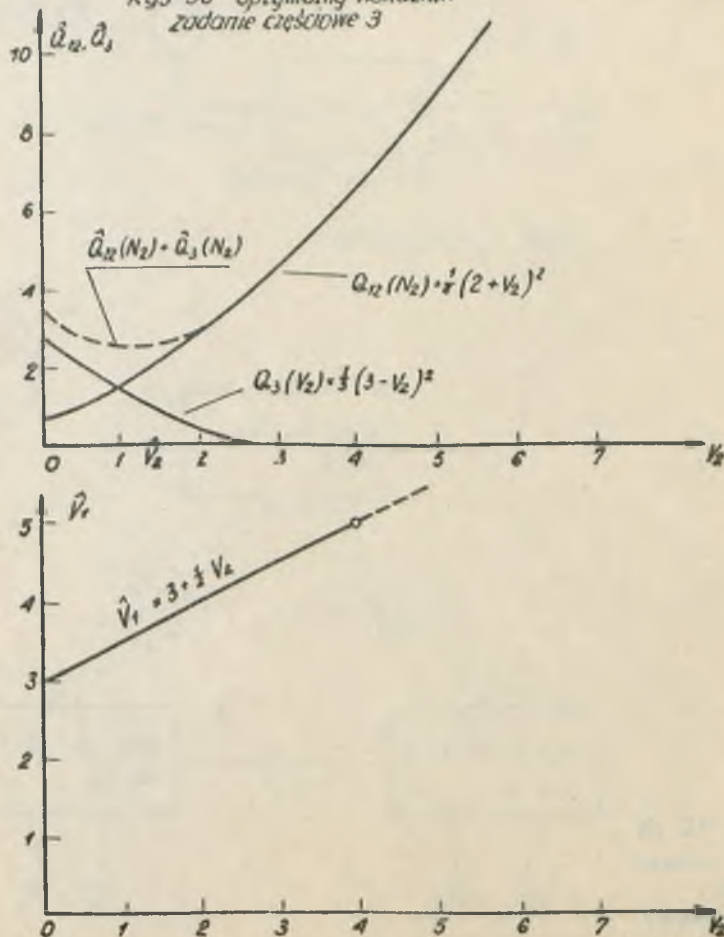
Rys. 28 Optymalna wartość wskaźnika jakości
zadanie częściowe 1



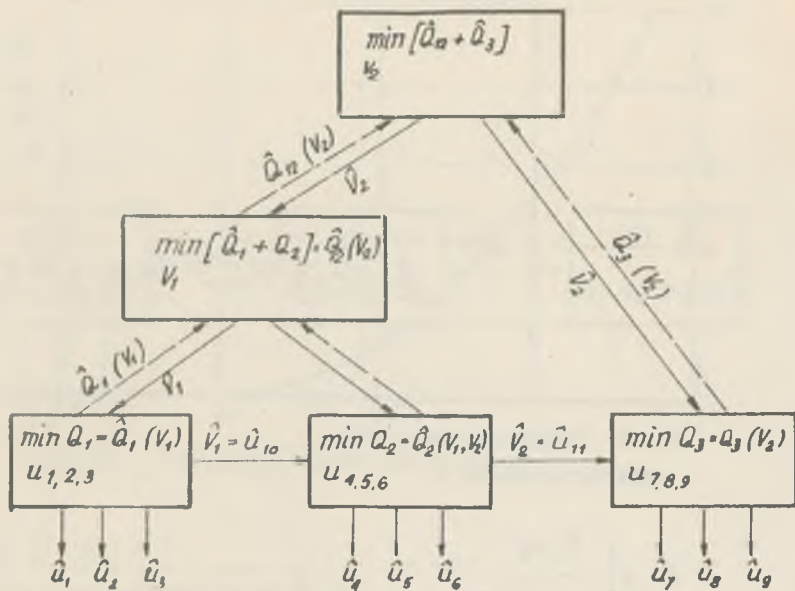
Rys 29 Optymalny wskaźnik
zadanie częściowe 2



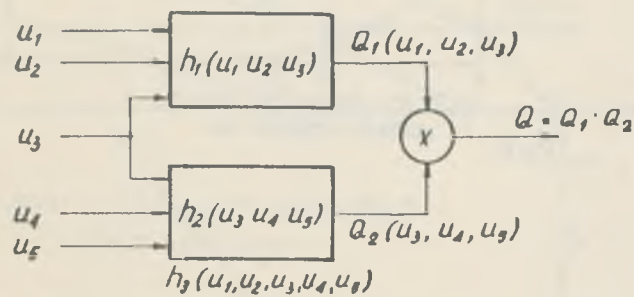
Rys 30 Optymalny wskaźnik
zadanie częściowe 3



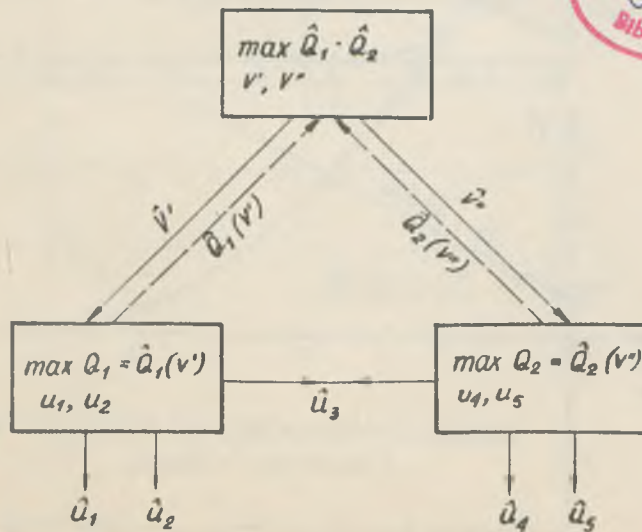
Rys 31 Suma wskaźników, jako podstawa decyzji trzeciego poziomu (u góry)
oraz reguła decyzyjna, dla poziomu drugiego



Rys. 32 Schemat struktury trzypoziomowej



Rys. 33 Przykład



Rys. 34. Struktura rozwiązania

BIBLIOTEKA GŁÓWNA
Politechniki Warszawskiej

C. 120254



400000000090017

**BG Magazyn
Podstawowy**