



**Krzysztof Malinowski
Piotr Tatjewski**

Podstawy Automatyki

Wersja 2, 2016



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓŁNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Projekt współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego

Przedmowa

Podręcznik niniejszy powstał w wyniku zgromadzenia wieloletnich doświadczeń autorów na Wydziale Elektroniki i Technik Informacyjnych Politechniki Warszawskiej, przy prowadzeniu przedmiotu „Podstawy Sterowania”, a następnie realizowanego do chwili obecnej przedmiotu „Podstawy Automatyki”. Prezentowany podręcznik jest podsumowaniem wielu poprzednich opracowań materiałów wykładowych, sukcesywnie doskonalonych i na bieżąco dostępnych dla studentów na wydziałowej stronie przedmiotu.

Pierwsza pełna wersja niniejszego podręcznika została opracowana w 2012 roku, w ramach Programu Rozwojowego Politechniki Warszawskiej – za wsparcie Programu autorzy pragną wyrazić gorące podziękowanie. Obecna wersja różni się od tej oryginalnej jedynie korektami i pewnymi niezbyt dużymi zmianami w rozdziale ósmym dotyczącym regulacji przemysłowej.

Autorzy chcieliby w tym miejscu podziękować studentom aktywnie uczestniczącym w prowadzonych zajęciach, za zgłaszaną uwagi do sposobu prezentacji, za wykrywane usterki edycyjne. Uwagi te w istotnym stopniu przyczyniały się do powstawania kolejnych, coraz lepszych, wersji niniejszego opracowania, a w efekcie prezentowanego podręcznika.

Krzysztof Malinowski

Piotr Tatjewski

Warszawa, 2016

Spis treści

1. Wstęp	7
1.1. Sterowanie, układy automatyki	7
1.2. Układ sterowania jako system informacyjny i informatyczny.....	10
1.3. Układy terowania logicznego i układy regulacji automatycznej	11
2. Sterowniki przemysłowe, sterownik PLC.....	15
2.1. Sterowniki przemysłowe	15
2.2. Programowalny sterownik logiczny (PLC).....	16
Literatura	18
3. Podstawowe rodzaje regulacji	19
3.1. Regulacja ciągła.....	22
3.2. Regulacja dwupołożeniowa.....	23
3.3. Regulacja trójpołożeniowa z korekcją	27
3.4. Regulacja ekstremalna.....	30
4. Modelowanie obiektów dynamicznych	33
4.1. Modele obiektów (procesów).....	33
4.2. Modele matematyczne; rodzaje i budowa.....	34
4.2.1. Modele teoretyczne (fenomenologiczne - fizyczne, fizyczno-chemiczne).....	35
4.2.2. Modele empiryczne (behawioralne)	42
4.2.3. Modele konceptualne (konceptualno-emeryczne).....	43
4.2.4. Identyfikacja modelu	44
4.3. Punkty równowagi, charakterystyki statyczne	46
4.4. Linearyzacja i skalowanie modelu	47
4.4.1. Linearyzacja lokalna.....	47
4.4.2. Linearyzacja poprzez (pomocnicze) sprzężenie zwrotne.	51
4.4.3. Skalowanie modelu.....	52
Zadania	52
5. Analiza liniowych modeli dynamicznych	56
5.1. Analiza w dziedzinie czasu	56
5.1.1. Rozwiązywanie równań modelu -przykład	56
5.1.2. Model liniowy w postaci ogólnej	60
5.1.3. Stabilność układu dynamicznego	62
5.2. Analiza w dziedzinie zmiennej zespolonej (metoda operatorowa).....	64
5.2.1. Transformata Laplace'a, transmitancja układu dynamicznego	64
5.2.2. Odpowiedzi skokowe i transmitancje podstawowych członów dynamicznych....	70
5.2.3. Przekształcanie schematów blokowych liniowych członów dynamicznych.....	75
5.2.4. Związki biegunów transmitancji z dynamiką układu, warunki stabilności.....	79
5.2.5. Algebraiczne kryterium stabilności Hurwitza	81
Zadania	83
6. Uchyby ustalone w układach regulacji	86
6.1. Obiekty i regulatory statyczne (bez całkowania)	87
6.1.1. Uchyb po skoku wartości zadanej, bez zakłócenia.....	87

6.1.2. Uchyb po skoku zakłócenia, przy zerowej wartości zadanej	88
6.2. Obiekty i regulatory z całkowaniem.....	89
6.2.1. Uchyby po zmianach wartości zadanej, bez zakłócenia	89
6.2.2. Uchyby po zmianach zakłócenia, przy zerowej wartości zadanej	92
6.3. Układy regulacji z dowolnym miejscem wpływu zakłócenia	95
Zadania	98
7. Analiza i korekcja układów regulacji w dziedzinie częstotliwości.....	100
7.1. Transmisja sygnału okresowego przez układ liniowy.....	100
7.2. Charakterystyki częstotliwościowe	102
7.3. Charakterystyki częstotliwościowe prostych członów dynamicznych	104
7.4. Kryterium Nyquist'a, zapasy stabilności.....	113
7.5. Projektowanie układów regulacji w dziedzinie częstotliwości	122
7.5.1. Zadanie projektowania podstawowego układu regulacji	122
7.5.2. Projektowanie serwomechanizmu.....	124
Zadania	129
8. Regulacja przemysłowa	132
8.1. Zadanie regulacji przemysłowej.....	132
8.2. Struktury i charakterystyki regulatora PID	132
8.2.1. Standardowe struktury i charakterystyki regulatora PID	132
8.2.2. Zmodyfikowane struktury regulatora PID	139
8.2.3. Ograniczenie całkowania	140
8.3. Modele obiektów dla doboru nastaw regulatorów PID.....	142
8.4. Dobór nastaw regulatorów PID	146
8.4.1. Kryteria doboru nastaw	146
8.4.2. Tabelaryczne reguły doboru nastaw.....	147
8.4.3. Dobór nastaw metodą SIMC	155
8.5. Regulacja kaskadowa, kompensacja zakłócenia mierzonego	159
8.6. Regulacja predykcyjna	161
Literatura	171
Zadania	171
9. Cyfrowa realizacja algorytmów regulacji	173
9.1. Transmitancja dyskretna.....	174
9.2. Metoda emulacji	175
9.2.1. Dyskretyzacja metodami bezpośredniej aproksymacji	176
9.2.2. Dyskretyzacja metodą Tustina	178
9.3. Dobór okresu próbkowania	179
Literatura	181
Zadania	182
Dodatek: Przekształcenie Laplace'a	183

1. Wstęp

Przedmiot Podstawy Automatyki stanowi wprowadzenie w zagadnienia związane z analizą i projektowaniem realizujących różnorodne zadania układów sterowania automatycznego, ze szczególnym uwzględnieniem podstawowych zadań regulacji. Nasze rozważania wypada poprzedzić sprecyzowaniem ważniejszych pojęć, przede wszystkim tytułowego terminu sterowanie.

1.1. Sterowanie, układy automatyki

Niech przedmiotem naszego zainteresowania będzie dany obiekt fizyczny i związany z tym obiektem pewien proces opisany poprzez odpowiednie wielkości. Prostym przykładem takiego obiektu może być poruszający się pojazd; przyjmijmy, że interesuje nas prędkość tego pojazdu i jej zmiany w pewnym przedziale czasu. Inny przykład, to obwód elektryczny, w którym interesują nas prądy płynące w poszczególnych gałęziach oraz napięcia na wybranych elementach.

Rozpatrywany, z punktu widzenia interesującego nas procesu, obiekt znajduje się pod wpływem różnych czynników. W szczególności wypada wyróżnić w tym miejscu dwa spośród tych czynników: oddziaływanie zewnętrzne stanowiące wielkości wejściowe sterowane, których wartości w czasie można kształtować i narzucać oraz oddziaływanie swobodne, których wartości są od nas niezależne. W tym kontekście możemy mówić o sterowaniu. W przypadku wspomnianego wyżej pojazdu wielkością sterowaną wpływającą na zmiany prędkości jest położenie pedału akceleratora, natomiast wielkości swobodne, niesterowane, to: aktualne nachylenie drogi, chropowatość nawierzchni oraz inne, w szczególności zależne od wiatru, opory ruchu.

W ogólnym przypadku wielkości sterowane zmieniane są przy pomocy odpowiednich urządzeń wykonawczych podczas procesu sterowania.

Sterowaniem nazywamy kształtowanie – zgodnie z określona procedurą i w czasie rzeczywistym - przeznaczonych do tego celu wielkości sterowanych wejściowych danego obiektu lub procesu w celu uzyskania pożądanego zachowania się tego obiektu (procesu), na ogół w obecności swobodnych (niesterowanych) oddziaływań na obiekt (proces) ze strony otoczenia.

Sterowanie jest przeważnie realizowane przy pomocy funkcjonalnie wyodrębnionego układu, nazywanego układem sterującym. Układ sterowania funkcjonujący bez czynnego udziału człowieka nazywamy układem sterowania automatycznego.

Układ sterowania automatycznego może być wbudowany w obiekt lub może być fizycznie wyodrębniony z tego obiektu i umieszczony poza nim. Warto w tym miejscu wymienić kilka przykładów często spotykanych układów sterowania; należą do nich:

- Sterowanie pralką automatyczną.
- Sterowanie prędkością obrotową talerza pamięci dyskowej.
- Stabilizacja położenia głowicy pamięci dyskowej.
- Utrzymywanie pożądanej temperatury w danym pomieszczeniu lub innym urządzeniu.
- Stabilizacja (nastawianie) przepływu zasilającego reaktor chemiczny.

Warto także podać przykłady zadań wymagających wprowadzenia złożonych układów sterowania (systemów sterowania); należą do nich:

- Sterowanie ruchem pociągów metra.
- Sterowanie pracą kolumny destylacyjnej.
- Stabilizacja położenia wieży i działa czołgu podczas jazdy w terenie.

- Sterowania zespołem kocioł-turbina w elektrowni.
- Sterowanie pracą zespołu instalacji – ciągu technologicznego; na przykład sterowanie pracą wydziału olejowego rafinerii.

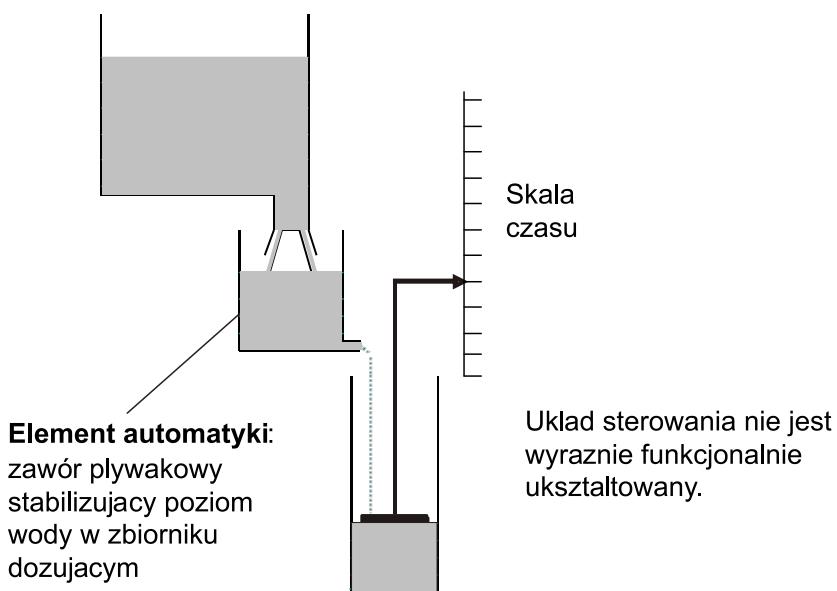
W tym miejscu warto poświęcić chwilę uwagi pojęciom automatyka i układ automatyki. Otóż, automatyka zajmuje się budową urządzeń oraz tworzeniem i badaniem układów (systemów) umożliwiających właściwe, możliwe w danych warunkach najlepsze, funkcjonowanie różnorodnych obiektów i procesów bez udziału człowieka lub przy minimalizacji tego udziału.

Automatyka obejmuje w szczególności:

- Budowę specjalizowanych komponentów obiektu realizujących funkcje sterowania i nadzoru nad działaniem tego obiektu,
- Projektowanie i budowę układów sterowania i nadzoru,
- Projektowanie i budowę systemów wspomagania decyzji do celów sterowania i zarządzania.

Jak widać, kompleks zagadnień związanych z pojęciem automatyka obejmuje, poza problematyką sterowania jako takiego, szereg innych kwestii, w tym dotyczących pomiarów różnych wielkości, nadzorowania biegu procesu, wprowadzania i stosowania blokad i zabezpieczeń. Układ automatyki zawiera elementy związane z wymienionymi funkcjami swojego działania; sterowanie jest tylko jedną z tych funkcji, aczkolwiek mającą zasadnicze znaczenie i stanowiącą, na ogólny, główny przedmiot zainteresowania projektanta.

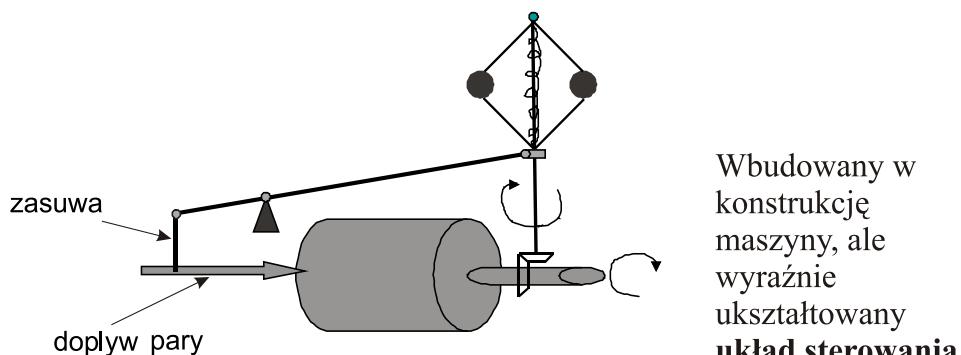
Historyczne początki związane z pojawieniem się prostych elementów lub układów automatyki sięgają głęboko w przeszłość. Przedstawiony na rysunku 1.1, w uproszczony oczywiście sposób, zegar wodny, którego konstrukcja przypisywana jest Ktesibiosowi, zawiera element automatyki - zawór pływkowy, zapewniający stabilizację poziomu wody w zbiorniku dozującym; celem jest oczywiście zapewnienie równomiernego wypływu wody z tego zbiornika - po to, by pływk połączony z wskaźnikiem czasu podniósł się ruchem jednostajnym.



Rys. 1.1. Konstrukcja zegara wodnego Ktesibiosa (ok. 300 p.n.e.).

Jako ciekawostkę warto zauważyć, że w starożytności, inaczej niż obecnie, dzielono każdy dzień od świtu do zachodu słońca na dwanaście godzin; podobnie na dwanaście godzin dzielono noc. Stąd długość godziny była różna w różnych porach roku. Konstrukcja zegara mechanicznego, która przy stałej, jak to ma miejsce obecnie, długości godziny leżałyby, w świetle obecnej wiedzy o ich osiągnięciach, jak najbardziej w zasięgu możliwości starożytnych Greków czy Rzymian, była w tej sytuacji bardzo trudna. Łatwo natomiast zauważać, że zegar wodny pozwalał stosunkowo łatwo rozwiązać problem zmiennej długości godziny; wystarczyło, na przykład, odpowiednio stopniowo unosić lub opuszczać zbiornik dozujący, zmieniając w ten sposób poziom cieczy a w konsekwencji intensywność wypływającego strumienia wody.

Wyraźnie wyodrębniony układ automatyki, w postaci regulatora prędkości obrotowej wału maszyny pojawił się w maszynie parowej Jamesa Watta (Rys 1.2). W istocie to właśnie wprowadzenie tego regulatora stanowiło główny wynalazek i osiągnięcie Watta; maszyna parowa, jako taka, była znana znacznie wcześniej, natomiast jej praktyczne zastosowanie napotykało na wielkie trudności dopóki nie udało się znaleźć sposobu na stabilizację prędkości obrotowej przy zmianach obciążenia na wale maszyny.



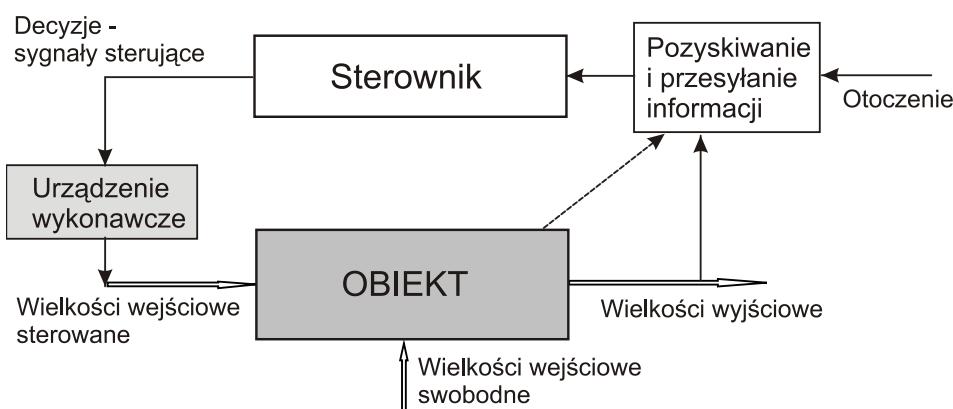
Rys. 1.2. Regulator odśrodkowy prędkości obrotowej maszyny parowej (J. Watt, 1788 r.).

W ogólnym przypadku obiekt wyposażony w układ sterowania automatycznego, przedstawiony na Rys. 1.3, zawiera następujące elementy:

- urządzenia (czujniki) pomiarowe pozwalające na pozyskanie informacji o zachowaniu się obiektu oraz urządzenia służące do przesłania tej informacji do sterownika (w naszym przykładzie jest to mechaniczny pomiar prędkości obrotowej wału maszyny parowej wyrażający się położeniem pierścienia unoszonego przez wirujący układ kul),
- sterownik, w którym następuje wyznaczenie wartości sygnału sterującego kształtującego wartość wielkości wejściowej na którą możemy i chcemy oddziaływać (w przypadku maszyny parowej rolę sterownika odgrywa bezpośrednio układ pomiaru prędkości obrotowej; wartość tego pomiaru w postaci położenia pierścienia stanowi wielkość sterującą określającą pożąданie położenie zasuwy na dopływie pary),
- urządzenie wykonawcze, które nastawia pożądaną, określzoną przez sterownik, wartość wielkości wejściowej (w maszynie parowej rolę tę pełni zasuwa na dopływie pary).

Wpływ otoczenia na zachowanie obiektu wyraża się poprzez różnorakie oddziaływanie swobodne, niesterowane; w przypadku maszyny parowej może to być zarówno zmienne obciążenie wału maszyny, do którego podłączane są różne urządzenia, jak i zmienne parametry pary. Wielkości swobodne często mają charakter zakłóceń; określamy je tym

mianem zwłaszcza wtedy, gdy mówimy o przypadkowych odchyłkach danego oddziaływania od jego określonej wartości, na przykład od odchyłkach obciążenia na wale od danej wartości nominalnej. Terminu „zakłócenie” będziemy stale używać omawiając zagadnienia związane z projektowaniem układów regulacji; tym niemniej należy pamiętać, że w wielu przypadkach istnienie wejść swobodnych wręcz warunkuje celowość zajmowania się danym obiektem i jego sterowaniem. Jako przykład, poza powyżej omawianą maszyną parową, która ma napędzać różnorakie urządzenia obciążające w zmienny sposób jej wał, warto rozważyć zbiornik retencyjny na rzece (Rys. 1.5); w tym przypadku wielkością swobodną wejściową jest dopływ wody do zbiornika – trudno jest traktować ten dopływ za zakłócenie w potocznym tego słowa znaczeniu.



Rys. 1.3. Obiekt i elementy układu sterowania.

1.2. Układ sterowania jako system informacyjny i informatyczny

Układ sterowania, niezależnie od występowania w nim różnego rodzaju elementów fizycznych i różnych nośników energii i sygnałów, można rozpatrywać jako przykład systemu informacyjnego. Stąd warto poświęcić chwilę uwagi temu, co rozumiemy poprzez to pojęcie.

Systemy informacyjne można ogólnie charakteryzować poprzez ich funkcje; na te funkcje składają się:

- Pozyskiwanie informacji (pomiary),
- Przesyłanie informacji (np. z wykorzystaniem sieci danych),
- Przechowywanie informacji (bazy danych),
- Przetwarzanie informacji (np. podejmowanie decyzji sterujących)
- oraz, często, także
- Prezentacja informacji (np. wizualizacja stanu obiektu).

Stąd możemy mówić o systemach informacyjnych jako systemach typu 4P lub 5P. Poszczególne, wymienione powyżej funkcjonalności, występują w poszczególnych rodzajach systemów informacyjnych w różnym nasileniu. W systemach pomiarowych dominuje oczywiście pozyskiwanie informacji, w systemach telekomunikacyjnych przesyłanie informacji, w systemach wspomaganego komputerem projektowania przetwarzanie i prezentacja graficzna danych (informacji). Warto podkreślić, że układy i systemy sterowania stanowią kategorię systemów informacyjnych, w której obecne są w znaczającej mierze wszystkie z przywołanych funkcjonalności.

Obecnie, w skutek powszechnego rozwoju techniki komputerowej, szczególną rolę wśród systemów informacyjnych, w tym systemów sterowania, odgrywają systemy, które możemy

określać mianem systemów informatycznych. Przez system informatyczny rozumiemy w tym miejscu system informacyjny zrealizowany w oparciu o urządzenia i środki informatyczne, tj.: komputery, sieci danych i odpowiednie oprogramowanie.

Współczesne układy (systemy) sterowania są w większości przypadków realizowane właśnie jako systemy informatyczne. Centralną część takiego układu sterowania stanowi sterownik – urządzenie cyfrowe i odpowiedni algorytm - wypracowujący w czasie rzeczywistym decyzje sterujące – sterowania - określające wartości wielkości wejściowych mogących podlegać odpowiedniemu kształtowaniu w celu uzyskania pożądanego zachowania się obiektu.

Cechy charakterystyczne (informatycznych) systemów sterowania automatycznego, wyróżniające je w ogromnej liczbie różnorodnych systemów informatycznych, są określane między innymi poprzez istotne wymagania, do których należą:

1. Bezpieczeństwo i niezawodność działania; wymagane jest bardzo staranne programowanie oraz rzetelne testowanie wszystkich urządzeń i oprogramowania przed uruchomieniem systemu sterowania, dla podwyższenia bezpieczeństwa i niezawodności stosowana jest redundancja sprzętowa i programowa - często wielokrotna; wykorzystywane są urządzenia o wysokiej wytrzymałości, odporności na działanie w wysokiej (lub niskiej) temperaturze, odporności na wstrząsy, zakłócenia elektromagnetyczne, i.t.p.
2. Przestrzeganie ograniczeń czasowych; poszczególne czynności – w trybie normalnym lub awaryjnym - muszą być wykonywane w określonym reżimie czasowym; wykorzystywane są w tym celu systemy operacyjne czasu rzeczywistego.
3. Złożone systemy sterowania pracują z reguły pod bieżącym nadzorem człowieka lub nawet zespołu operatorów; na przykład system sterowania zespołem kocioł-turbina nadzorowany jest przez co najmniej dwóch dyżurnych operatorów.

Systemy sterowania mają często, niemal zawsze w przypadku obiektów przemysłowych, charakter warstwowy; w poszczególnych warstwach realizowane są różne funkcje. Dla przykładu w systemie sterowania klimatyzacją w nowoczesnym budynku biurowym lub hotelowym możemy w warstwie niższej, zwanej warstwą sterowania bezpośredniego, wyodrębnić układy regulacji związane, w szczególności, z utrzymywaniem zadanych wartości temperatur i wilgotności w poszczególnych pomieszczeniach. Z kolei, warstwa wyższa takiego układu sterowania odpowiada za określenie tychże zadanych wartości; ta funkcja może być realizowana przez człowieka lub odpowiedni algorytm; celem działania algorytmu decyzyjnego może być minimalizacja kosztów energii przy jednoczesnym zapewnieniu odpowiednich warunków klimatycznych w określonych okresach czasu. W ramach tego wykładu interesować nas będą głównie proste układy sterowania, odpowiadające za realizację podstawowych funkcji sterowania bezpośredniego.

1.3. Układy sterowania logicznego i układy regulacji automatycznej

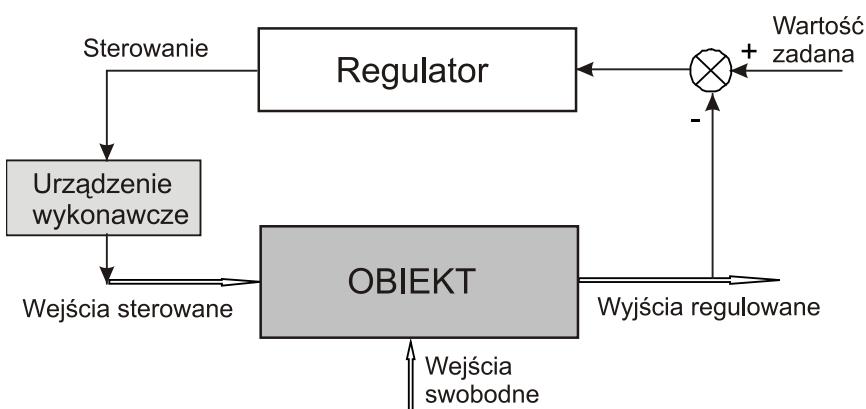
Podstawowe układy sterowania stanowiące tzw. warstwę sterowania bezpośredniego, będące głównym przedmiotem naszego zainteresowania w kolejnych rozdziałach, możemy podzielić na:

- **Układy sterowania logicznego (binarnego); podejmowane są w nich decyzje typu włącz-wyłącz,** odpowiednio uszeregowane i uwarunkowane upływem czasu lub informacją o określonych zdarzeniach. Jako przykłady służby mogą: sterowanie cyklem prania w pralce automatycznej, sterowanie linią napełniania butelek, sterowanie ruchem składów pociągów metra.

- **Układy regulacji; decyzje sterujące dotyczą wówczas wejść sterowanych w postaci wielkości ciągły**, takich jak położenia, prędkości, przepływy, napięcia, prądy itp.; cel działania stanowi nadążanie odpowiednich wyjść obiektu (wielkości wyjściowych procesu sterowanego) za zadanymi przebiegami tych wielkości, w szczególnym przypadku utrzymywanie poziomów danych wielkości na zadanych wartościach. Przykłady: utrzymywanie zadanej temperatury w pomieszczeniu, utrzymywanie (stabilizacja) prędkości obrotowej turbiny, naprowadzanie lufy działa na cel o zmieniającym się położeniu. Dwa pierwsze z wymienionych stanowią przykłady jednowymiarowej regulacji stałwartościowej, trzeci to przypadek znacznie bardziej zaawansowanej wielowymiarowej regulacji nadążnej.

Układy sterowania logicznego omawiane są w rozdziale drugim. Ich cechę charakterystyczną stanowi założenie o natychmiastowej reakcji obiektu na zmiany wartości sygnałów sterujących, dynamika tych zmian jest pomijana. Działanie obiektu kojarzone jest funkcjami realizowanymi w różnych stanach; liczba tych stanów może być znaczna ale jest skończona. Jako przykład może nam posłużyć w tym miejscu system sygnalizacji ruchem ulicznym, proces prania w pralce automatycznej, montaż na liniach produkcyjnych w fabrykach i wiele, wiele innych. Warto podkreślić, że systemy sterowania logicznego są obecnie szeroko rozpowszechnione i wykorzystywane w ogromnej liczbie zastosowań.

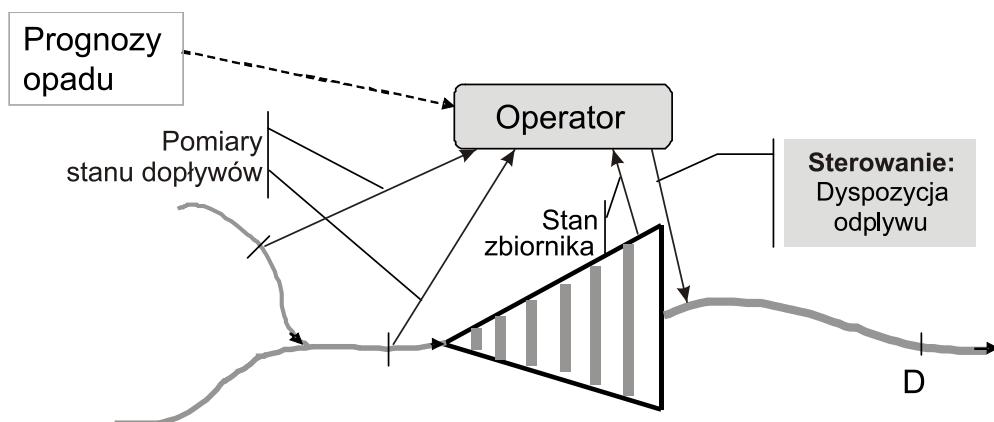
W przypadku układów regulacji interesuje nas, z kolei, właśnie dynamika zmian rozważanych wielkości, w szczególności wielkości regułowanej. Kształtowanie tej dynamiki, tj. uzyskanie odpowiednich przebiegów interesujących wielkości stanowi zadanie innej natury niż programowanie sterowników logicznych. Wymaga ono analizy różnych czynników i własności układów dynamicznych, w tym badania ich stabilności. Dlatego też badaniu i projektowaniu układów regulacji poświęcamy większość tego wykładu; traktuję o tym rozdział od trzeciego do dziewiątego.



Rys. 1.4. Układ regulacji automatycznej.

W podstawowym układzie regulacji pokazanym na Rys. 1.4 wyróżniamy szczególną postać sterownika, na który składa się węzeł porównawczy, zwany także węzłem sumującym, w którym następuje porównywanie wartości zadanej wielkości regułowanej z aktualnym pomiarem tej wielkości. Otrzymany sygnał uchybu, będący różnicą tych wartości, stanowi następnie podstawę do wyznaczenia przez regulator sygnału sterującego kształtującego wartość wielkości wejściowej – wejścia sterowanego. W istocie, w konstrukcji tego układu wyraża się podstawowa koncepcja sprzężenia zwrotnego, stanowiąca podstawę działania zdecydowanej większości układów sterowania automatycznego.

Należy podkreślić, że układy regulacji automatycznej stanowią nader rozpowszechnioną, podstawową, ale w istocie tylko jedną z możliwych form systemów sterowania procesów ciągłych. W celu pokazania przykładu układu sterowania niebędącego układem regulacji, a także niebędącego, w ścisłym tego słowa znaczeniu, układem sterowania automatycznego, rozważmy system sterowania, w którym ma miejsce wspomaganie decyzji związanych z zarządzaniem podczas powodzi usytuowanym na rzece zbiornikiem retencyjnym. Dyspozytor (operator) pełniący główną rolę w tym systemie może korzystać z różnych narzędzi umożliwiających ocenę wartości przyszłych dopływów do zbiornika (prognozy opadów, pomiary stanu wody) oraz ocenę konsekwencji możliwych działań – dyspozycji odpływu, natomiast decyzje podejmuje sam i na własną odpowiedzialność! Dopiero wówczas, gdy reguła podejmowania decyzji dotyczącej dyspozycji odpływu zostanie ujęta w postaci sprecyzowanego algorytmu i zastosowana bez udziału operatora – oraz bez odpowiedzialności z jego strony za skutki podejmowanych przez algorytm działań – system stanie się systemem sterowania automatycznego; nie będzie to jednak prosty układ regulacji - możemy go wówczas określić jako system sterowania optymalizującego.



Rys. 1.5. Przykład systemu wspomagania decyzji – sterowanie (zarządzanie) zbiornikiem retencyjnym podczas powodzi.

Na zakończenie wprowadzenia warto zauważyć, że rozwój środków i urządzeń automatyki można, jak to już zostało wyżej wspomniane, śledzić na przestrzeni wielu wieków, poczynając od czasów starożytnych; tym niemniej, szybki rozwój teorii, metod i urządzeń związanych z sterowaniem, w tym sterowaniem automatycznym, miał miejsce głównie w ostatnim stuleciu, w szczególności współczesne informatyczne, komputerowe układy sterowania rozwinięły się powszechnie po roku 1969, w którym wprowadzono do użytku mikroprocesory i inne układy scalone stanowiące obecnie podstawę konstrukcji sterowników.

Przytoczone poniżej daty określają poprzedzające kolejne etapy, prowadzące do tego współczesnego okresu:

- ok. 300 p.n.e. – stabilizacja poziomu wody, zegar wodny; Ktesibos,
- 1788 – regulator odśrodkowy prędkości obrotowej maszyny parowej; J. Watt,
- 1868 – artykuł „On governors” J. C. Maxwella, analiza dynamiki i stabilności regulatora odśrodkowego,
- 1877 – algebraiczne kryteria stabilności (E. J.Routh, A. Hurwitz)
- 1893 – A.M. Lyapunov: definicja stabilności, twierdzenia o stabilności

- 1900-1920 - żyroskop i autopilot dla statków (Sperry Gyroscope Company),
- 1927 – wzmacniacz elektroniczny ze sprzężeniem zwrotnym; H. S. Black (Bell Telephone Laboratoires),
- 1930 – teoria procesów stochastycznych (H. Wiener),
- 1936 – regulator PID (Callender et al.),
- 1932 - kryterium stabilności na podstawie charakterystyki częstotliwościowej (H. Nyquist)
- 1942 – reguły nastawiania regulatorów PID (Ziegler i Nichols),
- 1945 – kompletne reguły projektowania wzmacniaczy ze sprzężeniem zwrotnym metodą częstotliwościową (H. W. Bode, Bell Telephone Laboratories),
- 1948 – metoda linii pierwiastkowych (W. R. Evans),
- 1956 – zasada maksimum (L. S. Pontryagin); programowanie dynamiczne (R. Bellman) - poczatkują gwałtowny rozwój teorii sterowania optymalnego,
- 1960 – pierwszy kongres światowy IFAC (International Federation of Automatic Control)
- 1961, 1969 – loty Jurija Gagarina, pierwsze lądowanie na Księżycu – wielkie zastosowania teorii sterowania i sterowania optymalnego,
- **1969** – mikroprocesor; rozpoczyna się początkowo powolny, następnie coraz szybszy rozwój cyfrowych układów sterowania, obecnie informatycznych systemów sterowania.
- 1973,1976 – pierwsze zastosowania przemysłowe, pierwsze publikacje dotyczące algorytmów klasy MPC (Model-based Predictive Control)

2. Sterowniki przemysłowe, sterownik PLC

2.1. Sterowniki przemysłowe

Obiekty techniczne wymagające sterowania stanowią bardzo obszerny i zróżnicowany zbiór – od bardzo prostych urządzeń takich jak np. automat zamkający i otwierający bramę garażową, poprzez nieco bardziej złożone, jak np. automatyczna linia butelkowania, aż do bardzo skomplikowanych instalacji występujących w dużych zakładach przemysłowych, jak np. ciągi technologiczne związane z reaktorami w przemyśle chemicznym czy kolumnami destylacyjnymi w przemyśle petrochemicznym, linie produkcji i montażu w przemyśle maszynowym. Na całościowe sterowanie obiektem składa się zarówno **sterowanie logiczne** (zwane też binarnym) polegające na załączaniu czy wyłączaniu odpowiednich urządzeń w odpowiednich sytuacjach (stanach) wynikających z aktualnych pomiarów i zaprogramowanych następstw czasowych, jak i **automatyczna regulacja** (sterowanie w pętli sprzężenia zwrotnego) odpowiednich mierzonych wielkości wyjściowych obiektu wymagających stabilizacji na pewnych wartościach zadanych, ustalonych czy programowo zmieniających się w czasie.

W zależności od natury obiektu, rodzaju i stopnia złożoności zadania sterowania, może ono być realizowane przez:

- **Specjalizowane układy elektromechaniczne lub elektroniczne** (w tym mikroprocesorowe), najczęściej stosunkowo proste, wbudowane czy zintegrowane z obiektem sterowania (np. sterowanie bramą garażową, kuchenką mikrofalową, pralką automatyczną, wtryskiem paliwa do silnika spalinowego w samochodzie, itd.). Układy elektroniczne tego typu realizowane w technice mikroprocesorowej określa się jako **układy wbudowane** (*embedded systems*).
- **Programowalne sterowniki logiczne** (**PLC** – *Programmable Logic Controllers*), będące wyodrębnionymi, specjalizowanymi urządzeniami do realizacji zadań sterowania logicznego. Produkowane są jako rodziny urządzeń modułowych, od prostszych obsługujących jedynie niewiele sygnałów binarnych do bardzo złożonych. Mogą również zawierać moduły innego przeznaczenia niż tylko sterowanie logiczne i komunikacja, np. do obsługi sygnałów ciągłych dla potrzeb regulacji, ale nie jest to głównym ich zadaniem.
- **Programowalne sterowniki automatyki** (**PAC** – *Programmable Automation Controllers*, najmłodsza klasa urządzeń, termin wprowadzony w 2001). Produkowane jako urządzenia samodzielne (*standalone*) i rodziny urządzeń modułowych (*rack-mounted*). Mogą obsługiwać sygnały ciągłe jak i binarne, od PLC różnią się istotnie koncepcją oprogramowania systemowego i programowania, rolą w systemie sterowania. Przeznaczone do realizacji różnych zadań w otwartych, rozproszonych systemach automatyki, charakteryzują się wielofunkcyjnością, otwartymi standardami komunikacji, wielozadaniowością i modularną architekturą (wieloprocesorowość), otwartością środowiska programowania. Sterowniki PLC są obecnie wypierane przez sterowniki PAC, przede wszystkim w zastosowaniach szerszych niż tylko do sterowania liganego, jak zadania regulacji.
- **Regulatory przemysłowe** – samodzielne urządzenia sterujące konstruowane przede wszystkim dla realizacji pętli regulacji, z reguły umożliwiając również obsługę sygnałów binarnych. Konstruowane często jako urządzenia wielofunkcyjne (*multipurpose controllers*), umożliwiające równoczesną realizację kilku pętli regulacyjnych oraz zadań sterowania binarnego (mają wejścia/wyjścia zarówno ciągłe

jak i binarne). Regulatory przemysłowe są często specjalizowane do określonych klas zastosowań (np. ciepłownictwo – regulatory temperatury).

- **Sterowniki przemysłowe** – dysponujące dużymi mocami obliczeniowymi urządzenia sterujące będące elementami złożonych, kompleksowych systemów sterowania przemysłowego, umożliwiające efektywną realizację wielu złożonych zadań zarówno sterowania logicznego jak i regulacji oraz komunikacji, związanych z obsługą nawet tysięcy zmiennych binarnych i dziesiątek czy setek zmiennych ciągłych (powiązanych z pętlami regulacji).
- **Komputery PC** – zadania sterowania binarnego czy regulacji realizowane są również obecnie przez typowe komputery PC, najczęściej w tzw. wykonaniu przemysłowym, wyposażone w odpowiednie karty wejścia/wyjścia umożliwiające obsługę sygnałów pomiarowych (binarnych i ciągłych) i wprowadzanie sygnałów sterujących.

W epoce przed wprowadzeniem techniki mikroprocesorowej regulatory przemysłowe były jedynymi urządzeniami realizującymi zadania regulacji automatycznej. Produkowane były dawniej w technologii mechanicznej, a potem w analogowej technice elektronicznej. Obsługiwały jedynie jedną pętlę regulacji, nieco później, w miarze rozwoju elektroniki, kilka pętli. Oczywiście, wszystkie wymienione wyżej klasy urządzeń sterujących występujące obecnie, prostsze czy bardziej złożone, wielofunkcyjne, są urządzeniami konstruowanymi w technice komputerowej.

Programowalne sterowniki logiczne (PLC) są aktualnie grupą urządzeń sterujących powszechnie występujących, stąd poświęcimy im nieco uwagi w dalszej części rozdziału.

2.2. Programowalny sterownik logiczny (PLC)

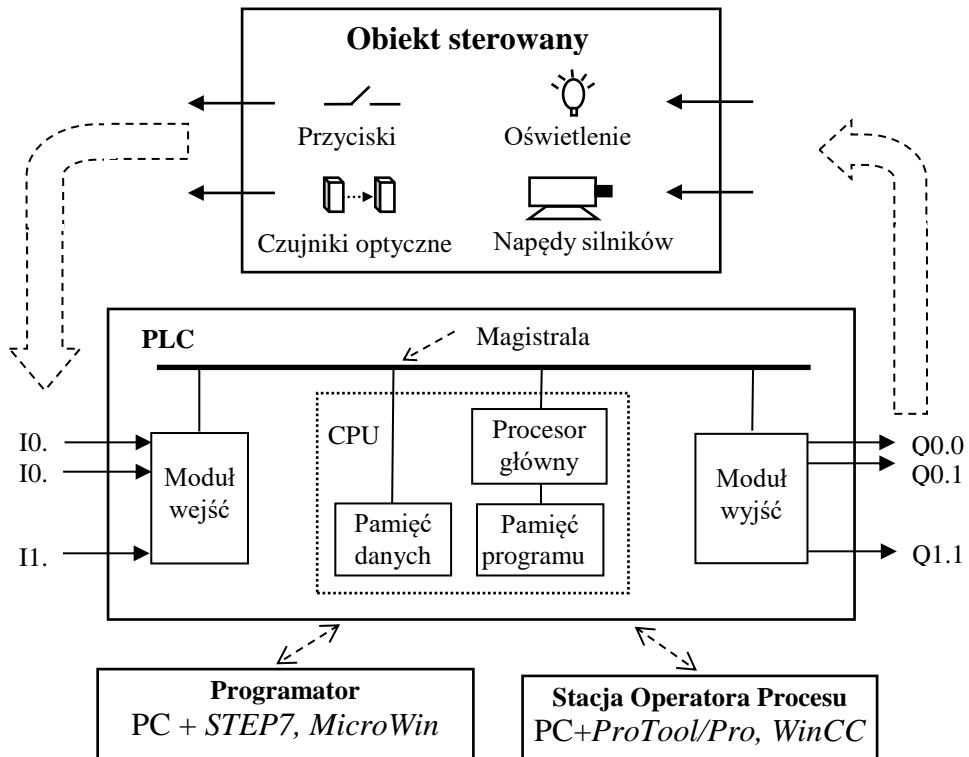
Sterownik PLC jest specjalizowanym komputerem. Podstawowe zespoły sterownika: procesor i pamięć wewnętrzna, nie różnią się istotnie od analogicznych zespołów komputera osobistego. Duże różnice występują natomiast w konstrukcji układów wejścia-wyjścia i w oprogramowaniu systemowym. Ponieważ w większości zastosowań sterownik pracuje autonomicznie, bez interwencji operatorów, nie posiada więc na ogół ani konsoli operatorskiej (ekranu i klawiatury), ani zawodnej pamięci masowej. Oprogramowanie sterownika jest w całości umieszczone w pamięci operacyjnej, zbudowanej z układów pamięci stałej i pamięci zapisywanej z podtrzymaniem baterijnym. Z drugiej strony, sterownik jest wyposażony w rozbudowane układy wejść i wyjść dwustanowych (i wejście-wyjście analogowych, jeśli ma być wykorzystywany również w układach sterowania ciągłego) umożliwiające połączenie go z instalacją sterowaną. Mechaniczna i elektryczna konstrukcja sterownika zapewnia możliwość pracy w obecności zakłóceń i zanieczyszczeń.

Oprogramowanie systemowe sterownika wykonuje **cyklicznie** następujące działania: odczytanie sygnałów wejściowych, wykonanie programu użytkowego obliczającego wartości sygnałów sterujących i wysłanie do obiektu obliczonych sygnałów. Operacje wejścia i wyjścia są wykonywane automatycznie. Zadaniem programisty jest opracowanie programu użytkowego, napisanego w **języku programowania** sterownika. Program użytkowy musi oczywiście realizować wybrany algorytm sterowania instalacją.

Architektura sterownika

Przykładowa architektura małego sterownika PLC wraz z urządzeniami towarzyszącymi jest przedstawiona na Rys. 2.1. Sterownik zawiera moduł wejść, jednostkę centralną złożoną z procesora, pamięci programu i pamięci danych oraz moduł wyjść. Sygnały wyjściowe obiektu opisujące stan sterowanego procesu są doprowadzone do **wejść sterownika**, a obliczone sygnały sterujące są przekazywane do obiektu poprzez **wyjścia sterownika**.

Ze względu na brak typowych urządzeń operatorskich, program użytkowy sterownika musi być utworzony w odrębnym systemie pełniącym rolę **programatora**, a następnie przesłany do sterownika PLC. Rolę programatora często pełni komputer IBM PC z oprogramowaniem narzędziowym, dostarczonym przez producenta sterownika. Na przykład, dla używanych podczas ćwiczenia sterowników Simatic S7-214 firmy Siemens, oprogramowaniem narzędziowym jest pakiet *STEP7 MicroWin*, działający pod kontrolą systemu Windows.



Rys. 2.1. Architektura sterownika PLC i urządzenia współpracujące.

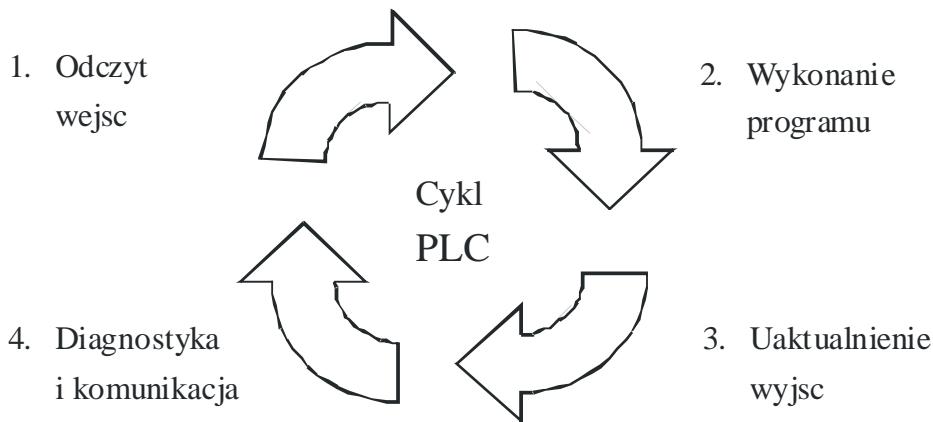
Język *STEP 7* umożliwia przygotowanie programu użytkowego, komplikację oraz załadowanie skompilowanego programu do pamięci operacyjnej sterownika. Komunikacja z komputerem PC odbywa się poprzez standardowy interfejs szeregowy RS232 lub RS485.

Ostatnim elementem systemu pokazanym na Rys. 2.1 jest **stacja operatora procesu** umożliwiająca komunikację operatora z wykonywanym programem, tzn. wyświetlanie informacji procesowych i wprowadzanie nowych wartości. Rolę stacji operatora może pełnić specjalny pulpit dołączany do sterownika lub komputer PC z odpowiednim oprogramowaniem, np. pakiet *WinCC* firmy Siemens działający pod kontrolą systemu Windows.

Wykonanie programu

Oprogramowanie sterownika binarnego tworzy jedną, cyklicznie obiegającą pętlę, przedstawioną symbolicznie na Rys. 2.2. Na każdy obieg pętli składa się operacja odczytania stanu wszystkich wejść sterownika, wykonanie programu użytkowego oraz operacja wysłania obliczonych wartości na wyjścia sterownika. Ostatnią część pętli zajmuje wykonanie programów diagnostycznych, sprawdzających poprawność pracy sprzętu, i funkcji

komunikacyjnych. Odczytanie stanu wejść, ustawienie wyjść oraz komunikacja (np. ze stacją operatora) są realizowane **poza programem użytkownika** przez oprogramowanie systemowe sterownika.



Rys. 2.2. Pętla programowa PLC.

W fazie wykonania programu użytkowego, program jest wykonywany sekwencyjnie od pierwszej do ostatniej instrukcji. W treści programu nie występują żadne wewnętrzne pętle ani rozgałęzienia. Ponieważ wartości sygnałów wejściowych są odczytywane i zapisywane w pamięci zawsze **przed** rozpoczęciem wykonania programu, więc zmiany wartości wejść, które nastąpią w obiekcie po rozpoczęciu wykonywania programu, zostaną przez program „zauważone” dopiero w trakcie **następnego** cyklu wykonywania. Dzięki temu wartości wejść są w całym przebiegu programu takie same.

Analogicznie, obliczone przez program wartości wyjściowe są przekazywane na wyjścia sterownika dopiero **po** zakończeniu wykonania całego programu — nawet jeśli zostaną obliczone już w pierwszych instrukcjach programu. Dzięki temu nie jest możliwa kilkukrotna zmiana wartości jakiegoś wyjścia w tym samym cyklu wykonania programu.

Czas trwania jednego obiegu pętli zależy od rozmiaru programu, liczby wejść i wyjść oraz od intensywności komunikacji sterownika ze światem zewnętrznym. W praktyce wynosi on zwykle od kilku do kilkuset milisekund.

Połączony z programatorem sterownik może pracować w dwóch trybach, oznaczanych symbolicznie: **STOP** i **RUN**. W trybie **STOP** sterownik komunikuje się z programatorem, za pomocą którego można ładować program do pamięci sterownika oraz przeglądać i zmieniać wartości jego rejestrów wewnętrznych. W trybie **RUN** sterownik obiega opisaną w tym rozdziale pętlę wykonania programu.

W normie IEC-61131-3 zdefiniowano 5 języków programowania sterowników PLC, pokrywających potencjalne zastosowania do procesów dyskretnych jak i ciągłych. Najbardziej rozpowszechnionym jest język schematów drabinkowych LD (*Ladder Diagram*), nawiązujący bezpośrednio do klasycznych stykowych obwodów przekaźnikowych. Po bliższe informacje dotyczące języka LD oraz przykłady programowania zadań sterowania logicznego dla prostych obiektów odsyłamy czytelnika do skryptu [1], skąd też zaczerpnięto zamieszczony wyżej opis sterownika PLC.

Literatura

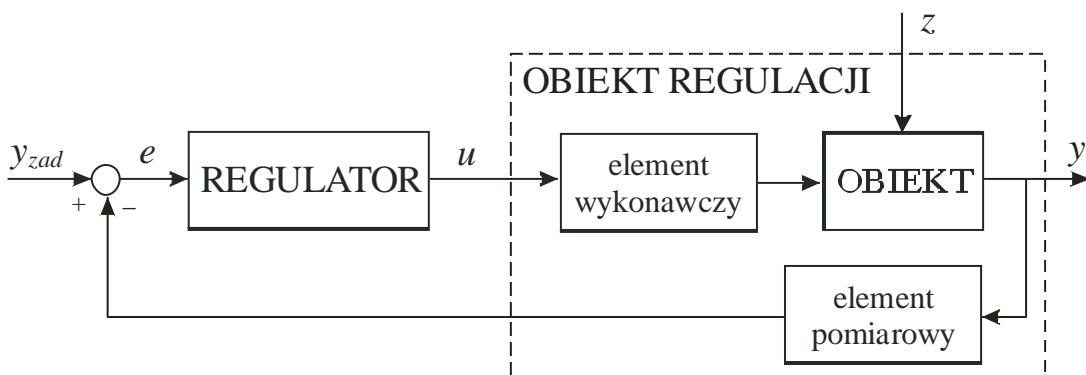
- U. Kręglewska i in.: Podstawy sterowania – ćwiczenia laboratoryjne. Skrypt OW PW, 2003.

3. Podstawowe rodzaje regulacji

Regulacja to sterowanie, realizowane w pętli sprzężenia zwrotnego, mające na celu utrzymanie wielkości wyjściowej regułowanej y obiektu na pożądanej wartości zadanej y_{zad} (stałej lub zmiennej w czasie).

W rozdziale niniejszym i następnych będziemy się zajmować regulacją obiektów ciągłych, tzn. obiektów, których wielkości wyjściowe regułowane mogą przyjmować dowolne wartości z pewnych fizycznie sensownych zakresów. Są to takie wielkości jak np. temperatura, ciśnienie, stężenie, położenie, prędkość, itp. Ogólna struktura układu regulacji przedstawiona jest na Rys. 3.1, gdzie

- y oznacza zmienną regułowaną (regułowane wyjście obiektu),
- u to sygnał sterujący z regulatora,
- $e = y_{zad} - y$ to uchyb regulacji,
- z oznacza zakłócenia (zmienne wejściowe niesterowane) oddziałujące na obiekt.



Rys. 3.1. Struktura układu regulacji z wyróżnieniem podstawowych elementów składających się na obiekt regulacji.

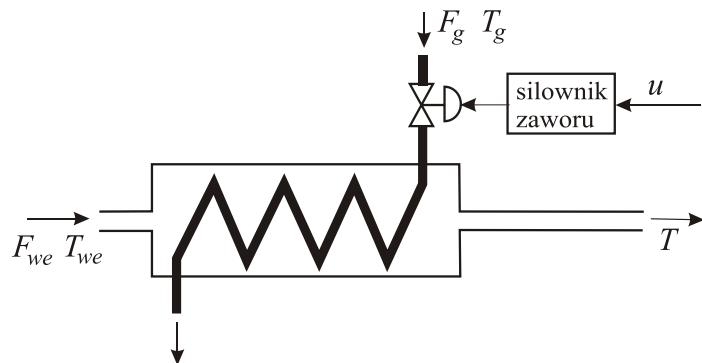
Na Rys. 3.1, obok regulatora i obiektu właściwego, wyróżniono element wykonawczy oraz element pomiarowy (czujnik) wielkości regułowanej – zawsze występujące w pętli regulacji elementy techniczne – pominięto natomiast przetworniki pomiarowe, występujące w praktyce jako osobne elementy w bardzo wielu istniejących instalacjach technicznych. We współczesnych realizacjach bowiem sygnały sterujący z regulatora i pomiarowy z czujnika przesyłane są w sposób cyfrowy, a przetworzenie sygnału sterowania na sygnał przestawiający element wykonawczy czy sygnału fizycznego z czujnika na sygnał przesyłany do regulatora są realizowane z reguły bezpośrednio w elemencie wykonawczym czy elemencie pomiarowym, wyposażonych w odpowiednie układy cyfrowe czy mikroprocesorowe. Z punktu widzenia regulatora obiektem regulacji jest wszystko, co znajduje się między generowanym przez regulator sygnałem sterującym u , a odbieranym przez regulator sygnałem pomiarowym wielkości regułowanej y , a więc kaskada: element wykonawczy (z odpowiednim przetwornikiem) – obiekt właściwy – czujnik pomiarowy (z odpowiednim przetwornikiem). W szczególności, element wykonawczy może w istotny sposób modyfikować własności dynamiczne obiektu regulacji, wprowadzając ograniczenia amplitudy oraz szybkości zmiany sygnału sterującego, związane z fizycznymi możliwościami realizacji oddziaływania na obiekt (ograniczone zakresy położień zaworu i szybkości jego

przestawiania, ograniczona prędkość silnika wykonawczego, itp.). Z drugiej strony, czas pomiaru wielkości wyjściowej może być istotny, nie do pominięcia (np. analiza składu produktu), wprowadzając opóźnienia sygnału pomiarowego dodatkowo modyfikujące dynamikę pętli regulacji.

Regulacja obiektów ciągłych może być realizowana jako:

- Regulacja ciągła – zmienna wyjściowa regulatora jest ciągła, tzn. może przyjmować dowolną wartość z odpowiedniego zakresu,
- Regulacja dwupołożeniowa – zmienna wyjściowa regulatora może przyjmować jedną z dwóch możliwych wartości, lub trójpołożeniowa -- zmienna wyjściowa regulatora może przyjmować jedną z trzech możliwych wartości,
- Regulacja krokowa – tzn. trójpołożeniowa z korekcją, jest to sposób (przybliżony) realizacji regulacji ciągłej z elementem wykonawczym trójpołożeniowym (tani element wykonawczy).

Przedstawimy wymienione rodzaje regulacji na przykładzie regulacji temperatury wody w przepływowym, przeciwprądowym wymienniku ciepła, przedstawionym schematycznie na Rys. 3.2. Do wymiennika dopływa woda chłodna o temperaturze T_{we} i natężeniu przepływu F_{we} , woda ta ulega podgrzaniu w wymienniku opuszczając go z temperaturą T . Medium grzewcze (np. woda gorąca) przepływa przez wymiennik wężowniczą, ma ono temperaturę wejściową T_g i natężenie przepływu F_g . Wielkością bezpośrednio sterującą natężeniem przepływu F_g jest położenie zaworu, przestawianego za pomocą pewnego silownika (silnika), którym może być np. silnik elektryczny. Sygnałem sterującym element wykonawczy obiektu (zawór z silownikiem) jest sygnał sterujący silownikiem, oznaczony przez u . Stąd, natężenie przepływu wody grzewczej zależy od u , $F_g = F_g(u)$. Celem regulacji będzie zapewnienie, aby temperatura T wody opuszczającej wymiennik miała pożądaną wartość T_{zad} .



Rys. 3.2. Schemat wymiennika ciepła.

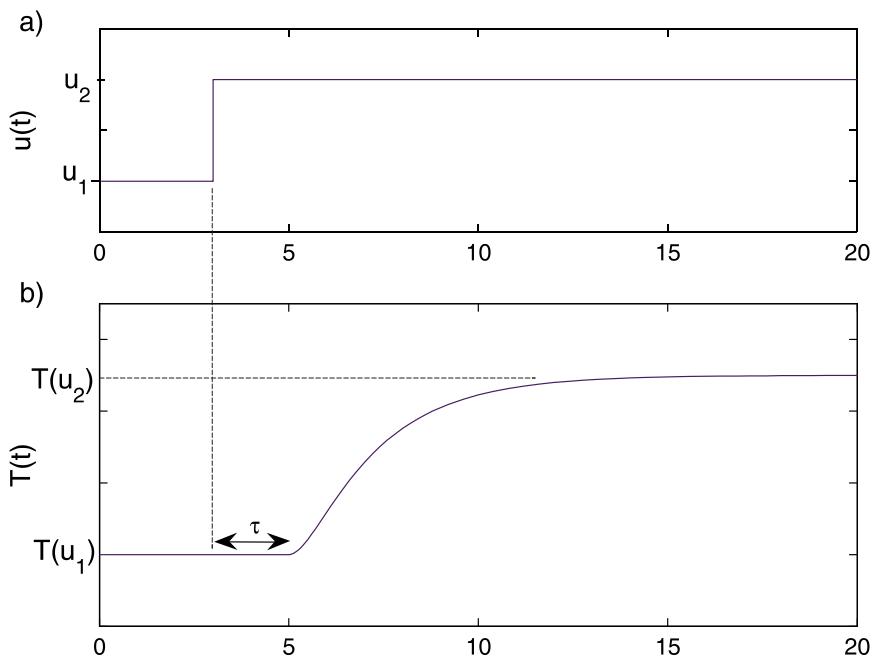
Stąd, z punktu widzenia układu regulacji:

- Temperatura T jest zmienną regulowaną,
- Sygnał u sterujący silownikiem zmieniającym położenie zaworu jest zmienną sterującą,
- Natężenie przepływu F_{we} oraz temperatury T_{we} i T_g są zakłóceniami.

Dla zaprojektowania układu regulacji jest potrzebny model dynamiki obiektu regulacji, tzn. pewien użyteczny sposób przedstawienia zależności wartości i zmian wielkości regulowanej od wartości i zmian sterowania i zakłóceń:

$$T(t) = f(F_g(u(t)), F_{we}(t), T_{we}(t), T_g(t)) \quad (3.1)$$

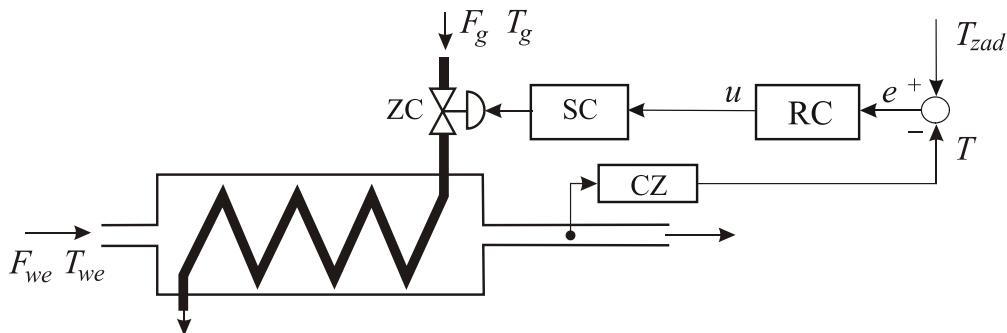
gdzie f oznacza odwzorowanie (funkcję) modelu. Rozważany wymiennik ciepła jest obiektem o dość typowej dynamice, wynikającej z bezwładności cieplnej. Jeśli, przy ustalonych wartościach zakłóceń i sterowania $u = u_1$, dokonamy skokowej zmiany wartości u (tzn. w efekcie zmiany położenia zaworu) od wartości u_1 do wartości u_2 , jak to przedstawiono na Rys. 3.3a, to wyjście naszego obiektu (temperatura $T(t)$) będzie się zmieniać tak, jak to przedstawiono na Rys. 3.3b.



Rys. 3.3. Skokowy przebieg sterowania $u(t)$ (górną krzywą) i odpowiadający mu przebieg temperatury wyjściowej $T(t)$ obiektu (tzw. odpowiedź skokowa, dolna krzywa).

Wartość temperatury T zareaguje na zmianę u dopiero po pewnym czasie τ , zwany czasem opóźnienia, narastając potem stopniowo do nowej wartości $T(u_2)$ (jest to narastanie o charakterze tzw. inercyjnym). Krzywa przedstawiona na Rys. 3.3b to tzw. odpowiedź skokowa obiektu regulacji. Jest to pewien prosty i łatwy do wyznaczenia sposób charakteryzowania własności dynamicznych obiektu, pewien model obiektu. Podkreślimy, że jest to model uzyskany w określonych warunkach (wartości sterowań u_1 i u_2 oraz pewne ustalone wartości zakłóceń F_{we} , T_{we} i T_g). Jeśli przedstawiony eksperyment powtórzymy dla innych ustalonych wartości zakłóceń i sterowań, to otrzymamy podobną krzywą, ale o nieco innych parametrach (w ogólności inne wartości czasu opóźnienia τ , ustalonej wartości końcowej $T(u_2)$, nieco inny charakter narastania). Dopiero odpowiednio pełna rodzina takich krzywych będzie w miarę wyczerpująco charakteryzować dynamikę obiektu. Jednakże, układ regulacji projektujemy najczęściej do pracy w otoczeniu pewnego punktu równowagi zwanego punktem pracy, charakteryzowanego wybranymi ustalonimi wartościami sterowania i zakłóceń, wykorzystując w tym celu model obiektu wyznaczony w tym punkcie pracy i tak projektując układ regulacji, aby pracował skutecznie w jego otoczeniu (tzn. przy pewnych odchyleniach sterowania i zakłóceń od wartości ustalonych definiujących punkt pracy).

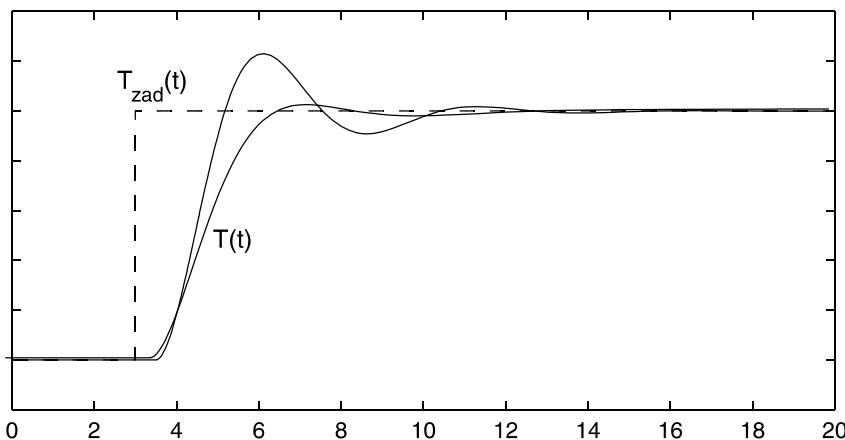
3.1. Regulacja ciągła



Rys. 3.4. Schemat układu regulacji ciągłej temperatury wody opuszczającej wymiennik ciepła (SC – siłownik zaworu o działaniu ciągłym, RC – regulator ciągły, CZ – czujnik pomiarowy, ZC – zawór o działaniu ciągłym).

Schemat układu regulacji ciągłej wymiennika ciepła przedstawia Rys. 3.4, gdzie założono, że siłownik ciągły (SC) i czujnik pomiarowy (CZ) wyposażone są w odpowiednie przetworniki sygnałowe. Regulator ciągły, tzn. generujący sygnał sterujący $u(t)$ mogący przyjmować dowolne wartości z pewnego zakresu, działa zgodnie z zaprojektowanym algorytmem swego działania, który np. dla popularnego regulatora PID można ogólnie przedstawić jako pewną funkcję dynamiczną r definiującą zależność sygnału $u(t)$ od wartości, pochodnej i całki uchybu regulacji $e(t) = T_{zad}(t) - T(t)$,

$$u(t) = r(e(t), \frac{de(t)}{dt}, \int_{t_0}^t e(s)ds) \quad (3.2)$$



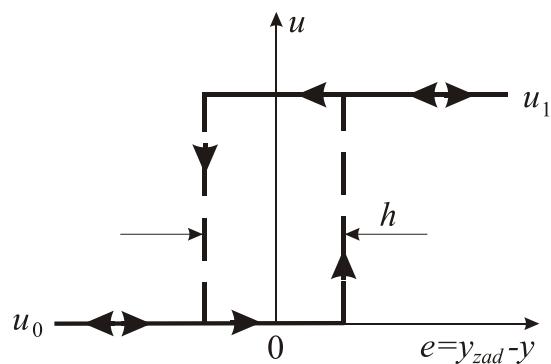
Rys. 3.5. Dwa typowe przebiegi wartości zmiennej regułowanej $T(t)$ uzyskane w układzie regulacji ciągłej przy dwóch różnych zestawach wartości parametrów regulatora PI, dla skokowej zmiany wartości zadanej $T_{zad}(t)$ (linia przerywana).

Na rysunku 3.5 przedstawiono skokowy przebieg wartości zadanej $T_{zad}(t)$ (górny wykres) i dwa typowe przebiegi zmiennej regułowanej $T(t)$ w układzie regulacji ciągłej z Rys. 3.4 dla obiektu o odpowiedzi skokowej z Rys. 3.3 i z regulatorem PI (tzn. PID z wyłączoną zależnością od pochodnej uchybu), uzyskane dla dwóch różnych zestawów

parametrów tego regulatora. Porównując przebiegi temperatury z Rys. 3.5 i Rys. 3.3 zauważmy, że przebiegi w układzie regulacji dochodzą w pobliżu ustalonej wartości docelowej znacznie szybciej, tzn. dynamika układu regulacji jest szybsza od dynamiki samego obiektu. Zauważmy także, że wartość regułowana osiąga w sposób trwały pożdaną wartość zadaną, po okresie przejściowym wywołanym zmianą wartości zadanej.

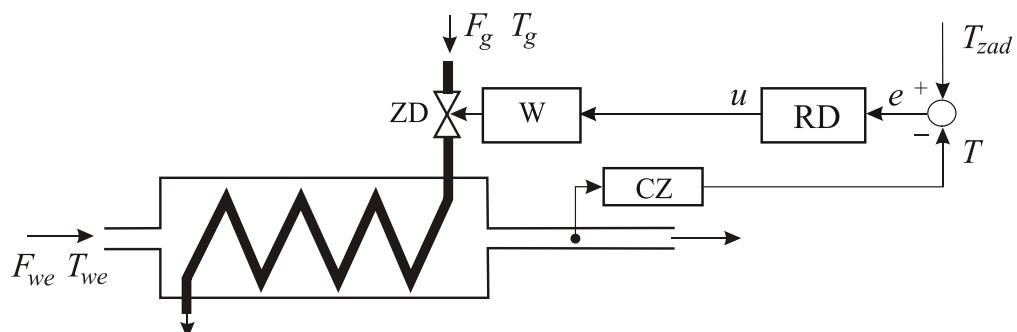
3.2. Regulacja dwupołożeniowa

Układem regulacji dwupołożeniowej nazywamy układ regulacji, w którym sygnał sterujący może przyjmować tylko dwie wartości (dwa "położenia"). Układy takie istnieją od bardzo dawna i są bardzo rozpowszechnione, szczególnie w układach regulacji temperatury; nazywane bywają też układami regulacji dwustawnej. Elementem generującym sygnał sterujący, pełniącym rolę tzw. regulatora dwupołożeniowego, jest element nieliniowy o charakterystyce przedstawionej na Rys. 3.6.



Rys. 3.6. Charakterystyka elementu dwupołożeniowego.

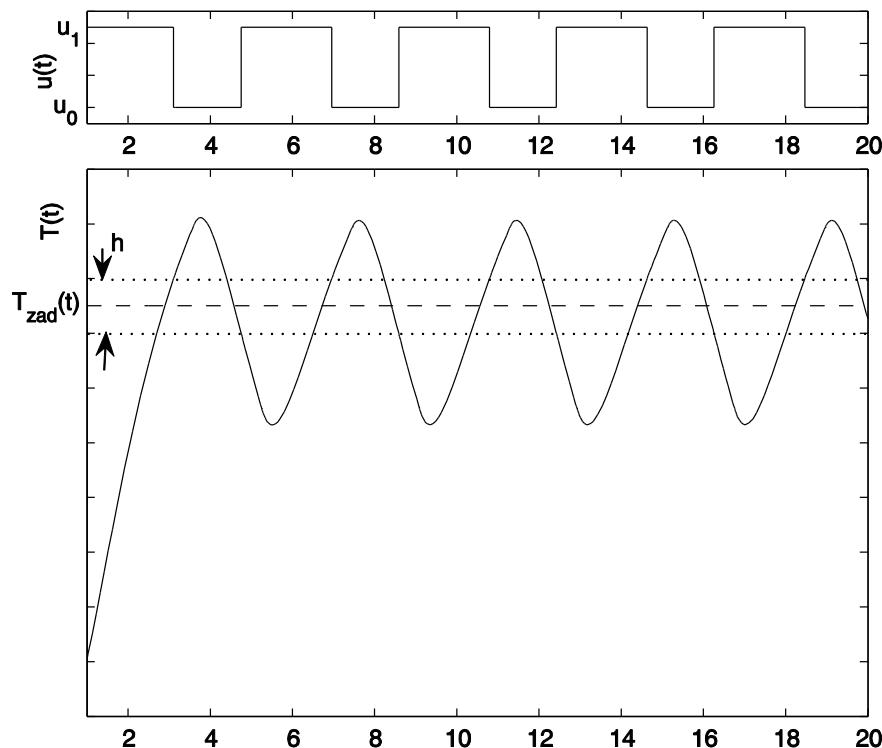
Element ten scharakteryzowany jest trzema wartościami: dwoma możliwymi wartościami sterowania u (u_0 i u_1) i wartością tzw. strefy histerezy h . Strefa histerezy charakteryzuje nieczułość w okolicy zera: zmiana sygnału sterującego z u_0 na u_1 (przy wzroście wartości uchybu e) zachodzi nie dla $e = 0$, ale dla $e = 0 + h/2$, podobnie zmiana sygnału sterującego z u_1 na u_0 (przy maleniu wartości uchybu e) zachodzi nie dla $e = 0$, ale dla $e = 0 - h/2$.



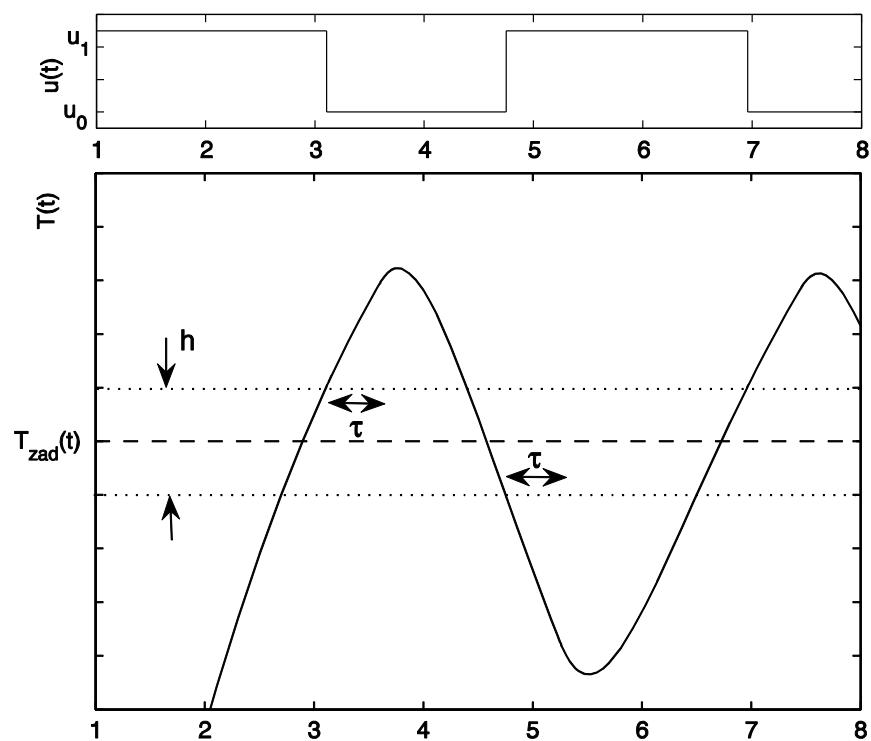
Rys. 3.7. Schemat układu regulacji dwupołożeniowej temperatury wody opuszczającej wymiennik ciepła (RD – regulator dwupołożeniowy, CZ – czujnik pomiarowy, ZD – zawór dwustawny, W – wzmacniacz).

Układ regulacji dwupołożeniowej wymiennika ciepła jest przedstawiony na Rys. 3.7, gdzie RD to regulator dwupołożeniowy o charakterystyce z Rys. 3.6 (wymieniony na tym rysunku odrębny blok wzmacniacza W nie wystąpi, jeśli moc sygnału z regulatora RD wystarcza do bezpośredniego wysterowania zaworu dwustawnego ZD).

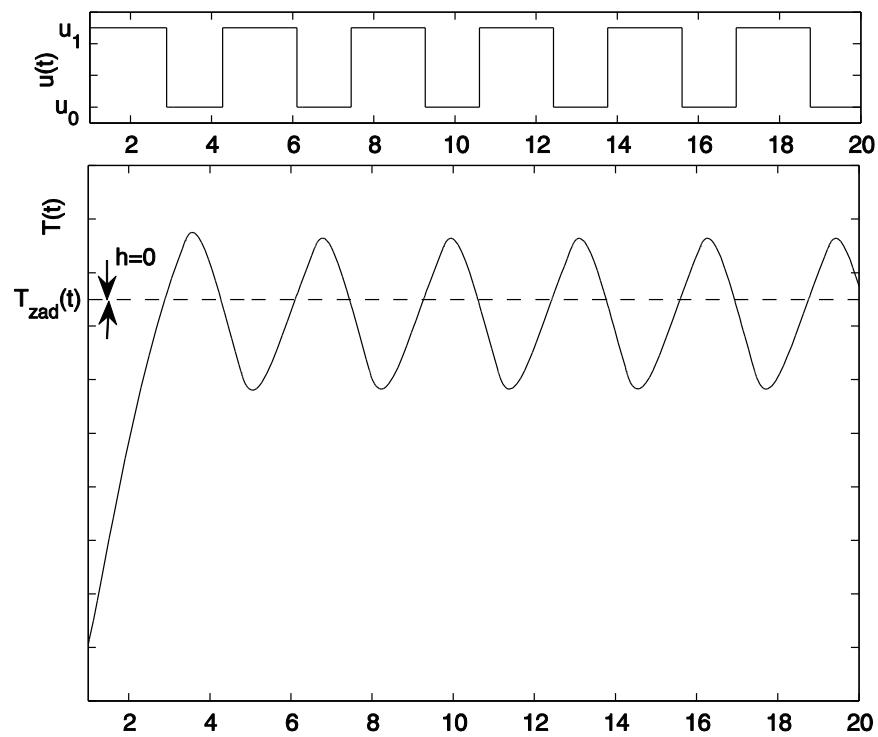
Przykładowe przebiegi temperatury $T(t)$ obiektu o charakterystyce skokowej z Rys. 3.3, wraz z odpowiadającymi im przebiegami sterowania $u(t)$, w układzie regulacji dwupołożeniowej z Rys. 3.7 przedstawiono na Rys. 3.8, 3.9, 3.10 i 3.11, po skokowym wzroście temperatury zadanej do wartości ustalonej pokazanej na tych rysunkach linią przerywaną. Rysunek 3.8 przedstawia stan ustalony, jaki został osiągnięty po wzroście temperatury $T(t)$ do wartości bliskich wartości zadanej $T_{zad}(t)$, w układzie z niezerową wartością histerezy h . Dzięki występowaniu istotnego opóźnienia τ (zob. Rys. 3.3) oraz niezerowej strefy histerezy w układzie powstaje stabilny przebieg oscylacyjny w otoczeniu wartości zadanej. Szczegółowy mechanizm jego powstawania dobrze ilustruje Rys. 3.9, będący powiększonym fragmentem rysunku poprzedniego, łącznie z Rys. 3.6, ilustrującym działanie regulatora dwupołożeniowego. Zmiana sterowania następuje nie bezpośrednio po zmianie znaku uchybu regulacji $e(t) = T_{zad}(t) - T(t)$, ale dopiero po osiągnięciu przez uchyb wartości $+h/2$ (przy wzroście uchybu od wartości ujemnych) lub $-h/2$ (przy maleniu uchybu od wartości dodatnich). Ponadto, wyjście obiektu zaczyna reagować na zmianę sterowania dopiero po czasie opóźnienia τ . Stąd długość okresu drgań ustalonych wyjścia zależy od wartości h i τ oraz od szybkości zmian wyjścia przy zastosowanych wartościach sterowania u_0 i u_1 . Przy występowaniu opóźnienia drgania występują nawet przy zerowej wartości strefy histerezy, co pokazuje Rys. 3.10 – mniejsze są wówczas jednakże amplituda i okres drgań.



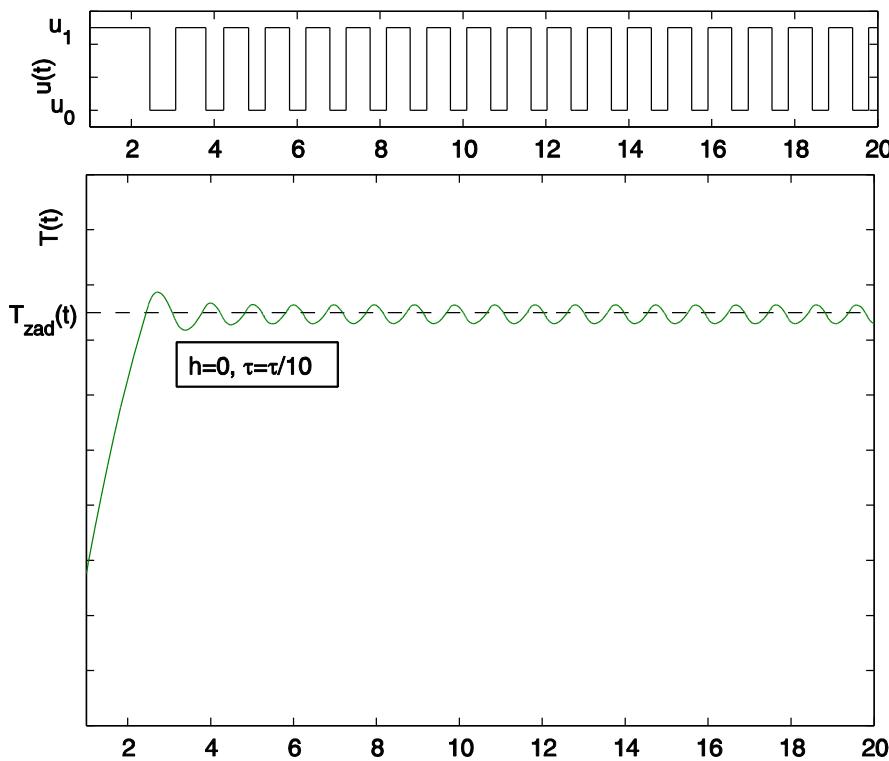
Rys. 3.8. Przebiegi w układzie regulacji dwupołożeniowej obiektu o charakterystyce skokowej przedstawionej na Rys. 3.3, dla niezerowej strefy histerezy.



Rys. 98H3.9. Powiększony fragment Rys. 3.8.



Rys. 3.10. Przebiegi w układzie regulacji dwupołożeniowej obiektu o charakterystyce skokowej przedstawionej na Rys. 3.3, dla zerowej strefy histerezy (element dwupołożeniowy bez histerezy).



Rys. 3.11. Przebiegi w układzie regulacji dwupołożeniowej z zerową strefą histerezy, z obiektem o charakterystyce skokowej z Rys. 3.3, ale z dziesięciokrotnie mniejszym opóźnieniem τ .

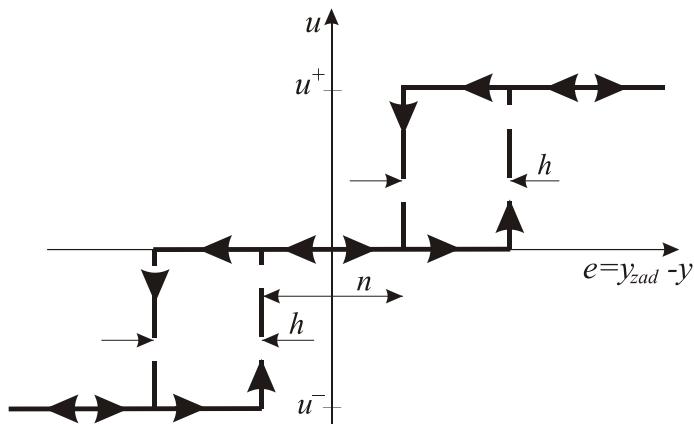
Układy regulacji dwupołożeniowej są układami prostymi i tanimi w realizacji, ze względu na prostotę realizacji sterowania dwupołożeniowego. Sterowanie to nie zapewnia jednak, z samej swej natury, dobrej jakości regulacji – w stanach ustalonych zmienna regulowana oscyluje wokół wartości zadanej. W wielu zastosowaniach, przede wszystkim w regulacji temperatury obiektów cieplnych, jest to akceptowalne i regulacja dwupołożeniowa jest bardzo rozpowszechniona. Dwie wartości sterowania odpowiadają bowiem w naturalny sposób sytuacjom włączenia bądź wyłączenia elementu grzejnego. Co więcej, oscylacja temperatury regulowanej wokół wartości zadanej bywa pożądana i celowo wykorzystywana. Powszechnie spotykanym przykładem jest tu regulacja temperatury w chłodziarkach, gdzie cykl oscylacji temperatury jest wykorzystywany do automatycznego odmrażania komory: jest on tak dobrany, aby w okresie wyższych wartości temperatury nastąpiło stopienie szronu i lodu zgromadzonego w okresie niskiej temperatury na elemencie chłodzącym (i odprowadzenie wody do parownika). Sterowanie w najpowszechniej spotykanych chłodziarkach sprężarkowych realizowane jest przez termostat, który włącza lub wyłącza agregat obwodu chłodzącego, pełniący rolę elementu wykonawczego.

Innym powszechnie spotykanym przykładem jest sterowanie dwupołożeniowe temperaturą żelazka elektrycznego. Elementem dwupołożeniowym jest tu np. odpowiednio wykonany przewodzący element bimetaliczny zmieniający swój kształt pod wpływem ciepła, który rozgrzewając się rozwiera styki obwodu elektrycznego odłączając element grzejny. Po odpowiednim ostygnięciu bimetal znowu powraca do kształtu zamkającego obwód grzejny, itd.

Inne powszechnie spotykane urządzenia codziennego użytku wykorzystujące regulację dwupołożeniową to termoalksy elektryczne, olejowe piecyki elektryczne czy kuchenki elektryczne ze sterowaniem termostatycznym.

3.3. Regulacja trójpołożeniowa z korekcją

Schemat regulacji trójpołożeniowej jest identyczny ze schematem regulacji dwupołożeniowej, jedynie elementem sterującym jest element trójpołożeniowy (zwany też elementem trójstawnym) przedstawiony na Rys. 3.12, w miejsce elementu dwupołożeniowego z Rys. 3.6.



Rys. 3.12. Charakterystyka elementu trójpołożeniowego (h – szerokość stref histerezy, n – szerokość strefy nieczułości).

Jednym z bardziej rozpowszechnionych zastosowań sterowania trójpołożeniowego są układy regulacji z tanimi stałoprędkościowymi elektrycznymi silnikami wykonawczymi, kiedy trzy możliwe stany sterowania to:

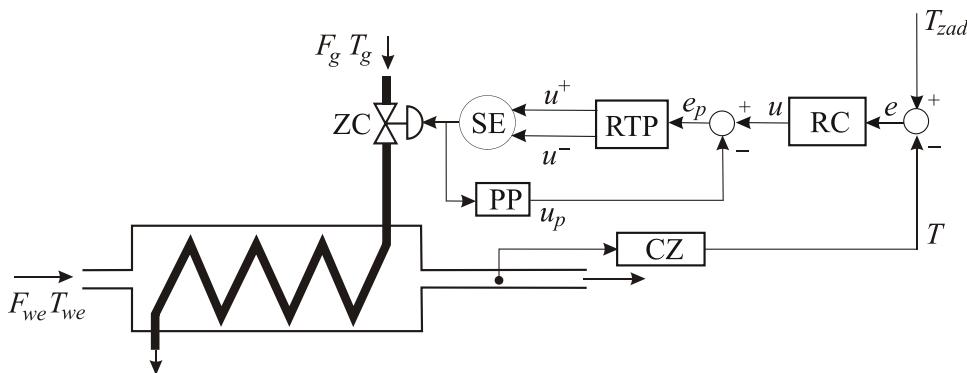
- załącz w prawo,
- stop (brak napięcia sterującego silnikiem),
- załącz w lewo (przeciwna polaryzacja napięcia sterującego silnikiem).

Zwróćmy uwagę, że w elemencie trójstawnym mamy dwa parametry nastawne: szerokość strefy nieczułości i szerokość stref histerezy. Dzięki występowaniu zerowego sterowania w zakresie niezerowej strefy nieczułości możliwe jest doprowadzenie układu regulacji do stanu równowagi, w którym nie występują oscylacje (z uchybem regulacji nie przekraczającym wartości $n/2 + h$) - ale warunkiem tego jest, aby środkowa ("zerowa") pozycja elementu trójpołożeniowego odpowiadała punktowi równowagi układu. W innych przypadkach stany równowagi w układzie regulacji trójpołożeniowej są oscylacyjne wokół wartości zadanej, podobne jak w układzie regulacji dwupołożeniowej.

Układy regulacji trójpołożeniowej z korekcją, zwane też układami regulacji krokowej, wykorzystują powyżej wymienioną możliwość uzyskiwania nieoscylacyjnego stanu równowagi z niewielkim uchybem regulacji, przy zastosowaniu trójpołożeniowego elementu sterującego stałoprędkościowym silnikiem elektrycznym (tani, niezawodny element wykonawczy) – dzięki wykorzystaniu w układzie regulacji dodatkowego sprzężenia zwrotnego i algorytmu regulacji ciągłej. Struktura takiego układu regulacji przedstawiona jest na Rys. 3.13, gdzie:

- RC – regulator (algorytm regulacji) o działaniu ciągłym,
- RTP – regulator trójpołożeniowy (o charakterystyce jak na rys. 3.12),
- SE – silnik elektryczny (stałoprędkościowy),

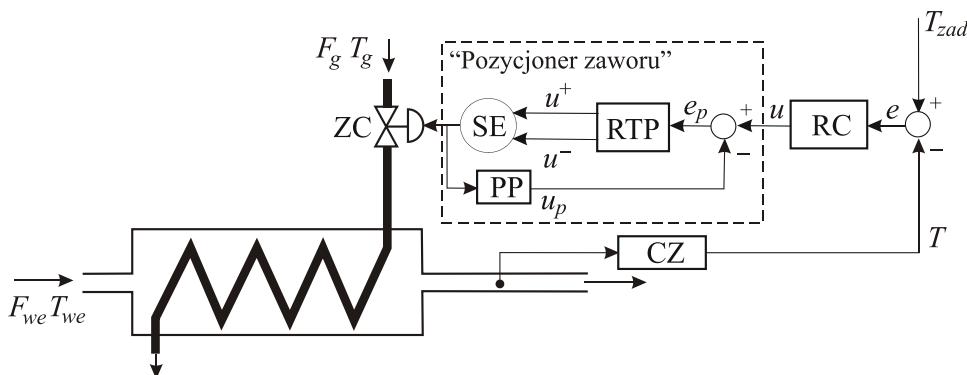
- PP – przetwornik pomiarowy (przetwarzający sygnał pomiarowy położenia zaworu na odpowiadający mu sygnał postaci u_p odpowiedniej do podania do regulatora ciągłego),
- CZ – czujnik pomiarowy wielkości regulowanej (temperatury),
- ZC – zawór regulacyjny o działaniu ciągłym.



Rys. 3.13. Struktura układu regulacji krokowej (regulacji trójpochodzeniowej z korekcją) wymiennika ciepła.

Celem działania wewnętrznej pętli regulacji trójpochodzeniowej jest ustawianie zaworu na pozycji odpowiadającej wartości wyjściowej regulatora ciągłego u , z dokładnością odpowiadającą strefie sygnału zerowego elementu trójpochodzeniowego. Nazwa "regulator krokowy" pochodzi od charakterystycznego zachowania się silnika w takim układzie – porusza on się "skokami", czy też "krokami" odpowiadającymi chwilom niezerowego sygnału wyjściowego elementu trójpochodzeniowego.

Regulacja krokowa umożliwia osiągnięcie dość wysokiej jakości regulacji, tak jak w regulacji ciągłej (ze strefą nieczułości regulatora ciągłego odpowiadającą strefie nieczułości elementu trójpochodzeniowego), mimo zastosowania silnika stałoprędkościowego sterowanego trójpochodzeniowo.

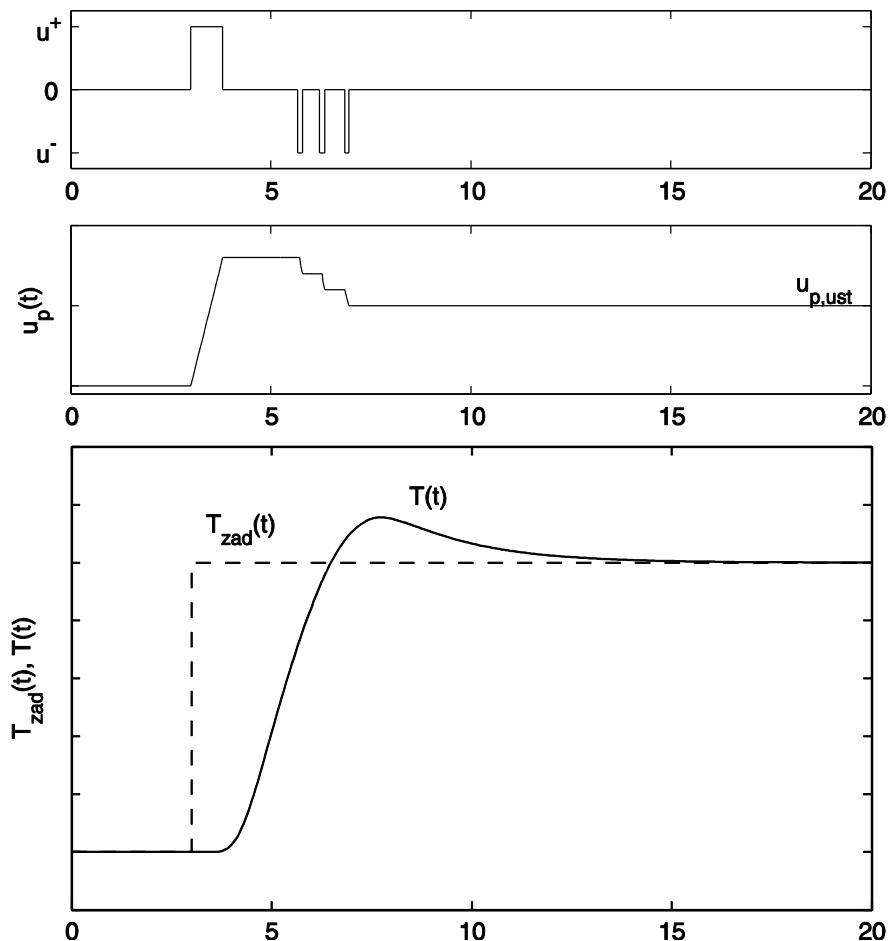


Rys. 3.14. Struktura układu regulacji z Rys. 3.13 widziana jako układ regulacji ciągłej z wewnętrzna pętlą regulacji trójpochodzeniowej traktowaną jako "pozycjoner zaworu"

Układ regulacji krokowej można bowiem traktować jako układ regulacji ciągłej, z wewnętrzna pętlą regulacji trójpochodzeniowej (regulacji pozycji zaworu) widzianą jako

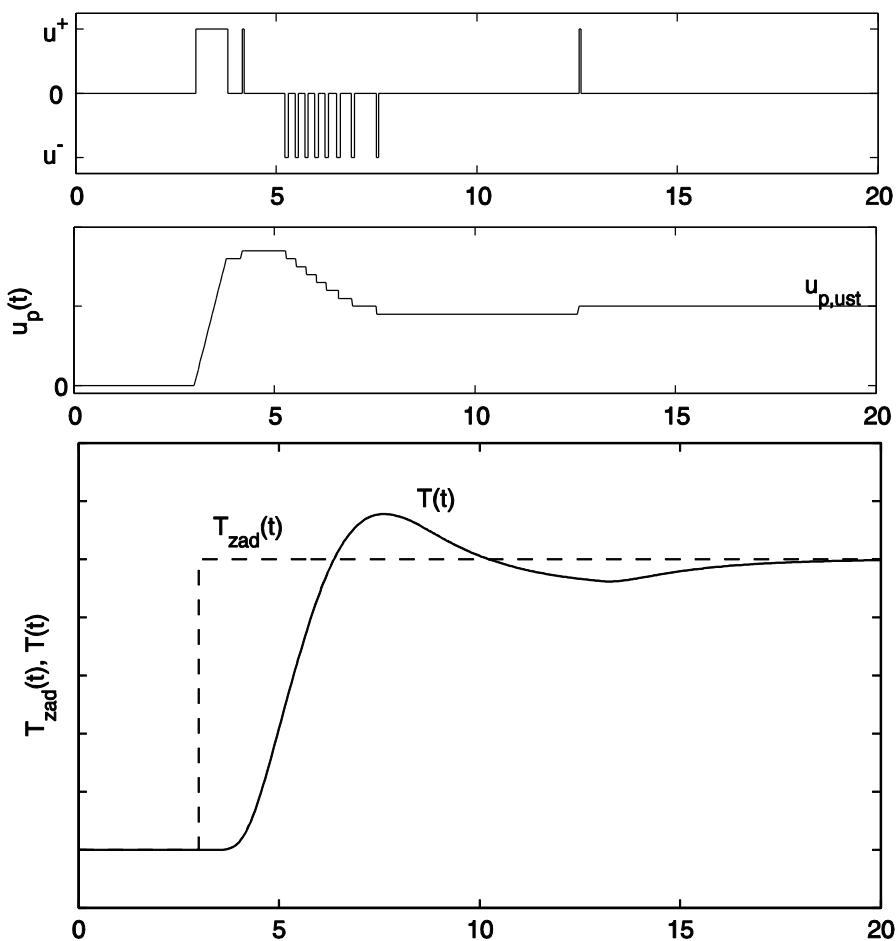
"pozycjoner zaworu" – tzn. elementem wykonawczym jest zawór o działaniu ciągłym wraz z tym pozycjonerem, jak to pokazano na Rys. 3.14. Przypomnijmy, że i w klasycznym układzie regulacji ciągłej z silnikiem (silnikiem) o działaniu ciągłym strefa nieczułości też jest typowo stosowana w regulatorze ciągłym, tak aby jego działanie nie było aktywowane w zakresie zmian uchybu odpowiadających dokładności pomiaru zmiennej regułowanej i pozycjonowania zaworu.

Przykładowe przebiegi wartości zadanej i wyjściowej temperatury, sygnału położenia zaworu $u_p(t)$ (zob. Rys. 3.13, 3.14) oraz sterowań dyskretnych u^+ , 0 i u^- na wyjściu regulatora krokkowego (elementu trójpołożeniowego) pokazane są na Rys. 3.15, dla regulatora ciągłego z algorytmem PI i elementu trójpołożeniowym z $n/2 = 0.1u_{p,ust}$, $h = n/2$, gdzie $u_{p,ust}$ jest wartością ustaloną sygnału $u_p(t)$, tj. również wartością ustaloną sterowania na wyjściu regulatora ciągłego (z dokładnością $n/2+h$) odpowiadającą ustalonej wartości zadanej temperatury T_{zad} .



Rys. 3.15. Przykładowe przebiegi w układzie regulacji krokowej (Rys. 3.13) z regulatorem ciągłym PI i elementem trójpołożeniowym o parametrach: $n/2 = 0.1u_{p,ust}$, $h=n/2$, gdzie:

- wykres górnny: sterowania dyskretne u^+ , 0, u^- na wyjściu elementu trójpołożeniowego,
- wykres środkowy: sygnał ciągły $u_p(t)$ odpowiadający pozycji zaworu na wejściu obiektu,
- wykres dolny: trajektorie wartości zadanej $T_{zad}(t)$ i aktualnej $T(t)$ na wyjściu układu.



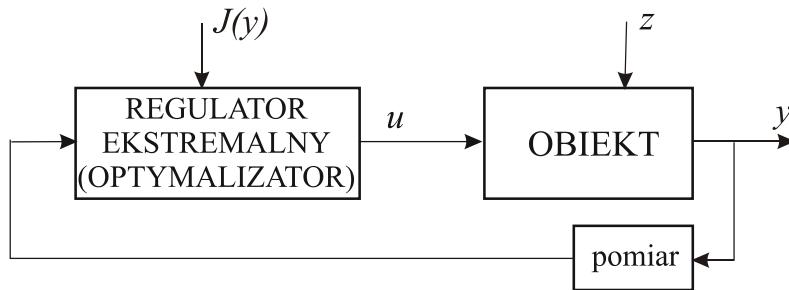
Rys. 3.16. Przykładowe przebiegi w układzie regulacji krokowej (Rys. 3.13) z regulatorem ciągłym PI i elementem trójpłozieniowym o parametrach $n/2 = 0.1u_{p,ust}$, $h = n/4$ (poszczególne wykresy jak na rysunku poprzednim).

Zwróćmy uwagę, że w wewnętrznej petli regulacji położenia zaworu może dojść do bardzo częstego przełączania elementu trójpłozieniowego, a w szczególności częstego przełączania na jednej z dwóch dwupołozieniowych części tego elementu odpowiadających odpowiednio sygnałom $u^+ \neq 0$ oraz $u^- \neq 0$, jeśli strefa histerezy elementu trójpłozieniowego jest zbyt mała. Strefa histerezy jest bowiem w istocie jedynym elementem przeciwdziałającym takiemu potencjalnie niekorzystnemu zjawisku, gdyż w rozważanej wewnętrznej pętli regulacji rolę "obiektu regulacji" pełni silnik, którego opóźnienie działania jest zwykle bardzo małe. Dla ilustracji, na Rys. 3.16 przedstawiono przebiegi w układzie regulacji identycznym jak odpowiadający Rys. 3.15, jedynie z dwukrotnie zmniejszoną strefą histerezy (do wartości $h = n/4 = 0.05u_{p,ust}$). Widać wyraźnie pojawienie się w przebiegach przejściowych efektu oscylacji na wyjściu elementu trójpłozieniowego, odpowiadających oscylacyjnemu przełączaniu na "dolnej" części dwupołozieniowej elementu trójpłozieniowego.

3.4. Regulacja ekstremalna

Regulacja ekstremalna jest szczególnym przypadkiem układu sterowania ze sprzężeniem zwrotnym, w którym celem działania regulatora jest wyznaczanie sygnałów sterujących obiektem w taki sposób, aby maksymalizować (lub minimalizować, co jest równoważne, bo $\min J(y) = -\max(-J(y))$) pewną funkcję, zwaną funkcją jakości działania układu. Schemat

regulacji ekstremalnej przedstawiony jest na Rys. 3.17.



Rys. 3.17. Struktura układu regulacji ekstremalnej.

Funkcja jakości działania układu $J(y)$, zależna od wyjścia układu y , może mieć charakter techniczny lub ekonomiczny. Wyjścia układu regulacji ekstremalnej y zależą od sterowań u , oznaczmy model tej zależności (tzn. model obiektu) przez F , $y = F(u)$, stąd w istocie regulator ekstremalny optymalizuje funkcję $J(F(u))$. W optymalizatorze może być znany lub nieznany model obiektu, jeśli jest nawet znany, to z reguły jest to niedokładne przybliżenie rzeczywistości, szczególnie zgrubne w zastosowaniach ekonomicznych. Stąd, jeśli regulator nawet wykorzystuje model obiektu, to wyznaczone przez niego sterowanie jest jedynie przybliżone i jest jedynie dobrym punktem startowym do bezpośredniego szukania na obiekcie jedną z metod optymalizacji bezpośredniej (tj. w istocie uporządkowaną metodą prób i błędów), które jest podstawą działania regulatora ekstremalnego, zwanego dlatego również optymalizatorem. Tzn., optymalizator generuje kolejne próbne wartości sterowania u , poprawiając je na podstawie obserwacji wyjścia obiektu i wynikających stąd wartości funkcji $J(y)=J(F(u))$ (zgodnie z optymalizacyjnym algorytmem przeszukiwania).

W działalności człowieka jest wiele przykładów regulacji ekstremalnej. W skali makroekonomicznej takim przykładem może być ustalanie optymalnych stawek podatkowych, np. akcyzowych na pewne produkty, przyjmując za funkcję jakości (do maksymalizacji) sumę wpływów z tego tytułu do budżetu państwa. Klasycznym przykładem może tu być stawka akcyzy na alkohol. Regulatorem ekstremalnym jest tu rząd (ministerstwo finansów). Działanie regulatora polega np. na stopniowym zwiększeniu stawki akcyzy i obserwacji wpływów z tego tytułu do budżetu państwa za pewien okres czasu. Jeśli wpływy zaczynają maleć (z tytułu przemytu, pokątnej produkcji w szarej strefie, itp., które to działania stają się coraz bardziej opłacalne przy wyższych stawkach akcyzy), to stawek dalej się nie zwiększa, a nawet nieco je obniża – aby osiągnąć maksymalizację sumy wpływów. Przypadek takiego działania wystąpił również w Polsce.

4. Modelowanie obiektów dynamicznych

Przedstawimy teraz ważne dla analizy i projektowania układów sterowania sposoby opisu zachowania się interesujących nas obiektów lub procesów. Podstawową rolę odgrywają tutaj modele matematyczne, umożliwiające zarówno formalną jak i wspomagana komputerem analizę działania tych obiektów. W szczególności, w oparciu o przykłady, zilustrowane zostanie postępowanie przy budowie modeli matematycznych, ich typowe postacie. Następnie omówimy podstawowe dla dalszych rozważań zagadnienie identyfikacji oraz linearyzacji modeli.

4.1. Modele obiektów (procesów).

Model jest zobrazowaniem zachowania się obiektu lub procesu w wygodnej dla nas formie. Możemy budować modele fizyczne (w odpowiedniej skali), modele analogowe (np. z wykorzystaniem układów elektronicznych – maszyny analogowe) lub modele matematyczne. Modele matematyczne stanowią podstawę do realizowanych w postaci procedur symulacyjnych modeli komputerowych.

W powyższym zdaniu mówimy o obiektach lub procesach. Rozróżnienie pomiędzy obiektem i procesem nie jest szczególnie istotne z punktu widzenia samego modelowania; jednak w pewnych sytuacjach jest potrzebne dla jasnego wyrażenia, czy chodzi nam o opis zachowania się jakiegoś wyodrębnionego fizycznie fragmentu środowiska, na przykład pojazdu mechanicznego, czy też interesujemy się jedynie pewnym zjawiskiem, procesem, w ramach danego obiektu, na przykład procesem wymiany ciepła pomiędzy danym pomieszczeniem a jego otoczeniem. W dalszym tekście będziemy używać terminów obiekt bądź proces w zależności od kontekstu; przy omawianiu teorii będzie na ogół stosowane pojęcie obiekt.

Sztuka budowy odpowiednich, zachowujących tzw. skalę dynamiczną, **modeli fizycznych**, na przykład modeli obiektów wielokrotnie pomniejszonych w stosunku do właściwego obiektu, wymaga szczególnych umiejętności; modele takie tworzone są zwłaszcza w tych sytuacjach, w których nie potrafimy stworzyć odpowiednio wiarygodnych modeli innego typu lub wtedy, gdy konieczne jest wykonanie eksperymentów w oparciu o obiekt fizyczny, zaś nie można, z różnych powodów, wykonać takich eksperymentów z wykorzystaniem rzeczywistego pełnowymiarowego rozpatrywanego obiektu (procesu). W ramach niniejszego wykładu nie będziemy rozważać budowy oraz wykorzystania modeli fizycznych. Należy jednak podkreślić, że w wielu ważnych przypadkach, na przykład przy projektowaniu samolotów, prowadzenie eksperymentów w oparciu o modele fizyczne jest niezbędne.

Model analogowy tworzony jest z kolei także jako model w istocie fizyczny, ale wykorzystujący obiekt fizyczny innego rodzaju niż obiekt (proces) rzeczywisty do zobrazowania działania tegoż obiektu. W szczególności, modele analogowe na ogół budowano w oparciu o specjalnie do tego przystosowane elementy elektroniczne będące na wyposażeniu tzw. maszyn analogowych. Użycie w tym miejscu czasu przeszłego uzasadnione jest tym, że obecnie, wobec bardzo daleko posuniętego rozwoju maszyn cyfrowych, komputerów, praktycznie zarzucono posługiwanie się maszynami i modelami analogowymi. Nie będziemy się też nimi dalej zajmować.

Model matematyczny, jako forma najbardziej obecnie rozpowszechniona, tworzony jest w postaci zależności matematycznych, najczęściej równań algebraicznych lub równań różniczkowych z towarzyszącymi im – w razie potrzeby – ograniczeniami nałożonymi na wartości odpowiednich zmiennych. Model taki można używać po przeprowadzeniu jego strojenia, w szczególności po dokonaniu identyfikacji i weryfikacji wartości parametrów modelu, a także – w razie potrzeby – po przeprowadzeniu linearyzacji. Modele matematyczne

są obecnie, w zdecydowanej większości przypadków, przekształcane następnie do postaci procedur obliczeniowych pozwalających na dokonywanie obliczeń symulujących działanie rozważanego obiektu bądź procesu.

4.2. Modele matematyczne; rodzaje i budowa

Modele matematyczne możemy, z grubsza, podzielić na:

- **Modele teoretyczne** (fenomenologiczne, *first principle models*)

Zależności składające się na taki model wynikają z dobrego rozumienia istoty działania obiektu; na podstawie fizycznych jego właściwości (wykorzystujemy prawa fizyki, opis dynamiki reakcji chemicznych, itp.) tworzymy spójny zestaw relacji (równań) pozwalający na wyznaczanie interesujących nas wielkości w oparciu o wielkości wejściowe oraz warunki początkowe (w przypadku obiektów o stałych rozłożonych, opisywanych przy pomocy równań różniczkowych cząstkowych, także warunki brzegowe). Dane pochodzące z obserwacji uzupełniają jedynie naszą wiedzę, w szczególności pozwalają ustalić lub, najczęściej, jedynie uściślić wartości parametrów modelu.

- **Modele empiryczne** (typu „czarnej skrzynki” - „*black box*” models)

Budując te modele nie wnikamy w fizyczną istotę (podstawy) danego zjawiska bądź obserwowanego zachowania się obiektu. Stosujemy różne modele bazowe, na przykład z góry założone liniowe lub nieliniowe zależności algebraiczne, różne typowe elementy dynamiczne, jak inercja z opóźnieniem, człon oscylacyjny, itp. Jedyną podstawę identyfikacji (strojenia) tych modeli, a następnie akceptacji modelu jako odpowiadającego naszym potrzebom, stanowią dane pochodzące z obserwacji zachowania się rzeczywistego obiektu bądź procesu.

- **Modele konceptualne (konceptualno-emeryczne)**

Tworząc taki model wyobrażamy sobie dany obiekt lub proces w postaci pewnego uproszczonego jego analogu fizycznego; na przykład odcinek rzeki wyobrażamy sobie jako kaskadę małych szeregowo połączonych zbiorników. Podstawę identyfikacji parametrów takiego modelu także stanowią dane pochodzące z obserwacji. Liczymy jednak na to, że liczba parametrów będzie w przypadku tego typu modelu mniejsza niż w sytuacji, gdy tworzymy model zewnętrzny. Model, który możemy określić mianem modelu konceptualnego, stanowi, w istocie, formę pośrednią pomiędzy modelem fizykalnym stworzonym w oparciu o zrozumienie i zapis praw fizycznych rządzących działaniem danego obiektu, a wspomnianym modelem w pełni zewnętrznym, zwany także modelem typu czarna skrzynka, w przypadku którego próbujemy po prostu dobrać taki opis matematyczny, który pozwoli na dostatecznie dobre dopasowanie wielkości obliczanych przy pomocy tego modelu do danych pochodzących z obserwacji.

Podstawowe pytanie poprzedzające tworzenie modelu musi dotyczyć tego, do czego jest on nam potrzebny ?

Punkt wyjścia stanowią wielkości, których zachowanie chcemy opisać, tj. których przebiegi w czasie, wyznaczane z określoną dokładnością, nas interesują; inne wielkości nie powinny nas na wstępie w ogóle sobą zajmować – nie należy modelować dla samego modelowania. Nawiązuje do tego podana poniżej ważna reguła ujęta jako zasada modelu wystarczającego.

Podstawowe reguły postępowania przy budowie modelu są następujące:

- Model powinien być najprostszy jak to tylko w danej sytuacji jest możliwe (dopuszczalne); ten postulat można wyrazić jako zasadę modelu wystarczającego – nie należy wprowadzać niepotrzebnych komplikacji wówczas, gdy prostszy opis jest wystarczający.
- Możliwości identyfikacji parametrów modelu powinny być uwzględnione na etapie projektowania i tworzenia tego modelu; trzeba brać pod uwagę zarówno dostępne dane, które można będzie wykorzystać do strojenia parametrów modelu, jak i wymaganą dokładność modelu.

4.2.1. Modele teoretyczne (fenomenologiczne - fizyczne, fizyczno-chemiczne)

Model teoretyczny tworzymy wykorzystując prawa fizyki, bilanse energii i materiału, modele reakcji chemicznych, itp.

Model teoretyczny na ogół wymaga wprowadzenia mniejszej liczby parametrów niż model konceptualny, a tym bardziej model empiryczny. Pewne parametry modelu teoretycznego, na przykład stałe fizyczne, mogą być przy tym niezależne od rozpatrywanego przypadku (np. przyspieszenie ziemskie g).

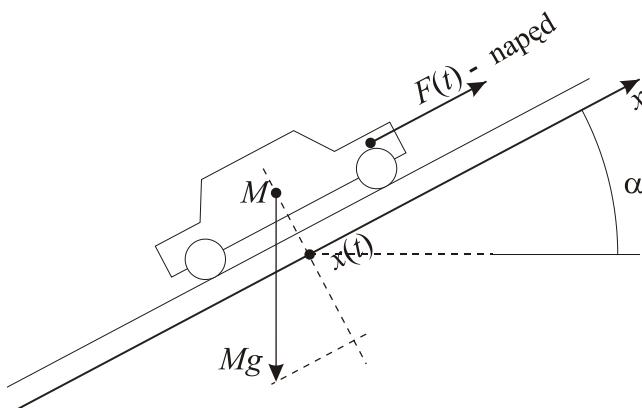
Przykład 4.1 (dynamika pojazdu).

Rozważmy opis ruchu pojazdu poruszającego się po prostoliniowej nachylonej pod kątem α do poziomu drogi. Położenie pojazdu oznaczamy w danej chwili czasu t jako $x(t)$. Interesuje nas stworzenie opisu zmian tego położenia w zależności od warunków, tj. nachylenia drogi i oporów ruchu, a także od innych oddziaływań, w tym czynności podejmowanych przez osobę kierującą pojazdem.

Korzystając z drugiej zasady dynamiki w ruchu postępowym; możemy napisać następujące równanie:

$$M\ddot{x}(t) = F(t) - Mg \sin \alpha - b\dot{x}(t), \quad (4.1)$$

gdzie: M – masa pojazdu, g – przyspieszenie ziemskie, α – kąt nachylenia drogi do poziomu, b – współczynnik oporów ruchu przy założeniu liniowej zależności siły oporu (tarcia potoczystego oraz oporu powietrza) od prędkości $\dot{x}(t) = dx(t)/dt$ $\dot{x}(t) = dx(t)/dt$.



Rys. 4.1. Pojazd poruszający się po drodze prostoliniowej.

Jeśli wartość $F(t)$ jest traktowana jako narzucona (na przykład określona przez układ sterowania), to nie ma potrzeby dalszego modelowania: $F(t)$ jest wielkością wejściową; M , g , α , b są parametrami modelu. Może tak być, że $\alpha = \alpha(t)$, $b = b(t)$, ale nadal będą to

(zmienne w czasie) parametry modelu lub, ewentualnie, wielkości wejściowe niesterowane, na które nie mamy wpływu. Przyjmując

$$\begin{aligned}x_1(t) &\stackrel{df}{=} x(t) \\x_2(t) &\stackrel{df}{=} \dot{x}(t)\end{aligned}$$

Otrzymujemy model w postaci układu równań

$$\begin{cases}\dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = \frac{1}{M} F(t) - g \sin \alpha - \frac{b}{M} x_2(t)\end{cases} \quad (4.2)$$

Jest to układ równań stanu w postaci normalnej; stan obiektu określają w tym przypadku dwie wielkości: aktualne położenie $x_1(t)$ oraz aktualna prędkość $x_2(t)$. Warunki początkowe są w postaci: $x_1(t_0) = x(t_0)$, $x_2(t_0) = \dot{x}(t_0)$, gdzie t_0 – chwila początkowa; na ogół przyjmujemy takie pozycjonowanie zera na osi czasu, że $t_0 = 0$.

Niech $F(t) = a \cdot u(t - T_0)$, gdzie T_0 – opóźnienie, $u(t - T_0)$ – oddziaływanie kierowcy, z uwzględnieniem opóźnionej reakcji, zaś $a \neq 0$ – odpowiednia stała. Zakładamy, że $b \neq 0$. Wówczas drugie z równań, opisujące zmiany prędkości możemy zapisać jako osobny model prędkości samochodu; prędkość $x_2(t) = \dot{x}(t)$, oznaczmy jako $y(t)$,

$$\frac{M}{b} \ddot{y}(t) + y(t) + \frac{M}{b} g \sin \alpha = \frac{a}{b} \cdot u(t - T_0)$$

Przyjmując oznaczenia $T = M/b$, $a_0(\alpha) = (M/b)g \sin \alpha$, $k_0 = a/b$, możemy powyższe równanie zapisać w postaci:

$$T \cdot \ddot{y}(t) + y(t) + a_0(\alpha) = k_0 \cdot u(t - T_0) \quad (4.3)$$

Jest to równanie różniczkowe pierwszego rzędu z opóźnieniem w funkcji wejściowej (sterowaniu).

Następnie, przyjmując stałą w czasie wartość sterowania \bar{u} oraz zerową wartość przyspieszenia, $\ddot{y}(t) = 0$, otrzymujemy równanie ruchu ustalonego, wyrażające zależność prędkości ustalonej \bar{y} od stałego sterowania \bar{u} (z parametrem nachyleniu drogi α):

$$\bar{y} = k_0 \bar{u} - a_0(\alpha) \quad (4.4)$$

Zależność tego typu nazywana jest charakterystyką statyczną obiektu. W naszym przykładzie jest to prosta funkcja afiniczna.

Wprowadzamy teraz zmienne przyrostowe, jako odchyłki odpowiednich wielkości od wyznaczonego powyżej stanu ustalonego:

$$\begin{aligned}y(t) &= \bar{y} + \Delta y(t) \\u(t) &= \bar{u} + \Delta u(t)\end{aligned}$$

Po wstawieniu tych wielkości do równania opisującego zmiany prędkości otrzymujemy model opisujący zmiany tych odchyłek w czasie

$$T \cdot \Delta \ddot{y}(t) + \Delta y(t) = k_0 \cdot \Delta u(t - T_0) \quad (4.5)$$

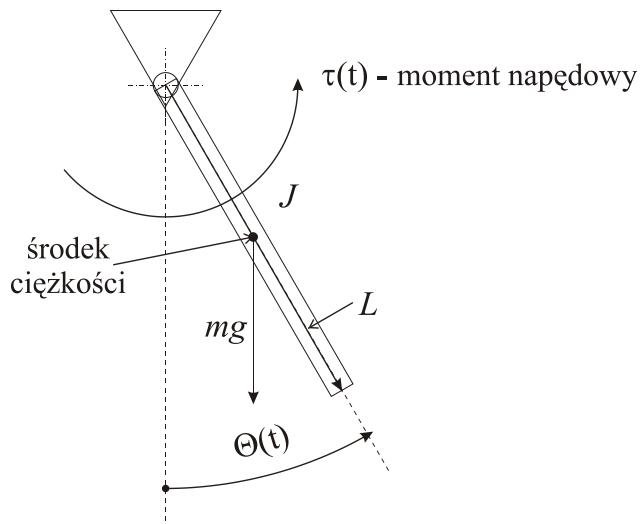
Jest to szeroko wykorzystywany model liniowy pierwszego rzędu, określany mianem obiektu Inercja+Opóźnienie (I+O), w literaturze anglojęzycznej spotykamy często skrót FOPD (*First*

Order Plus Delay).

□

Przykład 4.2 (wahadło fizyczne).

Jako drugi obiekt rozważmy wahadło fizyczne z napędem w postaci momentu siły przykładowego w przegubie obrotowym.



Rys. 4.2. Wahadło fizyczne z napędem.

Interesują nas wartości kąta wychylenia $\Theta(t)$, powiedzmy dla $t \geq t_0$; w przypadku obiektów o własnościach niezmiennych w czasie, jak w tej sytuacji, możemy zawsze chwilę początkową przyjmować jako $t_0 = 0$. Korzystamy z drugiej zasady dynamiki w ruchu obrotowym; równanie wyrażające zastosowanie tej zasady w naszym przypadku ma postać:

$$J \cdot \ddot{\Theta}(t) = -mgl \cdot \sin \Theta(t) + \tau(t) - b \cdot \dot{\Theta}(t)$$

gdzie

$b \cdot \dot{\Theta}(t)$ – opory ruchu (tarcie lepkie),

$\tau(t)$ – moment napędowy względem osi obrotu, kształtowany zewnętrznie,

J – moment bezwładności masy ruchomej względem osi obrotu,

$m, l=L/2, g$ – odpowiednio: masa części ruchomej, odległość środka ciężkości tej masy od osi obrotu, przyspieszenie ziemskie.

Równanie powyższe, po niewielkim przekształceniu, przyjmuje standardową postać modelu dynamicznego drugiego rzędu w postaci równania różniczkowego:

$$\ddot{\Theta}(t) + \frac{b}{J} \dot{\Theta}(t) + \frac{mgl}{J} \sin \Theta(t) = \frac{1}{J} \tau(t) \quad (4.6)$$

dla którego warunkami początkowymi są $\Theta(t_0), \dot{\Theta}(t_0)$.

Powyższe równanie można natychmiast przekształcić do układu równań w postaci normalnej, wprowadzając zmienne stanu odpowiadające wychyleniu wahadła $x_1(t)$ oraz prędkości kątowej tego wychylenia $x_2(t)$:

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = -\frac{mgl}{J} \sin x_1(t) - \frac{b}{J} x_2(t) + \frac{1}{J} \tau(t) \end{cases} \quad (4.7)$$

Warunkami początkowymi są $x_1(t_0)$, $x_2(t_0)$. Dwuwymiarowy stan obiektu $x(t)$ ma postać

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \in \mathbb{R}^2 \quad \square$$

Układ równań dynamiki w postaci normalnej

W ogólnym przypadku układ równań stanowiący model obiektu (procesu) dynamicznego w postaci normalnej dany jest jako:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), u(t)) \quad (4.8)$$

gdzie $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{x}(t_0)$ – warunek początkowy.

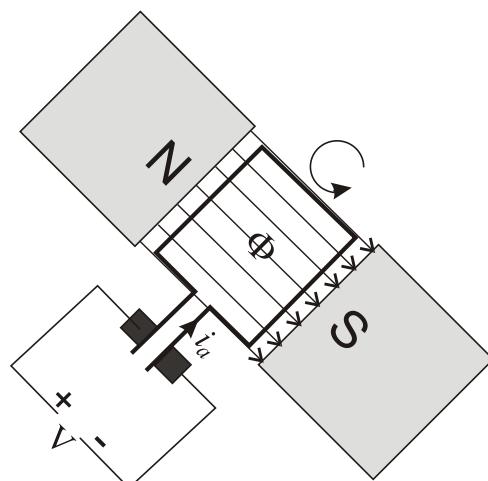
Uwaga

Model w takiej postaci możemy przekształcić do komputerowego modelu symulacyjnego. W tym celu potrzebny jest schemat numerycznego rozwiązania z czasem dyskretnym. Najprostsza wersja dyskretyzacji w czasie polega na rozważeniu sekwencji chwil czasu t_i , gdzie $t_i = t_0 + ih$, $i=0,1,2,\dots$, h jest stałym małym krokiem, a następnie przybliżenia poszukiwanego rozwiązania powyższego układu równań różniczkowych w tych chwilach poprzez ciąg wartości $\mathbf{x}_p(t_i)$ wyznaczonych w następujący sposób:

$$\mathbf{x}_p(t_{i+1}) = \mathbf{x}_p(t_i) + h \cdot \mathbf{f}(\mathbf{x}_p(t_i), u(t_i)), \quad \mathbf{x}_p(t_0) = \mathbf{x}(t_0) \quad (4.9)$$

Przedstawiony powyżej najprostszy schemat numerycznego, przybliżonego, rozwiązania układu równań różniczkowych w postaci normalnej nosi nazwę Metody Eulera; wymaga ona stosowania bardzo małe wartości kroku dyskretyzacji h dla zapewnienia stabilności i odpowiedniej dokładności. Tym niemniej, dzięki rosnącej mocy obliczeniowej komputerów, jej stosowanie wiąże się z coraz mniejszymi kosztami i metoda jest stosowana.

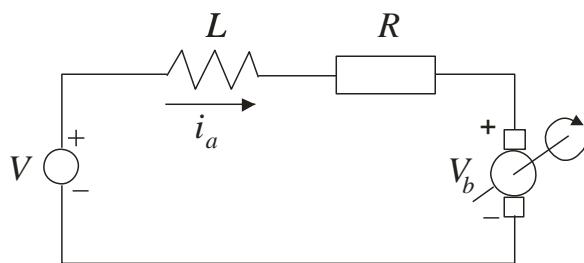
Przykład 4.3 (modelu silnika prądu stałego - obiektu mechaniczno-elektrycznego).



Rys. 4.3. Schemat konstrukcji silnika prądu stałego z magnesem trwałym.

Schemat silnika prądu stałego z trwałym magnesem stojana przedstawiony jest na rysunku 4.3. W uzwojeniu wirnika znajdującym się w stałym polu magnetycznym o indukcji Φ (wytworzonym przez magnes trwały) płynie prąd i_a wywołany przyłożonym napięciem V , co powoduje powstanie momentu obrotowego M proporcjonalnego do natężenia prądu, gdzie współczynnik proporcjonalności (stała mechaniczna) k_m zależy od strumienia Φ i liczby zwojów wirnika, $M = k_m i_a$, natomiast opory ruchu (momenty przeciwdziałające ruchowi) wynikają z momentu bezwładności wału silnika (z wirnikiem) J , oporów ruchu proporcjonalnych do prędkości obrotowej (tarcie) i momentu obciążenia M_{obc} . Stąd równanie dynamiki ruchu wirnika, z prędkością obrotową (kątową) oznaczoną przez ω :

$$k_m i_a = J \frac{d\omega}{dt} + b\omega + M_{obc} \quad (4.10)$$



Rys. 4.4. Schemat zastępczy obwodu elektrycznego wirnika silnika prądu stałego.

Schemat zastępczy obwodu elektrycznego wirnika przedstawiony jest na Rys. 4.4. Ruch w polu magnetycznym uzwojenia z płynącym prądem powoduje indukowanie się siły przeciwelektromotorycznej V_b , proporcjonalnej do prędkości obrotowej (i indukcji magnetycznej pola Φ oraz ilości zwojów wirnika – stałych), stąd $V_b = k_e \omega$ i równanie obwodu elektrycznego:

$$V = Ri_a + L \frac{di_a}{dt} + k_e \omega \quad (4.11)$$

gdzie współczynnik proporcjonalności k_e to tzw. stała elektryczna, a R i L to zastępcze oporność i indukcyjność uzwojeń wirnika, odpowiednio.

Równania (4.10) i (4.11) opisują dynamikę silnika. Oznaczając napięcie sterujące silnika V jako wejście sterujące u oraz wprowadzając zmienne stanu

$$x_1 \stackrel{\text{def}}{=} \omega, \quad x_2 \stackrel{\text{def}}{=} i_a$$

można równania te zapisać w postaci normalnej

$$\begin{aligned} \dot{x}_1(t) &= -\frac{b}{J} x_1(t) + \frac{k_m}{J} x_2(t) - \frac{1}{J} z(t), \\ \dot{x}_2(t) &= -\frac{k_e}{L} x_1(t) - \frac{R}{L} x_2(t) + \frac{1}{L} u(t), \end{aligned} \quad x(t_0) = [x_1(t_0) \ x_2(t_0)]^T, \quad (4.12)$$

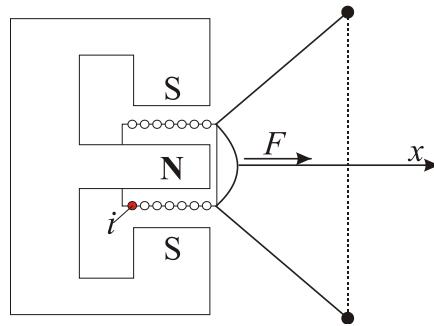
gdzie moment obciążenia M_{obc} oznaczono jako zakłócenie z .

Standardowo modele liniowe zapisujemy w postaci wektorowo-macierzowej, jak poniżej

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} -\frac{b}{J} & \frac{k_m}{J} \\ -\frac{k_e}{L} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} u(t) + \begin{bmatrix} -\frac{1}{J} \\ 0 \end{bmatrix} z(t) \Leftrightarrow \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) + \mathbf{f}z(t)$$

Model powyższy jest przykładem modelu liniowego o stałych współczynnikach. \square

Przykład 4.4 (model głośnika ze stałym magnesem).



Rys. 4.5. Budowa głośnika ze stałym magnesem.

Przedstawiony na rysunku obiekt składa się z: stałego magnesu, uzwojenia z płynącym prądem o zmiennym natężeniu i oraz zespółonej z uzwojeniem membrany zamocowanej na brzegu do obudowy głośnika. Uzwojenie zasilane jest ze źródła o zmiennym w czasie napięciu u ; napięcie to możemy traktować jako zmienną wejściową.

Budowę modelu rozpoczynamy od założenia, że membranę z uzwojeniem traktujemy jako sztywną strukturę o masie M posadowioną na zawieszeniu modelowanym jako sprężyna o współczynniku sprężystości k , przemieszczającą się w kierunku osi x i poddawaną działaniu siły F wywieranej na przewodnik z płynącym prądem znajdujący się w polu magnetycznym oraz napotykającej na opory ruchu o sile proporcjonalnej do prędkości. Korzystamy więc ponownie z drugiej zasady dynamiki w postaci:

$$M\ddot{x} = F - b \cdot \dot{x} - kx$$

W tym równaniu membrany, podobnie jak i w następnych zależnościach pominiemy oznaczanie zależności od czasu poszczególnych wielkości; przy pomocy kropki nad symbolem danej wielkości oznaczamy jej pochodną w czasie (prędkość zmian), zaś dwiema kropkami oznaczamy druga pochodna (przyspieszenie).

W powyższym równaniu membrany musimy opisać zmiany w czasie siły F ; b przyjmujemy jako stały w czasie współczynnik oporu ruchu. Wartość siły określamy wykorzystując znane z fizyki prawo, tzw. prawo motorowe, czyli zależność wyrażającą siłę działającą na przewodnik z płynącym prądem umieszczony w polu magnetycznym o indukcji B prostopadle do linii sił pola:

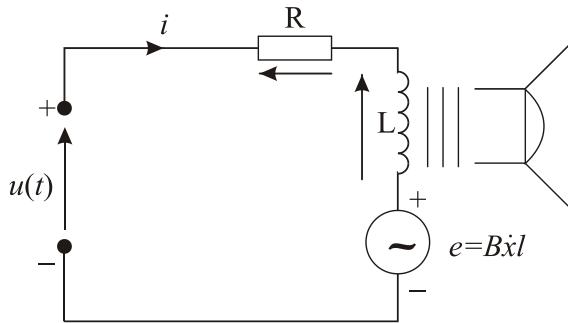
$$F = B \cdot i \cdot l,$$

gdzie l oznacza długość przewodnika; w rozważanym przypadku jest to długość uzwojenia równa $n \cdot \pi \cdot d$, n – liczba zwojów, d – średnica uzwojenia. Powyższa zależność wymaga wyznaczenia wartości prądu i ; musimy rozważyć obwód uzwojenia pobudzany zależnym od czasu napięciem $u(t)$.

W obwodzie tym wyróżniamy oporność R , może ona wyrażać sumę oporności uzwojenia

oraz ewentualnego oporu wewnętrznego źródła napięcia, indukcyjność uzwojenia L oraz indukowane dzięki ruchowi uzwojenia w polu magnetycznym napięcie e . Wartość e określa tzw. prawo generatora wyrażające, zgodnie z prawem Faradaya, wartość indukowanego napięcia jako

$$e = B\dot{x}l$$



Rys. 4.6. Obwód uzwojenia głośnika

Wykorzystując tę zależność możemy zapisać równanie bilansu napięć w obwodzie uzwojenia (drugie prawo Kirchoff'a) w postaci:

$$L \frac{di}{dt} + Ri + B\dot{x}l = u$$

Widzimy, że wszystkie wielkości zależne od innych zostały wyznaczone; łącząc powyższe zależności możemy pełny model wyrazić jako zespół dwóch równań:

$$M\ddot{x} + b\dot{x} + kx - B \cdot l \cdot i = 0$$

$$L \frac{di}{dt} + R \cdot i + Bl\dot{x} - u = 0$$

Jest to model trzeciego rzędu, o stałych skupionych; warunki początkowe odpowiadające stanowi obiektu w danej chwili początkowej t_0 to: $i(t_0)$, $x(t_0)$, $\dot{x}(t_0)$

Przyjmując następujące oznaczenia trzech współrzędnych stanu x_i , $i=1,2,3$,

$$\overset{\text{def}}{x_1} = x, \quad \overset{\text{def}}{x_2} = \dot{x}, \quad \overset{\text{def}}{x_3} = i$$

możemy zapisać równania modelu jako układ równań w postaci normalnej

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\frac{k}{M}x_1 - \frac{b}{M}x_2 - \frac{B \cdot l}{M}x_3 \\ \dot{x}_3 = -\frac{B \cdot l}{M}x_2 - \frac{R}{L}x_3 + \frac{I}{L}u \end{cases} \quad x(t_0) = \begin{bmatrix} x_1(t_0) \\ x_2(t_0) \\ x_3(t_0) \end{bmatrix} \quad (4.13)$$

Standardowo modele liniowe zapisujemy w postaci wektorowo-macierzowej, jak poniżej

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{k}{M} & -\frac{b}{M} & -\frac{B \cdot l}{M} \\ 0 & -\frac{B \cdot l}{L} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix} u \quad \Leftrightarrow \quad \dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{bu} \quad \square$$

4.2.2. Modele empiryczne (behawioralne)

Poświęćmy teraz chwilę uwagi budowie modelu empirycznego; możemy go także nazwać modelem behawioralnym. Są to modele deskryptywne. Rozważmy jako przykład proces sprzedaży dwóch produktów adresowanych do licznego grona klientów. Mogą to być, w szczególności, dwa rodzaje benzyny sprzedawane na stacji paliw, powiedzmy benzyna bezołowiowa 95 i benzyna typu super 98. Obserwujemy zależność ilości w litrach sprzedawanego paliwa od ceny, dajmy na to w skali tygodnia (lub dnia) przy poziomie cen p_{95} i p_{98} odnosimy do nich wielkości sprzedaży, odpowiednio, S_{95} oraz S_{98} . Oczywiście, na poziom sprzedaży ma wpływ wiele innych czynników, takich jak warunki pogodowe, różne organizowane w okolicy imprezy, itp. Przyjmujemy jednak, że w modelu sprzedaży uwzględniamy jedynie zależność od cen.

Jaka może być postać, kształt, takiej zależności? Nie wiemy, w jaki sposób informacja o cenie wpływa na decyzje zbioru klientów mogących korzystać z danej stacji paliw; musimy zatem założyć, przyjąć z góry, postać zależności. Na przykład możemy rozważyć najprostsze modele liniowe w postaci

$$\begin{aligned} S_{95} &= a_1 p_{95} + b_1 \quad \text{oraz} \\ S_{98} &= a_2 p_{98} + b_2 \end{aligned} \quad (4.14)$$

lub, w nieco bardziej rozbudowanej postaci uwzględniającej wzajemny wpływ cen na wielkość sprzedaży,

$$\begin{aligned} S_{95} &= a_{11} p_{95} + a_{12} p_{98} + b_1 \\ S_{98} &= a_{21} p_{95} + a_{22} p_{98} + b_2 \end{aligned} \quad (4.15)$$

Modele te będą, oczywiście, obowiązywać jedynie w określonym, niewielkim, zakresie zmian cen i wymagają dobrania właściwych wartości parametrów, których, jak widać, liczba szybko rośnie w przypadku rozbudowy strukturalnej równań. Na przykład wprowadzenie zależności sprzedaży benzyny Pb95 w okresie k od jej ceny w tym i kilku, powiedzmy r , poprzednich okresach, a więc uwzględnienie dynamiki procesu kształtowania się percepcji danej stacji wśród jej klientów, prowadzi natychmiast do znacznego wzrostu liczby parametrów, mianowicie mamy wówczas:

$$S_{95}(k) = a_{11,0} p_{95}(k) + a_{11,1} p_{95}(k-1) + \dots + a_{11,r} p_{95}(k-r)$$

Model liniowy może nie być wystarczający; w szczególności może nie ujmować asymetryczności wpływu na wielkość sprzedaży zmniejszenia oraz zwiększenia danej ceny. Pomocne w takiej sytuacji jest skorzystanie z innych prostych modeli, na przykład z modelu o stałej elastyczności cenowej. Elastyczność e_i wolumenu S_i sprzedaży danego produktu względem zmian jego ceny p_i określamy, w danym punkcie odpowiadającym tej cenie, jako:

$$e_i = \lim_{\Delta p_i \rightarrow 0} \frac{\frac{\Delta S_i}{S_i}}{\frac{p_i}{\Delta p_i}} = \lim_{\Delta p_i \rightarrow 0} \frac{p_i}{S_i} \cdot \frac{\Delta S_i}{\Delta p_i} = \frac{p_i}{S_i} \frac{dS_i}{dp_i}$$

Ważną rolę w modelowaniu ekonomicznym pełnią modele nieliniowe o stałej elastyczności; w takim przypadku w rozważanym przedziale zmienności cen, bądź innych instrumentów wejściowych, powyżej określona elastyczność przyjmowana jest jako wielkością stałą.

Prosty model o stałej elastyczności ma postać:

$$\begin{aligned} S_{95} &= \alpha_1 p_{95}^{b_1} \\ S_{98} &= \alpha_2 p_{98}^{b_2} \end{aligned} \quad (4.16)$$

gdzie b_1, b_2 – stałe elastyczności. Istotnie, łatwo pokazać, że w takim przypadku, na przykład

$$e_{95} = \frac{p_{95}}{S_{95}} \frac{dS_{95}}{dp_{95}} = \frac{p_{95}}{\alpha_1 p_{95}^{b_1}} \alpha_1 b_1 p_{95}^{b_1 - 1} = b_1$$

Model powyższy zawiera cztery parametry i wyraża wzajemnie niezależne zależności sprzedaży odpowiednich rodzajów paliw od ich cen.

Jeśli taki prosty model nie jest wystarczający, na przykład nie pozwala on wyrazić wyraźnej i ważnej z praktycznego punktu widzenia skrośnej zależności pomiędzy wielkością sprzedaży benzyny 95 a ceną benzyny 98, to możemy zaproponować bardziej rozbudowany model uwzględniający także oddziaływanie skrośne, w postaci:

$$\begin{aligned} S_{95} &= \alpha_1 p_{95}^{b_{11}} p_{98}^{b_{12}} \\ S_{98} &= \alpha_2 p_{95}^{b_{21}} p_{98}^{b_{22}} \end{aligned} \quad (4.17)$$

W modelu tym współczynniki b_{11} i b_{22} oznaczają elastyczności własne, zaś b_{12} i b_{21} są elastycznościami skrośnymi.

Jak łatwo zauważyc, podstawowy koszt związany z wprowadzaniem modeli tego typu przy wprowadzaniu kolejnych elementów komplikujących wiąże się ze wzrostem liczby parametrów, których wartości następnie trzeba ustalić. Na ogół liczba parametrów występująca w modelach empirycznych jest znacznie większa od liczby parametrów, które musimy wprowadzić tworząc modele teoretyczne (fizyczne) o podobnym stopniu skomplikowania i precyzji. Ponadto, jak to już zostało wspomniane, przynajmniej przybliżone wartości parametrów w modelach fizycznych są na ogół z góry znane. Stąd, z reguły, nie jest celowe konstruowanie modelu empirycznego w sytuacji, gdzie da się stworzyć i zastosować model teoretyczne.

4.2.3. Modele konceptualne (konceptualno-empiryczne)

Jak już zostało powiedziane, w pewnych sytuacjach może być celowe podejście do budowy modelu polegające na wyobrażeniu sobie prostszej rzeczywistości fizycznej jako pewnego analogu rozważanego procesu, a następnie wykorzystanie modelu tej uproszczonej rzeczywistości. W istocie, właściwie zawsze przy budowie modelu fizycznego stosujemy różne uproszczenia; na przykład wyobrażamy sobie opisując zjawiska zachodzące w reaktorze chemicznym, że w każdej chwili temperatura wsadu jest jednakowa w każdym punkcie, podczas gdy w rzeczywistości tak nie jest. W przypadku modelu konceptualnego idziemy jednak jakby dalej wprowadzając strukturalne różnice pomiędzy modelem i opisywaną przez niego rzeczywistością. Jako przykład można rozważyć podejście do budowy

modelu przepływu w danym długim odcinku rzeki polegające na wyobrażeniu sobie takiego odcinka jako kaskady połączonych ze sobą kilku bądź kilkunastu zbiorników, a następnie zaproponowaniu związków łączących poziomy retencji sąsiadujących ze sobą zbiorniczków z przepływem wody pomiędzy nimi. Inny przykład, to przystąpienie do opisu działania systemu wodociągów w dużej aglomeracji przy pomocy modelu uproszczonego, operującego poszczególnymi fragmentami tego systemu wyobrażanymi jako proste, punktowe, elementy gromadzące wodę i przekazujące ją do odbiorców lub innych podsystemów (elementów systemu).

Jako inny prosty przykład wprowadzenia modelu konceptualnego może posłużyć takie oto zadanie: chcemy opisać zmiany temperatury T w danym punkcie pewnego pomieszczenia w zależności od temperatury otoczenia oraz od stopnia otwarcia u zaworu doprowadzającego gorącą wodę do systemu grzejników w tym pomieszczeniu. Najprostszy model tych zmian możemy uzyskać przyjmując, że całość podlegających ogrzewaniu lub ochładzaniu elementów w rozważanym pomieszczeniu, wraz z powietrzem, uznamy za jedną masę o pojemności cieplnej c i temperaturze T oraz przyjmiemy stałą temperaturę otoczenia równą T_{ot} . Wówczas możemy natychmiast zaproponować model następujący:

$$c \frac{dT(t)}{dt} = ku(t) - b(T(t) - T_{ot})^\alpha$$

wyrażający zmiany gromadzonej energii cieplnej w zależności od natężenia dopływu tej energii z grzejników, opisanego jako $ku(t)$, oraz od różnicy temperatur pomiędzy pomieszczeniem i otoczeniem. Jeśli, dodatkowo, przyjmiemy, że wykładnik α jest równy jedności, to powyższe równanie odpowiada modelowi prostego układu inercyjnego. Jeśli, z kolei, uznamy za potrzebne uwzględnienie czynnika opóźnienia pomiędzy ustawieniem $u(t)$ zaworu w chwili t a zaistnieniem wpływu tego ustawienia na zmiany temperatury, to nasz model możemy przekształcić do postaci:

$$c \frac{dT(t)}{dt} = -b(T(t) - T_{ot}) + ku(t - T_0)$$

tj. do typowego modelu jednoinercyjnego z opóźnieniem. Wielkości c , b , k i T_0 są parametrami tego modelu.

Można, oczywiście, zadać pytanie, czy przedstawiony powyżej model, który określiliśmy mianem konceptualnego, nie jest w istocie modelem fizykalnym (teoretycznym), podobnie jak model ruchu samochodu lub model wahadła fizycznego. W rozważanym przypadku wydaje się, po prostu, że stopień przyjętych uproszczeń w wyobrażeniu fizyki procesów zachodzących w obiekcie jest na tyle poważny, że istnieje podstawa do uznania tego modelu bardziej jako konceptualnego niż fizykalnego, aczkolwiek granica pomiędzy modelami fizykalnymi i konceptualnymi jest z natury rzeczy rozmyta.

Wątku modelowania jako takiego nie będziemy dalej kontynuować. Kolejne punkty tego rozdziału poświęcimy postępowaniu prowadzącemu do identyfikacji parametrów modeli oraz dostosowaniu tych modeli do potrzeb projektowania układów regulacji, w szczególności poprzez linearyzację i skalowanie modeli.

4.2.4. Identyfikacja modelu

W najczęściej rozważanym przypadku identyfikacja modelu polega, po pierwsze, na dobraniu jak najlepszych, pod kątem danego kryterium, wartości przypisywanych parametrom tego modelu, a następnie na zweryfikowaniu, czy po takiej identyfikacji nasz model w dostatecznie dokładny dla stawianych celów sposób opisuje modelowaną rzeczywistość.

Istnieje wiele sposobów dostrajania wartości parametrów, o niektórych z nich będzie mowa w kolejnych rozdziałach. Podstawowe, uniwersalne w zastosowaniu, podejście polega na wprowadzeniu odpowiedniej miary odległości pomiędzy zaobserwowanym w rzeczywistości przebiegiem zmian danej wielkości procesowej, w szczególności wyjścia obiektu, a odpowiadającym mu przebiegiem zmian tejże wielkości wyznaczanej na podstawie modelu. Dobór wartości parametrów następuje w drodze minimalizacji tejże miary. Jako prosty przykład może posłużyć nam poszukiwanie wartości parametrów zaproponowanego wcześniej modelu

$$S_{95} = \alpha_1 p_{95}^{b_1}$$

wyrażającego zależność pomiędzy ceną, powiedzmy w danym dniu, benzyny 95, a wielkością dobową sprzedaży tego paliwa w rozpatrywanej stacji sprzedaży paliw. Przypuśćmy, że dysponujemy danymi dotyczącymi sprzedaży w porównywalnych dniach w przeszłości, na przykład w poprzednich takich samych dniach tygodnia. Te historyczne, rzeczywiste, wartości tworzą zestaw danych $S_{95,i}$, dla $i=1,\dots,I$; odpowiadające im ceny to $p_{95,i}$, $i=1,\dots,I$. Poszukujemy następnie takich wartości parametrów α_1 i b_1 dla których uzyskamy minimalną wartość kryterium identyfikacji określonego jako

$$J_I = \sum_{i=1}^I (S_{95,i} - \alpha_1 p_{95,i}^{b_1})^2$$

Oczywiście, nie można oczekwać, że minimalna otrzymana wartość wskaźnika J_I będzie równa zero; nasz model jest zawsze jedynie przybliżonym opisem rzeczywistości i nie uwzględnia wielu czynników mogących wpływać, w tym przypadku, na wielkość sprzedaży w kolejnych okresach. Jeśli wartość J_I , a właściwie wielkość uśredniona J_I/I będzie dostatecznie mała, to można zadać pytanie, czy wypada uznać, że rozważany model zasługuje na zaakceptowanie. W istocie, odpowiedź na to pytanie powinna zależeć od tego, czy model przejdzie weryfikację polegającą, na przykład, na obliczeniu, dla uzyskanych wartości α_1 i b_1 , wskaźnika J_L odpowiadającego innemu zestawowi danych, powiedzmy: $S_{95,j}$, $p_{95,j}$, $j=I+1,\dots,I+L$. Jeśli policzona dla tych danych wartość J_L/L będzie dostatecznie bliska J_I/I , to można będzie uznać wartości parametrów za dobrze dobrane. Warto w tym miejscu podkreślić znaczenie nawet tak prostej weryfikacji modelu przy pomocy niezależnego zestawu danych.

Przedstawiona w powyższym przykładzie postać wskaźnika „odległości” pomiędzy rzeczywistym przebiegiem, bądź sekwencją, wartości zmiennych procesowych, a ich modelowym opisem ma postać sumy kwadratów błędów; jest to najczęściej spotykana i stosowana postać kryterium jakości dopasowania modelu.

W podobny do omówionego sposobu można postępować w przypadku dowolnych modeli; pod warunkiem oczywiście, że będziemy dysponować dostateczną liczbą danych eksperymentalnych zarówno do celu strojenia parametrów modelu jak do jego weryfikacji.

Kończąc ten punkt warto zauważyc, że zadanie minimalizacji wskaźnika J_I względem parametrów jest zadaniem optymalizacji nieliniowej. Jeśli jednak dokonamy, poprzez obustronne logarytmowanie, prostego przekształcenia modelu wyrażającego zależność sprzedaży benzyny 95 od jej ceny do postaci

$$\ln S_{95} = \ln \alpha_1 + b_1 \ln p_{95},$$

to możemy zadanie doboru parametrów α_1 i b_1 wyrazić jako zadanie minimalizacji, względem wielkości $\beta_1 = \ln \alpha_1$ oraz b_1 , nieco innego wskaźnika, mianowicie

$$J_{I,\text{mod}} = \sum_{i=1}^I [\ln S_{95,i} - (\beta_1 + b_1 \ln p_{95,i})]^2$$

Jest to zadanie minimalizacji funkcji kwadratowej względem β_1 ($=\ln \alpha_1$) oraz b_1 , co znacznie ułatwia poszukiwanie minimum. Warto przy okazji podkreślić wyraźnie widoczną potrzebę występowania istotnej zmienności cen w ramach serii historycznych wartości wykorzystywanych do celów identyfikacji parametrów; w przeciwnym przypadku, w szczególności przy występowaniu stałej ceny $p_{95,i}$, minimalizacja wskaźnika $J_{I,\text{mod}}$ nie doprowadzi do ustalenia właściwej wartości każdego z obydwu poszukiwanych parametrów.

4.3. Punkty równowagi, charakterystyki statyczne

Charakterystykami statycznymi modelu nazywamy zależności algebraiczne opisujące *relacje między wejściami i wyjściami modelu w stanach ustalonych*, tzn. między ustalonimi wartościami wejść i wyjść modelu w stanie równowagi. Dla modelu w postaci normalnej

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), u(t), \mathbf{z}(t)) \\ y(t) &= g(\mathbf{x}(t))\end{aligned}$$

gdzie x jest stanem, u sterowaniem, z zakłóceniem, a y wyjściem modelu, charakterystyka statyczna $y = h(u, z)$ wynika z rozwiązania układu równań algebraicznych (w stanach ustalonych wszystkie pochodne są zerowe)

$$\begin{aligned}\mathbf{0} &= \mathbf{f}(\mathbf{x}, u, \mathbf{z}) \\ y &= g(\mathbf{x})\end{aligned} \Rightarrow y = h(u, \mathbf{z})$$

Przypomnijmy, że jeśli równanie stanu jest nieliniowe, to i charakterystyki statyczne będą w ogólności nieliniowe. Co więcej, dla danej ustalonej wartości wielkości wejściowych (u i z) może występować kilka różnych punktów równowagi (kilka różnych rozwiązań y podanego wyżej układu równań algebraicznych), tzn. charakterystyki statyczne mogą być niejednoznaczne. Natomiast jeśli równania stanu są liniowe

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{Ax}(t) + \mathbf{bu}(t) + \mathbf{Fz}(t) \\ y(t) &= \mathbf{c}^T \mathbf{x}(t)\end{aligned}$$

oraz macierz stanu A jest nieosobliwa, to charakterystyka statyczna jest jednoznaczna (tzn. każdej ustalonej wartości u i z odpowiada dokładnie jedna ustalona wartość wyjścia y) i dana jest wzorem

$$\begin{aligned}\mathbf{0} &= \mathbf{Ax} + \mathbf{bu} + \mathbf{Fz} \\ y &= \mathbf{c}^T \mathbf{x}\end{aligned} \Rightarrow y = -\mathbf{c}^T \mathbf{A}^{-1}(\mathbf{bu} + \mathbf{Fz})$$

Przykład 4.5.

Dla nieliniowego modelu

$$\begin{aligned}\dot{x}_1(t) &= 2x_1(t) + 2x_1(t)x_2(t) + u(t) \\ \dot{x}_2(t) &= -2x_2(t) + 2x_1(t)x_2(t), \quad y(t) = x_1(t) + 2x_2(t)\end{aligned}$$

wyznaczmy zbiór punktów równowagi (parametryzowany przez wartości u), a następnie charakterystyki statyczne.

Punkty równowagi wynikają z rozwiązania układu równań algebraicznych

$$\begin{aligned}0 &= 2x_1 + 2x_1x_2 + u \\0 &= -2x_2 + 2x_1x_2\end{aligned}$$

Odejmując powyższe równania stronami dostajemy

$$2x_1 + 2x_2 + u = 0 \quad \Rightarrow \quad x_2 = \frac{-u - 2x_1}{2} = -x_1 - \frac{u}{2}$$

Podstawiając ten wzór do drugiego równania dostajemy

$$0 = 2x_1 + u + x_1(-2x_1 - u) = (2x_1 + u)(1 - x_1)$$

Skąd wynika, że $x_1 = 1$ lub $x_1 = -u/2$. Stąd zbiór punktów równowagi:

$$x_1 = 1, x_2 = -(u/2) - 1 \quad \text{lub} \quad x_1 = -u/2, x_2 = 0$$

Podstawiając te zależności do równania wyjścia $y=x_1+2x_2$ dostajemy dwie charakterystyki statyczne (model nieliniowy !)

$$y = -u - 1 \quad \text{lub} \quad y = -u/2$$

□

Również w Przykładzie 4.1 podającym model dynamiki samochodu wyznaczyliśmy charakterystykę statyczną, traktując prędkość jako wielkość wyjściową, zob. wzór (4.4).

4.4. Linearyzacja i skalowanie modelu

Po uzyskaniu sprawdzonego, dostatecznie dobrego modelu rozważanego obiektu może wystąpić potrzeba dalszego uproszczenia tego modelu. W szczególności, bardzo ważną rolę odgrywa przekształcenie modelu do postaci liniowej, gdyż metody analizy obiektów liniowych oraz metody projektowania algorytmów sterowania dla takich obiektów są daleko bardziej zaawansowane i dopracowane w porównaniu do metod stosowanych do badania i projektowania układów nieliniowych. Jednocześnie, w wielu zastosowaniach istotnie wystarczy posługiwać się modelami liniowymi, w tym przy rozpatrywaniu działania obiektu w okolicach danego punktu pracy, co często ma miejsce w praktyce i jest powszechnie przyjmowane, w szczególności, przy projektowaniu algorytmów regulacji.

Do ilustracji postępowania przy uproszczeniu modelu polegającym na jego linearyzacji posłużymy się ponownie przykładem przedstawionego na rysunku 4.2 wahadła fizycznego z napędem w przegubie. Omówiony powyżej nieliniowy model tego wahadła wyraża się równaniem różniczkowym drugiego rzędu

$$\ddot{\Theta} + \frac{b}{J}\dot{\Theta} + \frac{mgl}{J}\sin\Theta = \frac{1}{J}\tau \quad (4.18)$$

Możemy rozważyć dwa znaczco różniące się rodzaje linearyzacji:

- 1) linearyzację lokalną,
- 2) linearyzację poprzez (pomocnicze) sprzężenie zwrotne.

4.4.1. Linearyzacja lokalna

Ten rodzaj linearyzacji jest powszechnie stosowany. Wykorzystuje się fakt, że w przypadku funkcji f , w przestrzeni n wymiarowej R^n , jeśli jest ona jest różniczkowalna i ma w otoczeniu danego punktu x_0 ciągłe pochodne względem swoich argumentów, można dokonać lokalnego przybliżenia wartości $f(x)$ w punkcie x bliskim x_0 w postaci zależności

$$f_{lin}(x) = f(x_0) + \frac{df(x)}{dx} \Big|_{x=x_0} \cdot (x - x_0),$$

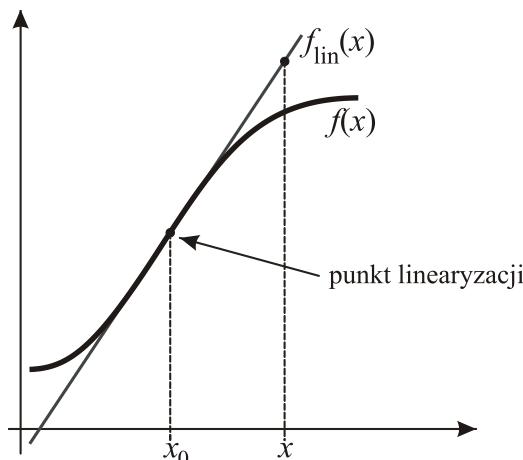
gdzie n wymiarowy wektor wierszowy $df(x)/dx$ zbudowany jest z elementów $df(x)/dx_i$ dla $i=1,\dots,n$. Dla przypadku jednowymiarowego ilustrację stanowi rysunek 4.5.

Oczywiście, zachodzi zależność

$$\lim_{x \rightarrow x_0} \frac{1}{|x - x_0|} |f(x) - f_{lin}(x)| \rightarrow 0^+.$$

Niech punkt równowagi $\Theta_0=0$, wówczas dokonując, dla niewielkich odchylen kąta Θ od zera, linearyzacji lokalnej funkcji sinus otrzymujemy liniowe przybliżenie nieliniowego członu w równaniu wahadła

$$\frac{mgl}{J} \sin \Theta \cong \frac{mgl}{J} \sin(0) + \frac{mgl}{J} \cos(0) \cdot (\Theta - 0) = \frac{mgl}{J} \Theta$$



Rys. 4.7. Linearyzacja funkcji różniczkowej wokół punktu x_0

Stąd zlinearyzowany model wahadła, dla małych wartości bezwzględnych kąta wychylenia Θ w otoczeniu zera, powiedzmy w granicach $\pm 0,1$ rad ($\pm 5^\circ$), ma postać:

$$\ddot{\Theta} + \frac{b}{J} \dot{\Theta} + \frac{mgl}{J} \Theta = \frac{\tau}{J}$$

Wyznaczmy z kolei wartość stałego momentu τ_0 pozwalającego utrzymać wychylenie w stanie ustalonym równe $\pi/4$ (45°), tj. takiego, że

$$\frac{mgl}{J} \sin \Theta_0 = \frac{\tau_0}{J}, \quad \Theta_0 = \frac{\pi}{4}$$

Mamy wówczas:

$$\tau_0 = mgl \frac{\sqrt{2}}{2}$$

oraz

$$\frac{mgl}{J} \sin \Theta \cong \frac{mgl}{J} \cdot \frac{\sqrt{2}}{2} + \frac{mgl}{J} \cdot \frac{\sqrt{2}}{2} \left(\Theta - \frac{\pi}{4} \right)$$

Wprowadzamy małe odchylenia θ i δ od wartości wychylenia i momentu w punkcie równowagi:

$$\Theta = \Theta_0 + \theta, \quad \tau = \tau_0 + \delta$$

Otrzymujemy model liniowy opisujący dobrze zmiany tych wielkości, przy założeniu, że mają one niewielkie wartości,

$$\ddot{\theta} + \frac{b}{J} \dot{\theta} + \frac{mgl}{J} \cdot \frac{\sqrt{2}}{2} \cdot \theta = \frac{1}{J} \delta$$

Warto zauważyć, że jest to inny model niż ten, który uzyskaliśmy dla wychyleń w otoczeniu punktu równowagi w zerze.

W ogólnym przypadku przypuśćmy, że dany jest układ opisany układem równań

$$\dot{x}_i = f_i(x_1, \dots, x_n, u), \quad i = 1, \dots, n$$

gdzie funkcje f_i mają ciągłe pochodne względem wszystkich zmiennych, i chcemy ten opis przybliżyć modelem liniowym w otoczeniu punktu równowagi x_0 odpowiadającemu ustalonej wartości $u=u_0$, tj. takiego punktu $\mathbf{x}_0=[x_{01}, \dots, x_{0n}]^T$, że

$$0 = f_i(x_{01}, \dots, x_{0n}, u_0), \quad i = 1, \dots, n \quad (4.19)$$

Wprowadźmy zmienne odpowiadające odchyleniom od punktu równowagi, mianowicie

$$\Delta x_i = x_i - x_{0i}, \quad i = 1, \dots, n, \quad \Delta u = u - u_0$$

Wówczas, rozwijając funkcje f_i w szereg Taylora w punkcie równowagi (\mathbf{x}_0, u_0) i uwzględniając definiujące punkt równowagi równości (4.19) dostajemy następujący model liniowy (zlinearyzowany)

$$\dot{\Delta x}_i = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0, u_0) \cdot \Delta x_j + \frac{\partial f_i}{\partial u}(\mathbf{x}_0, u_0) \cdot \Delta u, \quad i = 1, \dots, n \quad (4.20)$$

Model ten, wprowadzając macierz \mathbf{A} i wektor kolumnowy \mathbf{b} , złożone, odpowiednio, z elementów $\partial f_i / \partial x_j$ oraz $\partial f_i / \partial u$, zapisujemy w standardowej formie

$$\dot{\Delta \mathbf{x}}(t) = \mathbf{A} \cdot \Delta \mathbf{x}(t) + \mathbf{b} \cdot \Delta u(t)$$

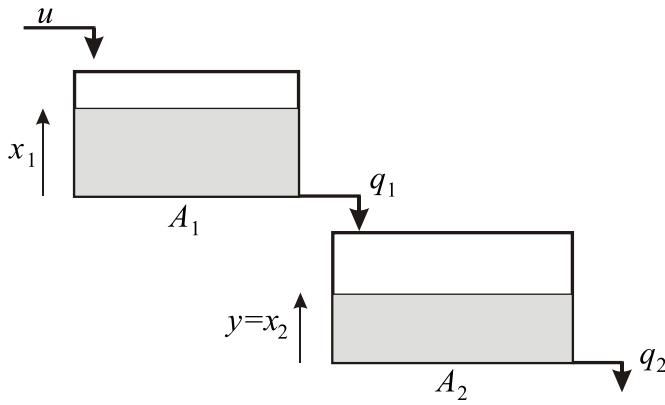
W większości zastosowań, rozpatrując powyższy zlinearyzowany model obiektu opuszczamy oznaczenia wielkości przyrostowych, z pomocą symbolu Δ , i model ten zapisujemy w powszechnie spotykanej postaci

$$\dot{\mathbf{x}}(t) = \mathbf{A} \cdot \mathbf{x}(t) + \mathbf{b} \cdot u(t)$$

Jeśli wielkości sterujących jest więcej niż jedna, to zamiast wektora \mathbf{b} w powyższym równaniu wystąpi macierz \mathbf{B} o liczbie kolumn równej liczbie wielkości sterujących.

Przykład 4.6. (kaskada zbiorników o swobodnych odpływach).

Rozważmy kaskadę dwóch zbiorników o swobodnym odpływach przedstawioną schematycznie na rysunku poniżej.



Rys. 4.8. Schemat obiektu w przykładzie 4.6.

Zbiorniki są o stałych przekrojach, odpowiednio A_1 i A_2 [m^2], do pierwszego dopływa woda z natężeniem u [m^3/s] i odpływa swobodnie z natężeniem odpływu q_1 [m^3/s] do drugiego zbiornika, z którego wypływa swobodnie z natężeniem odpływu q_2 [m^3/s], gdzie

$$q_1(t) = \alpha_1 \sqrt{x_1(t)}, \quad q_2(t) = \alpha_2 \sqrt{x_2(t)}$$

zaś α_1 i α_2 to współczynniki oporów odpływu. Wyznaczmy równania dynamiki, charakterystykę statyczną oraz model zlinearyzowany tego obiektu, traktując wysokość słupa wody w zbiorniku drugim jako wyjście $y(t)$.

Równania dynamiki odpowiadają bilansom objętości

$$\begin{aligned}\frac{d(A_1 x_1(t))}{dt} &= u(t) - \alpha_1 \sqrt{x_1(t)}, \\ \frac{d(A_2 x_2(t))}{dt} &= \alpha_1 \sqrt{x_1(t)} - \alpha_2 \sqrt{x_2(t)}\end{aligned}$$

Stąd równania w postaci normalnej

$$\begin{aligned}\frac{dx_1(t)}{dt} &= -\frac{\alpha_1}{A_1} \sqrt{x_1(t)} + \frac{1}{A_1} u(t), \\ \frac{dx_2(t)}{dt} &= \frac{\alpha_1}{A_2} \sqrt{x_1(t)} - \frac{\alpha_2}{A_2} \sqrt{x_2(t)}, \quad y(t) = x_2(t)\end{aligned}$$

Równanie charakterystyki statycznej (tzn. równanie punktów równowagi) dostajemy z równań algebraicznych (równań stanu przy zerowych pochodnych), podstawiając pierwsze do drugiego mamy

$$0 = \frac{u_0}{A_2} - \frac{\alpha_2}{A_2} \sqrt{x_{02}}, \quad \Rightarrow \quad y_0 = x_{02} = \frac{u_0^2}{\alpha_2^2}$$

Mimo nieliniowości, punkty równowagi (a więc i charakterystyka statyczna) są jednoznaczne.

Wykorzystując zależność (4.20) dostajemy model zlinearyzowany

$$\begin{aligned}\frac{d\Delta x_1(t)}{dt} &= -\frac{\alpha_1}{A_1 2\sqrt{x_{01}}} \Delta x_1(t) + \frac{1}{A_1} \Delta u(t), \\ \frac{d\Delta x_2(t)}{dt} &= \frac{\alpha_1}{A_2 2\sqrt{x_{01}}} \Delta x_1(t) - \frac{\alpha_2}{A_2 2\sqrt{x_{02}}} \Delta x_2(t), \quad \Delta y(t) = \Delta x_2(t)\end{aligned}$$

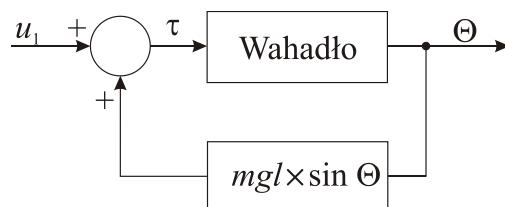
gdzie $\Delta x_1(t) = x_1(t)-x_{01}$, $\Delta x_2(t) = x_2(t)-x_{02}$, $\Delta y(t) = y(t)-y_0$, $\Delta u(t) = u(t)-u_0$. \square

4.4.2. Linearyzacja poprzez (pomocnicze) sprzężenie zwrotne.

Inny jakościowo rodzaj linearyzacji stanowi linearyzacja modelu układu nieliniowego, która może być uzyskana, dla pełnego zakresu zmienności występujących w modelu wielkości, przy pomocy realizowanego sprzężenia zwrotnego. W ogólnym przypadku obiektu nieliniowego opisanego układem równań różniczkowych nie jest łatwo ustalić, czy takie linearyzujące w globalnym sensie sprzężenie zwrotne jest możliwe do wprowadzenia, a także w jaki sposób to uczynić. Teoria sprzężeń linearyzujących operuje skomplikowanym aparatem matematycznym i wykracza poza ramy tego wykładu. Tym niemniej, w licznych interesujących przypadkach szczególnych można takie sprzężenie stosunkowo łatwo określić. Powróćmy, dla przykładu, do nieliniowego modelu wahadła

$$\ddot{\Theta} + \frac{b}{J} \dot{\Theta} + \frac{mgl}{J} \sin \Theta = \frac{\tau}{J}$$

Łatwo zauważyć, że po wprowadzeniu pokazanego na rysunku sprzężenia zwrotnego, dzięki któremu moment napędowy τ uzyskiwany jest w wyniku sumowania momentu zewnętrznego u_1 oraz komponentu pochodzącego ze sprzężenia o wartości $mgl\sin\Theta$, otrzymujemy zmodyfikowany obiekt, dla którego nowym sterowaniem jest u_1 , w postaci



Rys. 4.9. Linearyzujące sprzężenie zwrotne wahadła.

$$\ddot{\Theta} + \frac{b}{J} \dot{\Theta} + \frac{mgl}{J} \sin \Theta = \frac{1}{J} u_1 + \frac{mgl}{J} \sin \Theta.$$

Jest to model zlinearyzowany poprzez sprzężenie zwrotne, po oczywistym uproszczeniu przyjmuje on postać

$$\ddot{\Theta} + \frac{b}{J} \dot{\Theta} = \frac{1}{J} u_1$$

Linearyzujące sprzężenie samo w sobie jest układem nieliniowym. Układ sterowania realizujące takie sprzężenie musi być bardzo szybki jeśli trzeba zastosować układy z próbkowaniem; musi się w istocie zachowywać jak układ ciągły. Niezbędne obliczenia także powinny być prowadzone bardzo szybko, gdyż opóźnienia wprowadzą zniekształcenia. Rozwiążanie jest nowoczesne, a jego zastosowanie umożliwia dzisiejsza technologia. Zaletą tego układu jest to, że linearyzacja obowiązuje dla całego zakresu zmian kąta, w odróżnieniu od linearyzacji lokalnej nie występują przybliżenia. Warto zauważyć jednak, że w przypadku niedokładnej znajomości parametrów obiektu rzeczywista nieliniowość nie zostanie w całości wyeliminowana. Ponadto, wprowadzenie sprzężenia linearyzującego może, ułatwiając projektowanie układu sterowania, podnieść koszt działania takiego układu, a niekiedy także ograniczyć możliwe do uzyskania osiągi obiektu. Dlatego też stosowanie mechanizmu linearyzacji poprzez sprzężenie zwrotne nie jest powszechnie stosowanym rozwiązaniem,

warto jednak brać pod uwagę jego ewentualne wykorzystanie.

4.4.3. Skalowanie modelu.

Zarówno czytelność modelu jak i łatwość jego wykorzystania do obliczeń i projektowania mogą być znaczco poprawione poprzez odpowiednie skalowanie. Polega ono na doborze odpowiednich jednostek, w których wyrażane są występujące w modelu wielkości. W szczególności możemy wyróżnić:

- skalowanie amplitudy polegające na dokonaniu zmiany jednostek wartości zmiennych i parametrów,
- skalowanie czasu.

Skalowanie amplitudy jest kwestią dość oczywistą, aczkolwiek warto o niej pamiętać, gdyż dzięki użyciu właściwych jednostek unikniemy operowania zbyt dużymi lub zbyt małymi liczbami. Na przykład opisując ruch samochodu nie powinniśmy używać jednostek miary odległości takich jak milimetry, które z kolei są właściwe przy badaniu mikromechanizmów.

Skalowanie, tj. dobór właściwych jednostek czasu wymaga nieco większej uwagi, ze względu na to, że przy zmianie jednostek czasu, zmiennej niezależnej w równaniach modelu, łatwiej jest popełnić błąd. Rozważmy dla przykładu prosty model liniowy

$$\frac{d^2\Theta}{dt^2} + 15^2 \cdot 10^6 \Theta = 10^6 \tau$$

w którym poszczególne wielkości wyrażone są w następujących, standardowych jednostkach: czas - t - w sekundach, moment siły - τ - w Nm (newtonometrach), kąt - Θ - w radianach.

Współczynniki w równaniu mają duże wartości. Jeśli natomiast wyrazimy czas w milisekundach jako $t' = 1000t$, gdzie $t'[\text{ms}]$, $t[\text{s}]$, to druga pochodna kąta względem czasu wyrażonego w sekundach pozostaje w następującej relacji do tejże pochodnej wyliczonej względem czasu wyrażonego w nowych jednostkach (milisekundach):

$$\frac{d^2\Theta}{dt'^2} = \frac{d}{dt'} \left(\frac{d\Theta}{dt'} \right) = 1000 \frac{d}{dt'} \left(1000 \frac{d\Theta}{dt} \right) = 10^6 \frac{d^2\Theta}{dt'^2}$$

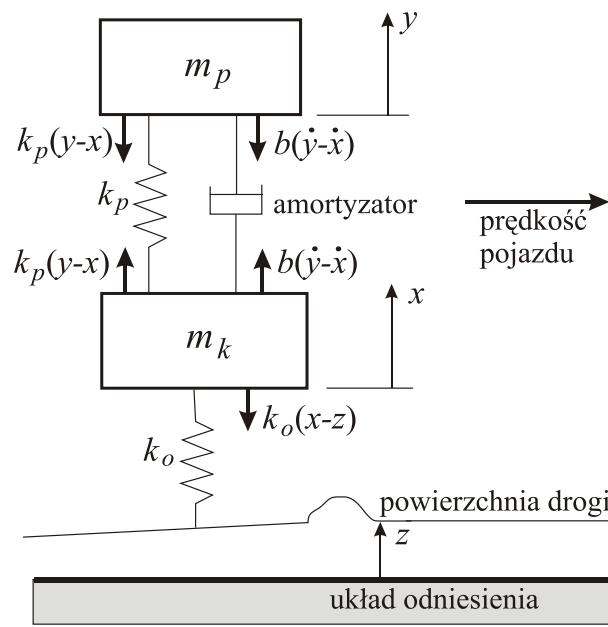
stąd, po obustronnym podzieleniu przez 10^6 , otrzymujemy równanie

$$\frac{d^2\Theta}{dt'^2} + 15^2 \cdot \Theta = \tau$$

w którym nie występują już współczynniki wyrażone bardzo dużymi liczbami.

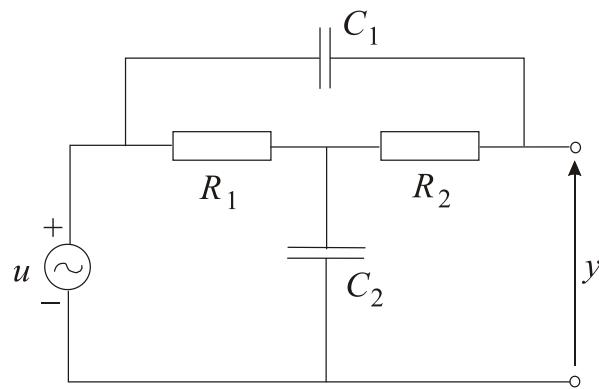
Zadania

Zadanie 4.1. Ułożyć model dynamiki pojedynczego punktu zawieszenia samochodu („quarter-car model”), przybliżony schematem w postaci dwóch mas: głównej pojazdu (m_p) i części jezdnej (koło z zamocowaniem (m_o), połączonych sprężyną liniową (o współczynniku sprężystości k_p) i amortyzatorem działającym proporcjonalnie (ze współczynnikiem proporcjonalności b) do prędkości ściskania. Ponadto, oddziaływanie opony z podłożem modelowane jest sprężyną liniową o współczynniku sprężystości k_o .



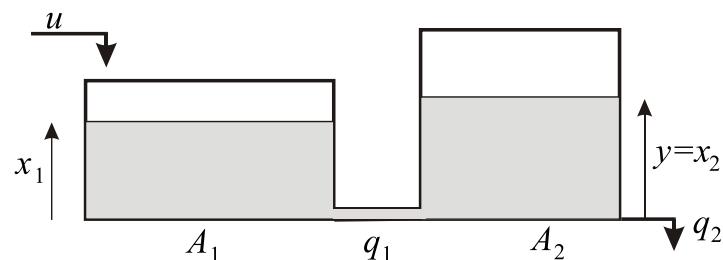
Wskazówki: Na rysunku zaznaczono położenie układów współrzędnych i siły działające na obie masy. Otrzymamy liniowy model dynamiki 4 rzędu.

Zadanie 4.2. Ułożyć model dynamiki w postaci równań stanu dla obwodu elektrycznego jak na rysunku, jako sterowanie u przyjmując napięcie źródła i jako wielkość wyjściową napięcie y jak zaznaczono na rysunku.



Wskazówka: Należy przyjąć jako zmienne stanu wielkości opisujące stan kondensatorów.

Zadanie 4.3. Ułożyć model dynamiki obiektu dwóch zbiorników jak na rysunku:



Zbiorniki są o stałych przekrojach, odpowiednio A_1 i A_2 [m^2], do pierwszego dopływa woda z natężeniem u [m^3/s], przepływa swobodnie między zbiornikami z natężeniem q_1 [m^3/s] i wypływa swobodnie z drugiego zbiornika z natężeniem odpływu $q_2(t) = \alpha_2 \sqrt{x_2(t)}$ [m^3/s], gdzie α_2 to współczynnik oporu odpływu, natomiast α_1 to współczynnik oporów przepływu między zbiornikami. Należy wyznaczyć równania dynamiki, charakterystykę statyczną oraz model zlinearyzowany tego obiektu, traktując wysokość słupa wody w zbiorniku drugim jako wyjście $y(t)$.

5. Analiza liniowych modeli dynamicznych

Po uzyskaniu wiarygodnego modelu rozważanego obiektu lub procesu możemy przystąpić do kolejnego etapu poprzedzającego zaprojektowanie układu sterowania. Etap ten określamy mianem analizy. Zarówno sposoby jak i zadania analizy są różnorodne. W niniejszym rozdziale uwaga zostanie skoncentrowana na analizie modeli liniowych, będącym jednak poruszać w wybranych miejscach kwestie analizy modeli nieliniowych w kontekście badania stabilności układów dynamicznych. Analiza stabilności odgrywa szczególną rolę, gdyż jednym z zawsze występujących wymagań przy projektowaniu układu regulacji automatycznej jest wymaganie zapewnienia asymptotycznej stabilności takiego układu; dodatkowo wymagane jest istnienie odpowiednich tzw. zapasów stabilności.

Analiza w dziedzinie czasu obejmuje takie zagadnienia jak: rozwiązywanie układu równań liniowych, określenie tzw. odpowiedzi impulsowej, oraz wspomniane powyżej badanie stabilności.

Analiza w dziedzinie zm. zespółonej, zwana także analizą częstotliwościową, bazuje na wykorzystaniu transformaty Laplace'a i wykorzystuje transmitancje operatorowe. Jest ona szczególnie odpowiednia do badania własności związanych z przenoszeniem sygnałów przez liniowy układ dynamiczny oraz do ujmowania ilościowych cech związanych z badaniem stabilności układu zamkniętego, tj. objętego pętlą sprzężenia zwrotnego.

Ogólnie rzecz ujmując zadanie analizy obejmuje badanie własności obiektu, w tym cech odpowiedzi układu na różne pobudzenia (na przykład odpowiedź na skok wartości wejścia), wrażliwości na wejścia niesterowane, wrażliwości na parametry modelu. Dalsze, wspomniane już cele analizy, to badanie stabilności układu, w szczególności cech stabilności układu zamkniętego (z regulatorem) w zależności od właściwości układu otwartego (przy otwartej pętli sprzężenia zwrotnego).

Jeszcze raz warto podkreślić, że analiza bazuje na wykorzystaniu zidentyfikowanego i zweryfikowanego modelu matematycznego; model stanowi zatem podstawę dla wszelkich dalszych rozważań. Stąd też analizie mogą także towarzyszyć elementy dalszej identyfikacji parametrów modelu.

Sposoby analizy w dziedzinie czasu obejmują, w pierwszym rzędzie, podejście bezpośrednie - badanie układu dynamicznego poprzez rozwiązywanie równań modeli; można w tym celu stosować metody analityczne, a także metody numeryczne. Możliwe jest także, w ograniczonym zakresie, badanie własności układu w oparciu o model w dziedzinie czasu, ale bez potrzeby rozwiązywania równań modelu, w szczególności dotyczy to badania stabilności.

5.1. Analiza w dziedzinie czasu

5.1.1. Rozwiązywanie równań modelu - przykład

Rozważmy, jako ilustrację takiego postępowania obiekt typu inercja z opóźnieniem: równanie opisujące działanie obiektu, w którym wielkość $y(t)$ oznacza jednocześnie stan i wyjście obiektu w chwili t , ma postać:

$$T\ddot{y}(t) + \dot{y}(t) = k_0 u(t - T_0) \quad (5.1)$$

gdzie $T > 0$, $T_0 > 0$, $k_0 > 0$. Przyjmujemy, że $t \geq 0$, $u(\tau) = 0$ dla $-T_0 \leq \tau \leq 0$, ponadto dany jest stan początkowy (warunek początkowy) $y_0 = y(0)$. Model określa zatem równanie różniczkowe pierwszego rzędu z opóźnieniem równym T_0 w sterowaniu.

Rozważmy rozwiązanie analityczne tego równania, przebiega ono w dwóch etapach:

1) Rozwiążanie równania jednorodnego

$$T\dot{y} + y = 0, \quad t \geq 0, \quad y(0) = y_0$$

Wykorzystujemy metodę rozdzielania zmiennych (z obustronnym całkowaniem):

$$T \frac{dy}{dt} = -y \Rightarrow T \frac{dy}{y} = -dt \Rightarrow T \ln y = -t + C_0.$$

Przyjmijmy teraz przedstawienie stałej całkowania C_0 jako logarytmu naturalnego z innej stałej o wartości dodatniej, tj. $C_0 = \ln C$, gdzie $C > 0$. Mamy wówczas, kontynuując powyższe przekształcenia,

$$\ln \frac{y}{C} = -\frac{t}{T} \Rightarrow y = Ce^{-\frac{t}{T}}, \quad t \geq 0. \quad (5.2)$$

2) Rozwiążanie równania niejednorodnego:

$$T\dot{y}(t) + y(t) = k_0 \cdot v(t),$$

gdzie $v(t) = u(t - T_0)$, $t \geq T_0$, oraz $v(t) = 0$ dla $0 \leq t \leq T_0$.

Stosujemy metodę uzmienniania stałej w rozwiązyaniu równania jednorodnego, tj. traktujemy C jako funkcję czasu, wówczas:

$$y(t) = C(t)e^{-\frac{t}{T}} \Rightarrow \dot{y}(t) = \dot{C}(t)e^{-\frac{t}{T}} + C(t)(-\frac{1}{T}e^{-\frac{t}{T}}).$$

Wstawiamy to wyrażenie w miejsce pochodnej stanu do wyjściowego niejednorodnego równania różniczkowego; otrzymujemy:

$$T \left[\dot{C}(t)e^{-\frac{t}{T}} - \frac{1}{T}C(t)e^{-\frac{t}{T}} \right] + C(t)e^{-\frac{t}{T}} = k_0 v(t),$$

stąd

$$\dot{C}(t) = \frac{1}{T}k_0 e^{\frac{t}{T}} \cdot v(t) \Rightarrow C(t) = C_1 + \int_0^t \frac{1}{T}k_0 e^{\frac{\tau}{T}} \cdot v(\tau) d\tau \Rightarrow y(t) = C_1 e^{-\frac{t}{T}} + \int_0^t \frac{1}{T}k_0 e^{-\frac{t-\tau}{T}} \cdot v(\tau) d\tau.$$

Następnie wyznaczamy C_1 :

$$y(0) = y_0 = C_1 e^{\frac{0}{T}} = C_1 \Rightarrow C_1 = y_0;$$

tak więc, ostatecznie, poszukiwane rozwiązanie naszego równania ma postać:

$$y(t) = y_0 e^{-\frac{t}{T}} + \int_0^t \frac{k_0}{T} e^{-\frac{t-\tau}{T}} u(\tau - T_0) d\tau, \quad t \geq 0, \quad T_0 \geq 0 \quad (5.3)$$

które dla $T_0 = 0$ redukuje się do postaci

$$y(t) = y_0 e^{-\frac{t}{T}} + \int_0^t \frac{k_0}{T} e^{-\frac{t-\tau}{T}} u(\tau) d\tau, \quad t \geq 0 \quad (5.4)$$

Pierwszy człon rozwiązań (5.3), czy (5.4), stanowi *składową swobodną*, pochodzącą od warunku początkowego, o wartości niezależnej od sterowania u . Natomiast drugi człon, zależny od przebiegu tego sterowania i niezależny od y_0 , nazywamy *składową wymuszoną*. Znając rozwiązanie, możemy przystąpić do badania własności układu dynamicznego (5.1).

Odpowiedź swobodna (układu autonomicznego)

Rozważamy zachowanie układu autonomicznego, tzn. dla $u(t) \equiv 0$. Wówczas z (5.3), czy z (5.4), wynika natychmiast

$$y(t) = y_0 e^{-\frac{t}{T}} \xrightarrow[t \rightarrow +\infty]{} 0^+$$

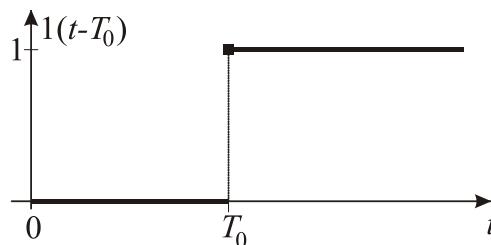
gdyż $T > 0$, a punkt zerowy jest stanem równowagi rozważanego układu dynamicznego z zerowym sterowaniem. Jednocześnie widzimy, że dla każdego $t \geq 0$ jest $|y(t)| \leq |y_0|$. Mamy więc do czynienia z przebiegiem jednostajnie ograniczonym i zbiegającym do punktu równowagi – takie przebiegi nazywamy asymptotyczne stabilnymi.

Odpowiedź skokowa

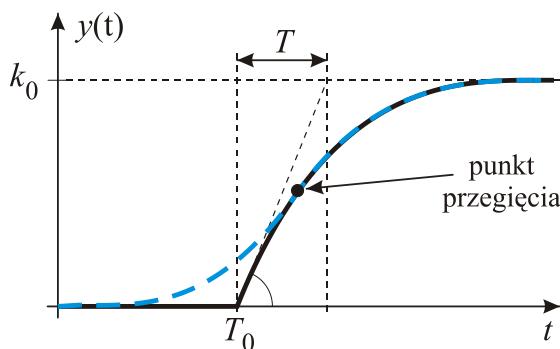
Rozważmy zachowanie naszego układu (5.1) dla $y_0 = 0$ i $u(t) = \mathbf{1}(t)$, tzn. $u(t) = 0$ dla $t < 0$ i $u(t) = 1$ dla $t \geq 0$; jest to sterowanie $u(t)$ w postaci tzw. *skoku jednostkowego* w chwili początkowej (zerowej). Na Rys. 5.2 przedstawiony jest skok jednostkowy opóźniony o czas T_0 , a przebieg wyjścia układu wynosi

$$y(t) = 0, \quad 0 \leq t < T_0,$$

$$y(t) = \int_{T_0}^t \frac{k_0}{T} e^{-\frac{t-\tau}{T}} d\tau = \frac{k_0}{T} e^{-\frac{t}{T}} \int_{T_0}^t e^{\frac{\tau}{T}} d\tau = \frac{k_0}{T} e^{-\frac{t}{T}} \left[T e^{\frac{\tau}{T}} \right]_{T_0}^t = k_0 \left[1 - e^{-\frac{t-T_0}{T}} \right], \quad t \geq T_0$$



Rys. 5.1. Skok jednostkowy opóźniony o czas T_0 .



Rys. 5.2. Odpowiedź układu inercyjnego z opóźnieniem na skok jednostkowy; przerywaną linią zaznaczono przykładowy przebieg przybliżany takim układem.

i jest przedstawiony na Rys. 5.2. Rysunek ten dobrze wyjaśnia, dlaczego rozważany układ dynamiczny nazywamy *inercyjnym z opóźnieniem*, a parametr T nazywamy *stałą czasową inercji*. Obliczając pochodną prawostronną rozwiązania w punkcie $t = T_0$ dostajemy

$$\left. \frac{dy(t)}{dt} \right|_{T_0^+} = k_0 \frac{1}{T}$$

Stąd wynika sposób wyznaczania stałej czasowej inercji T na podstawie wykresu odpowiedzi układu, pokazany na Rys. 5.2. Alternatywnym sposobem, na ogół dokładniejszym w sytuacji analizowania przebiegu wyjścia zakłóconego, jest skorzystanie z zależności:

$$y(T + T_0) = k_0(1 - e^{-1}) \approx 0.632k_0$$

Odpowiedź na wymuszenie ograniczone

Niech y_0 przyjmuje dowolną wartość, zaś przebieg sterujący spełnia warunek:

$|u(t)| \leq M$ dla $t \geq T_0$; czyli mamy ograniczony, jednostajnie, sygnał wejściowy. Wówczas

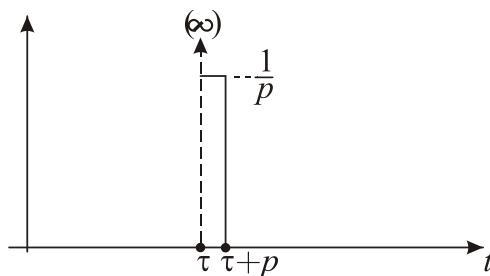
$$\begin{aligned} \forall_{t \geq 0} |y(t)| &\leq |y_0| + \left| \int_0^t \frac{k_0}{T} e^{-\frac{t-\tau}{T}} u(\tau - T_0) d\tau \right| \leq |y_0| + M \left| \int_{T_0}^t \frac{k_0}{T} e^{-\frac{t-\tau}{T}} d\tau \right| \leq \\ &\leq |y_0| + Mk_0 \left[1 - e^{-\frac{t-T_0}{T}} \right]_{t \geq T_0} \leq |y_0| + Mk_0 \end{aligned}$$

zatem także wielkość wyjścia jest jednostajnie ograniczona w czasie. W takim przypadku mówimy, że obiekt jest stabilny w sensie ograniczone wejście – ograniczone wyjście (stabilność OWOW – ang. BIBO, *bounded input – bounded output*).

Odpowiedź impulsowa

Rozważmy sygnał w postaci następującego impulsu, przedstawionego na Rys. 5.3:

$$h_p(t - \tau) = \begin{cases} 1/p, & \tau \leq t \leq \tau + p \\ 0, & t < \tau, t > \tau + p \end{cases}$$



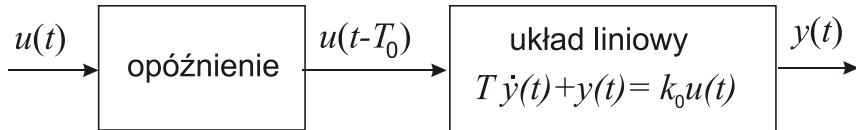
Rys. 5.3. Kształt impulsu przybliżającego (dla małej wartości p) impuls Diraca w chwili τ .

Wówczas, jak łatwo zauważyc

$$\int_{-\infty}^{\infty} h_p(t-\tau)dt = 1, \quad p \rightarrow 0^+ \Rightarrow h_p(t-\tau) \rightarrow \delta(t-\tau)$$

gdzie $\delta(t-\tau)$ to impuls Diraca w chwili $t = \tau$.

Rozważany układ dynamiczny jednoinercyjny z opóźnieniem można traktować jako złożenie dwóch układów: opóźnienia i jednoinercyjnego układu liniowego, jak to przedstawia Rys. 5.4



Rys. 5.4. Przedstawienie układu jednoinercyjnego z opóźnieniem jako szeregowego połączenia czystego opóźnienia i układu liniowego jednoinercyjnego.

Odpowiedzią impulsową nazywamy odpowiedź układu dynamicznego na impuls Diraca w chwili początkowej (zerowej), $\delta(t-0) = \delta(t)$. Wyznaczmy odpowiedź impulsową układu liniowego jednoinercyjnego, poprzez rozwiązywanie jego równania

$$T\dot{y}(t) + y(t) = k_0 \cdot u(t) \quad (5.5)$$

dla $u(t) = h_p(t)$ i przejście do granicy przy p dążącym do zera. Wykorzystując (5.4) mamy

$$\begin{aligned} y(t) &= \int_0^t \frac{k_0}{T} e^{-\frac{t-\tau}{T}} h_p(\tau) d\tau = \frac{k_0}{T} e^{-\frac{t}{T}} \int_0^t \frac{1}{p} e^{\frac{\tau}{T}} d\tau = \frac{k_0}{Tp} e^{-\frac{t}{T}} \left[T e^{\frac{\tau}{T}} \right]_0^t = \\ &= \frac{k_0}{p} e^{-\frac{t}{T}} (e^{\frac{t}{T}} - 1) = k_0 e^{-\frac{t}{T}} \frac{e^{\frac{t}{T}} - 1}{p} \xrightarrow[p \rightarrow 0]{} \frac{k_0}{T} e^{-\frac{t}{T}} \end{aligned}$$

$$\text{ponieważ } \lim_{p \rightarrow 0} \frac{e^{\frac{t}{T}} - 1}{p} = \lim_{p \rightarrow 0} \frac{\frac{1}{T} e^{\frac{t}{T}}}{1} = \frac{1}{T}.$$

Stąd odpowiedź impulsowa $g(t)$ układu (5.5) ma postać

$$g(t) = \frac{k_0}{T} e^{-\frac{t}{T}} \quad (5.6)$$

Zauważmy teraz, że ogólne rozwiązywanie równania (5.5) dane wzorem (5.4) można zapisać

$$y(t) = y_0 e^{-\frac{t}{T}} + \int_0^t g(t-\tau) u(\tau) d\tau, \quad t \geq 0 \quad (5.7)$$

tzn. całka określająca składową wymuszoną jest całką splotu *odpowiedzi impulsowej i funkcji wymuszenia* (sterowania). Równanie (5.3) można zapisać analogicznie, jedynie funkcja wymuszająca jest wówczas opóźnionym sterowaniem.

5.1.2. Model liniowy w postaci ogólnej

Rozważmy teraz przypadek ogólny układu liniowego, stacjonarnego, o stałych skupionych, o jednym wejściu i jednym wyjściu (SISO – *single input - single output*), zapisanego w

postaci normalnej

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) &= \mathbf{c}^T \mathbf{x}(t)\end{aligned}\tag{5.8}$$

gdzie $\mathbf{x} \in R^n$, $y \in R$, $u \in R$, $u(\tau) = 0$ dla $\tau < 0$, oraz dany jest stan początkowy $\mathbf{x}_0 = \mathbf{x}(0)$.

Zdefiniujmy macierz

$$e^{\mathbf{A}t} \stackrel{def}{=} \mathbf{I} + \mathbf{A}t + \frac{1}{2!} \mathbf{A}^2 t^2 + \frac{1}{3!} \mathbf{A}^3 t^3 + \dots = \sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{A}^i t^i$$

i zauważmy, że z powyższej definicji bezpośrednio wynika

$$\frac{d}{dt} e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t}$$

Biorąc powyższe pod uwagę, rozwiązanie równania jednorodnego $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ jest postaci

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0\tag{5.9}$$

albowiem $\dot{\mathbf{x}}(t) = \mathbf{A}e^{\mathbf{A}t} \mathbf{x}_0 = \mathbf{A}\mathbf{x}(t)$.

Natomiast pokażemy teraz, że rozwiązanie równania niejednorodnego $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$ jest postaci

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0 + e^{\mathbf{A}t} \int_0^t e^{-\mathbf{A}\tau} \mathbf{b}u(\tau) d\tau\tag{5.10}$$

Zróżniczkowanie powyższej zależności bezpośrednio udowadnia postawiona tezę:

$$\dot{\mathbf{x}}(t) = \mathbf{A}e^{\mathbf{A}t} \mathbf{x}_0 + \mathbf{A}e^{\mathbf{A}t} \int_0^t e^{-\mathbf{A}\tau} \mathbf{b}u(\tau) d\tau + e^{\mathbf{A}t} e^{-\mathbf{A}t} \mathbf{b}u(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$$

Podstawiając (5.10) do równania wyjścia dostajemy

$$y(t) = \mathbf{c}^T e^{\mathbf{A}t} \mathbf{x}_0 + \int_0^t \mathbf{c}^T e^{\mathbf{A}(t-\tau)} \mathbf{b}u(\tau) d\tau\tag{5.11}$$

Rozumując analogicznie, jak to uczyniliśmy dla układu jednoinercyjnego, można pokazać, że funkcja

$$g(t) = \mathbf{c}^T e^{\mathbf{A}t} \mathbf{b}\tag{5.12}$$

jest odpowiedzią impulsową układu dynamicznego danego przez (5.8), czyli rozwiązanie (5.11) można zapisać w równoważnej postaci

$$y(t) = \mathbf{c}^T e^{\mathbf{A}t} \mathbf{x}_0 + \int_0^t g(t-\tau) u(\tau) d\tau\tag{5.13}$$

Tak więc, dla każdego stacjonarnego układu liniowego, o stałych skupionych, odpowiedź wymuszona (tj. przebieg wyjścia przy zerowych warunkach początkowych) ma postać całki splotu funkcji odpowiedzi impulsowej $g(t)$ i wymuszenia $u(t)$.

Zwróćmy uwagę, że odpowiedź impulsowa i przebieg wyjścia układu jednoinercyjnego (5.5) są szczególnymi przypadkami zależności (5.12) i (5.13), dla $\mathbf{A} = -1/T$, $\mathbf{b} = k_0/T$, $\mathbf{c}=1$.

Zauważmy, że z liniowości operacji całkowania oraz postaci rozwiązania (5.11), czy (5.13), wynika bezpośrednio, że w układach liniowych obowiązuje *zasada superpozycji*

$$y(u_1(t) + u_2(t)) = y(u_1(t)) + y(u_2(t))$$

Dzięki dostępności rozwiązania analitycznego układu równań liniowych dowolnego rzędu badanie własności takiego układu w oparciu o to rozwiązanie jest możliwe; nie jest jednak zawsze wygodne, w szczególności wówczas, gdy chcemy badać zależność pomiędzy własnościami układu liniowego a własnościami tego układu po objęciu go pętlą sprzężenia zwrotnego.

5.1.3. Stabilność układu dynamicznego

Stabilność w sensie Lapunowa

Przypuśćmy, że dany jest układ dynamiczny stacjonarny, autonomiczny, tj. bez zewnętrznego pobudzenia:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) \quad (5.14)$$

gdzie $\mathbf{x} \in R^n$, $\mathbf{f}: R^n \rightarrow R^n$, $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n_x} (x_i)^2}$, $\bar{\mathbf{x}}$ – punkt równowagi, tj. $\mathbf{f}(\bar{\mathbf{x}}) = 0$.

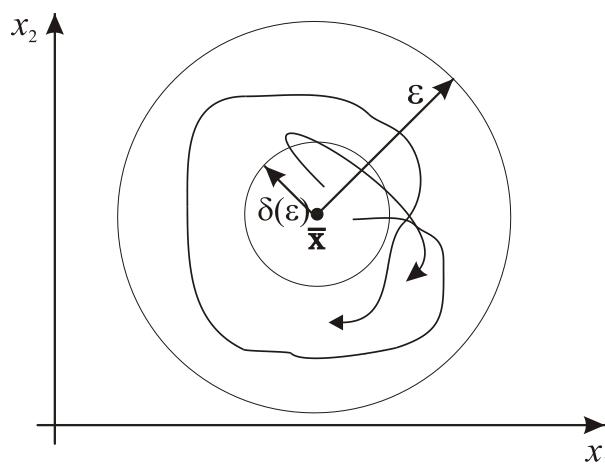
Definicja. Stabilność w sensie Lapunowa.

Punkt równowagi $\bar{\mathbf{x}}$ układu (5.14) jest stabilny w sensie Lapunowa, jeśli

$$\forall \varepsilon > 0 \exists \delta(\varepsilon) > 0 (\delta(\varepsilon) \leq \varepsilon) : \|\mathbf{x}_0 - \bar{\mathbf{x}}\| \leq \delta(\varepsilon) \Rightarrow \forall t \geq 0 \|\mathbf{x}(t) - \bar{\mathbf{x}}\| \leq \varepsilon$$

gdzie $\mathbf{x}(t)$ jest rozwiązaniem równania (5.14) przy warunku początkowym $\mathbf{x}(0) = \mathbf{x}_0$.

Rysunek 5.5 ilustruje warunek stabilności w przypadku, gdy $\mathbf{x} = (x_1, x_2)$: dla dowolnego otoczenia o promieniu $\varepsilon > 0$ punktu równowagi $\bar{\mathbf{x}}$ istnieje otoczenie tego punktu o promieniu $\delta(\varepsilon)$ ($0 < \delta(\varepsilon) \leq \varepsilon$), takie, że jeśli punkt początkowy \mathbf{x}_0 będzie do niego należał, to trajektoria układu wychodząca z punktu początkowego $\mathbf{x}(0) = \mathbf{x}_0$, takiego, że $\|\mathbf{x}_0 - \bar{\mathbf{x}}\| \leq \delta(\varepsilon)$, nie opuści otoczenia o promieniu ε , tj. $\forall t > 0 : \|\mathbf{x}_0 - \mathbf{x}(t)\| \leq \varepsilon$.



Rys. 5.5. Ilustracja stabilności punktu równowagi wg Lapunowa.

Wracając do przykładu obiektu inercji z opóźnieniem, bez wymuszenia zewnętrznego, tj.

$$\dot{y}(t) = -\frac{1}{T} y(t)$$

uzyskujemy punkt równowagi oraz wartości stałych ε i δ równe $\bar{y} = 0$, $\forall \varepsilon > 0 \ \delta = \varepsilon$. Tak więc układ jest stabilny wg Lapunowa.

Definicja. Stabilność asymptotyczna w sensie Lapunowa.

Punkt równowagi $\bar{\mathbf{x}}$ układu (5.14) jest stabilny asymptotycznie w sensie Lapunowa, jeśli jest stabilny w sensie Lapunowa oraz, dodatkowo, $\|\mathbf{x}(t) - \bar{\mathbf{x}}\| \xrightarrow[t \rightarrow +\infty]{} 0$.

Następujące twierdzenie, zwane *Pierwszym Twierdzeniem Lapunowa*, jest bardzo ważne:

Twierdzenie 5.1. *Niech będzie dany układ nieliniowy $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t))$, gdzie funkcja \mathbf{f} jest klasy C^1 ; niech $\bar{\mathbf{x}}$ będzie punktem równowagi tego układu, tj. $\mathbf{f}(\bar{\mathbf{x}}) = 0$. Jeśli układ zlinearyzowany lokalnie wokół tego punktu równowagi, tj. układ $\dot{\mathbf{v}} = \mathbf{A}\mathbf{v}$, $\mathbf{v} \in R^n$, gdzie $\mathbf{v} = \mathbf{x} - \bar{\mathbf{x}}$, $\mathbf{A} = \left[\frac{\partial f_i}{\partial x_j}(\bar{\mathbf{x}}) \right]$, jest stabilny asymptotycznie, to punkt $\bar{\mathbf{x}}$ jest stabilnym asymptotycznym punktem równowagi układu nieliniowego.*

Powyzsze twierdzenie jest *bardzo ważne*, gdyż, w istocie, to właśnie ono pozwala nam projektować układ regulacji posługując się opisem (modelem) liniowym obiektu otrzymanym w drodze linearyzacji lokalnej: *jeśli zapewniamy asymptotyczną stabilność układu regulacji z obiektem liniowym, to będziemy mieli pewność, że rzeczywisty (nieliniowy) układ regulacji będzie asymptotycznie stabilny przynajmniej w pewnym otoczeniu punktu równowagi*.

Warto zauważyć, że analogiczne twierdzenie dotyczące zwykłej stabilności, w miejscu stabilności asymptotycznej, nie jest – w ogólnym przypadku – prawdziwe. Na przykład układ nieliniowy $\dot{x} = x^2$, $x \in R$, jest zdecydowanie niestabilny w punkcie równowagi $\bar{x} = 0$, podczas gdy układ zlinearyzowany wokół tego punktu równowagi, tj. układ opisany równaniem $\dot{v} = 0$, jest stabilny wg Lapunowa – trzeba zauważyć, że nie jest to układ asymptotycznie stabilny!

W świetle powyższych uwag musimy zawsze wymagać asymptotycznej stabilności zaprojektowanego układu regulacji traktowanego jako autonomiczny układ dynamiczny.

Stabilność w sensie ograniczone wejście – ograniczone wyjście (OWOW, ang. BIBO – bounded input – bounded output)

Rozważmy sterowany układ dynamiczny

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), u(t)), \quad \mathbf{x} \in R^n, \quad u \in R$$

Definicja. Sterowany układ dynamiczny nazywamy globalnie stabilnym w sensie **OWOW**, jeśli dla dowolnego warunku początkowego \mathbf{x}_0 układ ten odpowiada jednostajnie ograniczonym rozwiązaniem $\mathbf{x}(t)$ na dowolne, jednostajnie ograniczone pobudzenie $u(t)$, tzn.

$$|u(t)| \leq M \quad \Rightarrow \quad \forall t \geq 0 \quad \|\mathbf{x}(t)\| \leq K(\mathbf{x}_0, M)$$

gdzie stała $K(x_0, M)$ może, jak zaznaczono, być zależna od x_0 i M .

Można udowodnić następujący ważny fakt:

Twierdzenie 5.2. *W przypadku układu liniowego stabilność asymptotyczna wg Lapunowa implikuje globalną stabilność OWOW.*

W odniesieniu do obiektu nieliniowego powyższa implikacja nie jest na ogół prawdziwa. Na przykład układ nieliniowy $\dot{y} = -y + y \cdot u$ jest przy braku sygnału pobudzającego ($u(t) \equiv 0$) stabilny w punkcie równowagi $y = 0$, i to stabilny asymptotycznie. Przy pojawienniu się odpowiednio silnego ograniczonego sygnału pobudzającego, np. stałego o wartości większej niż 1, otrzymujemy układ niestabilny: dla $u(t) = a$, $a > 1$ oraz $y_0 \neq 0$ zachodzi $|y(t)| \rightarrow \infty$.

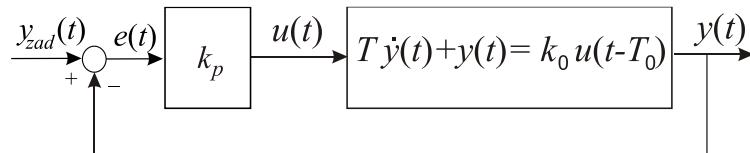
W rozdziale następnym podamy warunki asymptotycznej stabilności liniowego stacjonarnego układu dynamicznego.

5.2. Analiza w dziedzinie zmiennej zespolonej (metoda operatorowa)

Przypuśćmy, że chcemy dokonać analizy układów zamkniętych, w których jako obiekt występuje człon inercyjny z opóźnieniem. Rozważmy układ jak na rysunku 5.6. Jest to układ prostej regulacji proporcjonalnej, tj.

$$u(t) = k_p \cdot (y_{zad}(t) - y(t))$$

gdzie $y_{zad}(t)$ jest sygnałem zadanym.



Rys. 5.6. Układ regulacji członu inercyjnego z opóźnieniem, z regulatorem typu P.

Równanie układu zamkniętego jest następujące:

$$T\ddot{y}(t) + y(t) = k_0 k_p [y_{zad}(t - T_0) - y(t - T_0)].$$

Jest to równanie układu z opóźnieniem występującym w zmiennej stanu, tzw. równanie liniowe z odchylym argumentem, trudne do rozwiązania (choć w tym przypadku rozwiązanie analityczne jest jeszcze możliwe). Rozwiązywanie może być łatwo otrzymane metodami numerycznymi, ale tylko dla konkretnych wartości parametrów. Powstaje pytanie o to, jak badać własności rozwiązania powyższego równania, a także innych, często bardziej skomplikowanych, opisów innych układów zamkniętych, bez konieczności rozwiązywania tych równań? Możemy to czynić wykorzystując różne metody analizy nie wymagające bezpośredniego rozwiązywania równań modelu. Są to, m. in.:

- analiza z wykorzystaniem transformaty Laplace'a w płaszczyźnie zmiennej zespolonej s,
- analiza z wykorzystaniem charakterystyk częstotliwościowych,
- analiza w przestrzeni stanu (np. metody płaszczyzny fazowej).

5.2.1. Transformata Laplace'a, transmitancja układu dynamicznego

Dla przedstawionych dalej rozważań potrzebne jest przypomnienie faktów związanych z ważnym w wielu zastosowaniach pojęciem transformaty Laplace'a. Rozważmy mianowicie funkcję $f(t)$ jednej zmiennej rzeczywistej, o wartościach rzeczywistych. Funkcję tę nazywamy *oryginałem Laplace'a*, jeśli

$$f(t) = 0 \text{ dla } t < 0, \text{ jest przedziałami ciągła oraz } |f(t)| \leq M e^{\alpha t} \text{ dla } t \geq 0,$$

gdzie wartość stałej α może być dowolna (ma ona jednak istotne znaczenie). Dla takiej funkcji możemy zdefiniować przekształcenie całkowe, zwane transformatą Laplace'a:

$$\mathcal{L}[f(t)](s) = F(s) = \int_0^{+\infty} e^{-st} f(t) dt \quad (5.15)$$

gdzie $s \in \mathbb{C}$ (zm. zespolona) oraz $\operatorname{Re} s > \alpha$; należy podkreślić, że dla $\operatorname{Re} s \leq \alpha$ całka w powyższym przekształceniu może być rozbieżna, a stąd przekształcenie jest wówczas nieokreślone (zgodnie z tradycją dużą literą oznaczamy transformatę funkcji określonej małą literą).

Uwaga: nie są, w szczególności, określone wartości transformat Laplace'a funkcji $\sin \omega t$ oraz $\cos \omega t$ dla argumentu zespolonego $s = j\omega$!

Mamy natomiast:

$$f(t) = \sin \omega t \rightarrow F(s) = \frac{\omega}{s^2 + \omega^2} \text{ dla każdego } s \text{ takiego, że } \operatorname{Re} s > 0,$$

$$f(t) = \cos \omega t \rightarrow F(s) = \frac{s}{s^2 + \omega^2} \text{ dla każdego } s \text{ takiego, że } \operatorname{Re} s > 0.$$

Także

$$f(t) = t^n \rightarrow F(s) = \frac{n!}{s^{n+1}}, \quad n = 0, 1, 2, \dots \quad \operatorname{Re} s > 0. \quad (5.16)$$

Warto jeszcze przypomnieć następującą ważną własność transformaty:

$$\mathcal{L}[e^{-at} f(t)](s) = \int_0^{+\infty} e^{-st-at} f(t) dt = F(s+a), \quad (5.17)$$

gdzie a jest liczbą zespoloną; w szczególności, dla przykładu:

$$\mathcal{L}[e^{-at} 1(t)](s) = \int_0^{+\infty} e^{-st-at} dt = \frac{1}{s+a}. \quad (5.18)$$

Można zdefiniować transformatę odwrotną w postaci przekształcenia:

$$\mathcal{L}^{-1}[F(s)](t) = f(t) = \frac{1}{2\pi j} \int_{\sigma_c - j\omega}^{\sigma_c + j\omega} F(s) e^{st} ds. \quad (5.19)$$

W praktyce to trudne przekształcenie odwrotne nie jest wykorzystywane. Mając daną transformatę $F(s)$ odpowiadający jej oryginał $f(t)$ otrzymuje się w oparciu o odpowiednie przekształcenia transformaty (w szczególności rozkład na ułamki proste) oraz znajomość postaci transformat dla najczęściej spotykanych funkcji $f(t)$; niektóre z nich podane zostały powyżej.

Ważne są następujące własności przekształcenia Laplace'a:

– Bezpośrednio z definicji wynika addytywność:

$$\mathcal{L}[\beta f_1(t) + \gamma f_2(t)](s) = \beta F_1(s) + \gamma F_2(s) \quad (5.20)$$

– Jeśli oryginał Laplace'a $f(t)$ jest różniczkowalny (w zerze mamy tu na myśli pochodną prawostronną), to transformata pochodnej tej funkcji, która jak łatwo pokazać także jest oryginałem, dana jest w postaci

$$\mathcal{L}\left[\frac{df(t)}{dt}\right](s) = \int_0^{+\infty} e^{-st} \frac{f(t)}{dt} dt = sF(s) - f(0), \quad (5.21)$$

gdzie $F(s)$ jest transformata funkcji $f(t)$.

– Transformata całki funkcji $f(t)$ będącej oryginałem Laplace'a, dana jest w postaci

$$\mathcal{L}\left[\int_0^t f(\tau) d\tau\right](s) = \int_0^{+\infty} e^{-st} \left(\int_0^t f(\tau) d\tau\right) dt = \frac{1}{s} F(s), \quad (5.22)$$

gdzie, jak poprzednio, $F(s)$ jest transformata funkcji $f(t)$; oczywiście całka z funkcji będącej oryginałem Laplace'a także jest oryginałem.

– Transformata funkcji z opóźnionym argumentem

$$\mathcal{L}[f(t - T_0)](s) = e^{-sT_0} F(s) \quad (5.23)$$

– Jeśli istnieje granica oryginału w nieskończoności, to słuszne jest twierdzenie

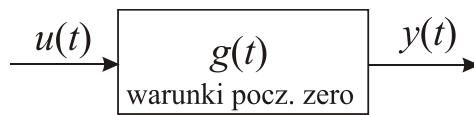
$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} [sF(s)] \quad (5.24)$$

Widzimy w szczególności, że operacje różniczkowania i całkowania w dziedzinie czasu odpowiadają prostym operacjom algebraicznym, odpowiednio, mnożenia lub dzielenia w dziedzinie zmiennej zespolonej s .

Podstawowym powodem, dla którego wykorzystanie przekształcenia Laplace'a jest niezmiernie wygodne w zastosowaniach do badania i projektowania układów liniowych jest bardzo cenna własność, wynikająca z twierdzenia Borela o splocie funkcji f i g , będących oryginałami Laplace'a.

$$\mathcal{L}\left[\int_0^t f(t - \tau) g(t) d\tau\right](s) = F(s)G(s) \quad (5.25)$$

Rozważmy układ liniowy z obiektem opisanym odpowiedzią impulsową $g(.)$, o której zakładamy, że jest oryginałem Laplace'a, przedstawiony na Rys. 5.7.



Rys. 5.7. Układ liniowy określony przy pomocy odpowiedzi impulsowej.

Wejście sterujące $u(.)$ niech będzie oryginałem Laplace'a. Wówczas wyjście układu $y(.)$ także jest oryginałem Laplace'a i, przy zerowych warunkach początkowych, ma uprzednio omawianą postać całki splotu:

$$y(t) = \int_0^t g(t - \tau) u(\tau) d\tau \quad (5.26)$$

Oznaczmy transformatę Laplace'a odpowiedzi impulsowej jako $G(s)$. Wówczas zachodzi następująca własność:

$$Y(s) = G(s) \cdot U(s) \quad (5.27)$$

Ta bardzo ważna, podstawowa dla dalszej analizy zależność obowiązuje, oczywiście, dla tych wartości zmiennej zespolonej s , dla których istnieje (jest określona) zarówno $G(s)$ jak i transformata $U(s)$ sygnału wejściowego. Widzimy, że całka splotu po dokonaniu przekształcenia Laplace'a przechodzi w iloczyn transformat. Jest to niezmiernie cenna właściwość przekształcenia operatorowego.

Definicja. Transmitancją $G(s)$ liniowego, stacjonarnego układu dynamicznego nazywamy transformatę Laplace'a jego odpowiedzi impulsowej $g(t)$.

Z zależności (5.27) wynika natychmiast

$$G(s) = \frac{Y(s)}{U(s)}$$

Definicja równoważna: Transmitancja $G(s)$ liniowego, stacjonarnego układu dynamicznego jest równa ilorazowi transformat Laplace'a, odpowiednio, przebiegu wyjścia i wejścia, przy zerowych warunkach początkowych.

Transmitancja układu liniowego danego równaniami stanu w postaci normalnej

Rozważmy układ o jednym wejściu i jednym wyjściu (SISO) stacjonarny liniowy, rzędu n , opisany układem równań normalnych

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) &= \mathbf{c}^T \mathbf{x}(t) + du(t) \end{aligned} \quad (5.28)$$

Po dokonaniu przekształcenia operatorowego tych równań, zgodnie z własnościami przekształcenia Laplace'a, przy zerowych warunkach początkowych, mamy

$$\begin{aligned} s \cdot X(s) - \mathbf{x}_0 &= \mathbf{A} \cdot X(s) + \mathbf{b} \cdot U(s) \\ Y(s) &= \mathbf{c}^T X(s) + d \cdot U(s) \end{aligned}$$

Stąd

$$\begin{aligned} X(s) &= (s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{x}_0 + (s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{b} \cdot U(s) \\ Y(s) &= \mathbf{c}^T \cdot (s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{x}_0 + [\mathbf{c}^T \cdot (s\mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{b} + d] \cdot U(s) \end{aligned} \quad (5.29)$$

Porównując (5.29) z (5.11) wnioskujemy, że

$$\mathcal{L}[e^{\mathbf{A}t}](s) = (s\mathbf{I} - \mathbf{A})^{-1} \quad (5.30)$$

Przyjmując zerowy warunek początkowy, z (5.29) wynika wzór na transmitancję obiektu

$$G(s) = \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} + d \quad (5.31)$$

Jest to wzór dla przypadku obiektu SISO, dla ogólnego obiektu o wielu wejściach i wyjściach (gdy u i y są wektorami), wektory \mathbf{b} i \mathbf{c}^T oraz skalar d stają się macierzami, odpowiednio, \mathbf{B} , \mathbf{C} i \mathbf{D} , o wymiarach wynikających z wymiarów wektorów \mathbf{u} i \mathbf{y} . W pracy niniejszej zajmujemy się tylko układami SISO, stąd stosujemy zapis odpowiedni do tego przypadku.

Transmitancja (5.31) jest funkcją wymierną, mianowicie

$$G(s) = \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} + d = \frac{L(s)}{M(s)} \quad (5.32)$$

Z definicji macierzy odwrotnej wynika, że $M(s) = \det[s\mathbf{I} - \mathbf{A}]$. Stąd pierwiastki $M(s)$, będące

biegunami transmitancji, są pierwiastkami równania $\det[s\mathbf{I} - \mathbf{A}] = 0$, zwanego *równaniem charakterystycznym* macierzy \mathbf{A} . Jeśli $d = 0$, to stopień wielomianu L jest mniejszy niż stopień wielomianu M ; mówimy wówczas, że transmitancja jest *właściwa*.

Jeśli z kolei opis układu liniowego dany jest w postaci równania różniczkowego, typu wejście-wyjście, rzędu m , tj.

$$\frac{d^m y}{dt^m} + a_{m-1} \frac{d^{m-1} y}{dt^{m-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = b_l \frac{d^l u}{dt^l} + \dots + b_1 \frac{du}{dt} + b_0 u \quad (5.33)$$

to transmitancja tego układu wyraża się zależnością

$$G(s) = \frac{b_l s^l + \dots + b_1 s + b_0}{s^m + a_{m-1} s^{m-1} + \dots + a_1 s + a_0} = \frac{\tilde{L}(s)}{\tilde{M}(s)} \quad (5.34)$$

gdzie mianownik transmitancji $\tilde{M}(s)$ jest *wielomianem charakterystycznym* stwarzyszonym z równaniem różniczkowym (a równanie $\tilde{M}(s) = 0$ to równanie charakterystyczne). Jeśli modele (5.28) i (5.33) opisują ten sam układ dynamiczny, to oczywiście $m = n$ i wielomian $\tilde{M}(s)$ jest tożsamy z wielomianem $M(s)$ (a równanie charakterystyczne $\tilde{M}(s) = 0$ to jednocześnie równanie $\det[s\mathbf{I} - \mathbf{A}] = 0$).

Jeśli w wejściu sterującym występuje opóźnienie, to znaczy zamiast $u(t)$ mamy $u(t-T_0)$, to, zgodnie z właściwościami transformaty Laplace'a, transformatę $U(s)$ będącego oryginałem sterowania $u(t)$ zastępujemy w powyższych zależnościach poprzez wyrażenie $U(s)\exp(-sT_0)$; inaczej mówiąc transmitancja $G_{op}(s)$ układu liniowego z opóźnieniem równym T_0 w sterowaniu wyraża się jako

$$G_{op}(s) = G(s)e^{-sT_0}.$$

Niech transmitancja obiektu będzie funkcją wymierną. Jeśli transformata sygnału wejściowego $u(t)$ jest również funkcją wymierną, na przykład jest sumą funkcji sinus lub cosinus lub funkcji typu t^n , to wówczas

$$Y(s) = G(s) \cdot U(s) = \frac{L_y(s)}{M_y(s)}$$

tj. transformata wyjścia jest także funkcją wymierną. Przebieg wyjścia jako funkcji czasu można w takim przypadku uzyskać przeprowadzając rozkład tej funkcji wymiernej na ułamki proste i odczytując z tablic transformat Laplace'a oryginały, tj. funkcje odpowiadające transmitancjom tych ułamków prostych. W ten sposób można wyznaczać np. odpowiedzi skokowe układów dynamicznych.

Przykład 5.1.

Rozważmy układ dynamiczny dany układem równań w postaci normalnej

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -a_{11} & 0 \\ a_{21} & -a_{22} \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} b_1 \\ 0 \end{bmatrix} u(t), \quad y(t) = x_2(t)$$

gdzie $a_{11} > 0$, $a_{21} > 0$, $a_{22} > 0$; reprezentuje on np. zlinearyzowany model kaskady dwóch zbiorników z Przykładu 4.5. Wyznaczmy ogólnie transmitancję tego układu oraz odpowiedź impulsową. Wyznaczmy ponadto macierz $e^{\mathbf{A}t}$.

Mamy

$$(s\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} s + a_{11} & 0 \\ -a_{21} & s + a_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{s + a_{11}} & 0 \\ \frac{a_{21}}{(s + a_{11})(s + a_{22})} & \frac{1}{s + a_{22}} \end{bmatrix}$$

Stąd z zależności (5.31) mamy

$$G(s) = [0 \quad 1] \begin{bmatrix} \frac{1}{s + a_{11}} & 0 \\ \frac{a_{21}}{(s + a_{11})(s + a_{22})} & \frac{1}{s + a_{22}} \end{bmatrix} \begin{bmatrix} b_1 \\ 0 \end{bmatrix} = \frac{a_{21}b_1}{(s + a_{11})(s + a_{22})}$$

Uzyskaliśmy transmitancję układu dwuineracyjnego.

Wyznaczmy odpowiedź impulsową. Dokonamy rozkładu transmitancji na ułamki proste

$$G(s) = \frac{a_{21}b_1}{(s + a_{11})(s + a_{22})} = \frac{a_{21}b_1}{a_{22} - a_{11}} \frac{1}{s + a_{11}} - \frac{a_{21}b_1}{a_{22} - a_{11}} \frac{1}{s + a_{22}}, \quad \text{skąd}$$

$$g(t) = \mathcal{L}^{-1}[G(s)](t) = \frac{a_{21}b_1}{a_{22} - a_{11}} (e^{-a_{11}t} - e^{-a_{22}t})$$

Korzystając z wzoru (5.30) oraz tabeli transformat Laplace'a mamy

$$\begin{aligned} e^{\mathbf{A}t} &= \mathcal{L}^{-1}[(s\mathbf{I} - \mathbf{A})^{-1}](t) = \mathcal{L}^{-1} \left[\begin{bmatrix} \frac{1}{s + a_{11}} & 0 \\ \frac{a_{21}}{(s + a_{11})(s + a_{22})} & \frac{1}{s + a_{22}} \end{bmatrix} \right] (t) = \\ &= \begin{bmatrix} e^{-a_{11}t} & 0 \\ \frac{a_{21}}{a_{22} - a_{11}} (e^{-a_{11}t} - e^{-a_{22}t}) & e^{-a_{22}t} \end{bmatrix} \end{aligned}$$

Zwróćmy uwagę, że korzystając teraz z wzoru $g(t) = \mathbf{c}^T e^{\mathbf{A}t} \mathbf{b}$ dostaniemy identyczną postać odpowiedzi impulsowej jak uzyskana wyżej na podstawie transmitancji. \square

Przykład 5.2.

Wyznaczyć odpowiedź skokową dla układu z przykładu powyżej, ogólnie i dla $a_{11}=4$, $a_{21}=3$, $a_{22}=6$, $b_1=1$.

Odpowiedź skokowa to odpowiedź na impuls jednostkowy $\mathbf{1}(t)$, którego transformata jest $1/s$. Dokonując rozkładu na ułamki proste, mamy

$$\begin{aligned} Y(s) = G(s) \frac{1}{s} &= \frac{a_{21}b_1}{s(s + a_{11})(s + a_{22})} = \frac{A}{s} + \frac{B}{s + a_{11}} + \frac{C}{s + a_{22}} = \\ &= \frac{A(s + a_{11})(s + a_{22}) + Bs(s + a_{22}) + C(s + a_{11})}{s(s + a_{11})(s + a_{22})} \end{aligned}$$

skąd dostajemy układ równań liniowych na współczynniki A , B i C rozkładu na ułamki proste

$$\begin{aligned} A + B + C &= 0 \\ A(a_{11} + a_{22}) + Ba_{22} + Ca_{11} &= 0 \\ Aa_{11}a_{22} &= a_{21}b_1 \end{aligned}$$

Rozwiążając powyższy układ równań dostajemy

$$A = \frac{a_{21}b_1}{a_{11}a_{22}}, \quad B = -\frac{a_{21}b_1}{a_{11}(a_{22} - a_{11})}, \quad C = \frac{a_{21}b_1}{a_{22}(a_{22} - a_{11})}$$

Czyli poszukiwana odpowiedź skokowa jest w ogólnej postaci

$$y(t) = \left(\frac{a_{21}b_1}{a_{11}a_{22}} - \frac{a_{21}b_1}{a_{11}(a_{22} - a_{11})} e^{-a_{11}t} + \frac{a_{21}b_1}{a_{22}(a_{22} - a_{11})} e^{-a_{22}t} \right) \cdot \mathbf{1}(t)$$

która dla $a_{11}=4$, $a_{21}=3$, $a_{22}=6$, $b_1=1$, przyjmuje postać

$$y(t) = \left(\frac{1}{8} - \frac{3}{8} e^{-4t} + \frac{2}{8} e^{-6t} \right) \cdot \mathbf{1}(t)$$

□

5.2.2. Odpowiedzi skokowe i transmitancje podstawowych członów dynamicznych

Człon proporcjonalny

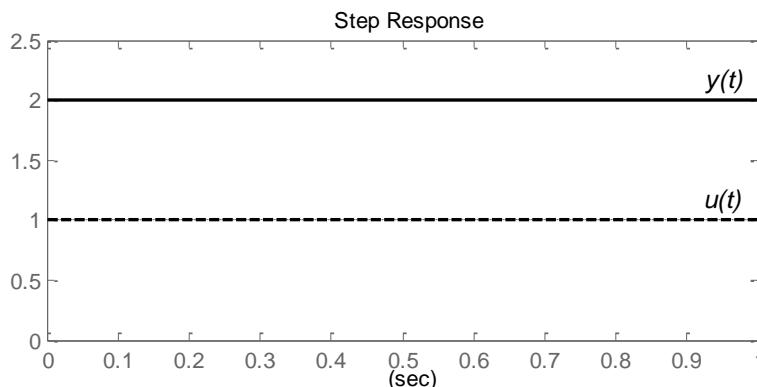
Równanie członu

$$y(t) = ku(t), \quad k > 0,$$

stąd bezpośrednio transmitancja:

$$G(s) = \frac{Y(s)}{U(s)} = k$$

Odpowiedź skokowa członu proporcjonalnego przedstawiona jest na Rys. 5.8.



Rys. 5.8. Odpowiedź skokowa członu proporcjonalnego o wzmacnieniu $k=2$ (linia przerywana – sygnał skoku jednostkowego $u(t)$ na wejściu członu).

Człon inercyjny z opóźnieniem

Równanie różniczkowe tego, uprzednio szczegółowo omawianego, członu dynamicznego:

$$T\ddot{y}(t) + y(t) = ku(t - T_0), \quad T > 0, \quad T_0 \geq 0, \quad k > 0$$

Dokonując przekształcenia Laplace'a, przy zerowych warunkach początkowych, mamy

$$sTY(s) + Y(s) = kU(s)e^{-sT_0},$$

skąd bezpośrednio

$$G(s) = \frac{Y(s)}{U(s)} = \frac{k}{1+sT} e^{-sT_0}$$

Transmitancję **członu inercyjnego** (bez opóźnienia) o wzmacnieniu k_0 dostajemy dla $T_0 = 0$:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{k}{1+sT}$$

Natomiast transmitancję samego **członu opóźniającego** dostajemy przy $T = 0$ (i $k = 1$):

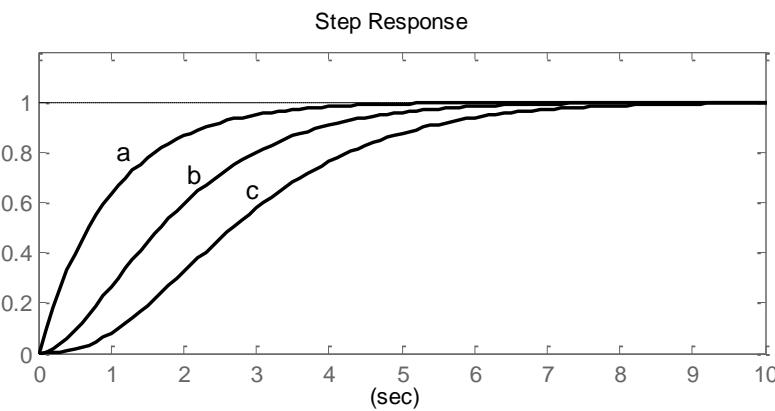
$$G(s) = \frac{Y(s)}{U(s)} = e^{-sT_0}$$

Odpowiedź skokowa członu inercyjnego z opóźnieniem podana była na Rys. 5.2. Latwo zmodyfikować ją do szczególnych przypadków odpowiedzi członów inercyjnego i opóźniającego. Natomiast na Rys. 5.9 przedstawiono odpowiedzi skokowe członów:

- a) Jednoinercyjnego o transmitancji $G(s) = \frac{k}{1+sT}$,
- b) Dwuinercyjnego o transmitancji $G(s) = \frac{k}{(1+sT)(1+sT)}$, który możemy traktować jako szeregowego połączenie dwóch członów inercyjnych (jednoinercyjnych):

$$G(s) = \frac{k_1}{1+sT} \cdot \frac{k_2}{1+sT}, \quad k_1 k_2 = k$$

- c) Trójinercyjnego o transmitancji $G(s) = \frac{k}{(1+sT)(1+sT)(1+sT)}$.



Rys. 5.9. Odpowiedzi skokowe członów inercyjnych dla $T = 1$ i $k = 1$.

Zwrócmy uwagę na pojawianie się charakterystycznego, opóźniającego kształtu początku narastania odpowiedzi skokowej wraz ze wzrostem liczby inercji.

Człoy różniczkujący i proporcjonalno-różniczkujący (idealne)

Równanie członu idealnego **różniczkującego**

$$y(t) = T\dot{u}(t), \quad T > 0,$$

odpowiedzią skokową tego członu jest impuls Diraca (impuls idealny) o polu T , stąd oczywiście jest on nierealizowalny.

Równanie różniczkowe członu idealnego **proporcjonalno-różniczkującego** jest sumą transmitancji członów proporcjonalnego i różniczkującego, postaci

$$\begin{aligned} y(t) &= ku(t) + T\dot{u}(t), \quad T > 0, \quad k > 0, \quad \text{lub} \\ y(t) &= k(u(t) + T\dot{u}(t)). \end{aligned}$$

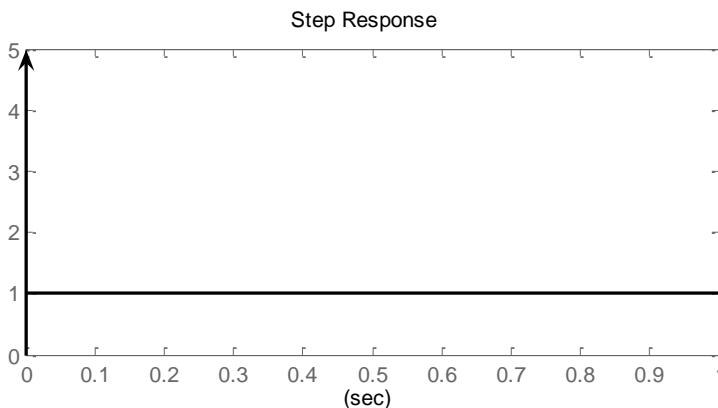
Dokonując przekształcenia Laplace'a, przy zerowych warunkach początkowych, mamy

$$Y(s) = k(U(s) + TsU(s))$$

skąd bezpośrednio

$$G(s) = \frac{Y(s)}{U(s)} = k(1 + sT).$$

Odpowiedź skokowa idealnego członu proporcjonalno-różniczkującego przedstawiona jest na Rys. 5.10. Zwróćmy uwagę na impuls Diraca w chwili zerowej, przedstawiony strzałką.



Rys. 5.10. Odpowiedź skokowa idealnego członu proporcjonalno-różniczkującego, $T=1$, $k=1$.

Człon proporcjonalno-różniczkujący rzeczywisty (realizowalny)

Równanie różniczkowe członu

$$y(t) + \frac{T}{\alpha} \dot{y}(t) = k(u(t) + T\dot{u}(t)), \quad T > 0, \quad k > 0, \quad \alpha > 1$$

Dokonując przekształcenia Laplace'a, przy zerowych warunkach początkowych, mamy

$$Y(s)(1 + s\frac{T}{\alpha}) = k(U(s) + TsU(s))$$

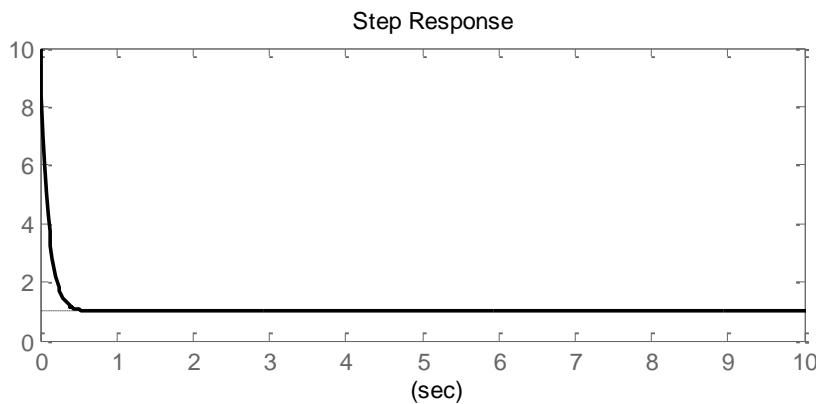
skąd bezpośrednio

$$G(s) = \frac{Y(s)}{U(s)} = \frac{k(1 + sT)}{1 + s(T/\alpha)}.$$

Transmitancję tę można przedstawić w postaci:

$$G(s) = \frac{Y(s)}{U(s)} = k(1 + sT) \cdot \frac{1}{1 + s(T/\alpha)}$$

Stąd interpretacja członu jako *szeregowego połączenia członów idealnego proporcjonalno-różniczkującego i inercyjnego*. Parametr α ustawiany jest typowo na wartości ok. 8 do 10 (tak aby działanie inercji było odpowiednio szybsze niż różniczkowania – nie dominowało nad działaniem różniczkującym). Odpowiedź skokowa członu przedstawiona jest na Rys. 5.11. Zwróćmy uwagę na początkowy fragment odpowiedzi, aproksymujący impuls idealny (Diraca) – im większą będzie wartość α , tym większy skok wartości w chwili zerowej (równy α) i szybsze opadanie odpowiedzi do wartości ustalonej.



Rys. 5.11. Odpowiedź skokowa rzeczywistego członu różniczkującego dla $T = 1$, $\alpha = 10$, $k = 1$.

Człon całkujący

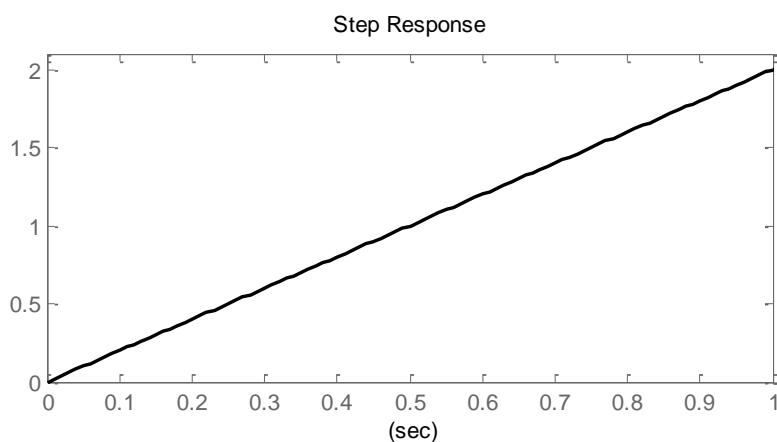
Równanie opisujące dynamikę członu całkującego z sygnałem wejściowym $u(t)$ ma postać

$$y(t) = k \int_0^t u(\tau) d\tau$$

Stąd bezpośrednio transmitancja Laplace'a

$$G(s) = \frac{Y(s)}{U(s)} = \frac{k}{s}$$

Odpowiedź skokowa członu przedstawiona jest na Rys. 5.12.



Rys. 5.12. Odpowiedź skokowa członu całkującego dla $k = 2$.

Człon drugiego rzędu (oscylacyjny)

Ważne miejsce w zestawie podstawowych liniowych członów dynamicznych zajmuje człon drugiego rzędu, zwany oscylacyjnym, który jednakże właściwości oscylacyjne wykazuje tylko dla pewnego zakresu swych parametrów – i jest to wówczas najprostszy i w wielu przypadkach wystarczający model układu drgającego, z tłumieniem drgań.

Równanie różniczkowe członu drugiego rzędu ma postać:

$$T^2 \ddot{y}(t) + 2\zeta T \dot{y}(t) + y(t) = ku(t), \quad k > 0, \quad \zeta > 0,$$

lub, oznaczając $\omega_n = 1/T$, przekształconą postać najczęściej spotykaną, wygodną dla analizy:

$$\ddot{y}(t) + 2\zeta\omega_n \dot{y}(t) + \omega_n^2 y(t) = k\omega_n^2 u(t)$$

Stąd, zakładając zerowe warunki początkowe, dostajemy transmitancję:

$$G(s) = \frac{k\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (5.35)$$

Rozpatrzmy dwa przypadki: $\zeta \geq 1$ i $0 < \zeta < 1$.

Jeśli $\zeta \geq 1$, to biegury $G(s)$ są rzeczywiste:

$$p_1 = -\zeta\omega_n + \omega_n \sqrt{\zeta^2 - 1}, \quad p_2 = -\zeta\omega_n - \omega_n \sqrt{\zeta^2 - 1}$$

Układ można wówczas przedstawić jako szeregowego połączenie dwóch członów inercyjnych:

$$G(s) = \frac{k_1}{1+sT_1} \cdot \frac{k_2}{1+sT_2}, \quad T_1 = -\frac{1}{p_1}, \quad T_2 = -\frac{1}{p_2}, \quad T_1 \geq T_2, \quad k_1 k_2 = k$$

Gdy dodatni współczynnik tłumienia ζ ma wartość mniejszą od jedności, tj. $0 < \zeta < 1$, to biegury transmitancji są zespolone o następujących wartościach:

$$p_1 = -\zeta\omega_n + \omega_n j \sqrt{1 - \zeta^2}, \quad p_2 = -\zeta\omega_n - \omega_n j \sqrt{1 - \zeta^2}$$

i członu tego nie można wówczas przedstawić w postaci połączenia szeregowego dwóch członów inercyjnych. Warto tutaj poświęcić nieco uwagi kształtowi odpowiedzi impulsowej. Transmitancję operatorową można zapisać w postaci:

$$G(s) = \frac{k\omega_n^2}{\omega_d} \cdot \frac{\omega_d}{(s + \zeta\omega_n)^2 + \omega_d^2}, \quad \omega_d = \omega_n \sqrt{1 - \zeta^2} \quad (5.36)$$

Jest to, jak widać z tabel transformat Laplace'a, transformata sinusoidy o pulsacji ω_d (pomnożona przez odpowiedni stały współczynnik), dzięki temu możemy natychmiast podać oryginał Laplace'a odpowiadający transmitancji $G(s)$, tj. odpowiedź impulsową:

$$g(t) = \frac{k\omega_n^2}{\omega_d} e^{-\zeta\omega_n t} \cdot \sin(\omega_d t) = \frac{k\omega_n}{\sqrt{1 - \zeta^2}} e^{-\zeta\omega_n t} \cdot \sin(\omega_d t) \quad (5.37)$$

Wielkość $\omega_d = \omega_n \sqrt{1 - \zeta^2}$ odpowiada pulsacji drgań tłumionych; wraz ze wzrostem współczynnika tłumienia pulsacja ta maleje. W przypadku braku tłumienia ($\zeta=0$) otrzymujemy układ drgający z naturalną pulsacją drgań równą ω_n .

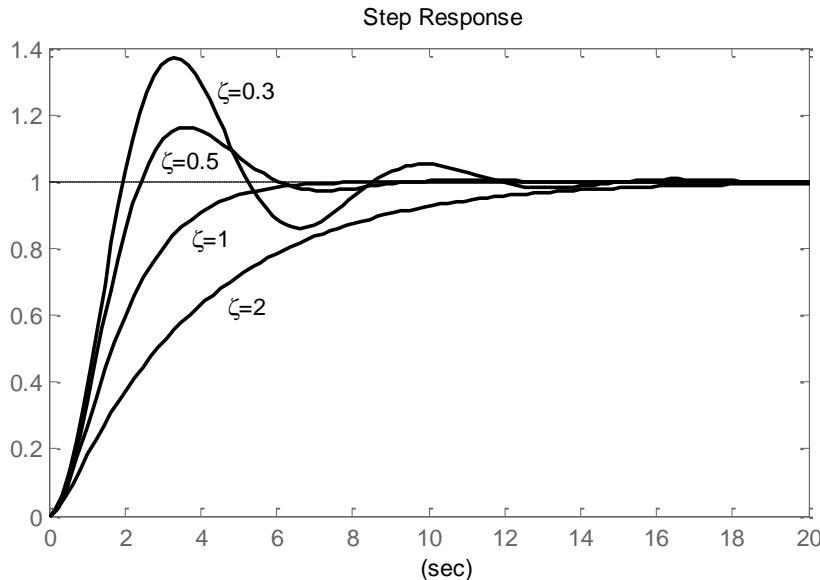
Natomiast odpowiedź skokową członu oscylacyjnego otrzymamy całkując odpowiedź impulsową lub dokonując odwrotnej transformaty Laplace'a

$$h(t) = \mathcal{L}^{-1} \left[\frac{k\omega_n^2}{\omega_d} \cdot \frac{\omega_d}{(s + \zeta\omega_n)^2 + \omega_d^2} \cdot \frac{1}{s} \right]$$

Można pokazać, że odpowiedź skokowa ma postać

$$h(t) = k \left[1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \cdot \sin(\omega_d t + \varphi) \right], \quad \varphi = \arctg \frac{\sqrt{1-\zeta^2}}{\zeta}.$$

Na Rys. 5.13 przedstawiono przebiegi kilku odpowiedzi skokowych członu drugiego rzędu, dla różnych wartości współczynnika tłumienia ζ .



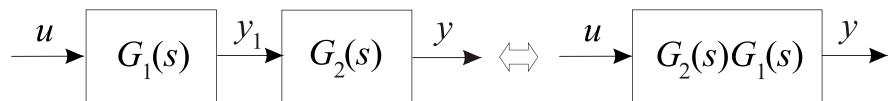
Rys. 5.13. Odpowiedzi skokowe członu drugiego rzędu dla kilku wartości współczynnika tłumienia ζ , przy $k = 1$, $\omega_n = 1$.

5.2.3. Przekształcanie schematów blokowych liniowych członów dynamicznych

Struktury układów sterowania przedstawia się często w postaci schematów blokowych, których elementami są człony dynamiczne. Jest to szczególnie dogodne w przypadku układów dynamicznych liniowych, przedstawianych w postaci transmitancji, gdyż łatwo jest dokonywać strukturalnych przekształceń takich schematów blokowych. Poniżej przedstawimy proste reguły elementarne takich przekształceń, a następnie podamy przykłady.

Połączenie szeregowe.

Połączenie szeregowo członów dynamicznych przedstawiono na Rys. 5.14.



Rys. 5.14. Połączenie szeregowo członów dynamicznych.

Z definicji bloków mamy

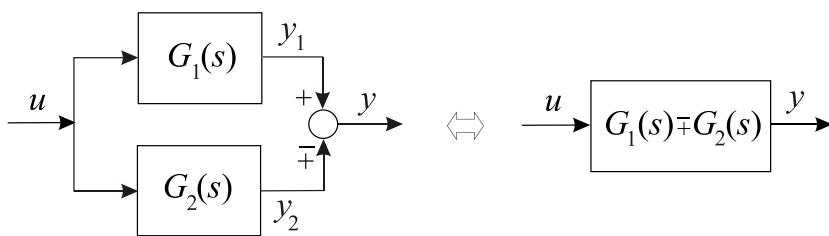
$$\begin{aligned} Y(s) &= G_2(s)Y_1(s) \\ Y_1(s) &= G_1(s)U(s) \end{aligned}$$

skąd bezpośrednio

$$Y(s) = G_2(s)G_1(s) \cdot U(s)$$

Połączenie równoległe.

Połączenie równoległe członów dynamicznych przedstawiono na Rys. 5.15.



Rys. 5.15. Połączenie równoległe członów dynamicznych.

Z definicji bloków i węzła mamy

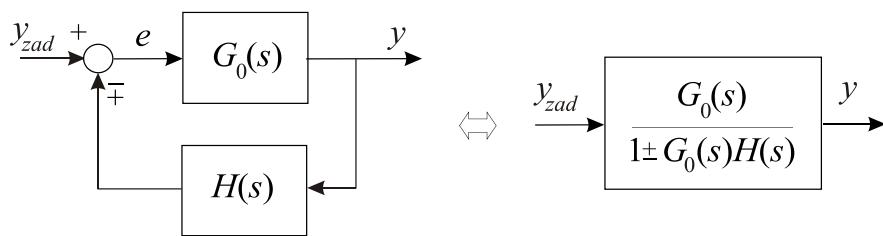
$$\begin{aligned} Y_1(s) &= G_1(s)U(s) \\ Y_2(s) &= G_2(s)U(s) \\ Y(s) &= Y_1(s) \mp Y_2(s) \end{aligned}$$

skąd bezpośrednio

$$Y(s) = [G_1(s) \mp G_2(s)]U(s)$$

Sprzężenie zwrotne.

Na rysunku 5.16 przedstawiono ogólny schemat układu ze sprzężeniem zwrotnym.



Rys. 5.16. Schemat ogólnego układu ze sprzężeniem zwrotnym.

Zapiszemy równania układu:

$$\begin{aligned} Y(s) &= G_0(s)[Y_{zad}(s) \mp H(s)Y(s)] \\ Y(s) \pm G_0(s)H(s)Y(s) &= G_0(s)Y_{zad}(s) \\ [1 \pm G_0(s)H(s)]Y(s) &= G_0(s)Y_{zad}(s) \end{aligned}$$

Skąd dostajemy równanie transmitancji zastępczej układu z członem dynamicznym o transmitancji $H(s)$ w pętli sprzężenia zwrotnego:

$$Y(s) = \frac{G_0(s)}{1 \pm G_0(s)H(s)} Y_{zad}(s)$$

Zwróćmy uwagę na przeciwnie znaki sumowania w węźle sumacyjnym i w mianowniku transmitancji zastępczej (tzn. ujemnemu sprzężeniu zwrotnemu odpowiada plus w mianowniku transmitancji).

Ważnym przypadkiem szczególnym jest *układ z jednostkowym sprzężeniem zwrotnym* (tzn. $H(s) \equiv 1$), wówczas

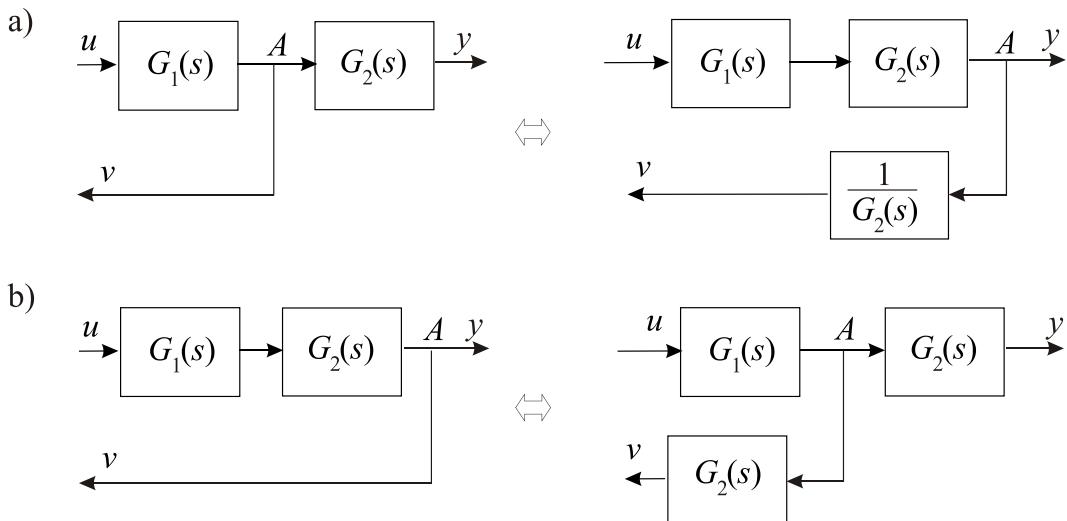
$$Y(s) = \frac{G_0(s)}{1 \pm G_0(s)} Y_{zad}(s)$$

Przesuwanie węzłów.

Przy przekształcaniu schematów blokowych pomocne jest przesuwanie węzłów sumacyjnych i informacyjnych, co może też być traktowane jako przesuwanie bloków przez węzły sumacyjne i informacyjne. Przesuwaniem tym rządzi ogólna reguła mówiąca, że

W schematach początkowym i przekształconym transmisja sygnałów, między punktami początkowymi i końcowymi przekształcanego i przekształconego fragmentu schematu (w tym w pętlach sprzężenia zwrotnego), nie może ulec zmianie.

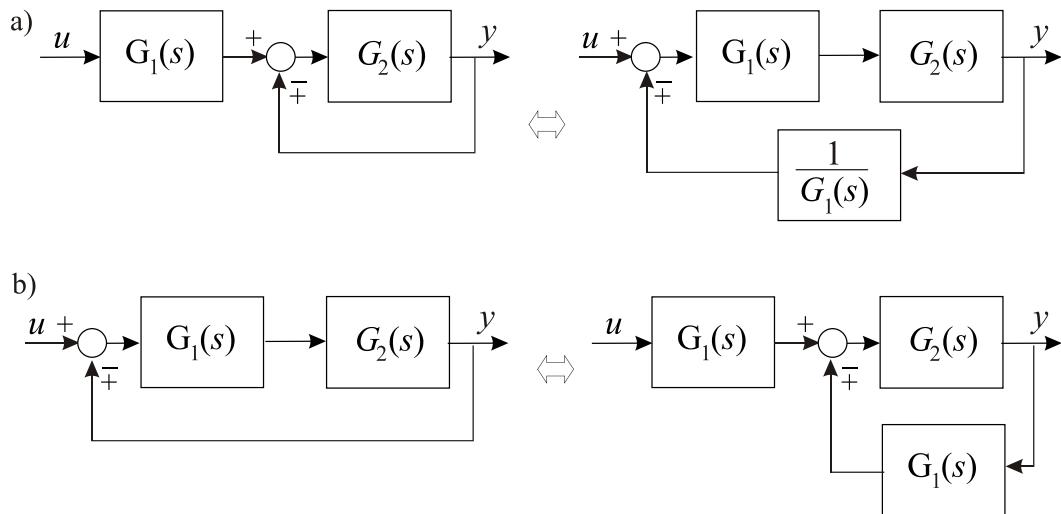
Na rysunku 5.17 zilustrowano przypadek przesuwania węzła informacyjnego, zgodnie z powyższą zasadą: transmitancje między sygnałami y i u oraz v i u są identyczne na schemacie pierwotnym i na każdym ze schematów przekształconych.



Rys. 5.17. Przesuwanie węzła informacyjnego A poza sąsiedni blok, a) w prawo i b) w lewo.

Na Rys. 5.18 zilustrowano przesuwanie węzła sumacyjnego. Zgodnie z ogólną zasadą: transmitancje między sygnałami y i u oraz w pętlach sprzężenia zwrotnego (co może być interpretowane jako „transmitancje między sygnałami y i y' ”) są identyczne na schemacie

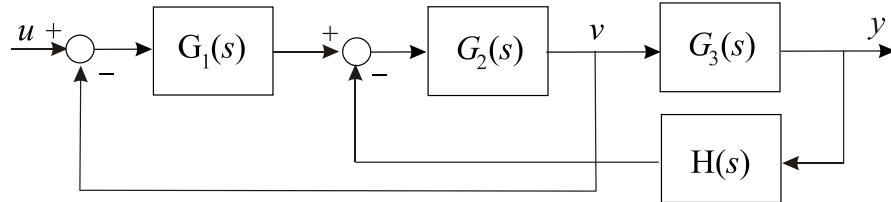
pierwotnym i na każdym ze schematów przekształconych.



Rys. 5.18. Przesuwanie węzła sumacyjnego.

Przykład 5.3.

Należy policzyć transmitancję zastępczą układu przedstawionego na Rys. 5.19.



Rys. 5.19. Struktura układu regulacji w Przykładzie 5.3.

Zadanie można łatwo wykonać przekształcając strukturę układu do prostszej, tak aby liczenie transmitancji zredukować do szeregowego liczenia transmitacji struktur podstawowych (połączenia szeregowe, równoległe i sprzężenie zwrotne). Przekształcenia można dokonać w naszym przypadku na dwa sposoby:

- poprzez przesunięcie wewnętrznego węzła sumacyjnego w lewo, lub
- poprzez przesunięcie węzła informacyjnego odpowiadającego sygnałowi v w prawo.

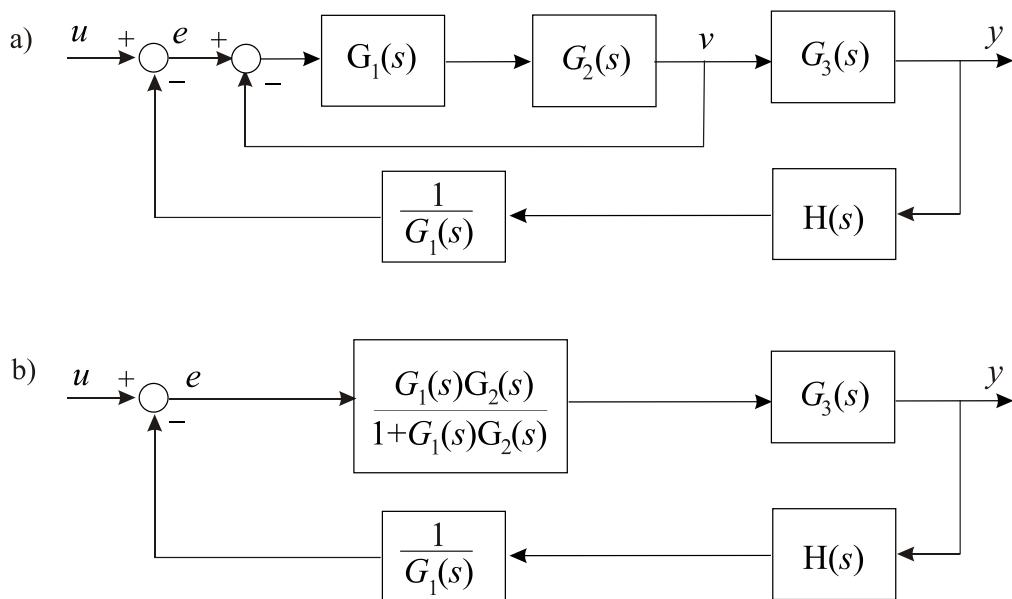
Rysunek 5.20 przedstawia kolejne przekształcenia struktury układu: na Rys. 5.20 a) mamy strukturę uzyskaną po przesunięciu węzła sumacyjnego. Możemy wówczas policzyć transmitancję wewnętrznej pętli sprzężenia zwrotnego:

$$V(s) = \frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)} E(s)$$

Rysunek 5.20 b) przedstawia równoważną strukturę regulacji ze zwiniętą wewnętrzną pętlą sprzężenia zwrotnego. Dla takiej struktury możemy już policzyć transmitancję zastępczą całego układu

$$\begin{aligned} \frac{Y(s)}{U(s)} &= \frac{\frac{G_1(s)G_2(s)}{1+G_1(s)G_2(s)}G_3(s)}{1+\frac{G_1(s)G_2(s)}{1+G_1(s)G_2(s)}G_3(s)\frac{H(s)}{G_1(s)}} = \frac{\frac{G_1(s)G_2(s)}{1+G_1(s)G_2(s)}G_3(s)}{1+\frac{G_2(s)G_3(s)H(s)}{1+G_1(s)G_2(s)}} = \\ &= \frac{G_1(s)G_2(s)G_3(s)}{1+G_1(s)G_2(s)+G_2(s)G_3(s)H(s)} = \frac{G_1(s)G_2(s)G_3(s)}{1+G_2(s)[G_1(s)+G_3(s)H(s)]} \end{aligned}$$

Alternatywne rozwiązywanie zadania poprzez przesunięcie węzła informacyjnego (a nie sumacyjnego) pozostawiamy jako ćwiczenie czytelnikowi.



Rys. 5.20. Kolejne przekształcenia struktury z Rys. 5.19 : a) przesunięcie węzła sumacyjnego, b) zwinięcie wewnętrznej pętli sprzężenia zwrotnego.

5.2.4. Związki biegunków transmitancji z dynamiką układu, warunki stabilności

Dla analizy charakteru przebiegów obiektu dynamicznego, wyboru rodzaju regulatora, istotna jest *wiedza jakościowa*, pozwalająca przyporządkować wartościom biegunków transmitancji charakter przebiegów czasowych obiektu. Takimi typowymi przebiegami są odpowiedź impulsowa i odpowiedź skokowa.

Przy zerowych warunkach początkowych, dla wymuszenia impulsowego ($u(t) = \delta(t)$, tzn. $U(s) = 1$) przebieg wyjścia równy jest odpowiedzi impulsowej, $y(t) = g(t) = \mathcal{L}^{-1}[G(s)](t)$. Czyli

$$Y(s) = \frac{L(s)}{M(s)} = k \frac{\prod_{i=1}^l (s - z_i)}{\prod_{i=1}^m (s - s_i)}$$

gdzie z_i to zera, a s_i biegunki transmitancji. Przyjmując, że wszystkie biegunki są jednokrotne,

z teorii funkcji wymiernych wynika rozkład na ułamki proste

$$Y(s) = \sum_{i=1}^{m_r} \frac{c_i}{s - p_i} + \sum_{i=m_r+1}^{m_r+m_z} \frac{c_i s + c_{i0}}{(s - \alpha_i)^2 + \omega_i^2}$$

gdzie m_r i m_z to liczby biegunów rzeczywistych p_i i par biegunów zespolonych $\alpha_i \pm \omega_i$ ($m_r + 2m_z = m$). Przebiegi czasowe będą stąd następujące

$$\begin{aligned} Y_i(s) &= \frac{c_i}{s - p_i} \Leftrightarrow y_i(t) = c_i e^{p_i t} \\ Y_i(s) &= \frac{c_i s + c_{i0}}{(s - \alpha_i)^2 + \omega_i^2} = \frac{c_i(s - \alpha_i)}{(s - \alpha_i)^2 + \omega_i^2} + \frac{c_{i0} + c_i \alpha_i}{\omega_i} \frac{\omega_i}{(s - \alpha_i)^2 + \omega_i^2} \Leftrightarrow \\ &\Leftrightarrow y_i(t) = c_i e^{\alpha_i t} \cos \omega_i t + \frac{c_{i0} + c_i \alpha_i}{\omega_i} e^{\alpha_i t} \sin \omega_i t \end{aligned}$$

Czyli przebieg czasowy odpowiedzi impulsowej będzie sumą przebiegów odpowiadających biegunom transmitancji (jak wyżej), z odpowiednimi współczynnikami. Dla biegunów wielokrotnych będzie podobnie, tylko przebiegi czasowe odpowiadające biegunom powtarzonym będą jeszcze mnożone przez zmienną czasu t , dla każdego zwielokrotnienia (powielenia) bieguna (zob. tabela transformat), np. dla bieguna rzeczywistego p_i krotności k :

$$\begin{aligned} Y(s) &= \frac{L(s)}{\cdots (s - p_i)^k \cdots} = \cdots + \frac{c_{i1}}{s - p_i} + \frac{c_{i2}}{(s - p_i)^2} + \cdots + \frac{c_{ik}}{(s - p_i)^k} \Leftrightarrow \\ &\Leftrightarrow y(t) = \cdots + c_{i1} e^{p_i t} + c_{i2} t e^{p_i t} + \cdots + \frac{c_{ik}}{(k-1)!} t^{k-1} e^{p_i t} + \cdots \end{aligned}$$

Ponieważ zmienność wykładnicza jest szybsza od potęgowej (tzn. e^{pt} maleje, dla $p < 0$, szybciej niż rośnie t^k , dla dowolnego k), to z powyższych rozważań wynika bezpośrednio:

Warunkiem koniecznym i dostatecznym stabilności asymptotycznej układu liniowego (autonomicznego) jest, aby części rzeczywiste wszystkich biegunów transmitancji były ujemne (bieguny leżały w lewej otwartej płaszczyźnie zmiennej zespolonej) – tylko wówczas bowiem funkcje wykładnicze, dominujące przebiegi czasowe w każdym ułamku prostym, będą malejące. Wykaźaliśmy więc:

Twierdzenie 5.3. *Jeśli wszystkie pierwiastki s_i mianownika transmitancji mają ujemne części rzeczywiste, to układ jest asymptotycznie stabilny (w sensie Lapunowa).*

Przypomnijmy, że liniowy układ dynamiczny SISO rzędu n może być opisany

- układem n równań pierwszego rzędu w postaci normalnej (5.28), lub równoważnie
- jednym równaniem różniczkowym (5.33) rzędu $m = n$.

Stąd, mianownik transmitancji $M(s)$ jest, w pierwszym przypadku, wielomianem $\det[s\mathbf{I} - \mathbf{A}]$; natomiast w drugim przypadku, wielomianem charakterystycznym równania różniczkowego (5.33) – oczywiście jest to ten sam wielomian (modelujemy ten sam układ). Stąd mamy następujące twierdzenie:

Twierdzenie 5.4. *Jeśli opis układu dynamicznego dany jest w postaci równań stanu*

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t), \quad \mathbf{x}(t) \in R^n \tag{5.38}$$

to układ ten jest stabilny asymptotycznie wtedy i tylko wtedy, gdy wszystkie wartości własne macierzy \mathbf{A} , tj. pierwiastki równania charakterystycznego $\det[s\mathbf{I}-\mathbf{A}] = 0$, mają ujemne części rzeczywiste.

W przypadku rozważanego uprzednio obiektu jednoinercyjnego, przy zerowym sterowaniu mamy równanie różniczkowe $\dot{y} + \frac{1}{T}y = 0$ oraz jego równanie charakterystyczne $s + \frac{1}{T} = 0$ (które jest równoważne równaniu $\det[s\mathbf{I}-\mathbf{A}] = 0$), stąd biegun $s_1 = -\frac{1}{T} < 0$ jeśli $T > 0$; układ jest więc stabilny asymptotycznie, co mogliśmy już uprzednio stwierdzić bezpośrednio.

Jeśli układ (5.38) jest stabilny, lecz nie jest asymptotycznie stabilny, to wszystkie wartości własne macierzy \mathbf{A} muszą mieć nieujemne części rzeczywiste; w drugą stronę stwierdzenie nie zawsze jest prawdziwe, tj. w przypadkach wielokrotnych wartości własnych mogą one mieć ujemne części rzeczywiste, a mimo to układ będzie niestabilny.

5.2.5. Algebraiczne kryterium stabilności Hurwitz'a

Do badania stabilności układu, bez potrzeby bezpośredniego obliczania pierwiastków równania charakterystycznego służą algebraiczne kryteria stabilności. Przedstawimy w tym miejscu najbardziej znane *kryterium Hurwitz'a*. Jest to kryterium algebraiczne określające warunki, jakie muszą spełniać współczynniki wielomianu charakterystycznego, aby jego wszystkie pierwiastki miały ujemne części rzeczywiste. Niech transmitancja ma postać:

$$G(s) = \frac{L(s)}{M(s)} = \frac{b_ls^l + b_{l-1}s^{l-1} + \dots + b_1s + b_0}{a_ms^m + a_{m-1}s^{m-1} + \dots + a_1s + a_0},$$

gdzie $l < m$.

Kryterium Hurwitz'a.

Przyjmijmy, że współczynnik a_m jest dodatni; najczęściej przyjmujemy $a_m=1$. Wówczas wszystkie pierwiastki wielomianu $M(s)$ mają ujemne części rzeczywiste wtedy i tylko wtedy, gdy:

1. wszystkie pozostałe współczynniki wielomianu są dodatnie, tj.

$$a_i > 0 \text{ dla } i = 0, \dots, m-1$$

2. wszystkie minory główne Δ_i , $i = 0, \dots, m-1$, macierzy Hurwitz'a \mathbf{H} są dodatnie, gdzie

$$\mathbf{H} = \begin{bmatrix} a_{m-1} & a_m & 0 & 0 & \cdots & 0 \\ a_{m-3} & a_{m-2} & a_{m-1} & a_m & \cdots & 0 \\ a_{m-5} & a_{m-4} & a_{m-3} & a_{m-2} & \cdots & 0 \\ a_{m-7} & a_{m-6} & a_{m-5} & a_{m-4} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & a_0 \end{bmatrix}_{m \times m}, \quad (5.39)$$

$$\Delta_{m-1} = a_{m-1}, \quad \Delta_{m-2} = \begin{vmatrix} a_{m-1} & a_m \\ a_{m-3} & a_{m-2} \end{vmatrix}, \quad \Delta_{m-3} = \begin{vmatrix} a_{m-1} & a_m & 0 \\ a_{m-3} & a_{m-2} & a_{m-1} \\ a_{m-5} & a_{m-4} & a_{m-3} \end{vmatrix}, \dots, \Delta_0 = \det \mathbf{H}.$$

Zauważmy, że $\Delta_0 = a_0\Delta_1$ i stąd nie trzeba osobno sprawdzać warunku $\Delta_0 > 0$.

Rozważymy warunki stabilności dla układów drugiego i trzeciego rzędu.

Układ 2 rzędu: Wielomian $M(s) = a_2 s^2 + a_1 s + a_0$, $m = 2$, warunki asymptotycznej stabilności są następujące:

$$1. \ a_1 > 0, \ a_0 > 0,$$

$$2. \ \mathbf{H} = \begin{bmatrix} a_1 & a_2 \\ 0 & a_0 \end{bmatrix} \quad \det \mathbf{H} = a_1 a_0 > 0 \Rightarrow a_0 > 0, a_1 > 0 \text{ lub } a_0 < 0, a_1 < 0.$$

Stąd układ jest asymptotycznie stabilny wtedy i tylko wtedy, gdy $a_2 > 0$, $a_1 > 0$ oraz $a_0 > 0$.

Układ 3 rzędu: Wielomian: $M(s) = a_3 s^3 + a_2 s^2 + a_1 s + a_0$, układ rzędu 3, warunki asymptotycznej stabilności są następujące:

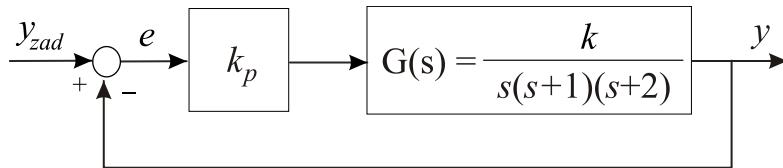
$$1. \ a_3 > 0, \ a_2 > 0, \ a_1 > 0, \ a_0 > 0,$$

$$2. \ \mathbf{H} = \begin{bmatrix} a_2 & a_3 & 0 \\ a_0 & a_1 & a_2 \\ 0 & 0 & a_0 \end{bmatrix}, \quad \text{stąd} \quad \begin{aligned} \Delta_2 &= a_2 > 0 \\ \Delta_1 &= a_2 a_1 - a_3 a_0 > 0 \\ \Delta_0 &= a_0 \cdot \Delta_1 \end{aligned}$$

Warunki asymptotycznej stabilności są teraz następujące: $a_3 > 0$, $a_2 > 0$, $a_1 > 0$, $a_0 > 0$ oraz $a_2 a_1 > a_3 a_0$.

Przykład 5.4.

Kryterium Hurwitza i wynikające z niego wymagania odnośnie współczynników mianownika transmitancji (wielomianu charakterystycznego układu) wykorzystamy do analizy stabilności układu regulacji pokazanego na rysunku, z obiektem trzeciego rzędu z całkowaniem i podwójną inercją oraz prostym regulatorem proporcjonalnym o wzmacnieniu k_p .



Rys. 5.21. Układ regulacji w przykładzie 5.4.

Transmitancję układu zamkniętego (od wejścia w postaci wartości zadanej $y_{zad}(t)$ do wyjścia $y(t)$) możemy łatwo wyznaczyć, mianowicie:

$$Y(s) = k_p G(s) e(s) = k_p G(s) (Y_{zad}(s) - Y(s)),$$

$$Y(s) = \frac{k_p G(s)}{1 + k_p G(s)} \cdot Y_{zad}(s),$$

Stąd transmitancja układu zamkniętego:

$$G_Z(s) = \frac{k_p G(s)}{1 + k_p G(s)} = \frac{\frac{kk_p}{s(s+1)(s+2)}}{1 + \frac{kk_p}{s(s+1)(s+2)}} = \frac{kk_p}{s^3 + 3s^2 + 2s + kk_p},$$

$$M_Z(s) = s^3 + 3s^2 + 2s + kk_p.$$

Korzystamy z kryterium Hurwitza; określamy po kolejnym wymagania (jest to układ 3. rzędu):

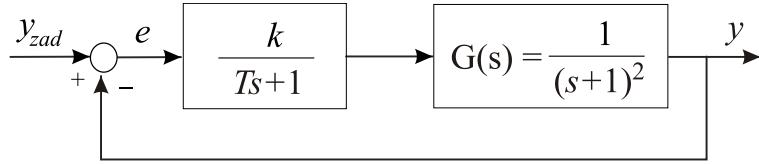
$$\begin{aligned} a_2 > 0, \quad a_1 > 0, \quad a_0 > 0 \Rightarrow kk_p > 0, \\ a_2 a_1 > a_0 \Rightarrow 3 \cdot 2 > kk_p \end{aligned}$$

Stąd poszukiwany warunek asymptotycznej stabilności układu zamkniętego: $0 < kk_p < 6$.

Zastosowanie kryterium Hurwitza nie pozwala bezpośrednio ocenić jakości stabilności; w szczególności nie pozwala, w ogólnym przypadku, wyznaczyć tzw. zapasów stabilności. Jednak w sytuacji takiej jak w powyższym przykładzie możemy łatwo określić wartość wzmacnienia k_p , oznaczmy ją jako k_{p6dB} , która dopiero w razie jej podwojenia doprowadzi do utraty stabilności. Ta wartość, zapewniająca tzw. zapas modułu (pojęcie to zostanie wyjaśnione później) równy 6 dB wynosi w powyższym przykładzie $k_{p6dB} = 3/k$. \square

Przykład 5.5.

Dla układu regulacji



Rys. 5.22. Układ regulacji w przykładzie 5.5.

należy określić zakres zmian parametrów k i T ($k>0$, $T>0$) gwarantujących stabilność układu. Postępując podobnie jak w przykładzie poprzednim mamy

$$\begin{aligned} G_Z(s) &= \frac{\frac{k}{Ts+1}G(s)}{1 + \frac{k}{Ts+1}G(s)} = \frac{\frac{k}{(Ts+1)(s+1)^2}}{1 + \frac{k}{(Ts+1)(s+1)^2}} = \\ &= \frac{k}{(Ts+1)(s+1)^2 + k} = \frac{k}{Ts^3 + (2T+1)s^2 + (T+2)s + k+1} \end{aligned}$$

Układ jest trzeciego rzędu, wykorzystując kryterium Hurwitza dostajemy warunek stabilności (wszystkie współczynniki są z założenia dodatnie)

$$a_2 a_1 - a_3 a_0 > 0 \Rightarrow (2T+1)(T+2) - T(k+1) > 0 \Rightarrow 2T^2 + 5T + 2 > T(k+1)$$

który można zapisać w postaci $k < 2T + 4 + \frac{2}{T}$. Zostawiamy czytelnikowi graficzne zilustrowanie obszaru stabilności na płaszczyźnie (T,k) . \square

Zadania

Zadanie 5.1. Układ dynamiczny opisany jest następującymi równaniami stanu

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) + u(t) \\ \dot{x}_2(t) &= -3x_1(t) - 4x_2(t), \quad y(t) = x_2(t) \end{aligned}$$

Zbadaj stabilność układu, wyznacz odpowiedź skokową.

Zadanie 5.2. Znajdź położenie i zbadaj stabilność punktu (punktów) równowagi następującego układu dynamicznego, dla $u(t)=0$,

$$\begin{aligned}\dot{x}_1(t) &= (x_1(t))^2 x_2(t) - \frac{\pi^2}{4} x_2(t) \sin(x_1(t)) \\ \dot{x}_2(t) &= x_1(t) - x_1(t)x_2(t) - 3x_2(t) + 3 + u(t)\end{aligned}$$

Wskazówka: wykorzystać równanie charakterystyczne układu liniowego $\det(\lambda I - A) = 0$.

Zadanie 5.3. Wyznacz położenie punktów równowagi następującego układu dynamicznego, dla $u(t)=u_0>0$

$$\begin{aligned}\dot{x}_1(t) &= -x_1(t)x_2(t) + u(t) \\ \dot{x}_2(t) &= x_1(t) - 2x_1(t)x_2(t), \quad y(t) = x_1(t)\end{aligned}$$

Zlinearyzuj układ wokół wybranego punktu równowagi, zbadaj stabilność układu zlinearyzowanego i wyznacz transmitancję tego układu. Czy na podstawie samej transmitancji można w tym przypadku wnioskować o stabilności układu?

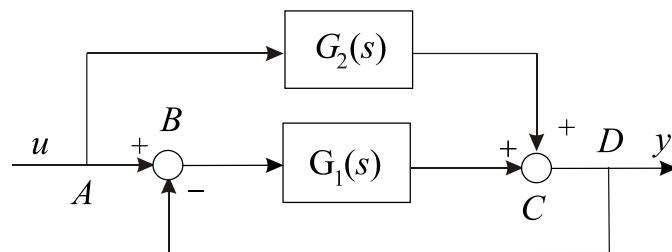
Zadanie 5.4. Układ dynamiczny opisany jest następującymi równaniami stanu

$$\begin{aligned}\dot{x}_1(t) &= x_1(t)^2 - x_1(t)x_2(t) \\ \dot{x}_2(t) &= 0.5x_1(t)x_2(t) - 4x_2(t) + u(t), \quad y(t) = x_2(t) - x_1(t)\end{aligned}$$

Wyznacz punkty równowagi dla $u(t) = 6$. Wyznacz modele zlinearyzowane w punktach równowagi, określ ich stabilność. Dla stabilnych punktów równowagi wyznacz odpowiedzi: impulsową, skokową i prędkościową (tzn. na wymuszenia: impulsowe, skokowe i liniowo rosnące (prędkościowe $u(t) = t\mathbf{1}(t)$)).

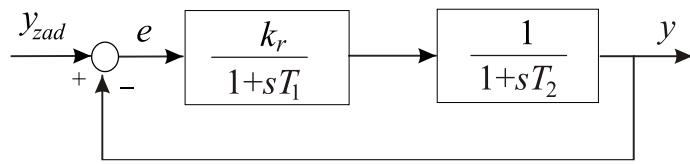
Zadanie 5.5. Wyznaczyć transmitancję zastępczą układu regulacji z Rys. 5.19 metodą przesunięcia węzła informacyjnego odpowiadającego zmiennej v .

Zadanie 5.6. Wyznaczyć transmitancję zastępczą układu regulacji jak na rysunku:



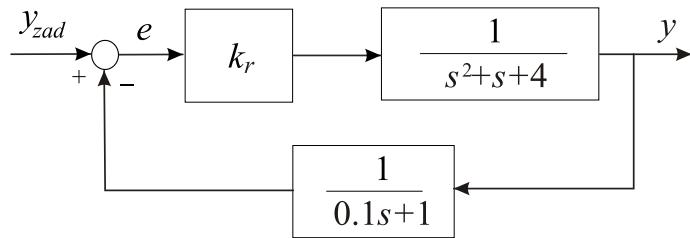
metodą przekształcenia schematu do prostszego umożliwiającego kolejne liczenie transmitancji schematów elementarnych.

Zadanie 5.7. Dla układu regulacji jak na rysunku:



wyznacz relację między wzmocnieniem k_r , a stałymi czasowymi T_1 i T_2 zapewniającą stabilne i nieoscylacyjne przebiegi odpowiedzi impulsowej i skokowej.

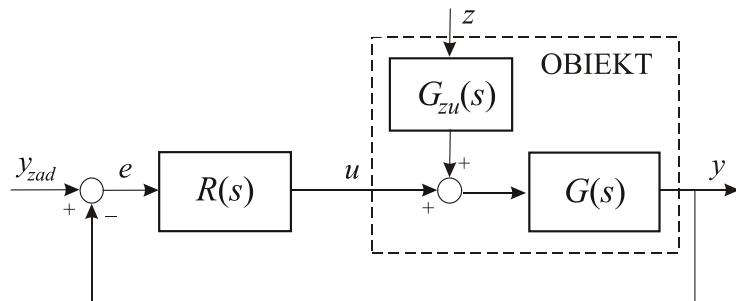
Zadanie 5.8. Dla układu regulacji jak na rysunku



wyznacz zakres wartości współczynnika wzmocnienia k_r zapewniający stabilność układu.

6. Uchyby ustalone w układach regulacji

Rozważmy układ regulacji przedstawiony na Rys. 6.1, gdzie wpływ zakłócenia jest sprowadzony do wejścia toru głównego (toru sterowania) obiektu – układy o innym sposobie oddziaływania zakłócenia przedstawimy w dalszej części rozdziału.



Rys.6.1. Struktura układu regulacji z zakłóceniem sprowadzonym do wejścia obiektu.

Mamy tu trzy transmitancje:

$$R(s) = \frac{U(s)}{E(s)} \quad \text{– transmitancja regulatora}$$

$$G(s) = \frac{Y(s)}{U(s)}_{z=0} \quad \text{– transmitancja toru sterowania obiektu}$$

$G_{zu}(s)$ – transmitancja toru zakłócenia sprowadzonego do wejścia obiektu

Przykład 6.1.

Model prędkości samochodu:

$$\frac{dy(t)}{dt} + \frac{b}{M} y(t) = \frac{c}{M} u(t) - g z(t)$$

$$Y(s)(s + \frac{b}{M}) = \frac{c}{M} U(s) - g Z(s)$$

$$\begin{aligned} Y(s) &= \frac{\frac{c}{M}}{s + \frac{b}{M}} \left(U(s) - \frac{Mg}{c} Z(s) \right) = \frac{\frac{c}{b}}{1 + s \frac{M}{b}} \left(U(s) - \frac{Mg}{c} Z(s) \right) = \\ &= \frac{k}{1 + sT} (U(s) - k_{zu} Z(s)) = G(s) (U(s) + G_{zu}(s) Z(s)) \end{aligned} \quad \square$$

Wyznaczamy transmitancje główną i uchybową układu regulacji z Rys. 6.1:

$$Y(s) = G(s) R(s) E(s) + G(s) G_{zu}(s) Z(s),$$

$$E(s) = Y_{zad}(s) - Y(s)$$

skąd

$$(1+G(s)R(s))Y(s)=G(s)R(s)Y_{zad}(s)+G(s)G_{zu}(s)Z(s)$$

$$Y(s)=\frac{G(s)R(s)}{1+G(s)R(s)}Y_{zad}(s)+\frac{G(s)G_{zu}(s)}{1+G(s)R(s)}Z(s) \quad (6.1)$$

Stąd można uzyskać bezpośrednio transmitancje dla uchybu $E(s)$

$$E(s)=\frac{1}{1+G(s)R(s)}Y_{zad}(s)-\frac{G(s)G_{zu}(s)}{1+G(s)R(s)}Z(s) \quad (6.2)$$

Transmitancję toru otwartego sterowania oznaczać będziemy dalej przez $G_0(s)$,

$$G_0(s)=G(s)R(s)$$

6.1. Obiekty i regulatory statyczne (bez całkowania)

Rozważymy na początku przypadek, kiedy transmitancje obiektu i regulatora nie zawierają całkowania, tzn. czynnika $1/s$. W takiej sytuacji:

$$\lim_{s \rightarrow 0} G(s) = G(0) = k < \infty, \quad \text{np. } G(s) = \frac{k}{1+sT}, \quad G(s) = \frac{k}{(1+sT_1)(1+sT_2)};$$

$$\lim_{s \rightarrow 0} G_{zu}(s) = G_{zu}(0) = k_{zu} < \infty,$$

$$\lim_{s \rightarrow 0} R(s) = R(0) = k_r < \infty, \quad \text{np. } R(s) = k_r, \quad R(s) = \frac{k_r(1+sT_1)}{1+sT_2}$$

6.1.1. Uchyb po skoku wartości zadanej, bez zakłócenia

W tej sytuacji mamy:

$$y_{zad}(t) = a \cdot \mathbf{1}(t),$$

$$z(t) \equiv 0$$

Stąd

$$Y(s) = \frac{G(s)R(s)}{1+G(s)R(s)} \cdot \frac{a}{s},$$

$$E(s) = \frac{1}{1+G(s)R(s)} \cdot \frac{a}{s}$$

Wykorzystując teraz twierdzenie o granicy oryginału przy $t \rightarrow \infty$ oraz zal. (6.1) mamy

$$y_{ust} = \lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} [sY(s)] = \lim_{s \rightarrow 0} s \frac{G(s)R(s)}{1+G(s)R(s)} \cdot \frac{a}{s} =$$

$$= \lim_{s \rightarrow 0} \frac{G(s)R(s)a}{1+G(s)R(s)} = \frac{kk_r \cdot a}{1+kk_r} \quad (6.3)$$

W przypadku idealnej realizacji wartości zadanej w stanie ustalonym byłoby $y_{ust} = a$, jednak w rozważanej sytuacji występuje zawsze pewien uchyb ustalony e_{ust} ,

$$e_{ust} = y_{zad} - y_{ust} = a - \frac{kk_r \cdot a}{1 + kk_r} = \frac{a}{1 + kk_r} \quad (6.4)$$

(zależność tę można też uzyskać bezpośrednio z równ.(6.2) dla transmitancji uchybowej). Tak więc dla każdej wartości $k_r > 0$ uchyb ustalony jest niezerowy, ale dąży on do zera przy $k \rightarrow \infty$.

6.1.2. Uchyb po skoku zakłócenia, przy zerowej wartości zadanej

W tej sytuacji mamy:

$$\begin{aligned} z(t) &= b \cdot \mathbf{1}(t), \\ y_{zad}(t) &\equiv 0 \end{aligned}$$

Korzystając tym razem bezpośrednio z transmitancji uchybowej (6.2) dostajemy

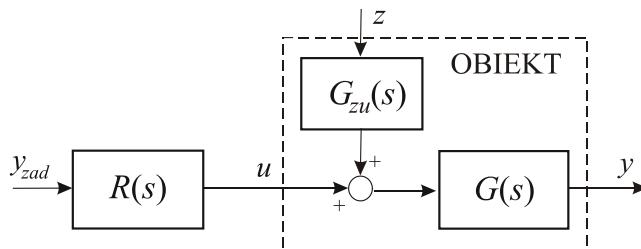
$$\begin{aligned} e_{ust} &= \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} [sE(s)] = \lim_{s \rightarrow 0} [s \frac{-G(s)G_{zu}(s)}{1 + G(s)R(s)} \cdot \frac{b}{s}] = \\ &= \lim_{s \rightarrow 0} \frac{-G(s)G_{zu}(s)b}{1 + G(s)R(s)} = \frac{-kk_{zu} \cdot b}{1 + kk_r} \end{aligned} \quad (6.5)$$

Dla porównania, w otwartym układzie regulacji (tzn. bez sprzężenia zwrotnego, jak to przedstawiono na Rys. 6.2, mielibyśmy

$$y_{ust} = \lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} [sY(s)] = \lim_{s \rightarrow 0} [sG(s)R(s) \cdot 0 + sG(s)G_{zu}(s) \cdot \frac{b}{s}] = kk_{zu}b \quad (6.6)$$

czyli

$$e_{ust} = y_{zad} - y_{ust} = -kk_{zu} \cdot b$$



Rys. 6.2. Otwarty układ sterowania.

Wnioski.

W statycznym układzie regulacji (tzn. bez całkowania w obiekcie i regulatorze):

- występuje uchyb ustalony nadążania za ustaloną (po skokowej zmianie) wartością zadaną, tym mniejszy im większe jest wzmacnienie regulatora k_r (dążąc do zera przy $k \rightarrow \infty$),
- wpływ oddziaływanego na wejście obiektu zakłócenia na wartość wielkości wyjściowej obiektu jest tłumiony $1 + kk_r$ razy, czyli tym silniej im większe jest wzmacnienie regulatora k_r .

6.2. Obiekty i regulatory z całkowaniem

W przypadku ogólnym, transmitancja zarówno regulatora $R(s)$ jak i obiektu $G(s)$ mogą zawierać biegunki w zerze, tzn. całkowania. Oznaczmy

$$\begin{aligned} R(s) &= \frac{1}{s^r} \cdot \tilde{R}(s), \quad \lim_{s \rightarrow 0} \tilde{R}(s) = \tilde{R}(0) = k_r \neq 0, \\ G(s) &= \frac{1}{s^l} \cdot \tilde{G}(s), \quad \lim_{s \rightarrow 0} \tilde{G}(s) = \tilde{G}(0) = k \neq 0 \end{aligned} \quad (6.7)$$

Na przykład

$$R(s) = \frac{1}{s^r} \cdot k_r, \quad R(s) = \frac{1}{s} \cdot \frac{k_r(1+T_1 s)}{(1+T_2 s)},$$

$$G(s) = \frac{1}{s^l} \cdot \frac{1}{(1+Ts)}, \quad G(s) = \frac{1}{s^2} \cdot \frac{(1+T_1 s)^2}{(1+T_2 s)}$$

Rozważmy teraz następujące wymuszenia standardowe (testowe), o amplitudzie a :

$$\begin{aligned} y_{zad}(t) &= a \cdot \mathbf{1}(t) \quad \Rightarrow \quad Y_{zad}(s) = \frac{a}{s}, \\ y_{zad}(t) &= at \cdot \mathbf{1}(t) \quad \Rightarrow \quad Y_{zad}(s) = \frac{a}{s^2}, \\ y_{zad}(t) &= at^2 \cdot \mathbf{1}(t) \quad \Rightarrow \quad Y_{zad}(s) = \frac{2a}{s^3} \end{aligned} \quad (6.8)$$

będące najważniejszymi przypadkami szczególnymi (dla $n = 0, 1, 2$) wymuszenia

$$y_{zad}(t) = at^n \cdot \mathbf{1}(t) \quad \Rightarrow \quad Y_{zad}(s) = a \frac{n!}{s^{n+1}} \quad (6.9)$$

6.2.1. Uchyby po zmianach wartości zadanej, bez zakłócenia

Zakładając $z(t)=0$, mamy z zal. (6.2)

$$E(s) = \frac{1}{1 + G(s)R(s)} Y_{zad}(s)$$

skąd, przyjmując $0^0 = 1$, dostajemy dla wymuszeń postaci (6.9)

$$\begin{aligned} e_{ust} &= \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} \left[s \frac{1}{1 + G(s)R(s)} \cdot \frac{an!}{s^{n+1}} \right] = \lim_{s \rightarrow 0} \left[\frac{1}{1 + \frac{1}{s^{l+r}} \tilde{G}(s)\tilde{R}(s)} \cdot \frac{an!}{s^n} \right] = \\ &= \lim_{s \rightarrow 0} \frac{s^{l+r-n} \cdot an!}{s^{l+r} + \tilde{G}(s)\tilde{R}(s)} = \begin{cases} 0 & \text{dla } l+r > n \\ a \frac{n!}{0^{l+r} + kk_r} & \text{dla } l+r = n \\ \infty & \text{dla } l+r < n \end{cases} \end{aligned} \quad (6.10)$$

Najważniejsze rezultaty podane powyższym wzorem ilustruje Tabela 6.1, gdzie wyszczególniono uchyby ustalone dla wartości $n = 0$ (wymuszenie skokowe) i $n = 1$ (wymuszenie liniowo zmienne w czasie, tzw. wymuszenie prędkościowe) i dla trzech najistotniejszych wartości sumy $l+r$, a mianowicie $l+r = 0, 1$ i 2 .

Tabela 6.1. Uchyby ustalone względem $y_{zad}(t)$ w układzie ze sprzężeniem zwrotnym, dla wymuszeń skokowego i prędkociowego i dla różnej liczby całkowań w torze sterowania

Wymuszenie \ l+r	$l+r = 0$	$l+r = 1$	$l+r = 2$
$n=0, \ y_{zad}(t) = a\mathbf{1}(t)$	$e_{ust} = \frac{a}{1+kk_r}$	$e_{ust} = 0$	$e_{ust} = 0$
$n=1, \ y_{zad}(t) = at\mathbf{1}(t)$	$e_{ust} \rightarrow \infty$	$e_{ust} = \frac{a}{kk_r}$	$e_{ust} = 0$
\vdots	\vdots	\vdots	\vdots
Układ regulacji <u>względem wartości zadanej $y_{zad}(t)$:</u>	statyczny	astatyczny 1 rzędu (klasy 1)	astatyczny 2 rzędu (klasy 2)

Wniosek.

O cechach ustalonego uchybu nadążania za zmianami wartości zadanej $y_{zad}(t)$ decyduje liczba całkowań (tj. biegunów w zerze) transmitancji otwartego toru sterowania, tzn. *suma całkowań w transmitancji regulatora i obiektu* – nie jest istotne, czy całkowania występują w transmitancji regulatora, czy obiektu.

Stąd np. chcąc zerowy uchyb statyczny nadążania po skokowej zmianie wartości zadanej dla *obiektu statycznego* (tzn. $l = 0$), trzeba i wystarczy zapewnić *jedno całkowanie w regulatorze*.

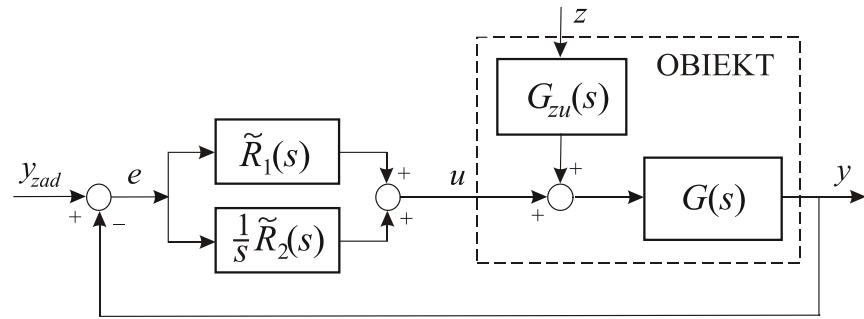
Uwaga. Nie jest konieczne, aby całkowanie występowało szeregowo przed całą pozostałą częścią transmitancji regulatora, tzn. aby $R(s) = \frac{1}{s^r} \cdot \tilde{R}(s)$. Jeśli transmitancja regulatora jest o strukturze równoległej, to wystarczy, aby całkowanie wystąpiło w jednej gałęzi. W najprostszym przypadku struktury składającej się z dwóch gałęzi równoległych, wystarczy, aby zachodziło

$$R(s) = \tilde{R}_1(s) + \frac{1}{s} \cdot \tilde{R}_2(s) \quad (6.11)$$

gdyż

$$\begin{aligned} e_{ust} &= \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} [s \frac{1}{1 + G(s)[\tilde{R}_1(s) + \frac{1}{s} \cdot \tilde{R}_2(s)]}] \cdot \frac{1}{s} = \\ &= \lim_{s \rightarrow 0} \frac{s}{s + G(s)\tilde{R}_1(s)s + G(s)\tilde{R}_2(s)} = 0 \end{aligned}$$

Analizowana sytuacja zilustrowana jest na Rys. 6.3. Intuicyjnie, w strukturze przedstawionej na tym rysunku w stanie ustalonym musi zachodzić $e(t) = e_{ust} = 0$, gdyż w przeciwnym razie dolny człon regulatora nadal całkowałby jakikolwiek niezerowy uchyb, aż do jego wyzerowania (zakładając stabilność asymptotyczną układu regulacji).



Rys. 6.3. Układ regulacji z regulatorem o strukturze równoległej z pojedynczym całkowaniem.

Przyjmując

$$\tilde{R}_1(s) = k_p, \quad \tilde{R}_2(s) = \frac{k_p}{T_i s}$$

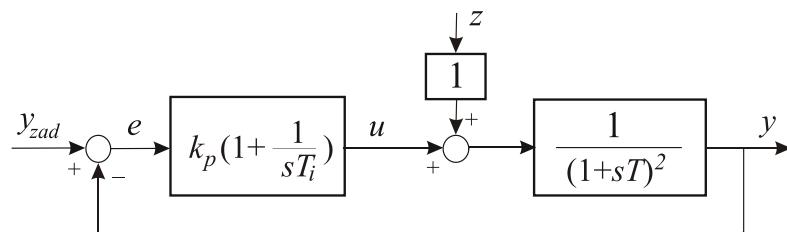
otrzymujemy transmitancję **regulatora PI** (*Proportional - Integral*):

$$R_{PI}(s) = k_p + \frac{k_p}{T_i s} = k_p \left(1 + \frac{1}{T_i s}\right) = \frac{k_p (1 + T_i s)}{T_i s}$$

Regulator PI to jeden z najpowszechniej stosowanych regulatorów przemysłowych, będący najważniejszym przypadkiem szczególnym regulatora PID.

Przykład 6.2.W układzie regulacji PI przedstawionym na Rys. 6.4, gdzie stała czasowa transmitancji obiektu wynosi $T = 0.4$, należy wyznaczyć zakres nastaw regulatora, tzn. wartości wzmacnienia k_p i czasu całkowania T_i , zapewniające:

1. Stabilność układu,

2. Uchyb ustalony nadążania za wartością zadaną $y_{zad}(t) = at \cdot \mathbf{1}(t)$ nie większy niż 5%.

Rys. 6.4. Układ regulacji w przykładzie 6.2.

ad 1. Stabilność. Transmitancja toru otwartego sterowania jest równa

$$G_0(s) = k_p \frac{sT_i + 1}{sT_i} \cdot \frac{1}{(1 + sT)^2}$$

Natomiast transmitancja zamkniętego układu regulacji

$$G_f(s) = \frac{G_0(s)}{1+G_0(s)} = \frac{k_p(1+sT_i)}{sT_i(1+sT)^2 + k_p(1+sT)} = \frac{L_f(s)}{M_f(s)}$$

skąd

$$M_f(s) = s^3 T_i T^2 + s^2 2T_i T + sT_i(1+k_p) + k_p$$

Zakładając $k_p > 0$ i $T_i > 0$ i korzystając z Kryterium stabilności Hurwitza dostajemy warunek stabilności układu regulacji w postaci

$$\begin{vmatrix} 2T_i T & T_i T^2 \\ k_p & T_i(1+k_p) \end{vmatrix} > 0$$

Stąd wynika

$$\begin{aligned} 2T_i T^2 (1+k_p) &> k_p T_i T^2 \\ 2T_i (1+k_p) &> k_p T \\ T_i &> \frac{T k_p}{2(1+k_p)} \end{aligned}$$

Podstawiając teraz wartość $T=0.4$ dostajemy warunek stabilności

$$T_i > \frac{0.2k_p}{(1+k_p)}$$

ad 2. Uchyb nadążania. Regulator jest z pojedynczym całkowaniem ($r = 1$), obiekt jest statyczny ($l = 0$), stąd układ regulacji jest astatyczny rzędu 1 (względem wartości zadanej). Korzystając z Tabeli 6.1 mamy dla wymuszenia prędkościowego ($n = 1$)

$$e_{ust} = \frac{a}{kk_r} = \frac{a}{1 \cdot \frac{k_p}{T_i}} = \frac{aT_i}{k_p}$$

ponieważ wzmacnienie k_r części statycznej $\tilde{R}(s)$ regulatora PI równe k_p/T_i :

$$R(s) = k_p \frac{sT_i + 1}{sT_i} = \frac{1}{s} \cdot \frac{k_p(sT_i + 1)}{T_i} = \frac{1}{s} \tilde{R}(s), \quad \frac{k_p(sT_i + 1)}{T_i} \xrightarrow[s \rightarrow 0]{} \frac{k_p}{T_i} = k_r$$

Uchyb ustalony powinien być nie większy niż 5%, stąd warunek na jakość regulacji

$$\frac{aT_i}{k_p} \leq 0.05a \quad \Rightarrow \quad T_i \leq 0.05k_p$$

□

6.2.2. Uchyby po zmianach zakłócenia, przy zerowej wartości zadanej

Rozważmy teraz uchyby powodowane wpływem zakłóceń oddziałyujących na wejście obiektu, tak jak to ilustruje Rys. 6.1. Rozważmy sytuację standardowych wymuszeń, tzn.

$$z(t) = bt^n \cdot \mathbf{1}(t),$$

$$y_{zad}(t) \equiv 0$$

gdzie b jest amplitudą wymuszenia. Będziemy nadal zakładać, że w transmitancji $G_{zu}(s)$ reprezentującej wpływ zakłócenia na wejście toru sterowania nie ma całkowania, tzn.

$$\lim_{s \rightarrow 0} G_{zu}(s) = G_{zu}(0) = k_{zu} < \infty$$

Uchyb ustalony w tym przypadku będzie następujący:

$$\begin{aligned}
 e_{ust} &= \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} [sE(s)] = \\
 &= \lim_{s \rightarrow 0} \left[s \frac{-G(s)G_{zu}(s)}{1 + G(s)R(s)} \cdot \frac{bn!}{s^{n+1}} \right] = \\
 &= \lim_{s \rightarrow 0} \left[\frac{-\frac{1}{s^l} \tilde{G}(s)G_{zu}(s)}{1 + \frac{1}{s^{l+r}} \tilde{G}(s)\tilde{R}(s)} \cdot \frac{bn!}{s^n} \right] = \\
 &= \lim_{s \rightarrow 0} \left[\frac{-s^{r-n} \tilde{G}(s)G_{zu}(s)}{s^{l+r} + \tilde{G}(s)\tilde{R}(s)} \cdot bn! \right] = \\
 &= \begin{cases} 0 & \text{dla } r > n \\ b \frac{-n!kk_{zu}}{0^{l+r} + kk_r} & \text{dla } r = n \\ \infty & \text{dla } r < n \end{cases} \quad (6.12)
 \end{aligned}$$

Zwróćmy uwagę, że tym razem na eliminację uchybu w stanie ustalonym wpływa jedynie całkowanie w regulatorze (tzn., przed węzłem dodającym zakłócenie). Jeśli ilość całkowań w regulatorze r jest równa stopniowi wymuszenia ($r = n$), to uchyb ustalony w (stabilnym) układzie regulacji będzie niezerowy, a jeśli $r < n$, to uchyb będzie nieskończoność narastała. Zależności te zilustrowano w Tabeli 6.2, gdzie wyszczególniono uchyby ustalone dla wartości $n = 0$ (wymuszenie skokowe) i $n = 1$ (wymuszenie liniowo narastające, tzw. wymuszenie prędkościowe) i dla trzech istotnych wartości r , $r = 0, 1$ i 2 .

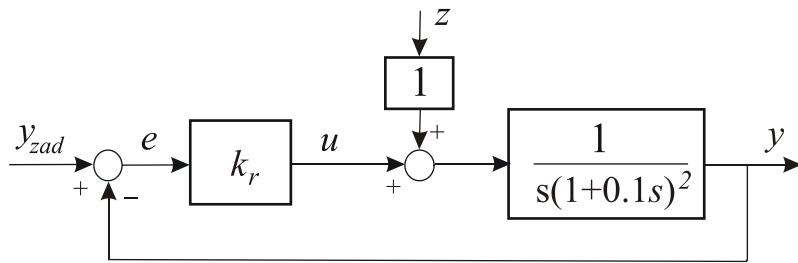
Tabela 6.2. Uchyby ustalone względem zakłócenia w układzie ze sprzężeniem zwrotnym, dla zmiany skokowej i prędkościowej $z(t)$ i dla różnej liczby r całkowań w regulatorze.

Wymuszenie \ r	$r = 0$	$r = 1$	$r = 2$
$n=0, \ z(t) = b\mathbf{1}(t)$	$e_{ust} = \frac{-bkk_{zu}}{0^l + kk_r}$	$e_{ust} = 0$	$e_{ust} = 0$
$n=1, \ z(t) = bt\mathbf{1}(t)$	$e_{ust} \rightarrow \infty$	$e_{ust} = \frac{-bk_{zu}}{k_r}$	$e_{ust} = 0$
\vdots	\vdots	\vdots	\vdots
Układ regulacji <u>względem zakłócenia $z(t)$ na wejściu:</u>	statyczny	astatyczny 1 rzędu (klasy 1)	astatyczny 2 rzędu (klasy 2)

Przykład 6.3.

Dla układu regulacji przedstawionego na Rys.6.5 należy wyznaczyć zakres wzmacnienia zapewniający:

1. Stabilność układu,
2. Tłumienie w stanach ustalonych zakłóceń skokowo zmiennych co najmniej dziesięciokrotne, tzn. $|e_{ust}| \leq 0.1|b|$, gdzie b jest amplitudą wymuszenia $z(t) = b\mathbf{1}(t)$.



Rys. 6.5. Układ regulacji w przykładzie 6.3.

ad 1. Stabilność. Transmitancja toru otwartego sterowania jest równa

$$G_0(s) = k_r \frac{1}{s(1+0.1s)^2}$$

Natomiast transmitancja zamkniętego układu regulacji

$$G_f(s) = \frac{G_0(s)}{1 + G_0(s)} = \frac{k_r}{s(1+0.1s)^2 + k_r} = \frac{k_r}{M_f(s)}$$

Stąd

$$M_f(s) = 0.01s^3 + 0.2s^2 + s + k_r$$

i korzystając z kryterium Hurwitza dostajemy, dla $k_r > 0$, warunek stabilności układu regulacji

$$\begin{vmatrix} 0.2 & 0.01 \\ k_r & 1 \end{vmatrix} > 0$$

Skąd wynika

$$0.2 - 0.01k_r > 0, \text{ czyli}$$

$$0 < k_r < 20$$

ad 2. Tłumienie zakłócenia. Regulator jest bez całkowania ($r = 0$), a obiekt jest astatyczny ($l = 1$), stąd układ regulacji jest statyczny (rzędu 0) względem wartości zadanej. Dla wymuszeń skokowych ($n = 0$) możemy odczytać uchyb z kolumny 1 Tabeli 6.2 .

Ten sam rezultat można również stosunkowo szybko uzyskać nie korzystając z Tabeli 6.2, bezpośrednio z transmitancji uchybowej (6.2):

$$\begin{aligned}
e_{ust} &= \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} [sE(s)] = \lim_{s \rightarrow 0} \left[s \frac{-G(s)G_{zu}(s)}{1+G(s)R(s)} \cdot \frac{b}{s} \right] = \lim_{s \rightarrow 0} \frac{\frac{-b}{s(1+0.1s)^2}}{1 + \frac{k_r}{s(1+0.1s)^2}} = \\
&= \lim_{s \rightarrow 0} \frac{-b}{s(1+0.1s)^2 + k_r} = -\frac{b}{k_r}
\end{aligned}$$

Z wymagania $|e_{ust}| \leq 0.1|b|$ wynika więc warunek jakości tłumienia zakłócenia

$$k_r \geq 10$$

Stąd łączne warunki na wzmacnianie regulatora zapewniające spełnienie przez układ regulacji postawionych wymagań:

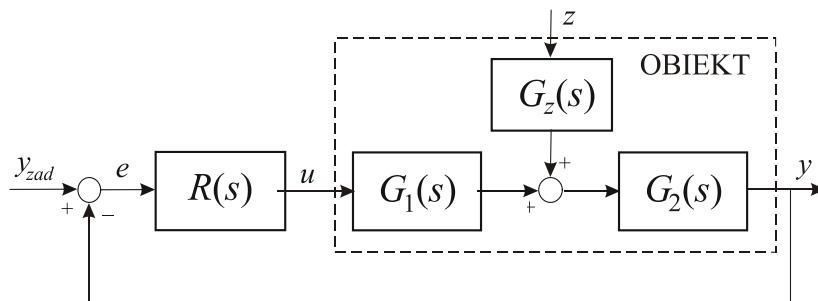
$$10 \leq k_r < 20 \quad \square$$

6.3. Układy regulacji z dowolnym miejscem wpływu zakłócenia

Nie zawsze naturalnie i korzystnie jest sprowadzać wpływ zakłócenia do wejścia obiektu, tak jak to pokazano na Rys. 6.1. W wielu sytuacjach struktura obiektu jest bardziej rozbudowana i zakłócenia oddziaływają wewnętrz tej struktury. Ta ogólna sytuacja przedstawiona jest na Rys. 6.6, gdzie obiekt składa się z kaskady dwóch transmitancji:

$$G(s) = G_1(s) \cdot G_2(s)$$

a zakłócenie dodawane jest do sygnału pomiędzy nimi.



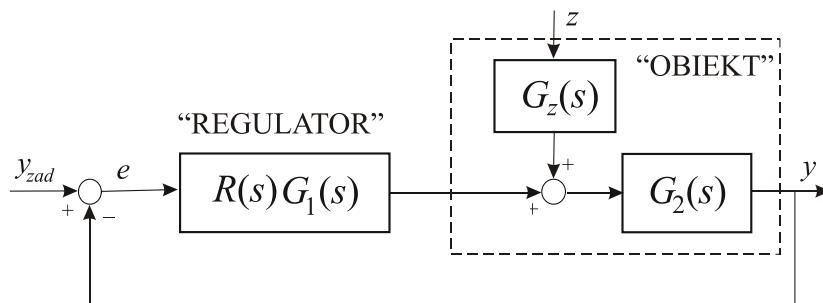
Rys. 6.6. Ogólny układ regulacji z kaskadową strukturą obiektu.

Struktura przedstawiona na Rys. 6.6 jest ogólna, obejmuje wszystkie trzy przypadki możliwych miejsc oddziaływanego zakłóceń na obiekt:

1. $G(s) = G_2(s)$ - zakłócenie oddziaływało na wejściu obiektu, najistotniejszy przypadek rozpatrzony dotychczas w tym rozdziale,
2. $G(s) = G_1(s)$ - zakłócenie oddziaływało na wyjściu obiektu, wpływ takiego zakłócenia na wyjście jest podobny jak wartości zadanej (prowadzi do takich samych zależności na uchyby ustalonej),
3. $G(s) = G_1(s) \cdot G_2(s)$ - zakłócenie oddziaływało wewnętrz obiektu, przypadek ten rozpatrzymy dalej.

Uwaga. Miejsce oddziaływanego zakłócenia nie ma wpływu na warunki stabilności zamkniętego układu regulacji. O stabilności decyduje bowiem transmitancja pętli otwartej układu regulacji, $G_0(s) = R(s)G(s)$.

Dla analizy wpływu zakłócenia w trzecim z wymienionych powyżej przypadków, przedstawmy strukturę układu regulacji z Rys. 6.6 w równoważnej postaci pokazanej na Rys. 6.7. Jeśli teraz porównamy strukturę z Rys. 6.7 z przedstawioną na Rys. 6.1 strukturą z oddziaływaniem zakłócenia na wejściu toru sterowania obiektu, to od razu zauważymy, że struktura z Rys. 6.7 jest równoważna tej ostatniej, jeśli tylko jako transmitancję „regulatora” potraktujemy transmitancję $R(s)G_1(s)$, a jako transmitancję „obiektu” transmitancję $G_2(s)$. Sprowadziliśmy tym samym przypadek z zakłóceniem oddziałującym wewnętrz toru sterowania obiektu do rozważonego już szczegółowo przypadku z oddziaływaniem zakłócenia na wejściu toru sterowania obiektu.

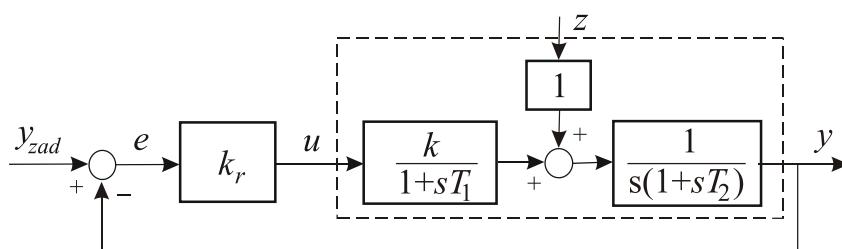


Rys. 6.7. Przekształcona struktura układu regulacji z Rys. 6.6.

Wniosek. O klasie układu (stopniu astatyzmu) względem zmian zakłócenia $z(t)$ decyduje ilość całkowań w transmitancji łącznej znajdującej się w torze sterowania otwartego układu regulacji przed węzłem wprowadzającym do tego toru sygnał zakłócenia.

Przykład 6.4.

Określić stopień astatyzmu (klasę) względem zmian wartości zadanej $y_{zad}(t)$ i zakłócenia $z(t)$ układu regulacji przedstawionego na Rys. 6.8. Zaproponować postać regulatora zapewniającego zerowe uchyby ustalone po skokowych zmianach zakłócenia.



Rys. 6.8. Układ regulacji w przykładzie 6.4.

Transmitancja toru otwartego sterowania

$$G_0(s) = \frac{k_r k}{s(1 + sT_1)(1 + sT_2)}$$

zawiera jedno całkowanie ($r+l = 1$). Stąd układ jest astatyczny 1 rzędu (klasy 1) względem wartości zadanej.

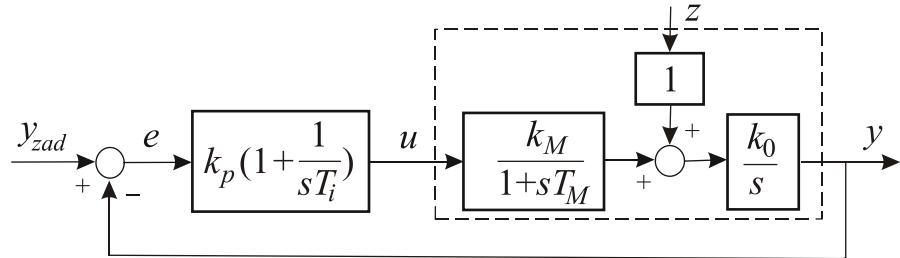
Natomiast transmitancja toru w części przed węzłem dodającym zakłócenie nie zawiera całkowania, stąd układ jest statyczny (klasy 0) względem zakłócenia. Aby zapewnić zerowy uchyb ustalony po skokowych zmianach zakłócenia, należy zastosować regulator zawierający całkowanie, np. regulator PI:

$$R(s) = k_p \left(1 + \frac{1}{T_i s}\right) = \frac{k_p (1 + T_i s)}{T_i s}$$
□

Przykład 6.5.

Dla układu regulacji przedstawionego na Rys. 6.9 podać wymagania na nastawy regulatora PI zapewniające:

1. Stabilność układu,
2. Zapewnienie w stanach ustalonych uchybu $|e_{ust}| \leq 0.05|b|$, gdzie $b[s^{-1}]$ jest amplitudą liniowo narastającego zakłócenia $z(t) = bt \cdot \mathbf{1}(t)$, dla parametrów obiektu: $k_0 = 10[s^{-1}]$, $k_M = 1$, $T_M = 2[s]$.



Rys. 6.9. Układ regulacji w przykładzie 6.5.

ad 1. Transmitancja układu otwartego:

$$G_0(s) = \frac{k_p (1 + sT_i)}{T_i s} \frac{k_M k_0}{s(1 + sT_M)}$$

Natomiast transmitancja zamkniętego układu regulacji (względem wartości zadanej):

$$G_f(s) = \frac{G_0(s)}{1 + G_0(s)} = \frac{k_p k_0 k_M (1 + sT_i)}{s^2 T_i (1 + sT_M) + k_p k_0 k_M (1 + sT_i)} = \frac{L_f(s)}{M_f(s)}$$

$$M_f(s) = s^3 T_i T_M + s^2 T_i + k_p k_0 k_M T_i s + k_p k_0 k_M$$

Stosując teraz kryterium Hurwitza, przy założeniu dodatniości wszystkich wzmacnień i stałych czasowych, otrzymujemy warunek stabilności

$$\begin{vmatrix} T_i & T_i T_M \\ k_p k_0 k_M & k_p k_0 k_M T_i \end{vmatrix} > 0$$

z którego wynika

$$k_p k_0 k_M T_i^2 - k_p k_0 k_M T_i T_M > 0$$

Czyli ostatecznie warunek stabilności postaci: $T_i > T_M = 2$.

ad 2. Transmitancja toru sterowania przed węzłem sumującym zakłócenie zawiera jedno całkowanie, stąd układ jest jedynie astatyczny 1 rzędu (klasy 1) względem zmian zakłócenia (a jest klasy 2 względem zmian wartości zadanej, bo cały tor sterowania zawiera 2 całkowania). Stąd dla zakłóceń zmieniających się liniowo w czasie występować będzie niezerowy uchyb ustalony, $e_{ust} = \frac{-bk_{zu}}{k_r}$ (patrz Tabela 6.2 dla $n = 1$ i $r = 1$), gdzie k_r to

wzmocnienie części statycznej „regulatora” – czyli w naszym przypadku transmitancji przed węzłem sumującym zakłócenie

$$\text{"}R(s)\text{"} = \frac{k_p(1+sT_i)}{T_is} \frac{k_M}{(1+sT_M)} = \frac{1}{s} \cdot \frac{k_p k_M (1+sT_i)}{T_i(1+sT_M)}$$

skąd $k_r = \frac{k_p k_M}{T_i} = \frac{k_p}{T_i}$, gdyż $k_M = 1$. Stąd, przyjmując $k_{zu} = 1$ (patrz rysunek układu regulacji) dostajemy

$$e_{ust} = \frac{-bk_{zu}}{k_r} = \frac{-bk_{zu}T_i}{k_p} = -b \frac{T_i}{k_p}$$

Warunek tłumienia zakłócenia $|e_{ust}| \leq 0.05|b|$ będzie więc spełniony, jeśli $T_i \leq 0.05k_p$.

Zależność na uchyb ustalony powodowany wpływem zakłócenia można też, alternatywnie nie korzystając z Tabeli 6.2, stosunkowo prosto uzyskać bezpośrednio z transmitancji uchybowej:

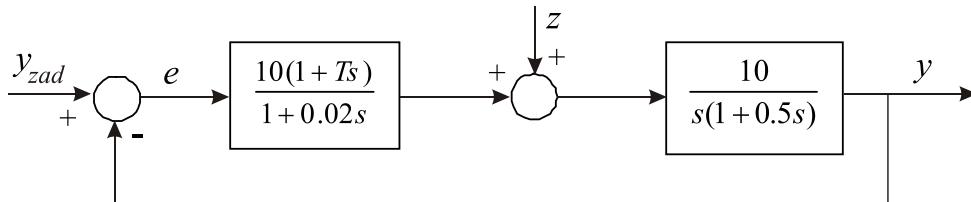
$$e_{ust} = \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} [sE(s)] = \lim_{s \rightarrow 0} [s \frac{-G_2(s)G_{zu}(s)}{1+G(s)R(s)} \cdot \frac{b}{s^2}] =$$

$$= \lim_{s \rightarrow 0} \frac{-bk_0}{\frac{s}{[1 + \frac{k_p k_M k_0 (1+sT_i)}{T_i s^2 (1+sT_M)}]s}} = \lim_{s \rightarrow 0} \frac{-bk_0}{s^2 + \frac{k_p k_M k_0 (1+sT_i)}{T_i (1+sT_M)}} = \frac{-bk_0 T_i}{k_p k_M k_0} = -b \frac{T_i}{k_p}$$

Reasumując, wszystkie wymagania zadania będą spełnione, jeśli $2 < T_i \leq 0.05k_p$. \square

Zadania

Zadanie 6.1. W układzie regulacji



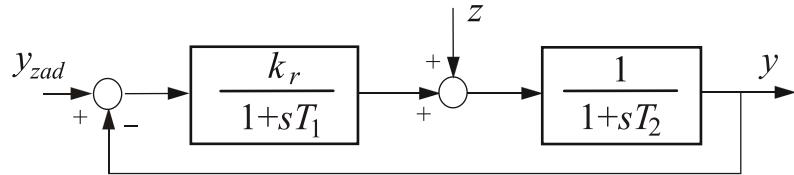
wyznacz zakres wartości stałej T , zapewniający asymptotyczną stabilność układu.

Jaki – przy spełnieniu wymagań stabilności – ustali się w tym układzie uchyb w stanie ustalonym:

- w odpowiedzi na skok wartości zadanej, $y_{zad}(t) = r_0 \cdot \mathbf{1}(t)$,

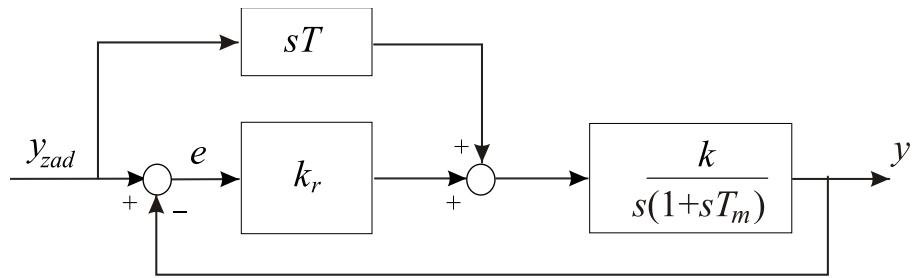
- w odpowiedzi na skok wartości zakłócenia na wejściu obiektu, $z(t) = z_0 \cdot \mathbf{1}(t)$.

Zadanie 6.2. Dla układu regulacji z zadania 5.5 o strukturze jak na rysunku



dobierz dodatkowo wzmacnienie k_r tak, aby wpływ ustalonego zakłócenia z był na wyjściu tłumiony co najmniej 100 razy.

Zadanie 6.3. W układzie regulacji



dobierz stałą czasową T tak, aby w stanie ustalonym uchyb regulacji był równy zero dla $y_{zad}(t) = at \cdot \mathbf{1}(t)$, gdzie a jest stałą (amplitudą wymuszenia prędkościowego).

7. Analiza i korekcja układów regulacji w dziedzinie częstotliwości

7.1. Transmisja sygnału okresowego przez układ liniowy

Rozważmy układ liniowy o transmitancji $G(s)$ pobudzony, począwszy od chwili zero, sterowaniem $u(t)$ – tj. $u(t) = 0$ dla $t < 0$, o transformacie $U(s)$ i przy niezerowych warunkach początkowych. Wówczas transformata wyjścia ma ogólną postać:

$$Y(s) = G(s) \cdot U(s) + \text{transformata składowej swobodnej}$$

Wiemy, że w przypadku układu liniowego asymptotycznie stabilnego składowa swobodna dąży z upływem czasu do zera. Zbadamy teraz zachowanie się składowej wymuszonej $Y_w(s)$, której transformata dana jest iloczynem $G(s)U(s)$.

Przyjmijmy następujące założenia:

1. $G(s) = \frac{L_G(s)}{M_G(s)}$, $L_G(\cdot)$ – wielomian stopnia l_G , $M_G(\cdot)$ – wielomian stopnia m_G .
2. $U(s) = \frac{L_U(s)}{M_U(s)}$, $L_U(\cdot)$ – wielomian stopnia l_U , $M_U(\cdot)$ – wielomian stopnia m_U .

Jednocześnie zakładamy, że $l_G + l_U < m_G + m_U$.

Przedstawiamy wielomiany w postaci kanonicznej; przyjmujemy, że wszystkie bieguny $G(s)$ i $U(s)$, tj. pierwiastki p_i wielomianu $M_G(s)$ oraz q_i wielomianu $M_U(s)$, są pojedyncze i wzajemnie różne, tj. $p_i \neq p_j \neq q_k$; $i \neq j$. Jest to założenie upraszczające dalsze rozumowanie, przedstawiony dalej wynik można także uzyskać w przypadku, gdy występują bieguny wielokrotne. Rozważaną transmitancję $G(s)$ oraz transformatę sterowania $U(s)$ można przedstawić w postaci

$$G(s) = k \frac{\prod_{i=1}^{l_G} (s - z_i)}{\prod_{i=1}^{m_G} (s - p_i)}, \quad U(s) = k_U \frac{\prod_{i=1}^{l_U} (s - v_i)}{\prod_{i=1}^{m_U} (s - q_i)}$$

Przy przyjętych założeniach możemy łatwo dokonać rozkładu transformaty $Y_w(s)$ składowej wymuszonej na ułamki proste:

$$Y_w(s) = \frac{C_1}{s - p_1} + \dots + \frac{C_{m_G}}{s - p_{m_G}} + \frac{C_{U1}}{s - q_1} + \dots + \frac{C_{Um_U}}{s - q_{m_U}}.$$

Pierwsza część transformaty składowej wymuszonej związana z biegunami transmitancji obiektu p_i , stanowi transformatę $Y_{kom}(s)$ składowej komutacyjnej,

$$Y_{kom}(s) = \frac{C_1}{s - p_1} + \dots + \frac{C_{m_G}}{s - p_{m_G}},$$

zaś druga część transformaty wyjścia odpowiada składowej transmisyjnej. Jeśli $\operatorname{Re} p_i < 0$, czyli bieguni transmitancji mają ujemne części rzeczywiste, to składowa komutacyjna, której dynamikę kształtują te bieguni, dąży z czasem, podobnie jak składowa swobodna, do zera. Wówczas dla odpowiednio dużych wartości czasu, $t \gg 0$, obserwujemy w istocie tylko składową transmisyjną, swoim charakterem odpowiadającą – pod względem dynamiki – przebiegowi wejściowemu; bieguni składowej transmisyjnej

$$Y_{tr}(s) = \frac{C_{U1}}{s - q_1} + \dots + \frac{C_{Um}}{s - q_m}$$

są jednocześnie biegunami transformaty sterowania.

Szczególnie ważną rolę odgrywają wejścia w postaci funkcji sinus lub cosinus; przypomnijmy, że ciągle funkcje okresowe można dowolnie dokładnie przybliżać sumą takich funkcji. Stąd warto zbadać składową transmisyjną przebiegu wyjścia wymuszonego funkcją o postaci $u(t) = A \cdot \sin(\omega t) \cdot \mathbf{1}(t)$.

Niech więc $u(t) = A \cdot \sin(\omega t)$ dla $t \geq 0$ ($u(t) = 0$ dla $t < 0$) będzie sygnałem podawanym na układ liniowy, o transmitancji $G(s)$. Transformata $u(t)$ jest określona dla wartości s takich, że $\operatorname{Re} s > 0$ i ma postać:

$$U(s) = A \frac{\omega}{s^2 + \omega} = \frac{A}{2j} \left[\frac{1}{s - j\omega} - \frac{1}{s + j\omega} \right].$$

$q_1 = j\omega$, $q_2 = -j\omega$ są biegunami transformaty sterowania. Wówczas transformata składowej transmisyjnej ma postać:

$$Y_{tr}(s) = \frac{C_{U1}}{s - j\omega} + \frac{C_{U2}}{s + j\omega},$$

gdzie współczynniki zespolone C_{U1} i C_{U2} wyznaczamy w następujący sposób:

$$\begin{aligned} C_{U1} &= \lim_{s \rightarrow j\omega} (s - j\omega) \cdot Y_{tr}(s) = \lim_{s \rightarrow j\omega} (s - j\omega) \cdot Y(s) = \lim_{s \rightarrow j\omega} (s - j\omega) G(s) U(s) = \\ &= \lim_{s \rightarrow j\omega} (s - j\omega) G(s) \frac{A}{2j} \left[\frac{1}{s - j\omega} - \frac{1}{s + j\omega} \right]_{s=j\omega} = \frac{A}{2j} G(j\omega) \end{aligned}$$

$$\text{podobnie } C_{U2} = \lim_{s \rightarrow -j\omega} (s + j\omega) \cdot Y_{tr}(s) = -\frac{A}{2j} G(-j\omega).$$

Stąd przebieg składowej transmisyjnej dla czasu $t > 0$ ma postać:

$$y_{tr}(t) = \frac{A}{2j} G(j\omega) e^{j\omega t} - \frac{A}{2j} G(-j\omega) e^{-j\omega t}.$$

Biorąc pod uwagę, że $G(s)$ jest funkcją wymierną o współczynnikach rzeczywistych zmiennej zespolonej mamy $G(-j\omega) = \overline{G(j\omega)}$. Ponieważ $G(j\omega) = |G(j\omega)| e^{j\varphi(\omega)}$, gdzie $\varphi(\omega) = \arg G(j\omega)$, to otrzymujemy

$$y_{tr}(t) = \frac{A}{2j} |G(j\omega)| e^{j(\omega t + \varphi)} - \frac{A}{2j} |G(j\omega)| e^{-j(\omega t + \varphi)}.$$

Po uproszczeniu wynika stąd podstawa dla analizy układu liniowego zależność:

$$y_{tr}(t) = A |G(j\omega)| \sin(\omega t + \varphi(\omega)) \quad (7.1)$$

Widzimy, że składowa transmisyjna wyjścia wymuszonego sygnałem sinusoidalnym o pulsacji ω i amplitudzie A jest sinusoidą o amplitudzie $A/G(j\omega)$ i pulsacji ω , przesuniętą w fazie o kąt $\varphi(\omega)$. Stąd wartości modułu i argumentu transmitancji mają podstawowe znaczenie przy badaniu przenoszenia przez układ liniowy sygnałów sinusoidalnych.

Przedstawione rozumowanie może być powtórzone dla przypadku, w którym transmitancja

obiektu ma postać funkcji wymiernej, jak wyżej, przemnożonej przez człon mnożnikowy $\exp(-sT_0)$ odpowiadający istnieniu opóźnienia w podawaniu sygnału sterującego na wejście obiektu. Otrzymany wynik, dotyczący składowej transmisyjnej wyjścia, jest identyczny, z tym, że przesunięcie fazy $\varphi(\omega)$ zawiera zależny liniowo od pulsacji składnik równy $-\omega T_0$. Oznacza to, że pomiędzy przesunięciem fazy wnoszonym przez dowolny układ liniowy bez opóźnienia w sterowaniu, $\varphi_{bez\cdot op}(\omega)$, a przesunięciem fazy wnoszonym przez ten sam układ po wprowadzeniu opóźnienia, $\varphi_{op}(\omega)$, zachodzi związek: $\varphi_{op}(\omega) = \varphi_{bez\cdot op}(\omega) - \omega T_0$.

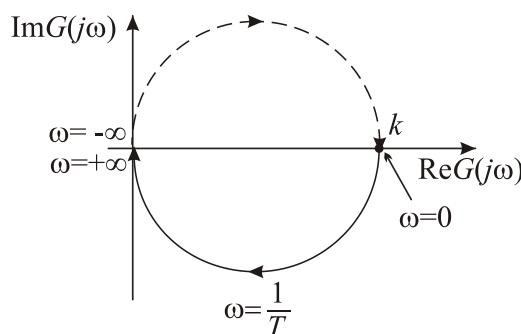
Należy na zakończenie wyraźnie podkreślić, że tylko dla układu asymptotycznie stabilnego składowa transmisyjna wyjścia określa w pełni jego zachowanie w dłuższej perspektywie czasowej; w przypadku układu nie asymptotycznie stabilnego składowa ta występuje na tle składowych swobodnej i komutacyjnej, i może być przez te składowe zdominowana.

Przedstawiona w poprzednich punktach analiza ukazuje potrzebę posługiwania się transmitancją obiektu jako funkcją pulsacji ω . Zmienność $G(j\omega)$ można obrazować albo w postaci tzw. *hodografu*, zwanego także *charakterystyką amplitudowo-fazową* transmitancji lub *krzywą Nyquista*, albo w postaci pary charakterystyk częstotliwościowych: amplitudowej i fazowej, zwanych też *charakterystykami (krzywymi) Bodego*.

7.2. Charakterystyki częstotliwościowe

Charakterystyka amplitudowo-fazowa (hodograf)

Biorąc pod uwagę, że $G(j\omega) = \text{Re } G(j\omega) + j \text{Im } G(j\omega)$, hodograf transmitancji jest odwzorowaniem na płaszczyźnie zbioru punktów $(\text{Re } G(\omega), \text{Im } G(\omega))$, parametryzowanych przez pulsację ω ; przy zmianie ω od $-\infty$ do $+\infty$, lub od 0 do $+\infty$ (relacja $G(-j\omega) = \overline{G(j\omega)}$) sprawia, że hodograf dla ujemnych pulsacji jest lustrzanym odbiciem hodografu dla dodatnich pulsacji). Przedstawiony na Rys. 7.1 przykład hodografu dotyczy układu inercyjnego o transmitancji $k/(sT+1)$.



Rys. 7.1. Hodograf (charakterystyka amplitudowo-fazowa) transmitancji członu inercyjnego, przy zmianie ω od $-\infty$ do 0 (linia przerywana) i od 0 do $+\infty$ (linia ciągła).

Hodograf transmitancji jest szczególnie pomocny przy badaniu relacji, jaka zachodzi pomiędzy układem otwartym o danej transmitancji i układem zamkniętym powstały po objęciu układu otwartego pętlą ujemnego sprzężenia zwrotnego. Nawiążemy do tego wątku przedstawiając w kolejnym punkcie ważne kryterium Nyquista.

Charakterystyki logarytmiczne amplitudowa i fazowa

Hodograf transmitancji, czyli charakterystyka amplitudowo-fazowa, przedstawia na płaszczyźnie pełen obraz zmienności tej funkcji, jednak pulsacja ω występuje tu jako

parametr towarzyszący punktom krzywej, a nie jako współrzędna niezależna na osi odciętych. Stąd, w szeregu zastosowań, w szczególności przy projektowaniu układu regulacji, wygodniej jest posługiwać się charakterystykami częstotliwościowymi transmitancji – charakterystyką amplitudową i charakterystyką fazy (fazową).

Charakterystyka amplitudowa przedstawia przebieg wielkości określonej jako $Lm(\omega)=20\log|G(j\omega)|$ [dB] w funkcji pulsacji ω , dla $0 < \omega < +\infty$; na osi pulsacji stosowana jest skala logarytmiczna, natomiast jednostkami na osi rzędnych są tzw. decyble. Miary w postaci decybeli używamy w sytuacji, gdy chcemy porównywać wielkości zmieniające się liniowo w bardzo szerokim zakresie; z taką właśnie sytuacją mamy do czynienia w przypadku zmian wzmacnienia amplitudy sygnału sinusoidalnego przenoszonego przez dynamiczny układ liniowy.

Warto w tym miejscu zauważyć, że rozważając zmienność takiej wielkości jak poziom mocy sygnału P stosujemy miarę zmienności pomiędzy, powiedzmy, poziomem mocy P_1 i P_2 jako wyrażoną w decybelach i równą $10\log(P_1/P_2)$; logarytm obliczany jest oczywiście przy podstawie 10. Opisując w naszym przypadku wzmacnienie $|G(j\omega)|$ jako stosunek amplitud sygnałów, w elektronice są to często amplitudy napięcia, stosujemy przy mnożeniu wartości logarytmu współczynnik równy 20 zamiast współczynnika 10. Wynika to stąd, że poziom mocy sygnału zależny jest od kwadratu amplitudy; zatem określona powyżej wielkość $Lm(\omega)=20\log|G(j\omega)|$ odpowiada w istocie wyrażonej w decybelach mierze zmienności (wzmacnienia) poziomów mocy rozważanych sygnałów. Oznaczenie Lm odpowiada innej często stosowanej nazwie charakterystyki amplitudowej jako logarytmicznej charakterystyki modułu. W sytuacji, gdy 10-krotny wzrost częstotliwości powoduje 10-krotny wzrost (spadek) wzmacnienia, mówi się o wzroście (spadku) 20 dB na dekadę lub, odpowiednio, o nachyleniu charakterystyki +1 (-1). Stukrotny wzrost (spadek) wzmacnienia na dekadę odpowiada, oczywiście nachyleniu charakterystyki +2 (-2).

Szczególnie ważnym powodem posługiwania się logarytmiczną charakterystyką modułu transmitancji jest to, że w przypadku funkcji transmitancji będącej iloczynem kilku transmitancji różnych członów połączonych szeregowo wypadkowa logarytmiczna charakterystyka modułu jest sumą charakterystyk modułu poszczególnych członów. Z taką sytuacją mamy do czynienia bardzo często przy analizie układów liniowych, w tym podczas projektowania układów regulacji.

Charakterystyka fazowa, zwana także charakterystyką fazy, przedstawia przebieg wielkości określonej jako $\varphi(\omega) = \arg G(j\omega)$ w funkcji pulsacji ω , dla $0 < \omega < +\infty$; także w tej sytuacji stosowana jest skala logarytmiczna na osi pulsacji. Wypada zauważyć, że również w przypadku charakterystyki fazowej dla funkcji transmitancji będącej iloczynem kilku transmitancji różnych członów wypadkowa charakterystyka fazy jest sumą charakterystyk fazy poszczególnych członów. Wynika to oczywiście stąd, że przy mnożeniu wielkości zespolonych argument iloczynu jest sumą argumentów poszczególnych czynników.

Prezentując charakterystyki częstotliwościowe podstawowych liniowych członów dynamicznych będziemy też omawiać technikę przedstawiania charakterystyk częstotliwościowych w postaci charakterystyk asymptotycznych, które są bardzo łatwe do konstrukcji i umożliwiają uproszczoną, ale bardzo skuteczną analizę dynamiki i wstępna syntezę postaci regulatora. Przedstawimy je na przykładzie członów proporcjonalno-różniczkującego (idealnego) i inercyjnego, najpierw jednak zaczniemy od najprostszego członu proporcjonalnego (dla którego charakterystyki dokładne i asymptotyczne są identyczne).

7.3. Charakterystyki częstotliwościowe prostych członów dynamicznych

Człon proporcjonalny

W tym najprostszym przypadku mamy do czynienia z czystym wzmacnieniem, $G(s) = k$, $k > 0$. Oczywiście charakterystyka amplitudowa ma stałą wartość, równą $20\log k$, wnoszone przesunięcie fazowe jest zerowe.

Człony różniczkujący i proporcjonalno-różniczkujący (idealne)

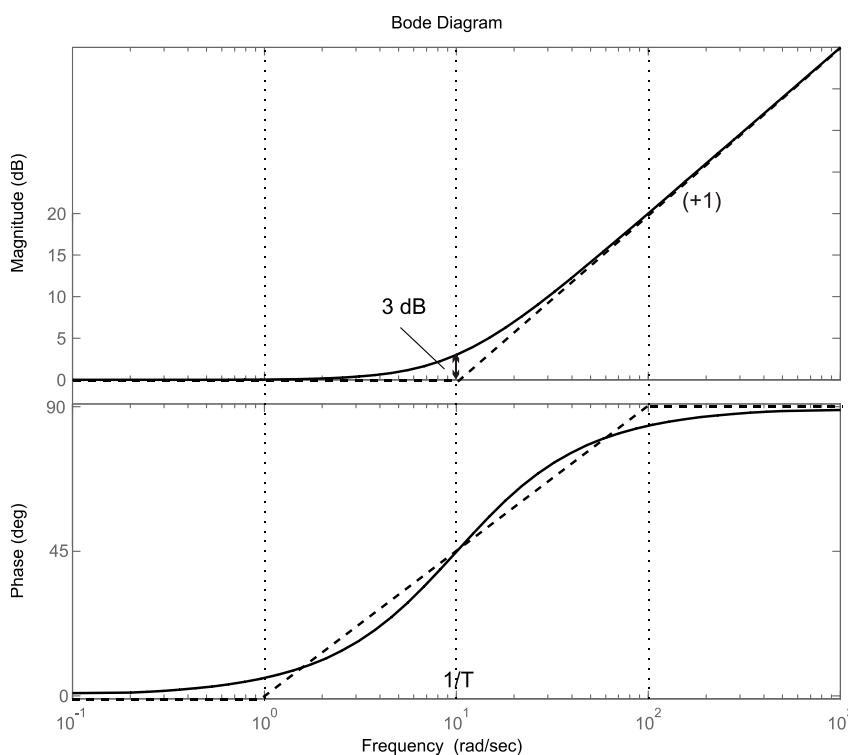
Dla członu idealnego proporcjonalno-różniczkującego, o jednostkowym wzmacnieniu, tzn. o transmitancji $G(s) = 1+sT$, charakterystyki dokładne (linie ciągłe) i asymptotyczne (linie przerywane) przedstawia Rys. 7.2.

Charakterystyka asymptotyczna modułu dla $s=j\omega$, $0 < \omega < +\infty$, przebiega następująco: dla wartości $\omega < 1/T$ wyrażenie

$$20\log|1+j\omega T| = 20\log\sqrt{1+\omega^2 T^2}$$

przyjmuje, w przybliżeniu, wartość równą零, natomiast dla $\omega > 1/T$ wyrażenie to przybliża się wartością $20\log(\omega T)$, która przy każdym dziesięciokrotnym zwiększeniu (tzn. o dekadę) wartości ω wzrasta o 20 dB – mówimy w takim przypadku, że wzrasta z nachyleniem +1.

Charakterystyka asymptotyczna fazy może być konstruowana na podstawie następującego rozumowania: przesunięcie fazowe $\varphi(\omega)$ wnoszone przez człon $(1+sT)$ wynosi, dla $s=j\omega$ (punkt załamania asymptotycznej charakterystyki modułu) $\text{arctg}(\omega T)$; stąd dla $\omega=1/(10T)$ $\varphi(1/(10T)) = \text{arctg}(0,1) = 0,1 \text{ rad } (5,7^\circ)$, zaś dla $\omega=10/T$ mamy $\varphi(10/T) = \text{arctg}(10) = 1,47 \text{ rad } (84,3^\circ)$. Stąd, w przybliżeniu, można przesunięcie fazowe powodowane przez ten człon przyjmować za równe zero dla $\omega < 1/(10T)$ oraz równe 1,57 rad (90°) dla $\omega > 10/T$, a w zakresie między $\omega=1/(10T)$ a $\omega=10/T$ konstruować charakterystykę asymptotyczną fazy jako odcinek funkcji liniowej zmieniającej wartość o 1,57 rad (o 90°).



Rys. 7.2. Charakterystyki częstotliwościowe członu o transmitancji $1+sT$ dla $T=0.1$ (liniami przerywanymi zaznaczono charakterystyki asymptotyczne).

Charakterystyka amplitudowo-fazowa członu $G(j\omega) = k(1+j\omega T)$ to dla pulsacji ω z przedziału od 0 do $+\infty$ pionowa półprosta dodatnia rozpoczęta na osi rzędnych w punkcie k (półprosta $\operatorname{Re}G(j\omega) = k$).

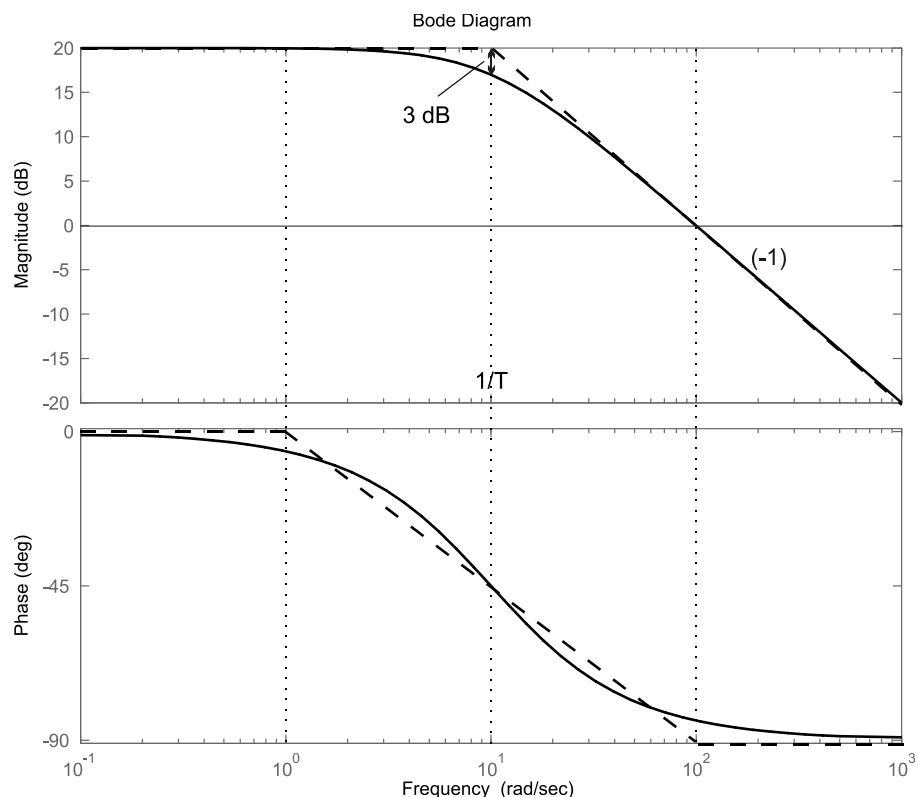
Idealny człon różniczkujący, o transmitancji $G(s) = sT$, jest szczególnym przypadkiem członu rozważonego powyżej, wnosi on stałe dodatnie przesunięcie fazowe równe $\pi/2$ rad ($+90^\circ$), a jego charakterystyka amplitudowa modułu jest linią prostą o nachyleniu +1 (20 dB na dekadę), przecinającą oś 0 dB w punkcie $\omega=1/T$. Jest on praktycznie niemożliwy do zrealizowania (a stąd również człon proporcjonalno-różniczkujący $1+sT$).

Człon inercyjny

Transmitancja ma postać $G(s) = \frac{k}{1+sT}$, $k>0$, $T>0$, stąd moduł i przesunięcie fazowe

$$|G(j\omega)| = \frac{k}{\sqrt{1+T^2\omega^2}}, \quad \varphi(\omega) = -\arctg(T\omega) \quad (7.2)$$

Charakterystyka amplitudowa i fazowa dla przypadku $T=0.1$, $k=10$ pokazane są na Rys. 7.3, linią przerywaną zaznaczono charakterystyki asymptotyczne.



Rys. 7.3. Charakterystyki częstotliwościowe członu inercyjnego dla $T=0.1$, $k=10$.

Przebieg asymptotycznej logarytmicznej charakterystyki modulu jest następujący: dla $\omega < 1/T$ ma ona wartość $20\log k$ [dB], natomiast dla pulsacji $\omega > 1/T$ charakterystyka ta opada z nachyleniem -1, tj. -20 dB na dekadę. W punkcie $\omega = 1/T$ odchylenie charakterystyki asymptotycznej od rzeczywistej jest maksymalne i wynosi 3 dB.

Oczywiście, dla członu $k/(1+sT)$ charakterystyka fazy jest lustrzanym odbiciem charakterystyki członu $1+sT$, asymptotyczna charakterystyka fazy będzie zmniejszać wartość o 90° , w przybliżeniu w zakresie pulsacji między $\omega=1/(10T)$ a $\omega=10/T$.

Charakterystyka amplitudowo-fazowa przedstawiona jest na Rys. 7.1, jest to półokrąg, ponieważ można łatwo pokazać, że części rzeczywista i urojona transmitancji, dane wzorem

$$G(j\omega) = \frac{k}{1 + j\omega T} = \frac{k}{1 + T^2\omega^2} - j \frac{k\omega T}{1 + T^2\omega^2}$$

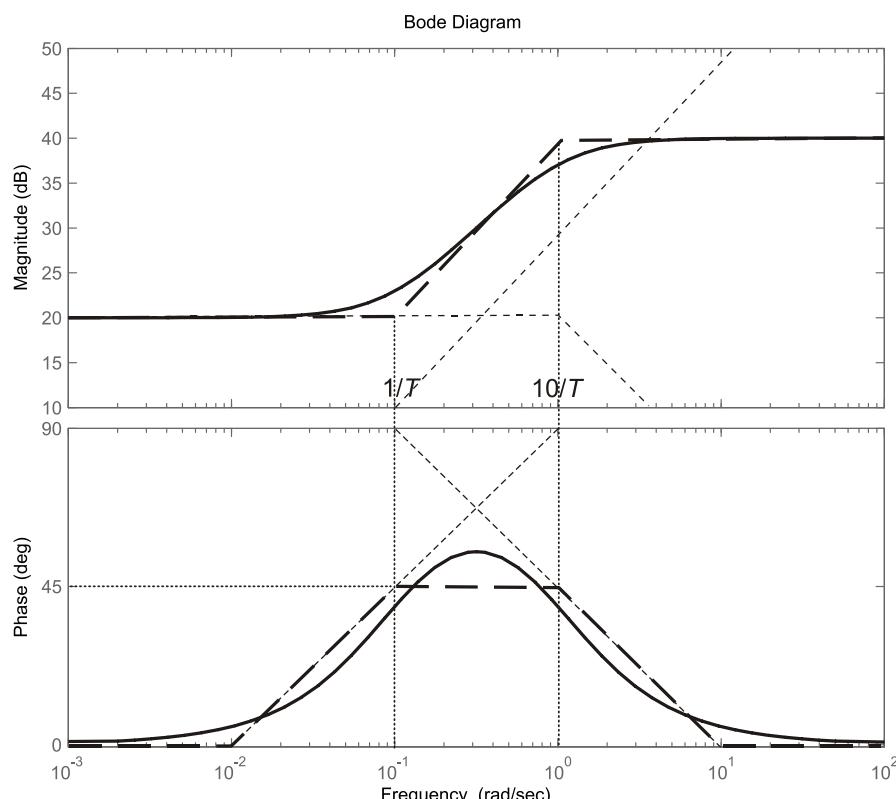
spełniają równanie półokręgu o promieniu $k/2$ i środku w punkcie $(k/2, 0)$.

Przed przystąpieniem do prezentacji kolejnych członów dynamicznych przypomnijmy, że przebiegi charakterystyk logarytmicznych modułu i fazy (zarówno dokładne jak i asymptotyczne) szeregowego połączenia dowolnej liczby członów dynamicznych, tj. iloczynu ich transmitancji, otrzymujemy sumując charakterystyki poszczególnych członów składowych.

Realizowalny człon proporcjonalno-różniczkujący

Idealne różniczkowanie nie jest realizowalne fizycznie, stąd stosowany jest realizowalny (rzeczywisty) człon proporcjonalno-różniczkujący, który występuje np. w postaci

$$G(s) = k \frac{1+sT}{1+s(T/\alpha)}, \quad \text{gdzie } \alpha > 1 \quad (7.3)$$



Rys. 7.4. Charakterystyki częstotliwościowe realizowanego członu proporcjonalno-różniczkującego $k=10$, $T=10$, $\alpha=10$).

lub w postaci $G(s) = k(1 + \frac{sT}{1+sT/\alpha})$, która jest bardzo podobna, ponieważ $G(s) = k(1 + \frac{sT}{1+sT/\alpha}) = k(\frac{1+sT(1+1/\alpha)}{1+sT/\alpha})$, a wartość $1+1/\alpha$ niewiele odbiega od jedności.

Zauważmy, że można potraktować powyższy człon jako szeregowego połączenie omówionych uprzednio członów proporcjonalno-różniczkującego (idealnego) i inercyjnego:

$$G(s) = k \frac{1+sT}{1+sT/\alpha} = (1+sT) \cdot \frac{k}{1+sT/\alpha}$$

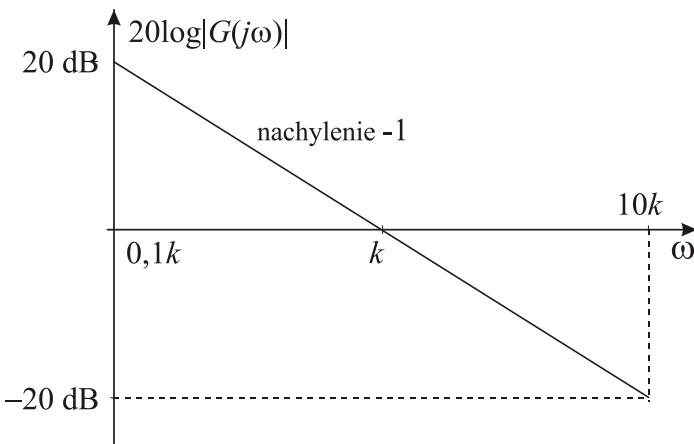
Stąd jego charakterystyki są zsumowaniem charakterystyk tych członów, ze stałymi czasowymi, odpowiednio, T i T/α (wartość α przyjmuje się, np. w regulatorach PID, ok. 10): Charakterystyki, zarówno dokładne jak i asymptotyczne, przedstawione są na Rys. 7.4, dla transmitancji $G(s)$ o parametrach $k=10$, $T=10$, $\alpha=10$. Zaznaczono też, cienkimi przerywanymi liniami, charakterystyki asymptotyczne członów składowych (przy czym charakterystyka fazowa członu inercyjnego jest przesunięta o 90° do góry). Człon wprowadza dodatnie przesunięcie fazy w zakresie średnich pulsacji, ale jednocześnie podbija (wzmacnia) amplitudę sygnału wejściowego dla wyższych pulsacji.

Człon całkujący

Transmitancja Laplace'a członu całkującego ma postać

$$G(s) = \frac{k}{s} \Rightarrow |G(j\omega)| = \frac{k}{\omega}, \quad \varphi(\omega) = -\frac{\pi}{2}$$

Widać natychmiast, że człon całkujący realizuje wzmacnianie amplitudy drgań sinusoidalnych odwrotnie proporcjonalne do pulsacji tych drgań i wnosi stałe ujemne przesunięcie fazowe równe $-\pi/2$ (-90°). Charakterystyka amplitudowa pokazana jest na rysunku; jest to prosta o nachyleniu -1, tj. -20 decybeli na dekadę.



Rys. 7.5. Charakterystyka amplitudowa członu całkującego.

Warto zauważyć, że człon całkujący jest układem stabilnym, ale nie jest asymptotycznie stabilnym. W tej sytuacji dla wejścia w formie sinusoidy, tj. $u(t) = A \cdot \sin(\omega t)$ dla $t \geq 0$ ($u(t) = 0$ dla $t < 0$) otrzymujemy, przy warunku początkowym y_0 , przebieg wyjściowy:

$$y(t) = y_0 + \frac{k \cdot A}{\omega} - \frac{k \cdot A}{\omega} \cdot \cos(\omega t)$$

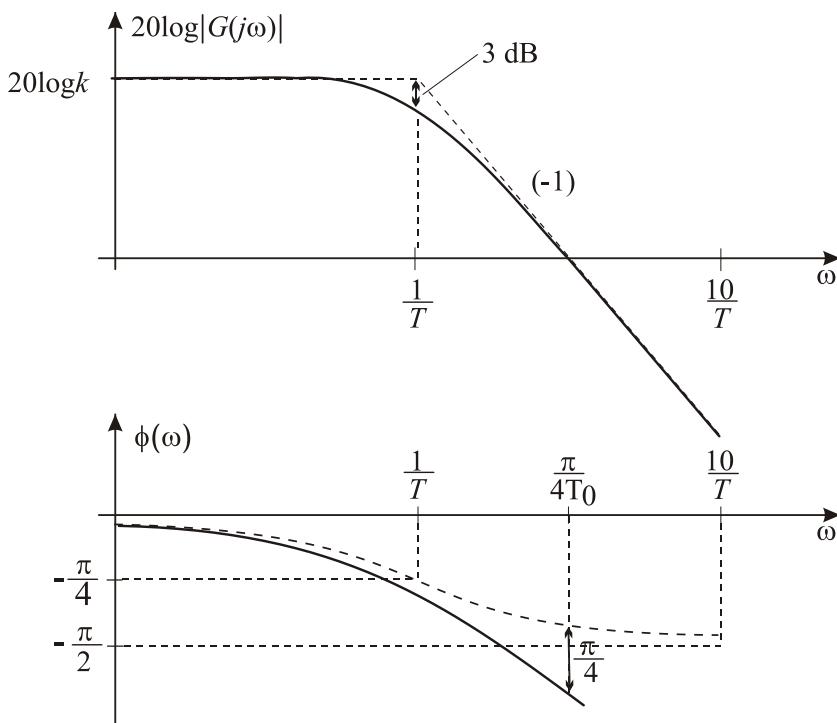
W tym przypadku składowa swobodna nie zanika z czasem – ma stałą wartość, podobnie, składowa komutacyjna także nie zanika i także ma wartość stałą, równą $\frac{k \cdot A}{\omega}$. Natomiast składowa transmisyjna ma postać: $-\frac{k \cdot A}{\omega} \cdot \cos(\omega t) = +\frac{kA}{\omega} \sin\left(\omega t - \frac{\pi}{2}\right)$.

Człon inercyjny z opóźnieniem

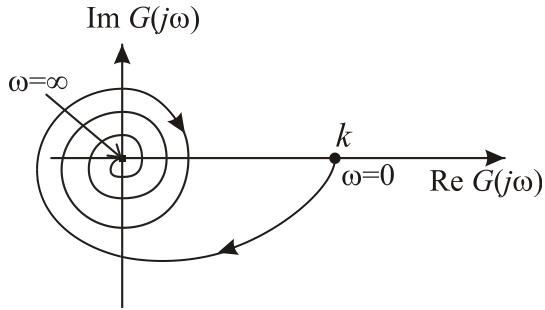
W tym przypadku mamy

$$G(s) = \frac{k}{sT+1} e^{-sT_0} \Rightarrow |G(j\omega)| = \frac{k}{\sqrt{1+T^2\omega^2}}, \quad \phi(\omega) = -\arctg(T\omega) - T_0\omega \quad (7.4)$$

Jak łatwo zauważyć, wprowadzenie opóźnienia nie ma wpływu na przebieg charakterystyki amplitudowej, natomiast istotnej zmianie w stosunku do przypadku bez opóźnienia ulega charakterystyka fazowa: opóźnienie powoduje zmniejszenie kąta fazowego o $-T_0\omega$, tzn. dla $\omega=\pi/2T_0$ o $-\pi/2$ (zob. Rys. 7.6), dla $\omega=\pi/T_0$ o $-\pi$, itd., jest to więc dla większych wartości pulsacji szybki dodatkowy spadek kąta fazowego, dążący do $-\infty$, jak to pokazuje hodograf przedstawiony na Rys. 7.7.



Rys. 7.6 Charakterystyki częstotliwościowe członu inercyjnego z opóźnieniem (dla $T_0=T/4$; linią przerywaną zaznaczono charakterystykę fazową członu bez opóźnienia, opóźnienie nie wpływa na charakterystyki amplitudowe).



Rys. 7.7. Hodograf członu inercyjnego z opóźnieniem.

Człon drugiego rzędu (oscylacyjny)

Standardowa postać transmitancji członu (zob. rozdz. 5.2.3):

$$G(s) = \frac{k\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \quad k > 0, \quad \zeta > 0. \quad (7.5)$$

Jeśli $\zeta \geq 1$, to biegunki $G(s)$ są rzeczywiste,

$$p_1 = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}, \quad p_2 = -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1},$$

i układ można wówczas przedstawić jako szeregowego połączenie dwóch członów inercyjnych:

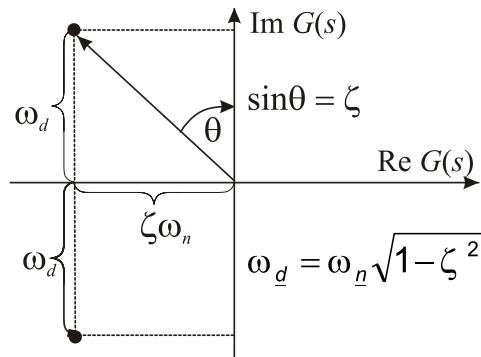
$$G(s) = \frac{k_1}{T_1 s + 1} \cdot \frac{k_2}{T_2 s + 1}, \quad T_1 = -\frac{1}{p_1}, \quad T_2 = -\frac{1}{p_2}, \quad T_1 \geq T_2, \quad k_1 k_2 = k.$$

Człon inercyjny był już rozważany, stąd przypadkiem tym nie będziemy się dalej zajmować.

Z bardziej interesującą sytuacją mamy do czynienia wtedy, gdy dodatni współczynnik tłumienia ζ ma wartość mniejszą od jedności, tj. $0 < \zeta < 1$. Wówczas biegunki transmitancji są zespolone o następujących wartościach:

$$p_1 = -\zeta\omega_n + \omega_n j\sqrt{1-\zeta^2}, \quad p_2 = -\zeta\omega_n - \omega_n j\sqrt{1-\zeta^2}$$

jak to przedstawiono na Rys. 7.8. Członu tego nie można wówczas przedstawić w postaci połączenia szeregowego dwóch członów inercyjnych.

Rys. 7.8. Położenie biegunków członu oscylacyjnego ($\zeta < 1$).

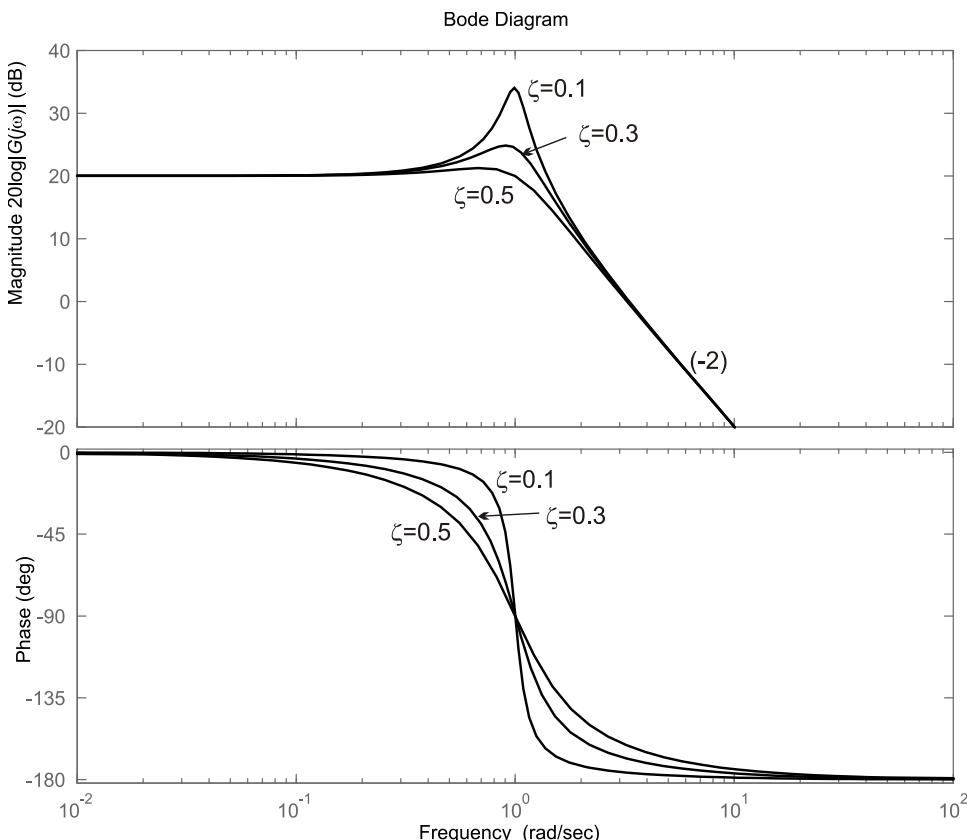
Transmitancja i jej moduł dla $s=j\omega$ są w tym przypadku następujące:

$$G(j\omega) = \frac{k\omega_n^2}{-\omega^2 + j2\zeta\omega_n\omega + \omega_n^2}, \quad |G(j\omega)| = \frac{k\omega_n^2}{\sqrt{(\omega_n^2 - \omega^2)^2 + 4\zeta^2\omega_n^2\omega^2}}.$$

Jak łatwo zauważyc, dla $\omega \ll \omega_n$ moduł transmitancji jest w przybliżeniu stały i równy k , natomiast dla $\omega \gg \omega_n$ otrzymujemy przybliżenie w postaci $k(\omega_n/\omega)^2$, wówczas wartość modułu maleje stukrotnie przy dziesięciokrotnym wzroście pulsacji – charakterystyka modułu (logarytmiczna) opada z nachyleniem -2 (- 40 dB na dekadę). W ten sposób otrzymujemy asymptotyczną charakterystykę modułu. W okolicy wartości $\omega = \omega_n$ kształt charakterystyki modułu jest istotnie zależny od wartości współczynnika tłumienia ζ , mamy bowiem

$$G(j\omega) = \frac{k}{1 - \frac{\omega^2}{\omega_n^2} + j2\zeta\frac{\omega}{\omega_n}}, \quad \text{dla } \omega \approx \omega_n : G(j\omega) \approx \frac{k}{-j2\zeta\frac{\omega}{\omega_n}}, \quad |G(j\omega)| = \frac{k}{2\zeta\frac{\omega}{\omega_n}}$$

Widac stąd, że wartość modułu może wzrastać bardzo znacznie dla małych wartości ζ . Badając ten przebieg analitycznie, uzyskujemy dla wartości dodatnich $\zeta < 0,708$ tzw. szczyt rezonansowy, tj. podniesienie wartości ponad poziom $20\log k$; na przykład dla $\zeta = 0.1$ – co odpowiada bardzo małemu tłumieniu – otrzymujemy przewyższenie tej wartości o ok. 14dB, oznacza to pięciokrotne wzmacnianie amplitudy podanej na wejście układu sinusoidy o pulsacji rezonansowej $\omega_r = \omega_n\sqrt{1-2\zeta^2} = \omega_n\sqrt{0,98} \approx 0,99\omega_n$, w porównaniu do wzmacniania przebiegów o małych pulsacjach.



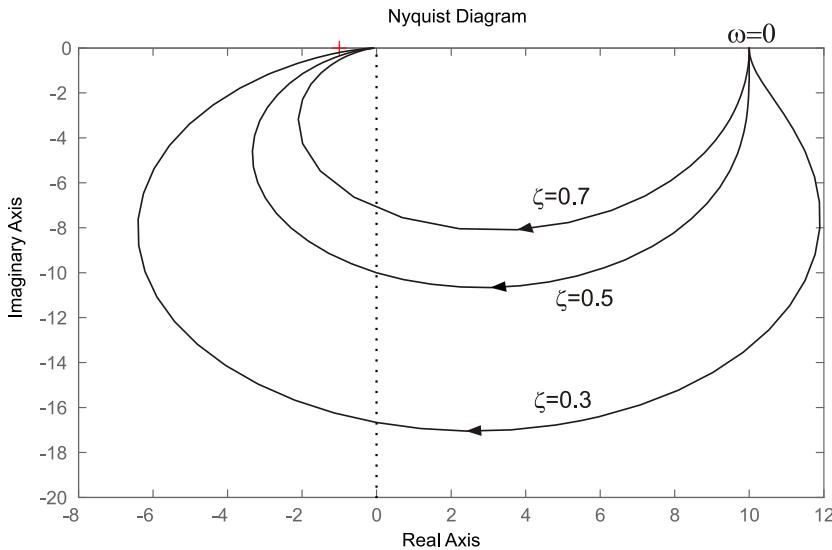
Rys. 7.9. Charakterystyki częstotliwościowe członu oscylacyjnego ($k = 10$, $\omega_n = 1$, $\zeta < 0.7$).

Tak więc należy zwracać uwagę, zwłaszcza w przypadku małej wartości współczynnika tłumienia, na różnice pomiędzy przebiegiem charakterystyki asymptotycznej i rzeczywistej członu oscylacyjnego w okolicy pulsacji ω_n ; różnice te mogą być, jak pokazano powyżej, bardzo znaczące.

Charakterystyka fazowa może być wyznaczona w oparciu o zmianę argumentu mianownika $G(j\omega)$; mianowicie:

$$\arg G(j\omega) = \begin{cases} -\arctan \frac{2\zeta\omega_n\omega}{\omega_n^2 - \omega^2} & \text{dla } 0 < \omega < \omega_n \\ -\frac{\pi}{2} & \text{dla } \omega = \omega_n, \\ -\pi + \arctan \frac{2\zeta\omega_n\omega}{\omega^2 - \omega_n^2} & \text{dla } \omega > \omega_n \end{cases}$$

Przebieg charakterystyki jest, pomimo tak złożonego opisu, regularny. Charakterystyki logarytmiczne członu oscylacyjnego, dla $k=10$, $\omega_n=1$ i trzech różnych wartości współczynnika tłumienia ζ przedstawiono na rys. 7.9. Widac, że dla małych wartości ζ , $\zeta < 1$, przebieg obu charakterystyk znacznie odbiega od charakterystyk asymptotycznych – nie są one w tym przypadku użyteczne. Natomiast charakterystyki amplitudowo-fazowe (hodografy, krzywe Nyquista) przedstawia Rys. 7.10, dla $0 \leq \omega < +\infty$.



Rys. 7.10. Hodografy członu oscylacyjnego dla $0 \leq \omega < +\infty$ ($k = 10$, $\omega_n = 1$, $\zeta \leq 0.7$).

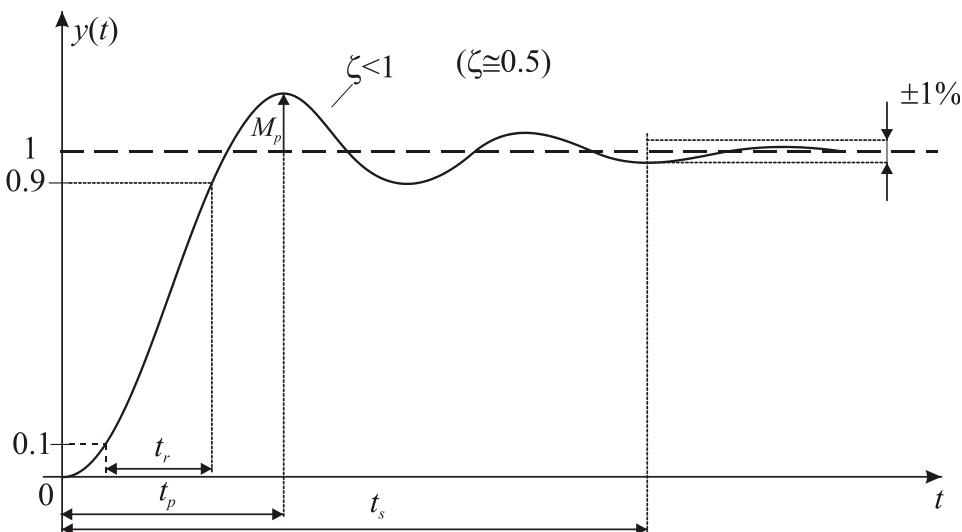
Przypomnijmy, że odpowiedź członu oscylacyjnego na skok jednostkowy $u(t) = \mathbf{1}(t)$ ma postać (zob rozdz.. 5.2.3):

$$h(t) = k \left[1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \cdot \sin(\omega_d t + \varphi) \right], \quad \varphi = \arctg \frac{\sqrt{1-\zeta^2}}{\zeta} \quad (7.6)$$

Założymy dla uproszczenia, że $k = 1$ i przeanalizujmy odpowiedź układu na skok jednostkowy

(inaczej ujmując, unormowanie odpowiedzi skokowej do jedności można rozważać jako wejście w postaci skoku o amplitudzie $1/k$). Na tej odpowiedzi zaznaczamy charakterystyczne wartości:

- t_s – tzw. czas ustalania (ang. *settling time*), tj. czas, po którym odpowiedź osiąga na stałe wartości różniące się od wartości ustalonej mniej (co do wartości bezwzględnej) niż ustalony poziom, np. 1% lub 5%,
- M_p – tzw. maksymalne (pierwsze) przeregulowanie, wartość M_p odniesioną do wzmacnienia nazywamy przeregulowaniem χ (ang. *overshoot*), $\chi = M_p \cdot 100\%$, jeśli M_p jest pikiem odpowiedzi skokowej unormowanej do jedności (Rys. 7.11)
- t_p – czas do osiągnięcia wartości maksymalnej, tj. szczytu pierwszego przeregulowania (ang. *peak time*),
- t_r – tzw. czas przestawiania lub wzrostu (ang. *rise time*).



Rys. 7.11. Odpowiedź (unormowana do jedności) członu oscylacyjnego ($\zeta < 1$).

Warto podać przykładowe zależności powyższych wielkości charakteryzujących odpowiedź na skok jednostkowy od parametrów modelu obiektu. Dla typowych wartości tłumienia, powiedzmy pomiędzy 0,3 a 0,7 otrzymujemy następujące zależności:

$$t_r \approx \frac{1,8}{\omega_n}, \quad \chi = e^{-\frac{\pi\zeta}{\sqrt{1-\zeta^2}}} \cdot 100\%, \quad t_s = \frac{4,6}{\zeta\omega_n} \quad (7.7)$$

Dla $\zeta=0,5$ otrzymujemy, w szczególności:

$$t_r = \frac{1,8}{\omega_n}, \quad \chi \approx 16\%, \quad t_s = \frac{9,2}{\omega_n} \quad (7.8)$$

Warto tu dodać jeszcze dwie liczby: $\zeta = 0,7 \rightarrow \chi \approx 5\%$, $\zeta = 0,3 \rightarrow \chi \approx 35\%$.

Przedstawione powyżej związki można wykorzystać do projektowania, do doboru parametrów układu – jeśli leży to w możliwościach i zamiarach projektanta. W szczególności,

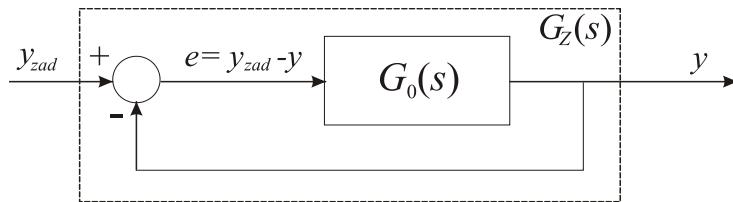
jeżeli zadane są specyfikacje określające dopuszczane przez projektanta maksymalne wartości t'_r, t'_s, χ' parametrów t_r, t_s, χ odpowiedzi skokowej, tj. wymagania są dane w postaci: $t_r \leq t'_r, t_s \leq t'_s, \chi \leq \chi'$, to należy w taki sposób dobrać wielkości ζ, ω_n , aby zachodziły następujące nierówności:

$$\omega_n \geq \frac{1,8}{t'_r} \quad \zeta \geq \zeta(\chi') \quad \zeta \omega_n \geq \frac{4,6}{t'_s},$$

gdzie $\zeta(\chi)$ oznacza funkcję odwrotną do podanej powyżej zależności χ od ζ .

7.4. Kryterium Nyquist'a, zapasy stabilności

Bardzo ważną rolę w analizie i projektowaniu układów regulacji odgrywa kryterium Nyquista; odnosi się ono do relacji układ otwarty \rightarrow układ zamknięty pętlą ujemnego sprzężenia zwrotnego o jednostkowym wzmacnieniu, taki jak pokazany na rysunku poniżej.



Rys. 7.12. Układ zamknięty ujemnym sprzężeniem zwrotnym o jednostkowym wzmacnieniu.

Postawmy następujące pytanie: czy możemy wnosić o zachowaniu się, a w szczególności stabilności i jej cechach, układu zamkniętego $G_Z(s)$ na podstawie zachowania się (charakterystyki) układu otwartego $G_0(s)$. Odpowiedź na to pytanie daje kryterium Nyquista. Wobec jego znaczenia przedstawimy wyprowadzenie tego kryterium.

Niech będzie dana reprezentująca układ otwarty transmitancja wymierna $G_0(s) = \frac{L_0(s)}{s^l \cdot M_0(s)}$,

gdzie l_0 – stopień licznika, m_0 – stopień mianownika. Wyróżniamy l -krotny biegum w zerze. Przyjmujemy następujące założenia:

$$l + m_0 > l_0, \quad \text{oraz} \quad L_0(0) \neq 0, M_0(0) \neq 0, \quad \forall \omega: M_0(j\omega) \neq 0 .$$

Transmitancja układu zamkniętego ma postać:

$$G_Z(s) = \frac{G_0(s)}{1 + G_0(s)} = \frac{L_0(s)}{L_0(s) + s^l \cdot M_0(s)} \quad (7.9)$$

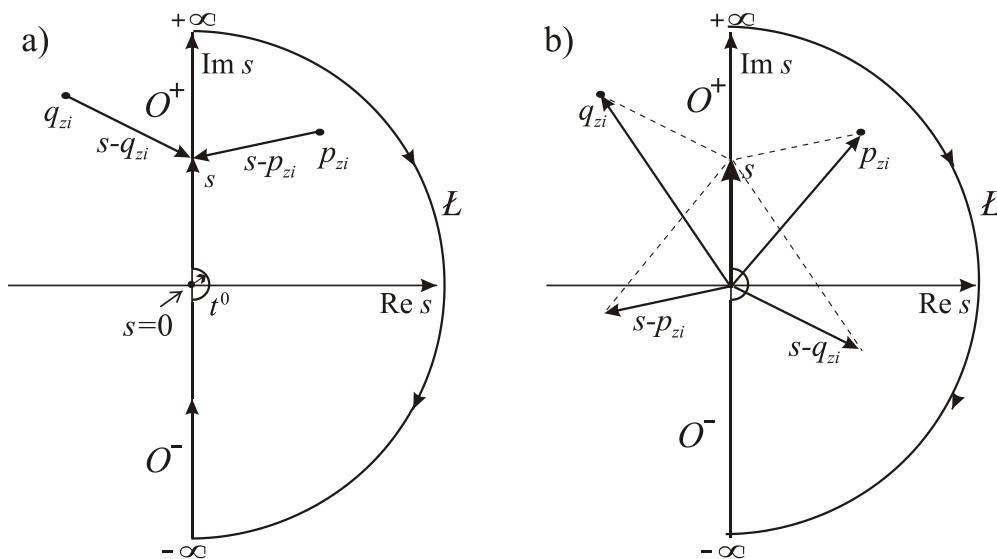
Biegumy układu zamkniętego są pierwiastkami równania $L_0(s) + s^l M_0(s) = 0$. Ponieważ $L_0(0) \neq 0, M_0(0) \neq 0$, to układ zamknięty nie ma bieguna w zerze dla każdego l . Ważną rolę w formułowaniu kryterium odgrywa wyrażenie określone mianem *różnicę zwrotną*:

$$1 + G_0(s) = \frac{L_0(s) + s^l M_0(s)}{s^l \cdot M_0(s)} = \frac{\prod_{i=1}^{n_z} (s - p_{zi}) \prod_{i=1}^{v_z} (s - q_{zi})}{\prod_{i=1}^{n_o} (s - p_{oi}) \prod_{i=1}^{v_o} (s - q_{oi})} \quad (7.10)$$

gdzie:

- n_z – liczba biegunów układu zamkniętego w prawej półpłaszczyźnie oraz, ewentualnie, na osi urojonej,
- v_z – liczba biegunów układu zamkniętego w lewej półpłaszczyźnie,
- n_o – liczba biegunów układu otwartego w prawej półpłaszczyźnie; biegun w zerze ma krotność l i nie zalicza się do prawej półpłaszczyzny, zakładamy brak biegunów układu otwartego na osi urojonej,
- v_0 – liczba biegunów układu otwartego w lewej półpłaszczyźnie, włącznie z biegunem (o krotności l) w zerze.

+ W liczniku wyrażenia określającego różnicę zwrotną występuje wielomian, którego pierwiastki są biegunami układu zamkniętego, natomiast w mianowniku występuje wielomian, którego pierwiastki są biegunami układu otwartego.



Rys. 7.13. a) Płaszczyzna zespolona z konturem zamkniętym Nyquista i z wektorami $s-p_{zi}$ i $s-q_{zi}$ zaczepionymi w punktach p_{zi} i q_{zi} ; b) konstrukcja wektorów $s-p_{zi}$ i $s-q_{zi}$ odpowiadających położeniu bieguna w prawej i lewej półpłaszczyźnie.

Wyobraźmy sobie teraz, że zmienna zespolona s przebiega wzdłuż konturu zamkniętego pokazanego na Rys. 7.13 a), zwanego konturem Nyquista: oś urojona od $-j\infty$ do $-jr$ (O^-), następnie prawy półokrąg t^0 o bardzo małym promieniu r i oś urojona od $+jr$ do $+j\infty$ (O^+) oraz prawy półokrąg L o nieskończonym promieniu (domykający kontur). Kontur jest poprowadzony tak, że nie obejmuje on bieguna położonego w zerze (jeśli wystąpią niezerowe bieguny położone na osi urojonej, to rozważania pozostaną ważne jeśli potraktujemy je jako położone w prawej półpłaszczyźnie – w przypadku takich biegunów układ dynamiczny nie jest asymptotycznie stabilny, podobnie jak w przypadku bieguna w prawej półpłaszczyźnie).

Zbadamy zakres zmian argumentu różnicy zwrotnej przy pełnym obiegu zmiennej s wokół rozpatrywanego konturu. Rozpatrujemy prawą stronę zależności określającej (7.10). Podstawę naszego rozumowania stanowi spostrzeżenie, zilustrowane na Rys. 7.13 b), że

- jeśli biegun układu zamkniętego (otwartego) p_{zi} jest położony w prawej półpłaszczyźnie, tj. wewnętrz nam konturu, to zmiana argumentu wyrażenia $s-p_{zi}$ przy pełnym obiegu

- konturu przez zmienną s jest równa -2π ,
- jeśli natomiast biegun układu zamkniętego (otwartego) q_{zi} leży w lewej półpłaszczyźnie, to odpowiednia zmiana argumentu wyrażenia $s-q_{zi}$ przy pełnym obiegu konturu przez zmienną s jest równa 0.

Spostrzeżenia te są najlepiej widoczne, jeśli zaczepimy wektory $s-p_{zi}$ i $s-q_{zi}$ w punktach, odpowiednio, p_{zi} i q_{zi} – wówczas końce tych wektorów poruszają się po konturze (rys. 7.13 a). Biorąc powyższe pod uwagę oraz uwzględniając, że zmiany argumentu czynników znajdujących się w mianowniku różnicy zwrotnej mają przeciwny znak w stosunku do zmian argumentów czynników znajdujących się w liczniku, otrzymujemy następującą zależność:

$$\Delta_{O^- \cup t^0 \cup O^+ \cup L} \arg(1 + G_0(s)) = -2n_z \pi + 2n_o \pi$$

Przypatrzmy się teraz lewej stronie tej zależności. Różnica zwrotnej na dużym łuku L przybiera wartość równą 1, stąd zmiana jej argumentu przy przebiegnięciu przez zmienną s tego łuku jest równa 0. Z kolei, jeśli mamy l -krotny biegun układu otwartego (stąd jednocześnie i różnicy zwrotnej) w zerze, to na małym łuku (półokręgu) t^0 wokół zera mamy $s^l = r^l e^{jl\varphi}$, $\varphi \in [-\pi/2, +\pi/2]$, czyli mamy zmianę argumentu różnicy zwrotnej o $-l\pi$ dla każdej wartości promienia r (dla bardzo małych r jest to zmiana argumentu w przybliżeniu wyrażenia $L_0(0)/(s^l M_0(0))$). Stąd w oparciu o obydwie powyższe zależności otrzymujemy

$$\begin{aligned} \lim_{r \rightarrow 0} \Delta_{O^- \cup t^0 \cup O^+ \cup L} \arg(1 + G_0(s)) &= \lim_{r \rightarrow 0} \Delta_{O^- \cup O^+} \arg(1 + G_0(s)) - l\pi = \\ &= \Delta_{-\infty \leq \omega \leq +\infty} \arg(1 + G_0(j\omega)) - l\pi \end{aligned}$$

Tak więc mamy ostatecznie

$$\Delta_{-\infty \leq \omega \leq +\infty} \arg(1 + G_0(j\omega)) - l\pi = 2(n_o - n_z)\pi \quad (7.11)$$

Zwróćmy uwagę, że

$$\frac{1}{s^l} = \frac{1}{r^l e^{jl\varphi}} = \text{Re}^{-jl\varphi}, \quad \text{gdzie } R = \frac{1}{r^l} \xrightarrow[r \rightarrow 0]{} \infty$$

czyli zmianie argumentu o $-l\pi$ towarzyszy nieskończony promień, stąd lewą stronę równości (7.11) należy rozumieć jako uzupełnienie hodografu $G_0(j\omega)$ wycinkiem okręgu o nieskończonym promieniu i kącie $-l\pi$ i, odpowiadającym wartości $\omega=0$ (l bieguna w zerze).

Do uzyskania asymptotycznej stabilności układu zamkniętego regulacji wymagamy, aby liczba biegunów układu zamkniętego w prawej półpłaszczyźnie była równa 0, tj. $n_z = 0$, stąd wymaganie, aby zachodził warunek, zwany **kryterium Nyquista**:

$$\Delta_{-\infty \leq \omega \leq +\infty} \arg(1 + G_0(j\omega)) - l\pi = 2n_o \pi \quad (\text{KN}) \quad (7.12)$$

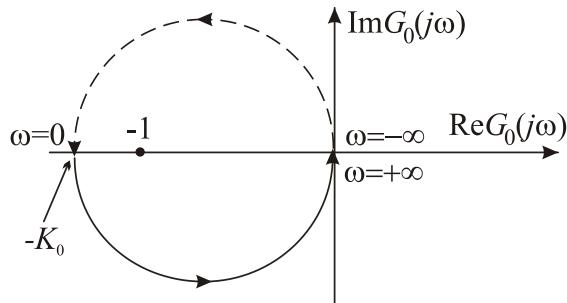
tzn. *Układ zamknięty regulacji jest asymptotycznie stabilny, gdy zmiana argumentu różnicy zwrotnej przy zmianie pulsacji ω od $-\infty$ do $+\infty$, uzupełniona o kąt $-l\pi$, jest równa $2\pi n_o$ (gdzie n_o jest liczbą biegunów układu otwartego położonych w prawej półpłaszczyźnie).*

Jeśli wektor różnicy zwrotnej $1+G_0(j\omega)$ jest zaczepiony w punkcie $(-1+j0)$, to zmiana argumentu tego wektora o 2π jest równoważna jego obrotowi o kąt 2π , czyli koniec wektora różnicy zwrotnej, wodzący po hodografie $G_0(j\omega)$, dokonuje pełnego obiegu punktu $(-1+j0)$. Dla interpretacji wpływu biegunów w zerze zauważmy ponadto, że jeśli symbolicznie oznaczyć przez $\omega=0^-$ wartość $\omega=0$ osiąganą od strony ujemnych wartości ω , a przez $\omega=0^+$ wartość $\omega=0$ osiąganą od strony dodatnich wartości ω , to wycinek okręgu o nieskończonym promieniu i kącie $-l\pi$ łączy „punkty” $\omega=0^-$ i $\omega=0^+$ hodografu – jak pokazano na Rys. 7.16 dla

$l=1$. Mając powyższe na uwadze, **kryterium Nyquista** może być sformułowane w następującej, równoważnej postaci:

Układ zamknięty jest asymptotycznie stabilny, jeśli hodograf układu otwartego $G_0(j\omega)$ wyznaczony dla zmiany pulsacji od $-\infty$ do $+\infty$ i uzupełniony wycinkiem okręgu o kącie $-l\pi$ i nieskończonym promieniu, dokonuje pełnego obiegu punktu krytycznego $(-1+j0)$ n_0 razy.

Rys. 7.14 przedstawia hodograf układu otwartego o transmitancji $G_0(s)=K_0/(sT-1)$, a więc niestabilnego (biegun w punkcie $1/T$) – ale z kryterium Nyquista wynika, że po zastosowaniu ujemnego sprzężenia zwrotnego uzyskamy układ regulacji stabilny, jeśli $K_0 > 1$ (jak na rysunku).



Rys. 7.14. Hodograf układu otwartego niestabilnego o transmitancji $G_0(s)=K_0/(sT-1)$.

Najczęściej układ otwarty $G_0(j\omega)$ jest stabilny, tj. $n_o = 0$ (nie ma biegów w prawej półpłaszczyźnie), wówczas kryterium Nyquista przyjmuje postać

$$\Delta_{-\infty \leq \omega \leq +\infty} \arg(1+G_0(j\omega)) - l\pi = 0 \quad (\text{KN}_0) \quad (7.13)$$

która jest szczególnie prosta w interpretacji: *Jeśli $n_o=0$, to układ zamknięty regulacji jest asymptotycznie stabilny, jeśli hodograf układu otwartego $G_0(j\omega)$, wyznaczony dla zmiany pulsacji od $-\infty$ do $+\infty$ i uzupełniony (domknięty) wycinkiem okręgu o kącie $-l\pi$ i nieskończonym promieniu, nie obejmuje punktu krytycznego $(-1+j0)$.*

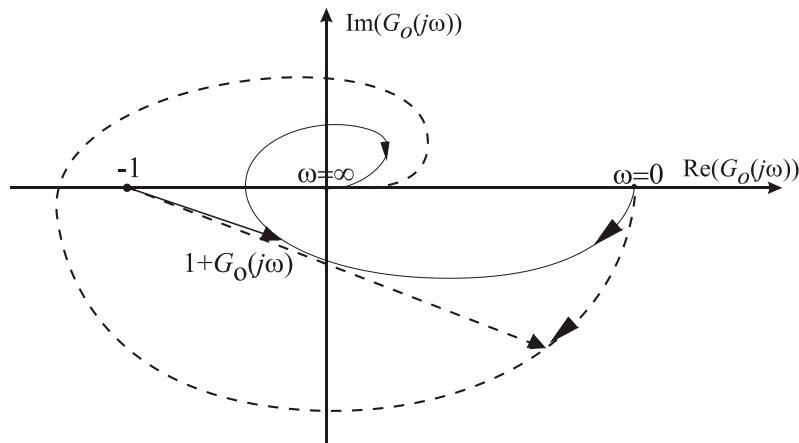
Wobec symetrii zmian argumentu względem punktu zerowego osi ujętej (zarówno biegony układu otwartego jak i biegony układu zamkniętego występują w parach zespolonych sprzężonych lub przyjmują wartości rzeczywiste) kryterium Nyquista formułuje się też dla zmiany pulsacji w zakresie wartości nieujemnych, w postaci:

$$\Delta_{0 \leq \omega \leq +\infty} \arg(1+G_0(j\omega)) - \frac{l}{2}\pi = n_o\pi \quad (\text{KN+}) \quad (7.14)$$

lub dla przypadku układu otwartego stabilnego:

$$\Delta_{0 \leq \omega \leq +\infty} \arg(1+G_0(j\omega)) - \frac{\pi}{2}l = 0 \quad (\text{KN}_0+) \quad (7.15)$$

Podobnie jak poprzednio, warunek ten możemy interpretować bezpośrednio lub jako odpowiednie obieganie (czy obejmowanie) przez hodograf układu otwartego punktu krytycznego $(-1+j0)$ przy zmianie ω od 0 do $+\infty$. Przypadek zastosowania Kryterium Nyquista w postaci (KN_0+) , dla obiektu czwartego rzędu statycznego (bez całkowania), stabilnego w układzie otwartym, przedstawia Rys. 7.15.

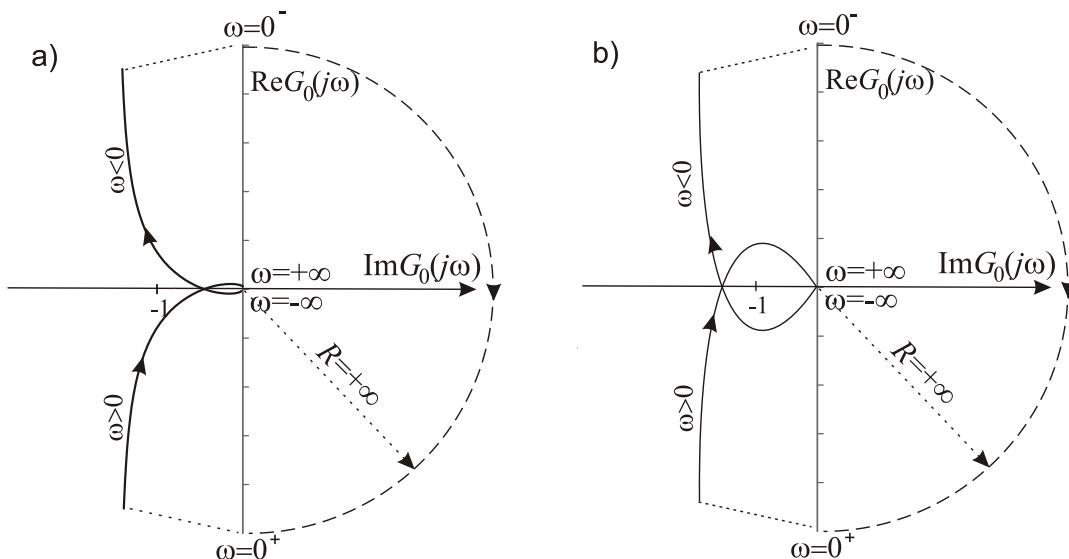


Rys. 7.15. Przebiegi hodografów transmitancji (dla $0 \leq \omega < +\infty$) układu otwartego czwartego rzędu stabilnego i bez całkowania ($l=0$) odpowiadającego stabilnemu (linia ciągła) lub niestabilnemu (linia przerwana) układowi zamkniętemu.

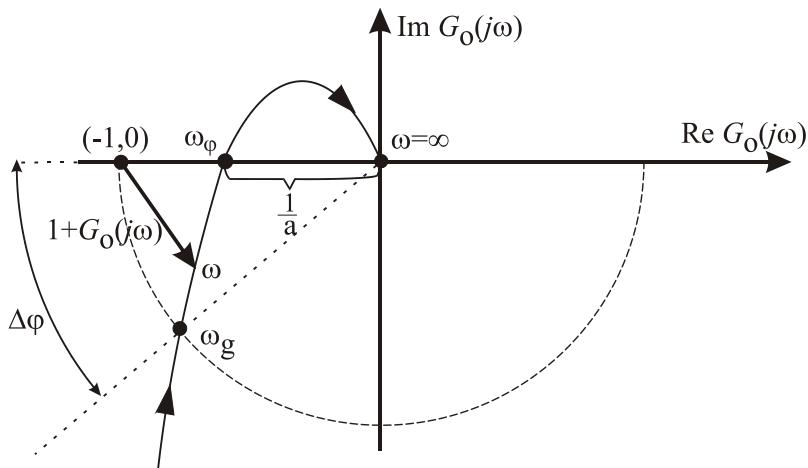
Z kolei Rys. 7.16 przedstawia pełne, hodografy układu otwartego stabilnego trzeciego rzędu, z pojedynczym biegunem w zerze, o transmitancji

$$G_0(s) = \frac{K_0}{s(Ts+1)^2}, \quad \text{gdzie } K_0 > 0, T > 0.$$

(na Rys. 7.16 a) $K_0=1$ i $T=1$). Zgodnie z kryterium Nyquista (KN₀), pełny kontur hodografu domykamy półokręgiem o nieskończonym promieniu (łączącym $\omega=0^-$ i $\omega=0^+$), wnioskując dalej o stabilności układu zamkniętego w przypadku a) i o niestabilności w przypadku b).



Rys. 7.16. Hodograf (pełny, domknięty) transmitancji układu astatycznego, o transmitancji $G_0(s) = K_0/s(Ts+1)^2$: a) układ zamknięty stabilny, b) układ zamknięty niestabilny.



Rys. 7.17. Przykładowy rzebieg hodografu transmitancji stabilnego układu otwartego w otoczeniu punktu krytycznego, układ zamknięty stabilny (obiekt z Rys. 7.16).

Na Rys. 7.17 pokazano przykładowy przebieg hodografu stabilnego układu otwartego jak i zamkniętego (Rys. 7.17 odpowiada powiększonemu fragmentowi rysunku poprzedniego, z układem trzeciego rzędu). W pokazanej na rysunku analizie przyjęte zostały następujące oznaczenia wyróżniające dwie ważne wartości pulsacji, tzw. pulsację odcięcia amplitudy ω_g oraz pulsację odcięcia fazy ω_ϕ , dla których:

$$\omega_g : |G_0(j\omega_g)| = 1 \text{ oraz, odpowiednio, } \omega_\phi : \arg G_0(j\omega_\phi) = -180^\circ.$$

Przy pomocy tych wartości możemy zdefiniować dwie ważne wielkości charakteryzujące od strony ilościowej stabilność układu zamkniętego:

$$\frac{1}{|G_0(j\omega_\phi)|} = a \quad \text{– zapas modułu} \quad (7.16)$$

$$180^\circ + \arg G_0(j\omega_g)^\circ = \Delta\varphi \quad \text{– zapas fazy} \quad (7.17)$$

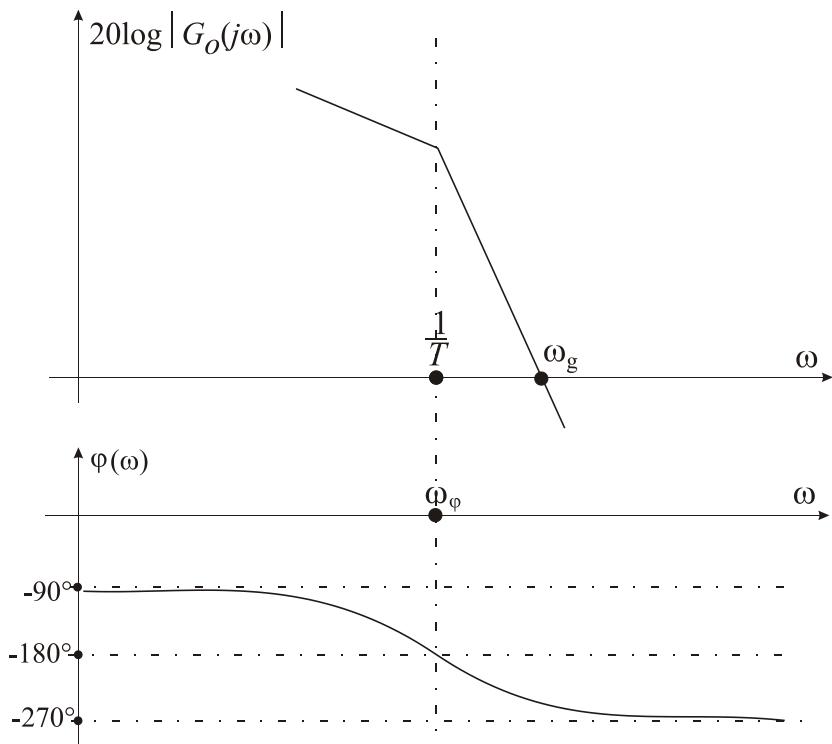
Zapas modułu a to taka wielkość, że $a \cdot |G_0(j\omega_\phi)| = 1$, tzn. pomnożenie przez a wzmacnienia (a więc i $|G_0(j\omega_\phi)|$) powoduje, że układ zamknięty znajduje się na granicy stabilności.

Odnieśmy teraz kryterium Nyquista w przypadku otwartego układu stabilnego ($n_o=0$, obiekt $G_0(s)=K_0/s(sT+1)^2$) do przebiegu logarytmicznych charakterystyk modułu i fazy. Ich przebieg odpowiadający przypadkowi układu zamkniętego niestabilnego jest pokazany na Rys. 7.18 – mamy tu relację $\omega_g > \omega_\phi$.

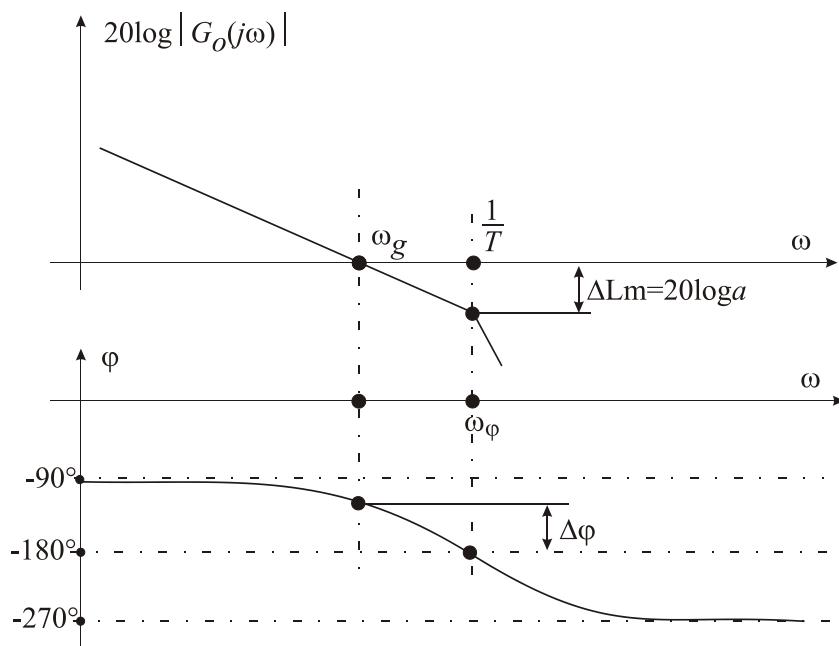
Natomiast w przypadku zamkniętego układu asymptotycznie stabilnego przebieg logarytmicznych charakterystyk modułu i fazy stabilnego układu otwartego jest taki, jak pokazany na Rys. 7.19 – mamy tu relację $\omega_g < \omega_\phi$. Na rysunku zaznaczono zapas modułu, wyrażony jako $\Delta Lm = 20\log(a)$, oraz zapas fazy $\Delta\varphi$. Interpretacja tych zapasów jest nader czytelna, jeśli pamiętać, że $a \cdot |G_0(j\omega_\phi)| = 1$ (tzn. $20\log a [\text{dB}] + 20\log |G_0(j\omega_\phi)| [\text{dB}] = 0 [\text{dB}]$).

Standardowe pożądane minimalne wartości zapasów modułu i fazy są następujące:

Zapas modułu:	6 – 8 dB
Zapas fazy:	30° – 45°



Rys. 7.18. Charakterystyki częstotliwościowe logarytmiczne stabilnego układu otwartego; układ zamknięty niestabilny ($\omega_g > \omega_\varphi$).



Rys. 7.19. Charakterystyki częstotliwościowe logarytmiczne stabilnego układu otwartego; układ zamknięty stabilny ($\omega_g < \omega_\varphi$).

Kryterium Nyquista stosuje się także w przypadku obiektu opisanego transmitancją o postaci:

$$G_0(s) = \frac{L_0(s)}{s^l \cdot M_0(s)} e^{-sT_0}$$

tj. obiektu z opóźnieniem. Człon opóźniający o transmitancji e^{-sT_0} wnosi oczywiście dodatkowe ujemne przesunięcie fazowe, równe $-\omega T_0$ dla pulsacji ω . Przesunięcie to, dla większych wartości opóźnienia, znacząco utrudnia projektowanie układu regulacji o odpowiednim stopniu stabilności.

Przykład 7.1.

Rozważamy obiekt o transmitancji

$$G_0(s) = \frac{4}{s(1+s)(1+0.2s)}$$

Korzystając z charakterystyk logarytmicznych asymptotycznych należy zaproponować korektę wzmacnienia K_0 tak, aby $\Delta\varphi \geq 30^\circ$, $\Delta Lm \geq 6\text{dB}$.

Dekomponujemy transmitancję na człony elementarne: wzmacniająco-całkujący i dwa inercyjne o jednostkowym wzmacnieniu:

$$G_0(s) = \frac{4}{s(1+s)(1+0.2s)} = \frac{4}{s} \cdot \frac{1}{1+s} \cdot \frac{1}{1+0.2s}$$

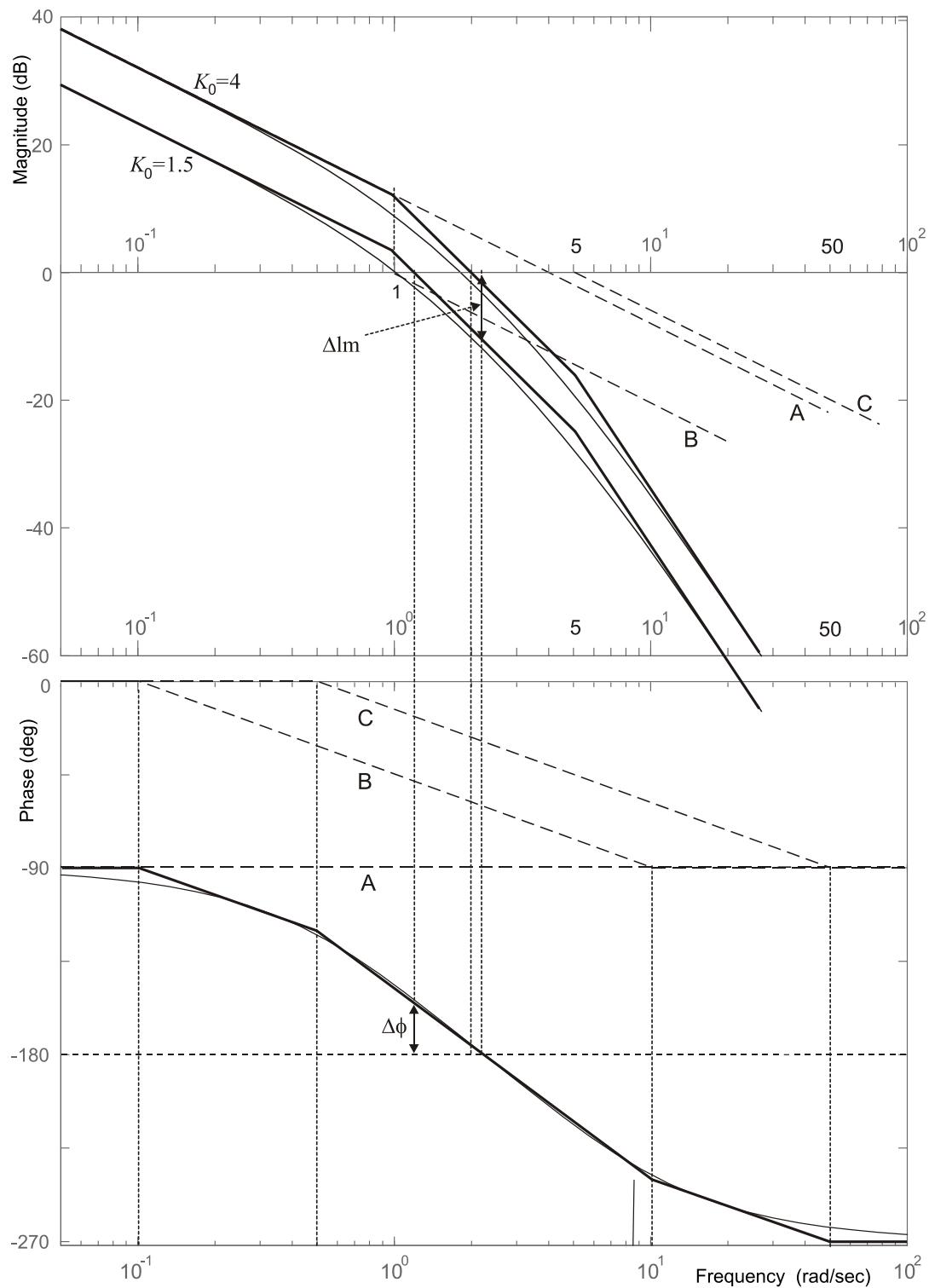
Charakterystyki częstotliwościowe (Bodego) są sumami charakterystyk składowych członów elementarnych

$$\begin{aligned} 20\log|G_0(j\omega)| &= 20\log\left|\frac{4}{j\omega}\right| + 20\log\left|\frac{1}{1+j\omega}\right| + 20\log\left|\frac{1}{1+0.2j\omega}\right| = A + B + C \\ \arg G_0(j\omega) &= -\frac{\pi}{2} + \arg\left(\frac{1}{1+j\omega}\right) + \arg\left(\frac{1}{1+0.2j\omega}\right) = A + B + C \end{aligned}$$

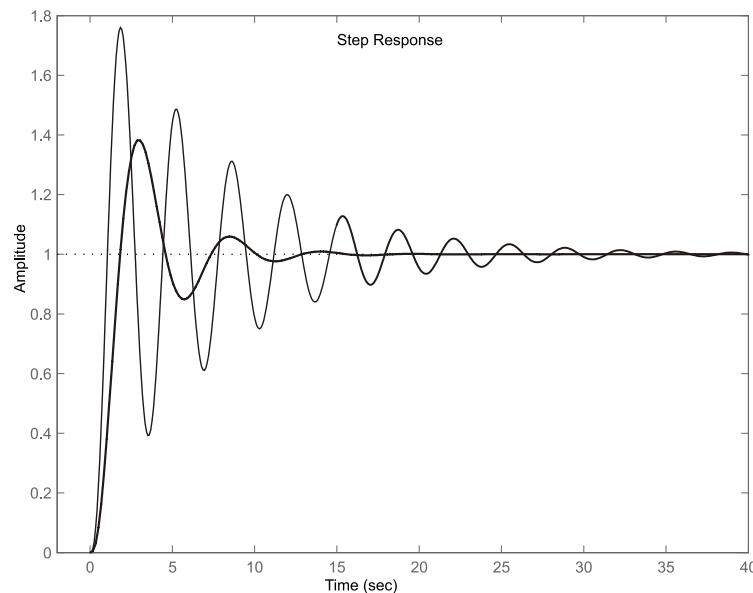
Na tej podstawie wykreślmy charakterystyki logarytmiczne asymptotyczne modułu i fazy, jak to przedstawiono na Rys. 7.20. Są one sumami charakterystyk trzech członów składowych pokazanych liniami przerywanymi, oznaczonych jako A, B i C (patrz wzory powyżej).

Widzimy, że układ ma bardzo mały zapas fazy – kilka stopni (dokładne sprawdzenie daje 9°), jak też zbyt mały zapas modułu. Zmniejszenie wzmacnienia z wartości 4 do 1.5 powinno zapewnić spełnienie wymagań – z charakterystyk asymptotycznych wynika, że zapas fazy powinien być wówczas ok. 30° , a zapas modułu wyraźnie większy od 6 dB (dokładne sprawdzenie pokazuje, że są to wartości 32.7° i 12 dB). Dla porównania, na rysunku przedstawiono również logarytmiczne charakterystyki dokładne (cieńszymi liniami) – widać dużą dokładność, użyteczność charakterystyk asymptotycznych zarówno amplitudowej jak i fazowej (w przykładzie nie występuje człon oscylacyjny o małym tłumieniu).

Natomiast na Rys. 7.21 przedstawiono odpowiedzi skokowe układu zamkniętego (ze sprzężeniem zwrotnym), z obiektami o charakterystykach z Rys. 7.20 ($K_0 = 4$ – cieńsza linia, $K_0 = 1.5$ – grubsza linia). Zmniejszenie wzmacnienia spowodowało spadek oscylacyjności odpowiedzi skokowej i silne tłumienie oscylacji (ok. pięciokrotne) – przebieg szybko się stabilizuje, ale przeregulowanie jest jeszcze duże (ok. 40%). Zwiększenie zapasu fazy, np. do 45° , spowoduje spadek przeregulowania i dalszy spadek oscylacyjności.



Rys. 7.20. Charakterystyki częstotliwościowe logarytmiczne w przykładzie 7.1.

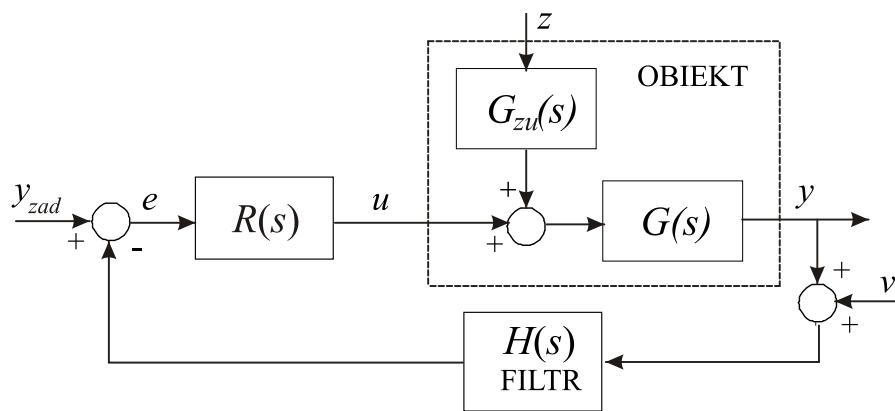


Rys. 7.21. Odpowiedzi skokowe układu zamkniętego z obiektami o charakterystykach z Rys. 7.20 w torze otwartym (cieńsza linia: $K_0 = 4$, grubsza linia: $K_0 = 1.5$).

7.5. Projektowanie układów regulacji w dziedzinie częstotliwości

7.5.1. Zadanie projektowania podstawowego układu regulacji

Projektowanie jest sztuką osiągania kompromisu. Rozważmy prosty układ regulacji pokazany na Rys. 7.22. $G(s)$ oznacza na nim transmitancję obiektu, $R(s)$ transmitancję regulatora (także zwanego, w tym przypadku, szeregowym członem korekcyjnym), $H(s)$ transmitancję filtra w gałęzi ujemnego sprzężenia zwrotnego, y_{zad} – wartość zadaną (może ona być zależna od czasu), z – zakłócenie działające na wejściu obiektu (tzw. zakłócenia obiektowe lub procesowe), v – zakłócenie pomiarowe (zwane także szumem pomiarowym).



Rys. 7.22. Układ regulacji.

Zadanie projektowania pokazanego na rysunku prostego układu regulacji dotyczy określenia (zaprojektowania) regulatora (korektora) $R(s)$ oraz, w ogólnym przypadku, filtru $H(s)$ w gałęzi ujemnego sprzężenia zwrotnego. W dalszych rozważaniach będziemy zakładać (domyślnie), że $H(s) \equiv 1$ lub, w szczególnych przypadkach, że filtr $H(s)$ został już zaprojektowany. Zadanie projektowania, które będziemy rozważać, polega na określeniu

transmitancji $R(s)$.

Wymagania stawiane projektowanemu układowi regulacji:

1. Układ zamknięty musi być asymptotycznie stabilny, przy czym potrzebne jest zapewnienie odpowiednich zabezpieczeń (np. zapasów) stabilności (por. rozdz. 7.1).
2. Sygnał $y(t)$ powinien jak najlepiej (jak najdokładniej) nadążać za przebiegiem wartości zadanej wielkości regulowanej $y_{zad}(t)$ (wielkość zadana może być zależna od czasu).
3. Należy w jak największym możliwym stopniu eliminować wpływ wielkości $z(t)$ (zakłóceń obiektowych) i $v(t)$ (zakłóceń pomiarowych) na wartość wyjścia regulowanego $y(t)$.
4. Przebiegi przejściowe (będące skutkiem różnych od zera warunków początkowych bądź wynikające z załączenia układu lub ze zmiany wartości zadanej – na przykład zmiany o charakterze skokowym) powinny mieć odpowiedni kształt.

Trzeba koniecznie podkreślić, że powyższe wymagania nie są wzajemnie zgodne. Projektowanie jest właściwie sztuką osiągania kompromisu pomiędzy tymi wymaganiami; z tym, że wymaganie asymptotycznej stabilności układu regulacji musi być zawsze spełnione, wymaganie dokładności stawiane jest zaraz po nim, na drugim miejscu.

Można łatwo wyprowadzić zależność określającą zależność transformaty wyjścia $Y(s)$ i uchybu $E(s)$ od transmitancji obiektu i regulatora, a także od transformat wielkości $y_{zad}(t)$, $w(t)$ i $v(t)$. Istotnie, podstawowe zależności są następujące:

$$\begin{aligned} Y(s) &= G(s)(R(s)E(s) + G_{zu}(s)Z(s)) \\ E(s) &= Y_{zad}(s) - Y(s) - V(s) \end{aligned}$$

Po ich prostym przekształceniu otrzymujemy poszukiwane wyrażenie:

$$Y(s) = \frac{G(s)R(s)}{1+G(s)R(s)} Y_{zad}(s) - \frac{G(s)R(s)}{1+G(s)R(s)} V(s) + \frac{G(s)G_{zu}(s)}{1+G(s)R(s)} Z(s) \quad (7.18)$$

$$E(s) = \frac{1}{1+G(s)R(s)} Y_{zad}(s) - \frac{1}{1+G(s)R(s)} V(s) - \frac{G(s)G_{zu}(s)}{1+G(s)R(s)} Z(s) \quad (7.19)$$

Wielkość $G_0(s) = R(s)G(s)$ będziemy, jak uprzednio, określać jako *transmitancję układu otwartego* (toru otwartego).

Jak łatwo zauważyc, $Y(s)$ (jak i $E(s)$) w taki sam sposób zależy od transformaty wartości zadanej jak i od transformaty zakłócenia pomiarowego $V(s)$ (wynika to też bezpośrednio z Rys. 7.22). Natomiast transmitancja od wejścia reprezentującego zakłócenie procesowe z do wyjścia regulowanego y jest inna i wyraźnie różna od transmitancji głównej układu zamkniętego (od y_{zad} do y).

Warto wyróżnić dwie ważne różniące się sytuacje:

- *Projektowanie serwomechanizmu*; dany jest obiekt minimalnofazowy (bez opóźnienia i zer w prawej półpłaszczyźnie), ponadto zakładamy, że istotne jest uzyskanie odpowiedniej dokładności odtwarzania sygnałów zmiennych w czasie, skoncentrowanych w zadanym paśmie roboczym, oraz zapewnienie tłumienie zakłóceń pomiarowych v ; wpływ zakłóceń procesowych z jest w zasadzie pomijalny.
- *Projektowanie układu regulacji przemysłowej*; dany jest obiekt minimalnofazowy lub, częściej, nieminimalnofazowy, ponadto istotne jest uzyskanie odpowiednio małego (bądź zerowego) uchybu w stanie ustalonym oraz możliwie najlepsze tłumienie zakłóceń procesowych z ; wpływ zakłóceń pomiarowych v jest najczęściej pomijalnie mały.

7.5.2. Projektowanie serwomechanizmu

W tym przypadku przyjmowane są następujące założenia:

- Pomijamy wpływ zakłóceń obiektowych $z(t)$.
- Przyjmujemy, że sygnały określające wartość zadaną $y_{zad}(t)$ mają składowe sinusoidalne o pulsacjach dodatnich ograniczonych poprzez wartość ω_r – określamy te pulsacje jako pulsacje robocze; sygnały zadane przedstawiamy w takiej sytuacji jako $y_{zad}(t) = R_0 \sin(\omega t)$, gdzie $\omega \leq \omega_r$.
- Sygnały określające zakłócenia pomiarowe $v(t)$ także mają przebiegi sinusoidalne: $v(t) = V_0 \sin(\omega_v t)$, gdzie $\omega_v \geq \omega_r$, zaś ω_v oznacza kres dolny pasma zakłóceń pomiarowych,
- Transmitancja obiektu $G(s)$ ma postać $G(s) = \frac{L(s)}{s^l M(s)}$, ponadto nie ma ona zer i biegów w prawej półpłaszczyźnie, tzn. zarówno wielomian licznikowy $L(s)$ nie ma pierwiastków (zwanych zerami transmitancji) w prawej półpłaszczyźnie jak i wielomian $M(s)$ nie ma pierwiastków (biegów transmitancji) w prawej półpłaszczyźnie. W takiej sytuacji charakterystyka fazowa ma, co do wartości bezwzględnej, przebieg minimalnie odległy od zera spośród wszystkich transmitancji posiadających taki sam kształt charakterystyki modułu (stąd nazwa: *obiekt minimalnofazowy*).

Sensowny układ sterowania da się zaprojektować wówczas, gdy: $\omega_v > \omega_r$, tzn. wtedy, gdy występuje rozdzielnosc pasm sygnałów zakłócających i sygnałów zadanych. Gdy te pasma na siebie wyraźnie nachodzą, konieczne jest poprawienie urządzeń pomiarowych; trzeba je tak zmienić, aby pasmo zakłóceń nie pokrywało się z pasmem pulsacji roboczych.

Standardowe wymagania projektowe są następujące:

1. *Stabilność*: bezwzględnie wymagamy uzyskanie asymptotycznej stabilności oraz zapewnienie odpowiednich zapasów modułu i fazy; przeważnie wymagamy, aby zapas modułu $\Delta L_m \geq 6$ dB, zapas fazy $\Delta\phi \geq 45^\circ$.
2. *Dokładność*: może być wyrażona jako wymaganie dokładności nadążania, w postaci żądania, aby $|E(j\omega)| \leq d|Y_{zad}(j\omega)|$, gdzie d jest odpowiednio małą liczbą dodatnią (na przykład $d = 0,01$, co odpowiada dokładności 1%), dla wszystkich wartości pulsacji w paśmie roboczym, tj. dla $0 \leq \omega \leq \omega_r$. Może temu towarzyszyć także wymaganie zapewnienia zerowego uchybu w stanie ustalonym. Analizując wymaganie dokładności nadążania wychodzimy z zależności

$$E(j\omega) = \frac{1}{1 + G(j\omega)R(j\omega)} Y_{zad}(j\omega) \quad (7.20)$$

z której wynika, że spełnienie wymagania $|E(j\omega)| \leq d|Y_{zad}(j\omega)|$ prowadzi do warunku

$$\left| \frac{1}{1 + G(j\omega)R(j\omega)} \right| \leq d \quad \Rightarrow \quad |1 + G(j\omega)R(j\omega)| \geq \frac{1}{d}, \quad \text{dla } \omega \leq \omega_r$$

Dla dużych wartości $1/d$ (np. $1/d=100$) możemy powyższą nierówność przybliżyć wymaganiem odpowiednio dużej wartości modułu transmitancji układu otwartego w paśmie roboczym.

$$|G(j\omega)R(j\omega)| \geq \frac{1}{d}, \quad \text{dla } \omega \leq \omega_r \quad (7.21)$$

3. *Tłumienie zakłóceń pomiarowych:* należy zapewnić jak najlepsze tłumienie zakłóceń $v(t)$ dla $\omega \geq \omega_v$ – oznacza to, zgodnie z (7.18), że pożądana jest jak najmniejsza wartość wzmocnienia w torze otwartym $|G(j\omega)R(j\omega)|$ dla $\omega \geq \omega_v$.

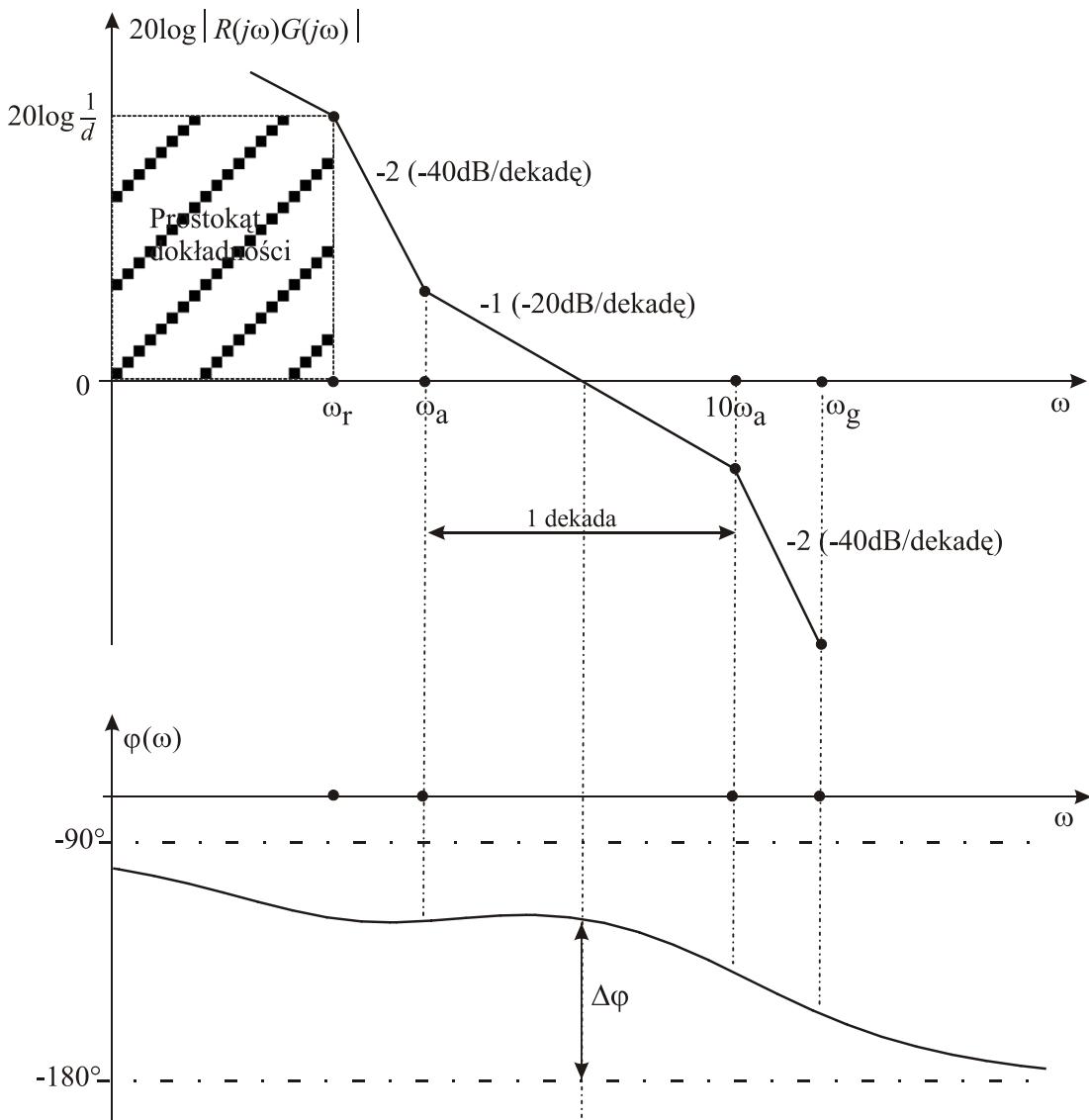
Projektowanie w takiej sytuacji polega na kształtowaniu przebiegu logarytmicznej charakterystyki modułu toru otwartego $20\log|G(j\omega)R(j\omega)|$ (tj. charakterystyki amplitudowej) przy pomocy odpowiedniego doboru transmitancji regulatora $R(s)$.

Posługujemy się asymptotyczną charakterystyką modułu; jej załamania odpowiadają punktom, w których zaczynają odgrywać rolę kolejne człony iloczynu określającego transmitancję, tj. czynniki typu $(1+jT\omega)$ lub $1/(1+jT\omega)$ (mogą także występować człony oscylacyjne; wówczas ich uwzględnienie wymaga większej uwagi, por. rozdz. 7.1). Przypomnijmy, że:

- w przypadku wpływu wyrażenia $(1+jT\omega)$ na przebieg logarytmicznej charakterystyki modułu przyjmujemy, że dla $\omega < 1/T$ logarytm modułu tego wyrażenia, który dodajemy do logarytmów pozostałych członów transmitancji, jest równy 0, natomiast dla $\omega > 1/T$ logarytm tego modułu, pomnożony przez wartość 20, tj. $20\log(\omega T)$ narasta liniowo, przy wzroście ω , od zera w góre z intensywnością (nachyleniem) 20dB na dekadę (dla $\omega = 10/T$ mamy $20\log(\omega T) = 20 \log(10) = 20$ dB). Nabylenie to oznaczamy jako +1.
- W przypadku działania członu $1/(1+jT\omega)$ przekroczenie wartości $\omega = 1/T$ powoduje zmniejszanie wartości logarytmu całej charakterystyki modułu (odejmujemy od niej $20\log(\omega T)$) z nabyleniem -1, tj. minus 20dB na dekadę.

Pożądany przebieg charakterystyki amplitudowej układu otwartego, pokazany na Rys. 7.23, jest następujący:

1. Logarytmiczna charakterystyka modułu układu otwartego musi przechodzić ponad tzw. prostokątem dokładności o wysokości $20\log(1/d)$, sięgającym na osi pulsacji do wartości ω_r , po to, by dla pulsacji $\omega \leq \omega_r$ spełnione było wymaganie $|G(j\omega)R(j\omega)| \geq \frac{1}{d}$.
2. Następnie, możliwie jak najbliżej wartości ω równej ω_r , nabylenie charakterystyki powinno osiągać wartość -2; zależy nam bowiem na tym, aby wzmocnienie toru otwartego malało dla pulsacji wyższych niż ω_r możliwie szybko w celu tłumienia zakłóceń pomiarowych.
3. Jednakże, musimy zapewnić asymptotyczną stabilność układu zamkniętego. W tym celu w okolicy pulsacji odcięcia amplitudy ω_g , gdzie $20\log|G(j\omega_g)R(j\omega_g)| = 0$, należy zapewnić nabylenie charakterystyki modułu równe -1, przez dekadę; spowoduje to, że przesunięcie fazy będzie w środku tego rejonu zbliżało się do $-\pi/2$ (do -90°). Nabylenie równe -1 musimy utrzymać na odcinku odpowiadającym mniej więcej jednej dekadzie, tj. od $\omega = \omega_a$ do $\omega = 10\omega_a$, przy czym należy zapewnić, aby, w przybliżeniu, $20\log|G(j\omega_a)R(j\omega_a)| \approx 8 \div 10$ dB.
4. Po osiągnięciu przez pulsację wartości równej $10\omega_a$ nabylenie charakterystyki modułu powinno znowu przyjąć wartość równą -2, aby nadal szybko zmniejszać wzmocnienie toru otwartego i tym samym zapewnić jak najlepsze tłumienie zakłóceń charakteryzujących się wysokimi częstotliwościami.



Rys. 7.23. Kształt charakterystyk modułu i fazy ukształtowanych zgodnie z zaleceniami projektowymi dla serwomechanizmów.

Przykład 7.2.

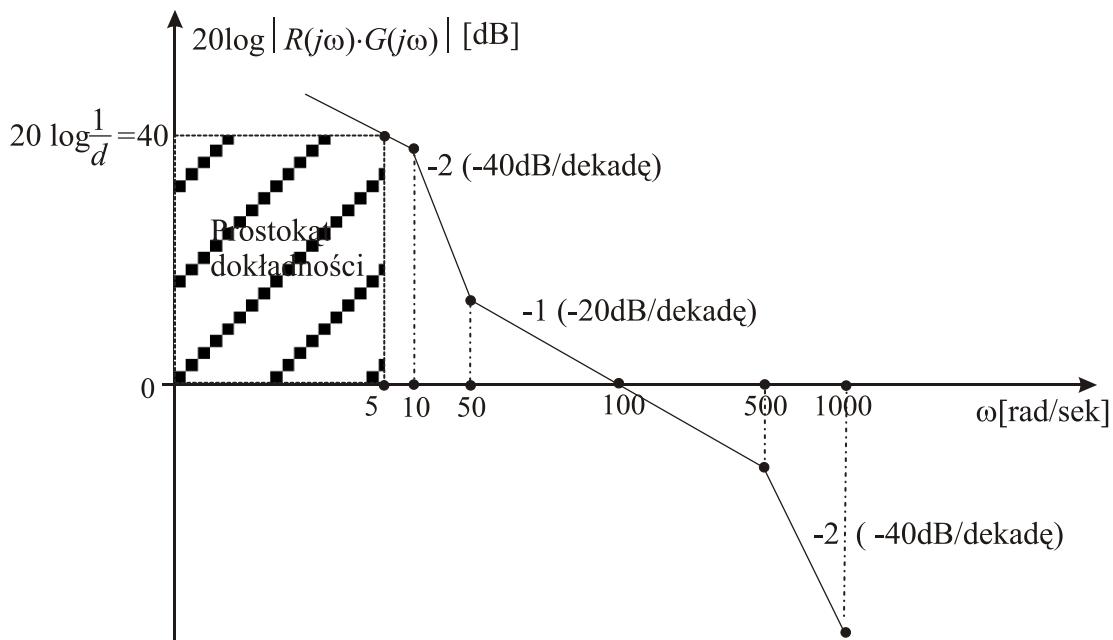
Przyjmijmy, że transmitancja obiektu dana jest w postaci:

$$G(s) = \frac{4}{s(0,1s + 1)}$$

Wymagania projektowe:

1. stabilność asymptotyczna, zapasy, odpowiednio: $\Delta\text{Im} \geq 6 \text{ dB}$, $\Delta\varphi \geq 45^\circ$,
2. uzyskanie zerowego uchybu w stanie ustalonym oraz dokładność nadążania określona przez wartość $d=0,01$ (1% dokładności) przy $\omega_r = 5 \text{ rad/sek}$,
3. tłumienie zakłóceń pomiarowych dla $\omega_v = 1000 \text{ rad/sek}$.

Odpowiednio kształtuujemy przebieg charakterystyki asymptotycznej modułu układu otwartego, jak to zilustrowano na Rys. 7.24.



Rys. 7.24. Projekt przebiegu charakterystyki układowu otwartego dla obiektu $G(s)=1/(s(1+0,1s))$.

W tym celu, po pierwsze zauważamy, że obiekt cechuje się astatyzmem (występuje w nim całkowanie). Przewidujemy więc transmitancję regulatora w postaci $R(s) = K \cdot D_{kor}(s)$, gdzie działanie członu korektora odpowiadające transmitancji $D_{kor}(s)$ nie powinno manifestować się w paśmie roboczym, tj. powinno być $|D_{kor}(j\omega_r)| = 1$. Następnie dobieramy taką wartość wzmacnienia K regulatora, aby dla pulsacji $\omega_r = 5$ zapewnić przejście charakterystyki modułu ponad prostokątem dokładności. Ponieważ

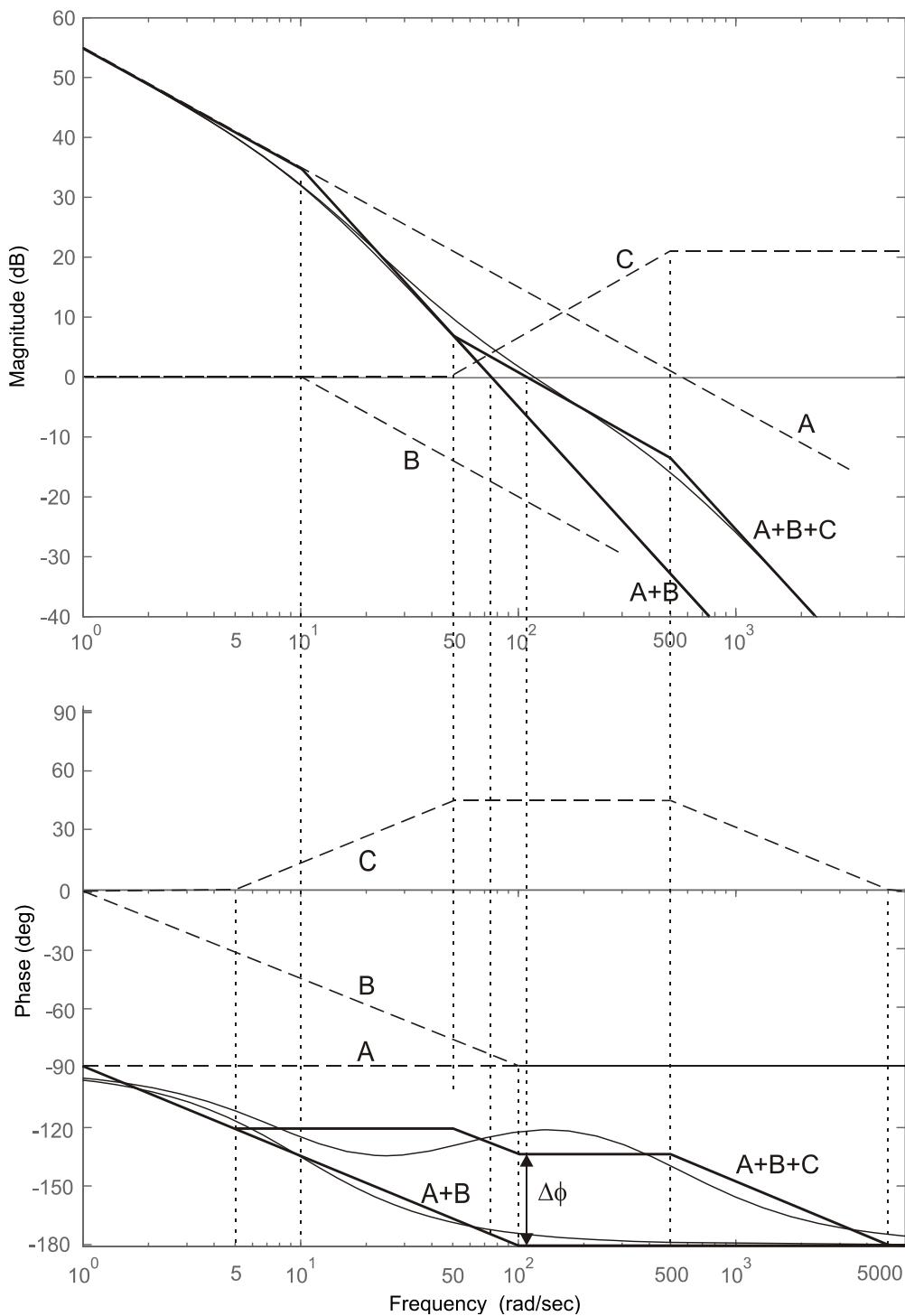
$$|G(j5)| = \frac{4}{5} \cdot \frac{1}{\sqrt{1,25}} = \frac{40}{55},$$

to prawidłowa wartość wzmacnienia K wynosi $100/(40/55)$, tj. można przyjąć $K = 140$. Warto zauważyc, że w rozważanym przypadku charakterystyka modułu z regulatorem proporcjonalnym ulega już przy pulsacji równej 10 przejściu do nachylenia -2; nie warto zatem odchyłać jej w dół przy pulsacji 5 po to, by potem znowu korygować to nachylenie (w przeciwnym kierunku) w punkcie $\omega = 10$. Opadająca z nachyleniem -2 charakterystyka modułu uzyskuje poziom 8 dB przy pulsacji zbliżonej do $\omega = 50$. Zgodnie z zaleceniami projektowymi należy w tym miejscu „włączyć” człon o postaci $(1+0.02s)$, aby uzyskać nachylenie -1 przy przejściu charakterystyki przez 0 dB. Pulsacja odcięcia przyjmie wówczas wartość ok. 100. Po utrzymaniu nachylenia -1 na odcinku jednej dekady, tj. do wartości pulsacji $\omega = 500$, charakterystyka modułu powinna zacząć ponownie opadać z nachyleniem -2; w tym celu trzeba zapewnić uaktywnienie członu o postaci $1/(1+Ts)$, w naszym przypadku członu $1/(1+0.002s)$. Dzięki temu po przekroczeniu pulsacji $\omega_v = 1000$ uzyskamy tłumienie zakłóceń pomiarowych co najmniej o 28 dB, tj. dwudziestopięciokrotne.

Uzyskany regulator ma ostatecznie postać częstokroć stosowanego w takich przypadkach realizowalnego regulatora PD (proporcjonalno-różniczkującego)

$$R(s) = 140 \frac{0.02s + 1}{0.002s + 1}.$$

Regulatory tego rodzaju spotykamy w wielu serwomechanizmach położeniowych.



Rys. 7.25. Charakterystyki częstotliwościowe logarytmiczne toru otwartego w przykładzie 7.2.

Transmitancja toru otwartego układu z korekcją i jej charakterystyka widmowa logarytmiczna amplitudowa mają postacie

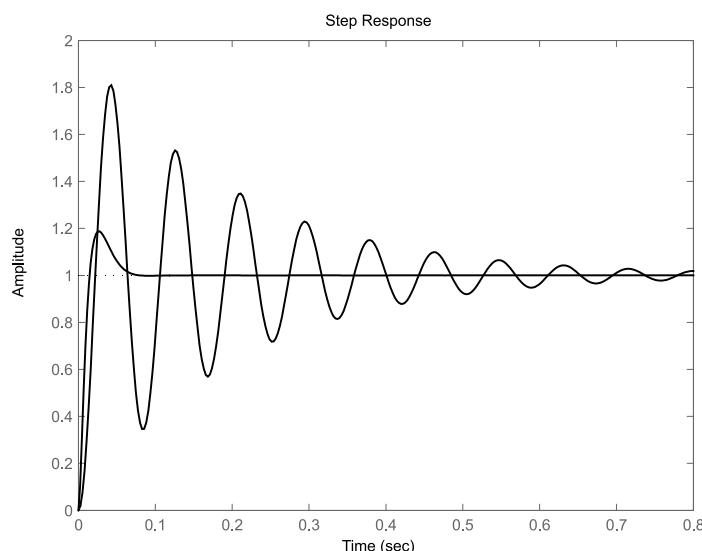
$$G_0(s) = \frac{0.02s + 1}{s(0.1s + 1)(0.002s + 1)} = \frac{560}{s} \cdot \frac{1}{0.1s + 1} \cdot \frac{0.02s + 1}{0.002s + 1}$$

$$20 \log |G_0(j\omega)| = 20 \log \frac{560}{j\omega} + 20 \log \left| \frac{1}{0.1j\omega + 1} \right| + 20 \log \left| \frac{0.02j\omega + 1}{0.002j\omega + 1} \right| = \mathbf{A} + \mathbf{B} + \mathbf{C}$$

przedstawione jako składająca się ze standardowych członów: proporcjonalno-całkującego, inercyjnego i proporcjonalno-różniczkującego rzeczywistego. Charakterystyki podanej wyżej transmitancji, jak i transmitancji przed korekcją członem proporcjonalno-różniczkującym pokazane są na Rys. 7.25 – charakterystyki asymptotyczne liniami ciągłymi pogrubionymi, a dokładne liniami cieńskimi. Pokazano również, liniami przerywanymi, charakterystyki asymptotyczne członów składowych.

Charakterystyki amplitudowe asymptotyczne są bardzo zbliżone do dokładnych, natomiast dokładność przybliżenia charakterystyki fazowej charakterystyką asymptotyczną jest mniejsza, niemniej nadal użyteczna. Zapasy fazy (dokładne) dla układów przed korekcją dynamiki (członem $D(s) = (0.02s+1)/(0.002s+1)$) i po korekcji wynoszą, odpowiednio, 7.6° i 58.6° . Zapas fazy na charakterystyce asymptotycznej (zaznaczony na rysunku) jest nieco mniejszy (błąd ok. dziesięć stopni).

Na rysunku 7.26 pokazano odpowiedzi skokowe układów regulacji z transmitancjami toru otwartego odpowiadającymi charakterystykom częstotliwościowym przedstawionym na rysunku poprzednim.



Rys. 7.26. Odpowiedzi skokowe układów regulacji z $G_0(s)$ z rysunku poprzedniego.

Zadania

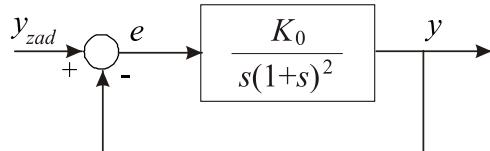
Zadanie 7.1. Narysuj charakterystyki częstotliwościowe asymptotyczne układów dynamicznych o transmitancjach

a) $\frac{1}{s(1+10s)^2}$ b) $\frac{0.01}{s(s+0.1)^2}$ c) $\frac{0.01}{s(s+0.01)(5s+1)}$ d) $\frac{100(1+s)}{s(1+10s)(1+0.02s)^2}$

oraz oszacuj zapasy stabilności dla każdego z tych układów.

Zadanie 7.2. Dla układu dynamicznego o transmitancji $G(s) = \frac{-1}{(s+1)(s+2)}$ wyznacz odpowiedź układu w stanie ustalonym (moduł i przesunięcie fazowe sygnału wyjściowego) na wymuszenie sinusoidalne o jednostkowej amplitudzie i pulsacji $\omega = 1$.

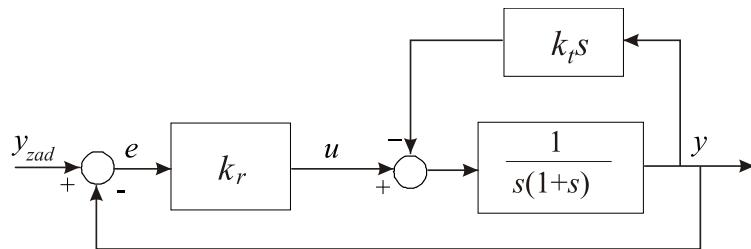
Zadanie 7.3. Dla układu regulacji



wyznacz wartość wzmocnienia $K_0 = K_{0kr}$, przy którym układ traci stabilność.

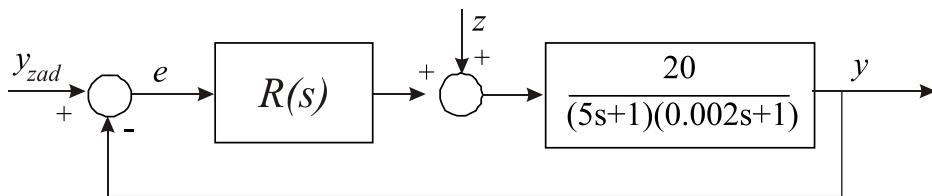
Przyjmując $K_0 = K_{0kr}/2$ oszacuj zapas fazy uzyskanego układu, korzystając z charakterystyk logarytmicznych modułu i fazy.

Zadanie 7.4. Dla układu regulacji



z obiektem objętym dodatkowym (wewnętrzny) sprzężeniem prędkościowym dobrać wzmocnienia k_r i k_t tak, aby ustalanie się przebiegów przejściowych było aperiodyczne (tzn. bez oscylacji) i aby uchyb ustalony dla $y_{zad}(t) = at\mathbf{1}(t)$ (a jest stałą) nie przekraczał 10% wartości amplitudy a . Dla wybranego zestawu parametrów należy wykreślić charakterystyki częstotliwościowe logarytmiczne (asymptotyczne) transmitancji układu otwartego.

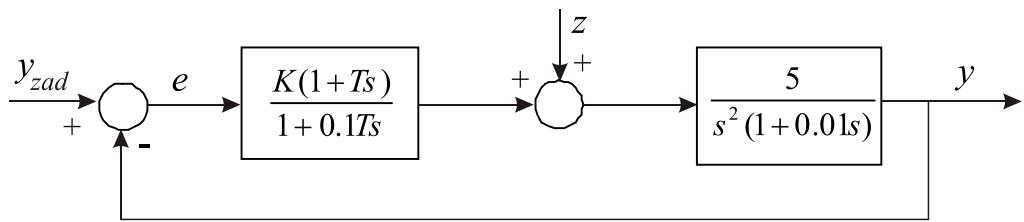
Zadanie 7.5. W układzie regulacji



dobierz możliwie prostą transmitancję regulatora $R(s)$ w taki sposób, aby zapewnić:

- odtwarzanie wartości zadanej $y_{zad}(t) = r_0 \sin(\omega t)$, dla $0 < \omega \leq 0,2$ z uchybem poniżej 0,1% oraz stabilność układu z zachowaniem wymagań projektowych dla serwomechanizmów.
- dodatkowo zerowy uchyb w stanie ustalonym dla zakłóceń stałych przez dłuższe odcinki czasu (skokowych).

W każdym z przypadków a) i b), zapas fazy powinien być co najmniej 45° .

Zadanie 7.6. W układzie regulacji

dobierz wzmocnienie regulatora K i wartość stałej T w taki sposób, aby zapewnić:

- odtwarzanie wartości zadanej $y_{zad}(t) = r_0 \sin(0.1t)$ z uchybem poniżej 1%,
- stabilność układu z zachowaniem wymagań projektowych dla serwomechanizmów.

Zadanie 7.7

W układzie regulacji z zadania 7.6 wyznacz zakres dodatnich wartości parametrów K i T , dla których układ ten jest asymptotycznie stabilny. Porównaj odległość zaprojektowanej w zadaniu 7.6 wartości K od wartości K_{kr} odpowiadającej wyznaczonej granicy stabilności (przy stałej zaprojektowanej w zadaniu 7.6 wartości T) z wyznaczonym w zadaniu 7.6 zapasem modułu.

8. Regulacja przemysłowa

8.1. Zadanie regulacji przemysłowej

Układami regulacji przemysłowej nazywamy układy regulacji takich wielkości jak temperatura, ciśnienie, przepływ, poziom, itp. występujące przede wszystkim w zakładach przemysłowych, zakładach energetycznych, itp. – w odróżnieniu od serwomechanizmów, gdzie wielkość regułowana ma charakter przesunięcia liniowego lub kątowego. Układy regulacji typu przemysłowego występują również w naszym bezpośrednim otoczeniu, np. układ regulacji temperatury w mieszkaniu czy domu jednorodzinnym, układ regulacji w węźle cieplnym budynku wielorodzinnego, układy regulacji ciśnień czy przepływów w miejskich sieciach wodociągowych, oczyszczalniach ścieków, itp.

Typowe są następujące cechy zadań regulacji przemysłowej:

- Sygnał wielkości zadanej jest najczęściej stały w dłuższych odcinkach czasu (tzw. zadnie regulacji stałwartościowej, zwane też zadaniem stabilizacji), o wartości zmienianej rzadko, np. skokowo przez operatora w celu dostosowania pracy układu do zmienionych wymagań. Sygnał wartości zadanej może być też zmienny w czasie, np. w z góry zaprogramowany sposób (tzw. regulacja programowa), czy też np. jako sygnał wyjściowy regulatora nadziednego zmienny w sposób wynikający z bieżącego działania tego regulatora (regulacja nadążna).
- Z reguły zadanie *tlumienia zakłóceń* jest najistotniejsze, znacznie bardziej niż zmniejszenie wpływu szumu pomiarowego.
- Standardem jest stosowanie *regulatora PID (Proportional-Integral-Derivative)*, przy czym najczęściej dysponujemy jedynie *bardzo uproszczonym modelem obiektu* utworzonym specjalnie dla zadania doboru nastaw regulatora.

8.2. Struktury i charakterystyki regulatora PID

8.2.1. Standardowe struktury i charakterystyki regulatora PID

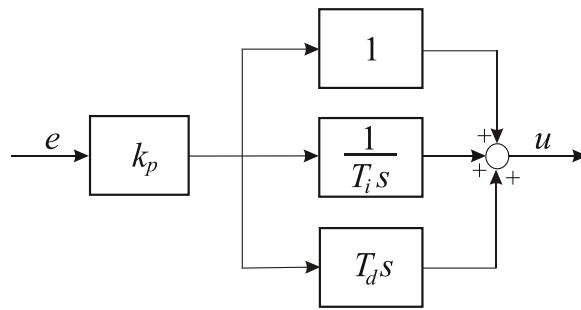
Działanie *standardowego algorytmu* regulatora PID opisuje w dziedzinie czasu następująca zależność:

$$u(t) = k_p \left(e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{de(t)}{dt} \right) \quad (8.1)$$

której odpowiada następująca postać transmitancji operatorowej:

$$R_{PID}(s) = k_p \left(1 + \frac{1}{T_i s} + T_d s \right) = k_p + \frac{k_p}{T_i s} + k_p T_d s \quad (8.2)$$

Jest to tzw. standardowa *struktura równoległa idealna regulatora PID*, gdyż działania wszystkich trzech rodzajów jego oddziaływań (proporcjonalnego, całkującego i różniczkującego) są równoległe, jak to bezpośrednio widać z postaci wzorów (8.1) czy (8.2), a co dodatkowo jest zilustrowane na Rys. 8.1.



Rys. 8.1. Standardowa równoległa struktura algorytmu PID (idealnego).

We wzorach powyższych występują trzy *parametry (nastawy) regulatora PID*:

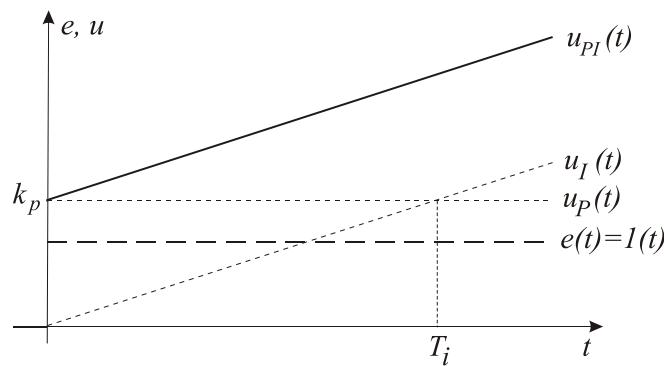
k_p – współczynnik wzmocnienia (wzmocnienie) regulatora; używany bywa też, alternatywnie w miejsce k_p , tzw. zakres proporcjonalności x_p definiowany jako

$$x_p = \frac{1}{k_p} \cdot 100\% ,$$

T_i – stała czasowa całkowania, zwana też czasem zdwojenia,

T_d – stała czasowa różniczkowania, zwana też czasem wyprzedzenia.

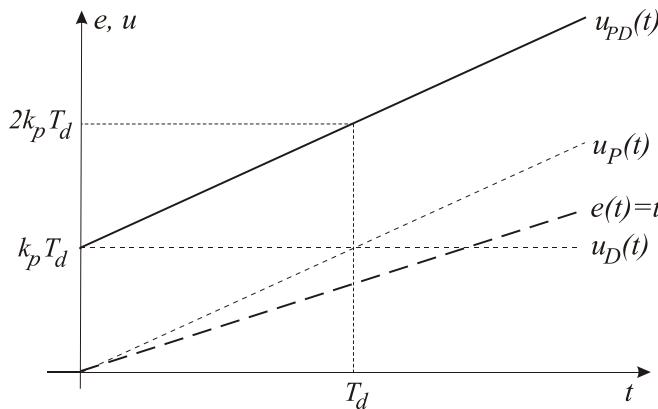
Działanie całkujące regulatora PID najlepiej wyjaśnia jego wersja PI (tzn. z wyłączeniem różniczkowania, $T_d = 0$), na Rys. 8.2 przedstawiono odpowiedź regulatora PI na jednostkowy skok uchybu $e(t)$ na jego wejściu. Odpowiedź ta składa się z dwóch składowych: odpowiedzi proporcjonalnej $u_P(t)$ będącej też sygnałem skokowym, jedynie o amplitudzie k_p , i odpowiedzi członu całkującego $u_I(t)$ w postaci sygnału liniowo narastającego w czasie. Rysunek ten wyjaśnia też tradycyjną nazwę *czas zdwojenia* używaną dla określenia stałej czasowej całkowania T_i ; po czasie T_i wartość sygnału na wyjściu członu całkującego osiąga wartość wyjścia członu proporcjonalnego (która jest stała), tzn. wartość sygnału wyjściowego regulatora PI (czy też PID) osiąga podwojoną wartość wzmocnionej (przez k_p) amplitudy sygnału skokowego na jego wejściu.



Rys. 8.2. Odpowiedź skokowa regulatora PI.

Z kolei działanie różniczkujące regulatora PID najlepiej wyjaśnić dla jego wersji PD (tzn. z wyłączeniem całkowania, $T_i = \infty$), dla wymuszenia sygnałem uchybu liniowo narastającym $t \cdot 1(t)$. Sytuacja taka przedstawiona jest na Rys. 8.3 – odpowiedź regulatora składa się z

dwóch składowych: odpowiedzi proporcjonalnej $u_P(t)$ będącej tez sygnałem liniowo narastającym (tylko o innym nachyleniu, zależnym od wartości k_p) oraz odpowiedzi członu różniczkującego $u_D(t)$ będącej sygnałem skokowym o amplitudzie skoku równej $k_p T_d$. Rysunek ten wyjaśnia też tradycyjną nazwę *czas wyprzedzenia* używaną dla określenia stałej czasowej różniczkowania T_d : po czasie T_d wartość składowej proporcjonalnej sygnału wyjściowego regulatora PD (czy też PID) osiąga wartość amplitudy sygnału odpowiedzi części różniczkującej (sygnał stały), tzn. ten ostatni sygnał wyprzedza o czas T_d reakcję członu proporcjonalnego.



Rys. 286H8.3. Odpowiedź prędkościowa (na sygnał o stałej pochodnej) regulatora PD.

Regulator PID opisany transmitancją (8.2) nazywany jest regulatorem *idealnym*, gdyż realizuje idealne działanie różniczkujące opisane członem $k_p T_d s$, oczywiście nierealizowalnym fizycznie. Struktura równoległa rzeczywista (*realizowalna, praktyczna*) regulatora PID może być zrealizowana następująco:

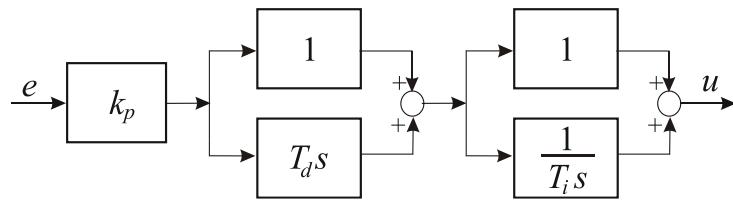
$$R_{PID}(s) = k_p \left(1 + \frac{1}{T_i s} + \frac{T_d s}{\alpha} \right) = k_p + \frac{k_p}{T_i} s + \frac{k_p T_d s}{\alpha} \quad (8.3)$$

gdzie zamiast idealnego członu różniczkującego występuje tzw. *rzeczywisty człon różniczkujący* powstały z tego poprzedniego poprzez dodanie działania inercyjnego ze stałą czasową inercji T_d/α . Wartość współczynnika α (zwanego niekiedy wzmocnieniem dynamicznym regulatora PID) wybierana jest z reguły z zakresu 5 do 20, typowa wartość to $\alpha = 10$.

Oprócz omówionej powyżej równoległej realizacji algorytmu regulatora PID stosowana bywa też w praktyce *szeregową strukturą algorytmu PID*, o standardowej transmitancji postaci (z idealnym członem różniczkującym)

$$R_{PID}^{sz}(s) = k_p \left(1 + \frac{1}{T_i s} \right) (1 + T_d s) \quad (8.4)$$

Zwrócmy uwagę, że (8.4) jest szeregowym połączeniem regulatorów równoległych PD i PI, jak to wyraźnie widać z samej postaci transmitancji, jak i na ilustrującym ją Rys. 8.4.



Rys. 8.4. Standardowa szeregowa struktura algorytmu PID (idealnego).

Postać *szeregowa realizowalna*, z operacją różniczkowania realizowaną uzyskaną w identyczny sposób jak powyżej dla wersji równoległej, jest następującą:

$$R_{PID}^{sz}(s) = k_p \left(1 + \frac{1}{T_i s} \right) \left(1 + \frac{T_d s}{\frac{T_d}{\alpha} s + 1} \right) \quad (8.5)$$

Można łatwo wyprowadzić zależności na wzajemne relacje między równoległą (8.2) i szeregową (8.4) postaciami transmitancji regulatora PID. Mamy mianowicie

$$\begin{aligned} R_{PID}^{sz}(s) &= k_p \left(1 + \frac{1}{T_i s} \right) (1 + T_d s) = k_p \left(1 + T_d s + \frac{1}{T_i s} + \frac{T_d s}{T_i s} \right) = \\ &= k_p \left(\frac{T_i s + T_d s}{T_i s} + T_d s + \frac{1}{T_i s} \right) = \\ &= k_p \frac{(T_i + T_d)}{T_i} \left(1 + \frac{1}{(T_i + T_d)s} + \frac{T_i T_d}{T_i + T_d} s \right) \end{aligned} \quad (8.6)$$

Z zależności powyższej wynika, że struktura szeregowa regulatora PID jest równoważna strukturze równoległej, a nastawy równoważnej postaci równoległej są kombinacjami nastaw struktury szeregowej. Natomiast, przekształcając zależność (8.6) do postaci

$$R_{PID}^{sz}(s) = k_p \left(1 + \frac{T_d}{T_i} \right) \left(1 + \frac{\frac{T_d}{1 + \frac{T_d}{T_i}} s + \frac{1}{(1 + \frac{T_d}{T_i}) T_i s}}{1 + \frac{T_d}{T_i}} \right) \quad (8.7)$$

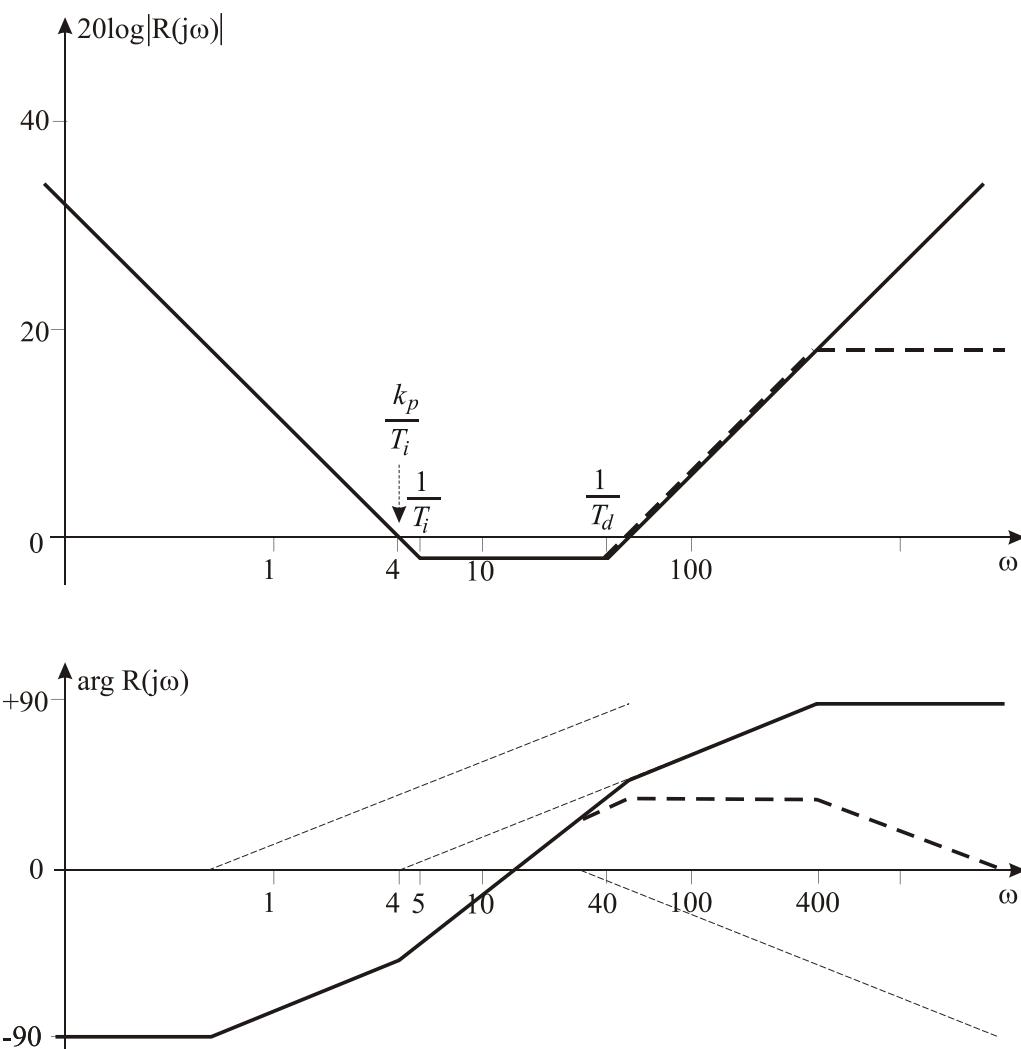
można łatwo zaobserwować, że im mniejsza jest wartość stałej czasowej różniczkowania T_d w stosunku do stałej czasowej całkowania T_i , tym bliższa jest struktura szeregowa strukturze równoległej z tymi samymi nastawami. Przypomnijmy, że zawsze $T_d < T_i$, a w istocie w praktyce stosuje się $T_d \ll T_i$. Zauważmy ponadto, że przy wyłączonym różniczkowaniu ($T_d = 0$), tzn. dla regulatora PI, standardowe struktury równoległa i szeregowa są identyczne.

Charakterystyki częstotliwościowe regulatora PID najwygodniej jest wyznaczyć dla struktury szeregowej (8.4) czy (8.5) przedstawionych, odpowiednio, w postaciach

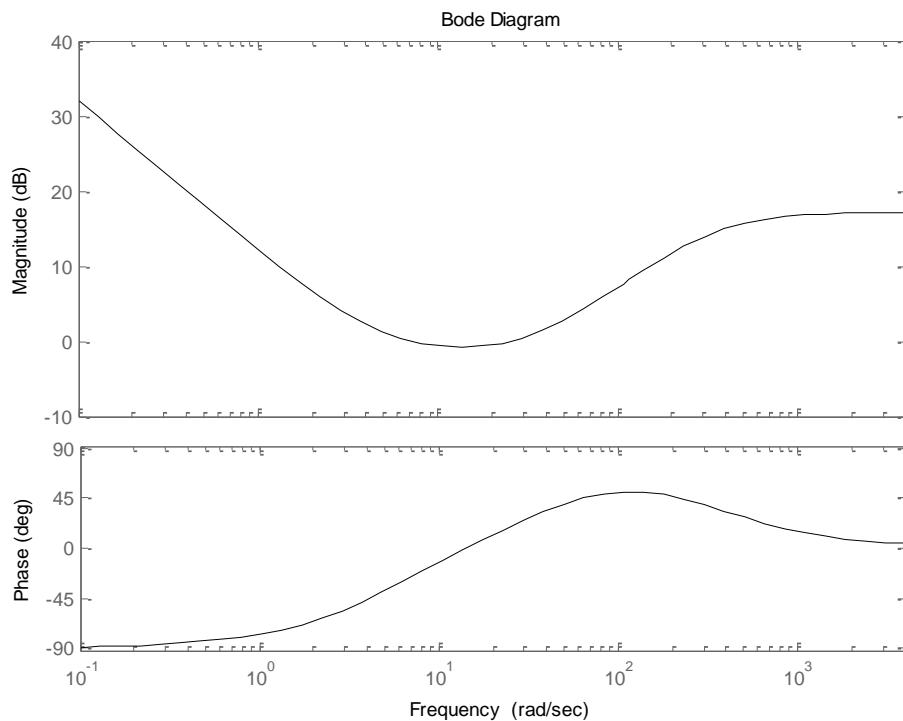
$$R_{PID}^{sz}(s) = \frac{k_p}{T_i s} (1 + T_i s)(1 + T_d s) \quad (8.8)$$

$$R_{PID}^{sz}(s) = \frac{k_p}{T_i s} (1 + T_i s) \frac{1 + T_d (1 + \frac{1}{\alpha}) s}{1 + \frac{T_d}{\alpha} s} \quad (8.9)$$

Charakterystyki częstotliwościowe Bodego asymptotyczne, amplitudową i fazową, dla obu tych przypadków przedstawiono na Rys. 8.5, dla regulatora PID (8.4) o nastawach $k_p=0.8$, $T_i=0.2$, $T_d=0.025$ (linia ciągła) i (8.5) z $\alpha = 8$ (linia przerywana w zakresie większych pulsacji). Dla ułatwienia rozumienia charakterystyki fazowej naszkicowano też (liniami cienkimi przerywanymi) zmiany fazę czynników składowych transmitancji. Natomiast na Rys. 8.6 przedstawiono, dla porównania, charakterystyki Bodego dokładne regulatora rzeczywistego (8.5).



Rys. 8.5. Charakterystyki częstotliwościowe logarytmiczne asymptotyczne przykładowego regulatora PID.



Rys. 8.6. Charakterystyki częstotliwościowe regulatora PID (rzeczywistego).

Z charakterystyk częstotliwościowych można bezpośrednio wywnioskować, dlaczego regulator PID stał się tak uniwersalnym w zadaniach regulacji przemysłowej. Ma on prostą strukturę, zapewniając przy tym realizację podstawowych celów: tłumienia zakłóceń przy poszerzeniu pasma przenoszenia skutującym dobrą szybkością regulacji również w zadaniu nadążania. Dzięki działaniu całkującemu uzyskujemy zerowe uchyby regulacji w stanach ustalonych (w zadaniach stabilizacji), natomiast zastosowanie działania różniczkującego wprowadza dodatkowe dodatnie przesunięcie fazowe w zakresie średnich częstotliwości skutkujące poszerzeniem pasma przenoszenia, a więc możliwością uzyskania większej szybkości działania układu regulacji.

Ze względu na równoważność struktur równoległej i szeregowej, zob. wzór (8.6) czy (8.7), charakterystyki częstotliwościowe regulatora PID o strukturze równoległej mają taki sam kształt, jedynie nieco inne są wartości pulsacji odpowiadających ich punktom charakterystycznym. Przy tym, różnice te są tym mniejsze im mniejsza jest wartość T_d w stosunku do wartości T_i , patrz zal. (8.7).

Uwagi dodatkowe*.

Niech k_p , T_i i T_d oznaczają nastawy regulatora PID o strukturze równoległej, oznaczmy natomiast nastawy równoważnego regulatora o strukturze szeregowej przez k_{pS} , T_{iS} i T_{dS}

- a) *Przejście od struktury szeregowej do równoległej.* Z wzoru (8.6) mamy bezpośrednio

* Materiał uzupełniający.

$$\begin{aligned}
 k_p &= k_{ps} \frac{(T_{is} + T_{ds})}{T_{is}} = k_{ps} \left(1 + \frac{T_{ds}}{T_{is}}\right) \\
 T_i &= T_{is} + T_{ds} = T_{is} \left(1 + \frac{T_{ds}}{T_{is}}\right) \\
 T_d &= \frac{T_{is}}{T_{is} + T_{ds}} T_{ds} = T_{ds} \frac{1}{\left(1 + \frac{T_{ds}}{T_{is}}\right)}
 \end{aligned} \tag{8.10}$$

Rozważmy stosunek stałych czasowych T_d do T_i . Z wzorów powyżej bezpośrednio wynika

$$\frac{T_d}{T_i} = \frac{T_{ds}}{T_{is}} \frac{1}{\left(1 + \frac{T_{ds}}{T_{is}}\right)^2} \tag{8.11}$$

Analizując przebieg tej funkcji względem argumentu T_{ds}/T_{is} łatwo sprawdzić, że przyjmuje ona wartość maksymalną dla $T_{ds} = T_{is}$, i że wartość ta równa się 0.25. Stąd wniosek, że niezależnie od wartości stałych czasowych regulatora PID o strukturze szeregowej, stałe czasowe równoważnej postaci równoległej spełniają warunek $T_d \leq 0.25T_i$.

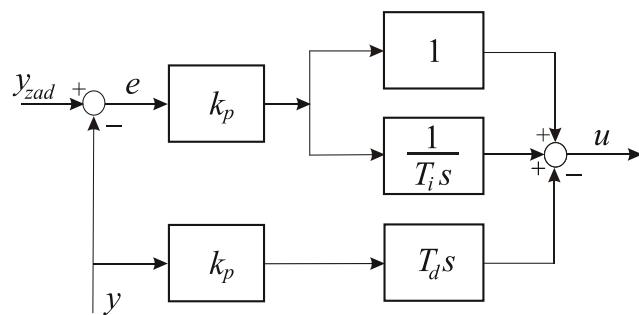
- b) *Przejście od struktury równoległej do szeregowej.* Należy wyznaczyć nastawy k_{ps} , T_{is} i T_{ds} równoważnej struktury szeregowej z wzorów (8.10), w funkcji nastaw k_p , T_i i T_d . Łatwo to wykonać zauważając, że z wzorów tych bezpośrednio wynika $T_{is}T_{ds} = T_iT_d$; w rezultacie otrzymujemy:

$$\begin{aligned}
 k_{ps} &= k_p \frac{T_{is}}{T_i} = \frac{1}{2} k_p \left(1 + \sqrt{1 - \frac{4T_d}{T_i}}\right) \\
 T_{is} &= \frac{1}{2} T_i \left(1 + \sqrt{1 - \frac{4T_d}{T_i}}\right) \\
 T_{ds} &= T_d \frac{2}{1 + \sqrt{1 - \frac{4T_d}{T_i}}} = \frac{1}{2} T_i \left(1 - \sqrt{1 - \frac{4T_d}{T_i}}\right)
 \end{aligned} \tag{8.12}$$

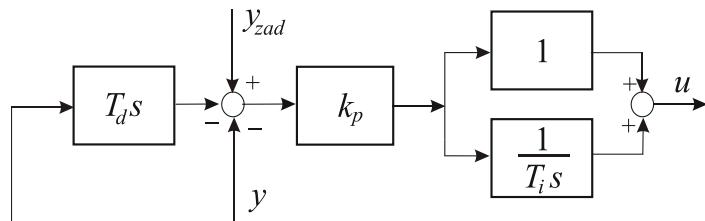
Z rozważań powyższych wynika, że chcąc mieć regulator równoległy PID z dwoma biegunami rzeczywistymi o przejrzystej strukturze takiej, jak regulator szeregowy, powinno się przyjmować $T_d \leq 0.25T_i$, gdyż $T_d = 0.25T_i$ odpowiada w równoważnej strukturze szeregowej przypadkowi granicznemu $T_{ds} = T_{is}$ (gdy $T_d < 0.25T_i$ to $T_{ds} < T_{is}$). W praktyce przyjmuje się $T_d = (0.25 \div 0.1)T_i$. Zauważmy, że gdy $T_d \ll T_i$, np. $T_d = 0.1T_i$, to jak wynika z wzorów (8.10) czy (8.12), wartości nastaw w strukturach równoważnych szeregowej i równoległej są bardzo zbliżone do siebie. Nadmienmy ponadto, że dla regulacji obiektów oscylacyjnych (o parze biegunów zespolonych) korzystne może być wykorzystanie regulatora PID o biegunach zespolonych – zgodnie z powyższymi uwagami może to być tylko regulator o strukturze równoległej.

8.2.2. Zmodyfikowane struktury regulatora PID

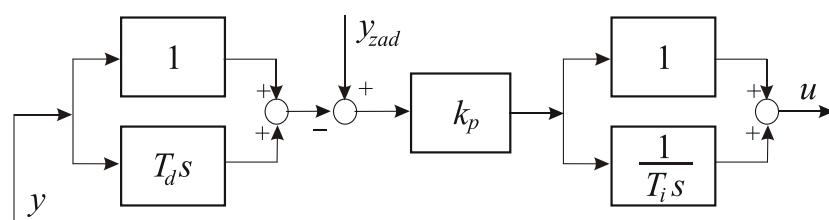
Przedstawione powyżej struktury równoległa i szeregowa regulatora PID to tzw. *struktury standardowe*, w których wejściem regulatora jest jedynie sygnał uchybu regulacji $e(t)$. Struktury te nie są polecane w sytuacjach, kiedy mogą następować nagłe, w szczególności skokowe zmiany sygnału uchybu regulacji powodowane nagłymi zmianami sygnału wartości zadanej, a w regulatorze występuje człon różniczkujący (regulator PID lub PD). Różniczkowanie sygnału uchybu powoduje wówczas występowanie w sygnale wyjściowym regulatora dużych, skokowo zmiennych wartości, co jest niekorzystne. W sytuacji takiej zalecane są tzw. *zmodyfikowane struktury regulatora PID*, w których sygnał uchybu regulacji poddawany jest działaniu proporcjonalnemu i całkującemu, a działaniu różniczkującemu podlega jedynie sygnał wyjściowy obiektu regulacji $y(t)$. Taka zmodyfikowana struktura równoległa regulatora PID przedstawiona jest na Rys. 8.7, natomiast odpowiednia zmodyfikowana struktura szeregowa przedstawiona jest na Rys. 8.8. Ta ostatnia może być przedstawiona w równoważnej postaci podanej na Rys. 8.9, wyraźniej pokazującej związek ze standardową strukturą szeregową przedstawioną uprzednio na Rys. 8.4.



Rys. 8.7. Zmodyfikowana równoległa struktura algorytmu PID (idealna).



Rys. 8.8. Zmodyfikowana szeregowa struktura algorytmu PID (idealna).



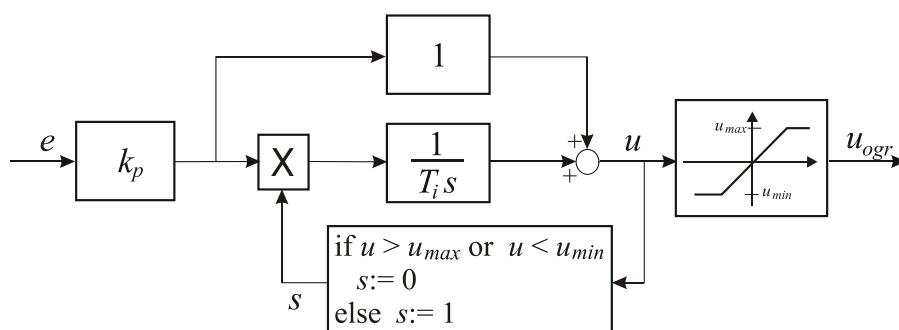
Rys. 8.9. Równoważne przedstawienie struktury z rysunku poprzedniego.

Zwróćmy uwagę, że w zadaniach stabilizacji (przy ustalonej wartości zadanej y_{zad}) struktury standardowe i zmodyfikowane działają identycznie, tzn. działają one identycznie przy tłumieniu zakłóceń (niezależnie od miejsca ich oddziaływanego na obiekt regulacji).

8.2.3. Ograniczenie całkowania

W sytuacji, kiedy podczas pracy regulatora z całkowaniem uchybu (PI lub PID) występują dłuższe okresy czasu z uchybem stałym na wartości maksymalnej (minimalnej lub maksymalnej) konieczna jest implementacja regulatora z tzw. ograniczeniem całkowania. W sytuacjach takich następuje bowiem silny wzrost amplitudy sygnału na wyjściu elementu całkującego (integratora) regulatora, stąd sygnał wyjściowy regulatora nierzadko wielokrotnie przekracza możliwości jego realizacji, tzn. wielokrotnie przekracza ograniczenia amplitudy sygnału sterującego obiektem wynikające z ograniczonej mocy elementu wykonawczego. Nie byłoby w tym nic złego, gdyby sygnał podawany na obiekt był ograniczany przez możliwości elementu wykonawczego, gdyby nie fakt długiego czasu potrzebnego na zredukowanie takiego „nacalkowanego” stanu integratora sygnałem uchybu regulacji o przeciwnym znaku, już po zmianie znaku tego uchybu. Powoduje to znaczne pogorszenie jakości regulacji, przejawiające się przede wszystkim znacznie dłuższym czasem regulacji i dłuższymi przeregulowaniami. Stąd standardem w przemysłowych regulatorach PID jest występowanie mechanizmów ograniczenia całkowania, różnego rodzaju.

Najprostszy mechanizm ograniczenia całkowania to zatrzymanie akcji całkowania w sytuacji, kiedy sygnał wyjściowy regulatora przekracza swoje dopuszczalne wartości maksymalne (wartości ograniczenia). Sposób ten zilustrowano na Rys. 8.10, gdzie element mnożący zatrzymuje działanie integratora regulatora PI (podając zero na jego wejście) w sytuacji przekroczenia ograniczeń. Układy ograniczenia całkowania regulatora PI czy PID nazywane są układami *anti-windup*.



Rys. 8.10. Przykładowa struktura regulatora PI z ograniczeniem całkowania.

Przykład 8.1.

Porównamy przebiegi w układach regulacji z regulatorem PI implementowanym

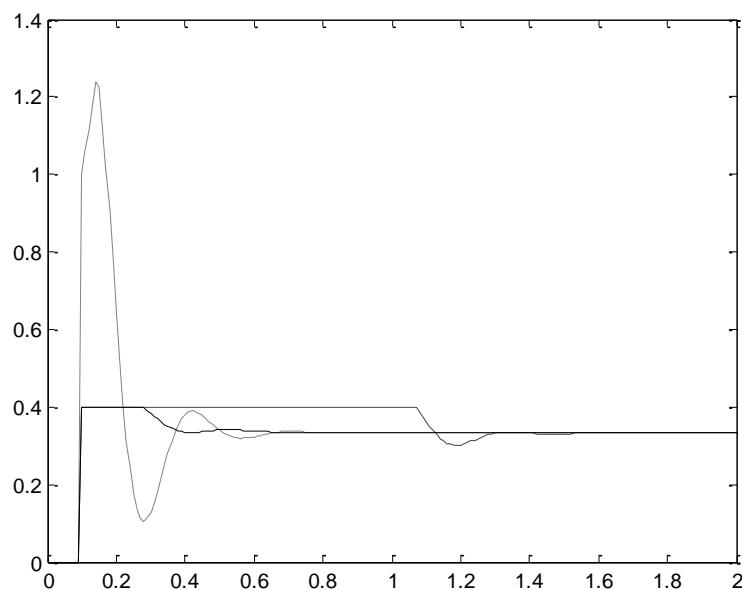
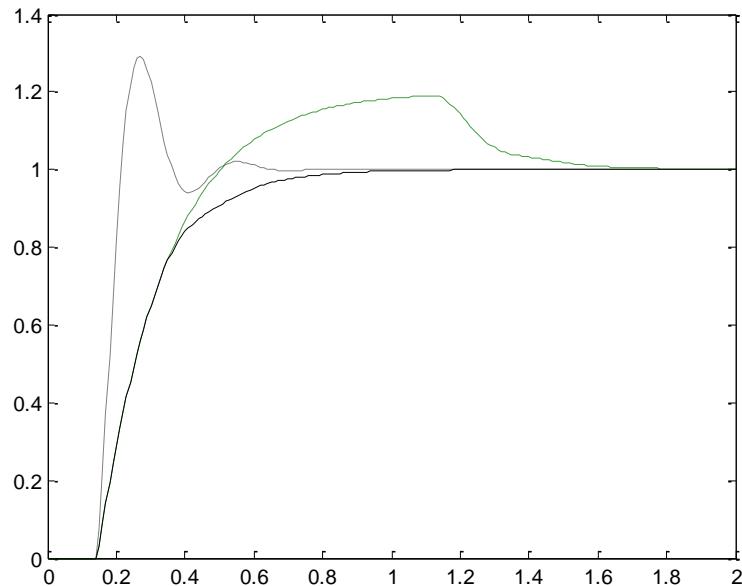
- bez ograniczenia sterowania,
- z wartością sygnału sterującego ograniczoną do zakresu $u_{min} = -0.4$, $u_{max} = +0.4$ i bez układu ograniczania całkowania,
- z układem ograniczania całkowania zrealizowanym jak na Rys. 8.10

dla obiektu o transmitancji

$$G(s) = \frac{3e^{-0.05s}}{1 + 0.2s}$$

i z regulatorem PI o nastawach $k_p=1$, $T_i=1/6$.

Na Rys. 8.11 przedstawiono przebiegi wyjścia (wielkości regulowanych) i wejścia (sterowań) obiektu, linia kropkowaną dla przypadku a), linią przerywaną dla przypadku b) i linią ciągłą dla przypadku c).



Rys. 8.11. Przebiegi wielkości regulowanej (rys. górny) i sterowań (rys. dolny) w układzie regulacji PI po skoku wartości zadanej w chwili $t = 0.1$: bez ograniczenia amplitudy sterowania (linia kropkowana) i z tym ograniczeniem (linie przerywana – bez układu *anti-windup*, ciągła – z układem *anti-windup*).

8.3. Modele obiektów dla doboru nastaw regulatorów PID

Typowym sposobem postępowania przed doborem nastaw regulatora PID jest aproksymacja dynamiki obiektu (w danym punkcie pracy) przy pomocy jednego ze standardowych prostych modeli dynamiki, dla których opracowano wzory lub tabele doboru nastaw.

Najprostsze i jednocześnie najważniejsze *proste modele obiektów typu przemysłowego* to:

1. Jednoinercyjny z opóźnieniem, tzw. FOPD (*first-order plus delay*), dla obiektów statycznych (inercyjnych):

$$G(s) = \frac{ke^{-sT_0}}{1+sT} \quad (8.13)$$

2. Całkujący z opóźnieniem, tzw. IPD (*integrating plus delay*), dla obiektów astatycznych, przyjmowany w jednej z postaci podanych poniżej,

$$G(s) = \frac{ke^{-sT_0}}{s} = \frac{e^{-sT_0}}{Ts} \quad (8.14)$$

3. Drugiego rzędu z opóźnieniem, tzw. SOPD (*second-order plus delay*):

- Dwuinercyjny z opóźnieniem, dla obiektów nieoscylacyjnych:

$$G(s) = \frac{ke^{-sT_0}}{(1+sT_1)(1+sT_2)} \quad (8.15)$$

- Oscylacyjny z opóźnieniem, dla obiektów o oscylacyjnej odpowiedzi skokowej:

$$G(s) = \frac{ke^{-sT_0}}{T^2 s^2 + 2T\zeta s + 1} = \frac{ke^{-sT_0} \omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}, \quad 0 < \zeta < 1 \quad (8.16)$$

Istnieją dwa podstawowe podejścia do uzyskiwania wymienionych powyżej prostych modeli dynamiki obiektów:

- *Metody bezpośrednie*: na podstawie cech przebiegu odpowiedzi skokowej (pomierzonej z obiektu lub zasymulowanej na modelu dokładniejszym).
- *Metoda optymalizacji parametrycznej*: najczęściej metoda minimalizacji sumy kwadratów (tzw. metoda optymalizacji średniokwadratowej) różnic między wartościami wyjść y^R z przebiegu referencyjnego (uzyskanego z pomiarów rzeczywistych lub z symulacji modelu dokładnego) i wyjść modelu prostszego y^G , dla tych samych wartości sterowań, np. odpowiadających skokowi sygnału wejściowego obiektu (odpowiedzi skokowe):

$$\min_P \sum_{k=1}^N (y^R(k) - y^G(k, P))^2 \quad (8.17)$$

gdzie P oznacza wektor parametrów optymalizowanych modelu uproszczonego; przy czym przy analizie odpowiedzi skokowej wzmacnienie k i opóźnienie T_0 dobieramy z reguły bezpośrednio na wstępnie, a optymalizujemy jedynie wartości stałych czasowych (ew. dostrajamy też opóźnienie).

Korzystając z odpowiedzi skokowej, wzmacnienie określamy z wzoru

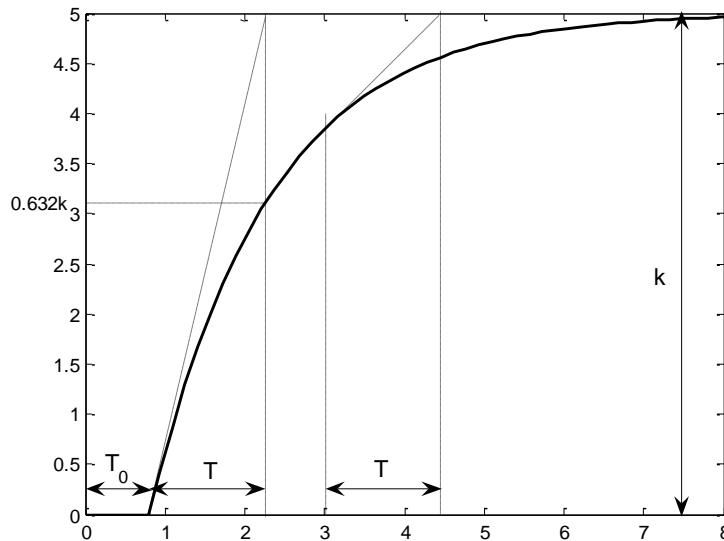
$$k = \frac{y^\infty - y_0}{\Delta u} = \frac{\Delta y^\infty}{\Delta u} \quad (8.18)$$

gdzie y_0 to początkowa ustalona wartość wyjścia obiektu (przed podaniem skoku sterowania), a y^∞ to wartość ustalona po podaniu skoku. Dla właściwego doboru nastaw wzmacnienie k

należy wyznaczyć jako bezwymiarowe, tzn. jako iloraz przyrostów unormowanych do zakresów zmienności wielkości wejściowej (sterującej) i wyjściowej (regulowanej) obiektu.

Identyfikacja modelu FOPD na podstawie odpowiedzi skokowej

Na Rys. 8.12 pokazano przebieg odpowiedzi skokowej obiektu o transmitancji FOPD, na wymuszenie skokiem wejścia w chwili $t = 0$, od wartości 0 do $u_0 = 1$.



Rys. 8.12. Wyznaczanie parametrów modelu FOPD z odpowiedzi skokowej.

Na rysunku zilustrowano dwie metody wyznaczania parametrów modelu:

- *metoda stycznej*, poprowadzonej z dowolnego punktu krzywej odpowiedzi, danej wzorem

$$\frac{dy}{dt}(t) = \frac{k}{T} e^{-\frac{t-T_0}{T}} \quad (8.19)$$

przy czym na rysunku przedstawiono dwie styczne, w punktach dla $t=T_0$ i $t=3$.

- *metoda przyrostu wartości wyjścia w czasie T* (w stosunku do całego przyrostu), gdyż

$$y(T_0 + T) = k(1 - e^{-1}) \cong k \cdot 0.632 \quad (8.20)$$

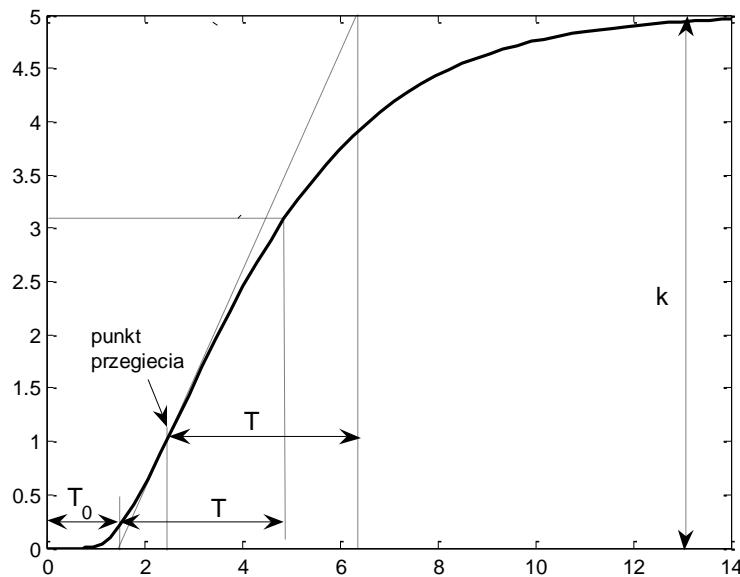
Ten ostatni sposób uważany jest za pewniejszy. Natomiast wzmacnienie k wyznaczamy jako iloraz przyrostów, jak to wyjaśniono powyżej.

W praktyce na ogół przybliżamy modelem FOPD transmitancje wyższych rzędów, najczęściej spotykany sposób postępowania pokazany na Rys. 8.13, gdzie modelem FOPD przybliżana jest odpowiedź skokowa obiektu trójinercynego o transmitancji

$$G(s) = \frac{5}{(1+2s)(1+1.8s)(1+0.1s)} \quad (8.21)$$

Odpowiedzi wieloinercyjne mają charakterystyczny kształt z punktem przegięcia. W punkcie tym poprowadzono na Rys. 8.13 styczną do wykresu – na osi rzędnych odcina ona zastępcze opóźnienie T_0 , natomiast jej przecięcie z wartością docelową wyjścia szacuje zastępczą stałą czasową T metodą stycznej. Sposób ten prowadzi jednak często do mniej dokładnej aproksymacji, lepsze wyniki uzyskujemy na ogół metodą przyrostu, tzn. stałą czasową T

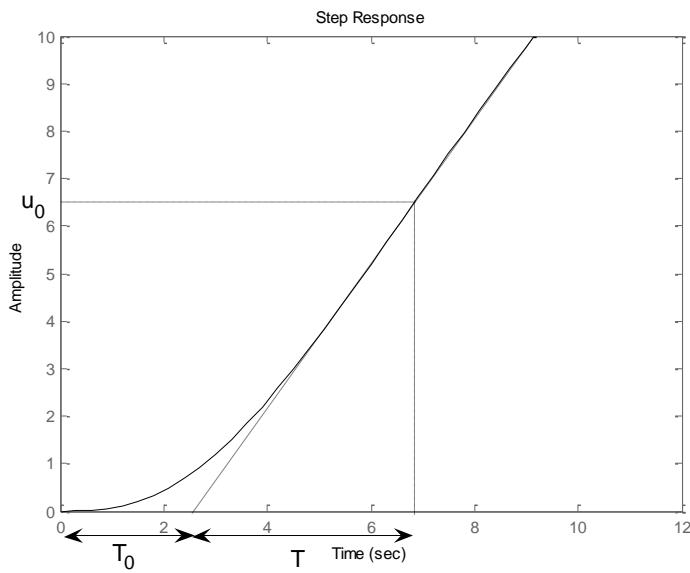
wyznaczamy jako czas, w którym skok wyjścia osiąga wartość $0.632k$, gdzie k jest amplitudą całkowitego przyrostu wartości wyjścia po skoku sterowania (też pokazane na Rys. 8.13).



Rys. 8.13. Wyznaczanie parametrów modelu FOPD z odpowiedzi skokowej obiektu wieloinercycznego metodą stycznej w punkcie przejęcia.

Identyfikacja modelu IPD na podstawie odpowiedzi skokowej

Metoda wyznaczania parametrów modelu IPD przedstawiona jest na Rys. 8.14.



Rys. 8.14. Wyznaczanie zastępczego modelu IPD z odpowiedzi skokowej obiektu regulacji.

Odwrotność T wzmacnienia prędkościowego k wyznaczamy jako czas, po którym wyjście obiektu osiąga wartość amplitudy u_0 skoku na jego wejściu (w jednostkach porównywalnych, bezwymiarowych). Sposób ten wynika bezpośrednio z postaci dynamiki obiektu danego

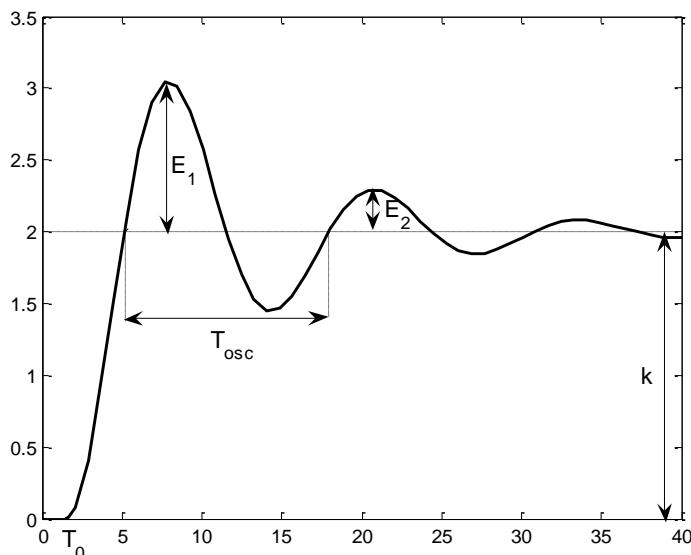
transmitancją (8.14), albowiem dla sygnału wejściowego skokowego o amplitudzie u_0 mamy (skok chwili $t=0$):

$$y(t) = \frac{1}{T} \cdot u_0 \cdot (t - T_0) \Rightarrow (y(t_1) = u_0) \Rightarrow (t_1 = T_0 + T) \quad (8.22)$$

Identyfikacja modelu SOPD oscylacyjnego na podstawie odpowiedzi skokowej

Istnieją odpowiednie zależności na wyznaczenie stałych czasowych modelu dwuinercyjnego z opóźnieniem bezpośrednio na podstawie odpowiedzi skokowej, ale są one bardziej złożone i obarczone niepewnością, dlatego pominiemy ich przedstawienie (optymalizacja parametryczna jest tu bardziej polecana), tym bardziej, że w tym przypadku możemy spróbować zastosować również model FOPD. Natomiast pokażemy, jak wyznaczać parametry modelu oscylacyjnego z opóźnieniem – sposób jest dość prosty i stosunkowo dokładny, a model ten jest ważny jako jedyny model uproszczony dla obiektów o oscylacyjnej odpowiedzi skokowej.

Odpowiedź skokowa obiektu o transmitancji SOPD oscylacyjnej z $k=2$, $T=4$, $\zeta=0.2$ i $T_0=1.5$ przedstawiona jest na Rys. 8.15.



Rys. 8.15. Przebieg odpowiedzi skokowej modelu SOPD oscylacyjnego.

Opisana jest ona zależnością (dla $T_0=0$):

$$y(t) = k \left[1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\omega_0 t} \sin(\omega_n \sqrt{1-\zeta^2} \cdot t + \varphi) \right] \cdot \mathbf{1}(t), \quad \varphi = \arctg \frac{\sqrt{1-\zeta^2}}{\zeta} \quad (8.23)$$

gdzie $\omega_n=1/T$ to pulsacja drgań własnych (swobodnych), ζ to współczynnik tłumienia względnego, a k współczynnik wzmacniania. Z zależności powyższej i parametrów zaznaczonych na rysunku wynikają następujące proste wzory:

$$\omega_n = \sqrt{\omega^2 + \alpha^2}, \quad \zeta = \frac{\alpha}{\omega_n}, \quad \text{gdzie} \quad \omega = \frac{2\pi}{T_{osc}}, \quad \alpha = \frac{1}{T_{osc}} \ln \frac{E_1}{E_2} \quad (8.24)$$

Istnieje jeszcze szereg innych modeli uproszczonych bardziej złożonych niż przedstawione powyżej, jak np. model wieloinercyjny z opóźnieniem o identycznych stałych czasowych, model FOPD z dodatkowym całkowaniem, model podwójnie całkujący z opóźnieniem, czy modele o bardziej rozbudowanym liczniku dla obiektów nieminimalnofazowych (z zerami transmitancji w prawej półpłaszczyźnie). Są to modele dla trudniejszych, mniej typowych obiektów. Zainteresowanego czytelnika odsyłamy do literatury, np. [3].

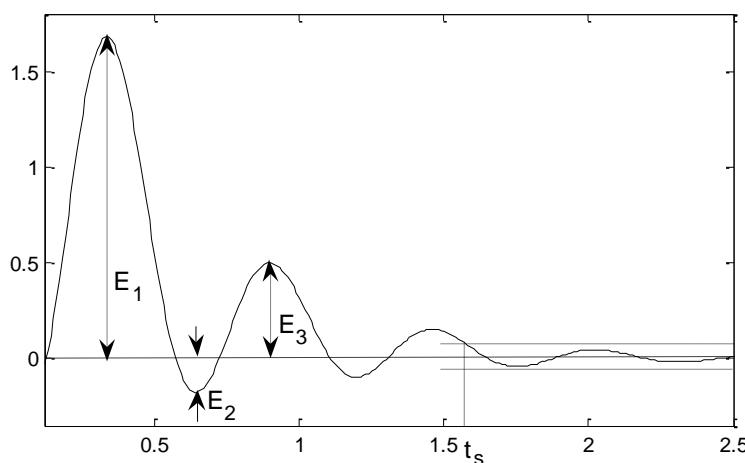
8.4. Dobór nastaw regulatorów PID

8.4.1. Kryteria doboru nastaw

Właściwości zamkniętego układu regulacji zależą od cech obiektu oraz od wartości nastaw regulatora – te ostatnie są w ręku projektanta układu. Nastawy regulatorów w układach regulacji przemysłowej dobierane są pod kątem uzyskania odpowiednich cech przebiegów przejściowych i stanów ustalonych. Jeśli dopuszczalne są niezerowe uchyby ustalone, to możliwe jest zastosowanie regulatora P, jak to ma często miejsce w układach regulacji poziomu cieczy w zbiornikach, kiedy istotne jest utrzymanie poziomu w założonych granicach, a nie jego dokładna stabilizacja na określonej wartości – przypomnijmy, że regulator P ma lepsze właściwości dynamiczne, stabilnościowe niż regulator PI, gdyż całkowanie ma działanie destabilizujące, wprowadzając ujemne przesunięcie fazowe w całym zakresie częstotliwości.

Istnieje szereg kryteriów doboru nastaw regulatorów P, PI czy PID, wynikających z różnych wymagań stawianych przed układami regulacji. Wymagania formułowane są najczęściej w języku pewnych wzorcowych przebiegów przejściowych w układzie regulacji (lub funkcji tych przebiegów), wynikających z pożądanej jakości regulacji, rozumianej jako odpowiednia jakość nadążania za wartościami zadanymi, tłumienia zakłóceń – przy zapewnieniu określonych cech niezawodnościowych (odpornej stabilności). Najczęściej stosowane przebiegi wzorcowe to:

- Przebieg oscylacyjny z czterokrotnym tłumieniem (klasyczne *kryterium QDR - quarter decay ratio*) – zilustrowany na Rys. 8.16. *Decay ratio* definiowany jest jako stosunek E_3 do E_1 , kryterium QDR zakłada, że stosunek ten powinien być 0.25 (*a quarter*).
- *Przebieg oscylacyjny z przeregulowaniem 20%* – przeregulowanie definiowane jest jako stosunek E_2 do E_1 , (patrz Rys. 8.16), i wyrażane jest zwykle w procentach.



Rys. 8.16. Przebieg uchybu regulacji (po skoku zakłócenia na wejściu obiektu).

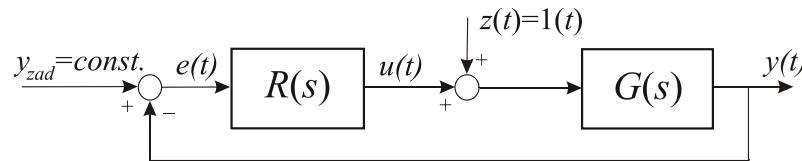
- Przebieg aperiodyczny krytyczny (bez przeregulowania na jego granicy), tzn. bez przeregulowania z minimalnym czasem regulacji, gdzie czas regulacji (czas ustalania się – *settling time* t_s , patrz Rys. 8.16) jest to czas, po którym uchyb regulacji stłumiony jest w założonej proporcji, np. nie przekracza $\pm 0.05 E_{ust}$, gdzie E_{ust} to uchyb ustalony jaki powstałby na wyjściu obiektu w układzie bez regulatora.
- Przebieg minimalizujący kryterium sformułowane jako całka z funkcji uchybu:
 - całka z modułu uchybu (*kryterium IAE – integral absolute error*):

$$J_{IAE} = \int_0^{\infty} |e(t)| dt \quad (8.25)$$

- całka całkę z kwadratu uchybu (*kryterium ISE – integral squared error*):

$$J_{ISE} = \int_0^{\infty} e(t)^2 dt \quad (8.26)$$

Zwróćmy uwagę, że całki J_{IAE} i J_{ISE} istnieją tylko dla przebiegów z zerowym uchybem ustalonym. Opracowanych jest szereg metod doboru nastaw regulatorów P, PI czy PID, realizujących założoną postać przebiegu przejściowego lub minimalizujących kryterium jakości regulacji, dla różnych struktur tych regulatorów. Uchyby regulacji mogą być powodowane zarówno zmianami wartości zadanej, jak i wpływem zakłóceń. Ponieważ w układach regulacji przemysłowej z reguły istotniejsze jest tłumienie zakłóceń, to najczęściej podawane są reguły dla przebiegu uchybu regulacji po skokowej zmianie zakłócenia oddziaływanego na wejście obiektu regulacji, jak w strukturze na rysunku poniżej. Reguły podane w następnych rozdziałach wyznaczane były na ogólnie w takiej strukturze.



Rys. 8.17. Struktura układu dla wyznaczania przebiegu uchybu regulacji dla doboru nastaw w regulacji przemysłowej (dla tłumienia zakłóceń).

8.4.2. Tabelaryczne reguły doboru nastaw

Reguły Zieglera-Nicholsa wg cech przebiegu krytycznego

Pierwsze i uważane obecnie za klasyczne, opracowane przez Zieglera i Nicholsa w 1942 r., są reguły doboru nastaw wg cech tzw. *przebiegu krytycznego*, tzn. w układzie regulacji na granicy jego stabilności [8]. Aby uzyskać parametry przebiegu krytycznego Ziegler i Nichols zaproponowali następujący eksperyment na obiekcie (rzeczywistym, lub symulacyjnie):

1. Ustawiamy regulator na działanie P, tzn. wyłączaemy całkowanie (lub ustawiamy T_i na maksymalnej wartości, idealnie $T_i = \infty$) i wyłączaemy różniczkowanie (ustawiając $T_d = 0$),
2. Zwiększamy stopniowo wzmacnienie regulatora k_p aż do momentu, w którym pojawią się oscylacje niegasnące (granica stabilności). Wzmocnienie odpowiadające uzyskanej granicy stabilności określone jest jako *wzmocnienie krytyczne* k_{kr} , a odpowiadający mu okres oscylacji niegaszących to *okres krytyczny* T_{kr} .

Dla obiektu scharakteryzowanego przez k_{kr} i T_{kr} Ziegler i Nichols wyznaczyli nastawy regulatorów P, PI i PID (według kryterium QDR), podane w Tabeli 8.1.

Tabela 8.1. Reguły Zieglera-Nicholsa (Z-N) wg cech przebiegu krytycznego (1942 r.).

Regulator	k_p	T_i	T_d
P	$0.5 k_{kr}$		
PI	$0.45 k_{kr}$	$T_{kr} / 1.2$	
PID _R	$0.6 k_{kr}$	$0.5T_{kr}$	$0.125T_{kr}$
PID _{Sz}	$0.3 k_{kr}$	$0.25T_{kr}$	$0.25T_{kr}$

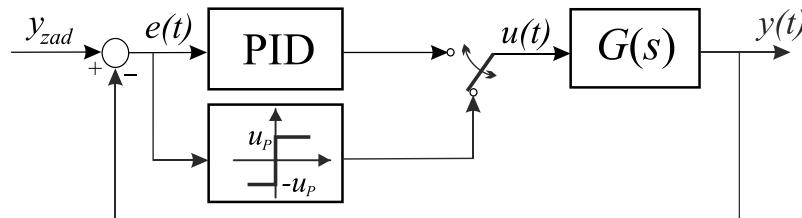
Zwróćmy uwagę na przejrzystą interpretację tych reguł. Dla regulatora P, $k_p = 0.5 k_{kr}$ oznacza po prostu przyjęcie zapasu fazy 6dB. Wprowadzenie całkowania (regulator PI) destabilizuje nieco układ, stąd mniejsze wzmacnienie k_p . Wprowadzenie różniczkowania odwrotnie, wprowadza dodatnie przesunięcie fazowe poszerzające pasmo przenoszenia układu, stąd możliwe zwiększenie wzmacnienia i zmniejszenie T_i (a więc zwiększenie siły całkowania). Cechą charakterystyczną omówionej metody doboru nastaw, i główną przyczyną jej popularności do dnia dzisiejszego, jest brak potrzeby tworzenia jakiegokolwiek modelu samego obiektu regulacji – modelujemy jedynie podstawowe parametry układu regulacji na granicy stabilności.

Ziegler i Nichols sformułowali swoje reguły bazując na szeregu eksperymentów praktycznych z wykorzystaniem regulatorów PID pneumatycznych firmy Taylor. Struktura algorytmu PID tych regulatorów była zbliżona do struktury równoległej – parametry w ostatnim wierszu Tabeli 8.1 powstały przez odpowiednie przeliczenie parametrów z wiersza poprzedniego (zob. rozdz. 8.2.1). Regulatory nastrojone wg powyższych reguł z reguły dają dobre przebiegi w układzie zamkniętym przy eliminacji wpływu zakłóceń oddziałujących na wejście obiektu, w zakresie dla $2 < k \cdot k_{kr} < 20$, dla procesów typu całkującego. Dla innych obiektów i wymuszeń, a szczególnie przy nadążaniu za skokowymi zmianami wartości zadanej, dla procesów statycznych z dominującym opóźnieniem, prowadzą na ogół do przebiegów bardziej agresywnych, tzn. oscylacyjnych z dość dużymi przeregulowaniami – co nie jest często pożądane i wymaga korekt nastaw Z-N (będących dobrym punktem startowym do korekt). Stąd istnieje szereg propozycji innych reguł doboru nastaw wg cech przebiegu krytycznego, np. reguły TLC (Tyreus-Luybena) [5] podane w Tabeli 8.2. Dają one przebiegi wolniejsze, z mniejszymi przeregulowaniami, zapewniające większe zapasy stabilności układu regulacji.

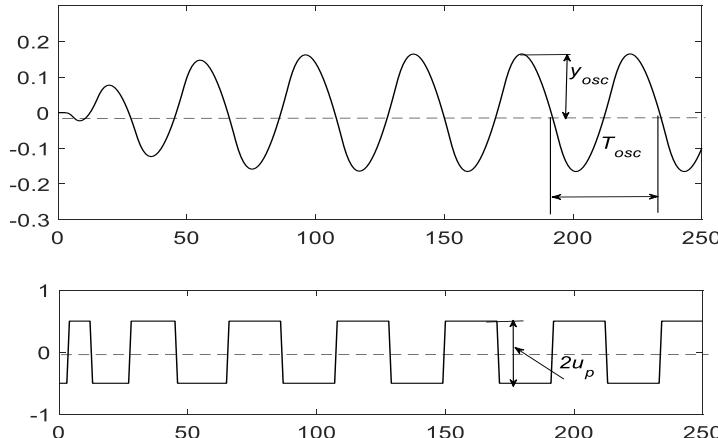
Tabela 8.2. Reguły Tyreusa-Luybena wg cech przebiegu krytycznego (1992 r.).

Regulator	k_p	T_i	T_d
PI	$k_{kr}/3.2$	$2.2T_{kr}$	
PID	$k_{kr}/2.2$	$2.2T_{kr}$	$T_{kr}/6.3$

Eksperyment Zieglera-Nicholsa z doprowadzeniem układu regulacji do granicy stabilności poprzez zwiększanie wzmacnienia regulatora P jest niekorzystny z dwóch powodów. Po pierwsze, niewygodne jest szukanie wzmacnienia krytycznego metodą prób i błędów. Po drugie, ten sposób prowadzi często do dużych amplitud sterowania i wyjścia. Stąd nie zawsze jest możliwy – takie wzbudzanie procesów przemysłowych może doprowadzić do znaczących strat produkcyjnych (odchodzenie od punktów pracy), a co ważniejsze, do awarii a nawet zniszczenia instalacji, do czego obsługa procesów nie może dopuścić ze względów bezpieczeństwa. Stąd duże znaczenie dla praktycznego wyznaczania parametrów przebiegu krytycznego ma zaproponowana w roku 1984 przez Åströma i Hägglunda *metoda przekaźnikowa* [1]. Eksperyment Åströma i Hägglunda polega na zastąpieniu regulatora PID przekaźnikiem dwupołożeniowym o określonej, niewielkiej amplitudzie wyjścia (co najmniej ok. czterokrotnie większej od poziomu szumów pomiarowych wyjścia), zob. Rys. 8.18.



Rys. 8.18. Struktura układu dla wyznaczania parametrów krytycznych metodą Åströma-Hägglunda.



Rys. 8.19. Przykładowe przebiegi w układzie regulacji z rysunku poprzedniego.

W takim układzie regulacji powstaje cykl graniczny, przy czym dla obiektów wieloinercyjnych przebieg wyjścia jest zbliżony do sinusoidy, jak to pokazano na Rys. 8.19 dla przykładowego obiektu o transmitancji

$$G(s) = \frac{5e^{-3s}}{(1+40s)(1+20s)}.$$

Mierząc amplitudę przebiegu wyjściowego y_{osc} i okres oscylacji T_{osc} można wyznaczyć parametry przebiegu krytycznego, z wzoru:

$$k_{kr} \approx \frac{4u_p}{\pi y_{osc}}, \quad T_{kr} = T_{osc}. \quad (8.27)$$

Wzór na k_{kr} powstaje przez porównanie amplitud sygnału wyjściowego (traktowanego jako sinusoidalny) z amplitudą pierwszej harmonicznej (o okresie T_{kr}) w rozkładzie prostokątnego sygnału sterującego z przekaźnika w szeregu Fouriera. Jeśli sygnał wyjściowy obiektu jest bardziej zbliżony do trójkątnego (piłokształtnego) niż sinusoidalnego, to wzmacnienie krytyczne uzyskane z wzoru (8.27) należy nieco zwiększyć (o ok. 20%); jeśli natomiast jest bardziej prostokątny niż sinusoidalny, to należy nieco zmniejszyć.

Reguły Zieglera-Nicholsa dla modelu obiektu typu FOPD

Eksperyment Zieglera-Nicholsa nie zawsze był możliwy do wykonania, szczególnie na przemysłowych obiektach rzeczywistych. Stąd już Ziegler i Nichols sformułowali swoje reguły (tzn. wg kryterium QDR) również dla standardowych modeli obiektów. W Tabeli 8.3 przytaczamy je dla obiektu FOPD, gdzie do klasycznych reguł dla regulatora PID równolegle dodaliśmy przeliczone reguły dla regulatora o strukturze szeregowej.

Tabela 8.3. Reguły Zieglera-Nicholsa dla obiektu FOPD.

Regulator	k_p	T_i	T_d
P	$\frac{T}{kT_0}$		
PI	$0.9 \frac{T}{kT_0}$	$\frac{T_0}{0.3}$	
PID _R	$1.2 \frac{T}{kT_0}$	$2T_0$	$0.5T_0$
PID _{Sz}	$0.6 \frac{T}{kT_0}$	T_0	T_0

Przebiegi uzyskiwane w układach zamkniętych z tak dobranymi nastawami odpowiadają przebiegom dla nastaw dobranych na podstawie eksperymentu Zieglera-Nicholsa, ze wszystkimi zaletami i wadami, dając na ogół poprawne rezultaty w zakresie $0.1 < T_0/T < 0.7$.

Przykładowy zestaw reguł dla obiektu IPD

Tabela 8.4. Reguły doboru nastaw dla obiektu astatycznego typu IPD.

Kryterium	Regulator	Optymalne nastawy		
		$k_p kT_0$	$\frac{T_i}{T_0}$	$\frac{T_d}{T_0}$
20% przeregulowania, minimum czasu regulacji	P	0.7		
	PI	0.7	3	
	PID	1.1	2	0.37
0% przeregulowania, minimum czasu regulacji	P	0.37		
	PI	0.46	5.75	
	PID	0.65	5	0.23
minimum ISE	PI	1.045	4.3	
	PID	1.365	1.6	0.51

Istnieje wiele innych prostych reguł doboru nastaw regulatorów PID, o różnej strukturze, zarówno dla dla obiektu FOPD jak i innych, wg różnych kryteriów, zob. np. [4,5]. Przykładowo, w Poradniku Inżyniera „Automatyka” [3] zamieszczono krzywe i tabele doboru nastaw regulatorów dla obiektów FOPD i IPD, wg kryteriów minimum czasu regulacji i przeregulowania 0% oraz 20%, jak i według kryterium ISE; reguły doboru nastaw (dla struktury równoległej) znaleźć można też w klasycznym polskim podręczniku [2], za którym podajemy Tabelę 8.4 z regułami dla obiektu IPD.

Regulację PID stosuje się w praktyce przemysłowej powszechnie, również do obiektów elektromechanicznych, do serwomechanizmów, robotów – wówczas nastawy dobiera się przede wszystkim pod kątem odpowiedniej jakości nadążania za wartością zadaną, tłumienie zakłóceń ma na ogół mniejsze znaczenie. Modelowanie obiektów jest też nieco odmienne – łatwiej jest uzyskiwać dobre modele fenomenologiczne (z praw fizyki), łatwiej o dobrą identyfikację modeli, opóźnienie odgrywa najczęściej mniejszą rolę. Dla potrzeb doboru nastaw regulatorów PID modele te najczęściej upraszcza się do postaci prostych, standardowych transmitancji [5].

Przykład 8.2.

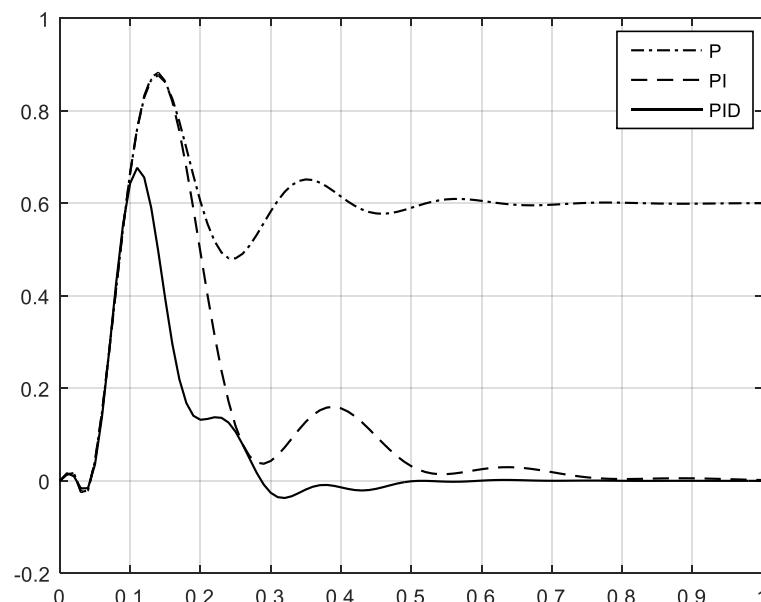
Dokonajmy doboru nastaw regulatorów, kolejno, P, PI i PID, metodą Zieglera-Nicholsa (tzn. wg kryterium QDR), dla obiektu o transmitancji FOPD postaci

$$G(s) = \frac{3e^{-0.05s}}{1 + 0.2s}$$

tzn. dla $k = 3$, $T_0 = 0.05$ i $T = 0.2$. Czyli mamy $T/T_0 = 4$.

Korzystając z Tabeli 8.2 wyznaczamy kolejno:

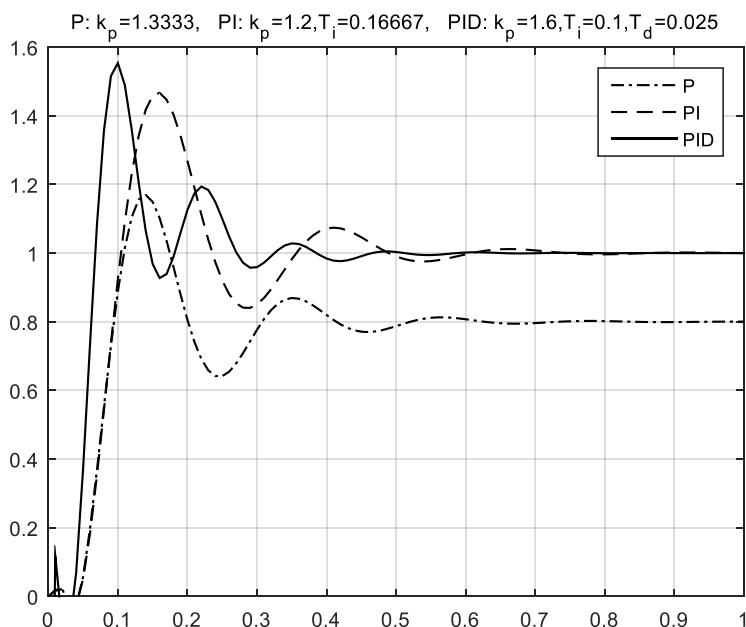
- dla regulatora P: $k_p = 4/3$,
- dla regulatora PI: $k_p = 1.2$, $T_i = 1/6$,
- dla regulatora PID: $k_p = 1.6$, $T_i = 0.1$, $T_d = 0.025$.



Rys. 8.20. Przebiegi wyjścia (modułu uchybu regulacji) po skoku jednostkowym zakłócenia na wejściu obiektu w układach regulacji z regulatorami P, PI i PID.

Tabela 8.5. Zapasy modułu i fazy układów regulacji generujących przebiegi podane na Rys. 8.20.

regulator	zapas modułu	zapas fazy
P	4.78 dB	49°
PI	4.84 dB	38°
PID	2.92 dB	49°

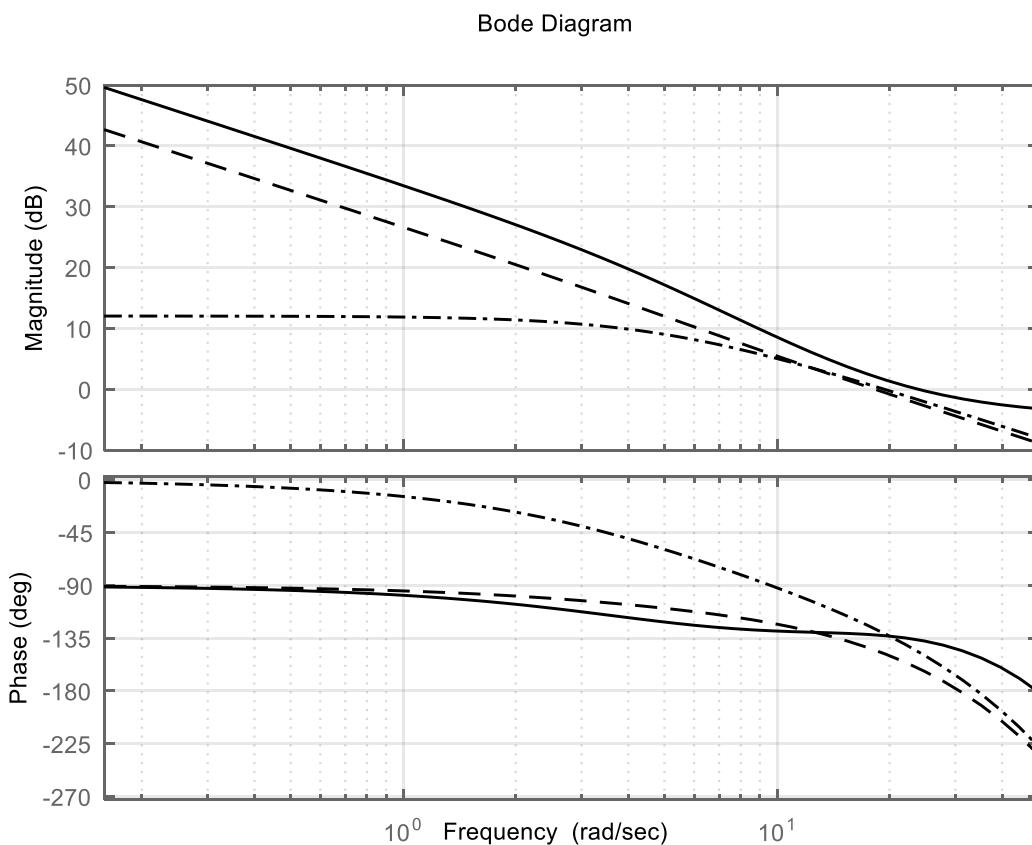


Rys. 8.21. Przebiegi wyjścia po skoku jednostkowym wartości zadanej w układach regulacji z regulatorami P, PI i PID.

Na Rys. 8.20 zamieszczono przebiegi wyjścia obiektu po skokowej (jednostkowej) zmianie zakłócenia na wejściu obiektu (w strukturze jak na Rys. 8.17), w układzie regulacji z obiektem $G(s)$ i kolejno, reglatorem P (krzywa przerywano-kropkowana), reglatorem PI (krzywa przerywana) i reglatorem PID (krzywa ciągła), dla wartości zadanej wyjścia równej zero. We wszystkich przypadkach następuje około czterokrotny zanik oscylacyjności (QDR równy w przybliżeniu 0.25), co potwierdza prawidłowość doboru nastaw. Przypomnijmy, że nastawy Z-N odpowiadają kryterium QDR, są to przebiegi na ogół dość silnie oscylacyjne, często nie spełniające typowych zaleceń dotyczących bezpiecznych zapasów stabilności. Jest tak i dla przebiegów przedstawionych na Rys. 8.20, dla których zapasy stabilności podane są w Tabeli 8.5.

Dla porównania, wykonano również symulację w wymienionych układach regulacji (przy niezmienionych nastawach regulatorów) dla jednostkowego skoku wartości zadanej. Wyniki pokazano na Rys. 8.21. Przebiegi są teraz silnie oscylacyjne, dla uzyskania bardziej tłumionych przebiegów nastawy powinny zostać dobrane nieco inaczej. Jednak przyjęte

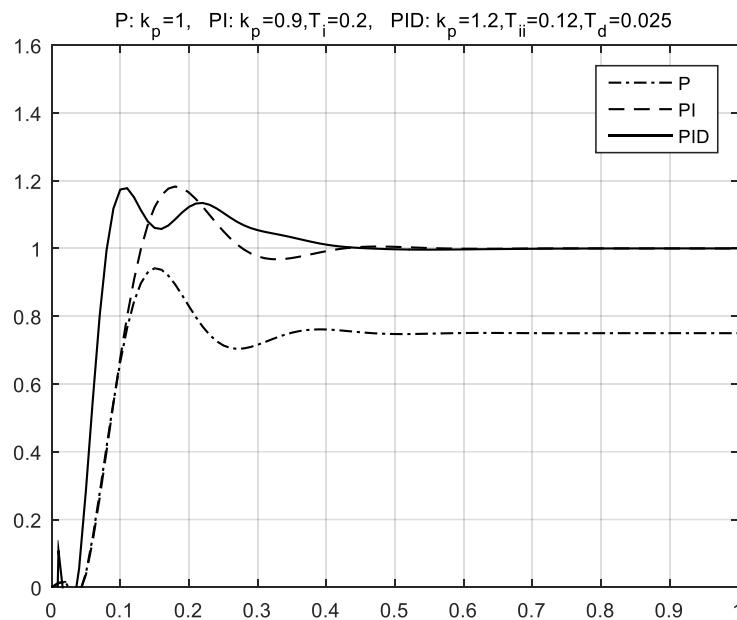
nastawy są dobrymi wartościami przybliżonymi (wyjściowymi dla korekty), wystarczy np. niewielkie stłumienie wzmacnienia regulatorów dla uzyskania mniejszych przeregulowań. Zwraca uwagę nieakceptowalna jakość regulacji z regulatorem P, z dużymi uchybami ustalonimi – na skutek braku całkowania w układzie regulacji i niemożliwości, ze względu na cechy obiektu, zastosowania dużego wzmacnienia regulatora. Widać też, że regulator PID szybciej sprowadza uchyb w otoczenie wartości zadanej. Uzasadnienie nie tego faktu można znaleźć na Rys. 8.20, gdzie zamieszczono charakterystyki częstotliwościowe transmitancji układu otwartego, dla wszystkich trzech rozważanych przypadków. Widać, że regulator PID, w stosunku do regulatora PI, ma nieco większe pasmo przenoszenia, ma też większy zapis fazy.



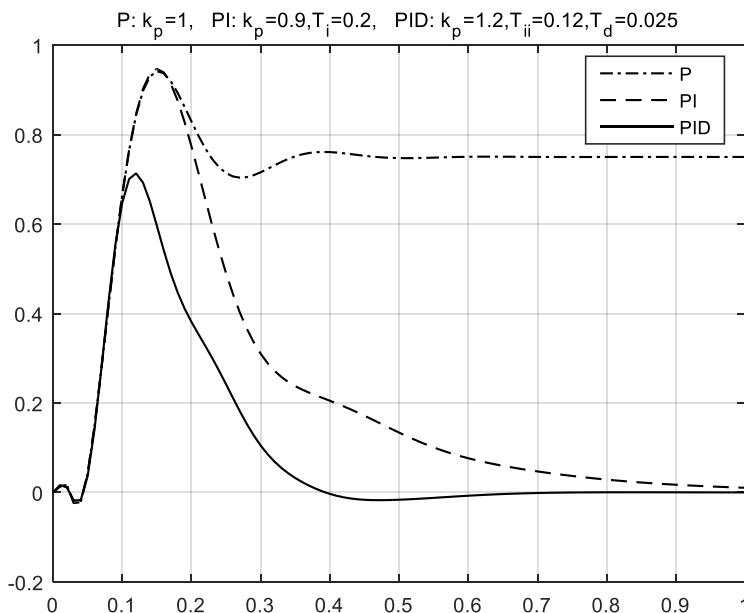
Rys. 8.22. Charakterystyki częstotliwościowe transmitancji układu otwartego rozważanych układów regulacji z regulatorami P (krzywe przerywano-kropkowane), PI (krzywe przerywane) i PID (krzywe ciągłe).

Na koniec, dokonano korekty nastaw dla uzyskania mniej oscylacyjnych (z przeregulowaniem 20%) przebiegów wyjścia obiektu po skoku wartości zadanej i jednocześnie poprawy zapisów stabilności. Skorygowano nastawy regulatorów do wartości:

- dla regulatora P: $k_p = 1$ (poprzednio $k_p = 4/3$),
- dla regulatora PI: $k_p = 0.9$, $T_i = 0.2$ (poprzednio $k_p = 1.2$, $T_i = 1/6$),
- dla regulatora PID: $k_p = 1.2$, $T_i = 0.12$, $T_d = 0.025$ (poprzednio $k_p = 1.6$, $T_i = 0.1$, $T_d = 0.025$).



Rys. 8.23. Przebiegi wyjścia po skoku jednostkowym wartości zadanej w układach regulacji P/PI/PID ze skorygowanymi nastawami.



Rys. 8.24. Przebiegi wyjścia (modułu uchybu regulacji) po skoku jednostkowym zakłócenia na wejściu obiektu w układach regulacji P/PI/PID ze skorygowanymi nastawami.

Odpowiedzi skokowe w układach regulacji ze skorygowanymi nastawami pokazano na Rys. 8.23 i Rys. 8.24, zaś zapasy modułu i fazy w Tabeli 8.5. Uzyskano założone, mniej oscylacyjne przebiegi wyjścia po skoku wartości zadanej. Natomiast skok zakłócenia na wejściu obiektu jest teraz wolniej tłumiony, por. Rys. 8.20 z Rys. 8.24. Zapasy stabilności są większe – zapasy modułu w zasadzie spełniają typowe kryterium, zaś zapasy fazy z nadmiarem.

Tabela 8.5. Zapasy modułu i fazy układów regulacji P/PI/PID ze skorygowanymi nastawami.

regulator	zapas modułu	zapas fazy
P	7.28 dB	68.9°
PI	7.34 dB	51.3°
PID	5.55 dB	53.7°

8.4.3. Dobór nastaw metodą SIMC[†]

W rozdziale tym przedstawimy prostą regułę doboru nastaw regulatorów PI i PID dla modeli inercyjnych pierwszego i drugiego rzędu z opóźnieniem (FOPD i SOPD), metodą SIMC (*simple internal model control*), będącą prostą realizacją ogólniejszej metody IMC (*internal model control*).

Metodę SIMC można stosować do obiektów inercyjnych (wieloinercyjnych) z opóźnieniem, została ona opublikowana w 2001 roku przez Skogestada w pracy pod znanym tytułem „*Probably the best simple PID tuning rules in the world*” [6].

Zakłada się, że dla doboru nastaw regulatora PID obiekt jest modelowany jako inercyjny drugiego rzędu z opóźnieniem (SOPD inercyjny, FOPD będzie przypadkiem szczególnym):

$$G(s) = \frac{ke^{-sT_0}}{(1+sT_1)(1+sT_2)} \quad (8.28)$$

Celem jest natomiast takie zaprojektowanie nastaw, aby układ zamknięty z regulatorem PID zachowywał się podobnie jak obiekt FOPD o zadanej stałej czasowej τ , tzn.

$$\left(\frac{Y}{Y_{zad}} \right)_{pozadane} = \frac{ke^{-sT_0}}{1+s\tau} \quad (8.29)$$

Stała czasowa τ jest tu *parametrem dostrajalnym*. Zalecane są wartości z zakresu $T_0 \leq \tau < \infty$, co powinno zapewniać wystarczające zapasy stabilności. Przy tych założeniach wyprowadzone są (co podamy dalej) następujące reguły doboru nastaw, nazwane *regułami SIMC (simple internal model control)*:

$$k_p = \frac{1}{k} \frac{T_1}{\tau + T_0} \quad (8.30)$$

$$T_i = \min\{ T_1, \frac{4T_1}{kk_p} \} = \min\{ T_1, 4(\tau + T_0) \} \quad (8.31)$$

$$T_d = T_2 \quad (8.32)$$

Parametr dostrajany τ jest typowym *parametrem kompromisu*:

- mniejsze wartości τ oznaczają szybsze przebiegi i lepsze tłumienie zakłóceń,

[†] Materiał dodatkowy

- zwiększanie wartości τ powoduje wolniejsze przebiegi, ale lepszą odporność (większe zapasy stabilności) i mniejsze zmiany (przyrosty) sygnału sterującego.

Jako dobry kompromis między szybkością i odpornością zapewniający jeszcze szybkie przebiegi autor zaleca $\tau = T_0$, z czego wynikają reguły:

$$k_p = \frac{0.5T_1}{kT_0}, \quad T_i = \min\{T_1, 8T_0\}, \quad T_d = T_2 \quad (8.33)$$

W omawianej metodzie, stosowanie drugiej stałej czasowej T_2 i regulatora PID zalecane jest jedynie wtedy, gdy jest ona istotna (co najmniej $T_2 > T_0$), w konsekwencji dla modeli typu FOPD zalecany jest regulator PI z nastawami dobieranymi jak powyżej (przy $T_d = 0$).

Wyprowadzenie reguł SIMC:

Oznaczając transmitancję regulatora przez $R(s)$, mamy

$$\frac{Y}{Y_{zad}} = \frac{G(s)R(s)}{1 + G(s)R(s)}$$

skąd bezpośrednio

$$R(s) = \frac{1}{G(s)} \frac{1}{\frac{1}{(Y/Y_{zad})} - 1} = \frac{(1+sT_1)(1+sT_2)}{k} \frac{1}{(\tau s + 1 - e^{-T_0 s})}$$

Zastępując teraz opóźnienie aproksymacją Taylora 1 rzędu

$$e^{-T_0 s} \approx 1 - \tau s$$

dostajemy

$$R(s) = \frac{(1+sT_1)(1+sT_2)}{k} \frac{1}{(\tau + T_0)s} = \frac{T_1}{k(\tau + T_0)} \left(1 + \frac{1}{T_1 s}\right) (1+sT_2) \quad (8.34)$$

Ostatni wzór przedstawia regulator szeregowy PID (idealny) z nastawami SIMC przy $T_i = T_1$. Modyfikacja tej zależności do postaci (8.31) wynika z wymagań lepszego tłumienia zakłóceń, zob. [6].

Autor metody SIMC zaleca następujący sposób *szybkiego upraszczania*, do postaci FOPD lub SOPD inercyjnej, transmitancji obiektu wyższego rzędu (nieoscylacyjnej), postaci

$$G(s) = k e^{-\theta s} \frac{1 - T_1 s}{\prod_{i=1}^n (1 + sT_{i0})} = \frac{k e^{-\theta s}}{(1 + sT_{10})} \frac{1 - T_1 s}{\prod_{i=2}^n (1 + sT_{i0})} \quad (8.35)$$

gdzie stałe czasowe indeksowane są od największej do najmniejszej.

Sposób upraszczania, zwany niekiedy „regułą połowy” („half rule”), jest następujący:

1. Aproksymacja stałych czasowych T_{i0} :

- dla modelu FOPD wybieramy:

$$T_1 = T_{10} + 0.5T_{20}, \quad T_0 = \theta + 0.5T_{20} + \sum_{i=3}^n T_{i0} \quad (8.36)$$

- dla modelu SOPD wybieramy pierwszą (większą) stałą czasową równą największej stałej czasowej modelu oryginalnego, a drugą stałą czasową i opóźnienie jak stałą czasową i opóźnienie w modelu FOPD:

$$T_1 = T_{10}, \quad T_2 = T_{20} + 0.5T_{30}, \quad T_0 = \theta + 0.5T_{30} + \sum_{i=4}^n T_{i0} \quad (8.37)$$

2. Uwzględnienie nieminimalnofazowego (dodatniego) zera (licznika) T_l – zwiększenie efektywnego opóźnienia:

$$T_0 = T_0 + T_l$$

W liczniku transmitancji (8.34) mogą również występować dodatkowe zera stabilne, tj. czynniki $(1+T_{lj})$. Wówczas małe wartości T_{lj} (mniejsze od połowy opóźnienia) odejmujemy od efektywnego opóźnienia wyliczonego w podany wyżej sposób, a większe odejmujemy od pierwszej stałej czasowej. Autor metody SIMC twierdzi, że taki sposób szybkiego upraszczania modelu daje w połączeniu z doborem nastaw regulatora PID metodą SIMC dobre rezultaty.

Metodą IMC (internal model control), tzn. zakładania pożądanej dynamiki układu zamkniętego i wynikającego stąd odpowiedniego doboru nastaw regulatora PID, można również dobierać nastawy regulatora PID dla obiektów innej postaci, np. dwuincrecyjnych oscylacyjnych [4].

Przykład 8.3.

Rozważany jest następujący model (dokładny) obiektu:

$$G(s) = k \frac{1 - 0.3s}{(1 + 2s)(1 + s)(1 + 0.4s)(1 + 0.2s)(1 + 0.05s)^2} \quad (8.38)$$

Aproxymując tę transmitancję modelem uproszczonym FOPD (regułą połowy) dostajemy:

$$T = 2 + 0.5 \cdot 1 = 2.5; \quad T_0 = 0.5 \cdot 1 + 0.4 + 0.2 + 2 \cdot 0.05 + 0.3 = 1.5$$

Następnie wyznaczamy z wzorów (8.30) i (8.31) nastawy regulatora PI:

$$kk_p = \frac{0.5T}{T_0} = \frac{1.25}{1.5} = 0.8333; \quad T_i = T = 2.5$$

Aproxymując transmitancję (8.38) modelem uproszczonym SOPD (regułą połowy) mamy:

$$T_1 = 2; \quad T_2 = 1 + 0.5 \cdot 0.4 = 1.2; \quad T_0 = 0.5 \cdot 0.4 + 0.2 + 2 \cdot 0.05 + 0.3 = 0.8$$

Wyznaczając teraz, z (8.30), (8.31) i (8.32) nastawy regulatora PID (szeregowego) dostajemy:

$$kk_p = \frac{0.5T_1}{T_0} = \frac{1}{0.8} = 1.25; \quad T_i = T_1 = 2; \quad T_d = T_2 = 1.2$$

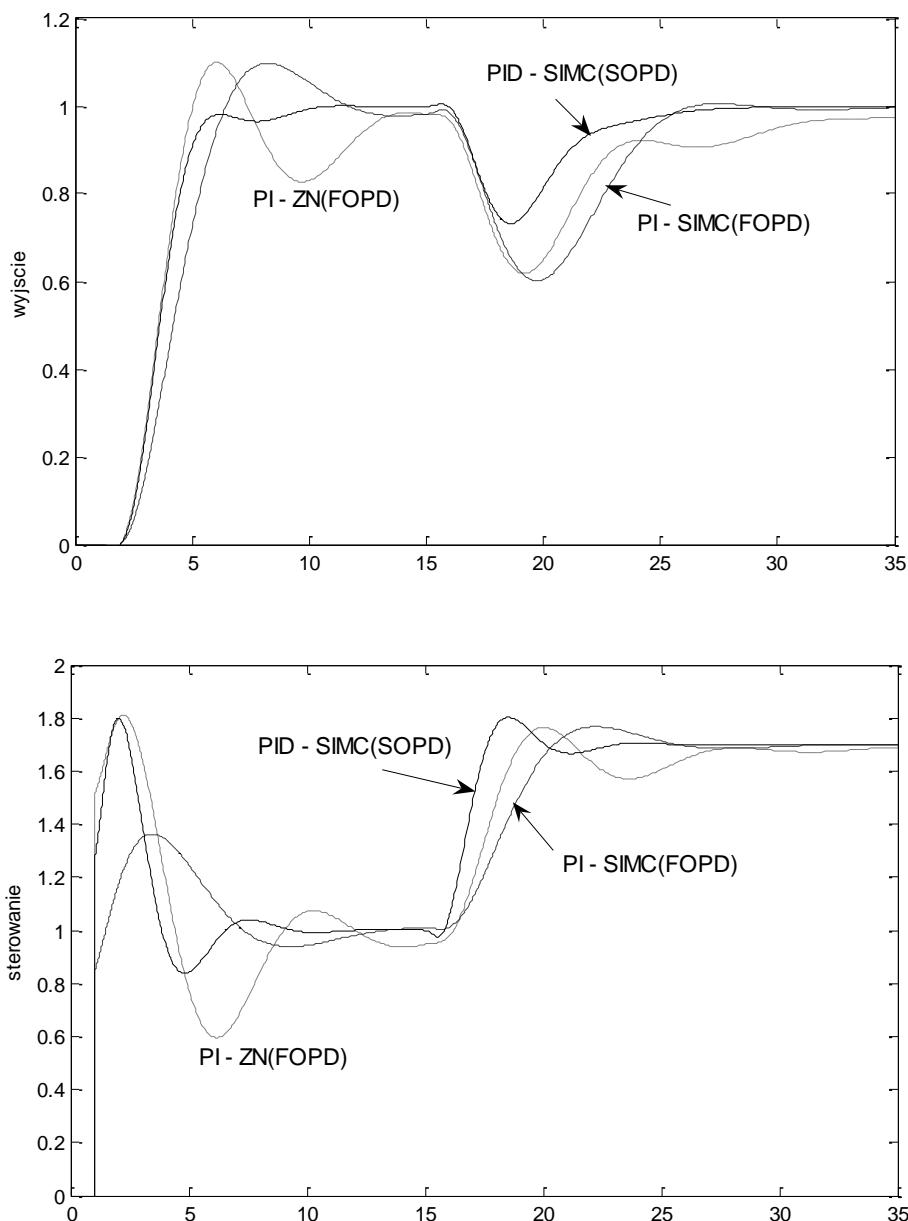
Dla porównania, dla obiektu FOPD dobrano też nastawy wg reguł Zieglera-Nicholsa dla regulatora PI: $kk_p = 1.5$, $T_i = 5$.

Na rysunku 8.25 przedstawiono wyniki symulacji zamkniętego układu regulacji z obiektem modelowanym jako transmitancja (8.38), oraz regulatorami PI, z nastawami SIMC i wg Zieglera-Nicholsa, oraz z regulatorem PID z nastawami SIMC dobranymi powyżej dla obiektu SOPD. Regulator PID zrealizowano w strukturze szeregowej zmodyfikowanej, tzn. bez różniczkowania wartości zadanej (uniknięcie uderzenia sterowania po skoku wartości zadanej), z realizowalnym różniczkowaniem ($\alpha=10$), tj. z regulatorem postaci

$$U(s) = k_p \frac{T_i s + 1}{T_i s} \left(Y_{zad}(s) - \left(1 + \frac{T_d s}{(T_d / \alpha) s + 1} \right) Y(s) \right)$$

W czasie trwania symulacji (od zera do 35 chwil czasowych) w chwili nr 1 nastąpił skok wartości zadanej wyjścia od zera do jedności, a w chwili nr 15 skok wartości zakłócenia od zera do -0.7 (zakłócenie oddziaływało na wejście obiektu).

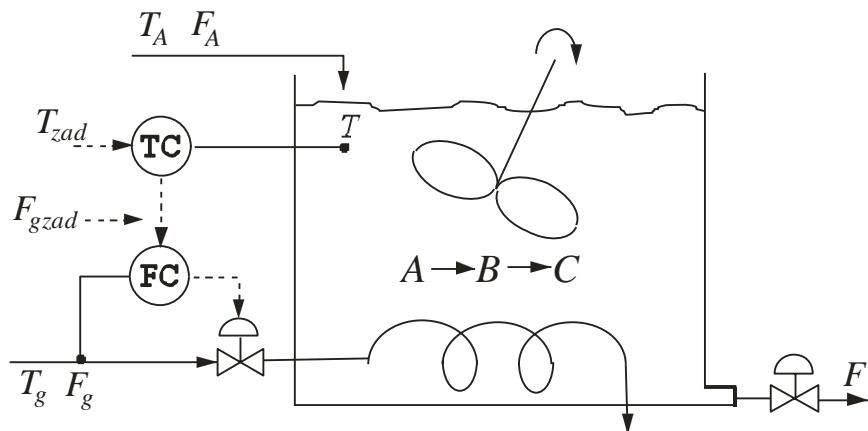
Oceniając przedstawione przebiegi trzeba stwierdzić, że wszystkie układy regulacji pracują prawidłowo, zarówno przy nadążaniu za wartością zadaną, jak i przy tłumieniu zakłócenia. Najlepsze rezultaty dało modelowanie SOPD „metodą połowy” i dobór nastaw regulatora PID (metodą SIMC). Z przebiegów z regulatorami PI, gorszy jest ten z nastawami Zieglera-Nicholsa, przebieg jest bardziej oscylacyjny i wolniej zeruje uchyb ustalony, ze względu na większą wartość stałej czasowej całkowania. Połączenie modelowania uproszczonego FOPD „metodą połowy” i dobór nastaw wg reguł Zieglera-Nicholsa nie dały w przykładzie najlepszych rezultatów.



Rys. 8.25. Przebiegi wyjścia i sterowania dla obiektu z przykładu 8.3, z regulatorami PI i PID.

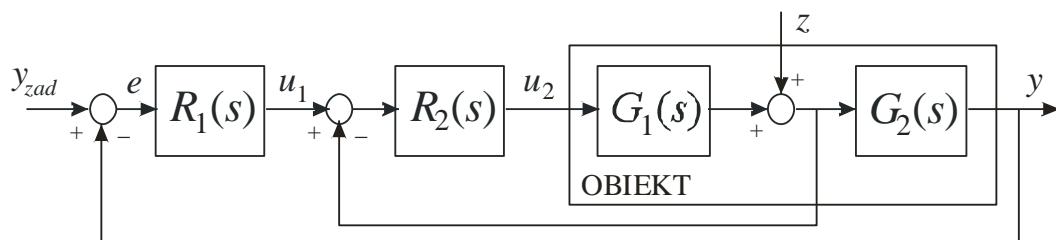
8.5. Regulacja kaskadowa, kompensacja zakłócenia mierzonego

W systemach regulacji obiektów przemysłowych powszechnie występują układy regulacji kaskadowej PID. Typowym przykładem jest tu regulacja temperatury w sytuacji zmienności natężenia przepływu (istotne zakłócenie) czynnika grzewczego, zilustrowana na Rys. 8.26.



Rys. 8.26. Regulacja kaskadowa temperatury w reaktorze.

Regulator nadrzędny temperatury (TC) ustawia wartość zadaną przepływu substancji grzejnej odpowiadającą temperaturze zadanej, a regulator podrzędny przepływu stabilizuje przepływ na tej wartości oddziaływaniem na zawór – przeciwdziałając w ten sposób wahaniom przepływu czynnika grzewczego znacznie szybciej, niż następowałoby to w układzie jednego regulatora (temperatury) oddziałyującego bezpośrednio na zawór. Struktura blokowa układu regulacji kaskadowej przedstawiona jest na Rys. 8.27, gdzie zaznaczono jedynie zakłócenie reguleowane w pętli wewnętrznej (inne zakłócenia, oddziałyujące na drugą część obiektu, też mogą występować). Regulatory R_1 i R_2 to zwykle regulatorzy typu PID.



Rys. 8.27. Struktura układu regulacji kaskadowej.

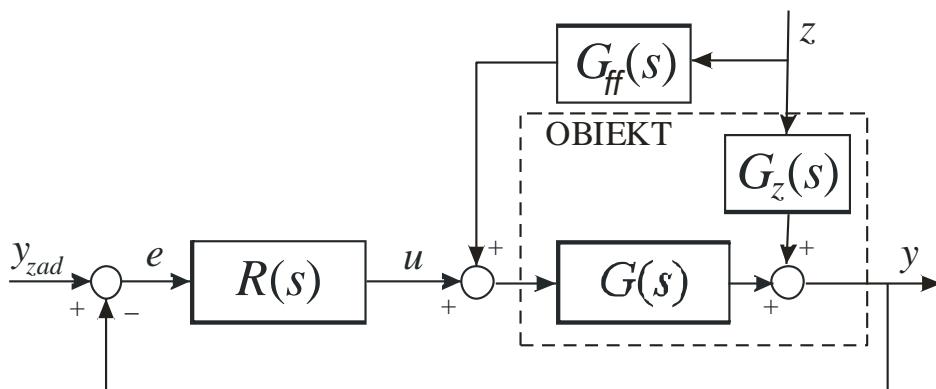
Jak wynika z przedstawionego przykładu i ogólnej struktury regulacji kaskadowej, stosuje się ją, gdy dostępna jest pomocnicza wielkość pomiarowa, wrażliwa na zmiany istotnego zakłócenia oddziałującego na pierwszą część obiektu – wówczas możemy tę część obiektu objąć dodatkowym wewnętrznym sprzężeniem zwrotnym, znacznie efektywniej kompensującym wpływ tego zakłócenia, niż czyniłby to jeden regulator (główny R_1). Ponadto, można pokazać, że działanie pętli wewnętrznej regulacji ma właściwości linearyzujące nieliniową charakterystykę części obiektu objętej tym sprzężeniem, tym silniejsze im większa jest wartość wzmacnienia regulatora R_2 .

Dobór nastaw regulatorów w strukturze kaskadowej przeprowadza się sekwencyjnie, w

sposób standardowy:

- najpierw dobieramy nastawy regulatora wewnętrznego (podrózdnego, R_2) dla modelu części obiektu G_1 ,
- potem dobieramy nastawy regulatora głównego R_1 dla modelu obiektu objętego wewnętrzną pętlą sprzężenie zwrotnego z regulatorem R_2 (tzn. obiektu zastępczego z wejściem u_1 i wyjściem y na Rys. 8.27).

W sytuacjach, kiedy istotne zakłócenie może być i jest mierzane oraz dysponujemy modelem toru zakłócenia (zakłócenie-wyjście), obowiązkowym rozszerzeniem podstawowej struktury regulacji PID jest dodanie toru bezpośredniej kompensacji wpływu tego zakłócenia, jak to pokazano na rys. Rys 8.28. Tor ten nazywany jest często torem „*feedforward*”, stąd cała struktura przedstawiona na rysunku określana bywa jako struktura „*feedback-feedforward*”. Przykładem mógłby tu być reaktor z Rys. 8.26 w sytuacji istotnie zakłócających pracę reaktora wahań temperatury T_A substancji dopływającej – mierząc tę temperaturę można by zaprojektować tor bezpośredni kompensacji tego zakłócenia, z sygnałem wyjściowym bloku *feedforward* dodawanym do wyjścia regulatora temperatury TC (głównego).



Rys. 8.28. Struktura układu regulacji z bezpośrednią kompensacją zakłócenia (struktura *feedback-feedforward*).

Idealną kompensację uzyskamy, gdy wpływ zakłócenia torem kompensacji, tj. przez G_{ff} i G , będzie kompensował wpływ zakłócenia przez transmitancję zakłócenioową G_z , tzn. kiedy

$$G_z(s)Z(s) + G(s)G_{ff}(s)Z(s) = 0 \quad (8.39)$$

skąd wynika wzór na transmitancję członu idealnie kompensującego:

$$G_{ff}(s) = -\frac{G_z(s)}{G(s)} \quad (8.40)$$

Idealna kompensacja możliwa jest jedynie wtedy, gdy realizowalna fizycznie i stabilna jest odwrotność transmitancji obiektu i, oczywiście, gdy model jest dokładnym opisem obiektu – obie te warunki na ogół nie są spełnione. W szczególności, nie można zrealizować fizycznie stabilnej odwrotności transmitancji obiektu zawierającej opóźnienie i/lub niestabilne zera (w praktyce odwracamy jedynie część transmitancji bez tych elementów), a modelowanie rzeczywistości zawsze jest obarczone błędami. Dlatego też tor bezpośredni kompensacji będzie w praktyce jedynie częściowo redukował wpływ zakłócenia – niemniej jest to bardzo pożąданie, gdyż na pętlę sprzężenie zwrotnego będzie wówczas oddziaływało zakłócenie zredukowane, a więc możliwe będzie znacznie szybsze, efektywniejsze jego tłumienie.

8.6. Regulacja predykcyjna[‡]

Zasada regulacji predykcyjnej

Pierwsze algorytmy regulacji predykcyjnej zostały zastosowane w latach siedemdziesiątych ubiegłego wieku i jako jedyne z tzw. zaawansowanych technik regulacji (*advanced control techniques*) – technik bardziej zaawansowanych niż oparte na standardowym algorytmie PID – odniosły olbrzymi sukces w aplikacjach praktycznych wywierając dominujący wpływ na kierunek rozwoju przemysłowych układów regulacji i sterowania w ostatnich dekadach. Regulacja predykcyjna jest techniką silnie opartą na modelu obiektu (ang. MPC – *Model-based Predictive Control*) i wymagającą znacznie większego nakładu obliczeń niż algorytmy PID, stąd mogła się rozwinąć dzięki postępu techniki mikroprocesorowej. Jest kilka przyczyn sukcesu regulacji predykcyjnej:

- Po pierwsze, algorytmy MPC w sposób naturalny potrafią uwzględniać ograniczenia zarówno sygnałów sterujących, jak i ograniczenia wielkości wyjściowych, decydujące często o jakości, efektywności i bezpieczeństwie produkcji.
- Po drugie, generują aktualne sterowania uwzględniając zarówno wspomniane ograniczenia jak i, dzięki bezpośredniemu wykorzystaniu modelu, interakcje wewnętrzne w obiekcie. Stąd w sposób naturalny można je stosować do obiektów wielowymiarowych, również przy nierównej liczbie wejść sterujących i wielkości regułowanych.
- Po trzecie, są to algorytmy, których zasada działania jest zrozumiała, stosunkowo łatwa do wyjaśnienia personelowi inżynierijnemu jak i operatorskiemu – aspekt bardzo istotny przy wprowadzaniu nowych technik do praktyki przemysłowej.

Zasadą działania regulatora predykcyjnego (MPC) jest wyznaczanie sterowania w taki sposób, aby minimalizować błędy regulacji na odcinku czasu od chwili obecnej do pewnego momentu w przyszłości (odcinek ten nazywamy *horyzontem predykcyjnym*), przewidując przebieg przyszłych wyjść obiektu w oparciu o model procesu – stąd jakość regulacji predykcyjnej dość silnie zależy od jakości wykorzystywanego modelu. Realizacja przedstawionej zasady działania regulatora MPC następuje poprzez rozwiązywanie zadania optymalizacji dynamicznej, w którym podstawowymi składnikami funkcji celu są kary za błędy regulacji na horyzoncie predykcyjnym. Jedną z typowych sformułowań zadania optymalizacji regulatora MPC jest zadanie następujące:

$$\min_{u(k|k), u(k+1|k), \dots, u(k+N-1|k)} \{ J(k) = \sum_{p=1}^N \| y^{zad}(k+p|k) - y(k+p|k) \|^2 + \lambda \sum_{p=0}^{N-1} \| \Delta u(k+p|k) \|^2 \}$$

przy ograniczeniach : (8.41)

$$u_{\min} \leq u(k+p|k) \leq u_{\max}, \quad p = 0, \dots, N-1$$

$$-\Delta u_{\max} \leq \Delta u(k+p|k) \leq \Delta u_{\max}, \quad p = 0, \dots, N-1$$

$$y_{\min} \leq y(k+p|k) \leq y_{\max}, \quad p = 1, \dots, N$$

gdzie:

- $y^{zad}(k+p|k)$ – wartość zadana wyjść regułowanych przewidywana na chwilę $k+p$, w chwili k ,
- $y(k+p|k)$ – wartość wyjść regułowanych przewidywana na chwilę $k+p$, w chwili k ,
- $u(k+p|k)$ – wartość sterowań (wyjść regulatora) ustalana na chwilę $k+p$, w chwili k ,

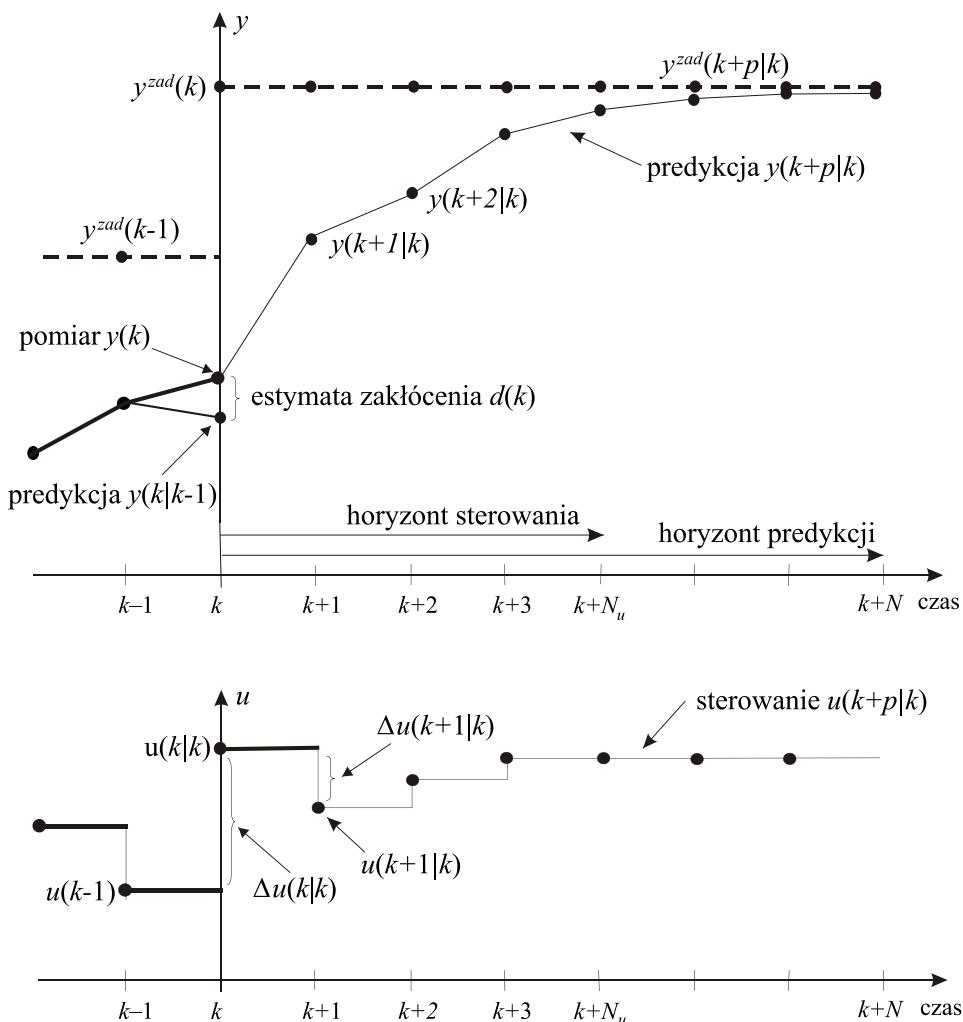
[‡] Materiał uzupełniający

- $\Delta u(k/k) = u(k/k) - u(k-1)$ – przyrost sterowania w chwili bieżącej k (do wyznaczenia), gdzie $u(k-1)$ jest sterowaniem wyznaczonym w chwili poprzedniej ($k-1$ szej) i zastosowanym do sterowania procesem,
- $\Delta u(k+p/k) = u(k+p/k) - u(k+p-1/k)$ – przyrosty sterowań w chwilach obecnej i przyszłych, $p = 0, \dots, N-1$.

Pierwsza suma w funkcji celu $J(k)$ to suma kwadratowych kar za uchyby regulacji $e(k+p/k)$ przewidywane na horyzoncie predykcji,

$$e(k+p|k) = y^{zad}(k+p|k) - y(k+p|k), \quad p=1, \dots, N \quad (8.42)$$

zaś druga suma reprezentuje sumę kar za przyrosty sterowań w kolejnych chwilach na horyzoncie predykcji, jest ona przy tym przemnożona przez współczynnik wagi λ , będący parametrem dostrajanym regulatora. Przyjęcie większej wartości λ oznacza wymaganie mniejszej zmienności sterowań, kosztem większych uchybów regulacji. Zadanie (8.41) sformułowane jest dla ogólnego przypadku obiektu MIMO (multi-input multi-output, tzn. o wielu wejściach i wielu wyjściach), kiedy wejście i wyjście procesu są w ogólności wektorami – stąd w funkcji celu kwadraty norm wektorów. Dla przypadku jednowymiarowego (obiektu SISO) zasada działania regulacji predykcyjnej zilustrowana jest na Rys. 8.29, zob. też [7].



Rys. 8.29. Zasada działania regulacji predykcyjnej.

W sformułowaniu zadania (8.41) sterowania optymalizowane są na całym horyzoncie predykcji, wyznaczane są ich optymalne wartości w chwilach $k, k+1, k+2, \dots, k+N-1$. Natomiast w praktyce, szczególnie jeśli horyzont predykcji N jest długi, to wyznaczamy sterowania jedynie na krótszym odcinku czasu zwany horyzontem sterowania, do chwili $k+N_u-1$, gdzie $N_u < N$, a w chwilach dalszych, pozostałych do końca horyzontu predykcji utrzymujemy wartość sterowań stałą równą wartości z chwili $k+N_u-1$, jak to pokazano na Rys. 8.29. Postępuje się tak, aby zmniejszyć wymiarowość zadania optymalizacji, a tym samym nakład obliczeń na jego rozwiązywanie.

Do sterowania procesem w chwili k , tzn. na odcinku czasu $[kT_p, (k-1)T_p]$, gdzie T_p jest okresem próbkowania, stosowany jest jedynie pierwszy element trajektorii optymalnych sterowań, wartość optymalna sterowania w chwili bieżącej $u(k) = u(k|k)$, zob. Rys. 8.29. W każdej z kolejnych dyskretnych chwil czasu cała procedura zadania regulatora predykcyjnego jest powtarzana: mierzona jest aktualna wartość wyjść procesu, aktualizowane jest i rozwiązywane zadanie optymalizacji dynamicznej (8.41). Kluczowym, najtrudniejszym do obliczenia elementem w tym zadaniu są wektory $y(k+p|k)$ wyjść prognozowanych na horyzoncie predykcji ($p=1,2,\dots,N$), obliczane z wykorzystaniem *modelu procesu regulowanego*, opisującego zależność funkcyjną wyjść od sterowań. W zależności od rodzaju użytego modelu procesu, uzyskuje się różne algorytmy typu MPC. W szczególności:

- Algorytm DMC (Dynamic Matrix Control) – jeden z pierwszych (opracowany w latach siedemdziesiątych XX wieku przez inżynierów firmy Shell) i nadal bardzo popularny w przemyśle algorytm predykcyjny wykorzystuje model liniowy w postaci dyskretnych odpowiedzi skokowych (dyskretne przebiegi wyjść odpowiadające dyskretnym skokom jednostkowym na wejściach procesu).
- Algorytm GPC (Generalized Predictive Control) – popularny, opracowany w latach osiemdziesiątych XX wieku algorytm wykorzystuje model liniowy w postaci transmitancji dyskretnych (równoważny modelowi w postaci liniowych równań różnicowych).
- Model w postaci liniowych lub nieliniowych równań stanu jest również często stosowany, w szczególności modele nieliniowe opisywane są w ten sposób. Pierwszym komercyjnym produktem z modelem w postaci równań stanu był algorytm SMOC (Shell Multivariable Optimizing Controller), opracowany w końcu dekady lat osiemdziesiątych XX wieku.

Zamiast wektora aktualnego i przyszłych sterowań, $u(k+p|k)$, $p = 0, \dots, N_u-1$, jako wektora zmiennych decyzyjnych zadania optymalizacji (8.41) wygodnie jest używać, równoważnie, wektora przyrostów sterowania, $\Delta u(k+p|k) = u(k+p|k) - u(k+p-1|k)$, $p = 0, \dots, N_u-1$, szczególnie przy stosowaniu liniowych modeli procesu. Oznaczmy taki wektor zmiennych decyzyjnych przez $\Delta U(k)$,

$$\Delta U(k) = [\Delta u(k|k) \quad \Delta u(k+1|k) \quad \cdots \quad \Delta u(k+N_u-1|k)]^T \quad (8.43)$$

Predykcja dla modelu w postaci równań stanu

Rozważmy model obiektu w postaci dyskretnych liniowych równań stanu:

$$\begin{aligned} x(k+1) &= \mathbf{A}x(k) + \mathbf{B}u(k) + v(k) \\ y(k) &= \mathbf{C}x(k) \end{aligned} \quad (8.44)$$

gdzie $v(k)$ reprezentuje niemierzalne zakłócenia stanu, w tym błędy modelowanie obiektu rzeczywistego (zakładamy, że równanie wyjścia $y(k) = \mathbf{C}x(k)$ jest dokładne, stąd nie wprowadzamy modelu zakłóceń wyjścia). Estymata zakłócenia jest formułowana w postaci:

$$v(k) = x(k) - [\mathbf{A}x(k-1) + \mathbf{B}u(k-1)] \quad (8.45)$$

tzn. jest to różnica miedzy wartością wyjścia $y(k)$ zmierzona w chwili aktualnej, a wartością wynikającą z prognozy wykonanej stosowanym modelem w chwili poprzedniej.

Jako zmienne zadania optymalizacji przyjmujemy przyrosty sterowania $\Delta u(k+p/k)$ w chwilach $p = 0, \dots, N_u-1$, co jest wygodniejsze przy formułowaniu zadania optymalizacji dynamicznej.

Rekurencyjne obliczanie predykcji stanu:

$$\begin{aligned} x(k+1|k) &= \mathbf{A}x(k) + \mathbf{B}u(k|k) + v(k) = \mathbf{A}x(k) + \mathbf{B}(\Delta u(k|k) + u(k-1)) + v(k) = \\ &= [\mathbf{A}x(k) + \mathbf{B}u(k-1) + v(k)] + \mathbf{B}(\Delta u(k|k)) = \\ &= x^0(k+1|k) + \mathbf{B}(\Delta u(k|k)), \end{aligned}$$

$$\begin{aligned} x(k+2|k) &= \mathbf{A}x(k+1|k) + \mathbf{B}u(k+1|k) + v(k) = \\ &= \mathbf{A}[x^0(k+1|k) + \mathbf{B}\Delta u(k|k)] + \mathbf{B}[u(k-1) + \Delta u(k|k) + \Delta u(k+1|k)] + v(k) = \\ &= [\mathbf{A}x^0(k+1|k) + \mathbf{B}u(k-1) + v(k)] + (\mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + \mathbf{B}(\Delta u(k+1|k)) = \\ &= x^0(k+2|k) + (\mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + \mathbf{B}(\Delta u(k+1|k)) \end{aligned}$$

$$\begin{aligned} x(k+3|k) &= \mathbf{A}x(k+2|k) + \mathbf{B}u(k+2|k) + v(k) = \\ &= \mathbf{A}[x^0(k+2|k) + (\mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + \mathbf{B}\Delta u(k+1|k)] + \\ &\quad + \mathbf{B}[u(k-1) + \Delta u(k|k) + \Delta u(k+1|k) + \Delta u(k+2|k)] + v(k) = \\ &= [\mathbf{A}x^0(k+2|k) + \mathbf{B}u(k-1) + v(k)] + \\ &\quad + (\mathbf{A}^2 + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + (\mathbf{A} + \mathbf{I})\mathbf{B}(\Delta u(k+1|k)) + \mathbf{B}(\Delta u(k+2|k)) = \\ &= x^0(k+2|k) + (\mathbf{A}^2 + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + (\mathbf{A} + \mathbf{I})\mathbf{B}(\Delta u(k+1|k)) + \mathbf{B}(\Delta u(k+2|k)) \end{aligned}$$

i tak dalej aż do predykcji stanu na chwilę $p = N_u$:

$$\begin{aligned} x(k+N_u|k) &= \mathbf{A}x(k+N_u-1|k) + \mathbf{B}u(k+N_u-1|k) + v(k) = \\ &= \mathbf{A}x(k+N_u-1|k) + \mathbf{B}[u(k-1) + \Delta u(k|k) + \dots + \Delta u(k+N_u-1|k)] + v(k) = \\ &= x^0(k+N_u|k) + (\mathbf{A}^{N_u-1} + \dots + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + \\ &\quad + (\mathbf{A}^{N_u-2} + \dots + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k+1|k) + \dots + \mathbf{B}(\Delta u(k+N_u-1|k)) \end{aligned}$$

Dla $p \geq N_u$ nie ma już zmiany sterowania ($\Delta u(k+p/k) = 0$ dla $p = N_u, \dots, N-1$), tzn. zestaw przyszłych przyrostów sterowania nie ulega powiększeniu:

$$\begin{aligned} x(k+N_u+1|k) &= \mathbf{A}x(k+N_u|k) + \mathbf{B}u(k+N_u-1|k) + v(k) = \\ &= \mathbf{A}[x(k+N_u|k) + \mathbf{B}[u(k-1) + \Delta u(k|k) + \dots + \Delta u(k+N_u-1|k)]] + v(k) = \\ &= x^0(k+N_u|k) + (\mathbf{A}^{N_u} + \dots + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + \\ &\quad + (\mathbf{A}^{N_u-1} + \dots + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k+1|k) + \dots + \mathbf{AB}(\Delta u(k+N_u-1|k)) \end{aligned}$$

itd. aż do $p = N$:

$$\begin{aligned} x(k+N|k) &= \mathbf{A}x(k+N-1|k) + \mathbf{B}u(k+N_u-1|k) + v(k) = \\ &= x^0(k+N-1|k) + (\mathbf{A}^{N-1} + \dots + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + \\ &\quad + (\mathbf{A}^{N-2} + \dots + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k+1|k) + \dots + \\ &\quad + (\mathbf{A}^{N-N_u} + \dots + \mathbf{A} + \mathbf{I})\mathbf{B}(\Delta u(k+N_u-1|k)) \end{aligned}$$

gdzie

$$x^0(k+p|k) = \mathbf{A}x^0(k+p-1|k) + \mathbf{B}u(k-1) + v(k), \quad p=1,\dots,N, \quad (8.46)$$

biorąc zaś pod uwagę równanie estymaty zakłócenia (8.45) dostajemy to równanie w postaci

$$\begin{aligned} x^0(k+p|k) &= \mathbf{A}x^0(k+p-1|k) + \mathbf{B}u(k-1) + v(k) = \\ &= \mathbf{A}x^0(k+p-1|k) + \mathbf{B}u(k-1) + x(k) - [\mathbf{A}x(k-1) + \mathbf{B}u(k-1)] = \\ &= \mathbf{A}x^0(k+p-1|k) + x(k) - \mathbf{A}x(k-1) \quad p=1,\dots,N, \end{aligned} \quad (8.47)$$

gdzie $x^0(k|k) = x(k)$.

Predykcję stanu na całym horyzoncie predykcji można zapisać również w postaci wektorowo-macierzowej:

$$\begin{bmatrix} x(k+1|k) \\ x(k+2|k) \\ \vdots \\ x(k+N_u|k) \\ x(k+N_u+1|k) \\ \vdots \\ x(k+N|k) \end{bmatrix} = \begin{bmatrix} x^0(k+1|k) \\ x^0(k+2|k) \\ \vdots \\ x^0(k+N_u|k) \\ x^0(k+N_u+1|k) \\ \vdots \\ x^0(k+N|k) \end{bmatrix} + \mathbf{M} \begin{bmatrix} \Delta u(k|k) \\ \Delta u(k+1|k) \\ \vdots \\ \Delta u(k+N_u-1|k) \end{bmatrix} \quad (8.48)$$

gdzie

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \\ \mathbf{M}_3 \\ \vdots \\ \mathbf{M}_{N_u} \\ \mathbf{M}_{N_u+1} \\ \vdots \\ \mathbf{M}_N \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{0} & \cdots & \mathbf{0} \\ (\mathbf{A}+\mathbf{I})\mathbf{B} & \mathbf{B} & \cdots & \mathbf{0} \\ (\mathbf{A}^2+\mathbf{A}+\mathbf{I})\mathbf{B} & (\mathbf{A}+\mathbf{I})\mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{A}^{N_u-1}+\cdots+\mathbf{A}+\mathbf{I})\mathbf{B} & (\mathbf{A}^{N_u-2}+\cdots+\mathbf{A}+\mathbf{I})\mathbf{B} & \cdots & \mathbf{B} \\ (\mathbf{A}^{N_u}+\cdots+\mathbf{A}+\mathbf{I})\mathbf{B} & (\mathbf{A}^{N_u-1}+\cdots+\mathbf{A}+\mathbf{I})\mathbf{B} & \cdots & (\mathbf{A}+\mathbf{I})\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{A}^{N-1}+\cdots+\mathbf{A}+\mathbf{I})\mathbf{B} & (\mathbf{A}^{N-2}+\cdots+\mathbf{A}+\mathbf{I})\mathbf{B} & \cdots & (\mathbf{A}^{N-N_u}+\cdots+\mathbf{A}+\mathbf{I}) \end{bmatrix} \quad (8.49)$$

Predykcja wyjścia bezpośrednio wynika z predykcji stanu, poprzez równanie wyjścia

$$\begin{aligned} y(k+p|k) &= \mathbf{C}x(k+p|k) = \mathbf{C}[x^0(k+p|k) + \mathbf{M}_p \Delta U(k)] = \\ &= y^0(k+p|k) + \mathbf{C}\mathbf{M}_p \Delta U(k), \quad p=1,\dots,N \end{aligned} \quad (8.50)$$

gdzie $\Delta U(k)$ jest wektorem zmiennych decyzyjnych (8.43) zadania optymalizacji regulatora predykcyjnego. Macierz \mathbf{M} zwana jest macierzą dynamiczną (*dynamic matrix*) i jest ona tzw. macierzą Toeplitza, tzn. ma specyficzną strukturę: następna kolumna (macierzowa) jest kolumną poprzednią przesuniętą w jeden wiersz (macierzowy) do dołu – stąd w istocie trzeba policzyć tylko pierwszą kolumnę (macierzową). Składowe $y^0(k+p|k) = \mathbf{C}x^0(k+p|k)$ nazywane bywają *składowymi swobodnymi* predykcji wyjścia ($p=1,\dots,N$), podczas gdy pozostałe składowiki (związane ze zmiennymi decyzyjnymi) to tzw. *składowe wymuszone*.

Zadanie optymalizacji regulatora predykcyjnego przyjmuje teraz postać:

$$\min_{\Delta U(k)} \{ J(k) = \sum_{p=1}^N \| y^{zad}(k+p|k) - y^0(k+p|k) - \mathbf{CM}_p \Delta U(k) \|^2 + \lambda \sum_{p=0}^{N_u-1} \| \Delta u(k+p|k) \|^2 \}$$

przy ograniczeniach :

$$\begin{aligned} u_{\min} &\leq u(k+p|k) \leq u_{\max}, & p = 0, \dots, N_u - 1 \\ -\Delta u_{\max} &\leq \Delta u(k+p|k) \leq \Delta u_{\max}, & p = 0, \dots, N_u - 1 \\ y_{\min} &\leq y^0(k+p|k) + \mathbf{CM}_p \Delta U(k) \leq y_{\max}, & p = 1, \dots, N \end{aligned}$$

Jest to zadanie z kwadratową, wypukłą funkcją celu i liniowymi ograniczeniami, czyli tzw. zadanie *programowania kwadratowego* (wypukłe), które z numerycznego punktu widzenia jest zadaniem umożliwiającym szybkie i niezawodne rozwiązywanie z wykorzystaniem standardowych algorytmów (procedur).

Przykład 8.4.

W przykładzie pokazane zostanie formułowanie zadania optymalizacji dynamicznej w algorytmie MPC z modelem liniowym w postaci dyskretnych równań stanu, dla przykładowego modelu obiektu SISO.

Rozważany model obiektu regulacji, łącznie z ograniczeniami:

$$\begin{aligned} x(k+1) &= \mathbf{Ax}(k) + \mathbf{Bu}(k) + v(k) = \begin{bmatrix} 0.4 & 0.1 \\ 0.3 & 0.5 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k) + v(k), \quad x, v \in R^2, u \in R, \\ y(k) &= \mathbf{Cx}(k) = [0 \ 1] x(k) = x_2(k), \\ 0 &\leq u(k) \leq 1, \\ y(k) &\geq 0 \end{aligned} \tag{8.52}$$

Parametry regulatora MPC: $N = 4$, $N_u = 2$, λ - parametr strojony. Jako zmienne decyzyjne przyjmiemy przyrosty sterowań

$$\begin{aligned} \Delta u(k|k) &= u(k|k) - u(k-1), \quad \Delta u(k+1|k) = u(k+1|k) - u(k|k), \\ \Delta U(k) &= [\Delta u(k|k) \ \ \Delta u(k+1|k)]^T \end{aligned}$$

Estymata zakłócenia:

$$v(k) = x(k) - [\mathbf{Ax}(k-1) + \mathbf{Bu}(k-1)]$$

Składowe swobodne predykcji stanu obliczamy zgodnie z (8.46), lub (8.47):

$$x^0(k+p|k) = \mathbf{Ax}^0(k+p-1|k) + x(k) - \mathbf{Ax}(k-1) \quad p = 1, \dots, N, \quad x^0(k|k) = x(k)$$

Korzystając teraz z zależności ((8.49) i ((8.50) mamy, kolejno dla $p = 1, 2, 3, 4$:

$$y(k+1|k) = \mathbf{C}[x^0(k+1|k) + \mathbf{B}\Delta u(k|k)] = y^0(k+1|k) + \Delta u(k|k),$$

$$\begin{aligned} y(k+2|k) &= \mathbf{C}[x^0(k+2|k) + (\mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + \mathbf{B}\Delta u(k+1|k)] = \\ &= y^0(k+2|k) + 1.5\Delta u(k|k) + \Delta u(k+1|k) \end{aligned}$$

$$\begin{aligned} y(k+3|k) &= \mathbf{C}[x^0(k+3|k) + (\mathbf{A}^2 + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + (\mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k+1|k)] = \\ &= y^0(k+3|k) + 1.78\Delta u(k|k) + 1.5\Delta u(k+1|k) \end{aligned}$$

$$\begin{aligned} y(k+4|k) &= \mathbf{C}[x^0(k+4|k) + (\mathbf{A}^3 + \mathbf{A}^2 + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k|k) + (\mathbf{A}^2 + \mathbf{A} + \mathbf{I})\mathbf{B}\Delta u(k+1|k)] = \\ &= y^0(k+4|k) + 1.974\Delta u(k|k) + 1.78\Delta u(k+1|k) \end{aligned}$$

Oznaczając

$$\begin{aligned} y_{kp}^{zad} &= y^{zad}(k+p|k), \quad y_{kp}^0 = y^0(k+p|k), \\ z_1 &= \Delta u(k|k), \quad z_2 = \Delta u(k+1|k) \end{aligned}$$

możemy sformułować funkcję celu zadania optymalizacji MPC w postaci

$$\begin{aligned} J(k) &= \sum_{p=1}^4 [y^{zad}(k+p|k) - y(k+p|k)]^2 + \lambda[(\Delta u(k|k))^2 + (\Delta u(k+1|k))^2] = \\ &= (y_{k1}^{zad} - y_{k1}^0 - z_1)^2 + (y_{k2}^{zad} - y_{k2}^0 - 1.5z_1 - z_2)^2 + (y_{k3}^{zad} - y_{k3}^0 - 1.78z_1 - 1.5z_2)^2 + \\ &\quad + (y_{k4}^{zad} - y_{k4}^0 - 1.974z_1 - 1.78z_2)^2 + \lambda(z_1)^2 + \lambda(z_2)^2 \end{aligned}$$

Oznaczając dalej

$$e_{kp}^0 = y_{kp}^{zad} - y_{kp}^0, \quad p=1,2,3,4$$

sformułowana funkcja celu przyjmuje postać

$$\begin{aligned} J(k) &= (e_{k1}^0 - z_1)^2 + (e_{k2}^0 - 1.5z_1 - z_2)^2 + (e_{k3}^0 - 1.78z_1 - 1.5z_2)^2 + (e_{k4}^0 - 1.974z_1 - 1.78z_2)^2 + \\ &\quad + \lambda(z_1)^2 + \lambda(z_2)^2 = \\ &= \alpha_{0k} + \alpha_{1k}z_1 + \alpha_{2k}z_2 + (\beta_1 + \lambda)z_1^2 + (\beta_2 + \lambda)z_2^2 + \beta_3z_1z_2 = \\ &= \alpha_{0k} + 0.5[z_1 \ z_2]^T \mathbf{H}(\lambda)[z_1 \ z_2]^T + \mathbf{f}_k^T[z_1 \ z_2]^T \end{aligned}$$

gdzie

$$\begin{aligned} \alpha_{0k} &= \sum_{p=1}^4 (e_{kp}^0)^2, \\ \alpha_{1k} &= -2(e_{k1}^0 + 1.5e_{k2}^0 + 1.78e_{k3}^0 + 1.974e_{k4}^0), \\ \alpha_{2k} &= -2(e_{k2}^0 + 1.5e_{k3}^0 + 1.78e_{k4}^0), \\ \beta_1 &= 1 + 1.5^2 + 1.78^2 + 1.974^2 = 10.315076, \\ \beta_2 &= 1 + 1.5^2 + 1.78^2 = 6.4184, \\ \beta_3 &= 2(1.5 + 1.5 \cdot 1.78 + 1.78 \cdot 1.974) = 7.68372 \\ \mathbf{H}(\lambda) &= \begin{bmatrix} 2(\beta_1 + \lambda) & \beta_3 \\ \beta_3 & 2(\beta_2 + \lambda) \end{bmatrix}, \quad \mathbf{f}_k = \begin{bmatrix} \alpha_{1k} \\ \alpha_{2k} \end{bmatrix} \end{aligned}$$

Dla zastosowania procedur standardowych programowania kwadratowego (*quadratic programming – QP*) nasze zadanie należy sformułować w standardowej postaci, np. dla procedury *quadprog* z pakietu Matlab jedną z postaci standardowych jest

$$\begin{aligned} \min\{ & 0.5z^T \mathbf{H}z + \mathbf{f}^T z \} \\ \text{przy ogr. } & \mathbf{A}_{og} z \leq \mathbf{b} \\ & z \leq \mathbf{Z}_{\max} \\ & z \geq \mathbf{Z}_{\min} \end{aligned} \tag{8.53}$$

gdzie \mathbf{z}_{\max} i \mathbf{z}_{\min} to ograniczenia górne i dolne zmiennych decyzyjnych (tzw. kostkowe). Zadanie optymalizacji dynamicznej w naszym przykładzie można zapisać

$$\min_{z_1, z_2} \left\{ 0.5z^T \mathbf{H}(\lambda)z + \mathbf{f}_k^T z \right\}$$

przy ograniczeniach :

$$\begin{aligned} 0 &\leq u(k-1) + z_1 \leq 1, \\ 0 &\leq u(k-1) + z_1 + z_2 \leq 1, \\ 0 &\leq y_{k1}^0 + z_1 \\ 0 &\leq y_{k2}^0 + 1.5z_1 + z_2 \\ 0 &\leq y_{k3}^0 + 1.78z_1 + 1.5z_2 \\ 0 &\leq y_{k4}^0 + 1.974z_1 + 1.78z_2 \end{aligned}$$

gdzie $z = [z_1, z_2]^T$. Przekształcając postać ograniczeń dostajemy to zadanie w postaci standardowej

$$\min_{z_1, z_2} \left\{ 0.5z^T \mathbf{H}(\lambda)z + \mathbf{f}_k^T z \right\}$$

przy ograniczeniach :

$$\begin{aligned} -z_1 - z_2 &\leq u(k-1) \\ z_1 + z_2 &\leq 1 - u(k-1) \\ -1.5z_1 - z_2 &\leq y_{k2}^0 \\ -1.78z_1 - 1.5z_2 &\leq y_{k3}^0 \\ -1.974z_1 - 1.78z_2 &\leq y_{k4}^0 \\ z_1 &\leq 1 - u(k-1) \\ z_1 &\geq \max\{-u(k-1), -y_{k1}^0\} \end{aligned}$$

tzn. w postaci (8.53), przy $\mathbf{H} = \mathbf{H}(\lambda)$, $\mathbf{f} = \mathbf{f}_k$, macierzy \mathbf{A}_{og} , wektorze \mathbf{b} i ograniczeniach kostkowych jak poniżej (wartości wektorów \mathbf{b} , \mathbf{z}_{\max} i \mathbf{z}_{\min} zależą od chwili k):

$$\begin{aligned} \mathbf{A}_{og} &= \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ -1.5 & -1 \\ -1.78 & -1.5 \\ -1.974 & -1.78 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} u(k-1) \\ 1 - u(k-1) \\ y_{k2}^0 \\ y_{k3}^0 \\ y_{k4}^0 \end{bmatrix}, \\ \mathbf{z}_{\max} &= \begin{bmatrix} 1 - u(k-1) \\ \text{Inf} \end{bmatrix}, \quad \mathbf{z}_{\min} = \begin{bmatrix} \max\{-u(k-1), -y_{k1}^0\} \\ -\text{Inf} \end{bmatrix} \end{aligned}$$

W wyniku rozwiązania zadania optymalizacji dynamicznej, w sformułowanej wyżej postaci standardowej QP (8.53), wyznaczana jest optymalna trajektoria sterowania

$$\hat{z}_1 = \Delta\hat{u}(k | k), \quad \hat{z}_2 = \Delta\hat{u}(k+1 | k)$$

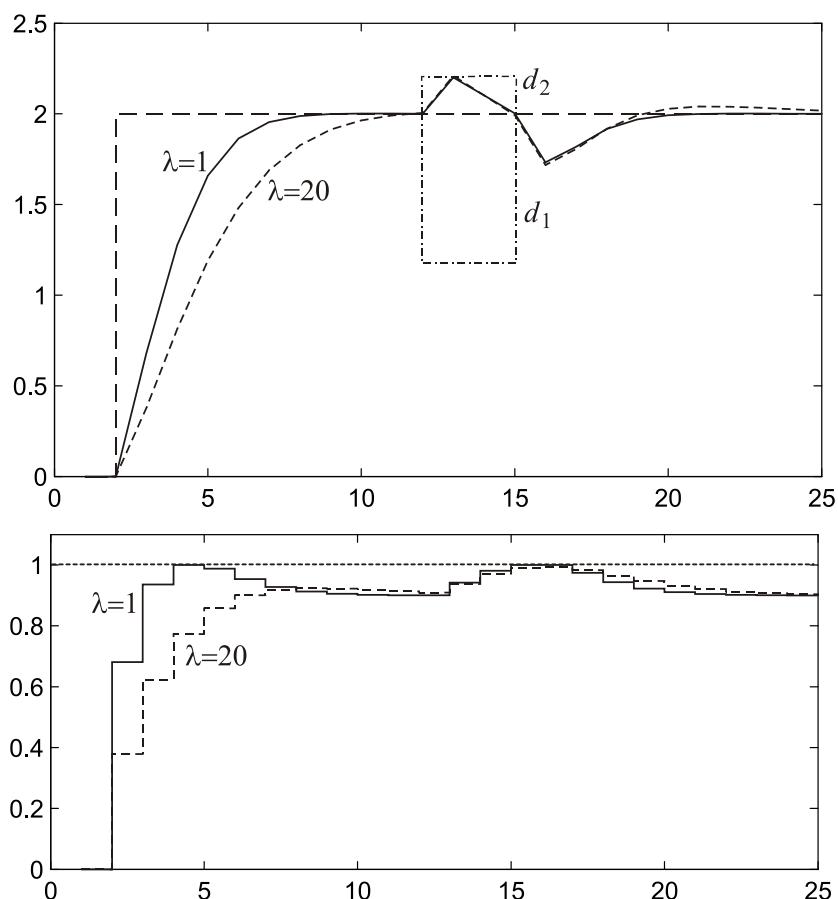
Do sterowania procesem w danej chwili k (tzn. na odcinku czasu $[kT_p, (k+1)T_p]$, gdzie T_p to okres próbkowania) wykorzystuje się jedynie pierwszy element trajektorii optymalnej odpowiadający bieżącej chwili czasu, tzn.

$$u(k) = u(k-1) + \Delta\hat{u}(k | k)$$

Po upływie okresu próbkowania, w chwili następnej ($k+1$) następuje dostarczenie kolejnych danych pomiarowych (w przykładzie powyżej pomiar stanu) i algorytm regulatora MPC jest powtarzany, itd. Tak więc, w każdej z kolejnych chwil próbkowania regulator MPC wykorzystuje dyskretnie sprzężenie zwrotne, ale nowe sterowanie wyznacza w układzie otwartym na horyzoncie predykcji (poprzez optymalizację przebiegu wartości sterowania na horyzoncie predykcji – nie poprzez bezpośrednie zastosowanie uprzednio zaprojektowanego prawa sterowania g , tzn. reguły typu $u(k) = g(x(k))$ lub $u(k) = g(y(k))$ lub podobnie, jak np. w przypadku regulacji PID).

Przykład 8.5.

Przedstawimy wyniki symulacji układu regulacji z regulatorem MPC wprowadzonym w przykładzie poprzednim, wyliczającym w każdym kroku przyrost sterowania poprzez rozwiązywanie zadania QP postaci (8.53), z parametrami wyznaczonymi w tym przykładzie dla modelu obiektu (8.52).



Rys. 8.30. Przebiegi wyjścia i sterowania w układzie regulacji predykcyjnej po skoku wartości zadanej w chwili $k=2$ i wpływie zakłóceń stanu d_1 i d_2 między chwilami $k=12$ i 14 .

Na Rys. 8.30 przedstawiono wyniki symulacji pętli regulacji z obiektem postaci

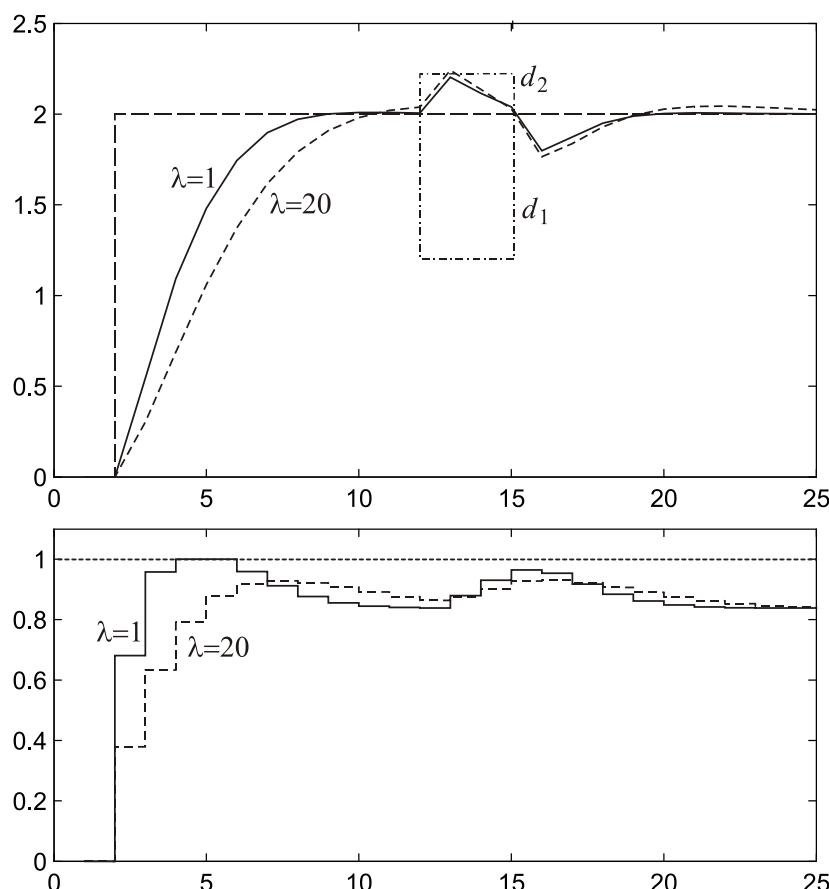
$$x(k+1) = \begin{bmatrix} 0.4 & 0.1 \\ 0.3 & 0.5 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k) + d(k), \quad y(k) = x_2(k) \quad (8.54)$$

gdzie $d(k)$ reprezentuje zakłócenia stanu obiektu – tzn. z obiektem o takiej samej strukturze i parametrami, jak w modelu użytym do projektowania regulatora. Przedstawiono trajektorie wyjścia i sterowania dla dwóch różnych wartości parametru λ (współczynnika wagowego w funkcji celu): $\lambda=1$ i $\lambda=20$; uzyskane w warunkach skoku wartości zadanej w chwili $k=2$ i prostokątnych zakłóceń o amplitudach $d_1 = -0.8$ i $d_2 = 0.2$ na odcinku czasu pomiędzy $k=12$ i $k=14$.

Natomiast na Rys. 8.31 przedstawiono wyniki symulacji z tym samym jak wyżej regulatorem MPC (zaprojektowanym dla modelu obiektu postaci (8.54)), ale w pętli regulacji z obiektem

$$x(k+1) = \begin{bmatrix} 0.3 & 0.15 \\ 0.3 & 0.6 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 0.8 \end{bmatrix} u(k) + d(k), \quad y(k) = x_2(k) \quad (8.55)$$

tz. o parametrach odchylonych w stosunku do przyjętych przy projektowaniu regulatora, co można traktować jak projektowanie dla modelu niedokładnie opisującego modelowany obiekt. Zauważmy, że zarówno w przypadku poprzednim działania jedynie zakłócenia (a modelu dokładnego), jak i obecnym regulator sprawdza wyjście do wartości zadanej – ma więc własności całkujące. W przypadku pierwszym estymata zakłócenia $v(t)$ wyliczana w regulatorze jest estymatą jedynie zakłócenia $d(t)$, natomiast w przypadku niedokładnego modelu jest ona estymatą zarówno zakłóceń $d(t)$ jak i zakłóceń powodowanych błędami modelowania obiektu. Zauważmy też, że zwiększenie wartości współczynnika wagowego λ powoduje mniejszą zmienność sterowania, wolniejsze przebiegi wyjścia regułowanego.



Rys. 8.31. Przebiegi wyjścia i sterowania w układzie regulacji predykcyjnej w warunkach jak na rysunku poprzednim, ale z odchylonymi parametrami obiektu (regulator niezmieniony).

Literatura

- [1] K. J. Åström, T. Hägglund: Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*, vol. 20, str. 645–651, 1984.
- [2] W. Findeisen: *Technika Regulacji Automatycznej*. PWN, Warszawa, 1978.
- [3] W. Findeisen, red.: *Poradnik Inżyniera: Automatyka*. WNT, 1969.
- [4] M. A. Johnson, M.H. Moradi, eds.: *PID Control*. Springer, London, 2005.
- [5] A. O'Dwyer: *Handbook of PI and PID Controller Tuning Rules*. Imperial College Press, London, 2003.
- [6] S. Skogestad: Probably the best simple PID tuning rules in the world. *AICHE Annual Meeting 2001*, Reno, NV, USA, paper 276h, 2001.
- [7] P. Tatjewski: *Sterowanie zaawansowane obiektów przemysłowych, Struktury i algorytmy*. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2002 (wydanie angielskie zmienione: *Advanced Control of Industrial Processes*. Springer, London, 2007).
- [8] J. G. Ziegler, N. B. Nichols: Optimum settings for automatic controllers. *Trans. of the ASME*, vol. 64, str. 759-768, 1942.

Zadania

Zadanie 8.1. Dany jest proces o transmitancji

$$G(s) = \frac{1}{(1+s)(1+0.2s)(1+0.05s)}$$

Należy wyznaczyć odpowiedź skokową obiektu (analitycznie lub korzystając z pakietu MATLAB), przybliżyć proces transmitancją FOPD i dobrać nastawy regulatorów PI i PID metodą Zieglera-Nicholsa. Przeprowadzić symulację uzyskanych układów regulacji dla skokowej zmiany wartości zadanej i zakłócenia na wejściu obiektu (komputerowo, wykorzystując MATLAB lub SIMULINK, realizując różniczkowanie jako rzeczywiste przy $\alpha = 10$). W razie potrzeby, skorygować nastawy tak, aby przeregulowanie nie przekraczało 10%.

Zadanie 8.2. Dany jest proces o transmitancji

$$G(s) = \frac{1}{(1+s)^4}$$

Stosując regulator proporcjonalny należy doprowadzić układ regulacji do granicy stabilności i wyznaczyć nastawy regulatorów PI i PID metodą Zieglera-Nicholsa wg cech przebiegu krytycznego (analitycznie lub korzystając z pakietu MATLAB). Przeprowadzić symulację uzyskanych układów regulacji dla skokowych zmian wartości zadanej i zakłócenia na wejściu obiektu (komputerowo, realizując różniczkowanie jako rzeczywiste z $\alpha = 10$). Dobrać również nastawy wg reguł Tyreusa-Luybena, porównać wyniki.

Zadanie 8.3. Dany jest proces o transmitancji

$$G(s) = G_1(s) \cdot G_2(s) = \frac{1}{(1+0.7\beta s)} e^{-0.3\beta s} \cdot \frac{1}{(1+0.7s)} e^{-0.3s}$$

Należy:

- Przybliżyć transmitancję obiektu postacią FOPD, a następnie dobrać nastawy regulatora PI dla obiektu $G(s)$ metodą Zieglera-Nicholsa, dla $\beta = 1$ i $\beta = 0.3$.
- Zaprojektować układ regulacji kaskadowej (przedstawiony na Rys. 8.27) dobierając najpierw nastawy regulatora PI $R_2(s)$ dla obiektu wewnętrznego $G_1(s)$, a następnie regulatora głównego PI $R_1(s)$ – metodą Zieglera Nicholsa, dla $\beta = 1$ i $\beta = 0.5$.
- Przeprowadzić symulację obu układów regulacji dla zakłócenia skokowego działającego na wyjściu obiektu $G_1(s)$, porównać jakość przebiegów przejściowych.

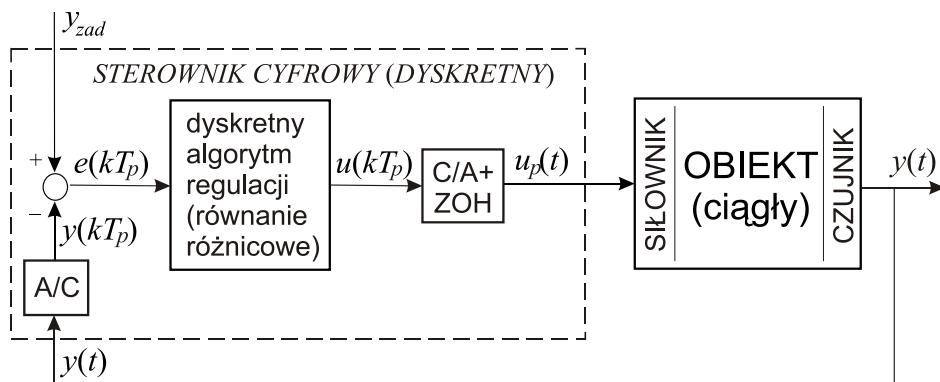
Zadanie 8.4. Dany jest proces o strukturze jak na Rys. 8.24 o transmitancjach głównej i zakłócieniowej

$$G(s) = \frac{2}{(1+s)} e^{-0.3s}, \quad G_z(s) = \frac{1}{(1+0.5s)} e^{-0.2s}$$

- Dla transmitancji głównej dobrać regulator PID metodą Zieglera-Nicholsa.
- Dodać człon *feedforward* o transmitancji $G_{ff}(s)$ zaprojektowany wg zależności (8.39), pomijając część nierealizowalną tej transmitancji.
- Przeprowadzić symulację układów regulacji *feedback* i *feedforward-feedback* przy skokowej zmianie zakłócenia, porównać wyniki.

9. Cyfrowa realizacja algorytmów regulacji

We współczesnych systemach sterowania dominuje niepodzielnie technika komputerowa, algorytmy regulacji i sterowania realizowane są programowo w sterownikach komputerowych (zob. rozdz. 2), pracujących z natury rzeczy z czasem dyskretnym. Okres czasu między kolejnymi zmianami wartości sterowania (i z reguły odczytami wielkości mierzonych w obiekcie) nazywany jest *okresem próbkowania (sampling period)*, jest to odpowiednia wielokrotność okresu zegara komputera sterującego. Okres próbkowania T_p jest dodatkowym, ważnym parametrem dobieranym przy dyskretnej implementacji algorytmu regulacji projektowanego z czasem ciągłym, czy przy projektowaniu bezpośrednio cyfrowego (dyskretnego) algorytmu regulacji. Struktura układu regulacji z komputerową (cyfrową) realizacją algorytmu regulacji pokazana jest na rys. 9.1, dla przypadku transmisji sygnału pomiarowego na wejście sterownika w postaci analogowej. Bardziej nowoczesnym rozwiązaniem jest transmisja cyfrowa sygnału pomiarowego (np. siecią polową), kiedy analogowy sygnał pomiarowy przetwarzany jest do postaci cyfrowej w samym czujniku czy przetworniku pomiarowym, bądź w koncentratorze danych.

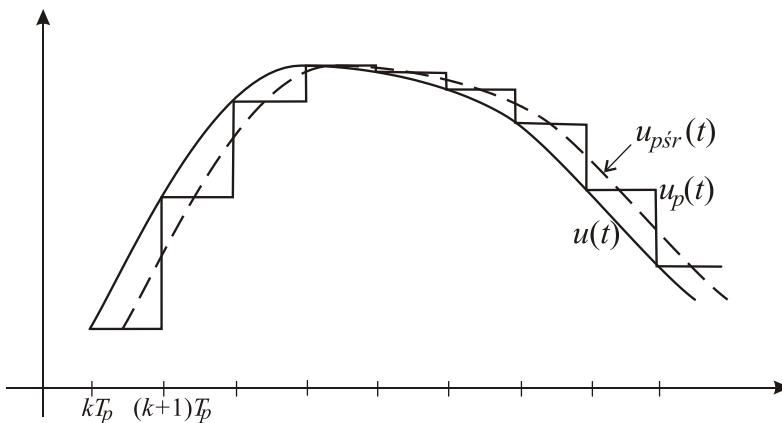


Rys. 9.1. Struktura układu regulacji z dyskretną realizacją algorytmu regulacji, z analogowym sygnałem pomiarowym na wejściu sterownika.

Blok C/A+ZOH reprezentuje na rysunku 9.1 przetwornik cyfrowo-analogowy wraz z ekstrapolatorem zerowego rzędu (ZOH - *Zero Order Hold*), tzn. na jego wejściu mamy sygnał dyskretny $u(kT_p)$, zaś na wyjściu odpowiadający mu sygnał ciągły schodkowy $u_p(t)$. Przetwarzanie dyskretnego sygnału sterującego $u(kT_p)$ na sygnał ciągły nie musi następować w sterowniku, element wykonawczy (siłownik) może posiadać własny sterownik z wejściem cyfrowym, wówczas transmisja między regulatorem a siłownikiem będzie cyfrowa.

Stosowanie w układzie regulacji sterowania dyskretnego z ekstrapolacją zerowego rzędu jest w przybliżeniu równoważne wprowadzeniu dodatkowego opóźnienia równego połowie okresu próbkowania, jak to poglądowo wyjaśniono na Rys. 9.2, gdzie przedstawiono sygnał ciągły $u(t)$, sygnał ten po dyskretyzacji z podtrzymaniem zerowego rzędu (A/C+ZOH) $u_p(t)$ oraz ten ostatni sygnał uśredniony $u_{psr}(t)$. Dzięki uśredniającym (całkującym, inercyjnym) własnościami zdecydowanej większości obiektów fizycznych obiekty te reagują na uśrednioną wartość sygnału dyskretnego – co odpowiada w przybliżeniu opóźnieniu sygnału sterującego o połowę okresu próbkowania, tzn.

$$u_{psr}(t) \approx u\left(t - \frac{T_p}{2}\right) \quad (9.1)$$



Rys. 9.2. Wpływ efektu dyskretyzacji na efektywny (uśredniony) sygnał sterowania.

Mamy dwa podstawowe podejścia do projektowania cyfrowego algorytmu regulacji:

- Metody emulacji: projektujemy algorytm regulatora w zakresie czasu ciągłego, np. w postaci opisu ciągłego w dziedzinie czasu czy w postaci transmitancji regulatora $R(s)$, a następnie przybliżamy go w postaci równania różnicowego. Sposób ten jest skuteczny i powszechnie stosowany dla małych okresów próbkowania T_p . Można go też zastosować dla nieco większych wartości T_p , zauważalnych przy porównaniu z dynamiką obiektu, dodając do opóźnienia obiektu połowę okresu próbkowania – ale w sytuacji większych okresów próbkowania zalecane jest projektowanie układu regulacji w dziedzinie czasu dyskretnego (projektowanie dyskretne, cyfrowe) [1,2].
- Metody projektowania dyskretnego: opisujemy obiekt modelem z czasem dyskretnym (przy wybranym okresie próbkowania) i dla takiego obiektu projektujemy bezpośrednio dyskretny algorytm regulacji, w postaci transmitancji dyskretnej czy bezpośrednio równania różnicowego [1,2,3,4].

9.1. Transmitancja dyskretna

Dla zdefiniowania transmitancji dyskretnej potrzebne jest przypomnienie pojęcia transformaty Z (Laurenta), która pełni analogiczną rolę przy opisie i analizie dynamicznych liniowych układów dyskretnych, jak transformata Laplace'a przy analizie liniowych układów ciągłych.

Dysponując sygnałem $f(k)$ określonym w dyskretnych chwilach $k = 0, 1, 2, \dots$ (przy czym $f(k) = 0$ dla $k < 0$), transformatę Laurenta definiuje się następująco

$$Z\{f(k)\} = F(z) = \sum_{k=0}^{\infty} f(k)z^{-k} \quad (9.2)$$

Ponieważ $f(k) = 0$ dla $k < 0$, to możemy dokonać następującego przekształcenia

$$\begin{aligned} Z\{f(k-1)\} &= \sum_{k=0}^{\infty} f(k-1)z^{-k} = \sum_{k=1}^{\infty} f(k-1)z^{-(k-1)}z^{-1} = \\ &= \sum_{k=0}^{\infty} f(k)z^{-k}z^{-1} = F(z)z^{-1} \end{aligned}$$

Skąd wynika, że z^{-1} to operator opóźnienia o jeden krok dyskretyzacji (tzn. o jeden okres próbkowania T_p). Stąd wynika prosty schemat przechodzenia od równania różnicowego liniowego układu dynamicznego do opisu w dziedzinie transformat Z, a stąd do transmitancji

dyskretnej. Rozważmy równanie różnicowe w następującej postaci (tzw. model ARX):

$$y(k) = -a_1 y(k-1) - a_2 y(k-2) - \cdots - a_M y(k-M) + b_0 u(k) + b_1 u(k-1) + \cdots + b_L u(k-L) \quad (9.3)$$

Korzystając z interpretacji operatora z^{-1} przejście do dziedziny zmiennej z (czy z^{-1}) jest natychmiastowe:

$$Y(z) = -a_1 Y(z)z^{-1} - a_2 Y(z)z^{-2} - \cdots - a_M Y(z)z^{-M} + b_0 U(z) + b_1 U(z)z^{-1} + \cdots + b_L U(z)z^{-L}$$

Wzór powyższy przepiszemy w postaci

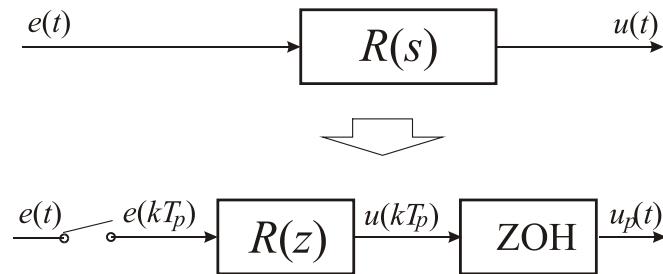
$$Y(z)(1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_M z^{-M}) = U(z)(b_0 + b_1 z^{-1} + \cdots + b_L z^{-L})$$

skąd natychmiast wynika postać *transmitancji dyskretnej*, równoważnej równaniu różnicowemu (9.3):

$$G(z) = \frac{Y(z)}{U(z)} = \frac{b_0 + b_1 z^{-1} + \cdots + b_L z^{-L}}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_M z^{-M}} \quad (9.4)$$

9.2. Metoda emulacji

Istota metody polega na zaprojektowaniu regulatora ciągłego, uzyskanego na ogół w postaci transmitancji ciągłej $R(s)$, a następnie opracowanie takiej jego dyskretnej realizacji $R(z)$, aby sygnał ciągły $u_p(t)$ na wyjściu ekstrapolatora możliwie dobrze przybliżała sygnał ciągły $u(t)$, który generowałby regulator ciągły – zasada ta przedstawiona jest na Rys. 9.3



Rys. 9.3. Zasada metody emulacji.

Metoda emulacji składa się z następujących etapów:

1. Zaprojektowanie regulatora (kompensatora) ciągłego, np. w postaci transmitancji ciągłej $R(s)$ metodą częstotliwościową.
2. Wybór okresu próbkowania i dyskretyzacja regulatora, do postaci transmitancji dyskretnej $R(z)$ czy bezpośrednio równania różnicowego.
3. Sprawdzenie projektu, poprzez eksperyment symulacyjny czy rzeczywisty, ew. skorygowanie projektu (parametrów regulatora).

Istnieje kilka metod dyskretyzacji ciągłego opisu obiektu dynamicznego danego w postaci równania różniczkowego lub transmitancji ciągłej. Omówimy poniżej najprostsze metody aproksymacji bezpośrednią, stosowaną do równania różniczkowego, czy różniczkowo-całkowego. Popularnymi, ogólnymi metodami dyskretyzacji modeli obiektów dynamicznych danych w postaci transmitancji ciągłych są metoda Tustina i metoda pokrycia zer i biegunów (*matched pole-zero metod*) [1,2].

9.2.1. Dyskretyzacja metodami bezpośredniej aproksymacji

Metoda ta polega na bezpośrednią aproksymacji pochodnych i całek za pomocą podstawowych schematów numerycznych:

- Metoda Eulera (wsteczna): aproksymujemy pochodną np. w punkcie $t = kT_p$ wstecznym ilorazem różnicowym

$$F(s)s \approx \frac{df(t)}{dt} = \frac{df(kT_p)}{dt} \cong \frac{f(kT_p) - f(kT_p - T_p)}{T_p} \quad (9.5)$$

(Nazwa metoda Eulera, gdyż formułę (9.5) można zapisać w postaci

$$f(kT_p) = f(kT_p - T_p) + T_p \frac{df(kT_p)}{dt}$$

w metodach numerycznych całkowania równań różniczkowych nazywanej wsteczną metodą Eulera – w odróżnieniu od metody Eulera, gdzie iloraz różnicowy występujący po prawej stronie wzoru (9.5) traktujemy jako progresywny, aproksymujący pochodną $df(kT_p - T_p)/dt$, tj. dokonujemy aproksymacji równania ciągłego w punkcie $t = kT_p - T_p$, na lewym końcu przedziału dyskretyzacji $[kT_p, kT_p + T_p]$).

- Metoda punktu środkowego: wzór (9.5) aproksymuje najlepiej wartość pochodnej w środku przedziału $[kT_p, kT_p + T_p]$, a nie na którymkolwiek z jego krańców (jak w metodach Eulera powyżej) – stąd dokonujemy aproksymacji w punkcie $t = kT_p - 1/2T_p$ (środek przedziału dyskretyzacji), tzn.

$$\frac{df(kT_p - \frac{T_p}{2})}{dt} \cong \frac{f(kT_p) - f(kT_p - T_p)}{T_p} \quad (9.6)$$

i konsekwentnie wartości zmiennych też aproksymujemy w punkcie środkowym, tzn. średnimi arytmetycznymi: $f(kT_p - T_p/2) \cong 1/2(f(kT_p) + f(kT_p - T_p))$, zob. Przykład 9.1.

Metody bezpośrednie są wygodne przede wszystkim dla równań różniczkowych pierwszego rzędu, bądź dla równań drugiego rzędu w postaci różniczkowo-całkowej (Przykład 9.2 poniżej) – pochodne (pierwszego rzędu) aproksymujemy jak wyżej, a całkę odpowiednio dobraną do punktu dyskretyzacji metodą trapezów, np. dla aproksymacji w punkcie $t = kT_p$ (metoda Eulera wsteczna):

$$\frac{F(s)}{s} \approx \int_0^t f(\tau) d\tau = \int_0^{kT_p} f(\tau) d\tau \cong \sum_{i=0}^{i=k} \frac{f(iT_p) + f(iT_p - T_p)}{2} T_p \quad (9.7)$$

Jeśli równanie różniczkowe jest drugiego rzędu, to należałoby aproksymować odpowiednim schematem różnicowym również pochodne drugiego rzędu, wygodniej jest wówczas na ogólnie zastosować inną popularną metodę ogólną, np. metodę Tustina.

Przykład 9.1.

Implementacja dyskretna korektora szeregowego (układ dynamiczny pierwszego rzędu)

$$R(s) = \frac{U(s)}{E(s)} = \frac{1 + 0.02s}{1 + 0.002s} \quad (9.8)$$

Przepisując powyższy wzór w postaci

$$U(s) + 0.002U(s)s = E(s) + 0.02E(s)s$$

a następnie przechodząc do dziedziny czasu dostajemy (przyjmując zerowe warunki początkowe, jak typowo przy podejściu transmitancyjnym)

$$u(t) + 0.002 \frac{du(t)}{dt} = e(t) + 0.02 \frac{de(t)}{dt}$$

Stosując teraz aproksymację dyskretną w punkcie $t = kT_p$, (metoda Eulera wsteczna, wzór (9.59.5)), dostajemy:

$$u(kT_p) + 0.002 \frac{u(kT_p) - u(kT_p - T_p)}{T_p} = e(kT_p) + 0.02 \frac{e(kT_p) - e(kT_p - T_p)}{T_p}$$

Porządkując odpowiednio powyższy wzór dostajemy dyskretne równanie różnicowe w dogodnej do implementacji postaci

$$u(k) = \frac{0.002}{T_p + 0.002} u(k-1) + \frac{T_p + 0.02}{T_p + 0.002} e(k) - \frac{0.02}{T_p + 0.002} e(k-1) \quad (9.9)$$

Stosując metodę punktu środkowego dostajemy:

$$\begin{aligned} \frac{1}{2}(u(kT_p) + u(kT_p - T_p)) + 0.002 \frac{u(kT_p) - u(kT_p - T_p)}{T_p} &= \\ &= \frac{1}{2}(e(kT_p) + e(kT_p - T_p)) + 0.02 \frac{e(kT_p) - e(kT_p - T_p)}{T_p} \end{aligned}$$

Skąd po uporządkowaniu dostajemy dyskretne równanie różnicowe

$$u(k) = -\frac{T_p - 0.004}{T_p + 0.004} u(k-1) + \frac{T_p + 0.04}{T_p + 0.004} e(k) + \frac{T_p - 0.04}{T_p + 0.004} e(k-1) \quad (9.10)$$

Postać wzorów uzyskanych dla obu metod pokazuje jawnie nieodłączną dla modeli dyskretnych zależność wartości współczynników równania różnicowego od długości okresu próbkowania T_p .

Przykład 9.2.

Dyskretna realizacja algorytmu regulatora PID (postać równoległa).

Algorytm regulatora ciągłego PID zapisany w dyskretnej chwili czasu $t = kT_p$ ma postać

$$u(kT_p) = k_p \left[e(kT_p) + \frac{1}{T_i} \int_0^{kT_p} e(\tau) d\tau + T_d \frac{de(kT_p)}{dt} \right] \quad (9.11)$$

Ze względu na całkę uchybu występującą w powyższym wzorze dogodnie jest wyznaczyć przyrostową wersję tego algorytmu. Uzyskujemy ją odejmując od wzoru (9.11) jego postać zapisaną dla chwili $t = kT_p - T_p$, mamy wtedy

$$u(kT_p) - u(kT_p - T_p) = k_p \left[e(kT_p) - e(kT_p - T_p) + \frac{1}{T_i} \int_{kT_p - T_p}^{kT_p} e(\tau) d\tau + T_d \frac{de(kT_p)}{dt} - T_d \frac{de(kT_p - T_p)}{dt} \right]$$

Stosując teraz aproksymacje (9.5) i (9.7) dostajemy

$$u(kT_p) = u(kT_p - T_p) + k_p \left[e(kT_p) - e(kT_p - T_p) + \frac{T_p}{2T_i} (e(kT_p) + e(kT_p - T_p)) + \right. \\ \left. + \frac{T_d}{T_p} (e(kT_p) - 2e(kT_p - T_p) + e(kT_p - 2T_p)) \right]$$

Wzór ten zapisany w dziedzinie czasu dyskretnego ma prostszą postać

$$u(k) = u(k-1) + k_p \left[e(k) - e(k-1) + \frac{T_p}{2T_i} (e(k) + e(k-1)) + \right. \\ \left. + \frac{T_d}{T_p} (e(k) - 2e(k-1) + e(k-2)) \right] \quad (9.12)$$

9.2.2. Dyskretyzacja metodą Tustina

Metoda polega na wykorzystaniu relacji między zmiennymi s i z przekształceń Laplace'a i Z, konstruowanej przez aproksymację trapezową operacji całkowania na odcinku czasu odpowiadającym okreowi próbkowania. Dla regulatora będącego czystym całkowaniem, o transmitancji Laplace'a: $R(s) = U(s)/E(s) = 1/s$, mamy bowiem wówczas

$$u(kT_p) = u(kT_p - T_p) + \int_{kT_p - T_p}^{kT_p} e(\tau) d\tau \cong u(kT_p - T_p) + \frac{e(kT_p - T_p) + e(kT_p)}{2} T_p$$

czyli w dziedzinie czasu dyskretnego

$$u(k) = u(k-1) + \frac{e(k-1) + e(k)}{2} T_p$$

Przechodząc do dziedziny zmiennej z mamy

$$U(z) = U(z)z^{-1} + \frac{T_p}{2} (E(z)z^{-1} + E(z))$$

skąd

$$\frac{U(z)}{E(z)} = \frac{T_2}{2} \frac{1+z^{-1}}{1-z^{-1}} = \frac{1}{2} \frac{1-z^{-1}}{1+z^{-1}}, \quad \text{przy trans mitancji Laplace'a: } \frac{U(s)}{E(s)} = \frac{1}{s}$$

Stąd dostajemy zależność na odpowiednik zmiennej s przekształcenia Laplace'a w dziedzinie przekształcenia Z:

$$s \approx \frac{2}{T_p} \frac{1-z^{-1}}{1+z^{-1}} = \frac{2}{T_p} \frac{z-1}{z+1} \quad (9.13)$$

Metoda Tustina polega na przejściu od transmitancji ciągłej do transmitancji dyskretnej poprzez zastosowanie podstawienia (9.13), które nazywane jest też transformacją biliniową.

Przykład 9.3

Dokonamy implementacji dyskretnej korektora szeregowego z przykładu 9.1,

$$R(s) = \frac{U(s)}{E(s)} = \frac{1+0.02s}{1+0.002s} \quad (9.14)$$

metodą Tustina. Stosując do tej transmitancji transformację biliniową uzyskujemy

$$U(z) + U(z) \frac{0.004}{T_p} \left(\frac{1-z^{-1}}{1+z^{-1}} \right) = E(z) + E(z) \frac{0.04}{T_p} \left(\frac{1-z^{-1}}{1+z^{-1}} \right)$$

Dalej mamy

$$\begin{aligned} U(z)(1+z^{-1}) + U(z) \frac{0.004}{T_p} (1-z^{-1}) &= E(z)(1+z^{-1}) + E(z) \frac{0.04}{T_p} (1-z^{-1}) \\ U(z) \frac{T_p + 0.004}{T_p} + U(z) \frac{T_p - 0.004}{T_p} z^{-1} &= E(z) \frac{T_p + 0.04}{T_p} + E(z) \frac{T_p - 0.04}{T_p} z^{-1} \end{aligned}$$

oraz finalnie w dziedzinie czasu dyskretnego:

$$u(k) = -\frac{T_p - 0.004}{T_p + 0.004} u(k-1) + \frac{T_p + 0.04}{T_p + 0.004} e(k) + \frac{T_p - 0.04}{T_p + 0.004} e(k-1) \quad (9.15)$$

Zauważmy, że uzyskaliśmy ten sam wzór, jak uprzednio w Przykładzie 9.1 w metodzie punktu środkowego.

9.3. Dobór okresu próbkowania

Dobór okresu próbkowania jest istotnym elementem projektowania dyskretnego algorytmu sterowania metodą emulacji. Twierdzenie Shannona mówi, że sygnał ciągły może być ponownie wiernie odtworzony z sygnału dyskretnego, jeśli był próbkiowany z częstotliwością co najmniej dwa razy większą od granicznej częstotliwości swego widma. Zakładając, że prawie całe widmo sygnału ciągłego mieści się w paśmie $10\omega_{bw}$, gdzie ω_{bw} to pasmo przenoszenia układu (przyjmowane do pulsacji spadku o 3dB charakterystyki amplitudowej układu zamkniętego), dostajemy

$$\frac{\omega_p}{2} \geq 10\omega_{bw} \quad \Rightarrow \quad \omega_p \geq 20\omega_{bw} \quad (9.16)$$

gdzie ω_p to pulsacja próbkowania, tzn. $\omega_p = 2\pi/T_p$. Przyjmuje się, że przy spełnieniu warunku (9.16) układ regulacji powinien się zachowywać stosunkowo poprawnie, niewiele odbiegając od zachowania układu ciągłego. Uważa się, że spełnienie nieco ostrzejszego warunku

$$\omega_p \geq 30\omega_{bw} \quad (9.17)$$

daje już bardzo dobrą aproksymację zachowania się układu ciągłego przez układ z dyskretną realizacją algorytmu sterującego. Nie powinno się nadmiernie zwiększać częstotliwości próbkowania, znacznie ponad wymagane wartości, gdyż prowadzi to do zwiększenia ilości obliczeń i mogą wówczas wystąpić negatywne efekty związane z błędami numerycznymi.

Przyjmowanie częstotliwości próbkowania poniżej zakresu danego warunkiem (9.16) jest możliwe, ale prowadzi do wzrostu oscylacyjności przebiegów przejściowych w układzie regulacji, a w skrajnym przypadku nawet do niestabilności (silniejszy negatywny wpływ opóźnienia wprowadzanego przez operację dyskretyzacji). Jeśli chcemy rozważać układ regulacji z częstotliwością próbkowania zbliżoną do wartości $10\omega_{bw}$ lub mniejszą, to metoda emulacji nie ma wówczas zastosowania, należy takie układy sterowania projektować w dziedzinie czasu dyskretnego.

Przykład 9.4.

Rozważmy obiekt o transmitancji

$$G(s) = \frac{560}{s(0.1s + 1)} \quad (9.18)$$

z korektorem szeregowym o transmitancji (9.14). Postać dyskretną tego korektora wyprowadziliśmy metodą aproksymacji bezpośredniej w przykładzie 9.1, jest ona dana zależnością (9.9). Jak łatwo sprawdzić, pulsacja graniczna pasma przenoszenia układu zamkniętego z transmitancją toru otwartego $G(s)R(s)$ to ok. 180 rd/sek. Przyjmując $\omega_p \approx 20\omega_{bw}$ dostajemy stąd $T_p = 2\pi/20\omega_{bw} \approx 0.0017$ (1.7 msec).

Równanie różnicowe korektora (9.9) uzyskane metodą bezpośrednią Eulera (wsteczną) przyjmuje w dziedzinie czasu dyskretnego postać, dla $T_p = 0.0017$:

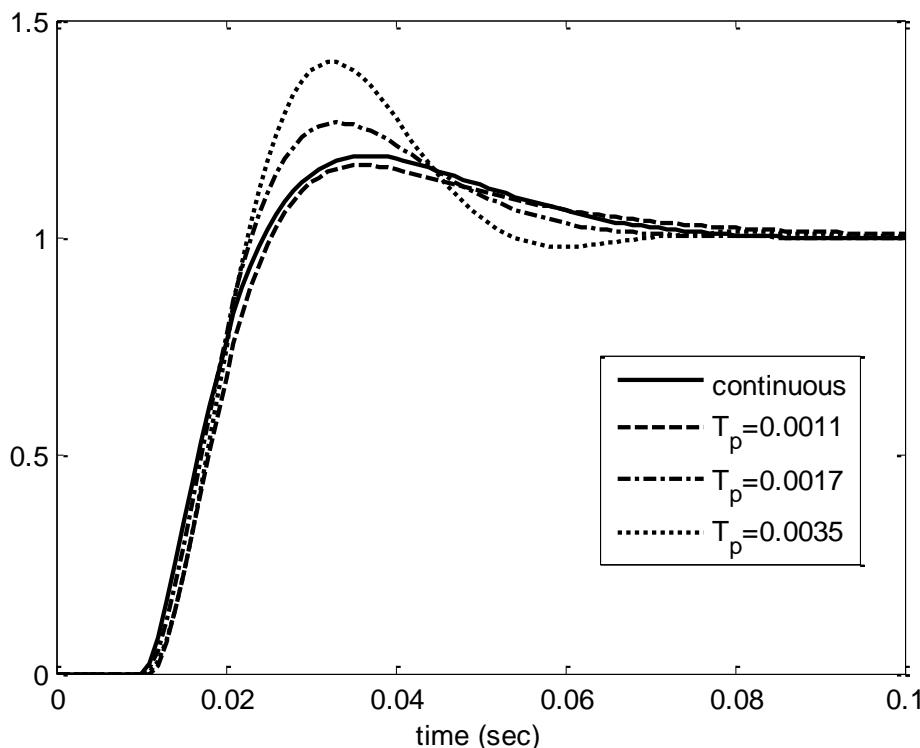
$$\begin{aligned} u(k) &= \frac{20}{37}u(k-1) + \frac{217}{37}e(k) - \frac{200}{37}e(k-1) \approx \\ &\approx 0.54u(k-1) + 5.865e(k) - 5.405e(k-1) \end{aligned}$$

Natomiast równanie to dla $T_p = 0.0011$ (1.1 msec, co odpowiada $\omega_p \approx 30\omega_{bw}$) ma postać

$$\begin{aligned} u(k) &= \frac{20}{31}u(k-1) + \frac{211}{31}e(k) - \frac{200}{31}e(k-1) \approx \\ &\approx 0.645u(k-1) + 6.806e(k) - 6.451e(k-1) \end{aligned}$$

a dla $T_p = 0.0035$ (3.5 msec, co odpowiada $\omega_p \approx 10\omega_{bw}$) dostajemy

$$\begin{aligned} u(k) &= \frac{20}{55}u(k-1) + \frac{235}{55}e(k) - \frac{200}{55}e(k-1) \approx \\ &\approx 0.364u(k-1) + 4.273e(k) - 3.636e(k-1) \end{aligned}$$

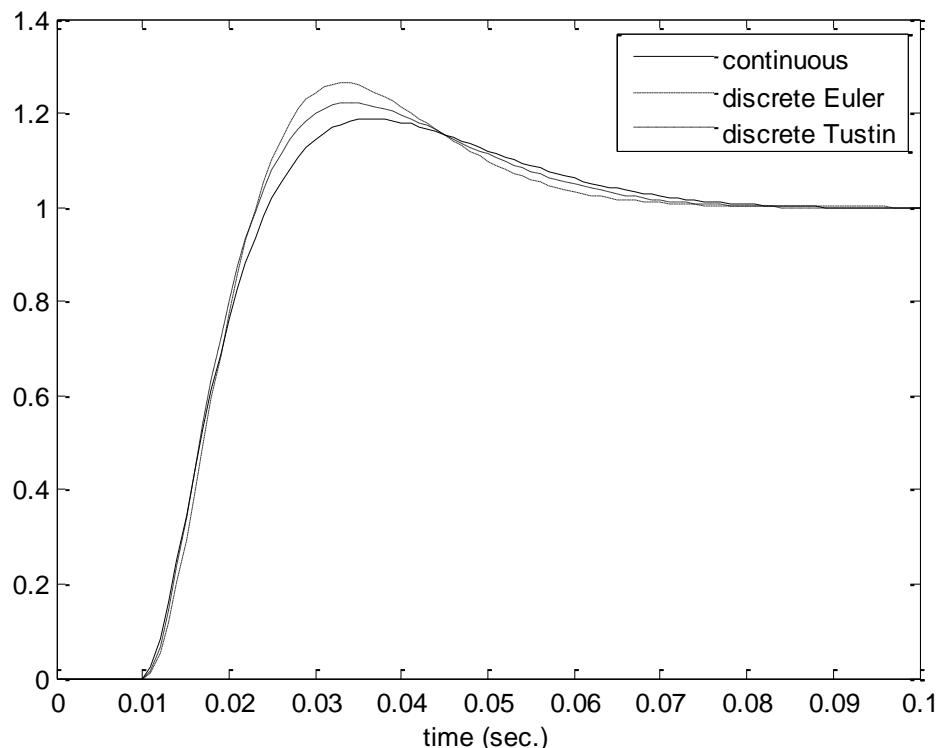


Rys. 9.4. Porównanie jakości regulacji przy różnych okresach próbkowania, dyskretyzacja metodą Eulera (wsteczną).

Na Rys. 9.4 przedstawiono porównanie wyników symulacji układu regulacji dla czterech przypadków: z obiektem ciągłym i regulatorem ciągłym oraz z tym samym obiektem ciągłym i korektorem dyskretnym uzyskanym metodą emulacji, dla trzech różnych okresów impulsowania omówionych powyżej. Wyniki te potwierdzają teoretyczne zasady postępowania: dla okresu próbkowania $T_p = 0.0035$ jakość regulacji jest nieakceptowana, za duże jest przeregulowanie (co odpowiada znacznie zmniejszonemu zapasowi fazy). Dla $T_p = 0.0017$ nie jest bardzo dobrze, ale układ jest do zaakceptowania, natomiast dla $T_p = 0.0011$ mamy bardzo zbliżone przebiegi w układach z korektorem ciągłym i dyskretnym – mimo zastosowania najprostszej metody dyskretyzacji bezpośredniej (metody Eulera). Jeśli zastosować lepszą metodę dyskretyzacji, np. popularna metodę Tustina (równoważną w tym przypadku metodzie punktu środkowego), to wyniki ulegają poprawie. Dla porównania, wstawienie $T_p = 0.0017$ do równania różnicowego korektora uzyskanego poprzez zastosowanie metody punktu środkowego (też metody Tustina) prowadzi do

$$u(k) = 0.404u(k-1) + 7.316e(k) - 6.719e(k-1)$$

i nieco lepszej jakości regulacji, co pokazuje porównanie z korektorem uzyskanym uprzednio metodą Eulera przedstawione na Rys. 9.5.



Rys. 9.5. Porównanie jakości regulacji dla różnych metod dyskretyzacji, $T_p = 0.0017$.

Literatura

- [1] K.J. Astrom, B. Wittenmark: *Computer Controlled Systems*. Prentice Hall, Upper Saddle River, New Jersey, 1997.
- [2] G.F. Franklin, J.D. Powell, A. Emami-Naeini: *Feedback Control of Dynamic Systems*. Addison Wesley, Reading, Massachusetts (several editions).

- [3] R. Gessing: *Podstawy Automatyki*. Wydawnictwo Politechniki Śląskiej, Gliwice, 2001.
- [4] G.C. Goodwin, S.F.Graebe, M.E.Salgado: *Control System Design*. Prentice Hall, Upper Saddle River, New Jersey, 2001.

Zadania

Zadanie 9.1. Zaprojektowano regulator PD rzeczywisty o transmitancji

$$R(s) = 10 \left(1 + \frac{0.5s}{1 + 0.05s} \right)$$

Wyznacz postacie cyfrowe tego regulatora (równania różnicowe) dla komputerowej realizacji regulatora z okresem próbkowania $T_p=0.01$ sek., stosując: metodę Eulera, metodę punktu środkowego i metodę Tustina.

Zadanie 9.2. Dla regulatora PID w postaci szeregowej

$$R(s) = k_p (1 + T_d s) \left(1 + \frac{1}{T_i s} \right)$$

wyznacz realizację dyskretną (równanie różnicowe) w postaci przyrostowej stosując metodę bezpośrednią Eulera (zastosowaną do równania różniczkowo-całkowego, z przybliżeniem całki metodą trapezów).

Zadanie 9.3. Dla obiektu rozważanego w przykładzie 7.1 (po korekcji wzmacnienia) dobierz okres próbkowania dla cyfrowej realizacji regulatora P (korektora wzmacnienia), przyjmując $\omega_p \approx 20\omega_{bw}$.

Dodatek: Przekształcenie Laplace'a

$F(s)$	$f(t), \quad t \geq 0$
1	$\delta(t)$
$\frac{1}{s}$	$1(t)$
$\frac{1}{s^2}$	t
$\frac{m!}{s^{m+1}}$	t^m
$\frac{1}{(s+a)}$	e^{-at}
$\frac{1}{(s+a)^2}$	te^{-at}
$\frac{1}{(s+a)^m}$	$\frac{1}{(m-1)!}t^{m-1}e^{-at}$
$\frac{b-a}{(s+a)(s+b)}$	$e^{-at} - e^{-bt}$
$\frac{a}{s^2+a^2}$	$\sin at$
$\frac{s}{s^2+a^2}$	$\cos at$
$\frac{b}{(s+a)^2+b^2}$	$e^{-at} \sin bt$
$\frac{s+a}{(s+a)^2+b^2}$	$e^{-at} \cos bt$

Niektóre własności przekształcenia Laplace'a:

$$\mathcal{L}[\alpha f_1(t) + \beta f_2(t)] = \alpha F_1(s) + \beta F_2(s)$$

$$\mathcal{L}[f(t - T_o)] = e^{-sT_o} F(s)$$

$$\mathcal{L}\left[\frac{d}{dt}f(t)\right] = sF(s) - f(0^+)$$

$$\mathcal{L}\left[\frac{d^m}{dt^m}f(t)\right] = s^m F(s) - \sum_{j=1}^m s^{m-j} \frac{d^{j-1}}{dt^{j-1}} f(0^+)$$

$$\mathcal{L}\left[\int_0^t f(\tau) d\tau\right] = \frac{F(s)}{s}$$

$$\mathcal{L}\left[\int_0^t f_1(t - \tau) f_2(\tau) d\tau\right] = F_1(s) \cdot F_2(s)$$

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} sF(s)$$