



CONTROLO DE SISTEMAS DINÂMICOS

Tempo contínuo e tempo discreto



Rui Neves-Silva



Índice

Índice	3
1. Introdução	8
2. Sistemas físicos e a modelação.....	12
2.1. Processo de modelação.....	12
2.2. Linearização de sistemas físicos.....	19
3. Resposta no tempo.....	26
3.1. Resposta livre	26
3.2. Degrau unitário.....	27
3.3. Resposta forçada.....	28
3.4. Impulso unitário.....	30
3.5. Resposta impulsiva	31
3.6. Transformação de Laplace.....	33
3.7. Produto de convolução.....	37
3.8. Solução da equação diferencial usando a Transformação de Laplace.....	38
4. Função de transferência e análise dinâmica.....	40
4.1. Função de transferência de um SLIT	40
4.2. Conversão direta para função de transferência.....	42
4.3. Álgebra de blocos	42
4.4. Polos e zeros da função de transferência.....	47
4.5. Estabilidade assintótica.....	49
4.6. Teorema do valor final.....	52
4.7. Teorema do valor inicial	53
4.8. Resposta no tempo de sistemas de ordem reduzida.....	53
5. Controlo por retroação.....	61
5.1. O sistema de controlo	61
5.2. Especificação de um sistema de controlo.....	63
5.3. Especificação no plano s	64

5.4.	Retroação proporcional da saída.....	65
5.5.	Diagrama do lugar de raízes	71
5.6.	O problema do seguimento da referência.....	78
6.	Resposta em frequência.....	82
6.1.	Sinal periódico sinusoidal	82
6.2.	Resposta no tempo a um sinal sinusoidal	83
6.3.	Critério de Nyquist.....	87
6.4.	Diagrama de Nyquist e a resposta em frequência	94
6.5.	Diagrama de Bode	95
6.6.	Construção aproximada do diagrama de Bode – método das assíntotas	96
6.7.	Diagrama de Bode real e assintótico	106
6.8.	Contribuição de polos instáveis e zeros de fase não mínima	109
6.9.	Construção do diagrama de Nyquist a partir do diagrama de Bode.....	112
7.	Especificação e projeto na frequência	116
7.1.	Robustez de estabilidade face a incerteza no ganho e na fase	116
7.2.	Efeito da introdução de atraso puro na estabilidade	119
7.3.	Crítica às margens de ganho e fase.....	121
7.4.	Largura de banda e rapidez de resposta	121
7.5.	Especificação no domínio da frequência	124
7.6.	Compensação de atraso	126
7.7.	Compensação de avanço.....	128
8.	Controlo PID.....	132
8.1.	Ajuste manual de controladores PID	135
8.2.	O controlador PID e o lugar das raízes.....	137
8.3.	Regras de Ziegler-Nichols	139
8.4.	Análise e modificação do método de Ziegler-Nichols	143
9.	Sistemas em tempo discreto.....	147
9.1.	Sistema de controlo por computador	148
9.2.	Degrau e impulso unitários.....	150

9.3.	Sistema discreto.....	150
9.4.	Equações às diferenças	153
9.5.	Solução das equações às diferenças	153
9.6.	Resposta impulsiva	156
9.7.	Cálculo da convolução	157
9.8.	Transformação em z	159
9.9.	Função de transferência.....	162
9.10.	Código do controlador	162
9.11.	Seleção do período de amostragem	164
9.12.	Teorema da amostragem (Nyquist-Shannon)	164
9.13.	Duas abordagens alternativas no projeto do controlador em tempo discreto	167
10.	Análise dinâmica do sistema discreto	169
10.1.	Álgebra de blocos.....	169
10.2.	Polos e zeros da função de transferência discreta.....	169
10.3.	Estabilidade assintótica	170
10.4.	Critério de estabilidade de Jury.....	172
10.5.	Equivalência entre respostas impulsivas contínuas e discretas	176
10.6.	Respostas de sistemas de primeiro grau	177
10.7.	Respostas de sistemas de segundo grau	180
11.	Métodos de discretização	182
11.1.	Discretização de controladores contínuos.....	182
11.2.	Métodos de discretização por aproximação	184
11.3.	Discretização de instalações	189
11.4.	Método de discretização ZOH.....	191
12.	Controlo por retroação em tempo discreto.....	193
12.1.	Especificação do sistema de controlo	193
12.2.	Diagrama de localização de raízes	196
13.	Resposta em frequência em tempo discreto.....	201

13.1.	Sinal discreto sinusoidal	201
13.2.	Resposta no tempo discreto a um sinal sinusoidal.....	202
13.3.	Critério de Nyquist	206
13.4.	Diagrama de Nyquist e a resposta em frequência.....	209
13.5.	Diagrama de Bode.....	209
13.6.	Construção aproximada do diagrama de Bode – método das assíntotas.....	210
13.7.	Projeto do controlador através da resposta em frequência	213
14.	Sistemas discretos em espaço de estados	214
14.1.	Vetor de estado e modelo de estado	215
14.2.	Matriz de transição de estado	218
14.3.	Solução geral da equação de estado.....	218
14.4.	Relação entre modelo de estado e função de transferência	219
14.5.	Cálculo da matriz de transição.....	220
14.6.	Conversão da função de transferência em modelo de estado.....	221
14.7.	Valores próprios e estabilidade.....	223
14.8.	Vetores próprios e desacoplamento dos modos dinâmicos	225
15.	Controlo em espaço de estados	230
15.1.	Controlabilidade completa	230
15.2.	Controlo por retroação linear de variáveis de estado	232
16.	Controlo com observadores de estado	236
16.1.	Observador assintótico	236
16.2.	Observabilidade completa	239
16.3.	Projeto do observador assintótico.....	241
16.4.	Teorema da separação dos modos.....	243
16.5.	Função de transferência do controlador.....	244

1. Introdução

As disciplinas de Controlo de Sistemas Dinâmicos atravessam horizontalmente as diversas engenharias tradicionais - como a eletrotécnica, mecânica ou química - e ainda outras áreas do conhecimento como, por exemplo, a biologia ou a economia.

A **Teoria do Controlo** tem dois objetivos de base:

- Permitir compreender a natureza dinâmica dos sistemas;
- Fornecer métodos para modificar o desempenho dinâmico dos sistemas.

Mas antes de mergulharmos nestas abordagens, vamos definir alguns dos conceitos importantes para a sua compreensão.

Com generalidade, define-se **sistema contínuo** como uma estrutura organizada de transferência de um sinal $u(t)$ noutra sinal $y(t)$, ambos existentes no tempo contínuo t . Se $T[.]$ representar simbolicamente o operador de transferência, então escreve-se

$$y(t) = T[u(t)]$$

A $u(t)$ dá-se o nome de entrada, excitação ou estímulo; e a $y(t)$ de saída, resposta ou reacção. Para início, neste curso, pensemos apenas em sistemas univariáveis, ou seja, com apenas uma entrada e uma saída. Estes sistemas são habitualmente designados por SISO (de *Single-Input Single-Output*). Mais tarde, veremos que todos estes conceitos são generalizáveis para o caso de múltiplas entradas e saídas (MIMO de *Multiple-Input Multiple-Output*).

O sistema pode ainda ser representado de uma forma esquemática através de um bloco que opera a entrada $u(t)$ na saída $y(t)$. Os percursos dos sinais são representados por setas como se observa na Figura 1.1.

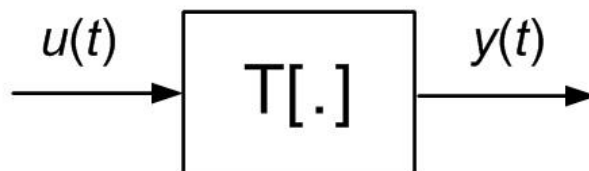


Figura 1.1 – Representação esquemática de um sistema.

O conceito mais importante associado à teoria dos sistemas dinâmicos é exatamente o conceito de **dinâmica**. Um sistema é dinâmico se a sua saída atual depende também das entradas passadas, i.e. se tem memória. Em oposição, se depender apenas da entrada atual, o sistema diz-se **estático**. A dinâmica de um sistema é a sua riqueza, ou melhor, a sua personalidade. Mais do que a forma como o sistema responde aos estímulos exteriores (a

partir da sua entrada), a análise dinâmica estuda a forma como o sistema reage a si próprio, ou à sua história passada. Por exemplo, o conceito de estabilidade assintótica, uma das características dinâmicas mais importantes a desenvolver adiante, depende essencialmente de características intrínsecas do sistema.

Na representação de sistemas físicos em tempo real é importante definir também o conceito de **causalidade**. Um sistema diz-se causal se a reação na saída não antecipa no tempo a excitação realizada à entrada. Um sistema causal (no sentido lato) reage, quanto muito, no mesmo instante em que sofre o estímulo. No entanto, refira-se que existem muitos sistemas com interesse para a engenharia que são não causais no tempo. Por exemplo, se os sinais estiverem pré-registados num qualquer tipo de suporte (e.g. fita magnética) é possível “olhar” o futuro realizando uma leitura mais adiante no conjunto de dados.

Para o desenvolvimento da disciplina, iremos nos focar na classe de sistemas dinâmicos e causais que, à primeira vista parecem bastante limitados na sua capacidade de representar a generalidade dos sistemas. No entanto, estes serão de grande utilidade no alcançar dos objetivos introduzidos no início deste capítulo. Falamos do conceito de **Sistema Linear e Invariante no Tempo** (ou SLIT).

A **invariância no tempo** está associada a uma certa constância das características dinâmicas do sistema ao longo do tempo. Dado um qualquer sinal de entrada $u(t)$ e a respectiva reacção $y(t)$, se a deslocação no tempo do sinal de excitação de um qualquer intervalo de tempo τ provocar a mesma resposta deslocada do mesmo intervalo, i.e.

$$y(t - \tau) = T[u(t - \tau)]$$

o sistema diz-se invariante no tempo.

Por outro lado, um sistema diz-se **linear** se a sobreposição escalada de quaisquer sinais à entrada resultarem sempre numa saída igual à sobreposição escalada das respostas de cada um dos sinais em separado, ou seja, no caso de apenas dois sinais,

$$y(t) = T[a \cdot u_1(t) + b \cdot u_2(t)] = a \cdot T[u_1(t)] + b \cdot T[u_2(t)]$$

A linearidade e a invariância no tempo irão permitir que, a partir do conhecimento da reacção do sistema a um sinal de entrada simples (e.g. um impulso), se possa extrapolar as conclusões sobre o comportamento dinâmico do sistema, para outros sinais mais elaborados em qualquer instante do tempo.

Na realidade, os sistemas físicos não são invariantes e poucos serão absolutamente lineares. No entanto, o conjunto de ferramentas que serão desenvolvidas a partir destas hipóteses permitem projetar sistemas de engenharia que em agregação com os sistemas físicos originais, conseguem apresentar previsibilidade no seu comportamento dinâmico.

Um dos objetivos deste texto é introduzir o conceito de **controlo automático**. O controlo prende-se com a modificação do desempenho dinâmico de sistemas físicos que iremos designar por instalações.

A **instalação** é um sistema físico cliente. O engenheiro responsável pelo projeto de um sistema de controlo não tem, normalmente, acesso a modificar a instalação. Esta instalação pode ser, por exemplo, uma máquina elétrica, ou um processo químico, ou uma caldeira a gás. O **controlador** é o sistema a projetar, exterior à instalação, que agregado com aquela resultará num **sistema de controlo** com o comportamento dinâmico desejado.

A Figura 1.2 representa um sistema de controlo por retroação para regulação.

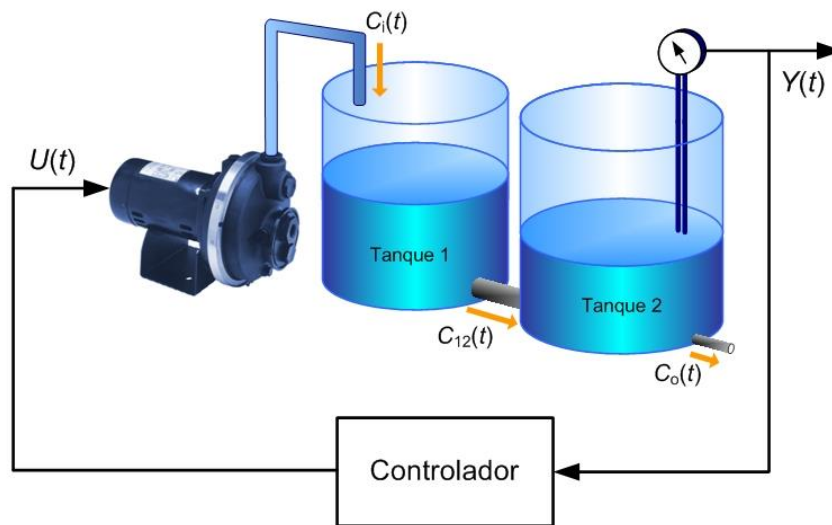


Figura 1.2 – Sistema de controlo por retroação.

A ação de controlo por retroação será desenvolvida mais à frente neste texto, mas realiza uma função básica que é natural para o ser humano. No caso da figura, se o nível do tanque 2 está mais alto que o valor desejado, reduz-se o caudal na bomba através do seu comando; se o nível se encontra abaixo do valor desejado aumenta-se a entrada de água. Fácil de apresentar, mas existe uma infinidade de comportamentos dinâmicos diferentes que uma instalação pode apresentar e a forma de compreender esses comportamentos e o subsequente projeto da unidade controladora serão o objeto deste texto.

2. Sistemas físicos e a modelação

O controlo, motivo deste texto, exerce-se sobre um objeto que é a instalação a controlar. Um passo importante para a realização do sistema de controlo é a realização de um modelo da instalação e sua compreensão. Embora as ferramentas de projeto de controladores que iremos ver à frente tenham um alcance mais largo, podendo ser aplicadas em áreas tão diversas como a biologia ou a economia, concentrarmo-nos, para já, na realização de modelos de sistemas físicos com interesse para a engenharia.

2.1. Processo de modelação

Antes de se descrever o processo de modelação é importante sublinhar a diferença entre o **sistema físico** e um seu eventual **modelo**.

O sistema físico a controlar, que designaremos por instalação, é real e, em geral, formado por componentes infinitamente complexas.

Um modelo matemático é uma descrição matemática das características dinâmicas do sistema que permita prever o seu desempenho antes mesmo do sistema ser realizado ou construído.

O modelo matemático é uma aproximação do sistema físico, ou seja, não é o sistema físico! Sendo uma aproximação, o modelo tem limites de validade. Por exemplo, uma montagem eletrónica com um amplificador operacional (ou *ampop*) pode ser observada a diversos níveis de complexidade, desde a simples relação de ganho (Figura 2.1),

$$V_o = A(V_+ - V_-)$$

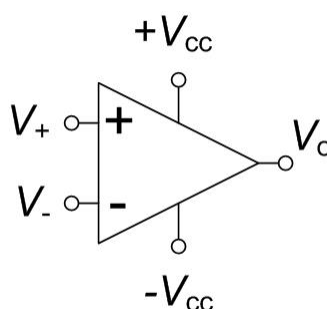


Figura 2.1 – Representação esquemática de alto nível de um amplificador operacional.

passando pela consideração dos transístores que integram o amplificador, até, no limite, à consideração de todos os efeitos capacitivos e indutivos existentes em cada micrómetro do circuito integrado. As diversas formas de abordar o problema irão resultar em modelos de complexidade crescente. Cabe ao engenheiro dedicado ao problema da construção do modelo encontrar o melhor compromisso entre simplicidade e exatidão.

Os passos para a construção de um modelo de um sistema físico são os apresentados nos parágrafos seguintes.

Inicia-se pelo desenho de um esquema da instalação, definindo as variáveis físicas relevantes, nomeadamente, as entradas de atuação, saídas medidas por sensores e outras variáveis que representem a acumulação de energia nos diversos componentes.

De seguida, escrevem-se as equações para cada componente usando as leis físicas adequadas (e.g. a lei de Newton para uma massa sob a ação de uma força ou a lei característica de um componente electrónico).

Combinam-se as equações obtidas de acordo com o primeiro esquema de forma a obter o modelo matemático completo.

Se possível, verifica-se a validade do modelo, comparando resultados obtidos experimentalmente na instalação com a solução das equações do modelo matemático sujeitos aos mesmos sinais de excitação.

Se o resultado final não for satisfatório devemos reiniciar o processo realizando uma descrição mais detalhada das componentes dominantes.

No entanto, nem sempre é possível ou praticável realizar a modelação a partir de leis físicas de todas as componentes. Podemos, nesses casos, socorrer-nos de modelos fornecidos pelos fabricantes das partes ou pela análise direta dos sinais de excitação-reação. Este último processo é designado por **identificação do modelo**.

Os exemplos seguintes ilustram o processo de modelação em algumas áreas diferentes da engenharia.

Exemplo 2.1

Considere a instalação para o comando da posição φ de uma antena cujo esquema tecnológico está representado na Figura 2.2.

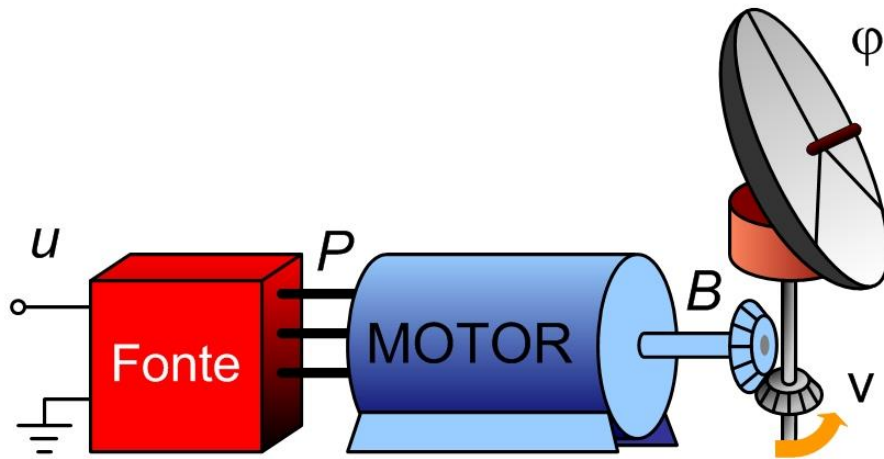


Figura 2.2 – Esquema tecnológico da instalação do Exemplo 2.1.

O primeiro passo será então identificar as principais variáveis físicas que intervêm no funcionamento da instalação. São estas (ver Figura 2.2):

- a tensão de comando $u(t)$ à entrada da fonte de potência;
- a potência eléctrica $P(t)$ que é fornecida ao motor;
- o binário mecânico $B(t)$ que é transmitido do motor para a estrutura móvel da antena;
- a posição azimutal $\varphi(t)$ da antena.

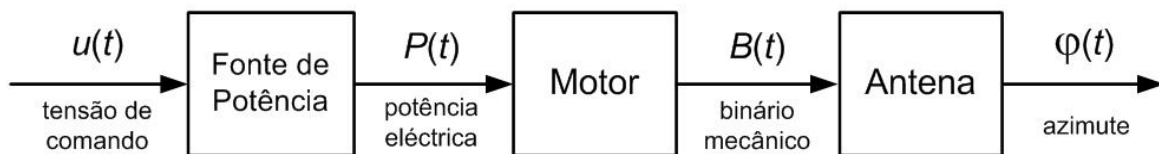


Figura 2.3 – Esquema funcional da instalação do Exemplo 2.1.

Os passos seguintes passariam por encontrar, para cada um dos três componentes, a relação entre as variáveis físicas à entrada e à saída. Este exemplo será deixado por aqui. No entanto, existe uma questão fundamental ilustrada neste exemplo. Embora as variáveis u , P , B e φ estejam presentes nas Figura 2.2 e Figura 2.3, são em cada uma das figuras entidades diferentes! Na primeira, são variáveis físicas: uma tensão de comando u , em volt, uma potência eléctrica P , em watt, etc. Na segunda figura, as variáveis, que representam as variáveis físicas, são funções, ou melhor, são sinais, no tempo contínuo. Os sinais são transmitidos entre os blocos funcionais sem carregarem os blocos que os transmitem e não representam a fluidez de qualquer entidade física como e.g. massa ou energia. A fluência é apenas de informação.

■

Exemplo 2.2

Considere agora, como instalação, um espaço habitacional e a modelação do seu comportamento térmico (representado esquematicamente na Figura 2.4).

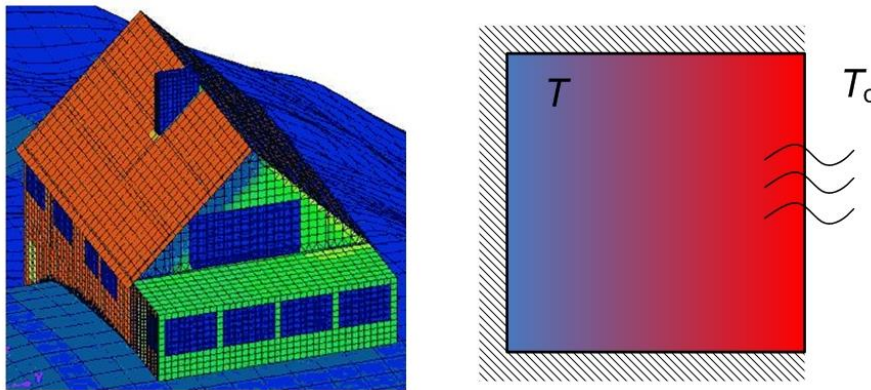


Figura 2.4 – Esquema da instalação do Exemplo 2.2.

Por simplicidade iremos representar a habitação por apenas uma divisão, de volume V , e considerar que a temperatura T_i é homogénea no interior desse espaço. A divisão está, do ponto de vista térmico, bem isolada com exterior à temperatura T_o , exceto numa das fachadas, em vidro, com área A e coeficiente de transmissão de calor κ , em $WK^{-1}m^{-2}$.

Faça-se então o modelo desta instalação. O esquema funcional (Figura 2.5) resume-se a um bloco de entrada $T_o(t)$, a influência do exterior (**perturbação**), e saída $T_i(t)$, a variável física (agora, já o sinal que a representa) que reflete a acumulação de energia térmica no espaço.



Figura 2.5 – Esquema funcional da instalação do Exemplo 2.2.

O segundo passo é usar a lei física do princípio de conservação de energia para escrever

$$\frac{dE(t)}{dt} = \kappa A(T_o(t) - T_i(t))$$

Esta equação estabelece que a variação, ao longo do tempo, da energia térmica $E(t)$ armazenada no espaço, só existe se houver transmissão de calor através da superfície vidrada que separa os espaços às temperaturas T_i e T_o . Por outro lado, a temperatura T_i é uma medida desta energia acumulada $E(t)$ através da relação

$$E = \rho_{ar} V S_{ar} T_i$$

onde ρ_{ar} é a densidade do ar, em kgm^{-3} , e S_{ar} é o calor específico do ar, em $\text{Jkg}^{-1}\text{K}^{-1}$. Substituindo esta relação na equação de conservação anterior, resulta (verifique como exercício)

$$\frac{\rho_{ar}VS_{ar}}{\kappa A} \cdot \frac{dT_i(t)}{dt} + T_i(t) = T_o(t)$$

tal que, para que esta equação seja dimensionalmente correta, a constante

$$\tau = \frac{\rho_{ar}VS_{ar}}{\kappa A}$$

tem a dimensão de um tempo. O parâmetro τ é a constante de tempo da habitação e é uma quantificação de quão rápida será a evolução da temperatura interior face a uma variação no exterior. Assim, o modelo da instalação é dado na forma de uma equação diferencial (linear) de 1º grau, dada por

$$\tau \cdot \frac{dT_i(t)}{dt} + T_i(t) = T_o(t)$$

No capítulo seguinte teremos oportunidade de estudar a evolução temporal deste sistema e assim, estarmos habilitados a comparar a solução desta equação com resultados experimentais disponíveis.

■

Exemplo 2.3

Este outro exemplo considera um sistema clássico massa, mola e amortecedor, como o existente na suspensão de um automóvel vulgar (Figura 2.6). Por simplicidade, considere-se que o pneu faz uma ligação rígida ao solo.

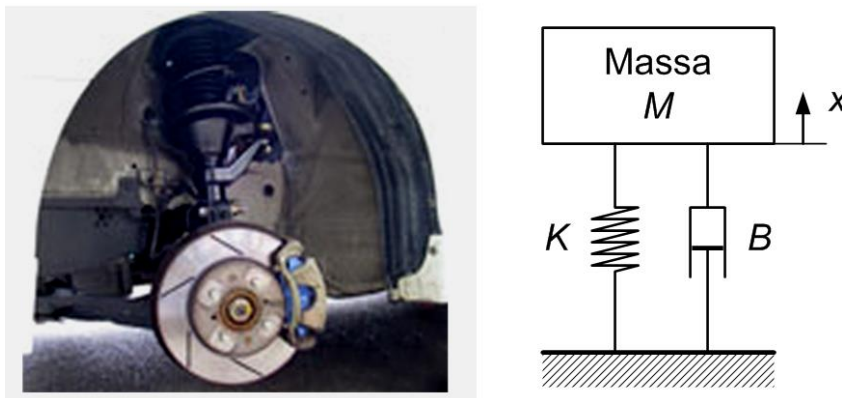


Figura 2.6 – Esquema da instalação do Exemplo 2.3.

As características dos elementos mola e amortecedor são conhecidas e dadas, respetivamente, pelas expressões das forças

$$F_m = -K(x - x_0)$$

$$F_a = -B \left(\frac{dx}{dt} - \frac{dx_0}{dt} \right) = -B \frac{dx}{dt}$$

onde x_0 é a posição em repouso (equilíbrio entre a força gravítica F_g e a tensão na mola); K é a constante da mola em Nm^{-1} ; e B é a constante do amortecedor em Nsm^{-1} . Agora, podemos utilizar a lei de Newton para escrever

$$M \frac{d^2x(t)}{dt^2} = F_m(t) + F_a(t) - F_g(t)$$

ou, por substituição,

$$M \frac{d^2x(t)}{dt^2} = -K(x(t) - x_0) - B \frac{dx(t)}{dt} - Mg$$

Mas, como pela condição do repouso, $Kx_0 = Mg$, obtém-se

$$M \frac{d^2x(t)}{dt^2} = -Kx(t) - B \frac{dx(t)}{dt}$$

resultando no modelo final de 2ª ordem

$$\frac{d^2x(t)}{dt^2} + \frac{B}{M} \cdot \frac{dx(t)}{dt} + \frac{K}{M} \cdot x(t) = 0$$

■

Note-se que o sistema do exemplo anterior não é excitado por qualquer variável exterior (à parte da força da gravidade que é constante e, portanto, não realiza excitação). Um sistema sem excitação exterior diz-se **autónomo**.

Exemplo 2.4

Mudemos agora para algo tecnologicamente muito diferente. Consideremos o comportamento elétrico de uma antena que tem no circuito RLC série um esquema equivalente (Figura 2.7).

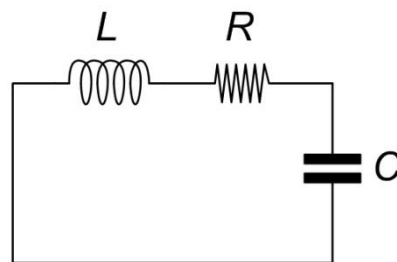


Figura 2.7 – Esquema da instalação do Exemplo 2.4.

As características dos elementos indutor, resistência e condensador são dadas, respetivamente, pelas expressões

$$V_L = L \frac{dI_L}{dt}$$

$$V_R = R \cdot I_R$$

$$I_C = C \frac{dV_C}{dt}$$

onde L é a indutância do indutor em H; R é a resistência em Ω ; e C é a capacidade do condensador em F. Usando as leis das malhas e dos nós podemos escrever

$$V_L + V_R + V_C = 0$$

$$I_L = I_R = I_C$$

e substituindo as características dos elementos obtém-se

$$L \frac{dI_C}{dt} + R \cdot I_C + V_C = 0$$

ou

$$L \frac{d}{dt} \left[C \frac{dV_C}{dt} \right] + R \left[C \frac{dV_C}{dt} \right] + V_C = 0$$

resultando no modelo final de 2ª ordem

$$\frac{d^2 V_C(t)}{dt^2} + \frac{R}{L} \cdot \frac{dV_C(t)}{dt} + \frac{1}{LC} \cdot V_C(t) = 0$$

também autónomo. ■

Note-se que apesar das diferenças tecnológicas entre os sistemas físicos dos Exemplo 2.3 e Exemplo 2.4, as equações diferenciais resultantes

- sistema massa, mola e amortecedor: $\frac{d^2 x(t)}{dt^2} + \frac{B}{M} \cdot \frac{dx(t)}{dt} + \frac{K}{M} \cdot x(t) = 0$
- sistema RLC série: $\frac{d^2 V_C(t)}{dt^2} + \frac{R}{L} \cdot \frac{dV_C(t)}{dt} + \frac{1}{LC} \cdot V_C(t) = 0$

são semelhantes à parte dos parâmetros. Isto significa que do ponto de vista temporal as variáveis x (a posição da massa) e V_C (a tensão no condensador) terão exatamente o mesmo comportamento se

$$\frac{B}{M} = \frac{R}{L} \quad \text{e} \quad \frac{K}{M} = \frac{1}{LC}$$

Na realidade, é possível fazer mais que isso se estabelecermos uma relação entre os tempos de simulação dos dois sistemas, por exemplo, se quisermos que um milissegundo no sistema eletrónico corresponda ao tempo de um segundo do sistema mecânico, bastará estabelecer as relações

$$\frac{R}{L} = 10^3 \frac{B}{M} \quad \text{e} \quad \frac{1}{LC} = 10^6 \frac{K}{M}$$

O ponto importante é que através da modelação matemática podemos estabelecer uma plataforma comum de análise de sistemas, independentemente da sua origem tecnológica. No entanto, o engenheiro nunca deve perder a distância ao sistema físico original e das limitações inerentes ao processo de modelação. O modelo matemático não é a instalação física, tal como uma maquete realizada por um arquiteto não é o edifício construído pelo engenheiro civil.

2.2. Linearização de sistemas físicos

Um **sistema linear** em tempo contínuo pode ser sempre colocado na forma de uma equação diferencial como

$$\begin{aligned} \frac{d^n y(t)}{dt^n} + a_1 \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + a_{n-1} \frac{dy(t)}{dt} + a_n y(t) \\ = b_0 \frac{d^m u(t)}{dt^m} + b_1 \frac{d^{m-1} u(t)}{dt^{m-1}} + \dots + b_{m-1} \frac{du(t)}{dt} + b_m u(t) \end{aligned}$$

ou seja, uma combinação linear das derivadas da ordem 0 até à ordem n para a saída y ; e uma combinação linear das derivadas da ordem 0 até à ordem m para a entrada u .

No entanto, os sistemas físicos não são, em geral, lineares. É vasta a literatura que se dedica ao seu estudo. Qual é então o interesse do estudo dos sistemas lineares e o que fazer quando o sistema físico resultar num modelo não linear?

O interesse advém de duas razões. A primeira prende-se com a relativa facilidade com que desenvolvem ferramentas para tratar os sistemas lineares, que resultam da própria característica linear (sobreposição e escalamento). No entanto, não chega termos um martelo na mão para que todos os problemas sejam pregos. O segundo motivo, mais importante, está relacionado com a finalidade deste texto, o controlo. O controlo por retroação realiza uma linearização do sistema, ou seja, uma instalação não linear colocada em anel fechado será mais linear. Assim, parte-se do modelo para a instalação não linear; lineariza-se em torno de um ponto de operação adequado; projeta-se o controlador para o modelo linearizado e, no final, o controlador garantirá a manutenção do funcionamento da instalação em redor do ponto de operação em que o modelo foi linearizado.

Exemplo 2.5

Considere-se a instalação esquematizada na Figura 2.8. Uma bomba comandada pelo sinal de **atuação** $u(t)$ fornece água, através do tanque 1, a um sistema de dois tanques ligados por um tubo de ligação de secção conhecida. O tanque 2, à direita, tem ainda um tubo de extração,

por onde a água escoa livremente. Este tanque está instrumentado com um sensor de nível de água, produzindo o sinal de saída $y(t)$. Os tanques têm ambos área da base A .

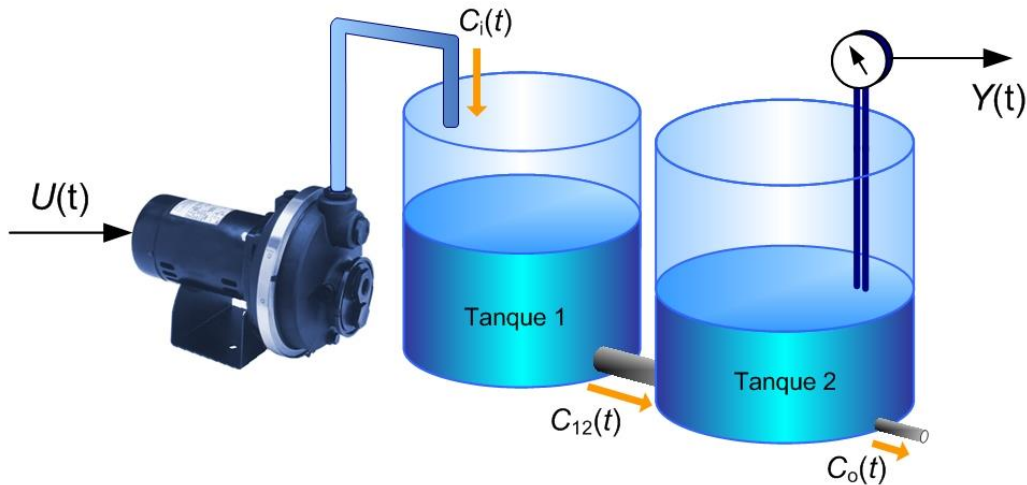


Figura 2.8 – Esquema da instalação do Exemplo 2.5.

A primeira tentativa para identificar a entrada e a saída da instalação é observar por onde a água entra e sai! No entanto, como não somos canalizadores, devemos identificar as entradas e saídas, não de água, mas de sinal. À análise do comportamento dinâmico dos sistemas interessa a relação entre sinais. Sinais de entrada (comando de velocidade da bomba) que provocam variações no sinal de saída (medida do nível no tanque). Esta distinção é crucial ao entendimento dos sistemas dinâmicos.

Façamos então o modelo matemático desta instalação começando por desenhar o respetivo esquema funcional (ver Figura 2.9). Explicando o esquema, temos uma função, representada pelo bloco *Bomba* que relaciona o sinal de comando da bomba $U(t)$ com o caudal volumétrico $C_i(t)$ que entra no tanque 1. Por outro lado, existirá um caudal $C_{12}(t)$ no tubo de ligação do tanque 1 para o tanque 2, positivo se for do tanque 1 para o tanque 2, negativo se for no sentido inverso. A diferença entre aqueles caudais $C_i(t) - C_{12}(t)$ é o balanço do volume de água que entra no tanque 1 por unidade de tempo. Integrando esta diferença e dividindo pela área do tanque 1 (uniforme ao longo da altura do tanque) obtém-se o nível de água no tanque, ou seja

$$H_1(t) = \frac{1}{A} \int_{-\infty}^t (C_i(\sigma) - C_{12}(\sigma)) d\sigma$$

ou na versão diferencial

$$\frac{dH_1(t)}{dt} = \frac{1}{A} (C_i(t) - C_{12}(t))$$

Esta equação pode ser lida como, a variação do volume de água no tanque 1 dada por $A \cdot dH_1(t)/dt$, é igual à diferença de caudais que entram e que saem desse tanque, ou seja, $C_i(t) - C_{12}(t)$.

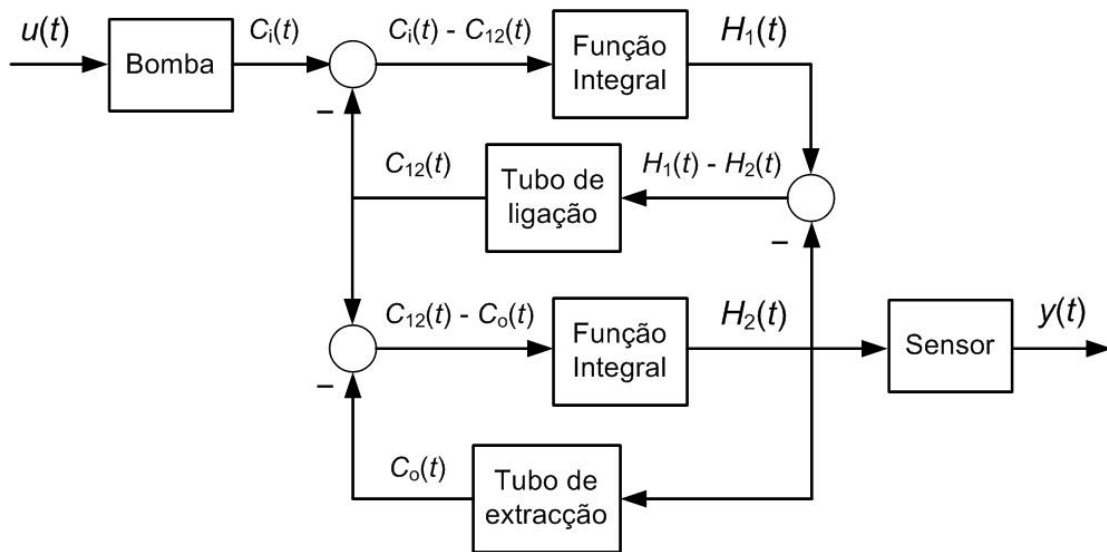


Figura 2.9 – Esquema funcional da instalação do Exemplo 2.5.

A mesma lei pode ser escrita relativamente ao tanque 2,

$$\frac{dH_2(t)}{dt} = \frac{1}{A} (C_{12}(t) - C_o(t))$$

Pensemos ainda que, para simplificar, a bomba de água tem um comportamento linear dado por

$$C_i(t) = \beta \cdot U(t)$$

Até este ponto estamos na presença de um sistema descrito por duas equações diferenciais lineares. No entanto, a relação entre os caudais e os níveis nos tanques é não linear e necessitamos de realizar um procedimento de linearização para obter um modelo linear da instalação.

Naturalmente, o caudal $C_{12}(t)$ que se estabelece entre os dois tanques dependerá da diferença de níveis entre estes. A lei de Bernoulli estabelece que o caudal é proporcional à área do orifício de passagem a_{12} e à raiz quadrada da diferença de níveis de água. O caudal volumétrico que se estabelece entre os dois tanques é então dado por

$$C_{12}(t) = a_{12} \sqrt{2g} \sqrt{H_1(t) - H_2(t)}$$

onde g é a aceleração gravítica. Do mesmo modo, o caudal de saída $C_o(t)$ depende do nível do tanque 2. De novo, através da aplicação da lei de Bernoulli resulta que

$$C_o(t) = a_o \sqrt{2g} \sqrt{H_2(t)}$$

Estas duas equações que estabelecem os caudais são não lineares na sua dependência com as variáveis níveis de água. No entanto, se estabelecermos um ponto de funcionamento (em redor de um ponto de equilíbrio) para a instalação, é possível obter um modelo linear válido para *pequenas* variações em redor do ponto de operação.

Antes, estabeleçam-se os possíveis pontos de equilíbrio. Estes são caracterizados por não apresentarem variações do valor das variáveis ao longo do tempo, ou seja,

$$\frac{dH_1(t)}{dt} = \frac{1}{A} (C_i(t) - C_{12}(t)) = 0$$

$$\frac{dH_2(t)}{dt} = \frac{1}{A} (C_{12}(t) - C_o(t)) = 0$$

implicando que, se o comando da bomba for constante e igual a U_{eq}

$$C_o(t) = C_{12}(t) = C_i(t) = \beta \cdot U_{eq}$$

Significa que o mesmo caudal (constante) estabelecido pelo comando da bomba $\beta \cdot U_{eq}$ atravessa o sistema de dois tanques até à saída de água do tanque 2; e os níveis (também constantes) são dados por

$$H_2(t) = H_{2,eq} = \left[\frac{C_o(t)}{a_o \sqrt{2g}} \right]^2 = \frac{\beta^2}{2a_o^2 g} U_{eq}^2$$

$$H_1(t) = H_{1,eq} = \left[\frac{C_{12}(t)}{a_{12} \sqrt{2g}} \right]^2 + H_{2,eq} = \left(\frac{1}{a_{12}^2} + \frac{1}{a_o^2} \right) \cdot \frac{\beta^2}{2g} U_{eq}^2$$

Para cada valor constante U_{eq} estabelecido no comando da bomba existirão pontos de equilíbrio diversos com níveis nos dois tanques também constantes.

A linearização surge agora de pensarmos em variações na vizinhança deste **ponto de equilíbrio** (ou ponto de operação). Quanto mais no afastarmos deste ponto, menos válido será o modelo linear obtido.

De seguida, definimos variáveis incrementais que representam o desvio face ao ponto de equilíbrio como

$$u(t) = U(t) - U_{eq}$$

$$h_1(t) = H_1(t) - H_{1,eq}$$

$$h_2(t) = H_2(t) - H_{2,eq}$$

para expandir em série de Taylor até ao 1º grau as expressões dos caudais

$$C_i(t) = \beta U(t) = \beta U_{eq} + \beta u(t)$$

$$C_{12}(t) \approx \beta U_{eq} + \left. \frac{dC_{12}}{dH_1} \right|_{eq} h_1(t) + \left. \frac{dC_{12}}{dH_2} \right|_{eq} h_2(t)$$

$$C_o(t) \approx \beta U_{eq} + \left. \frac{dC_o}{dH_2} \right|_{eq} h_2(t)$$

Finalmente, substituindo as expressões lineares aproximadas nas equações de balanço de volume obtém-se

$$\frac{dH_1(t)}{dt} = \frac{dh_1(t)}{dt} = \frac{1}{A} \left(\beta U_{eq} + \beta u(t) - \beta U_{eq} - \left. \frac{dC_{12}}{dH_1} \right|_{eq} h_1(t) - \left. \frac{dC_{12}}{dH_2} \right|_{eq} h_2(t) \right)$$

$$\frac{dH_2(t)}{dt} = \frac{1}{A} \left(\beta U_{eq} + \left. \frac{dC_{12}}{dH_1} \right|_{eq} h_1(t) + \left. \frac{dC_{12}}{dH_2} \right|_{eq} h_2(t) - \beta U_{eq} - \left. \frac{dC_o}{dH_2} \right|_{eq} h_2(t) \right)$$

ou (demonstre como exercício)

$$\frac{dh_1(t)}{dt} = -\frac{\alpha}{A} h_1(t) + \frac{\alpha}{A} h_2(t) + \frac{\beta}{A} u(t)$$

$$\frac{dh_2(t)}{dt} = \frac{\alpha}{A} h_1(t) - \frac{\alpha + \alpha_o}{A} h_2(t)$$

onde

$$\alpha = \left. \frac{dC_{12}}{dH_1} \right|_{eq} = -\left. \frac{dC_{12}}{dH_2} \right|_{eq} = \frac{a_{12}\sqrt{g}}{\sqrt{2(H_{1,eq} - H_{2,eq})}}$$

$$\alpha_o = \left. \frac{dC_o}{dH_2} \right|_{eq} = \frac{a_o\sqrt{g}}{\sqrt{2H_{2,eq}}}$$

ou seja, parâmetros que só dependem dos parâmetros físicos da instalação e do ponto de operação selecionado. ■

Note-se que no problema dos tanques é necessário garantir uma componente contínua que posiciona o atuador no ponto de funcionamento adequado. Esta componente pode ser fornecida diretamente de modo constante, ou ver-se-á mais tarde, ser integrada no controlador de forma dinâmica com o efeito integral. Quanto ao sinal de saída, como a subtração da constante é realizada também no sinal de referência a seguir, o erro de controlo é exatamente igual à situação linear. Resumindo, em geral, não será preciso tomar cuidados especiais na realização tecnológica do sistema de controlo para garantir este posicionamento da operação da instalação. Ter-se-á, no entanto, que inicializar a memória interna (integrador) do controlador relacionada com o valor constante inicial.

Assim, quando neste texto se representar a variável de saída $y(t)$ inicialmente igual a zero, significa apenas que partimos do equilíbrio em redor do ponto de linearização. Quanto mais nos afastarmos deste ponto, maiores serão as diferenças entre as dinâmicas do modelo linear e do sistema físico não linear representado.

3. Resposta no tempo

Como referido anteriormente, este texto concentra o seu estudo nos sistemas lineares invariantes no tempo (SLIT).

Um primeiro objetivo na modelação de sistemas dinâmicos é a possibilidade de simular no domínio do tempo a evolução do sinal de saída $y(t)$. Esta variação do sinal ao longo do tempo pode derivar de condições iniciais (fora do equilíbrio), do efeito de um sinal de excitação particular na entrada do sistema sob estudo ou de ambos.

3.1. Resposta livre

A resposta de um sistema a um conjunto de condições iniciais não nulas mas sem excitação na entrada ($u(t) = 0$), designa-se por **resposta livre** $y_L(t)$.

Exemplo 3.1

Considere-se o sistema caracterizado pela seguinte equação diferencial

$$\tau \frac{dy(t)}{dt} + y(t) = u(t)$$

com $\tau = 4.0$ s, entrada identicamente nula, $u(t) = 0$, e condição inicial $y(0) = 2.0$.

Propõe-se como solução para a equação diferencial um sinal exponencial, dado que a função exponencial é caracterizada por ter uma derivada igual a si própria. Assim, a solução livre toma a forma

$$y_L(t) = a \cdot e^{bt}$$

que substituída na equação diferencial resulta em

$$\tau \cdot a \cdot b + a = 0$$

ou

$$b = -\frac{1}{\tau} = -\frac{1}{4.0} = -0.25$$

e, da condição inicial,

$$y_L(0) = a \cdot e^{b \cdot 0} = a = 2.0$$

A solução é então dada por

$$y_L(t) = 2.0 \cdot e^{-0.25t}$$

representada na Figura 3.1.

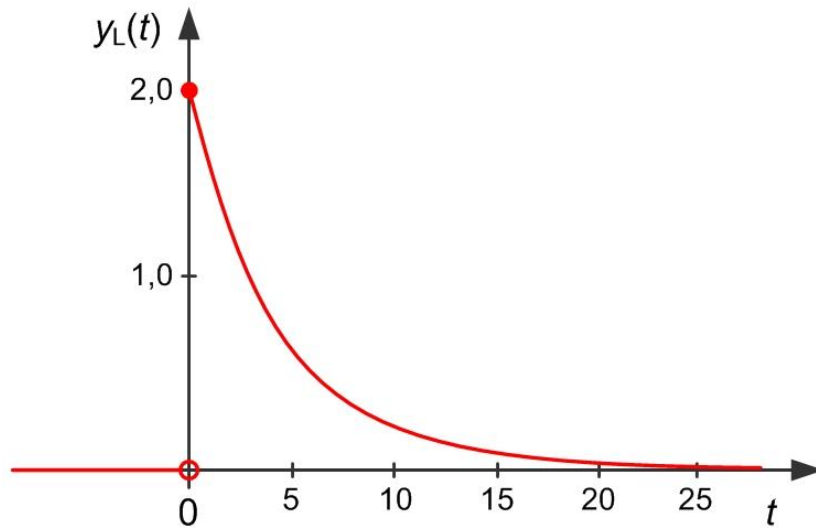


Figura 3.1 – Resposta livre (Exemplo 3.1).

■

Existe um conjunto de sinais preferenciais, frequentemente usados para a análise dinâmica de um sistema dinâmico. Começemos por introduzir o sinal em degrau unitário.

3.2. Degrâu unitário

Um sinal em degrau representa uma transição instantânea entre dois valores constantes num dado instante de tempo particular. O degrâu unitário é um caso particular daquele, transitando do valor 0 para o valor 1 no instante de tempo $t = 0$. O degrâu unitário define-se então por

$$\varepsilon(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

e a sua representação temporal é apresentada na Figura 3.2.

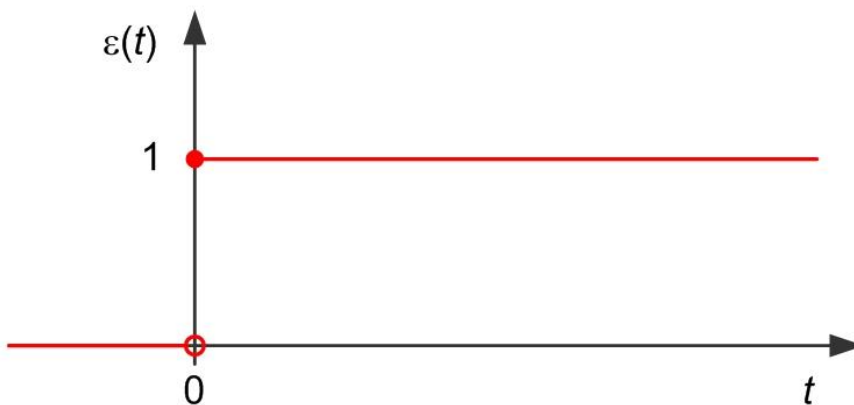


Figura 3.2 – Sinal degrâu unitário.

3.3. Resposta forçada

A resposta de um sistema a um sinal de excitação na entrada designa-se por **resposta forçada** $y_F(t)$.

Exemplo 3.2

Considere-se, de novo, o sistema caracterizado pela equação diferencial

$$\tau \frac{dy(t)}{dt} + y(t) = u(t)$$

com $\tau = 4.0$ s, condição inicial nula $y(0) = 0$ e entrada em degrau unitário $u(t) = \varepsilon(t)$.

Como no Exemplo 3.1, a solução é proposta baseada num sinal exponencial. No entanto, é necessário considerar uma componente constante que equilibre os dois lados da equação.

Assim, a solução forçada toma a forma proposta dado por

$$y_F(t) = a \cdot e^{bt} + c$$

que substituída na equação diferencial resulta para $t > 0$ em

$$\tau \cdot a \cdot b \cdot e^{bt} + a \cdot e^{bt} + c = 1.0$$

Daqui se conclui que

$$b = -\frac{1}{\tau} = -\frac{1}{4.0} = -0.25$$

$$c = 1.0$$

e, da condição inicial nula,

$$y_F(0) = a \cdot e^{b \cdot 0} + c = 0.0$$

ou seja,

$$a = -c = -1.0$$

A solução final é então dada por ($t > 0$)

$$y_F(t) = 1.0 - e^{-0.25t}$$

representada na Figura 3.3.

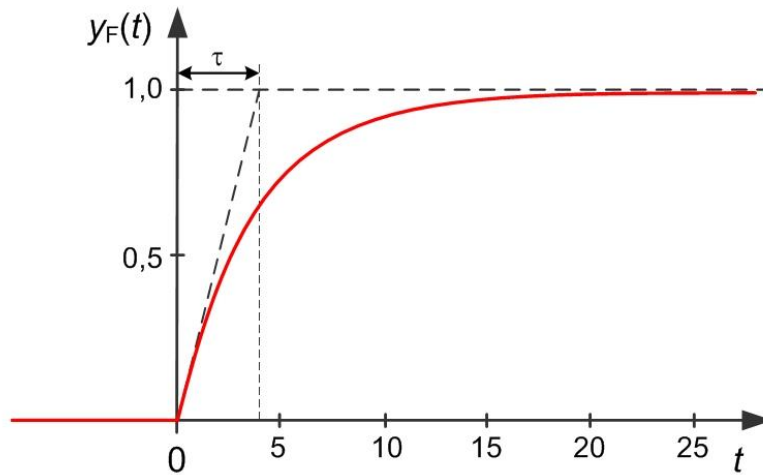


Figura 3.3 – Resposta forçada ao degrau (Exemplo 3.2).

■

Exemplo 3.3

Considere-se, ainda de novo, o sistema caracterizado pela equação diferencial

$$\tau \frac{dy(t)}{dt} + y(t) = u(t)$$

com $\tau = 4.0$ s, condição inicial $y(0) = 2$ e entrada em degrau unitário $u(t) = \varepsilon(t)$.

Porque o sistema é linear, a resposta surge da sobreposição das respostas livre e forçada encontradas nos exemplos anteriores. A solução é então dada por ($t > 0$)

$$y(t) = y_L(t) + y_F(t) = 1.0 + e^{-0.25t}$$

representada na Figura 3.4.

Note-se que, fazendo uso da definição de degrau unitário, é possível representar analiticamente a solução como

$$y(t) = (1.0 + e^{-0.25t}) \cdot \varepsilon(t)$$

ficando a expressão imediatamente válida para qualquer valor do tempo t anterior ou posterior a $t = 0$.

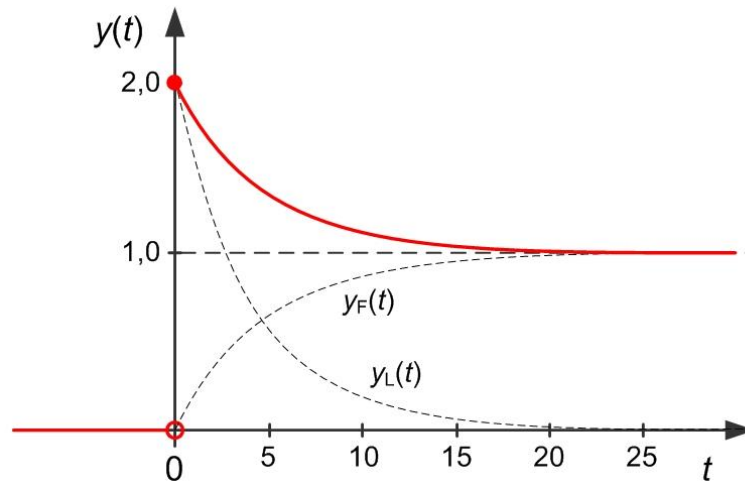


Figura 3.4 – Resposta a excitação em degrau com condições iniciais não nulas (Exemplo 3.3).

3.4. Impulso unitário

Outro sinal muito importante utilizado na análise de sistemas dinâmicos é o impulso unitário. O sinal impulso unitário é uma abstração matemática que representa a entrega ao sistema, pela entrada, de uma quantidade finita de energia num tempo infinitesimal. (Não confundir a energia de um sinal de informação com a energia em W de um sinal físico.) Na prática, o significado de infinitesimal será o de ser muito menor que as constantes de tempo naturais do sistema a excitar.

Considere-se o sinal retangular definido por

$$\delta_{\Delta}(t) = \begin{cases} 1/\Delta & 0 \leq t \leq \Delta \\ 0 & \text{c. c.} \end{cases}$$

representado na Figura 3.5.

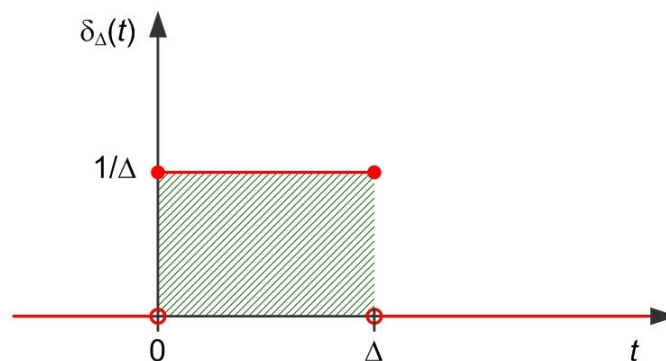


Figura 3.5 – Sinal retangular de duração Δ .

Note-se que independentemente do valor da largura do sinal retangular, a área definida na figura (energia do sinal) é sempre de valor unitário.

O impulso unitário é obtido a partir do limite, quando a largura Δ vai para zero, do sinal $\delta_{\Delta}(t)$, ou seja (ver Figura 3.6)

$$\delta(t) = \lim_{\Delta \rightarrow 0} \delta_{\Delta}(t)$$

tal que

$$\int_{-\infty}^{+\infty} \delta(\sigma) d\sigma = 1$$

No limite, o impulso unitário representa um sinal com amplitude infinita e duração nula mas cuja energia é unitária.

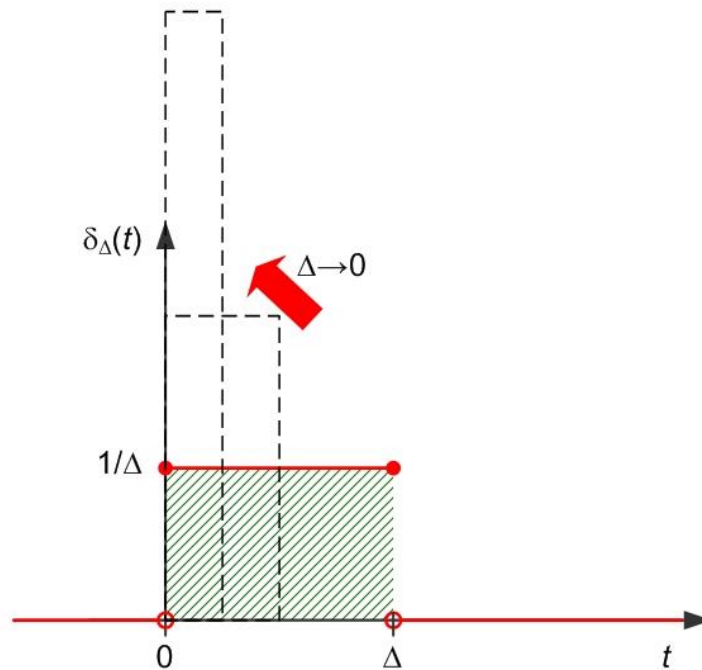


Figura 3.6 – Impulso unitário como o limite do sinal retangular.

Do facto do impulso unitário ter energia unitária pode estabelecer-se uma relação entre o impulso unitário e o degrau unitário como

$$\varepsilon(t) = \int_{-\infty}^t \delta(\sigma) d\sigma$$

ou de forma inversa

$$\delta(t) = \frac{d\varepsilon(t)}{dt}$$

3.5. Resposta impulsiva

A resposta de um sistema dinâmico a um impulso unitário na sua entrada designa-se por resposta impulsiva $h(t)$. Adiante, ver-se-á que a resposta impulsiva caracteriza de uma forma única um sistema linear invariante no tempo.

Exemplo 3.4

Considere-se mais uma vez a equação diferencial de 1º grau dada por

$$\tau \frac{dy(t)}{dt} + y(t) = u(t)$$

com $\tau = 4$ s, condição inicial nula $y(0) = 0$ e, agora, entrada em impulso unitário dada por $u(t) = \delta(t)$.

Sendo o sistema linear, a solução pode ser obtida da resposta ao degrau unitário (Exemplo 3.2) e da relação estabelecida entre o impulso e o degrau e impulso (derivada). Se o novo sinal de excitação (impulso) resulta da derivação do anterior (degrau), então a resposta impulsiva pode ser obtida derivando a resposta ao degrau unitário, obtendo-se o sinal

$$y(t) = h(t) = \frac{d}{dt} \{ (1.0 + e^{-0.25t}) \cdot \varepsilon(t) \} = 0.25 \cdot e^{-0.25t} \cdot \varepsilon(t)$$

representado na Figura 3.7.

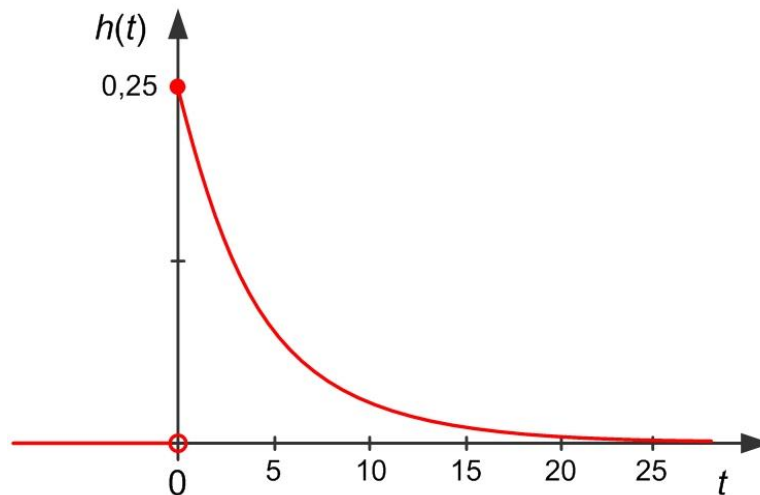


Figura 3.7 – Resposta impulsiva (Exemplo 3.4).

■

A propriedade de um sistema ser invariante no tempo significa que uma translação no tempo do sinal de entrada, resulta numa translação no tempo da mesma magnitude do sinal de saída, como se ilustra no exemplo seguinte.

Exemplo 3.5

Considere-se novamente a equação diferencial de 1º grau

$$\tau \frac{dy(t)}{dt} + y(t) = u(t)$$

Com $\tau = 4$ s, condição inicial nula $y(0) = 0$ e entrada impulsiva $u(t) = 6 \cdot \delta(t - 5)$.

Sendo o sistema linear e invariante no tempo, a solução pode ser obtida da resposta impulsiva $h(t)$ (obtida no Exemplo 3.4). A resposta do sistema é dada pelo sinal

$$y(t) = 6 \cdot h(t - 5) = 1.5 \cdot e^{-0.25t+1.25} \cdot \varepsilon(t - 5)$$

representado na Figura 3.8

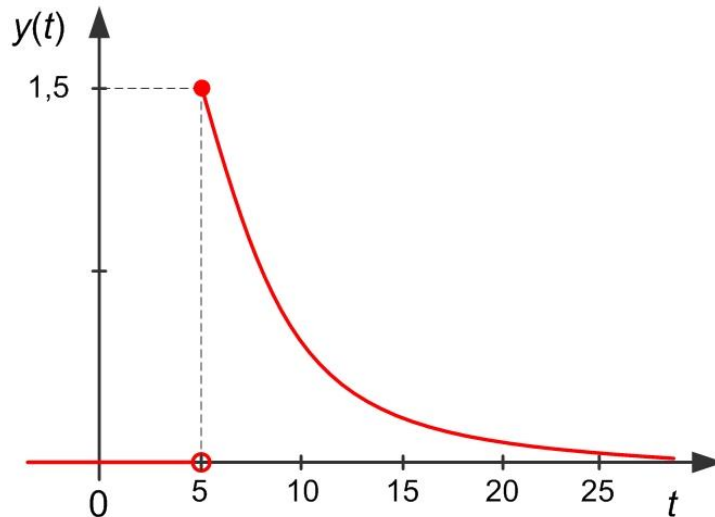


Figura 3.8 – Resposta do sistema invariante no tempo (Exemplo 3.5).

■

Apesar daquilo que os exemplos apresentados neste capítulo nos poderem levar a pensar, a resolução de equações diferenciais de ordem n não é em geral fácil. De seguida, apresenta-se uma forma alternativa de encontrar estas soluções recorrendo à **Transformação de Laplace**. Por conveniência para o estudo dos sistemas causais, este texto limitar-se-á à Transformação de Laplace unilateral direita.

3.6. Transformação de Laplace

A Transformação de Laplace é um operador que transforma um sinal $x(t)$ no domínio do tempo contínuo t , numa função $X(s)$ no domínio da variável complexa de Laplace s . A operação é definida por

$$X(s) = \int_0^{+\infty} x(\sigma) \cdot e^{-s\sigma} d\sigma$$

A Transformação de Laplace (ou Laplaciana) aparece representada na literatura em diversas formas equivalentes tais como

$$TL\{x(t)\} = \mathcal{L}\{x(t)\} = X(s)$$

A partir da definição é possível calcular a Transformação de Laplace de qualquer sinal, embora isso possa não ser simples. Os exemplos seguintes calculam as Transformações de Laplace para o impulso unitário e para o degrau unitário.

Exemplo 3.6

A Transformação de Laplace do impulso unitário $\delta(t)$ é dada por

$$\Delta(s) = \int_0^{+\infty} \delta(\sigma) \cdot e^{-s\sigma} d\sigma = 1 \quad (\text{Justifique!})$$

■

Exemplo 3.7

A Transformação de Laplace do degrau unitário $\varepsilon(t)$ é dada por

$$E(s) = \int_0^{+\infty} \varepsilon(\sigma) \cdot e^{-s\sigma} d\sigma = \int_0^{+\infty} e^{-s\sigma} d\sigma = \left[\frac{e^{-s\sigma}}{-s} \right]_0^{+\infty} = \frac{1}{s}$$

■

Adicionalmente, a Transformação de Laplace goza de um conjunto de propriedades que derivam da própria definição. Uma propriedade fundamental é a linearidade do operador de Transformação de Laplace, ou seja, sejam dois sinais genéricos $x(t)$ e $z(t)$, e duas constantes a e b tais que o sinal $w(t) = a \cdot x(t) + b \cdot z(t)$, então

$$W(s) = \text{TL}\{w(t)\} = \text{TL}\{a \cdot x(t) + b \cdot z(t)\} = a \cdot \text{TL}\{x(t)\} + b \cdot \text{TL}\{z(t)\} = a \cdot X(s) + b \cdot Z(s)$$

As tabelas das páginas seguintes apresentam as Transformações de Laplace dos sinais mais usuais e as propriedades da Transformação de Laplace.

Tabela de Transformações de Laplace

Sinal	Transformação de Laplace	
1	$\delta(t)$	1
2	$\varepsilon(t)$	$\frac{1}{s}$
3	$e^{-at} \cdot \varepsilon(t)$	$\frac{1}{s+a}$
4	$t \cdot \varepsilon(t)$	$\frac{1}{s^2}$
5	$t^n \cdot \varepsilon(t)$	$\frac{n!}{s^{n+1}}$
6	$t \cdot e^{-at} \cdot \varepsilon(t)$	$\frac{1}{(s+a)^2}$
7	$(1 - e^{-at}) \cdot \varepsilon(t)$	$\frac{a}{s(s+a)}$
8	$\sin at \cdot \varepsilon(t)$	$\frac{a}{s^2 + a^2}$
9	$\cos at \cdot \varepsilon(t)$	$\frac{s}{s^2 + a^2}$
10	$\sin(at + \varphi) \cdot \varepsilon(t)$	$\frac{(\sin \varphi)s + (\cos \varphi)a}{s^2 + a^2}$
11	$e^{-at} \cdot \sin bt \cdot \varepsilon(t)$	$\frac{b}{(s+a)^2 + b^2}$
12	$e^{-at} \cdot \cos bt \cdot \varepsilon(t)$	$\frac{s+a}{(s+a)^2 + b^2}$

Tabela de Propriedades da Transformação de Laplace

Propriedade		
P1	Definição	$X(s) = \text{TL}\{x(t)\} = \int_0^{+\infty} x(\sigma) \cdot e^{-s\sigma} d\sigma$
P2	Inversa	$x(t) = \text{TL}^{-1}\{X(s)\} = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} X(\sigma) e^{-\sigma t} d\sigma$
P3	Linearidade	Se $w(t) = a \cdot x(t) + b \cdot z(t)$, então: $W(s) = a \cdot X(s) + b \cdot Z(s)$
P4	Derivação	$\text{TL}\left\{\frac{dx(t)}{dt}\right\} = s \cdot X(s) - x(0)$
P5	Derivação de 2ª ordem	$\text{TL}\left\{\frac{d^2x(t)}{dt^2}\right\} = s^2 \cdot X(s) - s \cdot x(0) - x'(0)$
P6	Derivação de ordem n	$\text{TL}\left\{\frac{d^n x(t)}{dt^n}\right\} = s^n \cdot X(s) - s^{n-1}x(0) \dots - x^{(n)}(0)$
P7	Integração	$\text{TL}\left\{\int_0^t x(\tau) d\tau\right\} = \frac{1}{s} X(s)$
P8	Convolução	$\text{TL}\left\{\int_{-\infty}^{+\infty} x(\sigma)h(t - \sigma) d\sigma\right\} = X(s) \cdot H(s)$
P9	Escalamento no tempo	$\text{TL}\{x(at)\} = \frac{1}{a} X\left(\frac{s}{a}\right)$
P10	Translação no tempo	$\text{TL}\{x(t - \tau)\} = e^{-\tau s} \cdot X(s)$
P11	Translação em s	$\text{TL}\{e^{at} \cdot x(t)\} = X(s - a)$

3.7. Produto de convolução

Com o conhecimento adquirido nas secções anteriores deste capítulo, estamos em condições de determinar a solução de uma equação diferencial a um sinal de excitação genérico $u(t)$.

Considere-se um sistema linear invariante no tempo caracterizado pela equação diferencial

$$\begin{aligned} \frac{d^n y(t)}{dt^n} + a_1 \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + a_{n-1} \frac{dy(t)}{dt} + a_n y(t) \\ = b_0 \frac{d^m u(t)}{dt^m} + b_1 \frac{d^{m-1} u(t)}{dt^{m-1}} + \dots + b_{m-1} \frac{du(t)}{dt} + b_m u(t) \end{aligned}$$

e o sinal $u(t)$ representado, de forma aproximada, por um trem de sinais retangulares

$$u(t) \approx \sum_{i=-\infty}^{+\infty} u(t_i) \cdot \delta_{\Delta}(t - t_i) \cdot \Delta \quad \text{com} \quad t_i = t_{i-1} + \Delta$$

como representado na Figura 3.9.

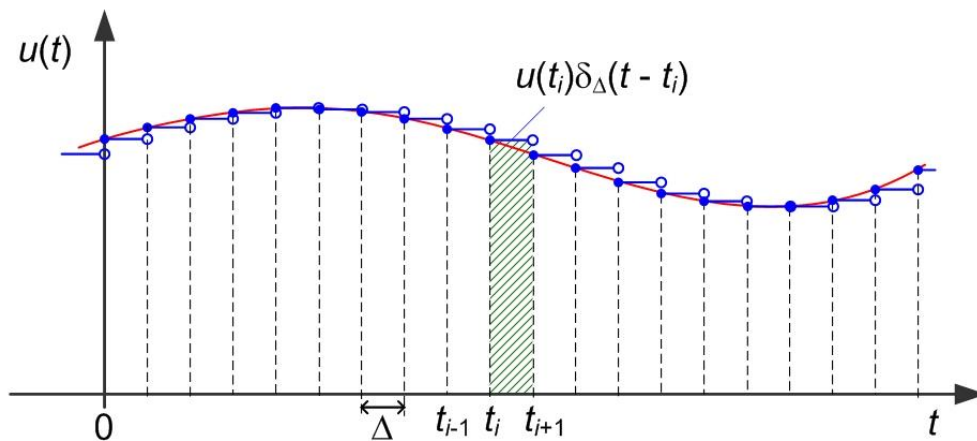


Figura 3.9 – Trem de sinais retangulares.

O ponto revelado aqui é que qualquer sinal de entrada pode ser representado por uma soma (mesmo que infinita) de sinais retangulares escalados de forma adequada. Se a largura dos sinais retangulares Δ for levada para zero, os sinais retangulares (em número crescente) convergirão para impulsos unitários, e a soma para a integração resultando em

$$u(t) = \int_{-\infty}^{+\infty} u(\sigma) \cdot \delta(t - \sigma) d\sigma$$

Conhecendo a resposta impulsiva do sistema $h(t)$, as propriedades da linearidade e da invariância no tempo do sistema implicam que a resposta a cada um dos termos da soma $u(\sigma) \cdot \delta(t - \sigma)$ será dada por $u(\sigma) \cdot h(t - \sigma)$ e

$$y(t) = \int_{-\infty}^{+\infty} u(\sigma) \cdot h(t - \sigma) d\sigma$$

A operação da equação anterior é designada por **produto de convolução** e é representada de forma abreviada por

$$y(t) = u(t) * h(t)$$

Interessante mesmo é que, aplicando a Transformação de Laplace ao produto de convolução resulta em

$$\begin{aligned} Y(s) &= \int_0^{+\infty} y(\sigma) \cdot e^{-s\sigma} d\sigma = \int_0^{+\infty} \left[\int_{-\infty}^{+\infty} u(\tau) \cdot h(\sigma - \tau) d\tau \right] \cdot e^{-s\sigma} d\sigma = \dots \\ &\dots = \left[\int_0^{+\infty} u(\sigma) e^{-s\sigma} d\sigma \right] \cdot \left[\int_0^{+\infty} h(\tau) e^{-s\tau} d\tau \right] = U(s) \cdot H(s) \end{aligned}$$

ou seja, a aplicação da Transformação de Laplace ao produto de convolução resulta num produto simples de transformadas, repetindo

$$Y(s) = H(s) \cdot U(s)$$

Exemplo 3.8

Considere-se a repetição do Exemplo 3.3 com o sistema caracterizado pela equação diferencial

$$\tau \frac{dy(t)}{dt} + y(t) = u(t)$$

com $\tau = 4$ s, condição inicial nula $y(0) = 2.0$ e entrada impulsiva $u(t) = \varepsilon(t)$.

Aplicando a Transformação de Laplace à equação diferencial obtém-se (justifique)

$$\tau[sY(s) - y(0)] + Y(s) = U(s) = \frac{1}{s}$$

ou

$$Y(s) = \frac{\tau}{\tau s + 1} y(0) + \frac{1}{\tau s + 1} \cdot \frac{1}{s}$$

Aplicando a Transformação de Laplace inversa, obtém-se

$$y(t) = 2.0 \cdot e^{-0.25t} \cdot \varepsilon(t) + (1.0 - e^{-0.25t}) \cdot \varepsilon(t)$$

representada na Figura 3.4 referida anteriormente.

■

3.8. Solução da equação diferencial usando a Transformação de Laplace

Em geral, se as condições iniciais forem nulas, pode se obter a solução da equação diferencial a partir das Transformações de Laplace da resposta impulsiva do sistema $H(s)$ e do sinal de entrada $U(s)$, ou seja

$$y(t) = TL^{-1}\{H(s)U(s)\}$$

No entanto, as conclusões a retirar deste ensaio estão limitadas ao sinal de excitação escolhido, neste exemplo, o degrau unitário. Como ganhar generalidade na análise dos sistemas?

4. Função de transferência e análise dinâmica

O capítulo anterior terminou com o levantamento de uma questão relativa a como ganhar generalidade na análise dinâmica de um sistema, independentemente do sinal de excitação à entrada. Este aspeto divide dois tipos de análise dos sistemas dinâmicos: simulação do sistema *versus* análise das características intrínsecas do sistema.

4.1. Função de transferência de um SLIT

É importante poder caracterizar um sistema dinâmico pelas suas características intrínsecas, *i.e.*, independentes do sinal de entrada e das condições iniciais. Isto é possível para os sistemas lineares, onde

$$Y(s) = H(s)U(s)$$

Define-se a **função de transferência** de um sistema dinâmico linear invariante no tempo como a função que relaciona as transformadas de Laplace dos sinais de entrada e de saída do sistema, ou seja

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

Esta função de transferência só depende das características intrínsecas do sistema.

Analicamente, a função de transferência confunde-se com a Transformação de Laplace da resposta impulsiva

$$H(s) = \text{TL}\{h(t)\}$$

e isto sucede pelo facto da Transformação de Laplace do impulso unitário ser igual à unidade. No entanto, a função de transferência representa uma relação entrada-saída, independente do sinal de excitação.

Voltando aos dois tipos de análise dos sistemas. A análise da função de transferência do sistema permite tirar conclusões diretas sobre as características dinâmicas do sistema sem envolver a análise de respostas temporais. A simulação da resposta temporal a sinais específicos, tratada no capítulo anterior, é muito importante no contexto da análise da resposta a sinais externos, e.g. perturbações ou sinais de comando.

Exemplo 4.1

Considere-se o sistema representado pela seguinte equação diferencial de 1º grau

$$\tau \frac{dy(t)}{dt} + y(t) = u(t)$$

Aplicando a Transformação de Laplace à equação diferencial obtém-se ($y(0) = 0$)

$$\tau s Y(s) + Y(s) = U(s)$$

e a função de transferência é dada por

$$H(s) = \frac{Y(s)}{U(s)} = \frac{1}{\tau s + 1}$$

■

Note-se que para esta análise, consideraram-se condições iniciais nulas. De facto, devido à linearidade do sistema em causa, as condições iniciais não modificam as conclusões que se retiram da análise da função de transferência, que relaciona os sinais à entrada e saída do sistema, se este for linear.

Exemplo 4.2

Considere-se o sistema representado na Figura 4.1.

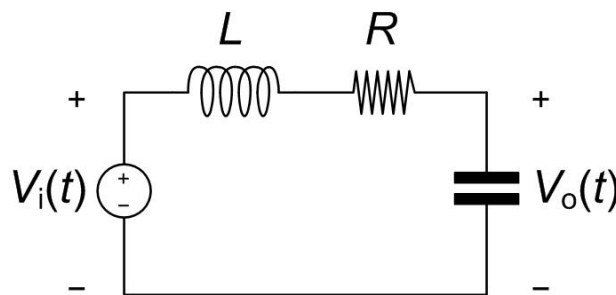


Figura 4.1 – Esquema da instalação do Exemplo 4.2.

Podemos escrever a equação diferencial que modela o comportamento do circuito eletrónico e converte-la na função de transferência, à semelhança do realizado no exemplo anterior. Vejamos, no entanto, uma forma alternativa de chegar ao mesmo resultado tirando partido de algum conhecimento da teoria dos circuitos elétricos.

A lei de Ohm diz-nos que existe proporcionalidade entre a diferença de tensão aos terminais de um elemento resistivo e a corrente que o atravessa. Mais, que esse fator de proporcionalidade define o valor da resistência, em Ohm, i.e.

$$\frac{v_R(t)}{i_R(t)} = R$$

As características dos elementos indutivo e capacitivo, dizem-nos, respetivamente, que

$$L \frac{di_L(t)}{dt} = v_L(t) \quad \text{e} \quad C \frac{dv_C(t)}{dt} = i_C(t)$$

Se aplicarmos a Transformação de Laplace às características dos três elementos (sem pré-história) obtém-se

$$\frac{V_R(s)}{I_R(s)} = R; \quad \frac{V_L(s)}{I_L(s)} = Ls \quad \text{e} \quad \frac{V_C(s)}{I_C(s)} = \frac{1}{sC}$$

Como a montagem da Figura 4.1 constitui um divisor de tensão (i.e. a corrente é estabelecida pela tensão de entrada e pela série das impedâncias dos três elementos; e a tensão de saída é igual ao produto da corrente pela impedância do elemento nesse ramo), a tensão de saída $V_o(s)$ é dada por

$$V_o(s) = \frac{1}{sC} \cdot \frac{1}{R + Ls + \frac{1}{sC}} \cdot V_i(s)$$

e logo, reordenando os termos, obtém-se para a função de transferência

$$H(s) = \frac{V_o(s)}{V_i(s)} = \frac{\frac{1}{LC}}{s^2 + \frac{R}{L}s + \frac{1}{LC}}$$

■

4.2. Conversão direta para função de transferência

A sequência de aplicação da Transformação de Laplace à equação diferencial com condições iniciais nulas pode ser automatizada num processo de conversão direta por inspeção. Se a equação diferencial estiver representada na forma

$$\begin{aligned} \frac{d^n y(t)}{dt^n} + a_1 \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + a_{n-1} \frac{dy(t)}{dt} + a_n y(t) \\ = b_0 \frac{d^m u(t)}{dt^m} + b_1 \frac{d^{m-1} u(t)}{dt^{m-1}} + \dots + b_{m-1} \frac{du(t)}{dt} + b_m u(t) \end{aligned}$$

a função de transferência resultante é dada por

$$H(s) = \frac{Y(s)}{U(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \dots + b_{m-1} s + b_m}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}$$

relacionando as Transformações de Laplace dos sinais de entrada $U(s)$ e de saída $Y(s)$.

4.3. Álgebra de blocos

Quando a complexidade de um sistema aumenta, a abordagem ao processo de modelação e análise passa, numa primeira fase, por sectionar esse processo em subsistemas mais simples, para depois integrar as funções de transferência resultantes. Para suportar essa aproximação representam-se as funções de transferência por blocos como representado na Figura 4.2.

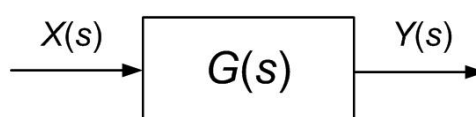


Figura 4.2 – Bloco simples de função de transferência.

O bloco é representado por um retângulo onde se inscreve a função de transferência representada. O sinal de entrada é representado por uma seta dirigida ao bloco e o sinal de saída por uma seta dirigida para o exterior. O esquema de blocos da Figura 4.2 significa então

$$Y(s) = G(s)X(s)$$

É agora possível descrever sistemas mais complexos por agregação de diversos blocos. De acordo como as entradas e saídas de os diversos blocos estão interligadas diferentes funções são realizadas pelo conjunto.

Se a saída de um bloco constituir a entrada de outro diz-se que temos os **blocos em cascata** (Figura 4.3).

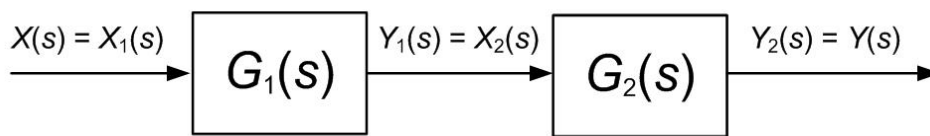


Figura 4.3 – Esquema de blocos em cascata.

Na situação da cascata, pode se escrever

$$Y(s) = G_2(s)X_2(s) = G_2(s)Y_1(s) = G_2(s)G_1(s) \cdot X(s)$$

ou seja, o esquema de blocos da Figura 4.3 é equivalente ao bloco da Figura 4.2 com

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

Note-se que no caso de se terem blocos univariáveis (i.e. uma entrada e uma saída) os blocos da Figura 4.3 são permutáveis, resultado do produto de escalares ser comutativo.

Na Figura 4.4 introduz-se um bloco diferente, designado por **bloco somador**, que realiza a soma dos diversos sinais dirigidos ao bloco. Neste caso especial o bloco é, normalmente, representado por um círculo (embora seja possível encontrar outras representações equivalentes na literatura).

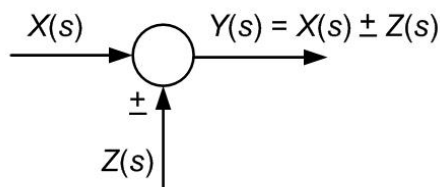


Figura 4.4 – Bloco somador de sinais.

As entradas do bloco somador podem ser afetadas de um sinal negativo. Por omissão as entradas são somadas positivamente.

Outra agregação básica de blocos é o esquema de **blocos em paralelo**, como representado na Figura 4.5. A entrada é comum aos dois blocos e as saídas respetivas são somadas, produzindo a saída do conjunto. Assim, temos a operação algébrica dada por

$$Y(s) = Y_1(s) + Y_2(s) = G_1(s)X(s) + G_2(s)X(s) = (G_1(s) + G_2(s)) \cdot X(s)$$

ou seja, o esquema de blocos da Figura 4.5 é equivalente ao bloco da Figura 4.2 com

$$G(s) = G_1(s) + G_2(s)$$

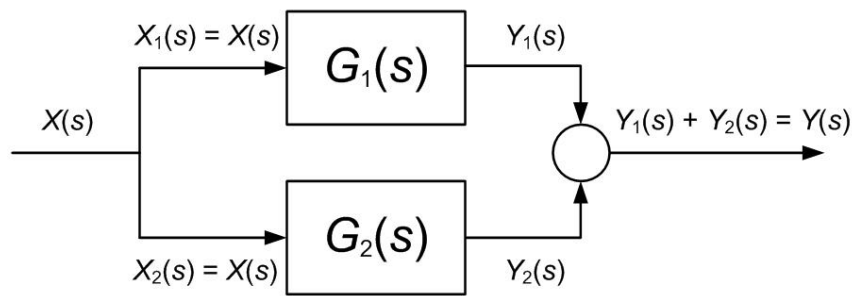


Figura 4.5 – Esquema de blocos em paralelo.

Uma questão que costuma gerar alguma dificuldade inicial na compreensão destes esquemas é a **derivação dos sinais**. Na Figura 4.5 o sinal $X(s)$ deriva para as entradas dos dois blocos em paralelo. Isto não significa que o sinal se divide, como um caudal mássico que se divide por dois canais de escoamento. A derivação não consome o sinal, que transporta apenas informação, e o sinal $X(s)$ entra completo e igualmente nos dois blocos.

Uma forma diferente de agregação básica, muito importante para o controlo, é o esquema de blocos em retroação (ver Figura 4.6). Aqui, a saída de um bloco constitui a entrada de um segundo e, por sua vez, a saída deste é usada para produzir a entrada do primeiro. Adicionalmente, é vulgar a existência de um bloco somador numa das ligações para que se possa introduzir uma entrada de **comando** do exterior.

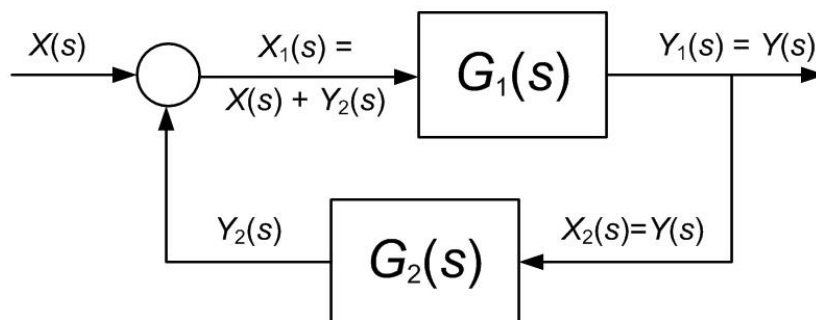


Figura 4.6 – Esquema de blocos em retroação positiva.

Assim, temos a operação algébrica dada por

$$Y(s) = G_1(s)X_1(s) = G_1(s)(X(s) + Y_2(s)) = G_1(s)X(s) + G_1(s)G_2(s)Y(s)$$

ou

$$Y(s) - G_1(s)G_2(s)Y(s) = G_1(s)X(s)$$

resultando em

$$Y(s) = \frac{G_1(s)}{1 - G_1(s)G_2(s)}X(s)$$

Assim, o esquema de blocos da Figura 4.6 é equivalente ao bloco da Figura 4.2 com

$$G(s) = \frac{G_1(s)}{1 - G_1(s)G_2(s)}$$

Como se verá adiante, é vulgar a retroação ser negativa como apresentado na Figura 4.7. Neste caso, fica como exercício mostrar que o esquema de blocos da Figura 4.7 é equivalente ao bloco da Figura 4.2 com

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

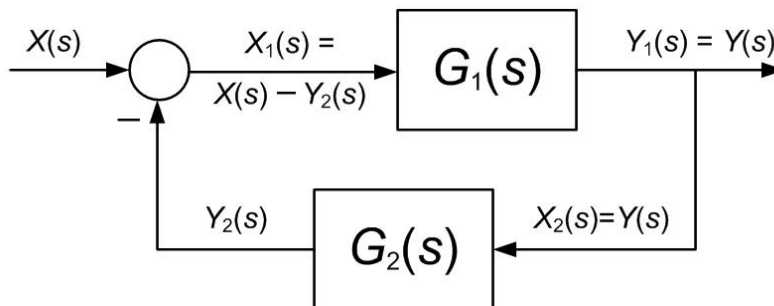


Figura 4.7 – Esquema de blocos em retroação negativa.

Este conjunto de operações algébricas permite a simplificação de esquemas de blocos que derivam da modelação dos sistemas, como se apresenta nos exemplos seguintes.

Exemplo 4.3

Considere-se o esquema de blocos representado na Figura 4.8.

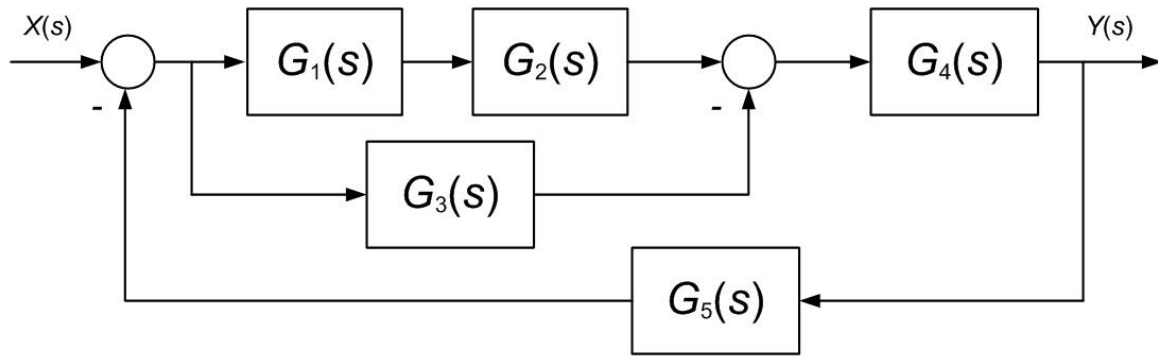


Figura 4.8 – Esquema de blocos do Exemplo 4.3.

O sistema representado neste esquema de blocos pode ser simplificado por agregação sucessiva de blocos. A Figura 4.9 identifica operações sucessivas de simplificação marcadas como A, B, C. O contorno **A** agrega em série os blocos relativos a $G_1(s)$ e $G_2(s)$ resultando em $G_2(s)G_1(s)$. Por sua vez, o contorno **B** agrega em paralelo este resultado com o bloco relativo a $G_3(s)$, resultando em $\{G_2(s)G_1(s) - G_3(s)\}$. O contorno **C** identifica novamente uma série, agora, equivalente a $G_4(s) \cdot \{G_2(s)G_1(s) - G_3(s)\}$. Finalmente, sobra a retroação negativa dos blocos no interior de **C** com o bloco relativo a $G_5(s)$, que tem como resultado final equivalente

$$G(s) = \frac{Y(s)}{X(s)} = \frac{G_4(s) \cdot \{G_2(s)G_1(s) - G_3(s)\}}{1 + G_4(s) \cdot \{G_2(s)G_1(s) - G_3(s)\} \cdot G_5(s)}$$

■

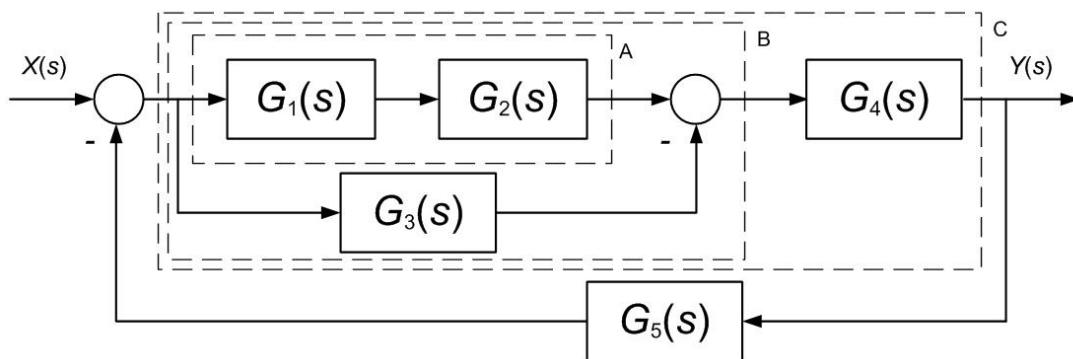


Figura 4.9 – Álgebra de blocos do Exemplo 4.3.

Por vezes, existe conveniência em realizar uma transposição de pontos de derivação ou de soma através dos blocos. Nesses casos, aplicam-se as equivalências óbvias apresentadas nas Figura 4.10 à Figura 4.13.

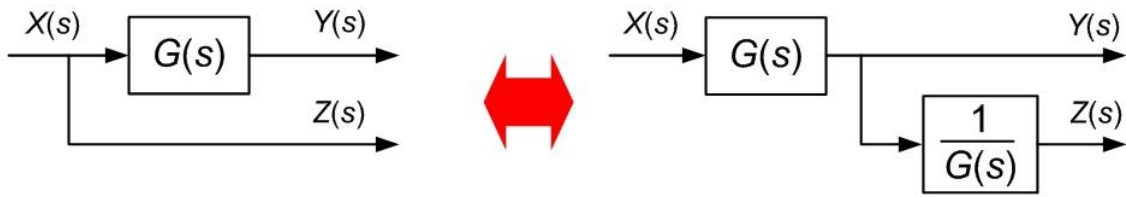


Figura 4.10 – Transposição para a frente de um ponto derivação.

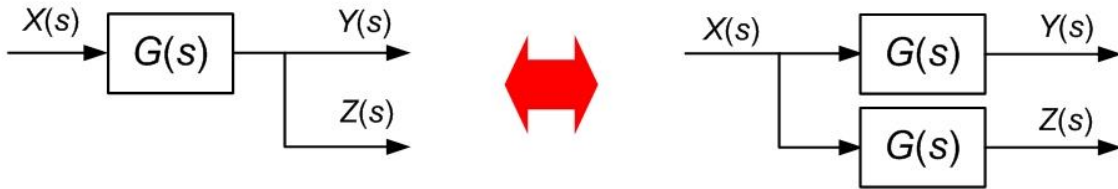


Figura 4.11 – Transposição para trás de um ponto derivação.

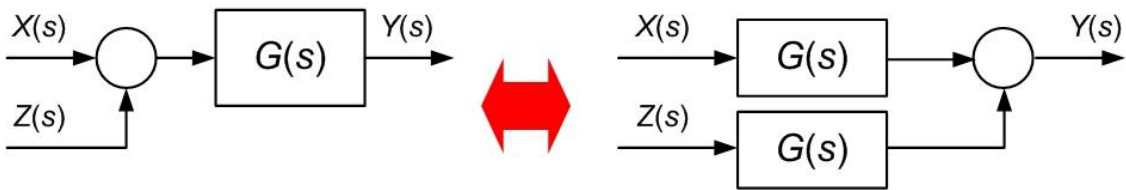


Figura 4.12 – Transposição para a frente de um bloco somador.

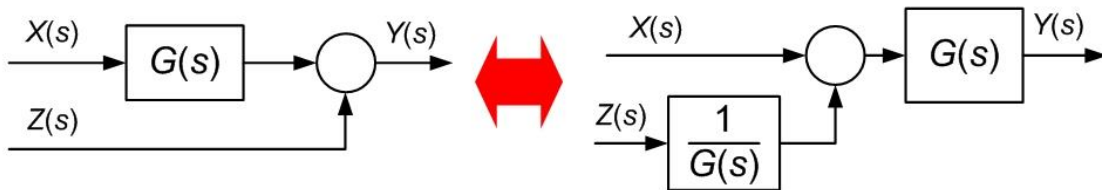


Figura 4.13 – Transposição para trás de um bloco somador.

4.4. Polos e zeros da função de transferência

A análise do comportamento dinâmico de um sistema através da sua função de transferência $G(s)$ é realizada observando a posição das raízes dos polinómios numerador $N(s)$ e denominador $D(s)$. Assim, seja $G(s)$ a função de transferência definida por

$$G(s) = \frac{N(s)}{D(s)}$$

onde o polinómio numerador, de ordem m , é dado por

$$N(s) = b_0s^m + b_1s^{m-1} + \dots + b_{m-1}s + b_m$$

e o polinómio denominador, de ordem n , é dado por

$$D(s) = s^n + a_1s^{n-1} + \dots + a_{n-1}s + a_n$$

Se o sistema for causal a ordem do denominador é superior ou igual à ordem do numerador, ou seja, se $m \leq n$. Se $m < n$, o sistema diz-se causal no sentido estrito. Se $m = n$, o sistema é causal no sentido lato.

As m raízes do numerador $N(s)$ designam-se por **zeros** da função de transferência e representam-se no plano complexo da variável s de Laplace por um zero - 'o'.

As n raízes do denominador $D(s)$ designam-se por **polos** da função de transferência e representam-se no plano complexo da variável s de Laplace por um xis - 'x'.

O lugar destas raízes, polos ou zeros, condiciona o comportamento dinâmico do sistema.

Exemplo 4.4

Considere-se o sistema descrito pela função de transferência

$$G(s) = \frac{2s^3 + 4s^2 + 8s + 16}{s^4 + 3s^3 - s^2 - 13s - 10}$$

O polinómio numerador tem as raízes (zeros da função de transferência)

$$N(s) = 2s^3 + 4s^2 + 8s + 16 = 0 \quad \left| \begin{array}{l} z_1 = -2.0 \\ z_2 = +2.0j \\ z_3 = -2.0j \end{array} \right.$$

e o polinómio denominador tem as raízes (polos da função de transferência)

$$D(s) = s^4 + 3s^3 - s^2 - 13s - 10 = 0 \quad \left| \begin{array}{l} p_1 = +2.0 \\ p_2 = -2.0 + 1.0j \\ p_3 = -2.0 - 1.0j \\ p_4 = -1.0 \end{array} \right.$$

Na Figura 4.14 observa-se o lugar destas raízes no diagrama polos-zeros.

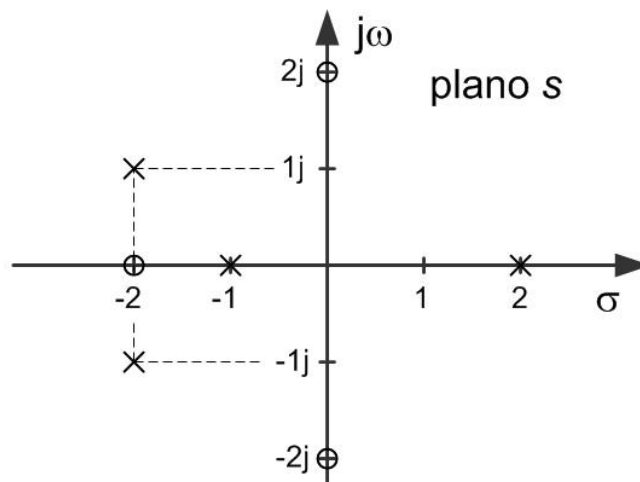


Figura 4.14 – Diagrama polos-zeros do Exemplo 4.4.



Note-se que a ordenação dada aos polos e zeros no exemplo anterior é completamente arbitrária. É de lembrar também que as raízes de polinómios com coeficientes reais são reais ou, quando complexas, em pares conjugados.

4.5. Estabilidade assintótica

A **estabilidade** é uma das características mais procuradas de um sistema dinâmico e é muito importante saber concluí-la a partir diagrama de polos-zeros. No capítulo de introdução referiu-se já, de forma breve, à estabilidade como a capacidade de um sistema voltar ao ponto de equilíbrio após sofrer uma perturbação. A estabilidade de um sistema linear é uma característica intrínseca do sistema e portanto não depende do sinal de excitação externo.

Um sistema diz-se estável no sentido assintótico se, dada a sua resposta impulsiva, existirem valores $\alpha, p > 0$, tal que, qualquer que seja o valor do tempo t

$$|h(t)| < \alpha e^{-pt}$$

Significa isto que a resposta impulsiva deve aproximar-se de zero (quando o tempo vai para infinito) mais depressa que uma exponencial decrescente com parâmetros positivos adequados (Figura 4.15). Uma consequência desta definição é que a resposta impulsiva de um sistema estável no sentido assintótico converge obrigatoriamente para zero.

Note-se que a estabilidade assintótica não é a única definição possível de estabilidade. Uma definição habitual alternativa é a estabilidade no sentido saída limitada para entrada limitada ou BIBO (de *Bounded Input Bounded Output*). Um sistema assintoticamente estável é estável no sentido BIBO, mas o contrário não se aplica. Por exemplo, um oscilador puro é estável no sentido BIBO, mas não converge para zero.

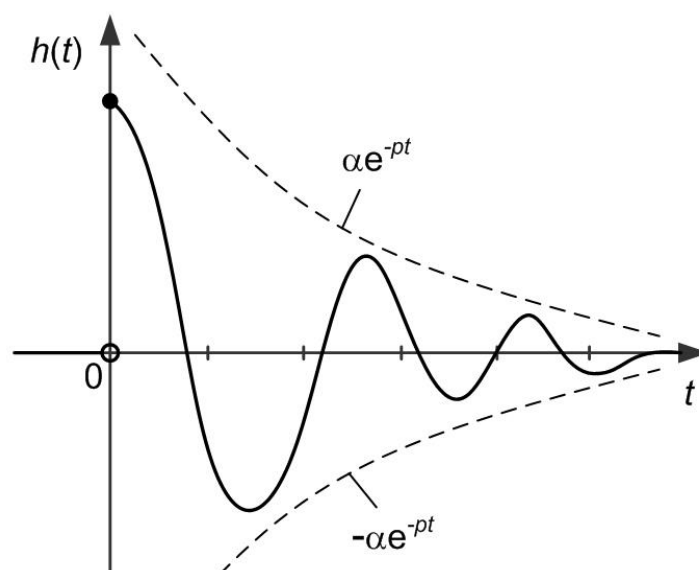


Figura 4.15 – Estabilidade assintótica.

Um sistema dinâmico representado pela sua função de transferência é estável no sentido assintótico se e só se todos os seus polos estão posicionados à esquerda do eixo imaginário, ou seja, se e só se

$$\text{Re}\{p_i\} < 0 \quad \text{para} \quad i = 1, \dots, n$$

A posição dos zeros não influencia a estabilidade do sistema, podendo estar localizados nos semiplanos complexos esquerdo ou direito.

Em oposição, um polo diz-se **instável** quando tem parte real positiva.

Exemplo 4.5

Considere-se de novo o sistema descrito pela função de transferência

$$G(s) = \frac{2s^3 + 4s^2 + 8s + 16}{s^4 + 3s^3 - s^2 - 13s - 10}$$

Viu-se já no Exemplo 4.4 que os polos da função de transferência são dados por

$$D(s) = s^4 + 3s^3 - s^2 - 13s - 10 = 0 \quad \left\{ \begin{array}{l} p_1 = +2.0 \\ p_2 = -2.0 + 1.0j \\ p_3 = -2.0 - 1.0j \\ p_4 = -1.0 \end{array} \right.$$

Como pelo menos um polo, p_1 , tem parte real positiva, o sistema não é estável no sentido assintótico (Figura 4.16).

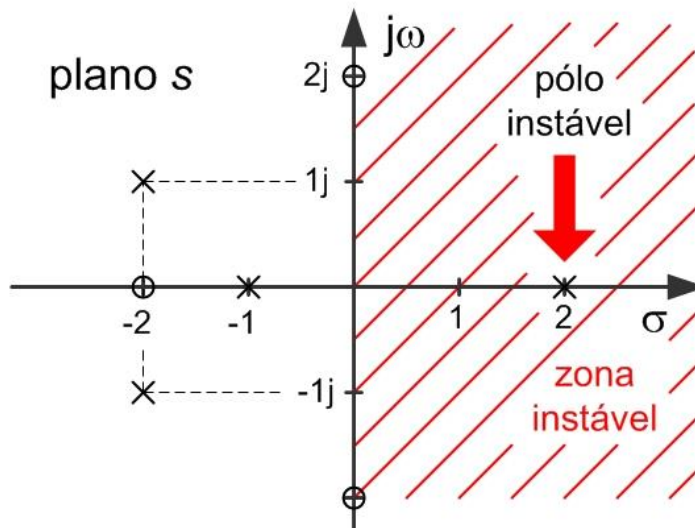


Figura 4.16 – Polo instável (Exemplo 4.5).

■

Exemplo 4.6

Considere-se o sistema descrito pela função de transferência

$$G(s) = \frac{5}{s^3 + 2s^2 + 4s + 8}$$

Neste caso os polos da função de transferência são dados por

$$D(s) = s^3 + 2s^2 + 4s + 8 = 0 \quad \left\{ \begin{array}{l} p_1 = -2.0 \\ p_2 = +2.0j \\ p_3 = -2.0j \end{array} \right.$$

Como os polos p_2 e p_3 têm parte real nula, o sistema não é estável no sentido assintótico. Nesta situação em que os polos estão sobre a fronteira da estabilidade (e não existe nenhum com parte real positiva) o sistema diz-se **marginalmente estável** (Figura 4.17).

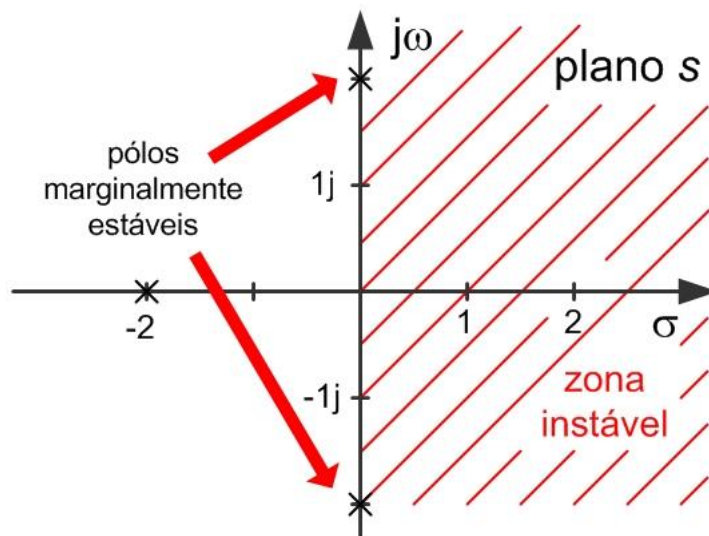


Figura 4.17 – Polos marginalmente estáveis (Exemplo 4.6).

■

O exemplo seguinte mostra, para um sistema de 1ª ordem sem zeros, o efeito da posição do único polo nos semiplanos complexos esquerdo ou direito nas respetivas respostas impulsivas.

Exemplo 4.7

Considere-se o sistema caracterizado pela equação diferencial

$$\frac{dy(t)}{dt} + ay(t) = u(t)$$

onde a é uma constante real. Aplicando a Transformação de Laplace à equação diferencial obtém-se a função de transferência

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

com um polo $p = -a$. A resposta impulsiva obtém-se aplicando a Transformação de Laplace inversa, resultando em

$$h(t) = \text{TL}^{-1} \left\{ \frac{1}{s+a} \times 1 \right\} = e^{-at} \cdot \varepsilon(t)$$

Se $a > 0$, o polo $p = -a$ terá parte real negativa, o sistema é estável e a resposta impulsiva converge para zero (Figura 4.18.a). Se $a < 0$, o polo $p = -a$ tem parte real positiva, o sistema é instável e a resposta impulsiva diverge (Figura 4.18.b). Na fronteira entre estes dois casos, com $a = 0$, o polo é nulo, o sistema é marginalmente estável e a resposta impulsiva, embora não divergindo, não converge para zero (Figura 4.18.c).

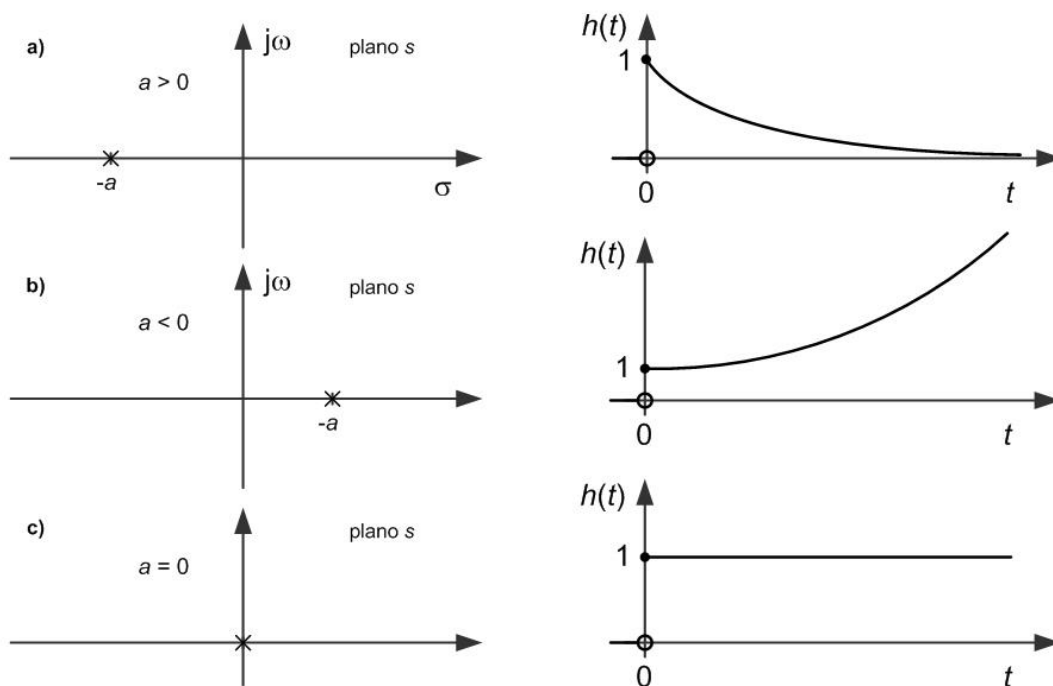


Figura 4.18 – Estabilidade de um sistema de 1ª ordem sem zeros (Exemplo 4.7).

■

4.6. Teorema do valor final

Um resultado de grande utilidade é o descrito pelo **teorema do valor final** com o enunciado seguinte. Seja $X(s)$ a Transformação de Laplace de um sinal $x(t)$, então o valor do limite do sinal quando o tempo vai para infinito é igual ao limite de $s \cdot X(s)$ quando s tende para zero, ou seja

$$\lim_{t \rightarrow \infty} x(t) = \lim_{s \rightarrow 0} s \cdot X(s)$$

Um corolário interessante deste resultado é que se o sinal em causa for a saída de um sistema descrito pela sua função de transferência $G(s)$, então o valor final da resposta a um degrau unitário é dado por

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} s \cdot Y(s) = \lim_{s \rightarrow 0} s \cdot G(s) \cdot \frac{1}{s} = \lim_{s \rightarrow 0} G(s)$$

sendo este valor designado por ganho estático do sistema.

4.7. Teorema do valor inicial

De um modo dual, enuncia-se o **teorema do valor inicial** do seguinte modo. Seja $X(s)$ a Transformação de Laplace de um sinal $x(t)$, então o limite do sinal no tempo quando o tempo se aproxima de zero (por valores à direita) é igual ao limite de $s \cdot X(s)$ quando s tende para infinito, ou seja

$$\lim_{t \rightarrow 0^+} x(t) = \lim_{s \rightarrow \infty} s \cdot X(s)$$

Novamente, se o sinal em causa for a saída de um sistema descrito pela sua função de transferência $G(s)$, então o valor inicial da resposta a um degrau unitário é dado por

$$\lim_{t \rightarrow 0} y(t) = \lim_{s \rightarrow \infty} s \cdot Y(s) = \lim_{s \rightarrow \infty} s \cdot G(s) \cdot \frac{1}{s} = \lim_{s \rightarrow \infty} G(s)$$

sendo este valor nulo sempre que o sistema for causal no sentido estrito ($n > m$). Se o sistema for causal no sentido lato ($n = m$), existe uma descontinuidade na origem na resposta ao degrau, e o valor inicial toma o valor do coeficiente b_0 da função de transferência.

Este teorema é ainda de grande utilidade para analisar a continuidade das primeiras derivadas na origem da resposta ao degrau. Assim, se $n - m = 1$, então a resposta ao degrau é contínua na origem, mas a derivada é descontínua apresentando um vértice. No caso em que $n - m \geq 2$, então a resposta ao degrau é contínua na origem, bem como a sua derivada, transitando suavemente a partir da condição inicial.

4.8. Resposta no tempo de sistemas de ordem reduzida

Compreendido o conceito de função de transferência, é possível agora retomar a questão da resposta no tempo com uma abordagem diferente. Do ponto de vista da resposta no tempo de um sistema estável, a dinâmica dominante é estabelecida pelo grupo de raízes (polos e zeros) mais próximos da origem (e à esquerda desta). Nesta secção, apresenta-se as respostas no tempo para um conjunto de sistemas básicos de primeira e segunda ordem que constituem na grande maioria dos casos a dinâmica dominante do sistema.

Considere-se, para começar, o sistema de 1ª ordem sem zeros com a função de transferência dada por

$$G(s) = \frac{Ka}{s + a} = \frac{K}{\tau s + 1}$$

Este sistema é caracterizado pelo ganho estático igual a K e um polo em (ver Figura 4.19)

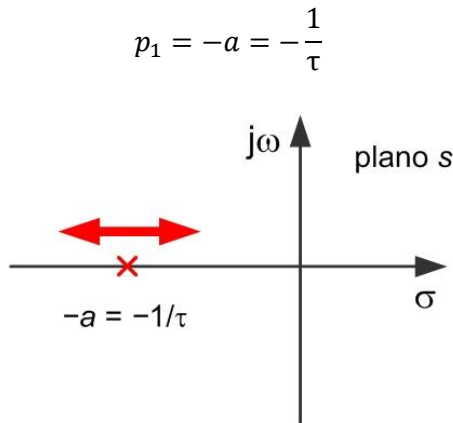


Figura 4.19 – Posicionamento do polo do sistema de 1ª ordem.

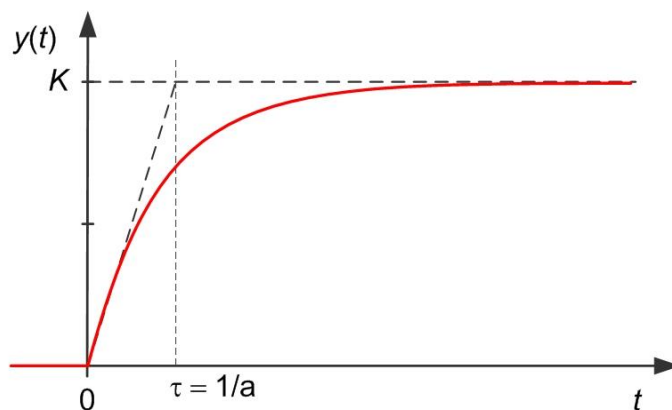


Figura 4.20 – Resposta ao degrau unitário do sistema de 1ª ordem.

Na Figura 4.20 representa-se a resposta no tempo deste sistema a um degrau unitário. A resposta será tanto mais rápida quanto maior for o polo em valor absoluto (mais afastado da origem para o lado esquerdo).

No caso particular em que o valor do polo é igual a zero, o sistema representado é o integrador puro e a função de transferência escreve-se

$$G(s) = \frac{K}{s}$$

O ganho estático é agora infinito significando que a resposta ao degrau unitário não estabilizará em nenhum valor finito, como se pode observar na Figura 4.21. Relembre-se que estando o polo da função de transferência do integrador sobre o eixo imaginário, o sistema já não é estável. No entanto, o integrador também não é instável e, como se verá no capítulo seguinte, é um bloco de grande utilidade no controlo pela sua capacidade de impor o seguimento sem erro permanente.

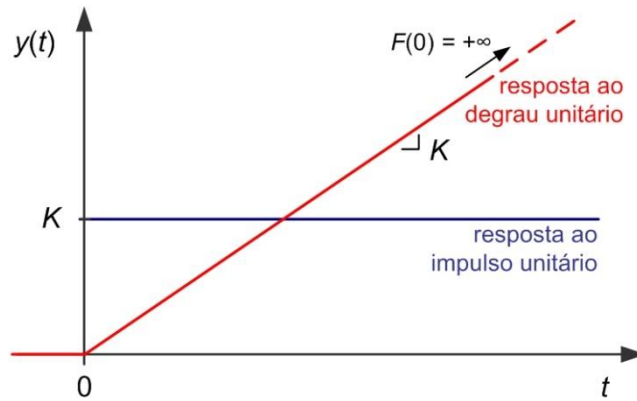


Figura 4.21 – Resposta ao degrau e ao impulso unitários do sistema integrador.

O sistema de 2ª ordem sem zeros é o bloco dinâmico mais importante por apresentar simultaneamente riqueza de comportamentos dinâmicos diferentes e coincidir frequentemente com a dinâmica dominante de sistemas de controlo em anel fechado.

Considere-se então a função de transferência do sistema de 2ª ordem sem zeros dada por

$$G(s) = \frac{b_0}{s^2 + a_1s + a_2} = \frac{K \cdot \omega_n^2}{s^2 + 2D\omega_n s + \omega_n^2}$$

A segunda forma de representação permite distinguir o efeito dos parâmetros da função de transferência no comportamento dinâmico. Definem-se assim as seguintes variáveis:

- Ganho estático K
- Frequência natural ω_n
- Fator de amortecimento D

Esta função de transferência (sem zeros) tem dois polos nas raízes do polinómio denominador, dados por (verifique como exercício)

$$p_{1,2} = -D\omega_n \pm \omega_n \sqrt{D^2 - 1}$$

Enquanto o fator de amortecimento estabelece a posição relativa dos dois polos, a frequência natural determina a distância à origem de ambos, e logo, irá ser determinante no estabelecimento da rapidez da resposta. Quanto maior for a frequência natural ω_n , mais rápida será a reacção do sistema a uma excitação na entrada.

Por outro lado, de acordo com o valor do fator de amortecimento D (positivo para ser estável), importa distinguir três casos possíveis

- Sistema sobre-amortecido quando o fator de amortecimento $D > 1$;
- Sistema criticamente amortecido quando o fator de amortecimento $D = 1$;

- Sistema sub-amortecido quando o fator de amortecimento $D < 1$.

Quando o fator de amortecimento é igual ou superior à unidade o sistema apresenta dois polos sobre o eixo real, e é equivalente a dois sistemas de 1º grau em série. A resposta no tempo ao degrau unitário não apresenta qualquer sobrelevação face ao valor final e o tempo de subida é imposto, essencialmente, pelo polo mais próximo da origem. No caso limite em que o fator de amortecimento é unitário, os dois polos ficam sobrepostos em $p_{1,2} = -\omega_n$ e temos a resposta mais rápida (com ω_n fixo) sem sobrelevação face ao valor final (ver Figura 4.22).

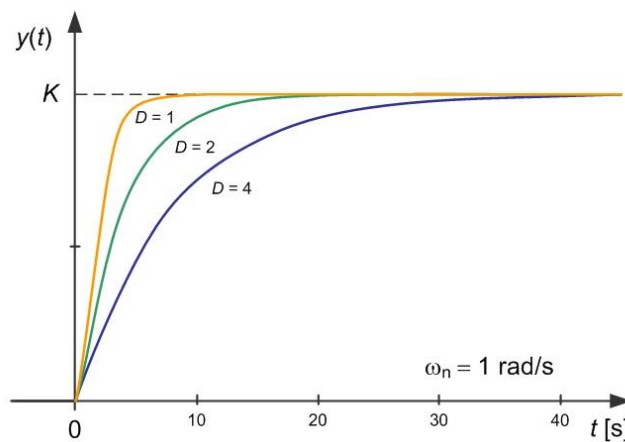


Figura 4.22 – Resposta ao degrau unitário do sistema de 2ª ordem ($D \geq 1$).

No entanto, o caso mais interessante do ponto de vista dinâmico é aquele que resulta de um fator de amortecimento entre zero e a unidade ($0 < D < 1$). Como $D^2 - 1 < 0$, as raízes serão complexas e dadas por

$$p_{1,2} = -D\omega_n \pm j\omega_n\sqrt{1 - D^2}$$

Como se observa na Figura 4.23, estes polos encontram-se a uma distância ω_n da origem e o factor de amortecimento D estabelece a relação entre as componentes real e imaginária dos polos. Quanto menor for o fator de amortecimento, maior é o peso da componente imaginária e logo menor é o ângulo ($\text{asin } D$) que os polos formam com o eixo imaginário. Antevê-se deste modo, que com a redução do fator de amortecimento, os polos aproximar-se-ão da zona instável (semiplano complexo direito).

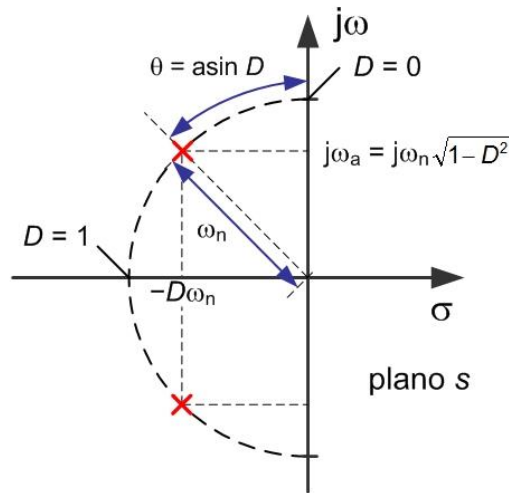


Figura 4.23 – Lugar dos polos em função do valor do fator de amortecimento.

Na Figura 4.24-a) mostra-se o lugar dos polos para diversos valores do fator de amortecimento e na Figura 4.24-b) regista-se a forma da resposta no tempo ao degrau unitário para os valores de D correspondentes. A expressão da resposta no tempo ao degrau unitário é dada por

$$y(t) = K \left\{ 1 - e^{-D\omega_n t} \left(\cos \omega_a t + \frac{D\omega_n}{\omega_a} \cdot \sin \omega_a t \right) \right\} \varepsilon(t)$$

onde $\omega_a = \omega_n \sqrt{1 - D^2}$. Ou seja, a resposta é dada pelo produto de: uma exponencial decrescente com o fator $(-D\omega_n)$, a parte real dos polos; por uma componente sinusoidal de frequência ω_a , a parte imaginária dos polos.

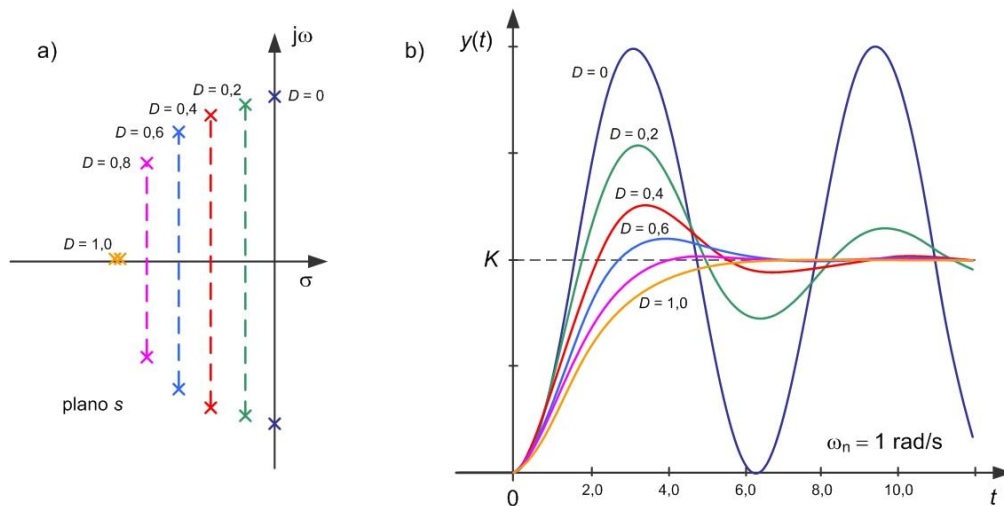


Figura 4.24 – a) Posição dos polos e b) resposta ao degrau unitário do sistema de 2ª ordem.

Quando D se aproxima da unidade a sobrelevação reduz-se, aproximando-se do caso limite da Figura 4.22. Na outra direção, à medida que o fator D se reduz o sistema é menos amortecido (mais oscilatório) apresentando valores de sobrelevação mais altos. No caso limite em que o fator de amortecimento se anula, os polos estão sobre o eixo imaginário em $p_{1,2} =$

$\pm j\omega_n$, o sistema deixa de ser assintoticamente estável e resulta num oscilador sinusoidal permanente.

Analisa-se de seguida o efeito na resposta no tempo da introdução de um zero na função de transferência dum sistema de 2ª ordem. Considere-se então a função de transferência do sistema de 2ª ordem com um zero dada por

$$G(s) = \frac{K \left(\frac{s}{\alpha D \omega_n} + 1 \right) \omega_n^2}{s^2 + 2D\omega_n s + \omega_n^2}$$

que, para além do lugar dos polos já analisados, tem um zero em

$$z_1 = -\alpha D \omega_n$$

No caso em que o fator α é positivo, o zero encontra-se no semiplano esquerdo (zero de fase mínima) como se representa na Figura 4.25-a). Na Figura 4.25-b) apresenta-se a resposta ao degrau unitário para diferentes valores de α ($D = 0.5$ e $\omega_n = 1.0$ rad/s). Para valores elevados de α o efeito do zero adicional não se faz sentir e a resposta é semelhante à do sistema sem zeros. Com a redução do fator, o zero aproxima-se da zona dos polos e observa-se um aumento de rapidez inicial da resposta bem como da sua sobrelevação.

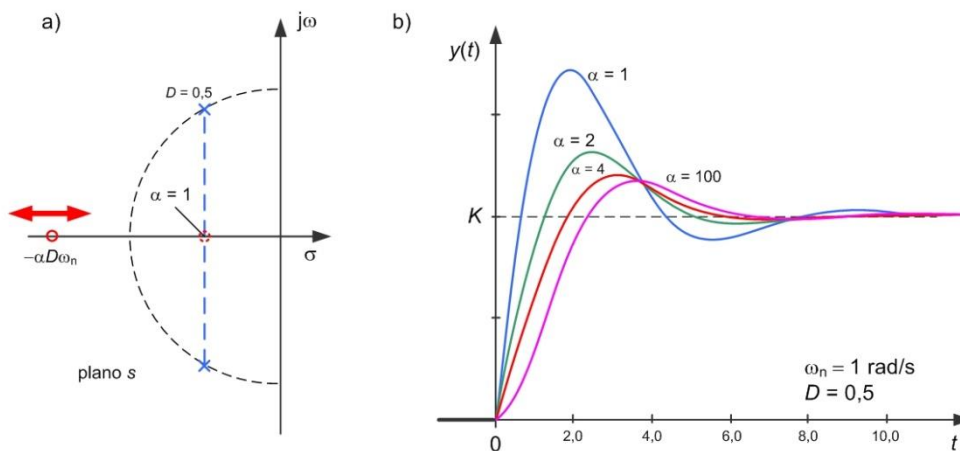


Figura 4.25 – a) Posição do zero de fase mínima e b) resposta ao degrau unitário.

No entanto, este efeito de sobrelevação tem uma origem diferente daquele já observado na função de transferência sem zeros e polos complexos. A sobrelevação originada pelos polos complexos surge de uma oscilação interna, independente da entrada como, por exemplo, um sistema pendular pouco amortecido que troca energia potencial e energia cinética. O efeito do zero efetua-se todo sobre o sinal de entrada. Recuando à equação diferencial que dá origem a esta função de transferência, teremos

$$\frac{d^2 y(t)}{dt^2} + 2D\omega_n \frac{dy(t)}{dt} + \omega_n^2 y(t) = \frac{K\omega_n}{\alpha D} \frac{du(t)}{dt} + K\omega_n^2 u(t)$$

Com a redução do fator α , aumenta-se o peso da derivada da entrada $u(t)$ na excitação do sistema, ou seja, no caso do degrau unitário $\varepsilon(t)$, o peso de um impulso unitário $\delta(t)$.

Se o fator α for negativo, o zero está localizado no semiplano complexo direito (zero de fase não mínima), a excitação impulsiva explicada no parágrafo anterior tem sinal contrário. A Figura 4.26-a) representa a localização do zero e na Figura 4.26-b) apresenta-se a resposta ao degrau unitário para diferentes valores de α negativo ($D = 0.5$ e $\omega_n = 1.0$ rad/s).

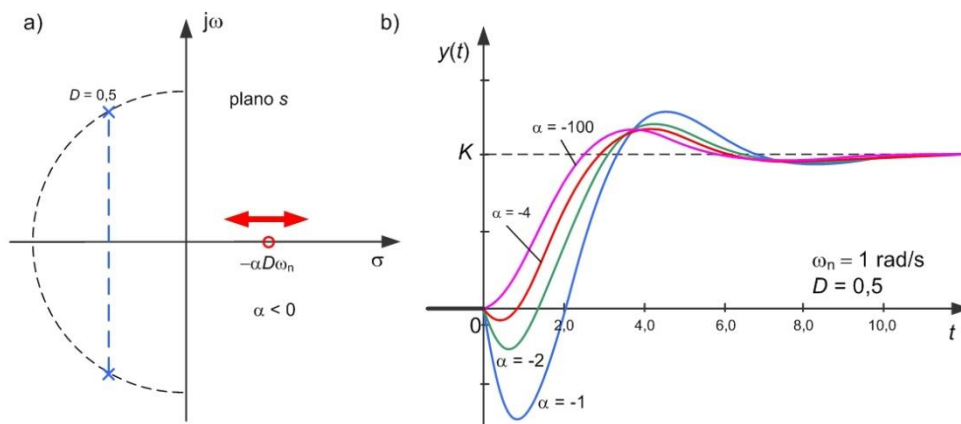


Figura 4.26 – a) Posição do zero de fase mínima e b) resposta ao degrau unitário.

Note-se que, com o afastamento do zero de fase não mínima para o lado direito ($\alpha = -100$), a resposta é semelhante à resposta do sistema de 2ª ordem sem zeros, sem se observar qualquer sublevação inicial.

Os zeros de fase não mínima, porque são zeros, não têm qualquer influência na estabilidade do sistema sob análise. No entanto, quando falarmos no controlo por retroação iremos observar que os sistemas de fase não mínima (ou seja, contendo zeros de fase no semiplano complexo direito) apresentam-se mais difíceis de controlar e com tendência para resultarem instáveis em anel fechado. Nesta fase, deixo à imaginação do leitor a dificuldade em controlar uma instalação à qual pedimos para se mover numa direção, e aquela, teimosamente, insiste em ir por uns momentos na direção contrária antes de obedecer ao comando. Um exemplo clássico deste tipo de sistemas é a reação na altitude de uma aeronave por atuação no leme de profundidade (Figura 4.27).

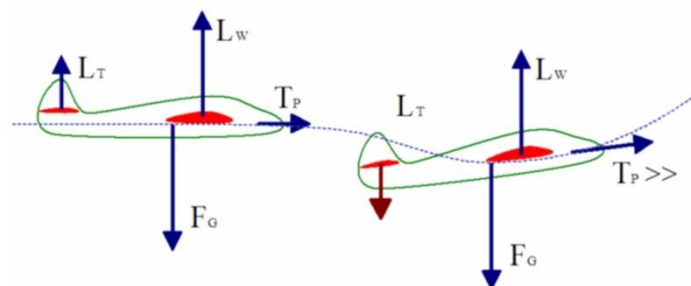


Figura 4.27 – Exemplo de um sistema de fase não mínima.

A atuação do leme de profundidade causa, no início da manobra de subida, uma perda de sustentação total da aeronave face ao seu peso. Assim, a aeronave começa por descer antes de, por reorientação do vetor velocidade, realizar a subida pretendida.

5. Controlo por retroação

Nos dois capítulos anteriores, fez-se a análise dos sistemas, primeiro a partir da equação diferencial e respetiva solução no tempo e, segundo, através da Transformação de Laplace e consequente função de transferência. Estes são dois passos básicos para compreender a engenharia associada à teoria dos sistemas e, em particular, do controlo.

Identificam-se dois posicionamentos possíveis na engenharia de sistemas. Podemos estar envolvidos no projeto e construção de uma instalação. Neste caso, temos como objetivo especificar a dinâmica pretendida para cumprir os objetivos da missão da instalação. Inicia-se pelo desenvolvimento de um modelo matemático de base física e participa-se na adequada seleção dos parâmetros físicos da instalação para cumprir a especificação.

Num posicionamento diferente, o engenheiro de sistemas é frequentemente chamado após o desenvolvimento da instalação. Neste caso, o envolvimento prende-se com o projeto e realização do sistema de controlo para a instalação. Inicia-se com a obtenção de um modelo da instalação (com componentes de base física e outras resultantes da identificação a partir de dados experimentais) e projeta-se o sistema de controlo para cumprir a especificação estabelecida pelo cliente. O projeto do sistema de controlo desdobra-se em dois passos de igual importância. A seleção da estrutura do controlador – o **tipo de controlador** - e a seleção dos **parâmetros do controlador** (vulgo ganhos).

Apesar desta visão alternativa, estas duas fases devem ser realizadas de forma integrada, se possível. A **engenharia do controlo** deve ser chamada a participar o mais cedo possível no dimensionamento da própria instalação, maximizando o desempenho global alcançável, mais tarde, com o projeto do sistema de controlo. O exemplo mais crítico desta aproximação acontece na indústria aeronáutica na qual o comportamento dinâmico é um dos primeiros objetivos a ser definido. Este objetivo é alcançado com o desenvolvimento integrado da aeronave e dos seus sistemas de controlo.

5.1. O sistema de controlo

Os sistemas de controlo dividem-se em dois grandes grupos:

- o **controlo por antecipação** – este tipo de controlo é efetuado em anel aberto (i.e. sem retroação) e baseia-se no conhecimento preciso dos efeitos que uma entrada de perturbação tem sobre a saída da instalação. Tendo acesso à medida da perturbação é possível injetar no atuador o sinal de compensação adequado para minimizar o efeito da perturbação na saída (ver Figura 5.1).

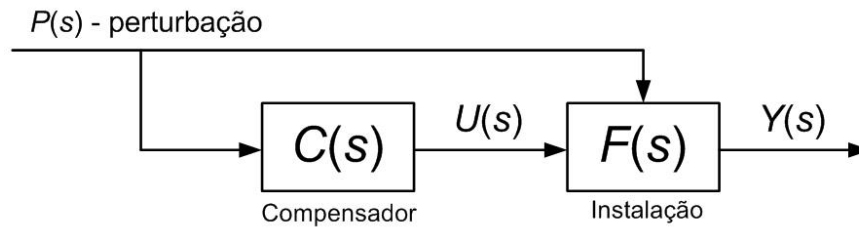


Figura 5.1 – Esquema de controlo por antecipação.

- o **controlo por retroação** – este tipo de controlo é efetuado em anel fechado (blocos em retroação) e rejeita os efeitos das perturbações através da medida dos seus efeitos na saída da instalação. O desvio da medida da saída face ao valor desejado é usado para calcular a ação de controlo sobre o atuador da instalação (ver Figura 5.2).

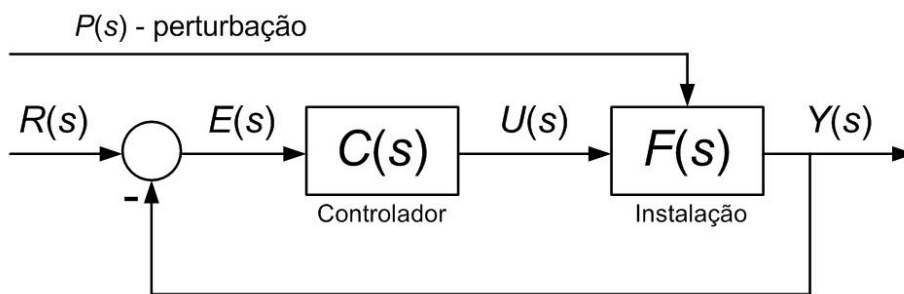


Figura 5.2 – Esquema de controlo por retroação.

Estas duas aproximações são complementares e compatíveis e os melhores resultados são conseguidos, frequentemente, pela melhor combinação dos dois tipos de controlo.

Os sistemas de controlo subdividem-se ainda de acordo com os objetivos, podemos ter

- **sistema de comando** – este tipo de controlo pode ser obtido com retroação ou apenas em antecipação. Neste caso, pretende-se apenas obter um comportamento dinâmico específico entre um sinal de comando (por exemplo, vindo de um *joystick* manual) e a saída da instalação.
- **sistema de seguimento** – este tipo de controlo é, quase sempre, alcançado por retroação, pretendendo-se que o erro entre a saída da instalação e o valor desejado para aquela, a entrada de referência, seja mínimo. O objetivo de um sistema de seguimento é fazer a saída da instalação y seguir a entrada de referência r .
- **sistema de regulação** – este tipo de controlo, também obtido por retroação da saída, designa sistemas em que o valor de referência permanece constante por longos períodos de tempo.

5.2. Especificação de um sistema de controlo

A forma mais simples (utilizada pela maioria dos clientes deste tipo de sistemas) de transmitir o resultado pretendido para o sistema de controlo é através do estabelecimento das especificações dinâmicas do sistema num conjunto de parâmetros associados à resposta temporal, por exemplo, na resposta a um sinal em degrau.

Na Figura 5.3 apresenta-se um conjunto habitual de parâmetros de especificação do comportamento dinâmico de um sistema de controlo através da sua resposta temporal. Estes parâmetros definem-se como

- Valor final da resposta y_∞ (a partir de um certo valor inicial y_0);
- Valor máximo da resposta y_{\max} que define a sobrelevação (*overshoot*) $s[\%]$;
- A sobrelevação $s[\%] = \frac{y_{\max} - y_\infty}{y_\infty - y_0} \times 100\%$
- Tempo de subida a 90% do valor final t_r (*rise time*);
- Tempo de estabelecimento a $\pm 5\%$ (ou a $\pm 1\%$) do valor final t_s (*settling time*);
- Tempo de pico t_p (*peak time*).

Por vezes, existe interesse em relacionar estes parâmetros de especificação com os parâmetros de uma função de transferência que sirva de modelo ao comportamento dinâmico do anel fechado. Se considerarmos a função de transferência de 2ª ordem sem zeros

$$H^*(s) = \frac{K \omega_n^2}{s^2 + 2D\omega_n s + \omega_n^2}$$

oferece-se, como modelo, um conjunto vasto de comportamentos dinâmicos disponíveis com a seleção adequada do ganho K , da frequência natural ω_n e do factor de amortecimento D .

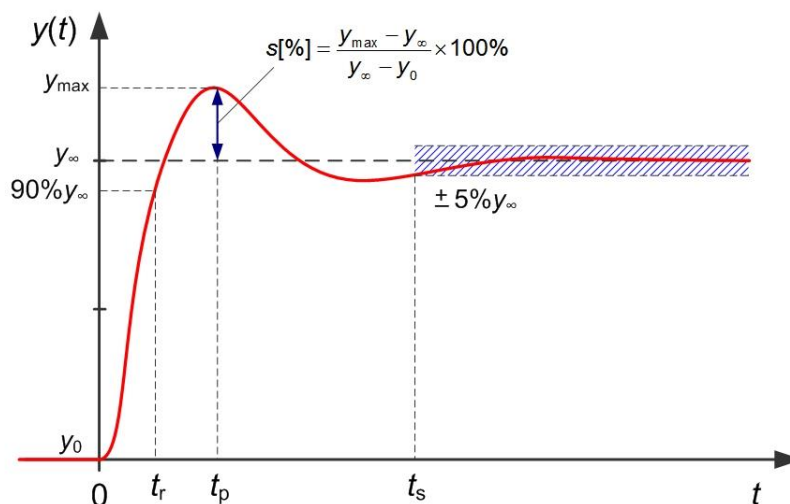


Figura 5.3 – Parâmetros de especificação no domínio do tempo (resposta ao degrau).

A lista seguinte apresenta um conjunto de fórmulas aproximadas que relacionam os parâmetros que caracterizam a resposta no tempo de um sistema de 2ª ordem com os seus coeficientes da função de transferência:

- Valor final da resposta ao degrau unitário

- $y_{\infty} = \lim_{s \rightarrow 0} H(s)$

- Tempo de estabelecimento a 1% e a 5%

$$t_{s[1\%]} \approx \frac{4.6}{D\omega_n} \quad \text{e} \quad t_{s[5\%]} \approx \frac{3}{D\omega_n}$$

- Tempo de subida a 90%

$$t_{r[90\%]} \approx \frac{3.7D}{\omega_n}$$

- Tempo de pico (fórmula exata)

$$t_p = \frac{\pi}{\omega_a} = \frac{\pi}{\omega_n \sqrt{1-D^2}}$$

- Sobrelevação

$$s[\%] = 100 e^{\frac{-D\pi}{\sqrt{1-D^2}}}$$

Note-se que esta função de transferência só tem três graus de liberdade (K , D e ω_n), sendo portanto necessário realizar algum compromisso no estabelecimento deste modelo.

5.3. Especificação no plano s

A especificação no domínio temporal, introduzida através de restrições nos parâmetros referidos na secção anterior, pode ser transportada para o domínio da variável s de Laplace. Considere-se, por exemplo, uma especificação envolvendo a sobrelevação, o tempo de estabelecimento e tempo de subida. A fórmula relativa à sobrelevação diz-nos que

$$s[\%] \leq S_{\max} \Rightarrow \text{asin } D \geq \text{asin } D_S$$

onde D_S é tal que

$$\frac{D_S}{\sqrt{1-D_S^2}} = -\frac{1}{\pi} \cdot \ln \frac{S_{\max}}{100}$$

Segundo, a fórmula (aproximada) para o tempo de estabelecimento resulta que

$$t_{s[5\%]} < T_s \Rightarrow \text{Re}\{p_i\} = -D_{S\max} \cdot \omega_n < -\frac{3}{T_s}$$

e, por último, a expressão para o cálculo aproximado do tempo de subida estabelece que

$$t_{r[90\%]} < T_r \Rightarrow |p_i| = \omega_n < -\frac{3.7D_s}{T_r}$$

A Figura 5.4 representa a sobreposição de cada uma destas restrições ao lugar dos polos do anel fechado p_i no plano s .

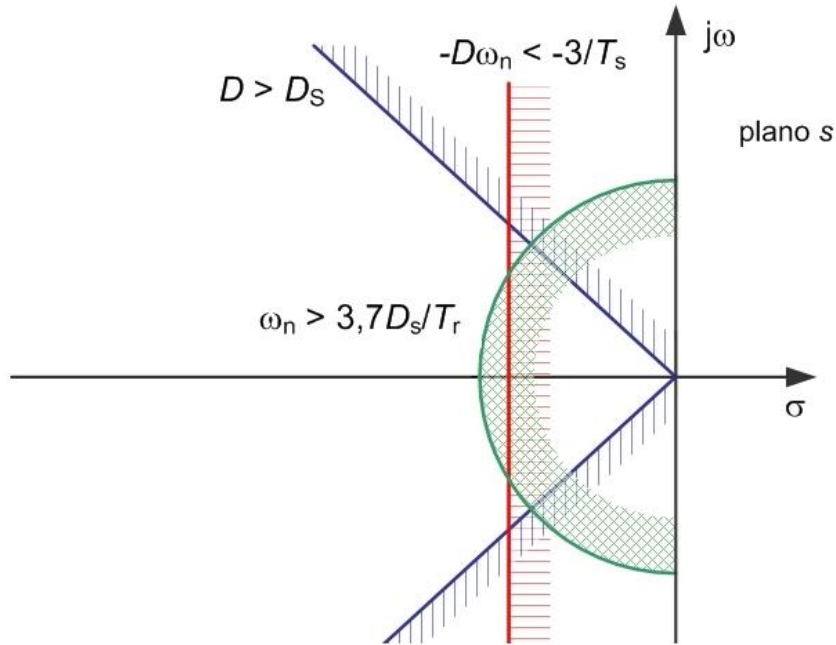


Figura 5.4 – Parâmetros de especificação no domínio da variável s .

5.4. Retroação proporcional da saída

Começemos por analisar a estrutura de controlo linear mais básica disponível, a retroação proporcional da saída. Referiu-se já no capítulo de introdução que a função básica do controlo consiste em atuar positivamente se a variável a regular tiver um valor aquém do valor de referência, e atuar negativamente se o valor estiver em excesso. Desejavelmente, queremos que a ação seja na direção certa, no sentido de contrariar a existência de erro, mas também que seja tanto maior quanto maior for este desvio.

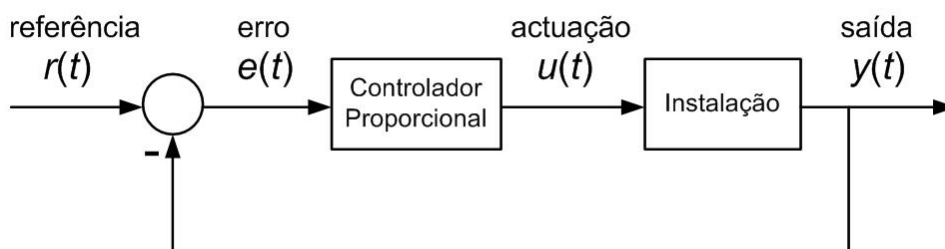


Figura 5.5 – Esquema tecnológico do sistema de controlo.

A retroação proporcional da saída atua sobre a instalação com um valor que é proporcional, através de um ganho K , ao erro entre a saída da instalação a controlar y e o seu valor

desejado, o sinal de referência r (ver Figura 5.5 e Figura 5.6). Ou seja, o erro de controlo é dado por

$$e(t) = r(t) - y(t)$$

e a atuação (também designada por ação de controlo) é calculada por

$$u(t) = K \cdot e(t)$$

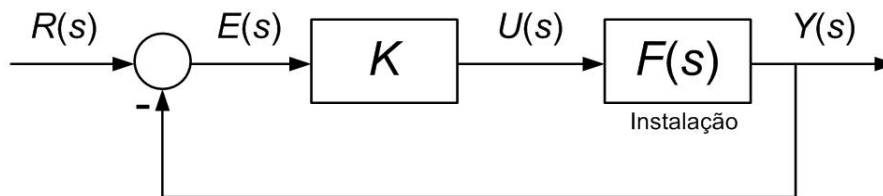


Figura 5.6 – Esquema de blocos do sistema de controlo.

De forma equivalente, as Transformações de Laplace destes sinais escrevem-se

$$E(s) = R(s) - Y(s)$$

e

$$U(s) = K \cdot E(s)$$

como representado no esquema de blocos na Figura 5.6.

Sabemos das regras da álgebra de blocos que o esquema de retroação da Figura 5.6 tem a função de transferência de $R(s)$ para $Y(s)$ dada por

$$H(s) = \frac{Y(s)}{R(s)} = \frac{K \cdot F(s)}{1 + K \cdot F(s)}$$

Esta função de transferência $H(s)$ é o objectivo final do projecto do sistema de controlo. Queremos, como resultado final, que a função de transferência do anel fechado $H(s)$, tanto quanto possível, cumpra a especificação definida previamente.

Exemplo 5.1

Considere-se de novo a instalação de dois tanques da Figura 5.7.

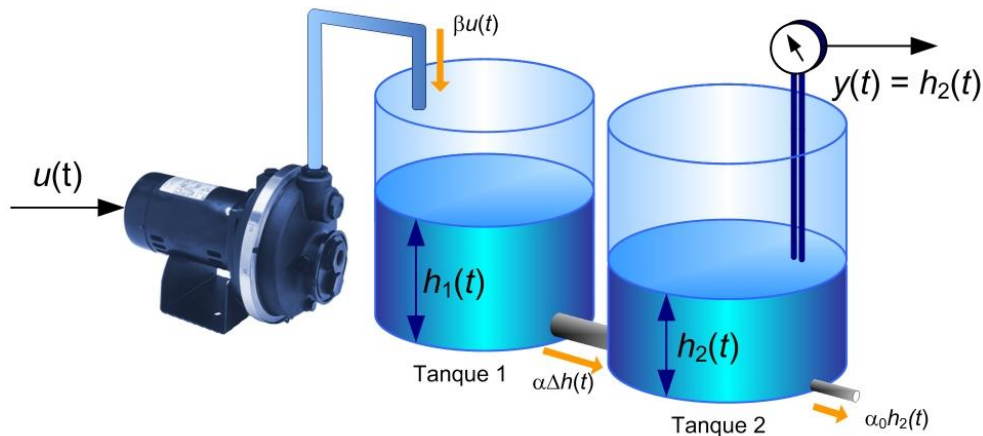


Figura 5.7 – Esquema da instalação do Exemplo 5.1.

O modelo linearizado desta instalação é dado pelo conjunto de equações que contabilizam a variação de volume de água em cada um dos tanques (ver Capítulo 2)

$$\frac{dh_1(t)}{dt} = -\frac{\alpha}{A}h_1(t) + \frac{\alpha}{A}h_2(t) + \frac{\beta}{A}u(t)$$

$$\frac{dh_2(t)}{dt} = \frac{\alpha}{A}h_1(t) - \frac{\alpha + \alpha_0}{A}h_2(t)$$

onde α , α_0 e β são constantes positivas; h_1 e h_2 são os níveis dos tanques 1 e 2 e a saída da instalação, $y(t) = h_2(t)$, é medida pelo sensor de nível.

Fica como exercício a demonstração que a função de transferência da entrada de atuação u para a saída y é dada por

$$F(s) = \frac{Y(s)}{U(s)} = \frac{\frac{\alpha \cdot \beta}{A^2}}{s^2 + \frac{2\alpha + \alpha_0}{A}s + \frac{\alpha \cdot \alpha_0}{A^2}}$$

Considere-se agora a instalação de uma forma mais concreta através da atribuição de valores numéricos aos parâmetros da instalação:

- Área dos tanques: $A = 1.0 \text{ m}^2$
- Abertura da união entre tanques: $\alpha = 2.5 \text{ m}^2\text{s}^{-1}$
- Abertura da saída do tanque 2: $\alpha_0 = 12.0 \text{ m}^2\text{s}^{-1}$
- Ganho comando-caudal da bomba: $\beta = 4.0 \text{ m}^3\text{s}^{-1}$

Neste caso, a função de transferência da instalação toma a forma

$$F(s) = \frac{Y(s)}{U(s)} = \frac{10}{s^2 + 17s + 30}$$

A função de transferência da instalação não tem zeros e tem polos em (Figura 5.8)

$$p_1 = -2$$

$$p_2 = -15$$

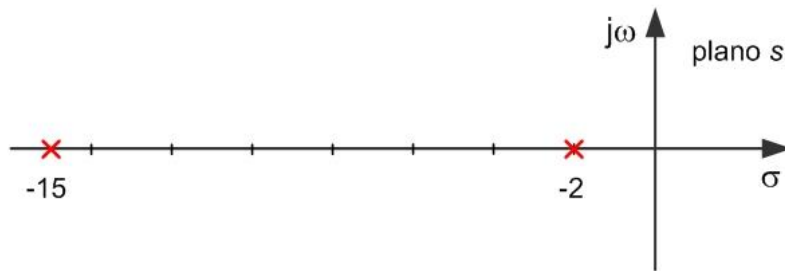


Figura 5.8 – Diagrama polos-zeros da instalação do Exemplo 5.1.

Surge então a questão: que dinâmicas estão disponíveis para o sistema de controlo (anel fechado da Figura 5.6) com retroação proporcional?

Calcule-se a função de transferência $H(s)$ do anel fechado (verifique como exercício) obtendo-se

$$H(s) = \frac{Y(s)}{R(s)} = \frac{K \cdot F(s)}{1 + K \cdot F(s)} = \frac{10K}{s^2 + 17s + (30 + 10K)}$$

sem zeros e com dois polos nas raízes de

$$D(s) = s^2 + 17s + (30 + 10K)$$

ou seja, em

$$p_{1,2} = -8.5 \pm \sqrt{42.25 - 10K}$$

Isto significa que a posição dos polos da função de transferência do anel fechado irá depender do valor **único** selecionado para o ganho K do controlador proporcional, oferecendo a possibilidade de comportamentos dinâmicos diversos. Continuando este exemplo, a tabela seguinte e a Figura 5.9 apresentam o valor e o lugar dos polos para um conjunto de ganhos positivos diferentes.

Tabela do lugar dos polos para diferentes valores do Exemplo 5.1

Ganho K	p_1	p_2
$K = 0$	-15.0	-2.0
$K = 2.0$	-13.2	-3.8
$K = 4.225$	-8.5	-8.5
$K = 8.0$	$-8.5 \pm j6.1$	
$K = 16.0$	$-8.5 \pm j10.9$	

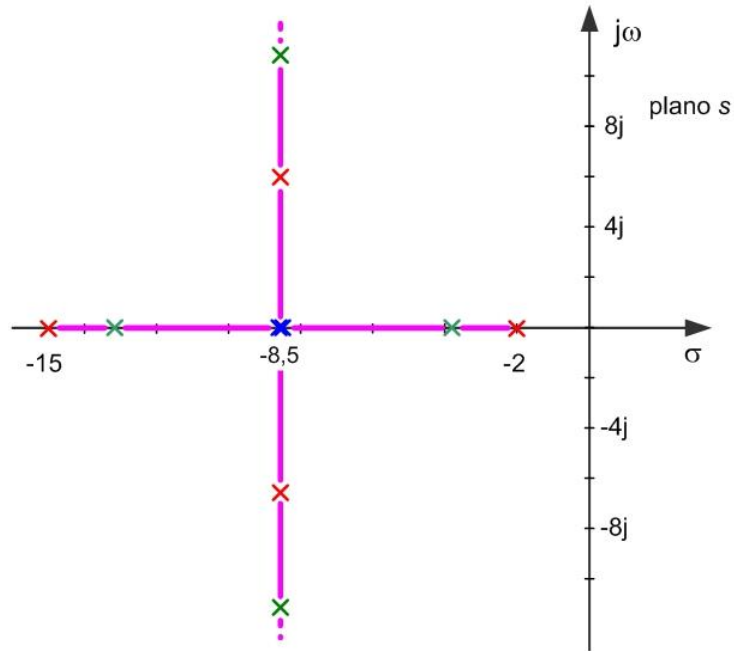


Figura 5.9 – Lugar de raízes do anel fechado para diferentes valores de K do Exemplo 5.1.

Viu-se já no capítulo anterior que o conhecimento da posição dos polos no plano s estabelece completamente o comportamento dinâmico do sistema sob análise. Neste caso, lembre-se que analisamos o sistema de controlo em anel fechado.

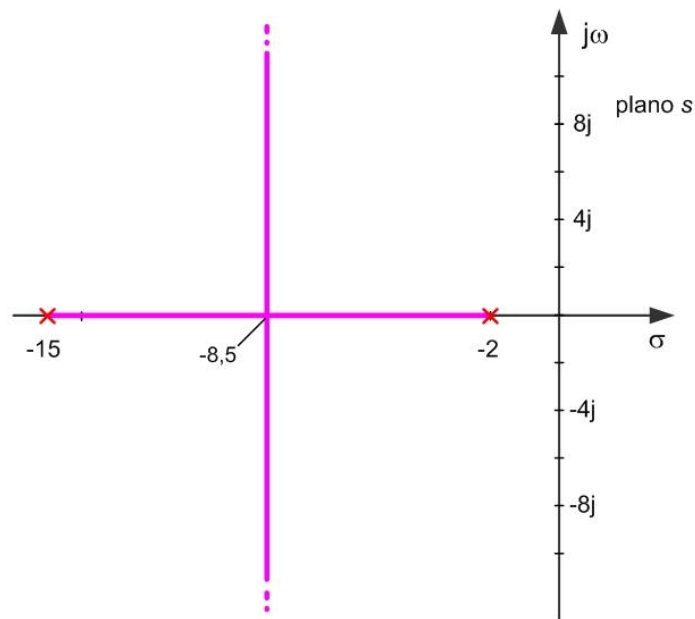


Figura 5.10 – Diagrama do lugar de raízes do Exemplo 5.1.

Se considerarmos uma variação contínua do ganho K positivo, o diagrama da Figura 5.9 pode ser substituído pela Figura 5.10. Fica conhecido assim, o conjunto de dinâmicas disponíveis para este sistema, nesta estrutura de controlo proporcional. Faltarà, a partir de uma

especificação para a dinâmica do sistema de controlo, seleccionar um valor específico para o ganho K .

Exemplo 5.2

Considere ainda a instalação do Exemplo 5.1 e introduza-se a seguinte especificação a cumprir pelo sistema em anel fechado com controlo proporcional:

$$s[\%] \leq 5\% \quad \text{e} \quad t_{s[5\%]} \leq 0.75 \text{ s}$$

Usando as fórmulas introduzidas acima, fica como exercício o cálculo demonstrando que as condições de especificação anterior são equivalentes a

$$D \geq 0.7 \quad \text{e} \quad D\omega_n \geq 4.0 \text{ rad/s}$$

respetivamente. Se sobrepuermos estas restrições ao diagrama representado na Figura 5.10, observa-se (ver Figura 5.11) que o **diagrama do lugar de raízes** fica restringido num ganho mínimo K_{\min} pela condição que derivou do tempo de estabelecimento e num K_{\max} pela condição que derivou da sobrelevação máxima. Note-se que a situação de sobrelevação igual a 5%, equivale aproximadamente a ter a parte imaginária dos polos igual à parte real. Uma vez que, para este exemplo, temos a expressão da localização dos polos do anel fechado, o cálculo dos valores dos ganhos máximo e mínimo pode ser feito a partir de

$$\sqrt{10K_{\max} - 42.25} = 8.5 \Rightarrow K_{\max} = 11.5$$

$$-8.5 + \sqrt{42.25 - 10K_{\min}} = -4.0 \Rightarrow K_{\min} = 2.2$$

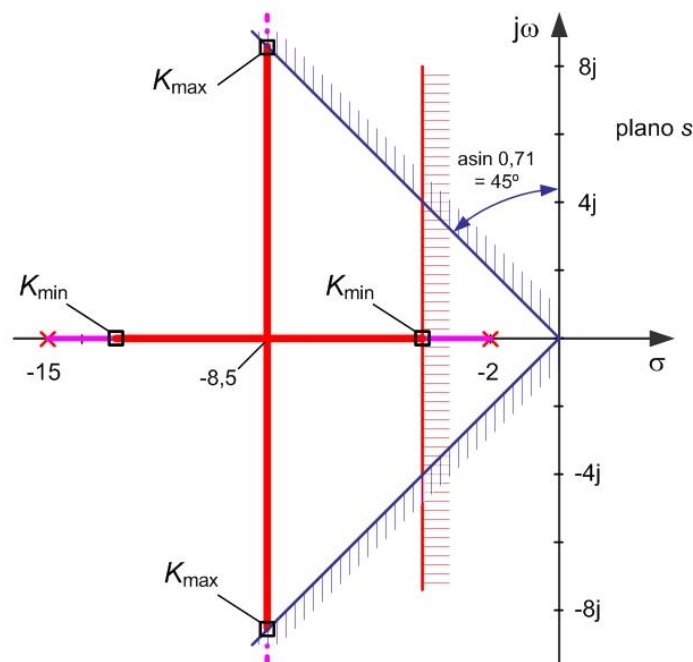


Figura 5.11 – Diagrama do lugar de raízes limitado pela especificação (Exemplo 5.2).

A resposta no tempo do sistema em anel fechado com os dois ganhos máximo e mínimo estão representadas na Figura 5.12. Note-se que o ganho máximo representa o cumprimento no limite da condição relativa à sobrelevação e o ganho mínimo representa o cumprimento no limite da condição relativa ao tempo de estabelecimento. As respostas no tempo para os ganhos compreendidos entre os dois extremos irão estar também compreendidas entre estas duas respostas limite. Portanto, todos os ganhos compreendidos no intervalo $K \in [K_{\min}; K_{\max}]$ cumprem a especificação.

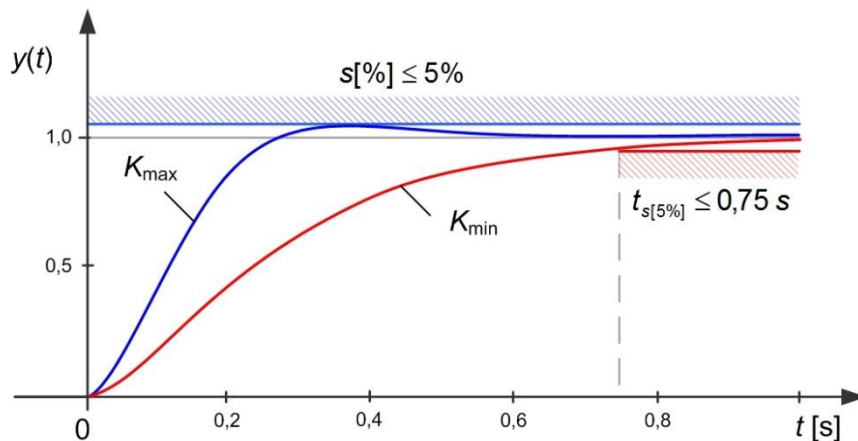


Figura 5.12 – Resposta no tempo dos sistema em anel fechado (Exemplo 5.2).

Observe-se ainda que a resposta no tempo relativa ao ganho máximo tem um tempo de subida mínimo. Mesmo que não tenha sido introduzida qualquer especificação relativa a este parâmetro, a resposta mais rápida entre as disponíveis seria a relativa ao ganho máximo K_{\max} e portanto essa seria a melhor seleção para o sistema de controlo.

■

5.5. Diagrama do lugar de raízes

Interessa agora generalizar o resultado do exemplo anterior introduzindo, formalmente, o **diagrama do lugar de raízes**.

O diagrama do lugar de raízes (ou *root-locus*) representa a localização geométrica das raízes (polos e zeros) do sistema em anel fechado no plano complexo s . O diagrama da Figura 5.10 relativa ao Exemplo 5.1 é um diagrama do lugar de raízes do respetivo sistema de controlo.

Atualmente, existem inúmeras ferramentas computacionais que nos permitem obter o traçado deste diagrama a partir da função de transferência do anel aberto. No entanto, é possível, através de um conjunto de regras simples, fazer o traçado do lugar de raízes diretamente numa folha de papel. Mais do que obter o diagrama perfeito (onde os computadores claramente vencem), é importante conhecer e saber aplicar estas regras para compreender o efeito da introdução de raízes adicionais no comportamento dinâmico final do anel fechado.

Assim, considere-se o anel aberto de um sistema de controlo dado pela função de transferência

$$G_0(s) = \frac{N_0(s)}{D_0(s)}$$

onde $N_0(s)$ e $D_0(s)$ são, respectivamente, os seus polinómios numerador e denominador. No Exemplo 5.1, o anel aberto resumia-se à função de transferência da instalação, no entanto, o anel aberto $G_0(s)$ pode ter uma componente dinâmica pertencente ao controlador. Nesse caso, o controlador não será apenas proporcional, mas incluindo polos e zeros.

Considere-se agora que este anel aberto está incluído numa estrutura de retroação com um ganho proporcional livre, como representado na Figura 5.6. A função de transferência do anel fechado será então dada por

$$H(s) = \frac{Y(s)}{R(s)} = \frac{K \cdot G_0(s)}{1 + K \cdot G_0(s)} = \frac{K \cdot \frac{N_0(s)}{D_0(s)}}{1 + K \cdot \frac{N_0(s)}{D_0(s)}} = \frac{K \cdot N_0(s)}{D_0(s) + K \cdot N_0(s)}$$

Os **zeros do anel fechado** são então dados pelas raízes solução de

$$K \cdot N_0(s) = 0$$

ou seja, são iguais aos zeros do anel aberto e os **polos do anel fechado** são dados pelas raízes solução de

$$D_0(s) + K \cdot N_0(s) = 0$$

ou seja, dependentes do valor do ganho K . Note-se que para $K = 0$, as raízes desta equação coincidem com os polos do anel aberto.

Fazendo um ponto da situação, conclui-se que os zeros permanecem imutáveis para diferentes valores do ganho K face ao seu valor para o anel aberto; mas os n polos mudam de lugar com a variação do ganho K a partir do lugar dos polos em anel aberto. Resta saber de que forma varia então o lugar dos polos para cada valor possível do ganho K que desenham os ramos do diagrama do lugar de raízes (*root-locus*)?

Como a função de transferência do anel fechado tem um denominador com coeficientes reais, este tem raízes reais ou complexas conjugadas e logo, o *root-locus* é simétrico em torno do eixo real.

Peguemos na equação que deve ser verificada pelos n polos r_i do anel fechado

$$D_0(r_i) + K \cdot N_0(r_i) = 0$$

que é equivalente a

$$K \cdot \frac{N_0(r_i)}{D_0(r_i)} = -1$$

Porque estamos no domínio dos números complexos, esta equação dá origem a duas condições:

- a condição de módulo

$$\left| \frac{N_0(r_i)}{D_0(r_i)} \right| = \frac{1}{K}$$

- e a condição de ângulo

$$\arg \left\{ \frac{N_0(r_i)}{D_0(r_i)} \right\} = \pm(2l + 1)\pi$$

Estas duas condições serão a base do conjunto de regras que permitem o traçado aproximado do diagrama de lugar de raízes.

Pegando na condição de módulo e levando o ganho K para infinito, ou seja

$$\left| \frac{N_0(r_i)}{D_0(r_i)} \right| = \frac{1}{K} \rightarrow 0$$

conclui-se que, ou

$$N_0(r_i) = 0$$

significando que, à medida que o ganho K aumenta, as soluções para os polos do anel fechado aproximar-se-ão ou dos zeros de $G_0(s)$ ou

$$D_0(r_i) \rightarrow \infty$$

significando que, à medida que o ganho K aumenta, as soluções para os polos do anel fechado aproximar-se-ão do infinito. Como o numerador $N_0(s)$ tem m raízes,

- m dos n ramos irão tender para os m zeros do anel aberto; e
- os restantes $(n - m)$ ramos irão afastar-se para infinito (em módulo).

Pegemos agora na condição de ângulo, que estabelece que qualquer lugar r_i pertencente ao *root-locus* está sujeito à restrição do somatório dos ângulos (argumentos) dos diversos fatores do numerador e do denominador

$$\sum_{j=1}^m \arg\{r_i - z_j\} - \sum_{j=1}^n \arg\{r_i - p_j\} = \pm(2l + 1)\pi$$

ou seja, tem de ser um múltiplo ímpar de 180° . A primeira conclusão é que se o lugar r_i , sob análise, pertencer ao eixo real, só pertencerão ao *root-locus* os troços do eixo real que tiverem

à sua direita um número ímpar de raízes (polos ou zeros que contribuem com $\pm 180^\circ$) do anel aberto. O número de raízes à esquerda (que contribuem com 0°) é irrelevante.

Façamos um exemplo com o que sabemos até este ponto.

Exemplo 5.3

Considere-se a função de transferência do anel aberto

$$G_0(s) = \frac{s^2 + 4.5s + 5.0}{s^4 + 11.5s^3 + 43.5s^2 + 54.0s}$$

com as raízes seguintes:

- zeros: $z_1 = -2.0;$ $z_2 = -2.5;$
- polos: $p_1 = 0.0;$ $p_2 = -3.0;$ $p_3 = -4.0;$ $p_4 = -4.5;$

Colocam-se então os polos e zeros no diagrama e marcam-se os troços do eixo real que têm à sua direita um número ímpar de raízes (ver Figura 5.13). Estes troços pertencem ao *root-locus*. Na mesma figura observa-se ainda que há 2 polos ligados por ramos a 2 zeros e outros dois polos ligados entre si à esquerda do plano. Viu-se já que estes dois polos terão que ir para infinito quando o ganho $K \rightarrow \infty$. Como não é possível que o façam pelo eixo real, porque os troços para infinito não pertencem ao traçado do *root-locus*, estes polos terão que sair do eixo real e desenhar ramos em direção a infinito. Esta situação foi observada já no Exemplo 5.1.

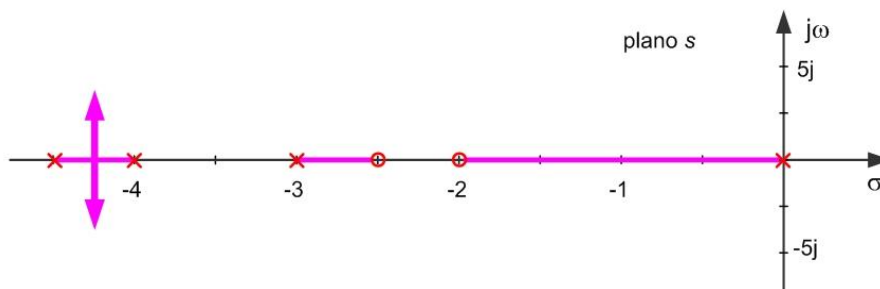


Figura 5.13 – Diagrama parcial do lugar de raízes do Exemplo 5.3.

■

Na sequência deste exemplo, interessa saber como os ramos se dirigem para infinito na situação habitual de existirem mais polos do que zeros (i.e. quando $n - m > 0$).

Deixa-se como exercício (mais difícil) a demonstração das duas regras seguintes:

- os $(n - m)$ polos do anel fechado dirigem-se para o infinito, quando $K \rightarrow \infty$, aproximando-se de $(n - m)$ assíntotas que fazem com o eixo real os ângulos

$$\phi_j = \frac{180^\circ}{n-m} (2j + 1) \quad \text{para } j = 0, \dots, n - m - 1$$

- estas assíntotas definem um centro assintótico em

$$\sigma_r = \frac{\sum_{j=1}^n p_j - \sum_{j=1}^m z_j}{n-m}$$

Exemplo 5.4

Repetindo o Exemplo 5.3 é possível dizer agora algo mais sobre o que acontece aos dois polos que sobram. Assim, sabemos que vão em direção a infinito com 2 assíntotas que fazem com o eixo real os ângulos

$$\phi_0 = \frac{180^\circ}{4-2} (2 \times 0 + 1) = +90^\circ; \quad \phi_1 = \frac{180^\circ}{4-2} (2 \times 1 + 1) = 270^\circ = -90^\circ$$

centradas em

$$\sigma_r = \frac{\sum_{j=1}^n p_j - \sum_{j=1}^m z_j}{n-m} = \frac{\{0 - 3.0 - 4.0 - 4.5\} - \{-2.0 - 2.5\}}{4-2} = -3.5$$

Na Figura 5.14 observa-se o traçado exato do *root-locus* da função de transferência do Exemplo 5.3 com a marcação das assíntotas verticais em -3.5 . Note-se que estas assíntotas não pertencem ao *root-locus*, sendo apenas linhas de construção para apoiar o traçado do diagrama.

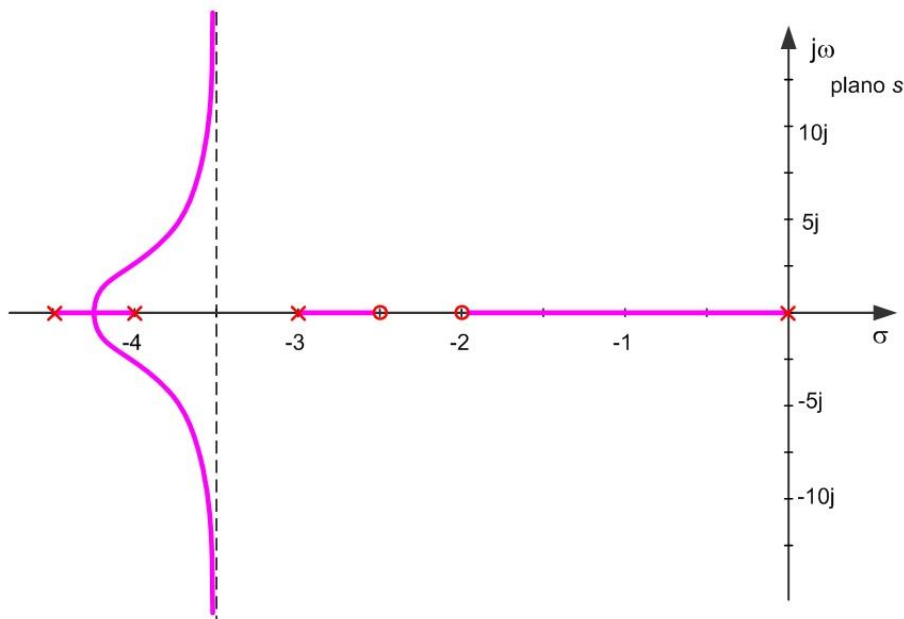


Figura 5.14 – Diagrama completo do lugar de raízes do Exemplo 5.3.

■

Note-se que é possível também, através da fórmula dos ângulos das assíntotas, preencher a tabela seguinte para os seis primeiros casos.

Tabela de ângulos das assíntotas do root-locus

Padrão	$n - m$	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4
+	0	sem assíntotas				
←+	1	+180°				
↑↓	2	+90°	+270°			
↖↗	3	+60°	+180°	+300°		
↘↙	4	+45°	+135°	+225°	+315°	
↖↗↘↙	5	+36°	+108°	+180°	+252°	+324°

Como se observa, as assíntotas dividem sempre a circunferência de 360° em fatias iguais. Quando a diferença do número de polos e zeros ($n - m$) é ímpar, existe uma assíntota em direção a $-\infty$ pelo eixo real. Quando esta diferença é par, a simetria do *root-locus* só deixa disponível uma solução. Assim, é relativamente fácil determinar mesmo sem cálculos o valor dos ângulos das assíntotas. No entanto, o centro assintótico é importante e tem de ser calculado para os casos com $(n - m) \geq 2$.

Observou-se no Exemplo 5.4 que os ramos do *root-locus* saem do eixo real e tendem assintoticamente para as duas assíntotas verticais bem identificadas pelo centro σ_r . No entanto, o traçado exato mostrou que os polos não saem do eixo real sobre a assíntota, mas num ponto diferente. A regra seguinte introduz uma forma de determinar os pontos de entrada e saída dos ramos no eixo real.

A condição a que estão sujeitos os polos do anel fechado

$$D_0(s) + K \cdot N_0(s) = 0$$

pode ser reescrita como

$$K = -\frac{D_0(s)}{N_0(s)}$$

significando que se pode definir uma função $k(x)$ real de variável real

$$k(x) = -\frac{D_0(x)}{N_0(x)}$$

Para cada saída $k(x) = K$ (interior ao subdomínio da função $k(x)$) a função irá ter como solução as raízes reais correspondentes aos polos do anel fechado sobre o eixo real para o valor K respectivo. Ou seja, os valores de K que resultam em polos complexos conjugados

não pertencem ao contradomínio da função $k(x)$. Como $k(x)$ é uma função contínua no seu domínio $\mathcal{R} \setminus \{z_i\}$, os pontos de entrada e saída dos ramos no eixo real irão coincidir com os mínimos e máximos relativos da função, respetivamente.

Assim, para determinar os pontos de entrada e saída no eixo real, define-se a função $k(x)$, e determinam-se os pontos onde a sua derivada se anula, ou seja

$$-\frac{d}{dx} \left(\frac{D_0(x)}{N_0(x)} \right) = 0$$

A determinação se o ponto é máximo ou mínimo pode ser realizada calculando a segunda derivada de $k(x)$. No entanto, se o objectivo é determinar se um ramo do *root-locus* está a entrar ou a sair do eixo real, a disposição dos polos e zeros do anel aberto é normalmente suficiente para o determinar. Em geral, entre dois zeros teremos pontos de entrada e entre dois polos teremos pontos de saída.

Uma regra adicional (sem demonstração aqui) estabelece que, quando diversos ramos (sempre em número par) chegam e partem de um ponto, fazem-no deixando ângulos iguais entre si. Assim, quando dois polos se encontram no eixo real saem para o domínio complexo fazendo ângulos de $\pm 90^\circ$. É apenas um detalhe de traçado.

Exemplo 5.5

Considere-se então, mais uma vez, o sistema do Exemplo 5.3 e defina-se a função $k(x)$ como

$$k(x) = -\frac{D_0(x)}{N_0(x)} = -\frac{x^4 + 11.5x^3 + 43.5x^2 + 54.0x}{x^2 + 4.5x + 5.0}$$

A determinação dos pontos de entrada e saída realiza-se agora derivando $k(x)$ e igualando a zero. O cálculo da derivada fica como exercício para o leitor, mostrando que os pontos de entrada e saída coincidem com as soluções de

$$2x^5 + 25x^4 + 123.5x^3 + 314.25x^2 + 435x + 270 = 0$$

que tem raízes: $r_1 = -2.31$; $r_2 = -3.25$; $r_3 = -4.25$; $r_{4,5} = -1.34 \pm j1.56$

A solução que procuramos é aquela que se encontra entre -4.5 e -4.0 ; logo r_3 , como se pode comprovar na Figura 5.14. Substituindo o valor de r_3 na função $k(x)$ obtém-se

$$k(r_3) = -\frac{D_0(-4.25)}{N_0(-4.25)} = 0.0843 > 0$$

sendo este o valor do ganho K correspondente ao ponto de saída do eixo real. A Figura 5.15 apresenta a forma da função $k(x)$ no intervalo $-4.5 \leq x \leq -3.0$. Note-se que os pontos -4.5 e -4.0 correspondem aos polos do anel aberto que anulam a função $k(x)$.

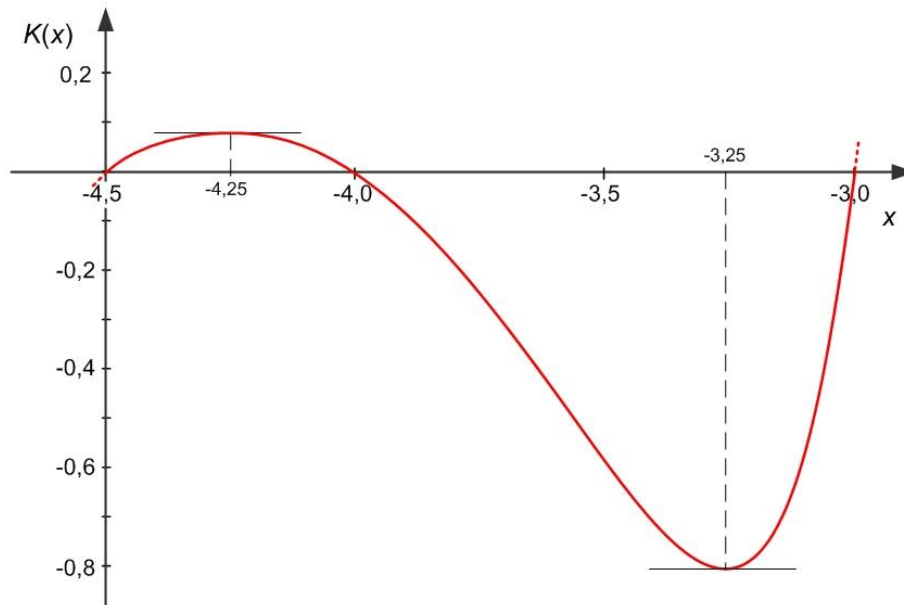


Figura 5.15 – Função $k(x)$ no intervalo $-4.5 \leq x \leq -3.0$ (Exemplo 5.5).

O diagrama do lugar de raízes foi traçado para $K > 0$. No entanto, pode fazer-se a mesma análise para ganho negativo. As soluções, também reais, r_1 e r_2 deste exemplo correspondem a pontos de entrada e de saída do *root-locus* para ganho negativo, nomeadamente, $K = k(r_1) = -100 < 0$ e $K = k(r_2) = -0.813 < 0$.

■

Existem ainda outras regras relativas aos ângulos que as raízes complexas fazem com os ramos de chegada/partida que derivam da condição de ângulo e que foram, antes do desenvolvimento das ferramentas computacionais, de grande utilidade. Atualmente, esse nível de detalhe no traçado deve ser obtido através dessas ferramentas.

5.6. O problema do seguimento da referência

Um esquema de **seguimento** tem por objetivo fazer a saída y da instalação seguir a entrada r de referência através da injeção do sinal de erro (desvio da saída ao valor de referência) no controlador.

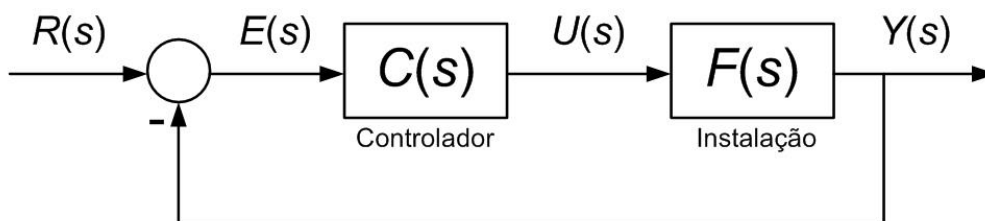


Figura 5.16 – Esquema de controlo de seguimento por retroação.

Note-se que o sinal de referência (também usualmente designado por *set-point*) pode ser gerado por um operador do sistema (e.g. uma pessoa numa habitação estabelece no sistema de ar condicionado o valor da temperatura desejada) ou por outros sistemas automáticos (e.g. um sistema integrado de gestão de energia de um edifício impõe valores de referência de temperatura para cada divisão aos diversos controladores dos aparelhos de climatização locais).

A dinâmica do seguimento do sinal de referência depende, como já se viu, do valor dos polos e zeros da função de transferência do anel fechado. Igualmente importante é o valor do erro quando o valor de referência é constante e diferente de zero. Este valor pode ser calculado através do Teorema do Valor Final (ver capítulo anterior) aplicado à resposta do sistema em anel fechado, ou seja

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} s \cdot E(s) = \lim_{s \rightarrow 0} s \cdot \left[R(s) - \frac{F(s)C(s)}{1 + F(s)C(s)} R(s) \right] = \lim_{s \rightarrow 0} \frac{s}{1 + F(s)C(s)} R(s)$$

Para analisar a capacidade de seguimento em regime forçado a um sinal de referência constante, analisa-se o valor final da resposta a um degrau de amplitude A

$$R(s) = \frac{A}{s}$$

resultando para o erro em regime estacionário

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} \frac{s}{1 + F(s)C(s)} \times \frac{A}{s} = \lim_{s \rightarrow 0} \frac{A}{1 + F(s)C(s)}$$

Para haver seguimento, o erro permanente deve ser o mais reduzido que for possível. Isto implica que o ganho estático do conjunto controlador e instalação, $F(s)C(s)$, deve muito maior que a unidade.

O exemplo seguinte ilustra as limitações do controlo proporcional na resolução do problema de seguimento.

Exemplo 5.6

Considere-se então o caso de um controlador de retroação proporcional $C(s) = K$ na controlo da instalação dada pela função de transferência

$$F(s) = \frac{1}{s^2 + 2s + 4}$$

Para o caso do controlo proporcional, o erro de seguimento a um sinal em degrau de amplitude A é dado por

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} \frac{A}{1 + KF(s)} = \frac{A}{1 + K \cdot \lim_{s \rightarrow 0} F(s)}$$

ou seja, dado um valor finito de ganho estático da instalação erro de seguimento será tanto menor quanto maior for o ganho de retroação.

Como neste exemplo o ganho estático da instalação vale

$$\lim_{s \rightarrow 0} F(s) = \frac{1}{4}$$

o erro permanente, em termos relativos, é dado por

$$\lim_{t \rightarrow \infty} \frac{e(t)}{A} = \frac{1}{1 + K \cdot \frac{1}{4}} = \frac{4}{4 + K}$$

A tabela seguinte mostra o valor do erro permanente em função de alguns valores do ganho K , válidos para este exemplo (ver Figura 5.17).

Tabela de ganhos do Exemplo 5.6.

K	$\lim_{t \rightarrow \infty} \frac{e(t)}{A}$
1	80%
4	50%
10	29%
100	4%

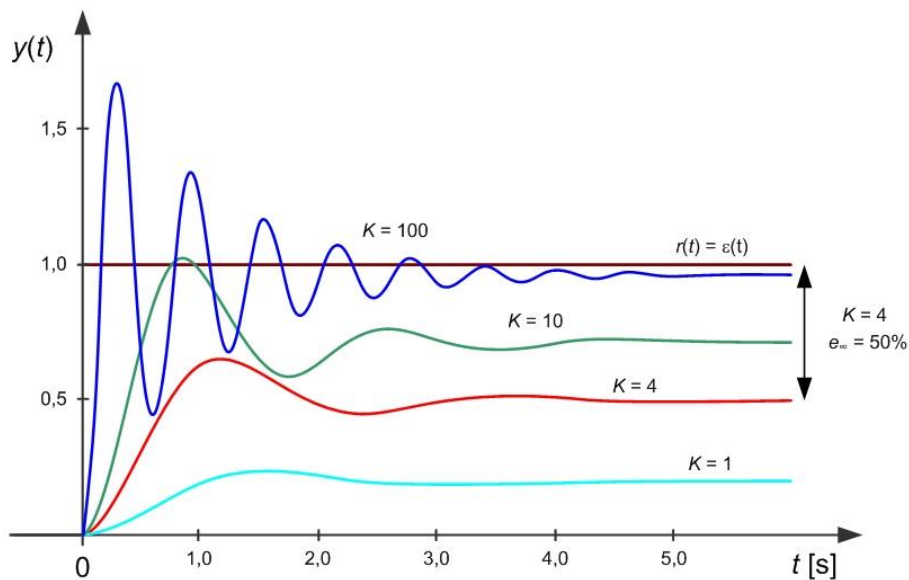


Figura 5.17 – Resposta ao degrau para diferentes valores do ganho K no Exemplo 5.6.

■

Note-se que o objetivo de levar o erro final para zero é equivalente a obter um ganho estático do sistema em anel fechado unitário, i.e.

$$\frac{\lim_{t \rightarrow \infty} y(t)}{\lim_{t \rightarrow \infty} r(t)} = \lim_{s \rightarrow 0} \frac{F(s)C(s)}{1 + F(s)C(s)} = 1$$

Esta condição exige que, de modo a levar o erro de seguimento para zero, o ganho estático do conjunto instalação/controlador seja infinito, ou seja

$$\lim_{s \rightarrow 0} F(s)C(s) = \infty$$

Assim, a condição para obter erro de seguimento nulo significa que o anel aberto do sistema tem de ter, pelo menos, um polo em $s = 0$ (efeito integral). Note-se que, a presença de um bloco integrador no anel aberto do sistema significa que, se o sistema for estável, o único ponto de equilíbrio possível equivale ao erro nulo.

No entanto, como se verifica em muitos aspetos da engenharia, a introdução de um benefício acarreta um custo e a anulação do erro permanente em sistemas de seguimento não é exceção. A introdução do efeito integral normalmente degrada a estabilidade do sistema em anel fechado.

Em geral, classifica-se um sistema quanto ao tipo consoante o número de pólos na origem, ou seja, um sistema de tipo 2 apresenta 2 pólos na origem. A tabela seguinte apresenta o erro em regime permanente para sistemas de controlo com retroacção unitária para referências em degrau, rampa ou parábola; para sistemas do tipo 0, tipo 1 e tipo 2.

Tabela de erros em regime permanente.

Referência	$r(t)$	$\lim_{t \rightarrow \infty} e(t)$		
		Tipo 0	Tipo 1	Tipo 2
Degrau	$\varepsilon(t)$	$\frac{1}{1 + F(0)C(0)}$	0	0
Rampa	$t \cdot \varepsilon(t)$	∞	$\frac{1}{F(0)C(0)}$	0
Parábola	$\frac{t^2}{2} \cdot \varepsilon(t)$	∞	∞	$\frac{1}{F(0)C(0)}$

Fica como exercício a sua verificação.

6. Resposta em frequência

Quando a complexidade das instalações é muito elevada, a obtenção da função de transferência pela via da modelação física do processo pode transformar-se numa tarefa hercúlea. Por exemplo, o engenheiro de controlo, deparando-se com uma grande instalação envolvendo processos termodinâmicos complexos (e.g. uma central de produção de energia elétrica a partir de combustíveis fósseis), não terá a possibilidade de modelar, usando apenas leis físicas e em tempo útil, todo o comportamento dinâmico importante para a realização do sistema de controlo (Figura 6.1). Nesta situação, as leis físicas de conservação (massa, energia, momento) são de grande utilidade, porque invioláveis, mas apenas explicam o que acontece em termos médios (i.e. nas variações lentas). O restante comportamento dinâmico da instalação face a variações rápidas, necessita de outras formas de identificação pela via de ensaios experimentais de recolha de sinais do sistema real.

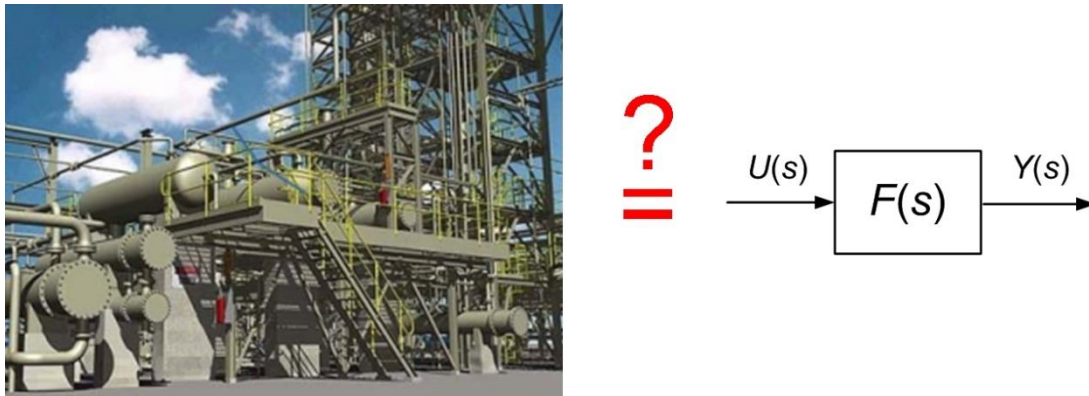


Figura 6.1 – Função de transferência equivalente a uma instalação real.

Os métodos de resposta em frequência constituem uma alternativa aos métodos apresentados no capítulo anterior, bastante populares em aplicações na indústria de processos devido aos bons resultados atingidos, na realização dos sistemas de controlo, mesmo em presença de incerteza no modelo da instalação. Outra vantagem está relacionada com a facilidade em obter, de forma experimental, modelos não paramétricos da instalação, ou seja, sem necessitar obter explicitamente a função de transferência do processo, mas suficientemente exatos para cumprir as especificações de controlo.

6.1. Sinal periódico sinusoidal

Antes de avançarmos para o controlo, importa introduzir (ou relembrar) alguns conceitos fundamentais da teoria dos sinais.

Um sinal $x(t)$ diz-se periódico sinusoidal de amplitude A , frequência ω e (desvio de) fase ϕ , se a sua variação no tempo contínuo t for representada pela função (Figura 6.2)

$$x(t) = A \sin(\omega t + \phi)$$

A frequência angular ω diz-nos quantos radianos são percorridos pelo argumento da função seno, por unidade de tempo, em segundos. É, portanto, medida em radianos por segundo, ou rad/s. Como a função seno é periódica com 2π (equivalente a 360°), define-se o período de tempo T como

$$\omega \cdot T = 2\pi$$

ou seja

$$\omega = \frac{2\pi}{T} = 2\pi f$$

onde f é a frequência em número de ciclos por unidade de tempo ou Hertz (Hz).

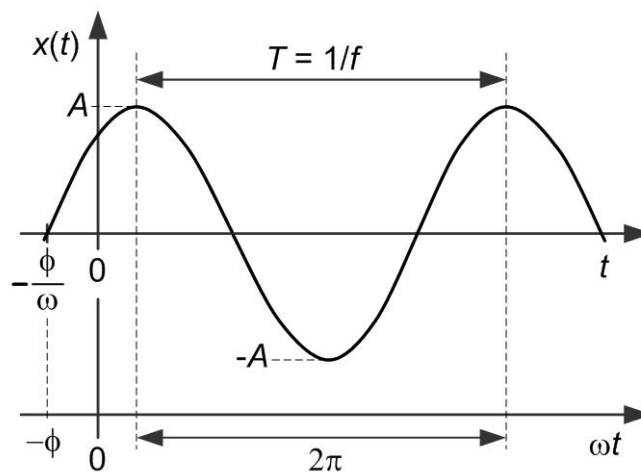


Figura 6.2 – Sinal sinusoidal de frequência angular ω .

A fase ϕ define o valor do argumento da função seno quando t é igual a zero e a amplitude A escala o valor máximo do sinal, visto que a saída da função seno varia no intervalo $[-1; +1]$.

Consultando a tabela das Transformações de Laplace obtém-se para a Transformação de Laplace do sinal $x(t)$

$$X(s) = A \frac{[\sin \phi]s + [\cos \phi]\omega}{s^2 + \omega^2}$$

6.2. Resposta no tempo a um sinal sinusoidal

Considere-se agora um sistema $G(s)$ excitado com um sinal sinusoidal $u(t)$ de frequência ω e amplitude U_0 . Se os n polos p_i de $G(s)$ forem todos estáveis e diferentes entre si, a saída

$$Y(s) = G(s)U(s)$$

admite a expansão em frações simples

$$Y(s) = \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} + \dots + \frac{r_n}{s - p_n} + \frac{r_0}{s + j\omega} + \frac{r_0^*}{s - j\omega}$$

onde r_i são os resíduos e r_0^* é o conjugado de r_0 . Se aplicarmos a Transformação de Laplace inversa a esta expressão obtém-se

$$y(t) = r_1 e^{p_1 t} + r_2 e^{p_2 t} + \dots + r_n e^{p_n t} + 2|r_0| \sin(\omega t + \phi)$$

Na Figura 6.3, representando a resposta $y(t)$, podem identificar-se duas componentes sobrepostas. A primeira, constituída pelos n termos exponenciais, é transitória porque se extingue à velocidade do polo mais lento. Passado este transitório resta o regime permanente dado por

$$y_p(t) = 2|r_0| \sin(\omega t + \phi)$$

com Transformação de Laplace

$$Y_p(s) = \frac{r_0}{s + j\omega} + \frac{r_0^*}{s - j\omega} = \frac{r_0(s - j\omega) + r_0^*(s + j\omega)}{s^2 + \omega^2} = \frac{(r_0 + r_0^*)s + (jr_0 + jr_0^*)\omega}{s^2 + \omega^2} = 2 \frac{\text{Re}\{r_0\}s + \text{Im}\{r_0\}\omega}{s^2 + \omega^2}$$

onde $\text{Re}\{r_0\}$ e $\text{Im}\{r_0\}$ representam as componentes real e imaginária de r_0 .

Resta saber quanto valerão as quantidades $|r_0|$ e ϕ .

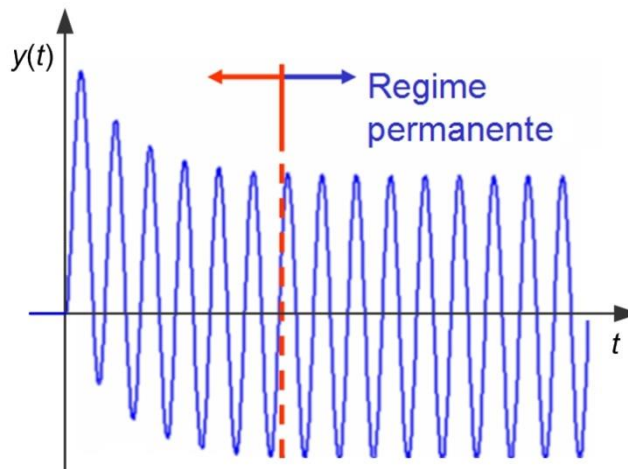


Figura 6.3 – Resposta a excitação sinusoidal em regime permanente.

O cálculo dos resíduos r_0 e r_0^* complexos conjugados, é feito a partir das expressões

$$r_0 = \lim_{s \rightarrow -j\omega} (s + j\omega)Y(s) = -\frac{U_0}{2j} \lim_{s \rightarrow -j\omega} G(s) = -\frac{U_0}{2j} G(-j\omega)$$

$$r_0^* = \lim_{s \rightarrow +j\omega} (s - j\omega)Y(s) = +\frac{U_0}{2j} \lim_{s \rightarrow +j\omega} G(s) = +\frac{U_0}{2j} G(+j\omega)$$

que substituído na expressão de $Y_p(s)$ resulta (demonstre como exercício)

$$Y_p(s) = U_0 \cdot \frac{\text{Im}\{G(j\omega)\} \cdot s + \text{Re}\{G(j\omega)\} \cdot \omega}{s^2 + \omega^2}$$

com Transformação de Laplace inversa

$$y_p(t) = U_0 \cdot |G(j\omega)| \cdot \sin(\omega t + \arg\{G(j\omega)\})$$

Assim, conclui-se que se um sistema linear invariante no tempo, estável, descrito pela sua função de transferência $G(s)$ for excitado por um sinal sinusoidal

$$u(t) = U_0 \sin(\omega t) \varepsilon(t)$$

responde, após a extinção de um transitório, com um regime permanente

$$y_p(t) = U_0 \cdot M_G(\omega) \sin(\omega t + \phi_G(\omega)) \varepsilon(t)$$

também sinusoidal de frequência ω igual mas com um ganho de amplitude dado por

$$M_G(\omega) = |G(j\omega)|$$

e desvio de fase dado por

$$\phi_G(\omega) = \arg\{G(j\omega)\}$$

A quantidade complexa $G(j\omega)$ é designada por resposta em frequência.

Um primeiro comentário ao que foi exposto é que a resposta em frequência de uma instalação pode ser obtida por via experimental. Excitando a entrada do sistema sinusoidalmente, fazemos variar a frequência ω por patamares, e regista-se o ganho de amplitude e o desvio de fase (por exemplo, com um osciloscópio) após a passagem do transitório, para cada valor de ω seleccionado.

Sendo uma função complexa, é possível representar $G(j\omega)$ graficamente de diversas formas, tais como:

- Representação direta de $G(j\omega)$ num plano complexo para todos os valores possíveis da frequência ω . Esta representação designa-se por diagrama polar e é a base para o traçado do **diagrama de Nyquist** a estudar mais à frente.
- Representação num diagrama cartesiano de $M_G(\omega)$ e $\phi_G(\omega)$ para todos os valores possíveis da frequência ω . Esta representação, que não será analisada neste texto, designa-se por diagrama de Nichols e permite-nos tirar conclusões semelhantes àquelas que tiraremos com o diagrama de Nyquist.
- Representação em separado das quantidades $M_G(\omega)$ e $\phi_G(\omega)$ em função da frequência ω . Esta representação é designada por **diagrama de Bode**. Porque a variável ω aparece explicitamente estes diagramas contém mais informação que as duas representações referidas antes.

O exemplo seguinte introduz estas representações diagramáticas possíveis, entre outras.

Exemplo 6.1

Considere-se a instalação de 2ª ordem com a função de transferência

$$F(s) = \frac{2}{(s + 1)^2}$$

A resposta em frequência da instalação é obtida substituindo a variável s por $j\omega$, ou seja

$$F(j\omega) = \frac{2}{(j\omega + 1)^2} = \frac{2}{(1 - \omega^2) + 2j\omega}$$

com ganho

$$M_F(\omega) = |F(j\omega)| = \frac{2}{\sqrt{(1 - \omega^2)^2 + 4\omega^2}}$$

e fase

$$\phi_F(\omega) = \arg\{F(j\omega)\} = -\operatorname{atan}\left(\frac{2\omega}{1 - \omega^2}\right)$$

O estudo destas duas funções da frequência angular ω pode fazer-se pela análise assintótica e pela substituição de certos valores notáveis. Iniciemos esta análise, pelo limite na baixa frequência ($\omega \rightarrow 0$), observando que

$$M_F(0) = 2$$

$$\phi_F(0) = 0^\circ$$

e na alta frequência ($\omega \rightarrow \infty$), através dos limites

$$\lim_{\omega \rightarrow \infty} M_F(\omega) = 0$$

$$\lim_{\omega \rightarrow \infty} \phi_F(\omega) = 180^\circ$$

Um ponto notável ocorre quando $\omega = 1$ rad/s porque resulta para a fase o arco cuja tangente é infinito, ou seja,

$$\phi_F(1) = -\operatorname{atan}(\infty) = -90^\circ$$

onde

$$M_F(1) = 1$$

Na Figura 6.4 observa-se o registo da resposta em frequência no plano complexo.

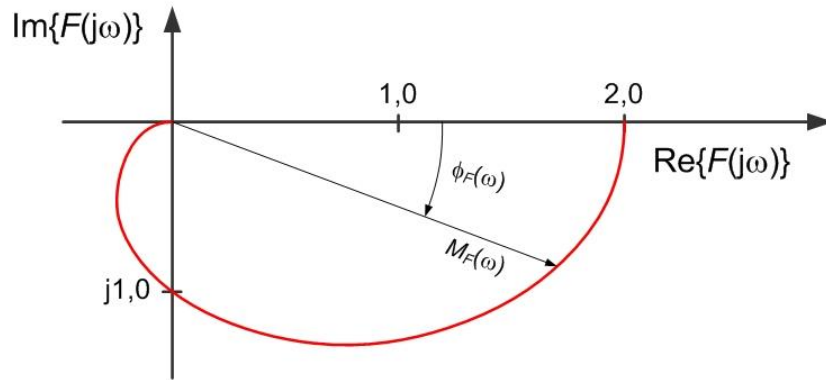


Figura 6.4 – Diagrama polar da resposta em frequência do Exemplo 6.1.

6.3. Critério de Nyquist

Até este ponto no capítulo, introduziu-se a resposta em frequência e foram discutidas possíveis representações diagramáticas. Nesta secção, aborda-se a influência da resposta em frequência no comportamento dinâmico do sistema sob análise. No entanto, diga-se já que, à semelhança do que se fez no diagrama do lugar de raízes, iremos usar a informação do sistema em anel aberto para retirar conclusões sobre o sistema em anel fechado, ou seja, o produto final do controlo por retroação.

Para fazer a análise dinâmica a partir da resposta em frequência, recordemos antes um teorema da teoria dos números complexos que estabelece o seguinte.

Considere-se um contorno fechado σ no plano complexo e uma função complexa de variável complexa $F(s)$ com um número Z de zeros e um número P de polos no interior do contorno σ . Então, o número N de envoltimentos (no mesmo sentido do contorno) da imagem desse contorno através da função F em redor da origem é igual à diferença entre o número de zeros e o número polos, ou seja, $N = Z - P$.

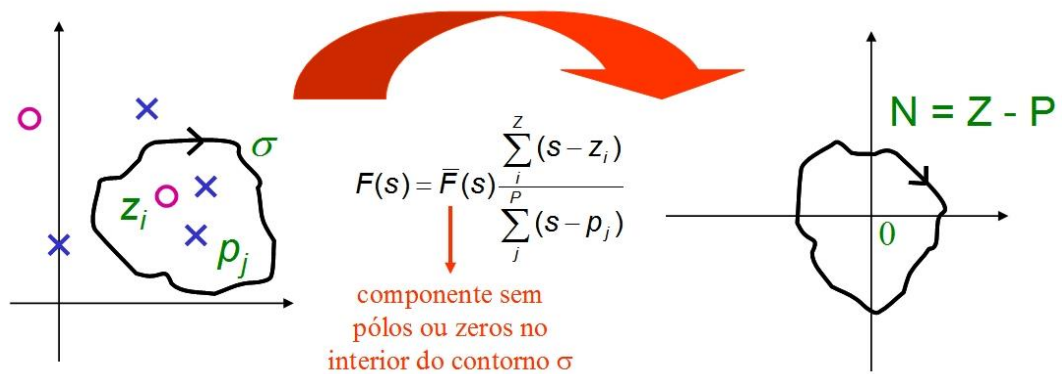


Figura 6.5 – Teorema dos Envoltimentos.

Exemplo 6.2

Considere-se a função complexa de variável complexa dada por

$$F(s) = s - (a + jb)$$

e duas circulações σ_1 e σ_2 que, respetivamente, envolvem e não envolvem a raiz da função $F(s)$, em $(a + jb)$, como representado no diagrama à esquerda da Figura 6.6. Como a função $F(s)$ é também uma função de translação de s para $s - (a + jb)$, o ponto $(a + jb)$ é trasladado para a origem e o contorno σ_1 , que envolvia a raiz, terá uma imagem através de $F(s)$, que envolverá a origem uma vez ($N = 1$) no mesmo sentido (ver Figura 6.6). Pelo mesmo motivo, a imagem do contorno σ_2 não envolverá a origem.

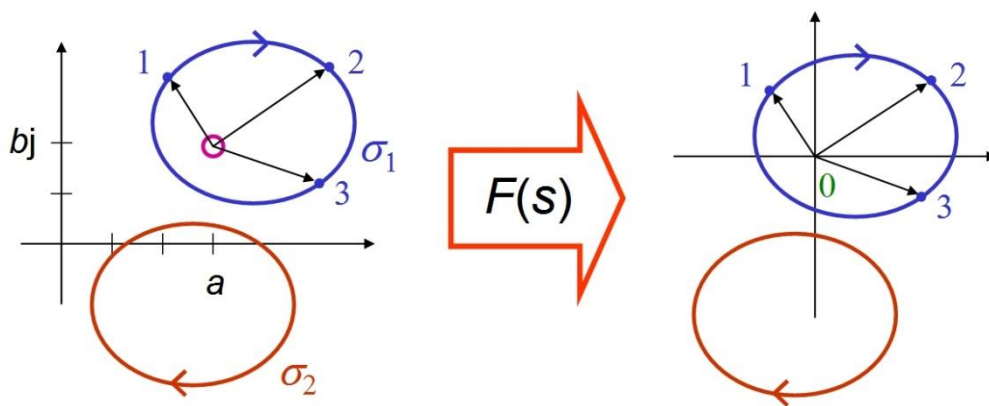


Figura 6.6 – Envolvimentos em redor da origem da imagem dos contornos σ_i por $F(s)$.

Note-se que se fosse considerado a função inversa

$$F(s) = \frac{1}{s - (a + jb)}$$

a imagem iria ser o inverso das obtidas no caso anterior, ou seja, o inverso da amplitude e o simétrico da fase. A imagem do contorno σ_1 iria envolver a origem uma vez no sentido inverso ($N = -1$) e a imagem do contorno σ_2 continuaria a não envolver a origem.

■

Apliquemos agora o teorema dos envoltivos à teoria do controlo. Considere-se o sistema de controlo em anel fechado da Figura 6.7.

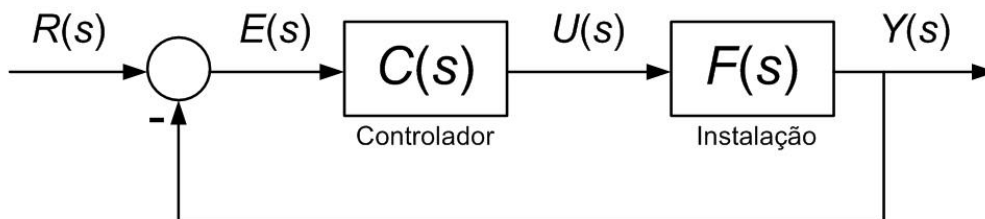


Figura 6.7 – Esquema de controlo de seguimento por retroação.

Se o anel aberto for dado por

$$G_0(s) = F(s)C(s) = \frac{N_0(s)}{D_0(s)}$$

a função de transferência do anel fechado será então dada por

$$H(s) = \frac{G_0(s)}{1 + G_0(s)} = \frac{\frac{N_0(s)}{D_0(s)}}{1 + \frac{N_0(s)}{D_0(s)}} = \frac{N_0(s)}{D_0(s) + N_0(s)}$$

Assim, os polos do anel aberto são dados pelas raízes de $D_0(s)$ e os polos do anel fechado são dados pelas raízes de $D_0(s) + N_0(s)$.

Considere-se agora a função (completamente artificial no sentido em que não representa nenhum sistema específico)

$$1 + G_0(s) = 1 + \frac{N_0(s)}{D_0(s)} = \frac{D_0(s) + N_0(s)}{D_0(s)}$$

e o contorno – Contorno de Nyquist – que envolve todo o semiplano complexo direito, ou seja a região instável do plano s , no sentido dos ponteiros do relógio.

Segundo o teorema acima, a imagem do contorno de Nyquist através da função

$$1 + G_0(s) = \frac{D_0(s) + N_0(s)}{D_0(s)}$$

terá um número de envolvimentos em redor da origem igual à diferença entre os polos instáveis (i.e. no interior do contorno de Nyquist) do anel fechado e os polos instáveis do anel aberto (ver Figura 6.8).

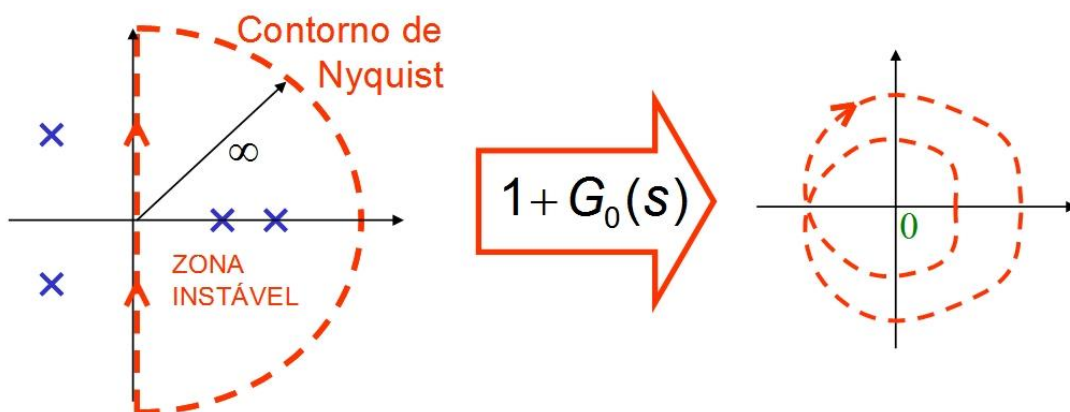


Figura 6.8 – Teorema dos Envolvimentos aplicado ao Contorno de Nyquist.

Sendo mais fácil analisar a imagem da função $G_0(s)$ (o anel aberto), a contagem de envolvimentos em redor da origem da função $1 + G_0(s)$ é equivalente à contagem de envolvimentos da função $G_0(s)$ em redor do ponto -1 , como se ilustra na Figura 6.9.

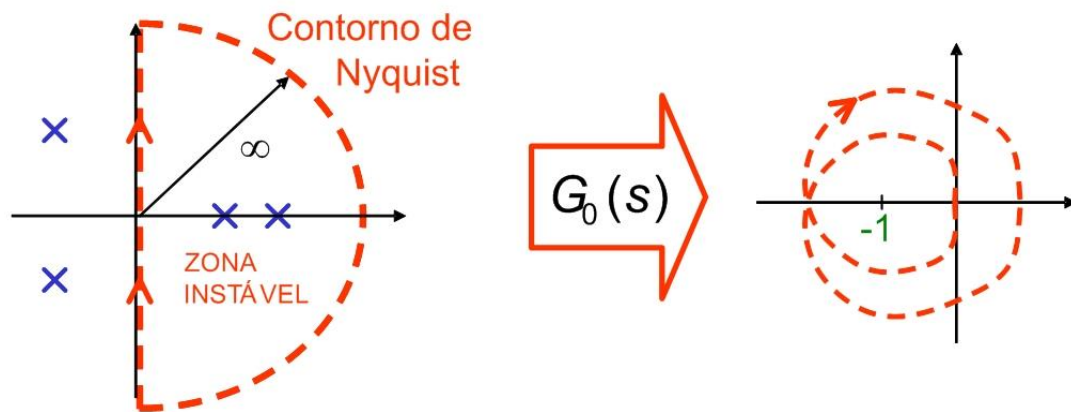


Figura 6.9 – Diagrama de Nyquist como a imagem do contorno de Nyquist através de $G_0(s)$.

Assim, o critério de Nyquist enuncia-se da seguinte forma. A imagem do contorno de Nyquist através da função $G_0(s)$ terá um número de envolvimentos em redor do ponto (-1) igual à diferença entre os polos instáveis do anel fechado e os polos instáveis do anel aberto, que equivale a escrever

$$N = Z - P$$

onde

- P – é o número de polos instáveis do anel aberto;
- Z – é o número de polos instáveis do anel fechado e
- N – é o número de envolvimentos em redor do ponto -1 no mesmo sentido do contorno de Nyquist (ver Figura 6.9).

O diagrama que se constrói através da imagem do contorno por $G_0(s)$ designa-se por diagrama de Nyquist de $G_0(s)$.

No caso mais geral, em que o anel aberto do sistema $G_0(s)$ é estável (i.e. $P = 0$), o critério de Nyquist resume-se a

$$Z = N$$

ou seja, para que o sistema em anel fechado resulte estável ($Z = 0$), o diagrama de Nyquist do anel aberto $G_0(s)$ não pode envolver o ponto -1 (i.e. $N = 0$). Se envolver o ponto -1 , o número N de voltas corresponde ao número de polos instáveis do sistema em anel fechado. Resulta disto, que para um anel aberto estável e no objetivo de ter um sistema em anel fechado estável, a circulação do ponto -1 deve ser evitada.

Considere-se agora o caso, menos frequente, em que o anel aberto do sistema $G_0(s)$ é instável (i.e. $P > 0$). Para que, novamente, o sistema em anel fechado resulte estável ($Z = 0$),

o diagrama de Nyquist do anel aberto $G_0(s)$ tem de envolver o ponto -1 , em sentido contrário ao do contorno de Nyquist, num número de voltas igual a

$$N = -P$$

(negativo porque no sentido contrário). Assim, para um anel aberto instável e no objetivo de ter um sistema em anel fechado, a circulação do ponto -1 tem de ocorrer no sentido contrário ao dos ponteiros do relógio P vezes.

Exemplo 6.3

Considere-se o sistema em anel aberto dado por

$$G_0(s) = \frac{1}{s + 1}$$

que é estável, pois o seu único polo vale -1 , e portanto $P = 0$. Se traçarmos o diagrama de Nyquist deste sistema, ou seja, o diagrama polar de

$$G(j\omega) = \frac{1}{j\omega + 1}$$

para valores de ω positivos e negativos, obtém-se o resultado da Figura 6.10.

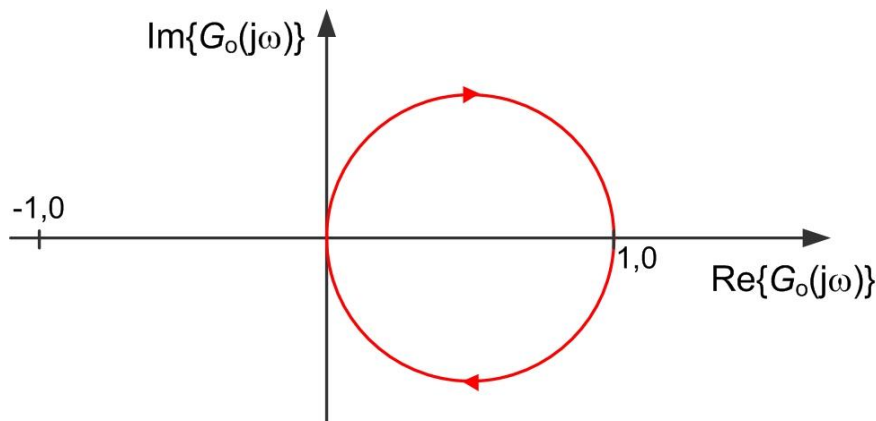


Figura 6.10 – Diagrama de Nyquist de $G_0(z)$ do Exemplo 6.3.

Este diagrama de Nyquist não realiza qualquer envolvimento em redor do ponto -1 , i.e. $N = 0$. Assim,

$$Z = N + P = 0$$

o anel fechado não terá qualquer polo instável. De modo a verificar este resultado, podemos calcular diretamente o anel fechado

$$H(s) = \frac{G_0(s)}{1 + G_0(s)} = \frac{\frac{1}{s + 1}}{1 + \frac{1}{s + 1}} = \frac{1}{s + 2}$$

que, com um polo em -2 , é estável.

■

Exemplo 6.4

Façamos agora a mesma análise para o anel aberto dado por

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

de segunda ordem e estável ($P = 0$). O traçado do diagrama de Nyquist está representado na Figura 6.11 e, como no exemplo anterior, $G_0(j\omega)$ não realiza qualquer envolvimento em redor do ponto -1 .

Conclui-se, portanto, que

$$Z = N + P = 0$$

e o anel fechado do sistema é estável (verifique como exercício!).

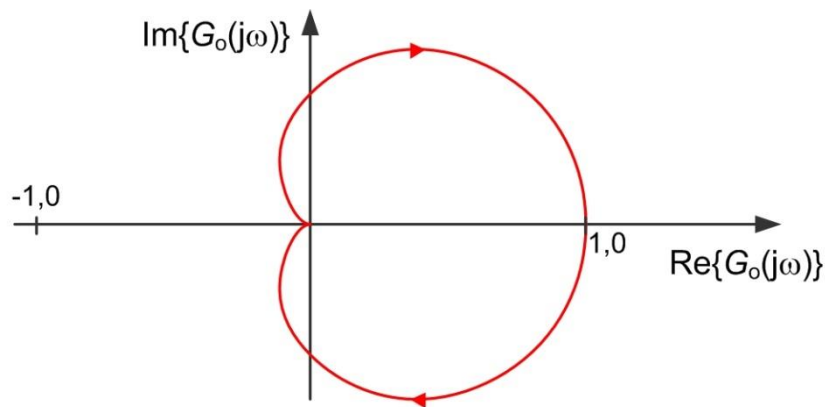


Figura 6.11 – Diagrama de Nyquist de $G_0(z)$ do Exemplo 6.4.

■

Exemplo 6.5

Considere-se um caso diferente com a análise do anel aberto dado por

$$G_0(s) = \frac{1}{(s^2 + 0.3s + 1)(s + 1)}$$

de terceira ordem e, ainda, estável ($P = 0$). O traçado do diagrama de Nyquist está representado na Figura 6.12, mas, neste caso, $G_0(j\omega)$ realiza dois envoltimentos em redor do ponto -1 , ou seja, $N = 2$.



Figura 6.12 – Diagrama de Nyquist de $G_0(s)$ do Exemplo 6.5.

Então, conclui-se que

$$Z = N + P = 2 + 0 = 2$$

e o anel fechado do sistema é instável com dois polos com parte real positiva. Verificando o resultado, se fecharmos o anel obtém-se

$$H(s) = \frac{1}{s^3 + 1.3s^2 + 1.3s + 1}$$

que tem polos em -1.40 (estável) e $0.05 \pm j1.20$ (instáveis!).

■

Exemplo 6.6

Finalmente, um caso com o anel aberto instável dado por

$$G_0(s) = \frac{2}{(s - 1)}$$

de primeira ordem ($P = 1$). O diagrama de Nyquist está representado na Figura 6.13, realizando um envolvimento em redor do ponto -1 mas no sentido contrário ao do contorno de Nyquist, ou seja, $N = -1$.

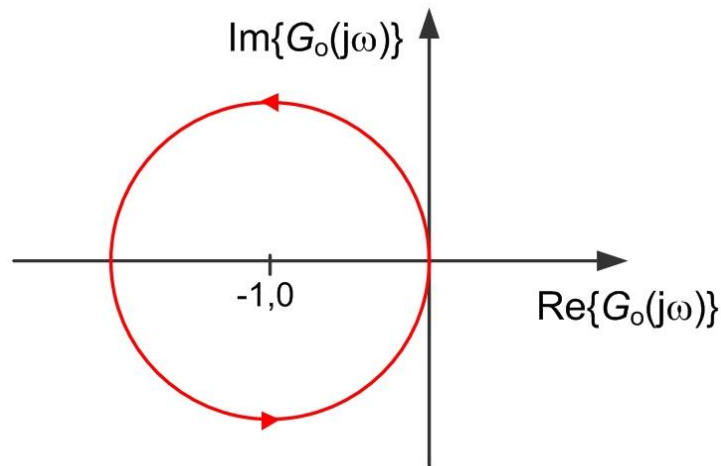


Figura 6.13 – Diagrama de Nyquist de $G_0(s)$ do Exemplo 6.6.

Então, conclui-se que

$$Z = N + P = -1 + 1 = 0$$

e o anel fechado do sistema é estável. Verifique-se o resultado calculando a função de transferência do anel fechado

$$H(s) = \frac{2}{s + 1}$$

que é estável.

■

6.4. Diagrama de Nyquist e a resposta em frequência

Podemos, nesta altura, perguntar qual a relação entre o diagrama de Nyquist e a resposta em frequência, também introduzida neste capítulo. Repare-se que o contorno fechado de Nyquist é constituído pelas seguintes três partes:

- o semieixo imaginário positivo, i.e. $s = j\omega$ com $0 \leq \omega < +\infty$;
- o arco infinito definido no limite de $s = \rho \cdot e^{j\theta}$ com $\rho \rightarrow +\infty$ e $-90^\circ < \theta < +90^\circ$;
- e o semieixo imaginário negativo, i.e. $s = j\omega$ com $-\infty < \omega \leq 0$.

O diagrama de Nyquist é a imagem destes três troços através de $G_0(s)$. Mas, a imagem do primeiro troço é o diagrama polar da resposta em frequência $G_0(j\omega)$. A imagem do arco infinito é, para um sistema com mais polos do que zeros, a origem coincidente com o limite de $G_0(j\omega)$ quando $\omega \rightarrow +\infty$. A imagem do terceiro troço é a resposta em frequência para frequências negativas $G_0(-j\omega)$, que resultará no diagrama polar da resposta em frequência rebatido em torno do eixo real.

6.5. Diagrama de Bode

O diagrama de Bode representa, explicitamente em dois diagramas paralelos, as funções de ganho ($M_G(\omega) = |G(j\omega)|$) e fase ($\phi_G(\omega) = \arg\{G(j\omega)\}$) em função da frequência angular ω . Este diagrama pode ser representado para o anel aberto do sistema, estando intimamente ligado ao diagrama de Nyquist, mas também para o anel fechado, fornecendo informações de carácter diferente.

Através da representação da resposta em frequência do anel aberto de sistemas estáveis ($P = 0$) é possível retirar a mesma informação que concluímos da análise do diagrama de Nyquist mais conclusões sobre a rapidez de resposta do sistema em anel fechado. No entanto, para sistemas em anel aberto instáveis ($P > 0$) o diagrama de Bode não dispensa a realização do Nyquist para verificação da estabilidade.

O diagrama de Bode do sistema em anel fechado concentra a informação importante no diagrama de ganho, resumindo-se esta à capacidade de seguimento e rapidez de resposta.

O diagrama de Bode é representado em escala logarítmica (de base 10) de frequência. A representação da frequência angular desenvolve-se assim em progressão geométrica.

Adicionalmente, a representação do ganho faz-se em escala logarítmica (de base 10) ou em escala linear recorrendo previamente à conversão do ganho para decibel (dB).

O ganho em dB é calculado como

$$G_{dB} = 20 \cdot \log_{10}|G|$$

Deste modo, a composição de fatores multiplicativos (por exemplo, de uma função de transferência) torna-se aditiva. Se o ganho

$$G = G_1 G_2 G_3$$

então o ganho em dB será dado por

$$\begin{aligned} G_{dB} &= 20 \cdot \log_{10}|G| = 20 \cdot \log_{10}|G_1 G_2 G_3| = \\ &20 \cdot \log_{10}|G_1| + 20 \cdot \log_{10}|G_2| + 20 \cdot \log_{10}|G_3| = G_{1,dB} + G_{2,dB} + G_{3,dB} \end{aligned}$$

A Figura 6.14 aponta alguns valores a memorizar, úteis para a manipulação de fatores de ganho em série. Note-se que o valor central correspondente à multiplicação pela unidade (não amplifica nem atenua) tem como equivalente a adição de 0 dB (que mantém o valor inalterado). O produto por fatores de 10 corresponde à adição de múltiplos de 20 dB; e logo a divisão por factores de 10 corresponde à subtração de múltiplos de 20 dB. Um valor de grande utilidade na construção do diagrama de bode é o +3 dB correspondente ao produto por $\sqrt{2}$. O motivo deste interesse será apresentado mais adiante no texto.

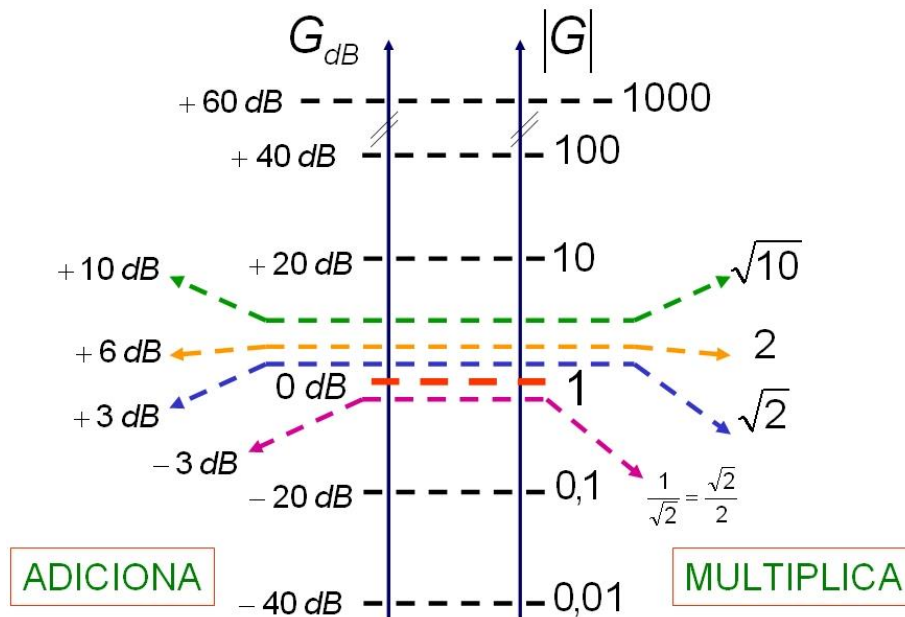


Figura 6.14 – Relação entre ganho absoluto e valor correspondente em dB.

6.6. Construção aproximada do diagrama de Bode – método das assíntotas

O traçado do diagrama de Bode pode ser realizado de forma aproximada tirando partido da escala logarítmica de ganho, que permite a sobreposição das diversas contribuições dos fatores, i.e.

$$\log(G_1(j\omega)G_2(j\omega)) = \log G_1(j\omega) + \log G_2(j\omega)$$

do facto do argumento do produto de dois números complexos ser igual à soma dos argumentos de cada um, ou seja

$$\arg\{G_1(j\omega)G_2(j\omega)\} = \arg\{G_1(j\omega)\} + \arg\{G_2(j\omega)\}$$

e ainda, da escala logarítmica de frequência, como ficará mais claro à frente.

Vejamos então a contribuição de cada fator de uma função de transferência para o traçado do diagrama de Bode.

Considere-se o fator simples de ganho

$$G(j\omega) = K$$

amplificador, se $K > 1$, ou atenuador, se $K < 1$. Então, a resposta em frequência é dada, no ganho de amplitude por

$$\log M_G(\omega) = \log|G(j\omega)| = \log|K|$$

e no desvio de fase por

$$\phi_G(\omega) = \arg\{K\} = 0^\circ \text{ se } K > 0$$

ou

$$\phi_G(\omega) = \arg\{K\} = 180^\circ \text{ se } K < 0$$

A Figura 6.15 representa o diagrama de Bode deste fator.

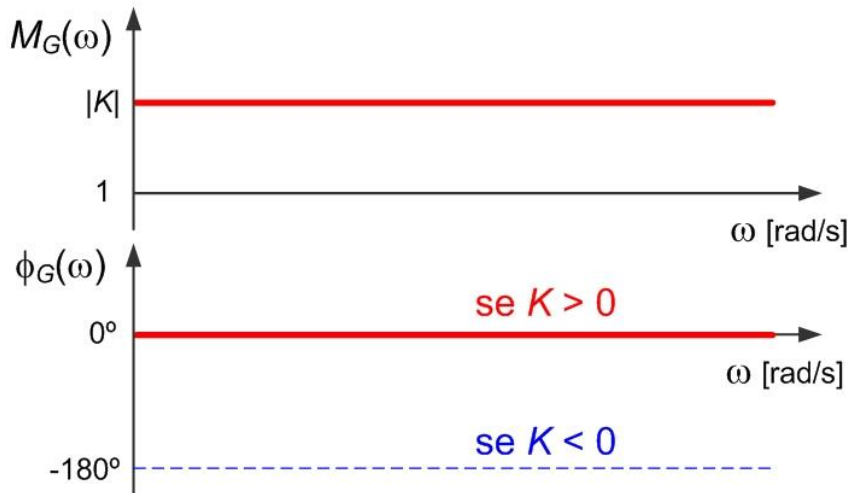


Figura 6.15 – Diagrama de Bode do fator $G(j\omega) = K$.

Note-se que um ganho K representa uma abstracção matemática pois nenhum sistema físico real pode apresentar uma função de transferência com ganho constante para frequências indefinidamente elevadas. Por exemplo, a realização eletrónica de um amplificador é feita com elementos eletrónicos ativos (e.g. transístor) que apresentam limitações a partir de um dado valor de frequência máximo. Na realização prática de um amplificador, apenas se tem de garantir que, na gama de frequências de interesse (adiante designada por largura de banda do anel fechado) o bloco amplificador mantém o valor do ganho constante.

Doravante neste texto considera-se, por omissão, o ganho K positivo sabendo que, se for negativo, basta subtrair (ou adicionar) 180° à fase obtida.

Considere-se agora um fator polinomial de 1ª ordem no numerador da função de transferência

$$G(j\omega) = K \left(1 + j \frac{\omega}{\omega_c} \right)$$

Este fator, viu-se já, está relacionado com a posição de um zero da função de transferência, neste caso, em $z = -\omega_c$.

A resposta em frequência é dada, no ganho de amplitude por

$$\log M_G(\omega) = \log|K| + \log \left| 1 + j \frac{\omega}{\omega_c} \right| = \log|K| + \log \sqrt{1 + \frac{\omega^2}{\omega_c^2}}$$

que, de forma assintótica para 0 e $+\infty$, se pode aproximar como

$$\log M_G(\omega) \approx \begin{cases} \log K + 0 & \Leftarrow \omega \rightarrow 0 \\ \log K + \log \frac{\omega}{\omega_c} & \Leftarrow \omega \rightarrow +\infty \end{cases}$$

Isto significa que à baixa frequência ($\omega \rightarrow 0$) o factor contribui com um valor constante e na alta frequência ($\omega \rightarrow +\infty$) a resposta de ganho evolui em linha recta com inclinação unitária. Para entender melhor esta última afirmação repare-se que

$$\log M_G(\omega) \approx \log K + \log \frac{\omega}{\omega_c} = \mathbf{1} \cdot \log \omega + (\log K - \log \omega_c)$$

e como a representação no diagrama de Bode é feita em escalas logarítmicas de ganho e frequência, i.e. $\log M_G(\omega)$ e $\log \omega$, a assíntota é uma reta de inclinação **1**. Na Figura 6.16 faz-se a representação das duas assíntotas e representa-se a curva real do ganho de amplitude deste fator. A indicação 1:1 representa que, para valores de frequência suficientemente elevados, o ganho aumenta uma década (dez vezes) por cada década de aumento da frequência. Como uma década de aumento no ganho equivale à adição de +20 dB, diz-se que o ganho cresce à razão de +20 dB por década, ou +20 dB/dec.

Por outro lado, a contribuição da fase é dada por

$$\phi_G(\omega) = \arg\{K\} + \arg\left\{1 + j\frac{\omega}{\omega_c}\right\} = 0^\circ + \operatorname{atan}\left\{\frac{\omega}{\omega_c}\right\}$$

que, de forma aproximada, se pode escrever

$$\phi_G(\omega) \approx \begin{cases} 0^\circ & \Leftarrow \omega \rightarrow 0 \\ 45^\circ & \Leftarrow \omega = \omega_c \\ 90^\circ & \Leftarrow \omega \rightarrow +\infty \end{cases}$$

Apesar da expressão anterior, os 45° à frequência ω_c são exactos.

Representamos então, também na Figura 6.16, a contribuição da fase. Aqui, repare-se que a diferença entre o diagrama assintótico e o traçado real faz-se notar desde uma década antes até uma década após a frequência ω_c . Mais à frente se fará, com detalhe, a quantificação destas diferenças entre os diagramas assintóticos e reais, quer na fase quer no ganho.

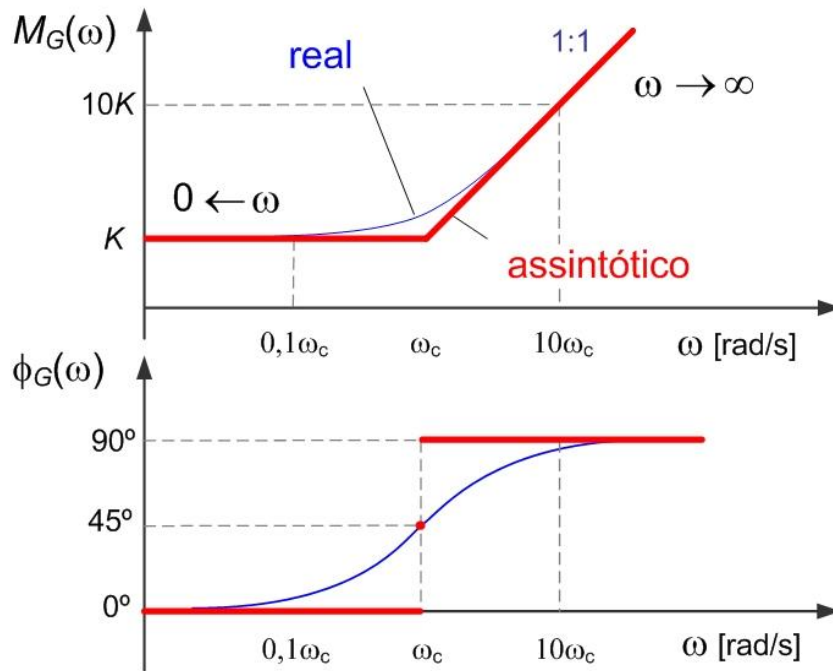


Figura 6.16 – Diagrama de Bode do fator $G(j\omega) = K \cdot (1 + j\omega/\omega_c)$.

Pensemos agora naquilo que muda para o caso anterior se considerarmos o fator

$$G(j\omega) = \frac{K}{1 + j\frac{\omega}{\omega_c}}$$

Como dividir no argumento do logaritmo é subtrair o logaritmo do inverso, resulta

$$\log M_G(\omega) = \log|K| - \log\left|1 + j\frac{\omega}{\omega_c}\right| = \log|K| - \log\sqrt{1 + \frac{\omega^2}{\omega_c^2}}$$

$$\approx \begin{cases} \log K - 0 & \Leftarrow \omega \rightarrow 0 \\ \log K - \log\frac{\omega}{\omega_c} & \Leftarrow \omega \rightarrow +\infty \end{cases}$$

ou seja, teremos da mesma forma uma reta mas de inclinação unitária negativa (ver Figura 6.17). De forma algo semelhante, como dividir por um numero complexo é inverter o ganho e subtrair a fase,

$$\phi_G(\omega) = \arg\{K\} - \arg\left\{1 + j\frac{\omega}{\omega_c}\right\} = 0^\circ - \text{atan}\left\{\frac{\omega}{\omega_c}\right\}$$

como representado na Figura 6.17.

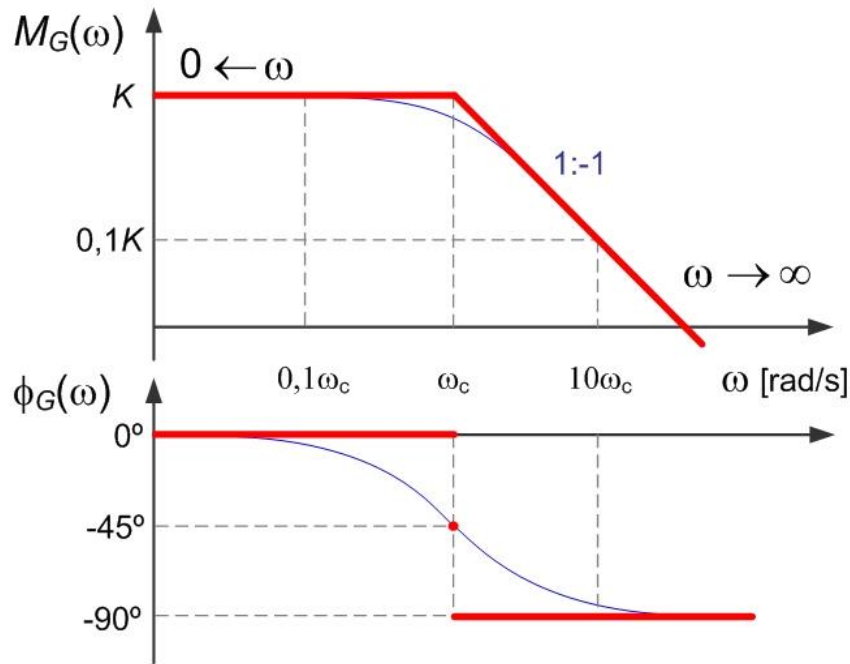


Figura 6.17 – Diagrama de Bode do fator $G(j\omega) = K \cdot (1 + j\omega/\omega_c)^{-1}$.

Note-se então que na representação em escalas logarítmicas as contribuições provenientes de zeros e polos da função de transferência resultam absolutamente simétricas.

Talvez o leitor já tenha reparado no facto que na escala logarítmica o valor 0 não aparece representado. Se formos em busca do zero, por exemplo na frequência, isso levar-nos-á para regiões do diagrama cada vez mais à esquerda. Quando temos o caso em que o zero está na origem o fator de Bode correspondente é

$$G(j\omega) = K \cdot (j\omega)$$

e logo

$$\log M_G(\omega) = \log K + \log|j\omega| = \log K + 1 \cdot \log \omega$$

$$\phi_G(\omega) = \arg\{K\} + \arg\{j\omega\} = 0^\circ + 90^\circ = +90^\circ$$

como representado na Figura 6.18. Esta representação é compatível com aquela feita na Figura 6.16 se atendermos ao facto que, neste caso, $\omega_c = 0$ rad/s e logo o zero não aparece representado.

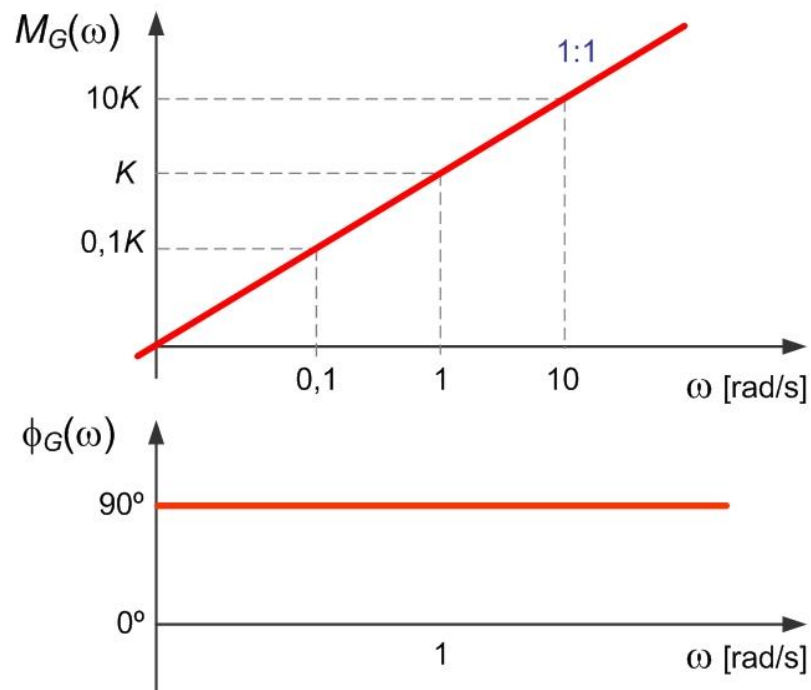


Figura 6.18 – Diagrama de Bode do fator $G(j\omega) = K \cdot (j\omega)$.

De forma análoga, temos para a contribuição resultante de um polo na origem (integrador)

$$G(j\omega) = \frac{K}{j\omega}$$

o ganho de amplitude e desvio de fase dados por

$$\log M_G(\omega) = \log K - \log|j\omega| = \log K - 1 \cdot \log \omega$$

$$\phi_G(\omega) = \arg\{K\} - \arg\{j\omega\} = 0^\circ - 90^\circ = -90^\circ$$

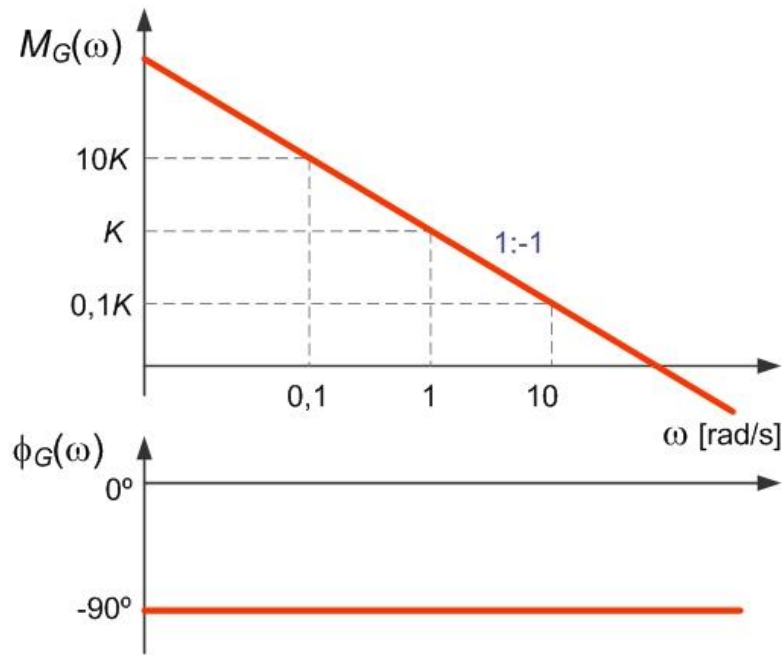


Figura 6.19 – Diagrama de Bode do fator $G(j\omega) = K/(j\omega)$.

Com o que vimos até agora já podemos realizar o traçado aproximado de diagramas de Bode de funções de transferência com polos e zeros reais.

Exemplo 6.7

Considere-se então o traçado do diagrama de Bode do sistema representado pela função de transferência

$$G(j\omega) = K_0 \frac{s + \omega_{c2}}{(s + \omega_{c1})(s + \omega_{c3})}$$

Substituindo a variável de Laplace s por $j\omega$ e colocando na forma de Bode obtém-se

$$G(j\omega) = K \frac{1 + j \frac{\omega}{\omega_{c2}}}{\left(1 + j \frac{\omega}{\omega_{c1}}\right) \left(1 + j \frac{\omega}{\omega_{c3}}\right)}$$

tal que (verifique como exercício)

$$K = K_0 \frac{\omega_{c2}}{\omega_{c1} \omega_{c3}}$$

Teremos então para o ganho e fase as contribuições (ver Figura 6.20 e Figura 6.21)

$$\log M_G(\omega) = \log|K| - \log \left| 1 + j \frac{\omega}{\omega_{c1}} \right| + \log \left| 1 + j \frac{\omega}{\omega_{c2}} \right| - \log \left| 1 + j \frac{\omega}{\omega_{c3}} \right|$$

$$\phi_G(\omega) = \arg\{K\} - \arg \left\{ 1 + j \frac{\omega}{\omega_{c1}} \right\} + \arg \left\{ 1 + j \frac{\omega}{\omega_{c2}} \right\} - \arg \left\{ 1 + j \frac{\omega}{\omega_{c3}} \right\}$$

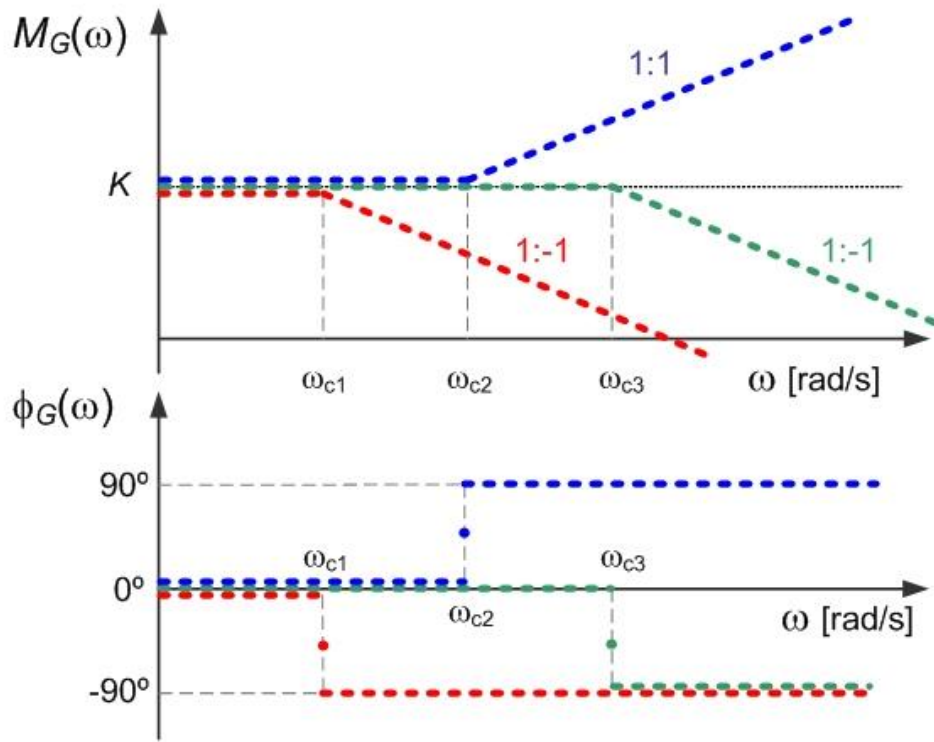


Figura 6.20 – Traçado assintótico dos fatores de Bode do Exemplo 6.7.

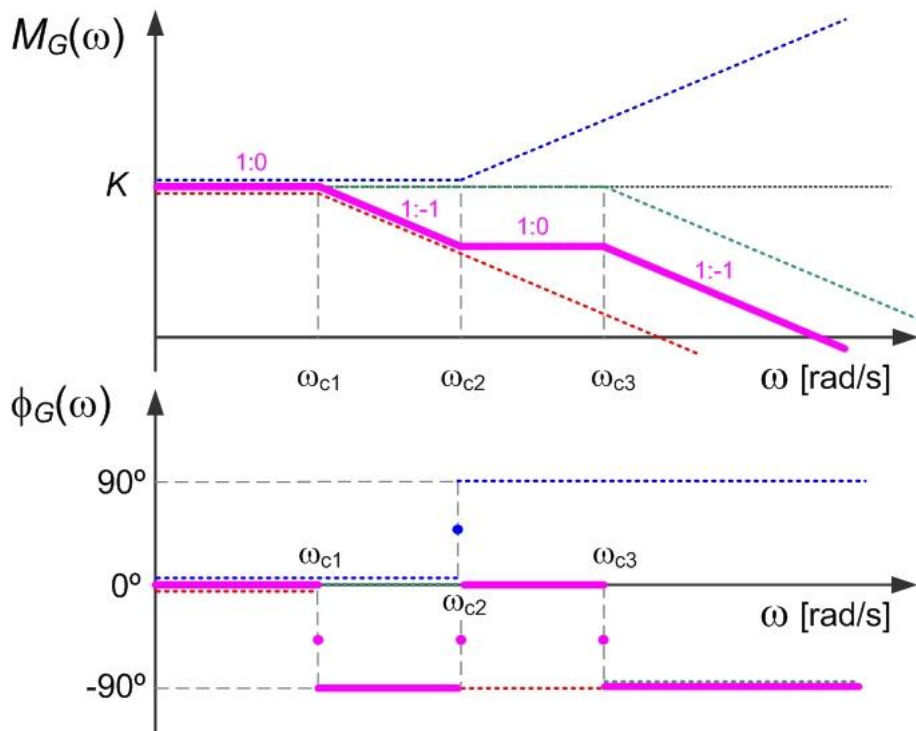


Figura 6.21 – Sobreposição dos fatores de Bode do Exemplo 6.7.

É agora possível estimar o comportamento real do diagrama como se representa na Figura 6.22.

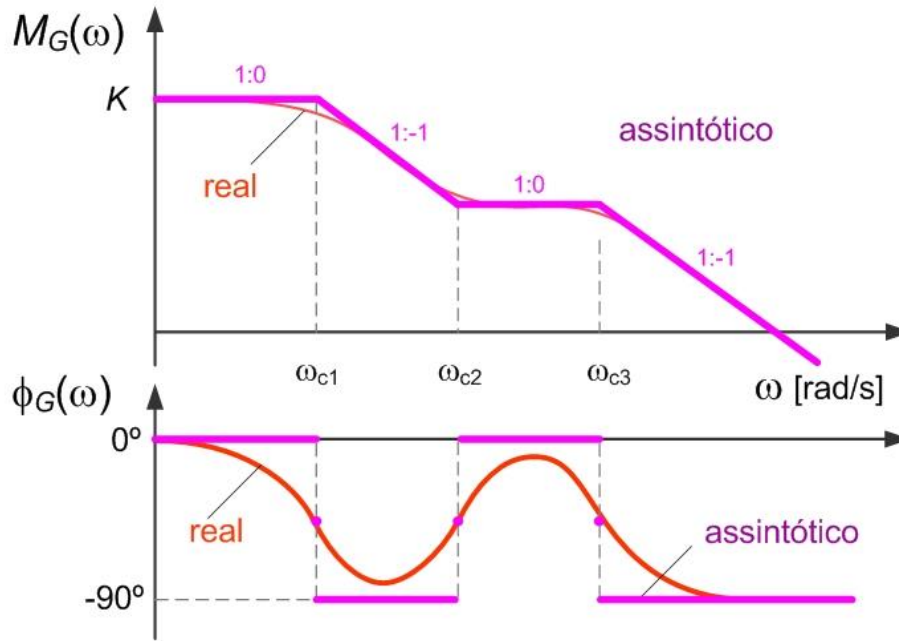


Figura 6.22 – Diagrama de Bode do Exemplo 6.7.

Exemplo 6.8

Considere-se agora o traçado do diagrama de Bode do sistema representado pela função de transferência

$$G(j\omega) = K_0 \frac{s + \omega_{c1}}{s \cdot (s + \omega_{c2})}$$

com um fator resultante de um polo na origem.

Temos então a resposta em frequência colocada na forma de Bode como

$$G(j\omega) = \frac{K}{j\omega} \frac{1 + j\frac{\omega}{\omega_{c1}}}{\left(1 + j\frac{\omega}{\omega_{c2}}\right)}$$

onde

$$K = K_0 \frac{\omega_{c1}}{\omega_{c2}}$$

e

$$\log M_G(\omega) = \log|K| - \log|j\omega| + \log \left| 1 + j\frac{\omega}{\omega_{c1}} \right| - \log \left| 1 + j\frac{\omega}{\omega_{c2}} \right|$$

$$\phi_G(\omega) = \arg\{K\} - \arg\{j\omega\} + \arg \left\{ 1 + j\frac{\omega}{\omega_{c1}} \right\} - \arg \left\{ 1 + j\frac{\omega}{\omega_{c2}} \right\}$$

Nas Figura 6.23 e Figura 6.24 representam-se dois passos da construção assintótica e na Figura 6.25 o diagrama final estimado. Neste exemplo, note-se que o fator de Bode relativo ao polo na origem (com o zero seria o mesmo) domina na baixa frequência.

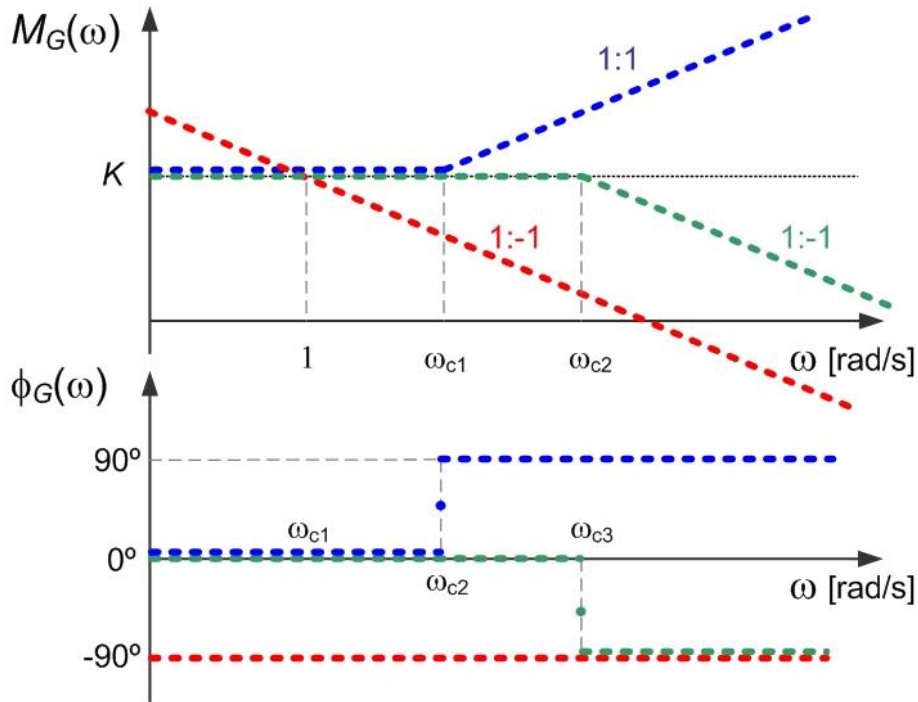


Figura 6.23 – Traçado assintótico dos fatores de Bode do Exemplo 6.8.

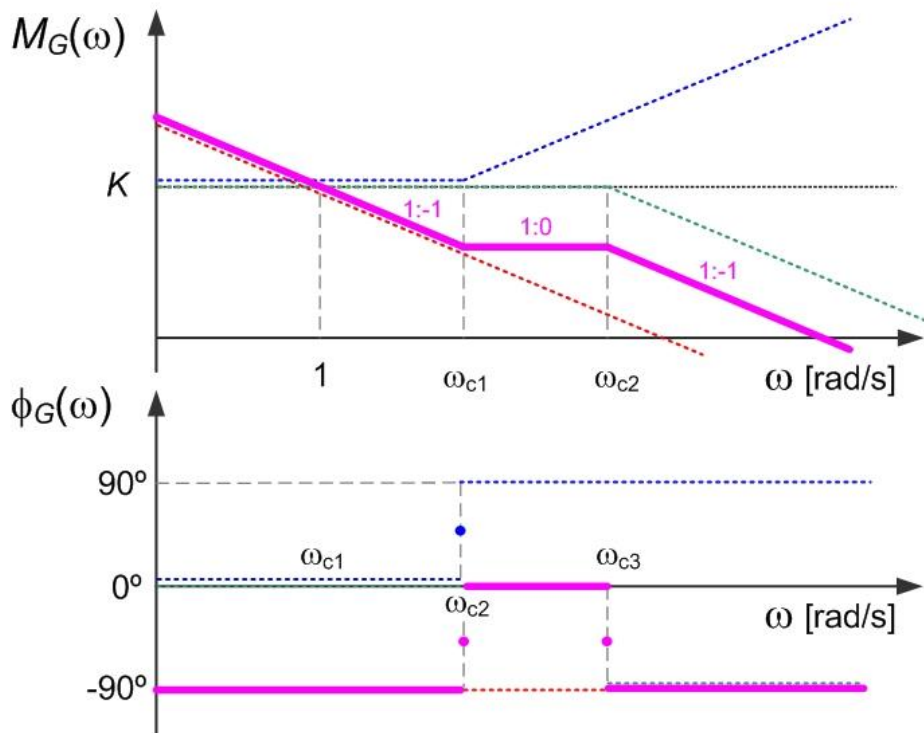


Figura 6.24 – Sobreposição dos fatores de Bode do Exemplo 6.8.

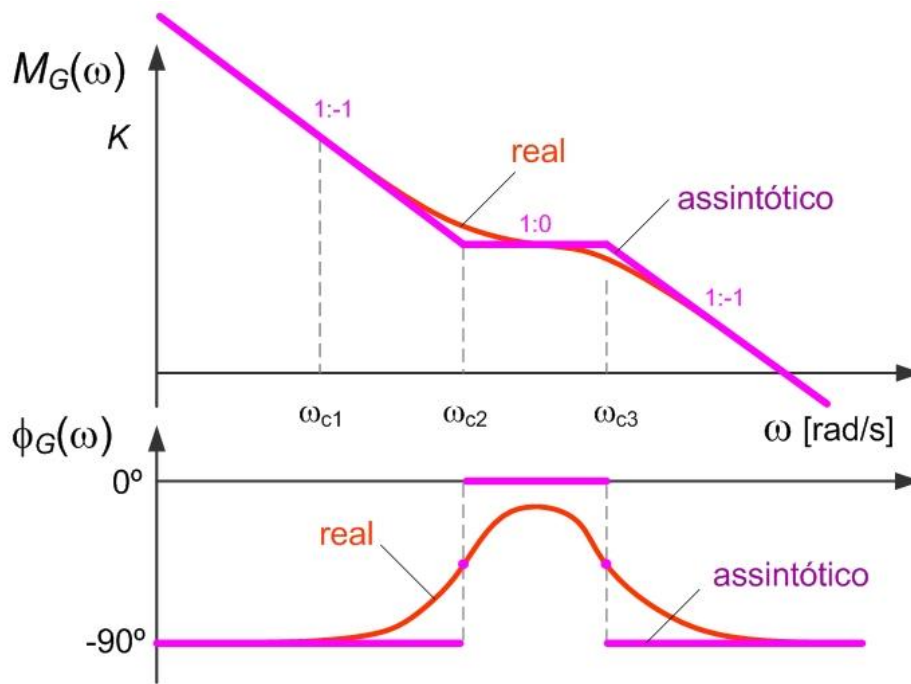


Figura 6.25 – Diagrama de Bode do Exemplo 6.8.

6.7. Diagrama de Bode real e assintótico

Vejamos então agora, com mais detalhe, qual a diferença entre o diagrama de Bode real e assintótico, para o ganho de amplitude e para o desvio de fase.

Dada a simetria da representação dos fatores resultante de polos e zeros estas observações são generalizáveis. Começemos pelo ganho de amplitude de fator resultante de um polo real em $-\omega_c$

$$G(j\omega) = \frac{K}{1 + j\frac{\omega}{\omega_c}}$$

O ganho à frequência de canto ω_c será dado por

$$M_G(\omega_c) = \frac{K}{|1 + j\frac{\omega}{\omega_c}|} = \frac{K}{|1 + j1|} = \frac{K}{\sqrt{2}} = 0.707K$$

ou seja

$$M_G(\omega_c)_{dB} = 20 \log_{10} \frac{K}{\sqrt{2}} = 20 \log_{10} K - 20 \log_{10} 1.414 = K_{dB} - 3.01 \text{ dB}$$

ou seja, à frequência de canto ω_c teremos um diferença de -3 dB , como representado na Figura 6.26. Nas frequências $\omega_c/3$ (antes) e $3\omega_c$ (depois e equidistante) o diagrama real vale -0.5 dB que o assintótico, ou seja, o ganho real é cerca de 95% daquele indicado pelo diagrama assintótico. No caso de um zero da função de transferência teremos à frequência

de canto +3 dB (relativamente ao diagrama assintótico) que equivale aproximadamente a 1.4 vezes o ganho de patamar às baixas frequências.

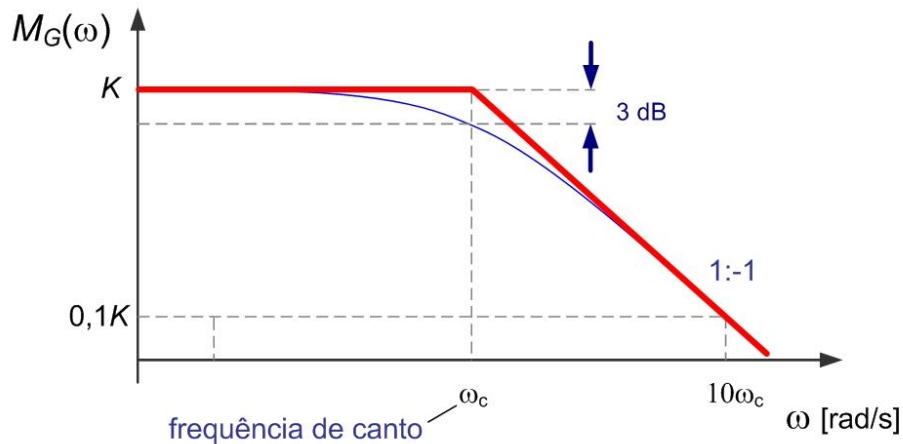


Figura 6.26 – Diagrama de Bode de ganho real de um polo sobre o eixo real.

Relativamente ao diagrama de fase, pode observar-se na Figura 6.27 o comportamento da curva real da contribuição de um zero (no semieixo real negativo). Os limites, quando a frequência vai para zero e infinito, são 0° e $+90^\circ$, respectivamente; e no valor da frequência de canto ω_c temos $+45^\circ$ de fase.

É possível calcular a inclinação da tangente à curva quando $\omega = \omega_c$ (que é deixado ao leitor como exercício) que resulta em 66° por década, cruzando as assíntotas em $0.2\omega_c$ e $5\omega_c$, como se observa na figura. Apesar de não coincidir com a tangente, uma melhor aproximação (com 6.5° de desvio máximo) surge de ligar as assíntotas desde $0.1\omega_c$ (uma década antes de ω_c) até $10\omega_c$ (uma década após ω_c).

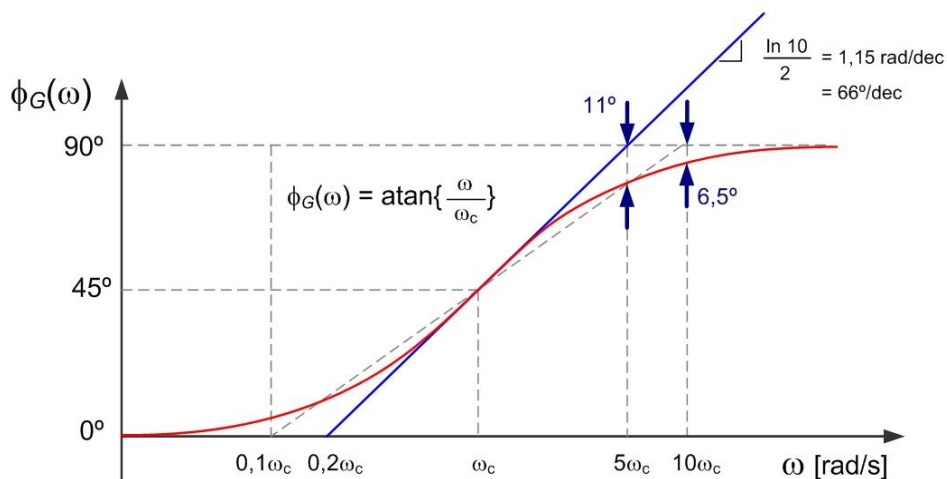


Figura 6.27 – Comportamento da fase real de um zero de fase mínima.

Na situação em que temos zeros (ou polos) reais duplos estes efeitos serão duplicados, como representado na Figura 6.28.

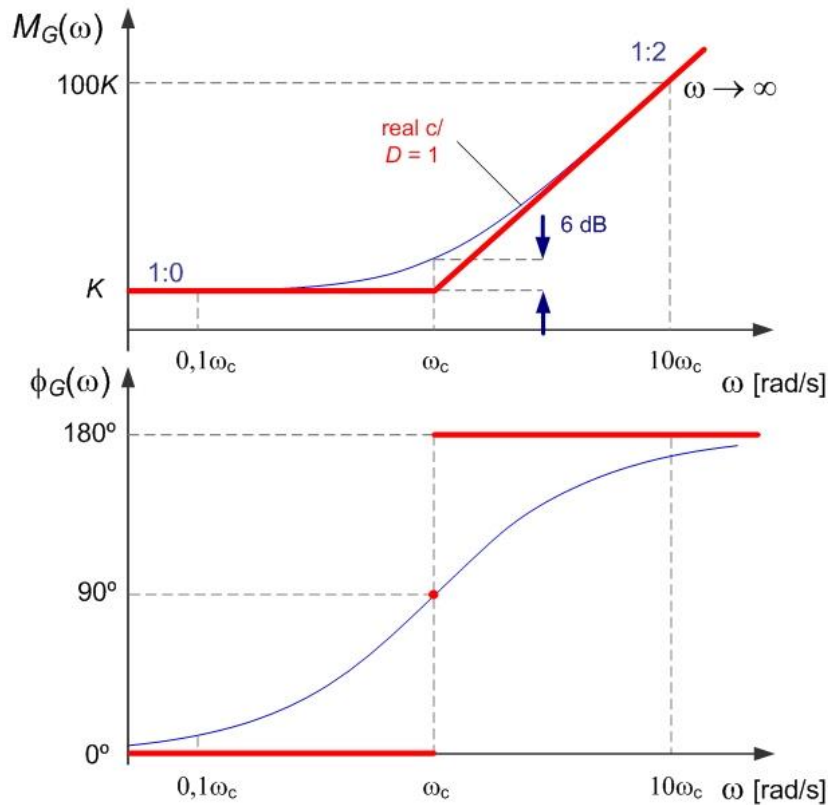


Figura 6.28 – Diagrama de Bode de um sistema de 2ª ordem com zero duplo.

Uma situação diferente é a que ocorre quando o fator de amortecimento de um facto de 2ª ordem é inferior à unidade e temos polos complexos conjugados.

Considere-se então um fator da função de transferência do tipo

$$G(s) = \frac{K \cdot \omega_n^2}{s^2 + 2D\omega_n s + \omega_n^2}$$

que colocado na forma de Bode resulta em

$$G(j\omega) = \frac{K}{\left(1 - \frac{\omega^2}{\omega_n^2}\right) + j \frac{2D\omega}{\omega_n}}$$

Note-se que, quando a frequência ω está na vizinhança da frequência natural ω_n , o módulo do denominador será tão reduzido quanto menor for o fator de amortecimento D . Se o amortecimento for nulo o ganho de amplitude é infinitamente elevado representando a situação de ressonância, bem conhecida dos engenheiros civis e mecânicos pela sua capacidade destrutiva das estruturas, e dos engenheiros eletrotécnicos na capacidade de colocar uma antena em sintonia com uma estação emissora. A Figura 6.29 apresenta a forma dos diagramas de Bode reais para diversos valores de amortecimento. Repare-se que com a diminuição do amortecimento a transição de fase vai sendo progressivamente mais rápida em redor do ponto médio em $\omega = \omega_c = \omega_n$.

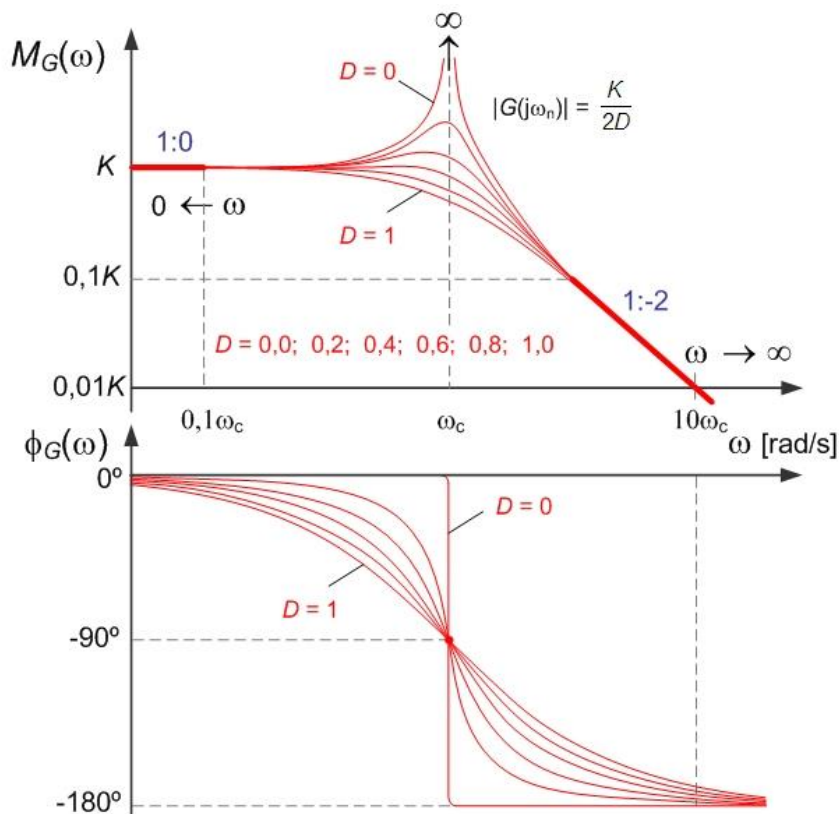


Figura 6.29 – Diagrama de Bode real para sistemas de 2ª ordem com polos complexos conjugados.

Mais uma vez, se em vez de polos tivermos zeros complexos conjugados os diagramas são verticalmente simétricos, apresentando um fenómeno de antirressonância. Os sistemas com zeros anti ressonantes não apresentam reação à saída (em regime permanente) quando excitados nestes valores de frequência.

6.8. Contribuição de polos instáveis e zeros de fase não mínima

Até este ponto, considerou-se apenas a contribuição de zeros e polos no semiplano complexo esquerdo, ou seja, zeros de fase mínima e polos estáveis ou marginalmente estáveis.

Um zero no semiplano direito (designado por zero de fase não mínima) contribui então com um fator de Bode

$$G(j\omega) = K \left(1 - j \frac{\omega}{\omega_c} \right)$$

Este fator difere daquele já visto para o zero de fase mínima no sinal associado à componente imaginária. Assim, esta diferença não afeta o ganho de amplitude igualmente dado por (ver Figura 6.30)

$$\log M_G(\omega) = \log|K| + \log \left| 1 - j \frac{\omega}{\omega_c} \right| = \log|K| + \log \sqrt{1 + \frac{\omega^2}{\omega_c^2}}$$

$$\approx \begin{cases} \log K + 0 & \Leftarrow \omega \rightarrow 0 \\ \log K + \log \frac{\omega}{\omega_c} & \Leftarrow \omega \rightarrow +\infty \end{cases}$$

mas a fase vem afetada de um sinal contrário, ou seja (Figura 6.30)

$$\phi_G(\omega) = \arg\{K\} + \arg\left\{1 - j\frac{\omega}{\omega_c}\right\} = 0^\circ - \text{atan}\left\{\frac{\omega}{\omega_c}\right\}$$

$$\approx \begin{cases} 0^\circ & \Leftarrow \omega \rightarrow 0 \\ -45^\circ & \Leftarrow \omega = \omega_c \\ -90^\circ & \Leftarrow \omega \rightarrow +\infty \end{cases}$$

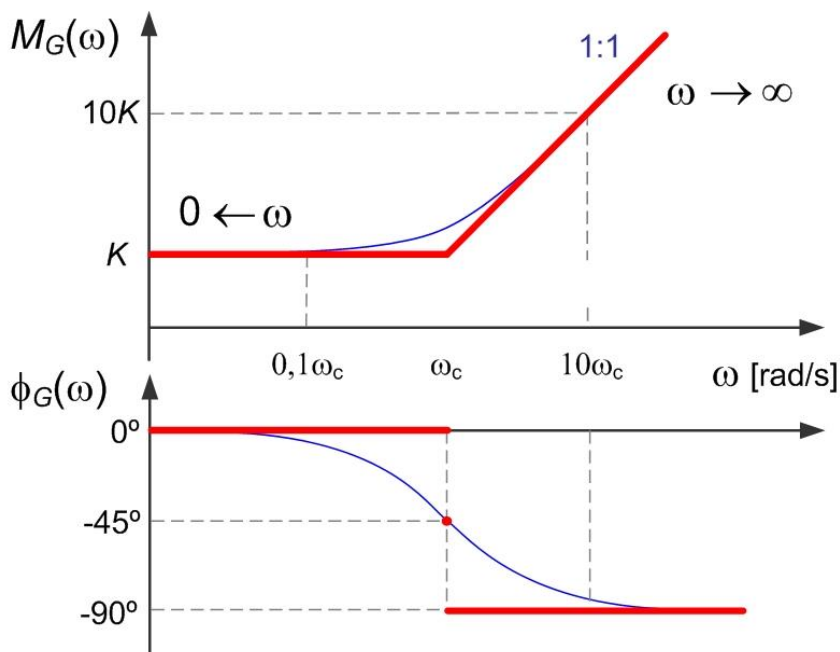


Figura 6.30 – Diagrama de Bode do fator $G(j\omega) = K \cdot (1 - j\omega/\omega_c)$.

O zero de fase não mínima comporta-se, na fase, da mesma forma que um polo estável, reduzindo a fase do conjunto. Isto constituirá um problema para o sistema de controlo pois facilitará a perda de estabilidade por envolvimento do ponto -1 no diagrama de Nyquist.

De forma análoga, podemos realizar o estudo da contribuição de um polo instável com o fator de Bode

$$G(j\omega) = \frac{K}{1 - j\frac{\omega}{\omega_c}}$$

Novamente, a diferença cinge-se à fase. O ganho é dado por (Figura 6.31)

$$\log M_G(\omega) = \log|K| - \log\left|1 - j\frac{\omega}{\omega_c}\right| = \log|K| - \log\sqrt{1 + \frac{\omega^2}{\omega_c^2}}$$

$$\approx \begin{cases} \log K - 0 & \Leftarrow \omega \rightarrow 0 \\ \log K - \log \frac{\omega}{\omega_c} & \Leftarrow \omega \rightarrow +\infty \end{cases}$$

e a fase

$$\phi_G(\omega) = \arg\{K\} - \arg\left\{1 - j\frac{\omega}{\omega_c}\right\} = 0^\circ + \operatorname{atan}\left\{\frac{\omega}{\omega_c}\right\}$$

$$\begin{cases} 0^\circ & \Leftarrow \omega \rightarrow 0 \\ +45^\circ & \Leftarrow \omega = \omega_c \\ +90^\circ & \Leftarrow \omega \rightarrow +\infty \end{cases}$$

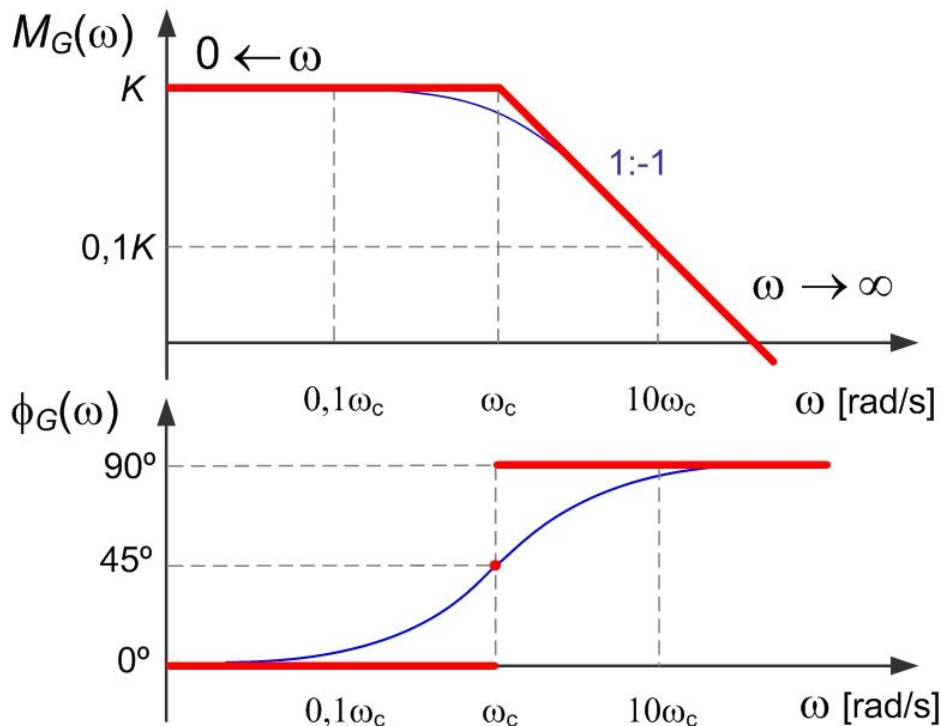


Figura 6.31 – Diagrama de Bode do fator $G(j\omega) = K \cdot (1 - j\omega/\omega_c)^{-1}$.

Dois comentários são importantes para este caso relativo ao polo instável. O primeiro vai no sentido de questionar a realização de um diagrama de Bode de um sistema instável, visto que a resposta no tempo em regime permanente sinusoidal ficará dominada pela componente exponencial crescente resultante da instabilidade. No entanto, o diagrama de Bode do anel aberto será um passo intermédio para a realização de um sistema em anel fechado estável e nesse sentido, aceita-se o traçado como ferramenta de projeto. O segundo comentário, surge na forma de adivinha: “se a adição de um zero de fase não mínima é prejudicial à estabilidade do anel fechado porque introduz mais fase negativa, isso significa que a adição de um polo instável é benéfica para a estabilidade do anel fechado porque introduz fase positiva?” A resposta a esta pergunta é: “não!” – como sugere o senso comum. A adição de polos instáveis no anel aberto não melhora, em geral, a estabilidade do anel fechado. Sugere-se ao leitor que

reflita sobre a justificação desta resposta tendo em conta os fundamentos do critério de Nyquist.

6.9. Construção do diagrama de Nyquist a partir do diagrama de Bode

Vimos já ser possível realizar o traçado do diagrama de Bode, de forma aproximada, através do método das assíntotas. Como o traçado do diagrama polar (ou do diagrama de Nyquist) não é tão simples, pode-se usar o diagrama de Bode como passo intermédio para o traçado do diagrama de Nyquist. Isto, sem desvalorizar o resultado que é o diagrama de Bode por si. O diagrama de Bode representa, ao longo da frequência angular ω , o comportamento do módulo e da fase da resposta em frequência. Recolhendo do diagrama de Bode um conjunto de pontos notáveis, podemos, com bastante aproximação, realizar o traçado do diagrama de Nyquist. O exemplo seguinte ilustra este procedimento.

Exemplo 6.9

Considere-se a função de transferência do sistema em anel aberto dada por

$$G(j\omega) = K_0 \frac{s + \omega_{c2}}{(s + \omega_{c1})(s + \omega_{c3})}$$

A Figura 6.32 representa o diagrama de Bode deste sistema onde foi marcado um conjunto de pontos notáveis pela facilidade de marcação no diagrama polar, nomeadamente, pontos com valores de fase igual a 0 , -45° ou -90° .

Este conjunto de (7) pontos está registado na tabela seguinte e têm correspondência com os pontos do diagrama polar da Figura 6.33. Na Figura 6.34 apresenta-se o diagrama de Nyquist completo, obtido por rebatimento em torno do eixo real do diagrama polar da Figura 6.33.

Conjunto de pontos notáveis para o traçado do Nyquist do Exemplo 6.9

	Frequência angular ω	$M_G(\omega)$	$\phi_G(\omega)$
1	$\omega \ll \omega_{c1}$	$\approx K_1$	$\approx 0^\circ$
2	$\omega = \omega_{c1}$	$\approx 0.7K_1$ (-3dB)	-45°
3	$\omega = \sqrt{\omega_{c1}\omega_{c2}}$	$= \sqrt{K_1K_2}$	$-90^\circ + \Delta$
4	$\omega = \omega_{c2}$	$\approx 1.4K_2$ (+3dB)	-45°
5	$\omega = \sqrt{\omega_{c2}\omega_{c3}}$	$\approx K_2$	$0^\circ - \Delta$
6	$\omega = \omega_{c3}$	$\approx 0.7K_2$ (-3dB)	-45°
7	$\omega \gg \omega_{c3}$	≈ 0 (-20 dB/dec)	$\approx -90^\circ$

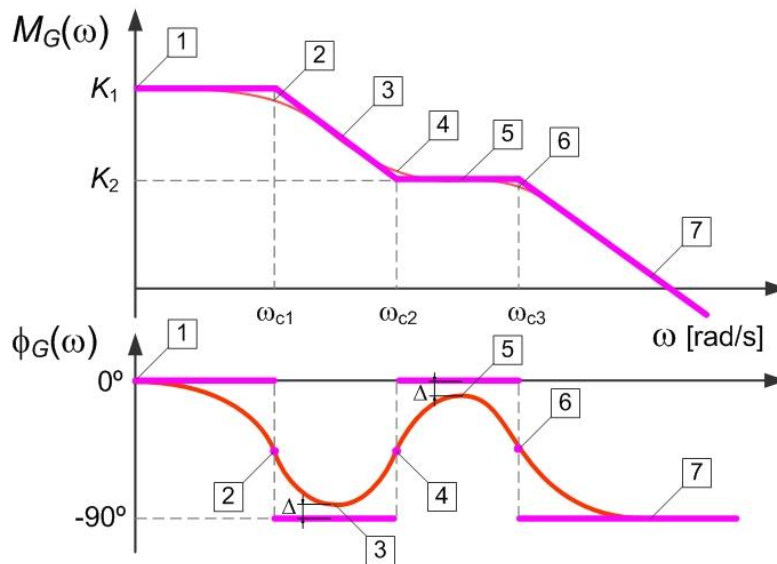


Figura 6.32 – Traçado do diagrama de Bode do Exemplo 6.9.

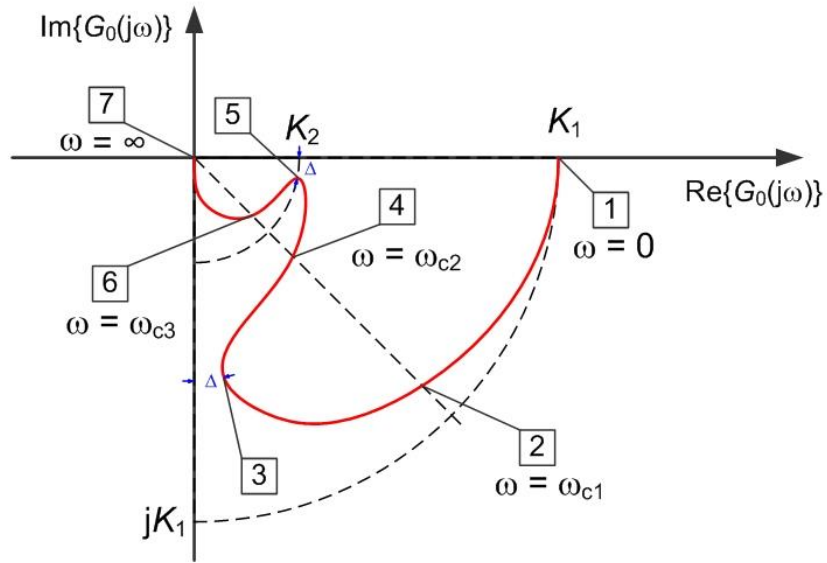


Figura 6.33 – Construção do diagrama polar de $G_0(j\omega)$ a partir do diagrama de Bode (Exemplo 6.9).

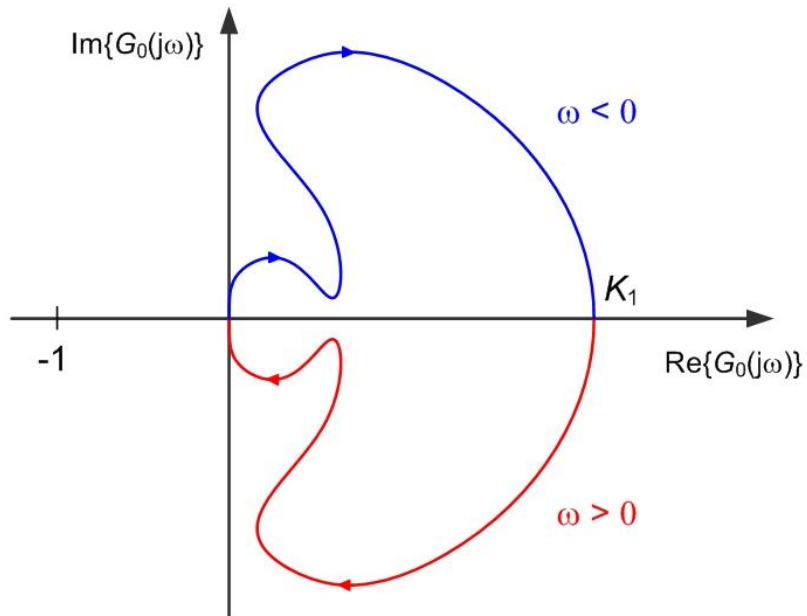


Figura 6.34 – Diagrama de Nyquist do Exemplo 6.9.

■

7. Especificação e projeto na frequência

Este capítulo debruça-se sobre um primeiro aspeto, mais básico, da especificação em frequência pela via da robustez do sistema em anel fechado a variações de ganho de amplitude ou a desvios de fase. Os dois parâmetros que daqui resultam, mais as respetivas frequências de cruzamento, permitem especificar o sistema com graus de liberdade equivalentes aos realizados no domínio do plano s , no capítulo 0.

No caso mais geral, em que o anel aberto do sistema $G_0(s)$ é estável (i.e. com $P = 0$), a especificação de estabilidade do anel fechado obriga a que

$$Z = N = 0$$

Para estes sistemas, estáveis em anel aberto, podem definir-se margens que meçam quanto longe está o sistema em anel fechado de se tornar instável. Em particular, é útil poder saber quanta fase ou quanto ganho pode ser introduzido no anel aberto do sistema sem que este fique instável, que é equivalente a definir uma margem de fase ϕ_{MF} e uma margem de ganho K_{MG} , respetivamente. A Figura 7.1 ilustra este conceito apontando os pontos do diagrama de Nyquist que são candidatos, pela adição de fase ou introdução de ganho, a envolver o ponto (-1) .

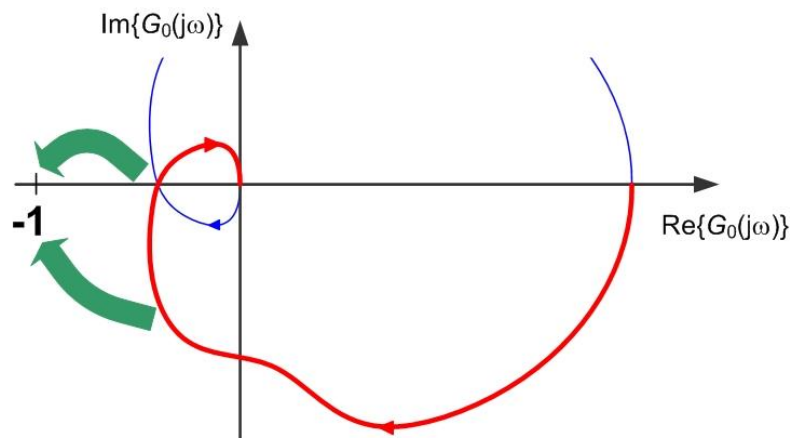


Figura 7.1 – Margens de estabilidade contra a adição de ganho e fase.

7.1. Robustez de estabilidade face a incerteza no ganho e na fase

Eventuais variações de fase e ganho estão associadas à incerteza na fase de modelação, à existência de dinâmica não modelada ou às não linearidades. Por exemplo, um atuador não linear pode apresentar variações de ganho de acordo com o respetivo ponto de funcionamento, por um lado, e a sua dinâmica, não considerada na fase de modelação da instalação, introduzir polos que atrasem mais a fase. A simples colocação de um sensor num sistema

com transporte de massa (e.g. o permutador de calor no pasteurizador) introduz atraso na medida que, como se verá adiante, insere fase negativa no ganho do anel.

Assim, na perspetiva independente da robustez à variação de fase ou de ganho, os dois pontos da resposta em frequência $G_0(j\omega)$, candidatos a serem responsáveis pela instabilidade são, por um lado

$$|G_0(j\omega_{cg})| = 1$$

$$\phi_{MF} = \arg\{G_0(j\omega_{cg})\} - (-180^\circ)$$

ou seja, o ponto que já está à distância unitária (tal como o ponto crítico (-1)), mas que lhe falta um determinado ângulo ϕ_{MF} para envolver o ponto crítico (ver Figura 7.2); e, por outro lado

$$\arg\{G_0(j\omega_{cf})\} = -180^\circ$$

$$K_{MG} = \frac{1}{|G_0(j\omega_{cf})|}$$

ou seja, o ponto que já se apresenta sobre o semieixo real negativo (tal como o ponto crítico (-1)), mas que lhe falta um determinado fator de ganho K_{MG} para envolver o ponto crítico (ver Figura 7.3).

Estas quantidades ϕ_{MF} e K_{MG} são designadas por margem de fase e margem de ganho (de estabilidade do anel fechado), respetivamente. Note-se que a margem de ganho também pode ser representada em dB, como

$$K_{MG(dB)} = -|G_0(j\omega_{cf})|_{dB} = -20 \log_{10}|G_0(j\omega_{cf})|$$

no entanto recomenda-se cautela visto que estas margens estão primeiramente ligadas ao traçado do diagrama de Nyquist onde a representação do ganho em dB não tem lugar.

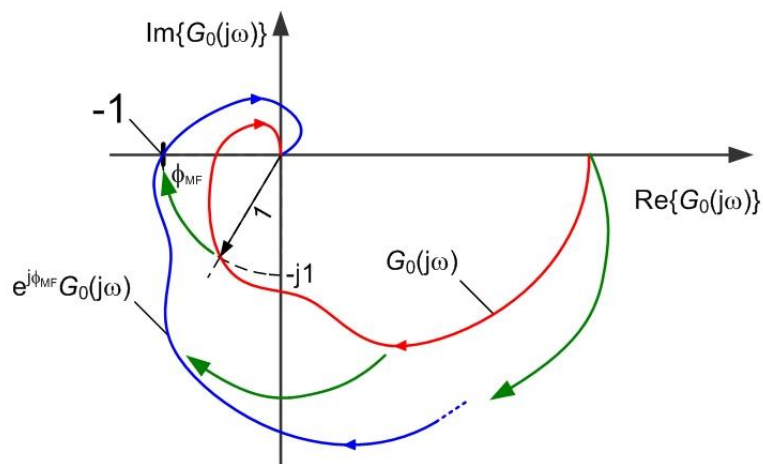


Figura 7.2 – Estabilidade marginal em consequência de adição de fase no diagrama de Nyquist.

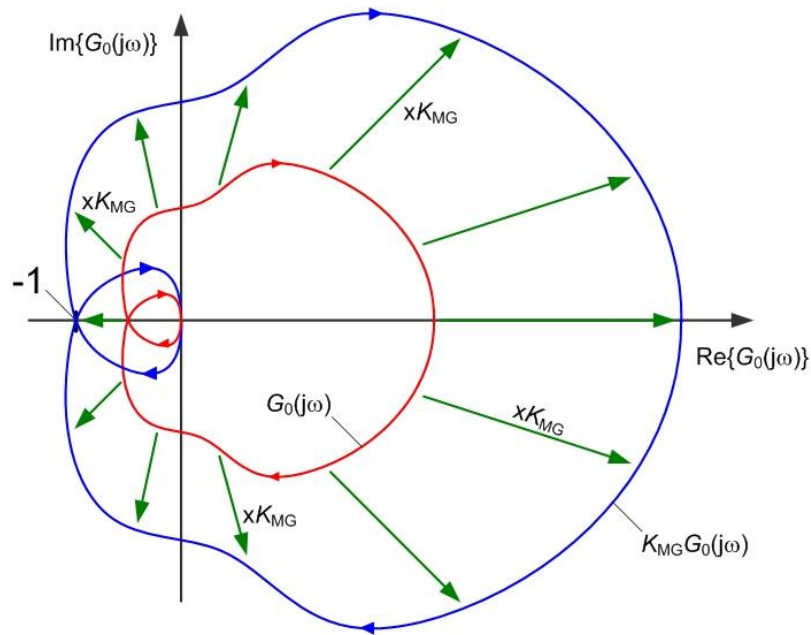


Figura 7.3 – Estabilidade marginal em consequência de adição de ganho no diagrama de Nyquist.

As margens de estabilidade de ganho e fase podem ser transportadas para o diagrama de Bode onde as suas representações e medidas aparecem mais simplificadas. Na Figura 7.4 ilustra-se a marcação das margens de fase e ganho no diagrama de Bode. Assim, para identificação da margem de fase, procura-se o ponto de cruzamento da curva do ganho com o ganho unitário, à frequência ω_{cg} , e regista-se como margem de fase a distância positiva até à fase -180° . Para identificação da margem de ganho, procura-se o ponto de cruzamento da curva de fase com o valor -180° , à frequência ω_{cf} , e regista-se como margem de ganho a “distância positiva” (no sentido logarítmico) até ao ganho unitário.

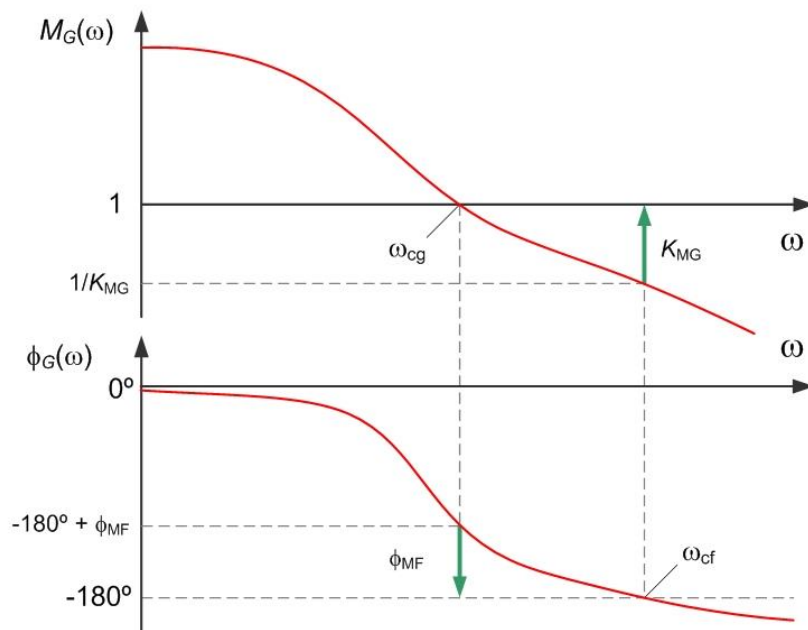


Figura 7.4 – Margens de ganho e de fase no diagrama de Bode.

O sistema em anel fechado será estável se estas margens forem ambas positivas, que implica que a frequência de cruzamento de ganho ω_{cg} deve ser inferior à frequência de cruzamento de fase ω_{cf} . Deve ser óbvio que as conclusões de estabilidade que se tiram da análise destes dois parâmetros são concordantes; o sistema em anel fechado não pode ser, simultaneamente, estável e não estável.

7.2. Efeito da introdução de atraso puro na estabilidade

Referiu-se já que uma causa comum de variações de fase (negativas) é a existência de um atraso puro no anel aberto do sistema. Para dar mais um exemplo, nos sistemas de controlo remoto cujo canal de transmissão introduz atraso (e.g. controlo de veículos submarinos usando ultrassons) é necessário realizar localmente o sistema de retroação que garante a estabilidade. Expliquemos o porquê.

Considere-se a introdução de um tempo-morto τ na dinâmica de um sistema, ou seja, de um bloco de atraso puro (Figura 7.5) entre a excitação na entrada e a resposta na saída. Se a entrada deste bloco for $y(t)$ a saída será

$$y_\tau(t) = y(t - \tau)$$

ao qual aplicando a Transformação de Laplace obtém-se

$$Y_\tau(s) = e^{-\tau s} \cdot Y(s)$$

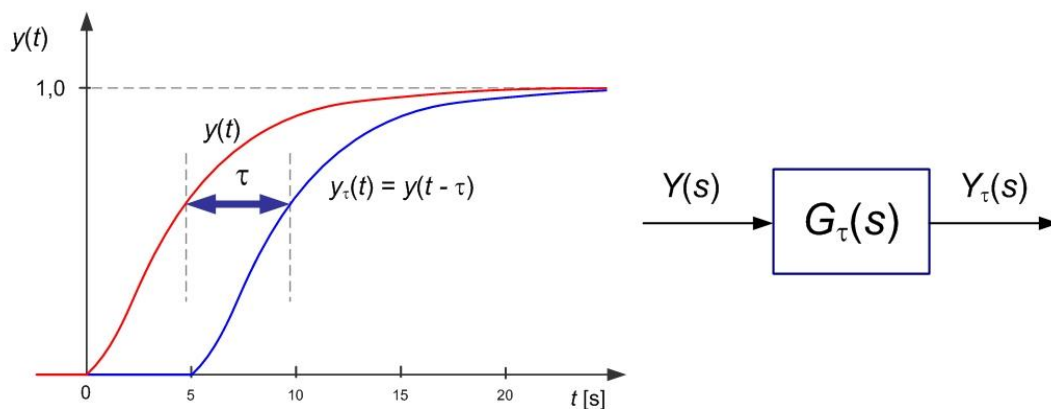


Figura 7.5 – Atraso puro.

A resposta em frequência de um atraso é, então, dada por

$$G_\tau(j\omega) = e^{-j\omega\tau}$$

com

$$M_{G_\tau}(\omega) = |e^{-j\omega\tau}| = 1 = 0 \text{ dB}$$

$$\phi_{G_\tau}(\omega) = \arg\{e^{-j\omega\tau}\} = -\omega\tau$$

ou seja, não tem qualquer efeito na resposta de ganho de amplitude, mas introduz fase negativa que reduz as margens de estabilidade do sistema em anel fechado. Esta degradação da fase é dada por

$$\Delta\phi_\tau = -\omega\tau = -10^{\log\omega} \cdot \tau$$

que no diagrama de Bode com escala logarítmica de frequência tem um crescimento negativo exponencial, como se ilustra na Figura 7.6.

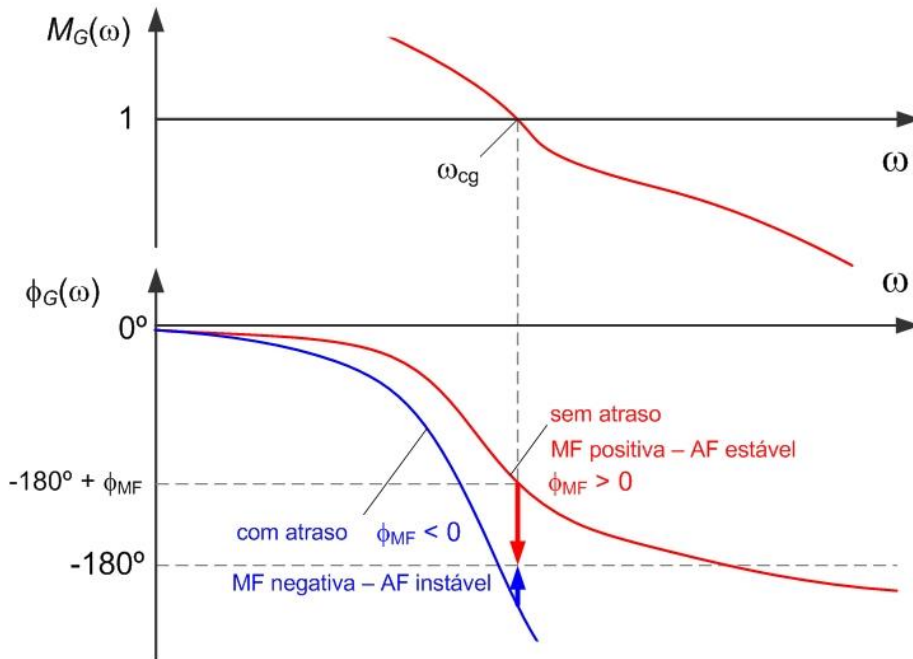


Figura 7.6 – Efeito possível na margem de fase pela introdução do atraso puro.

Exemplo 7.1

Considere-se o sistema com o anel aberto representado pela função de transferência

$$G_0(s) = F(s)C(s) = \frac{10^3}{s^2 + 100s}$$

O traçado do seu diagrama de Bode (que é deixado como exercício) resulta numa margem de fase com o valor aproximado de 85° (i.e. 1.5 rad) e uma frequência de cruzamento $\omega_{cg} \approx 10$ rad/s.

Um problema clássico trata de determinar qual o valor máximo de atraso τ que pode ser inserido no anel aberto sem o sistema em anel fechado perca a estabilidade. Neste caso, temos então

$$\tau_{max} = \frac{\phi_{max}}{\omega_{cg}} = \frac{1.5}{10} = 0.15 \text{ s}$$

que significa que, se a este sistema fosse introduzido um atraso de 0.15 s, o sistema em anel fechado estaria no limiar da estabilidade.

■

7.3. Crítica às margens de ganho e fase

Apesar de bastante úteis, as margens de ganho e de fase devem ser observadas com sentido crítico.

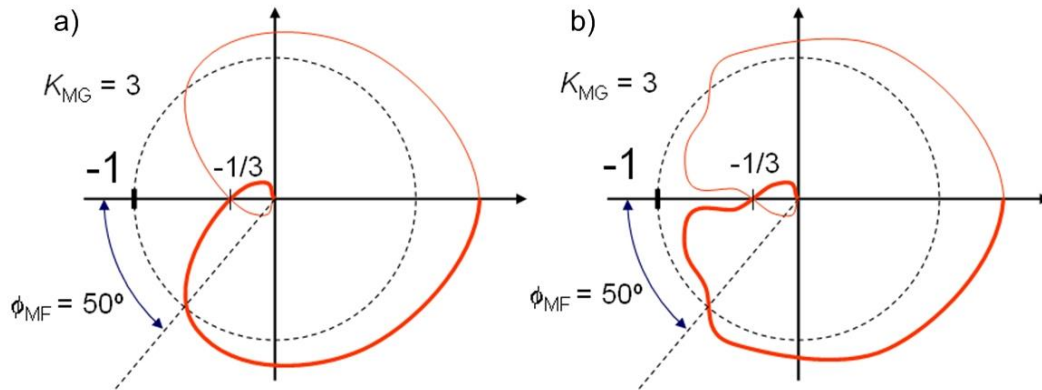


Figura 7.7 – Margens de ganho e fase iguais para diagramas de Nyquist diferentes.

Os dois exemplos da Figura 7.7, têm as mesmas margens de estabilidade mas o primeiro é mais tolerante a uma variação simultânea de ganho e fase pois todo o diagrama de Nyquist está mais afastado do ponto crítico.

7.4. Largura de banda e rapidez de resposta

Considerando a dependência do ganho de um sistema dinâmico com o valor da frequência, pode dividir-se a as gamas de frequências por bandas passantes e de rejeição, ou seja, gama de frequência em que o sinal não é atenuado (ou mesmo amplificado) e gama de frequência em que o sinal de entrada sofre uma forte atenuação para a saída.

Tipicamente, um sistema de controlo tem uma característica de ganho passa-baixo. Uma zona de baixas frequências em que o ganho é aproximadamente constante (se for de seguimento deseja-se unitário) e a partir de um dado valor de frequência uma zona onde o ganho apresenta um forte decaimento, como se representa na Figura 7.8.

Designa-se por largura de banda de um sistema o valor de frequência máximo da gama de frequências passantes.

A situação mais habitual é ter para o anel aberto ganhos elevados (amplificação) à baixa frequência e ganhos reduzidos (atenuação) à alta frequência. Isto reflete a incapacidade natural dos sistemas físicos responderem a solicitações progressivamente mais rápidas, ou seja, à inércia. A largura de banda de um sistema final em anel fechado pode ser prevista a

partir do anel aberto (controlador mais instalação) através da frequência de travessia do ganho unitário ω_{cg} (frequência de cruzamento de ganho). A tabela seguinte apresenta a relação entre os ganhos do anel aberto e anel fechado para as gamas de frequências muito menores e muito maiores que ω_{cg} .

Tabela de relações de ganhos do anel aberto e do anel fechado.

Gama de frequências	Anel aberto $G_0(s)$	Anel fechado $H(s)$
$\omega \ll \omega_{cg}$	$ G_0(j\omega) \gg 1$	$ H(j\omega) = \left \frac{G_0(j\omega)}{1 + G_0(j\omega)} \right \approx 1$
$\omega = \omega_{cg}$	$ G_0(j\omega_{cg}) = 1$	$ H(j\omega) = \left \frac{G_0(j\omega)}{1 + G_0(j\omega)} \right = \frac{1}{ 1 + e^{j \cdot \arg\{G_0(j\omega_{cg})\}} } \geq 0.5$
$\omega \gg \omega_{cg}$	$ G_0(j\omega) \ll 1$	$ H(j\omega) = \left \frac{G_0(j\omega)}{1 + G_0(j\omega)} \right \approx G_0(j\omega) \ll 1$

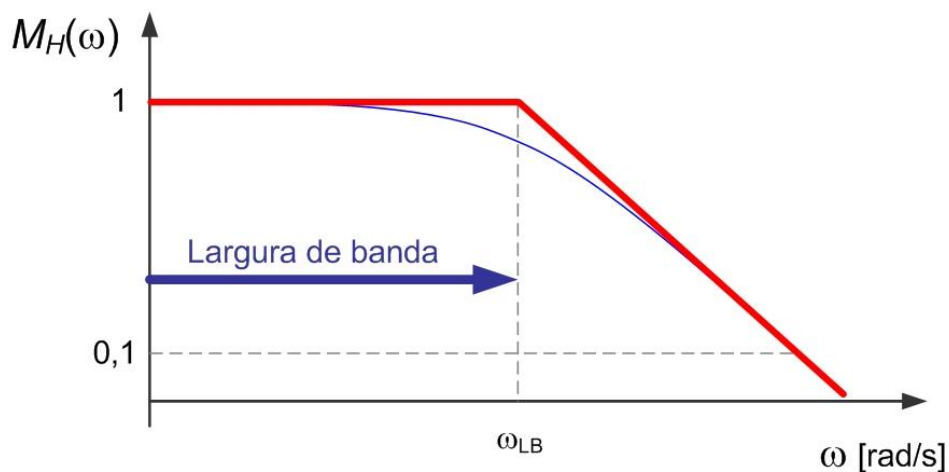


Figura 7.8 – Largura de banda no diagrama de Bode do anel fechado.

A largura de banda do anel fechado fica então estabelecida pela frequência de cruzamento de ganho do anel aberto ω_{cg} . Com o aumento da largura de banda do sistema em anel fechado a resposta no domínio do tempo fica mais rápida como se ilustra no exemplo seguinte.

Exemplo 7.2

Considere-se a instalação representada pela função de transferência

$$KG_0(s) = K \frac{2(s + 10)}{s(s + 1)(s + 100)}$$

A Figura 7.9 representa o diagrama de Bode de amplitude do anel aberto, com ganho K igual a 1; 5 e $5^2 = 25$, tal que a frequência de cruzamento de ganho aumenta progressivamente com K .

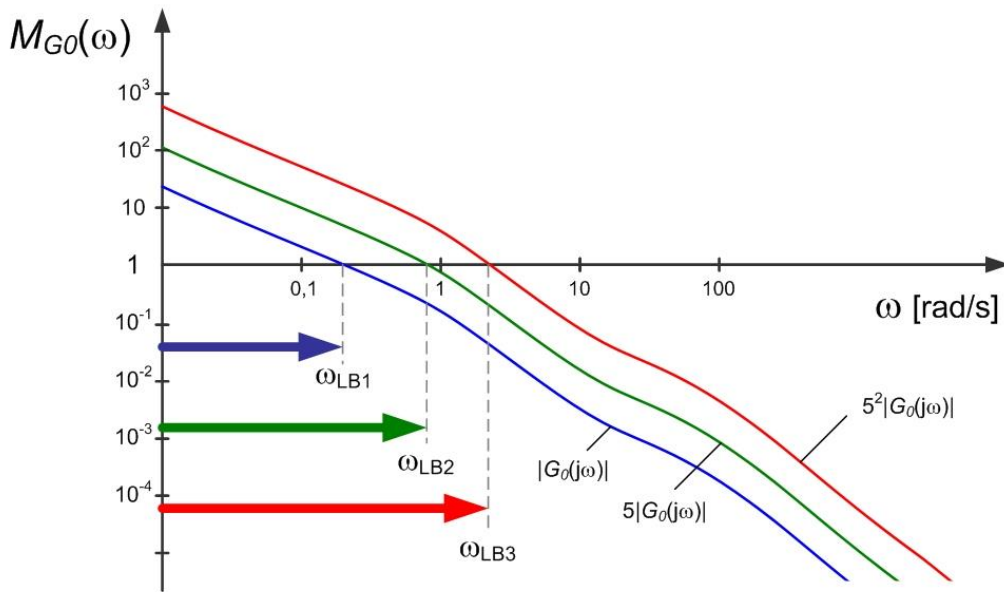


Figura 7.9 – Diagrama de Bode de amplitude relativo ao Exemplo 7.2.

Fechando o anel e excitando o sistema com um sinal de referência em degrau (unitário) obtêm-se as respostas no tempo representadas na Figura 7.10. Aqui, observa-se que o tempo de subida é progressivamente menor com o aumento da frequência de cruzamento de ganho, consequência do aumento do ganho K , no anel aberto. Simultaneamente, é possível observar que existe uma redução da estabilidade relativa. Apesar de não termos representado o diagrama de fase, normalmente, com o cruzamento da amplitude unitária a valores de frequência mais altos existe uma redução da margem de fase, que se reflete num aumento da sobrelevação na resposta temporal.

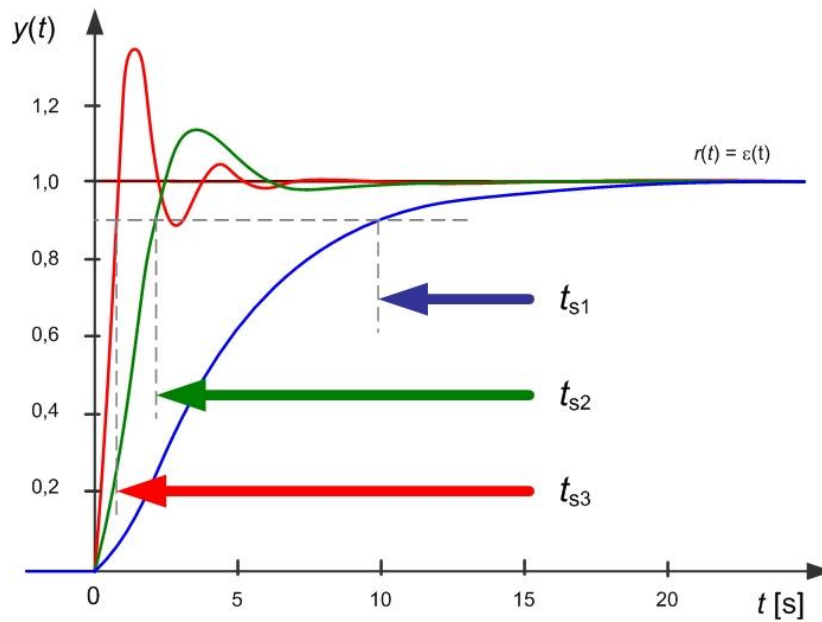


Figura 7.10 – Resposta ao degrau do sistema em anel fechado (Exemplo 7.2).

■

O exemplo anterior mostra o aumento da rapidez da resposta do sistema em anel fechado com o aumento da largura de banda, mas observa-se também alguma degradação da estabilidade do sistema. Este efeito já observado e explicado antes será desenvolvido mais à frente no contexto da resposta em frequência.

7.5. Especificação no domínio da frequência

O projeto de controladores no domínio da frequência tem, como anteriormente, um ponto de partida na especificação do comportamento pretendido para o anel fechado. Neste caso, a especificação tem de ser adaptada a condições verificáveis diretamente na resposta em frequência (diagrama de Bode do anel aberto).

- Seguimento – consegue-se estabelecendo condições de ganho do anel aberto suficientemente elevado na baixa frequência, i.e.

$$\lim_{\omega \rightarrow 0} |G_0(j\omega)| \gg 1$$

No capítulo 0 impôs-se condições para o seguimento sem erro da referência constante (degrau) que resultou na presença de efeito integral no anel aberto. O efeito integral garante ganho infinito à baixa frequência.

- Rapidez de resposta – viu-se na secção anterior estar ligada ao valor mínimo da largura de banda estabelecida pela frequência de cruzamento de ganho do anel aberto, i.e.

$$\omega_{cg} = \omega_{LB}$$

O tempo de subida (aproximadamente igual ao tempo de estabelecimento se o amortecimento for reduzido) é inversamente proporcional à largura de banda e da ordem de

$$t_r \approx \frac{2.5}{\omega_{LB}}$$

- Amortecimento e estabilidade – são garantidos por valores positivos mínimos das margens de fase e ganho, i.e.

$$\phi_{MF} \geq \phi_{MF,min} > 0$$

$$K_{MG}(dB) \geq K_{MG,min}(dB) > 0 \text{ dB}$$

As margens de estabilidade têm um efeito direto no amortecimento da resposta temporal. Quanto menor forem as margens de ganho e fase, menos amortecida é a resposta temporal do sistema. A fórmula seguinte (a esquecer!) relaciona a margem de fase de um sistema (estabelecida no anel aberto) com o fator de amortecimento D de um sistema de 2ª ordem

$$\phi_{MF} = \text{atan} \left\{ \frac{2D}{\sqrt{\sqrt{1 + 4D^4} - 2D^2}} \right\}$$

Isto significa que o sistema em anel fechado se comportará, no domínio do tempo, como um sistema de 2ª ordem com este fator de amortecimento. O comportamento desta relação está representado na Figura 7.11 e pode ser aproximado pela equação mais simples

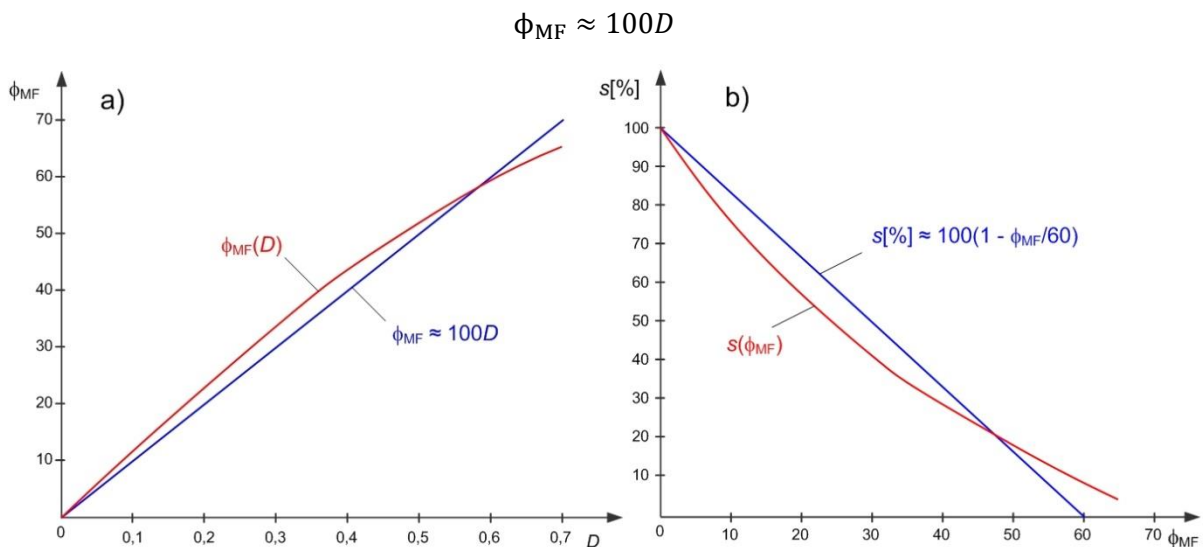


Figura 7.11 – Relação entre margem de fase, fator de amortecimento (a) e sobrelevação (b).

As secções seguintes apresentam dois mecanismos clássicos de compensação da instalação de modo a cumprir especificações básicas no domínio da frequência.

7.6. Compensação de atraso

A forma de compensação mais básica é a compensação de atraso que consiste em adicionar um par polo-zero, ambos no semiplano complexo esquerdo, tal que o polo se encontra mais próximo da origem.

A função de transferência do compensador de atraso é dada por

$$C_{at}(s) = \frac{K}{a} \cdot \frac{aTs + 1}{Ts + 1}$$

com $a < 1$. Nesta situação, o compensador apresenta um zero em $-1/(aT)$ e um polo em $-1/T$. Assim, no diagrama de Bode, a influência do polo realiza-se a um valor de frequência mais reduzido que o relativo à influência do zero. A Figura 7.12 representa do diagrama polos-zeros e o diagrama de Bode do compensador de atraso genérico.

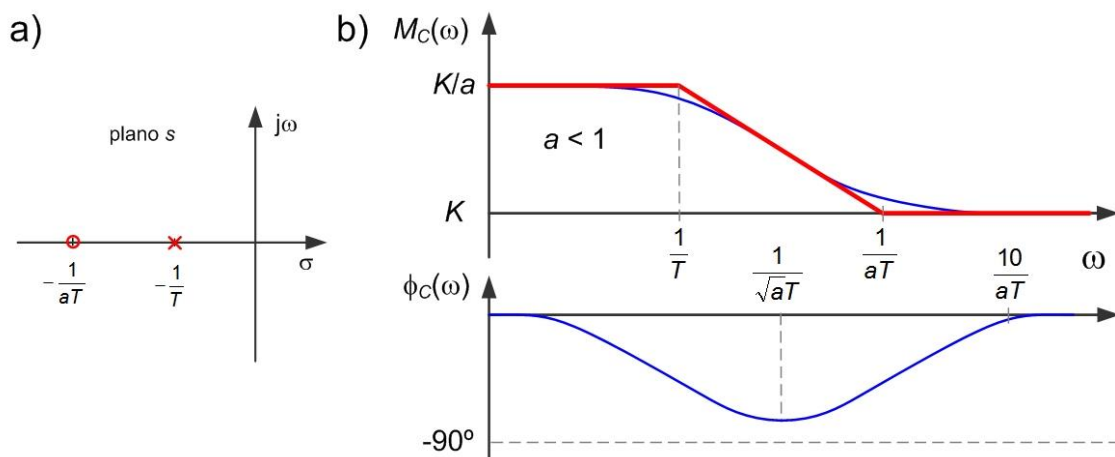


Figura 7.12 – Compensador de atraso. Diagrama polos-zeros (a) e Diagrama de Bode (b).

Este compensador realiza uma aproximação conservadora através do aumento do ganho à baixa frequência, sem modificar a frequência de cruzamento de ganho, ou seja, tem por objetivo cumprir especificações de seguimento mas sem aumentar a largura de banda. Este compensador é designado de atraso porque subtrai fase ao anel aberto.

A Figura 7.13 representa os vários passos do dimensionamento do compensador de atraso. O primeiro passo consiste no cumprimento duma especificação de margem de ganho. Procura-se o ponto do diagrama de fase que se encontra à distância ϕ_{MF}^* de -180° , ou seja, a frequência para qual a fase vale $-180^\circ + \phi_{MF}^*$. De seguida, corrige-se o ganho do anel aberto com o parâmetro K , do compensador, de forma a fazer, dessa frequência, a nova frequência de cruzamento ω_{cg}^* . Note-se que, no caso deste método de compensação, a largura de banda é consequência deste passo e, portanto, não se encontra livre para especificação.

O passo seguinte corrige o ganho estático do anel aberto de modo a cumprir uma especificação de seguimento. O ganho adicionado por este compensador vale K/a . Como K se encontra já dimensionado pelo passo anterior, determinamos o parâmetro a de modo a obter

$$C_{at}(0)F(0) = \frac{K}{a}F(0) = G_0^*(0)$$

onde $G_0^*(0)$ representa a especificação de ganho à baixa frequência.

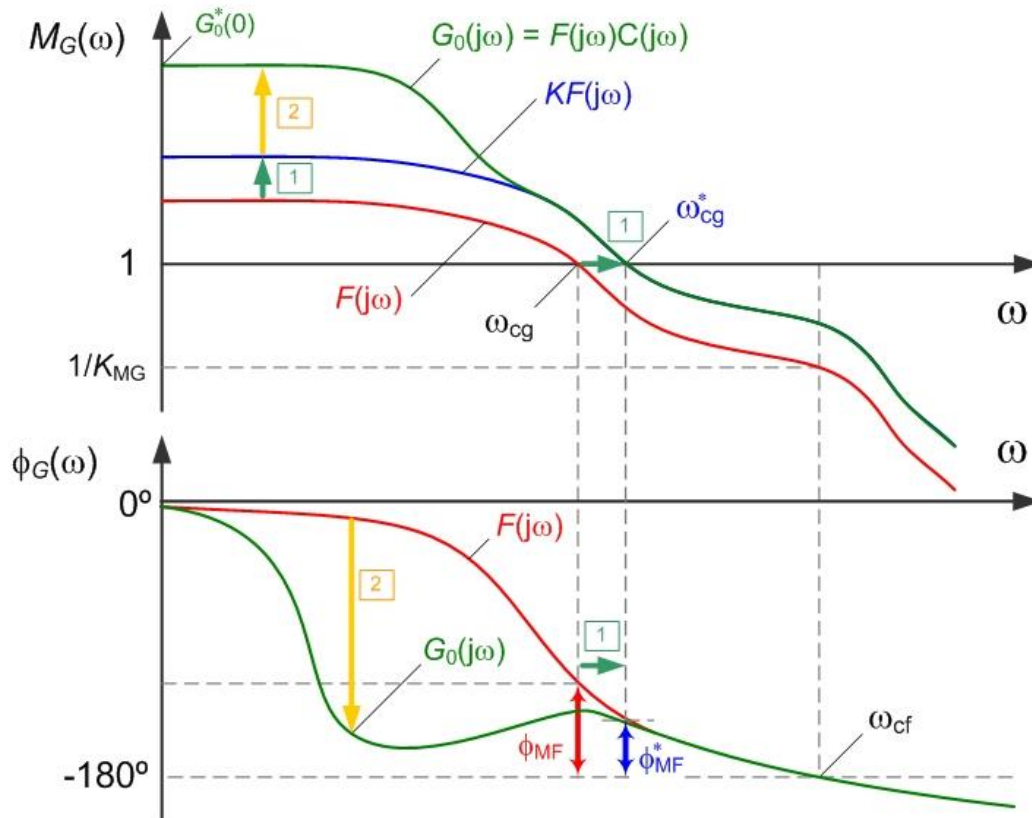


Figura 7.13 – Dimensionamento do compensador de atraso.

O dimensionamento do parâmetro T , que sobra, é feito de modo a não danificar o primeiro passo, i.e. sem modificar a frequência de cruzamento de ganho nem reduzir a margem de fase. Como a influência do zero se faz sentir até uma década acima da respetiva frequência de canto, temos de colocar o zero numa década abaixo de ω_{cg}^* , ou seja

$$\frac{1}{aT} = \frac{\omega_{cg}^*}{10}$$

Como o parâmetro a já foi seleccionado no passo anterior, determina-se o valor do tempo T .

Como já referido, a compensação de atraso não possibilita o dimensionamento independente da margem de fase de estabilidade do anel fechado e a da respetiva largura de banda. Normalmente, privilegia-se o primeiro.

7.7. Compensação de avanço

Nos casos em que o seguimento não é um problema (e.g. porque a instalação tem originalmente comportamento integral ou porque já foi compensada por um compensador de atraso), existe um método básico alternativo à compensação de atraso que permite estabelecer a margem de fase e a largura de banda de forma independente.

A função de transferência do compensador de avanço é dada por

$$C_{av}(s) = \frac{K}{\sqrt{a}} \cdot \frac{aTs + 1}{Ts + 1}$$

agora, com $a > 1$. Também aqui, o compensador apresenta um zero em $-1/(aT)$ e um polo em $-1/T$. No entanto, com $a > 1$, no diagrama de Bode, a influência do zero realiza-se a um valor de frequência mais reduzido que o relativo à influência do polo. A Figura 7.14 representa do diagrama polos-zeros e o diagrama de Bode do compensador de avanço genérico.

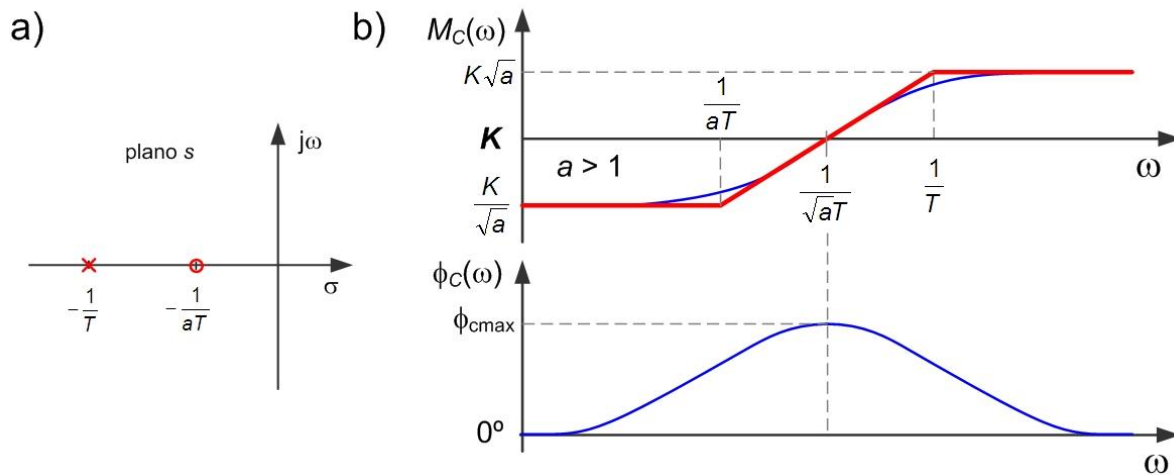


Figura 7.14 – Compensador de avanço. Diagrama polos-zeros (a) e Diagrama de Bode (b).

Este compensador tem uma aproximação ambiciosa através do aumento da fase à frequência de cruzamento de ganho de modo a possibilitar o aumento de largura de banda sem perda de estabilidade relativa. Este compensador é designado de avanço porque adiciona fase positiva ao anel aberto.

Na Figura 7.15 pode observar-se os vários passos do dimensionamento do compensador de avanço. O primeiro passo consiste no cumprimento dum especificação de largura de banda. Identifica-se o valor do ganho de amplitude à frequência correspondente à largura de banda pretendida, normalmente inferior à unidade, e estabelece-se o ganho K do compensador igual ao inverso desse ganho

$$K = \frac{1}{|F(j\omega_{cg}^*)|}$$

tal que o conjunto $KF(j\omega)$ tenha frequência de cruzamento igual a ω_{cg}^* . Este aumento de largura de banda é acompanhado por uma redução da margem de fase e é aqui que o compensador de avanço dá a sua contribuição principal. O passo seguinte é comparar o novo valor de margem de fase e verificar a diferença para a margem de fase especificada. Esta diferença vai ser solicitada ao compensador através do dimensionamento do parâmetro a . O valor máximo de fase adicionado pelo compensador de avanço vale (demonstre como exercício)

$$\phi_{c,\max} = a \sin \left\{ \frac{a-1}{a+1} \right\}$$

como se observa na Figura 7.16, ou, de forma inversa, o parâmetro a a seleccionar para obter um determinado valor de fase máxima $\phi_{c,\max}$ é dado por

$$a = \frac{1 + \sin \phi_{c,\max}}{1 - \sin \phi_{c,\max}}$$

onde (ver novamente a Figura 7.15)

$$\phi_{c,\max} = -180^\circ + \phi_{MF}^* - \phi_F(j\omega_{cg}^*)$$

Falta só, como último passo, garantir que a fase máxima é adicionada no valor de frequência correto, ou seja,

$$\frac{1}{\sqrt{a}T} = \omega_{cg}^*$$

Como o parâmetro a foi já estabelecido pelo passo anterior, esta expressão estabelece o valor do parâmetro T , ficando o compensador dimensionado.

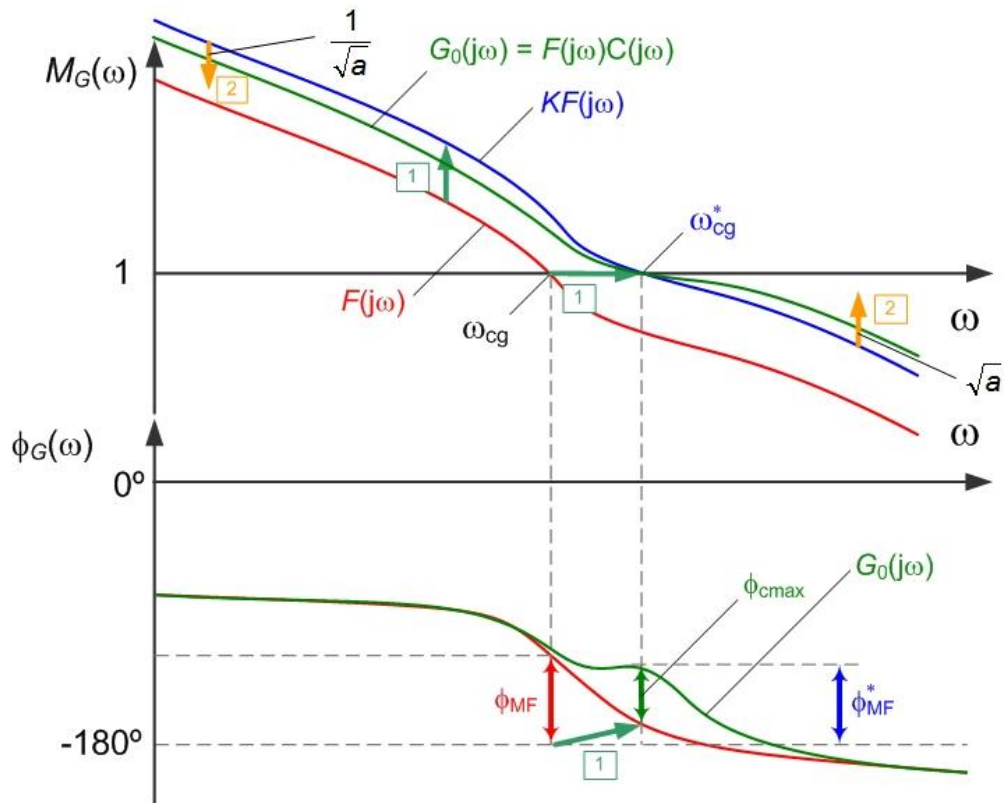


Figura 7.15 – Dimensionamento do compensador de avanço.

Como já referido, a compensação de avanço não possibilita o estabelecimento (independente) do ganho à baixa frequência para efeitos de especificação de seguimento.

Note-se ainda que, apesar de um conjunto zero-polo poder adicionar até 90° , o compensador de avanço não é eficaz na recuperação de valores de fase superiores a cerca de 45° ($a < 6$), como se observa na Figura 7.16. Para recuperar valores de fase superiores é preferível colocar dois compensadores em cascata distribuindo a fase a compensar.

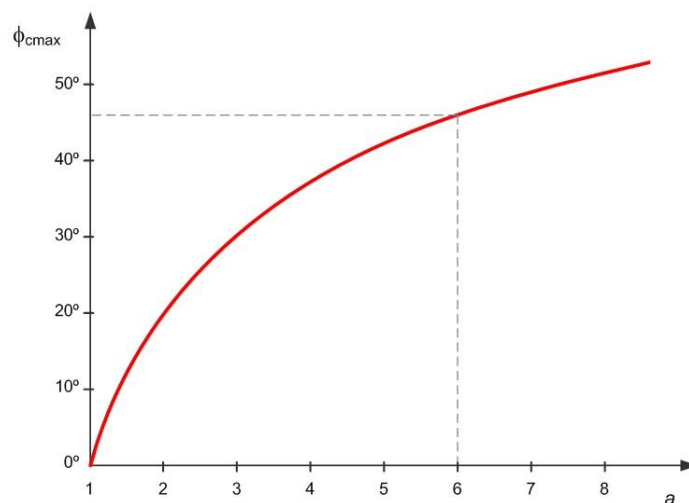


Figura 7.16 – Fase máxima do compensador de avanço em função do parâmetro a .

Note-se que o compensador reduz o ganho à baixa frequência e aumenta o ganho à alta frequência. Na frequência de travessia, enquanto a instalação contribui com ganho decrescente, o compensador contribui com ganho crescente (1:1 por cada compensador). Se a instalação tiver um decaimento de ganho fraco à frequência de travessia pode acontecer uma situação de patamar ou mesmo de ganho crescente com a frequência, aumentando a sensibilidade do sistema a variações simultâneas de ganho e fase.

8. Controlo PID

Existe uma estrutura de controlo clássica amplamente utilizada designada por controlo de três termos ou PID de proporcional, integral e derivativo. Os controladores baseados nesta estratégia cobrem uma parcela muito significativa na regulação dos processos e produtos industriais. Adicionalmente, existe uma bibliografia infindável sobre diversos aspetos deste tema.

O controlo PID resulta da ação combinada (em sobreposição) das três componentes seguintes:

- Proporcional – em que a ação de controlo é proporcional ao erro no instante de tempo atual;
- Integral – onde a ação de controlo é proporcional ao integral do erro passado; e
- Derivativa – cuja ação de controlo é proporcional à derivada (tendência) do erro no instante presente.

De forma analítica, a ação de controlo PID é calculada a partir da adição dos três termos correspondentes

$$u(t) = u_P(t) + u_I(t) + u_D(t)$$

definidos por

$$u_P(t) = K_P \cdot e(t)$$

$$u_I(t) = K_I \int_{-\infty}^t e(\sigma) d\sigma$$

$$u_D(t) = K_D \frac{de(t)}{dt}$$

onde

$$e(t) = r(t) - y(t)$$

é o erro de seguimento da referência r pela saída da instalação y .

Aplicando a Transformação de Laplace à lei de controlo PID, pode escrever-se

$$C_{PID}(s) = \frac{U(s)}{E(s)} = K_P + \frac{K_I}{s} + K_D s$$

como representado na Figura 8.1.

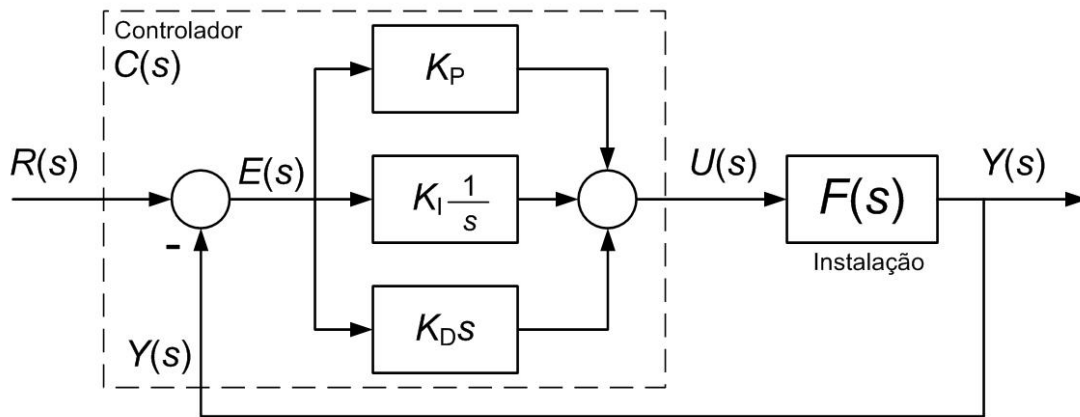


Figura 8.1 – Esquema de blocos do controlador PID em anel fechado.

Um formato alternativo com grande aceitação resulta de colocar o ganho proporcional em evidência e definir duas constantes de tempo T_i e T_d , designadas respectivamente por tempo integral e derivativo, como

$$T_i = \frac{K_P}{K_I} \quad \text{e} \quad T_d = \frac{K_D}{K_P}$$

tomando o controlador a forma

$$C_{PID}(s) = \frac{U(s)}{E(s)} = K_P \left(1 + \frac{1}{T_i s} + T_d s \right)$$

como representado Figura 8.2.

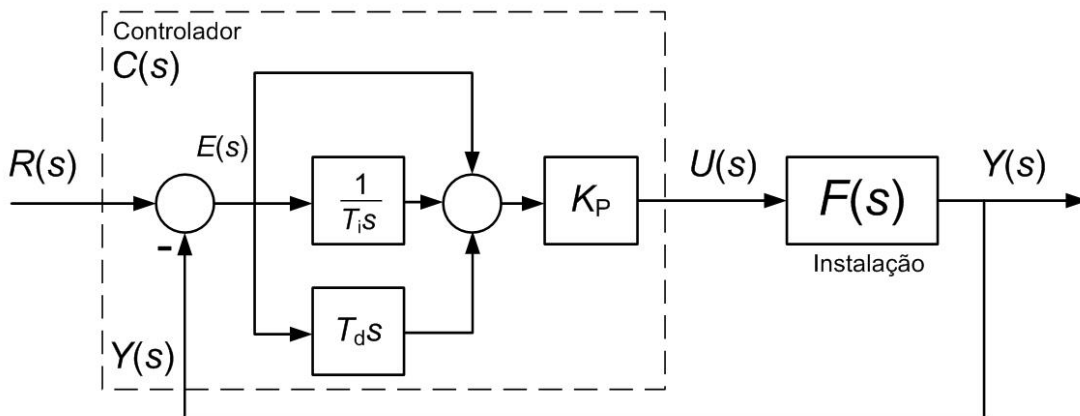


Figura 8.2 – Esquema de blocos do controlador PID com K_P em evidência.

Note-se que, relativamente à componente integral, a anulação do ganho K_I corresponde a levar o valor do tempo integral T_i para infinito.

O controlador PID pode aparecer truncado de alguma das suas componentes, embora seja comum manter sempre a componente proporcional. Assim, é possível encontrar aplicações para os quais a melhor resposta surge de controladores tipo PI, tipo PD ou apenas tipo P. Do

ponto de vista analíticos, estes são apenas casos especiais que surgem da anulação dos ganhos K_I e/ou K_D , no entanto do ponto de vista dinâmico implica a introdução de controladores de ordens diferentes.

Um primeiro comentário a fazer a este controlador com a componente derivativa ativa (T_d diferente de zero) é que não é causal. Na prática, a implementação do controlador exige que a componente derivativa seja aproximada, por adição de um polo de alta frequência, ou seja

$$C_{PID}(s) = \frac{U(s)}{E(s)} = K_P \left(1 + \frac{1}{T_I s} + \frac{T_d s}{\alpha T_d s + 1} \right)$$

onde α é uma constante positiva que estabelece o ganho máximo de alta frequência.

Definida a estrutura do controlador, como seleccionar o valor dos seus parâmetros (normalmente também designados por ganhos do PID)? Para fazer esta seleção, importa compreender o efeito que cada uma das três componentes tem na ação de controlo total.

A componente proporcional depende apenas do erro no instante t presente. No momento exato em que o erro é anulado a sua ação desaparece. O valor de K_P avalia a ação de controlo sobre a instalação por cada unidade presente no erro, positivo ou negativo, como referido nos capítulos anteriores.

A componente integral depende de todo o erro acumulado no passado até ao instante t actual. Esta componente dinâmica só estabiliza quando o erro for igual a zero. No entanto, mesmo com erro nulo à entrada a componente integral pode fornecer uma saída constante diferente de zero, servindo para estabelecer o ponto de operação da instalação.

A componente derivativa gera uma ação de controlo proporcional à derivada do erro, independentemente do seu valor absoluto. Ou seja, introduz capacidade preditiva (ver Figura 8.3) da tendência futura do erro antes que este ocorra e portanto tornando-o mais rápido abrindo a largura de banda do sistema.

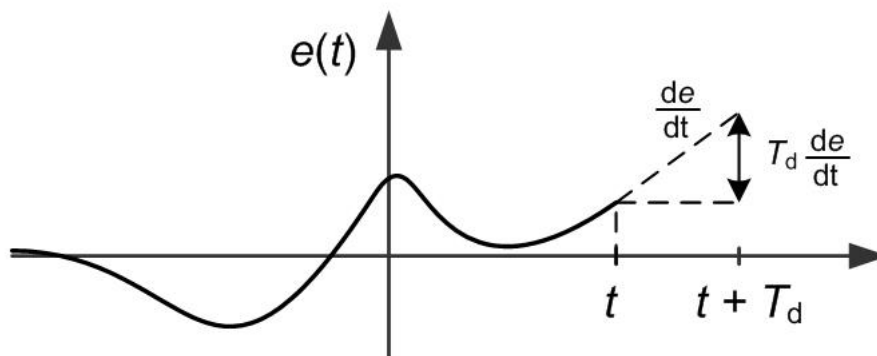


Figura 8.3 – Capacidade preditiva da componente derivativa.

8.1. Ajuste manual de controladores PID

Para além das capacidades desta estrutura de controlo, uma característica apreciada no meio industrial é a relativa facilidade do ajuste do controlador de forma manual, sem o conhecimento específico da função de transferência da instalação.

O procedimento para o ajuste manual desenrola-se em três passos (com a possibilidade de reajuste posterior) com a observação de variações em degrau no valor de referência:

Iniciar a componente proporcional com um valor reduzido do ganho K_P . Aumentar K_P até obter uma velocidade de resposta aceitável (tempo de subida t_r) tolerando alguma sobrelevação e algum erro estático;

Iniciar a componente integral com um valor elevado de T_i . Reduzir T_i até obter a remoção do erro estático num tempo aceitável (tempo de estabelecimento t_s) sem deixar o sistema perder a estabilidade.

Iniciar a componente derivativa com T_d igual a zero. Aumentar T_d até atingir um valor de sobrelevação aceitável ($s\%$).

O exemplo seguinte ilustra a aplicação destes passos a uma instalação a controlar.

Exemplo 8.1

Considere-se uma instalação a controlar com função de transferência desconhecida, mas acessível para ensaios. Coloca-se um controlador PID em série com a instalação e fecha-se o anel, como representado na Figura 8.4.

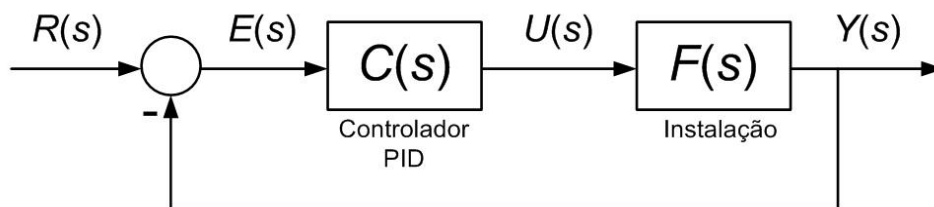


Figura 8.4 – Esquema de blocos do sistema de controlo (Exemplo 8.1).

Aplicando uma sequência de degraus, faz-se variar o ganho proporcional K_P desde 0 até 8.0. Na Figura 8.5 regista-se (em sobreposição) três respostas do sistema para os valores de ganho proporcional de $K_P = 0.5$; $K_P = 2.0$ e $K_P = 8.0$. Com este último valor decidiu-se parar devido ao tempo de subida satisfatório e à crescente sobrelevação. Observe-se que o valor do erro permanente foi diminuindo com o aumento do ganho proporcional, como era esperado.

Num segundo passo, com o valor do ganho proporcional fixado no passo anterior ($K_P = 8.0$), varia-se o valor do tempo integral T_i desde 100 s até 2.0 s. Na Figura 8.6 regista-se as respostas do sistema para os valores de tempo integral de $T_i = 100$ s; $T_i = 4.0$ s e $T_i = 2.0$ s.

Neste último valor parou-se devido ao tempo de estabelecimento satisfatório e, novamente, à oscilação crescente. O erro permanente é removido pela componente integral.

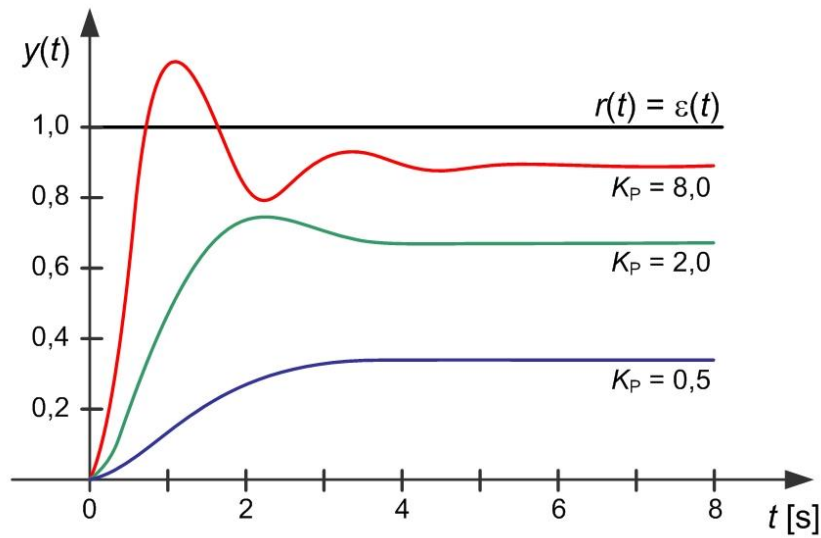


Figura 8.5 – Ajuste manual do PID – ganho proporcional (Exemplo 8.1).

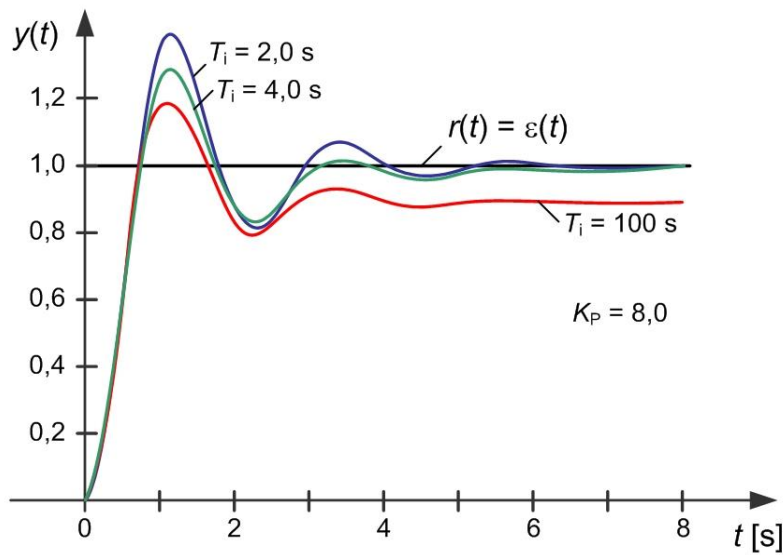


Figura 8.6 – Ajuste manual do PID – tempo integral (Exemplo 8.1).

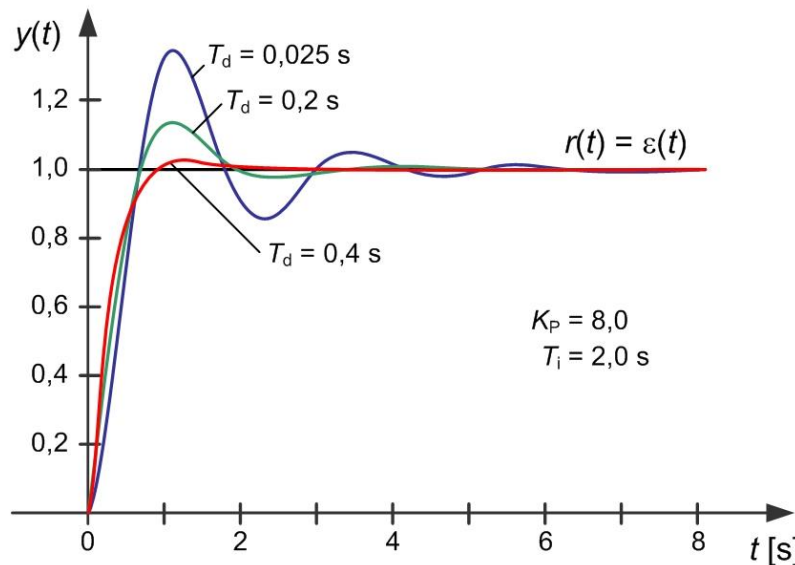


Figura 8.7 – Ajuste manual do PID – tempo derivativo (Exemplo 8.1).

Finalmente, com os valores fixados do ganho proporcional em $K_p = 8.0$ e do tempo integral em $T_i = 2.0$ s, varia-se o valor do tempo derivativo T_d desde 0 até 0.4 s. Na Figura 8.7 registam-se as respostas do sistema para os valores $T_d = 0.025$ s; $T_d = 0.2$ s e $T_d = 0.4$ s. A resposta final é rápida e suave e o controlador PID está ajustado com o conjunto de ganhos

$$K_p = 8.0; \quad T_i = 2.0 \text{ s}; \quad T_d = 0.4 \text{ s}$$

Este processo é repetido as vezes que forem necessárias para atingir o resultado pretendido. ■

Note-se que, no exemplo anterior, em momento algum foi definido qual era o resultado pretendido numa especificação clara. Termos difusos como rápido/lento, suave/oscilatório são empregues para descrever o nível de desempenho do sistema de acordo com os critérios do operador que procede a este ajuste.

8.2. O controlador PID e o lugar das raízes

Se a função de transferência da instalação for conhecida é possível realizar o ajuste do controlador PID de forma sistemática, encarando-o como uma função de transferência com os seus zeros e polos em série com a função de transferência da instalação. A função de transferência do controlador PID (não causal) é dada por

$$C_{PID}(s) = K_p \frac{T_d s^2 + s + 1/T_i}{s}$$

com um polo na origem ($s = 0$) e dois zeros em

$$z_{1,2} = -\frac{1}{2T_d} \pm \frac{\sqrt{1 - 4T_d/T_i}}{2T_d}$$

Através de uma análise do lugar de raízes é possível posicionar os polos do anel fechado para atingir uma determinada especificação de desempenho.

Exemplo 8.2

Considere-se a mesma instalação do Exemplo 8.1, antes desconhecida, agora descrita pela função de transferência

$$F(s) = \frac{1}{(s + 1)^2}$$

Estabeleça-se como especificação um tempo de subida a 90% da ordem de $t_r = 1$ s e sobrelevação mínima. Através de uma análise do lugar de raízes colocaram-se os zeros do PID em -0.8 e em -1.4 . Na Figura 8.8 apresenta-se o diagrama do lugar de raízes e o valor selecionado para o anel fechado.

A resposta final, representada na Figura 8.9 cumpre a especificação e o controlador PID é caracterizado pelo conjunto de ganhos

$$K_p = 4.4; \quad T_i = 1.96 \text{ s}; \quad T_d = 0.45 \text{ s}$$

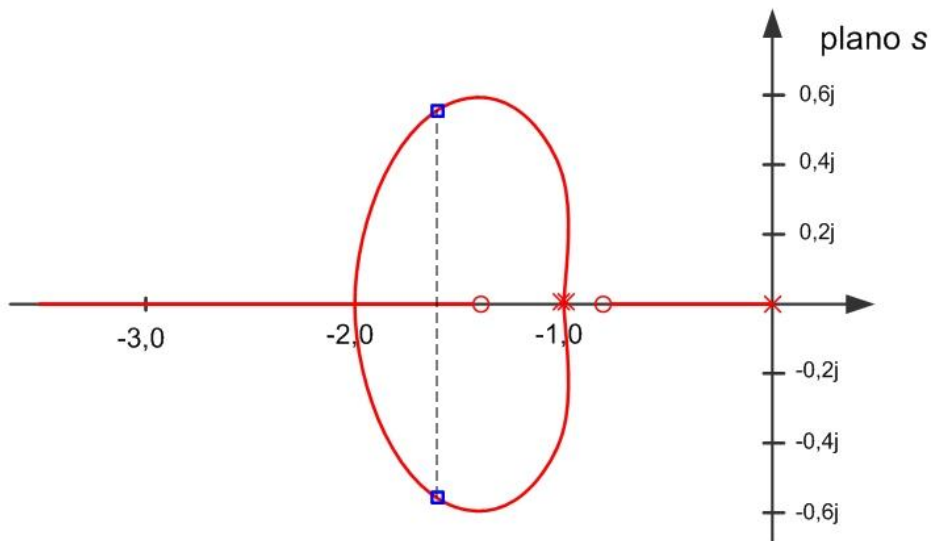


Figura 8.8 – Diagrama do lugar de raízes (Exemplo 8.2).

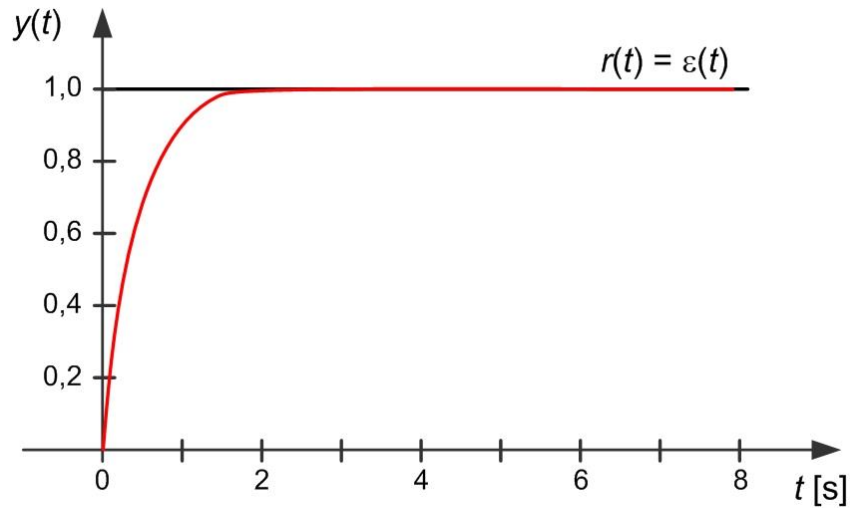


Figura 8.9 – Resposta do sistema de controlo PID (Exemplo 8.2).

■

No entanto, a principal vantagem do controlo PID sobressai quando existe informação reduzida sobre a instalação, não se sabendo com exatidão qual a sua função de transferência.

8.3. Regras de Ziegler-Nichols

Ziegler e Nichols desenvolveram no princípio da década de 40 (do século XX) um conjunto de regras de ajuste sistemático de controladores PID reunido em dois métodos equivalentes e alternativos.

O primeiro método é designado por método da curva de reação. Este método consiste na aplicação de um degrau de amplitude unitária (na prática, só irá interessar o ganho estático) e regista-se a resposta da instalação. Do registo da resposta extraem-se os valores do tempo-morto L , do tempo de subida T e do ganho estático K .

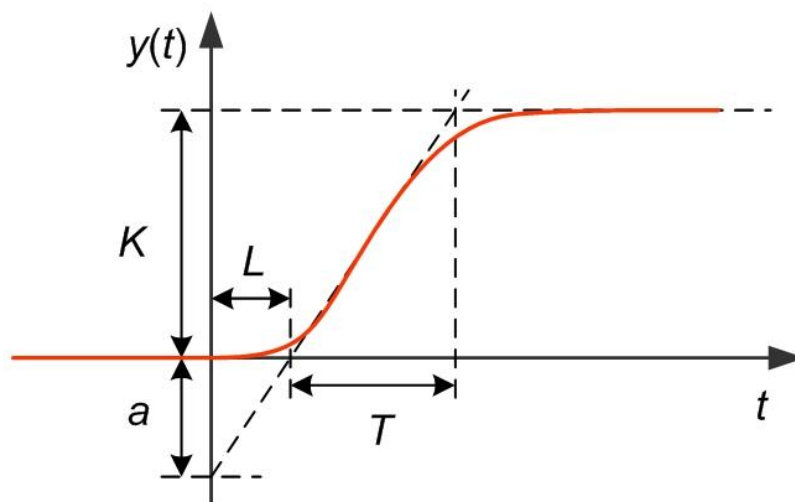


Figura 8.10 – Curva de reação da instalação ao degrau unitário.

A partir deste calcula-se o parâmetro

$$a = K \frac{L}{T}$$

e consulta-se a tabela seguinte para obter os parâmetros do controlador:

Tabela de regras de Ziegler-Nichols – método da curva de reação

Tipo	K_P	T_i	T_d
Controlador P	$\frac{1}{a}$		
Controlador PI	$\frac{0.9}{a}$	$3.3L$	
Controlador PID	$\frac{1.2}{a}$	$2.0L$	$0.5L$

Exemplo 8.3

Considere-se uma instalação à qual foi aplicado um degrau de amplitude unitária na sua entrada $u(t)$ no instante de tempo t igual a zero. A Figura 8.11 regista a curva de reação da instalação e de onde se extraem os valores

$$K \approx 4.0; \quad L \approx 1.0 \text{ s}; \quad \text{e} \quad T \approx 4.0 \text{ s}$$

e em consequência

$$a = K \frac{L}{T} \approx 1.0$$

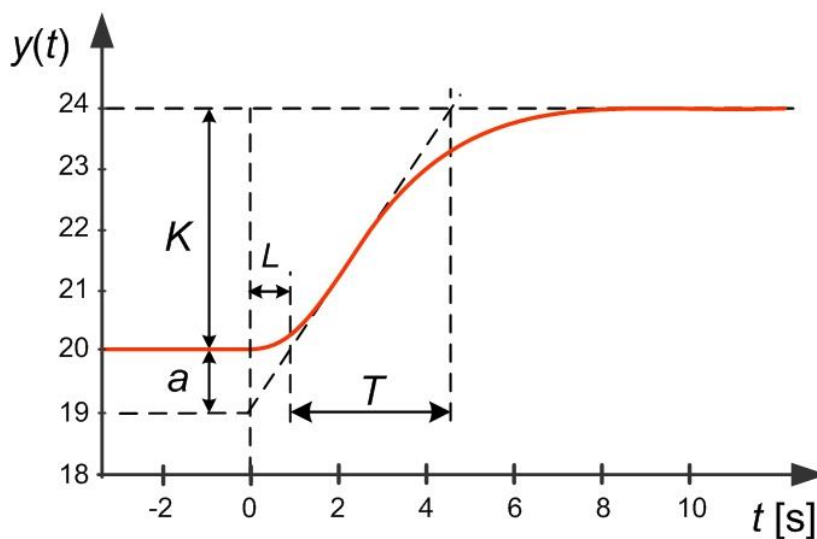


Figura 8.11 – Curva de reação da instalação registada no Exemplo 8.3.

Consultando a tabela anterior de ajuste de ganhos de Ziegler-Nichols, obtém-se

$$K_p = 1.2; \quad T_i = 2.0 \text{ s}; \quad T_d = 0.50 \text{ s}$$

que aplicados ao sistema em anel fechado, resultam na resposta registada na Figura 8.12. Para já, diga-se apenas que a sobrelevação apresentada por esta resposta (aproximadamente 35%) é normal como resultado da aplicação deste método.

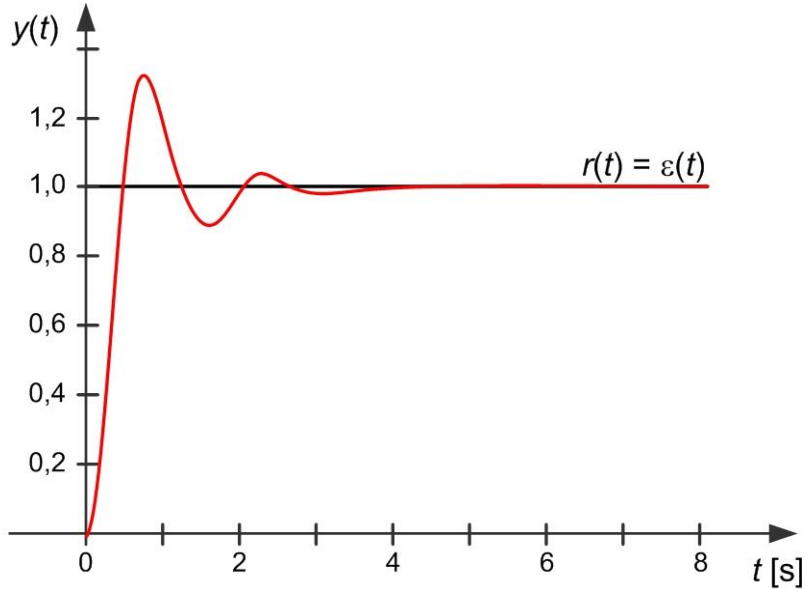


Figura 8.12 – Resposta do sistema em anel fechado (Exemplo 8.3).

■

Como se viu no Exemplo 8.3, não é fácil extrair estas medidas com exatidão.

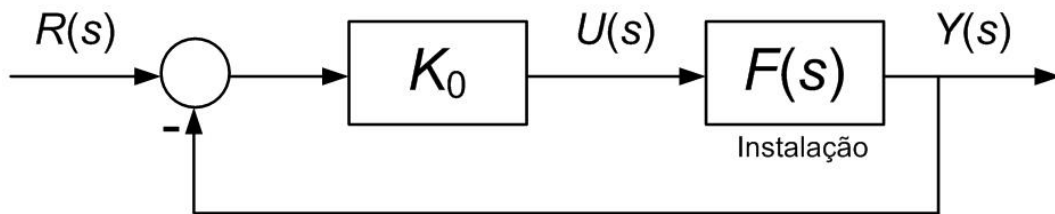


Figura 8.13 – Esquema do ensaio no método da sensibilidade última.

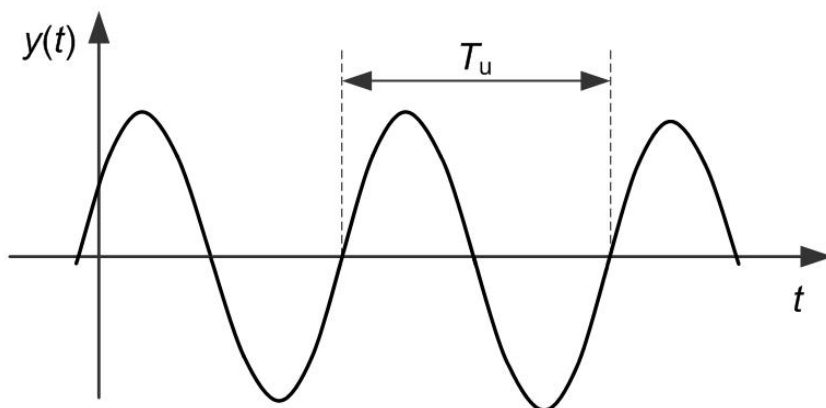


Figura 8.14 – Período crítico da oscilação permanente no método da sensibilidade última.

O método alternativo designa-se por método da sensibilidade última porque consiste em levar a instalação ao limiar da estabilidade. A instalação é colocada em anel fechado com um controlador proporcional, como na Figura 8.13. e aumenta-se o valor do ganho K_0 até atingir o valor crítico K_u correspondente ao limiar da estabilidade. Nesta situação, é normal observar uma oscilação permanente como caracterizada na Figura 8.14. Deste ensaio, extraem-se os valores do ganho crítico K_u e do período crítico T_u e consulta-se a tabela seguinte para obter os parâmetros do controlador:

Tabela de regras de Ziegler-Nichols – método da sensibilidade última

Tipo	K_P	T_i	T_d
Controlador P	$0.50K_u$		
Controlador PI	$0.45K_u$	$\frac{T_u}{1.2}$	
Controlador PID	$0.60K_u$	$\frac{T_u}{2.0}$	$\frac{T_u}{8.0}$

Exemplo 8.4

Considere-se a mesma instalação do Exemplo 8.3, agora colocada em anel fechado com um controlador proporcional. A instalação foi levada ao limiar da estabilidade, atingida com ganho crítico $K_u = 2$, registando-se a resposta representada na Figura 8.15.

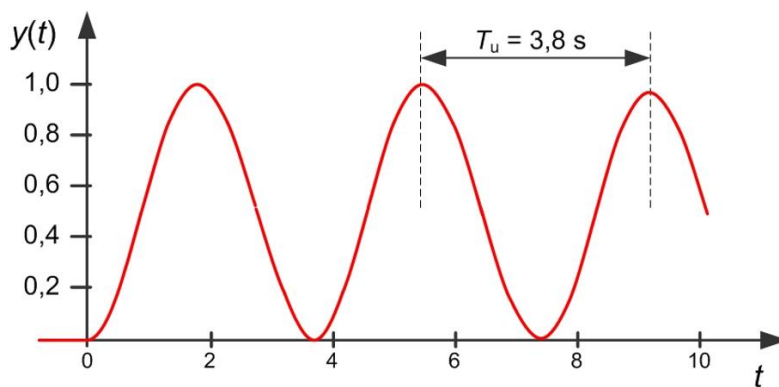


Figura 8.15 – Período crítico da oscilação permanente registada no Exemplo 8.4.

Deste registo extrai-se o valor do tempo crítico $T_u = 3.8$ s. Consultando a tabela anterior de ajuste de ganhos de Ziegler-Nichols, obtém-se

$$K_P = 1.2; \quad T_i = 1.9 \text{ s}; \quad T_d = 0.48 \text{ s}$$

Estes valores são semelhantes aos obtidos no Exemplo 8.3, que aplicados ao sistema em anel fechado, resultam numa resposta equivalente à já registada na Figura 8.12. ■

Como se pode observar nos exemplos os dois métodos são equivalentes. Uma das características dos controladores PID ajustados pelas regras de Ziegler-Nichols é o zero real duplo em

$$z_{1,2} = -\frac{1}{2T_d} \pm \frac{\sqrt{1 - 4T_d/T_i}}{2T_d} = -\frac{4}{T_u}$$

Note-se que o ensaio realizado para determinar os valores do ganho e período críticos é equivalente a encontrar o valor da margem de ganho da instalação K_{MG} e a frequência de cruzamento de fase ω_{cf} . Isto pode ser realizado excitando a instalação com um gerador de sinal sinusoidal de amplitude A_i e frequência ω . Aumenta-se o valor da frequência até a resposta se apresentar em oposição de fase com a entrada. Nesta situação regista-se a amplitude A_o da sinusoide (!) de saída e respetiva frequência. Esta está relacionada com o período crítico através

$$\omega_{cf} = \frac{2\pi}{T_u}$$

e o ganho entrada-saída é o inverso da margem de ganho, ou seja, do ganho crítico

$$K_u = K_{MG} = \frac{A_i}{A_o}$$

Este método requer algum equipamento adicional (menos disponível no tempo de Ziegler e Nichols) mas fornece algo que é crítico para o dono da instalação, a integridade da mesma. Na prática, existe um método que consiste em introduzir um relé limitador em anel fechado com a instalação que devolve com precisão os valores do período e ganho crítico. Deste método se falará mais tarde.

8.4. Análise e modificação do método de Ziegler-Nichols

Os métodos de Ziegler-Nichols são por vezes criticados por apresentarem sobrelevações muito elevadas, como se observou nos exemplos anteriores. No entanto, esta característica não é uma fraqueza do método mas algo que foi deliberadamente selecionado para responder às necessidades da indústria de processos onde a principal preocupação é a rejeição de perturbações com uma largura de banda alargada em detrimento das capacidades de seguimento, mais importantes em outras áreas de aplicação, e.g. aeronáutica ou robótica.

Observe-se que, à frequência de cruzamento de fase ω_{cf} , a resposta em frequência do anel aberto vale

$$G_0(j\omega_{cf}) = F(j\omega_{cf})C(j\omega_{cf}) = -\frac{K_P}{K_u} \left(1 + j \left(\omega_{cf} T_d - \frac{1}{\omega_{cf} T_i} \right) \right)$$

Substituindo nesta expressão os valores dos parâmetros do controlador sugeridos por Ziegler e Nichols (ver Tabela relativa à sensibilidade última), obtém-se

$$G_0(j\omega_{cf}) = F(j\omega_{cf})C(j\omega_{cf}) = -\frac{0.6K_u}{K_u} \left(1 + j \left(\frac{2\pi}{T_u} \cdot \frac{T_u}{8} - \frac{T_u}{2\pi} \cdot \frac{2}{T_u} \right) \right) =$$

$$-0.6 \left(1 + j \left(\frac{2\pi}{8} - \frac{2}{2\pi} \right) \right) = 0.66 \cdot e^{-j155^\circ}$$

Isto significa que, independentemente da instalação, o ponto de cruzamento do diagrama de Nyquist com o semieixo real negativo, em $-1/K_u$, é transferido para aquela posição, como ilustrado na Figura 8.16.

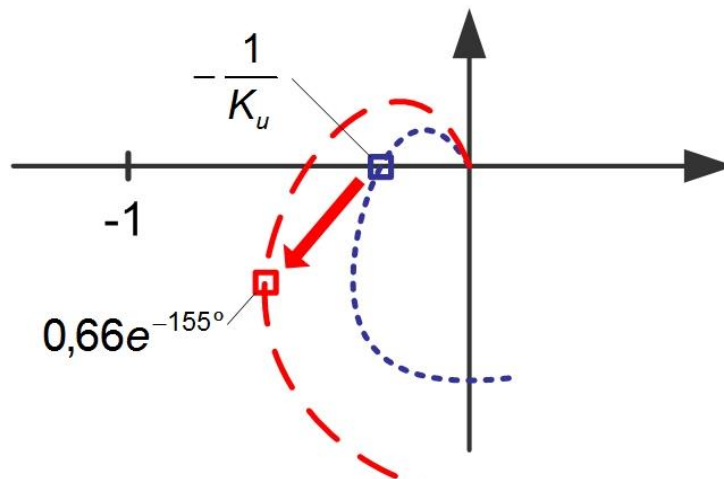


Figura 8.16 – Efeito da aplicação das regras de Ziegler-Nichols no diagrama de Nyquist.

Na realidade, nada obriga a que o ponto correspondente ao cruzamento de fase não seja transportado para outro valor. Mantendo a relação entre os tempos integral e derivativo igual a 4 (que já se viu colocar os zeros do PID sobrepostos no eixo real), pode construir-se um novo conjunto de regras com novos fatores de cálculo, como se apresenta na tabela seguinte:

Tabela de regras de Ziegler-Nichols modificadas (sensibilidade última)

Tipo	K_P	T_i	T_d
Controlador PID	αK_u	$4\beta T_u$	βT_u

Neste caso, o ponto de cruzamento do diagrama de Nyquist com o semieixo real negativo, em $-1/K_u$, é transferido para a posição

$$G_0(j\omega_{cf}) = -\frac{\alpha K_u}{K_u} \left(1 + j \left(\frac{2\pi}{T_u} \cdot \beta T_u - \frac{T_u}{2\pi} \cdot \frac{1}{4\beta T_u} \right) \right) = -\alpha \left(1 + j \left(2\beta\pi - \frac{1}{8\beta\pi} \right) \right)$$

Exemplo 8.5

Considere-se novamente a instalação do Exemplo 8.3. Usando fatores $\alpha=0.4$ e $\beta = 0.2$ (em lugar dos factores de Ziegler-Nicholds de $\alpha=0.6$ e $\beta = 0.125$) transporta-se o ponto de cruzamento para a posição

$$G_0 \left(j \frac{2\pi}{T_u} \right) = -\alpha \left(1 + j \left(2\beta\pi - \frac{1}{8\beta\pi} \right) \right) = -0.4 - j0.4231 = 0.58 \cdot e^{j133^\circ}$$

mais afastada do ponto -1 do diagrama de Nyquist.

Usando os mesmos valores do Exemplo 8.3 para o ganho e período crítico, Consulta-se a tabela anterior de ajuste de ganhos de Ziegler-Nichols modificada obtendo-se

$$K_p = 0.8; \quad T_i = 3.0 \text{ s}; \quad T_d = 0.75 \text{ s}$$

que aplicados ao sistema em anel fechado, resultam na resposta registada na Figura 8.17.

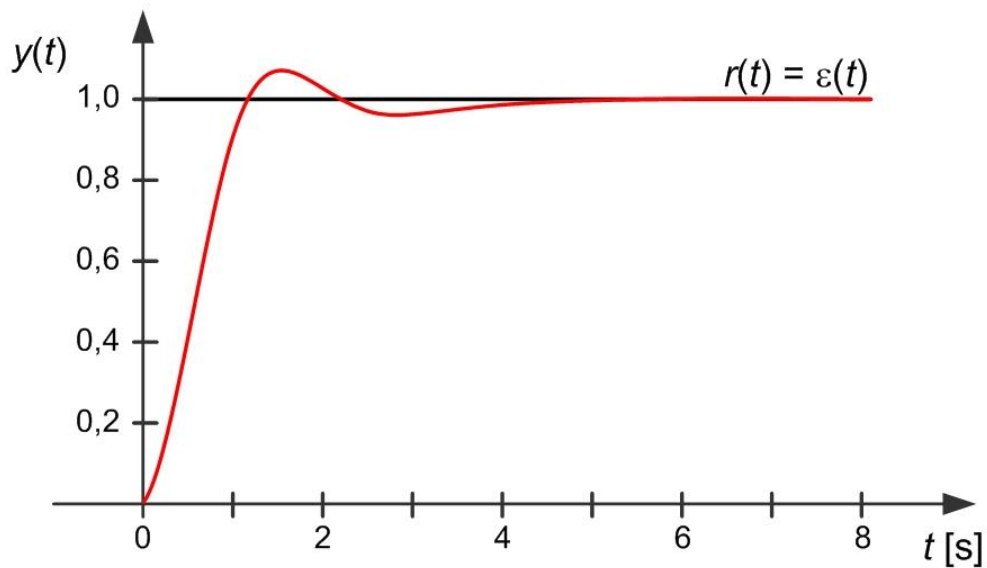


Figura 8.17 – Resposta do sistema em anel fechado (Exemplo 8.5).

■

9. Sistemas em tempo discreto

Neste texto, o desenvolvimento realizado até aqui foi sob o pressuposto que a realização do controlador seria através de uma tecnologia analógica, e.g. montagem eletrónica com amplificadores operacionais. Existem ainda muitas aplicações que requerem uma realização deste tipo, nomeadamente, por razões de contenção de custos. No entanto, atualmente, a realização dos sistemas de controlo é feita maioritariamente através de **sistemas digitais**, leia-se sistemas com microprocessadores.

Existem inúmeras vantagens na utilização da tecnologia digital face à analógica, nomeadamente:

- **a flexibilidade de desempenho.** Existe maior facilidade de adaptar a lei de controlo pela modificação do código e parâmetros do algoritmo de controlo. É mais simples modificar lógica programada (software) que lógica realizada em hardware, i.e. *hard-wired*.
- **a complexidade dos algoritmos.** A lógica programada permite um nível de complexidade nos algoritmos que seria muito difícil de alcançar com realizações em hardware. Pense-se, como exemplo, num controlador avançado que adapta os seus parâmetros, em linha, à medida que vai aprendendo mais sobre a instalação que está a controlar.
- **a exatidão do controlo.** Através do código programado eliminam-se as derivas da eletrónica e aumenta-se a insensibilidade ao ruído. Um valor armazenado numa memória é constante. Um valor estabelecido através da relação entre e.g. duas resistências elétricas é sensível aos mais diversos fatores, incluindo as flutuações no seu processo de fabrico.
- **a capacidade de memorização.** A memória do computador permite complementar o controlador com outras ações específicas, como o arranque de controlo, o registo de dados, entre muitas outras.

Por outro lado, sem contrabalançar as vantagens, podemos enumerar alguns pontos que requerem alguma atenção relativamente ao controlo digital. São estes:

- **a degradação dinâmica.** Nos picos de sinais rápidos a dinâmica do controlador degrada-se devido ao atraso resultante das conversões analógico-digital (A/D) e digital-analógico (D/A).
- **a fiabilidade dos sistemas.** Em tempo real pode ser difícil garantir a fiabilidade dos sistemas de controlo em ambientes industriais devido a problemas de compatibilidade

eletromagnética. Os sistemas digitais, operando com correntes fracas, são sensíveis a ambientes industriais populados de máquinas elétricas. No entanto, houve nos últimos anos um avanço enorme sobre o conhecimento relacionado com as formas de encapsular os sistemas, havendo standards e certificações que dão tranquilidade ao utilizador industrial.

- **a introdução de ruído de quantificação.** Os valores de amplitude não são registados na memória do computador como valores contínuos mas como valores quantificados cuja precisão analógica depende da gama considerada para a variável e o número de bits disponíveis. Por exemplo, uma gama de temperatura de 100 °C digitalizada com palavras digitais de 8 bits introduz um erro de quantificação dado por $100/(2^8) = 0.4$ °C.
- **a velocidade de processamento.** A rapidez de cálculo do computador pode ser insuficiente para produzir o valor da ação de controlo em tempo útil, nomeadamente quando se pretende realizar uma sequência de leitura do sensor, processamento do controlo e escrita para atuação; no mesmo instante de amostragem. Adicionalmente é necessário estabelecer tempos máximos de leitura (A/D) e escrita (D/A) no sistema de aquisição de dados.
- **a incompatibilidade e obsolescência de equipamentos.** Este é talvez o maior problema dos sistemas baseados em computador. A tecnologia digital evolui a um ritmo elevado e não é invulgar um equipamento ficar obsoleto em alguns anos. Adicionalmente, os sistemas operativos e as linguagens de programação vão sendo substituídas por outras mais evoluídas e a manutenção de sistemas mais antigos pode ser custosa. Existe pouco que o utilizador industrial possa fazer em relação a este problema...

Reforço novamente a ideia que, apesar dos pontos listados, as vantagens da tecnologia digital ultrapassam, em muito, os pontos que requerem algum cuidado.

9.1. Sistema de controlo por computador

Na Figura 9.1 apresenta-se a estrutura geral de um **sistema de controlo por computador**.

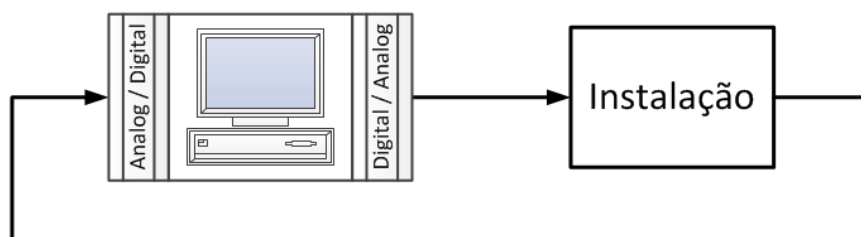


Figura 9.1 – Estrutura geral do sistema de controlo por computador.

Um sistema de controlo por computador é constituído pela instalação a controlar e pelo computador que executa a função do controlo, i.e. o controlador.

O sinal de entrada da instalação (atuação) provém da saída do computador; e o sinal de entrada do computador resulta da reacção (medida) à saída da instalação. Este fluxo de informação entre os dois elementos realiza-se através de um interface que converte os sinais analógicos em digitais (A/D na entrada do computador) e os sinais digitais em analógicos (D/A na saída do computador).

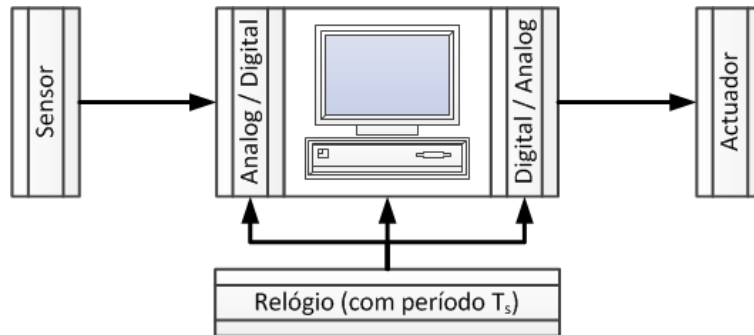


Figura 9.2 – Sincronização do sistema de controlo.

A instalação está equipada com sensores e atuadores, que convertem os sinais físicos (e.g. temperatura, pressão) em medidas elétricas contínuas no tempo; e convertem os valores de controlo (sinal elétrico) em variável física de atuação (e.g. posição de uma válvula, velocidade de um motor). As leituras das medidas realizadas no sensor, o seu processamento e a devolução do sinal de controlo devem ser sincronizadas por um relógio interno que garanta um ritmo constante no tempo (ver Figura 9.2). Este ritmo estabelece um intervalo de tempo que medeia o tratamento de duas amostras que se define como o período de amostragem, T_s (o índice s de *sampling*).

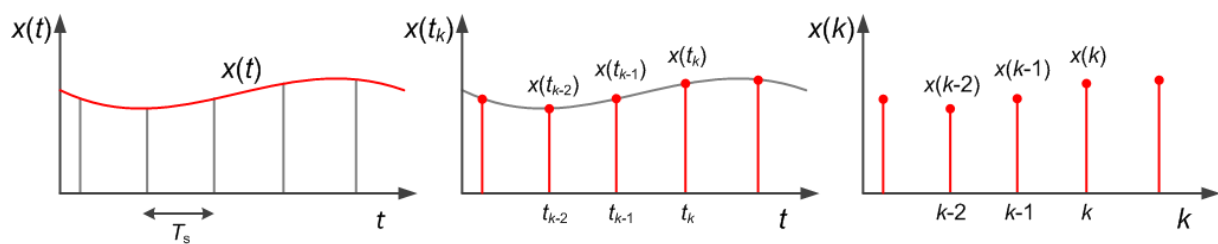


Figura 9.3 – Os sinais em tempo contínuo e em tempo discreto.

Estabelece-se assim uma ponte temporal entre o domínio do contínuo (fora do computador) e o domínio do digital (dentro do computador). O tempo deixa de existir de forma contínua, existindo operações apenas em instantes discretos dados por

$$t_k = k T_s$$

Referimo-nos assim ao **tempo discreto**. Para simplificar a notação, uma vez que o tempo de amostragem T_s é constante, não precisamos de nos estar a referir ao tempo discreto t_k mas

apenas ao seu índice k , com significado semelhante. Assim, a partir deste ponto referimos o tempo discreto através da variável índice da amostra k , até porque para o computador o “valor” do tempo não tem qualquer significado. O computador trata apenas sequência de amostras.

A Figura 9.3 ilustra a diferença entre o sinal em tempo contínuo; o sinal amostrado (ainda em tempo contínuo) e o sinal em tempo discreto. Este último perdeu completamente a informação sobre o que se passa entre amostras, e essa informação é irrecuperável.

O sinal em tempo discreto será daqui em diante referido apenas por **sinal discreto**.

9.2. Degrau e impulso unitários

Tal como nos sistemas contínuos, existem dois sinais preferenciais para a análise da dinâmica através da excitação dos sistemas. São estes o degrau unitário e o impulso unitário.

O sinal em **degrau unitário** define-se como

$$\varepsilon(k) = \begin{cases} 1 & k \geq 0 \\ 0 & k < 0 \end{cases}$$

e a sua representação temporal é apresentada na Figura 9.4

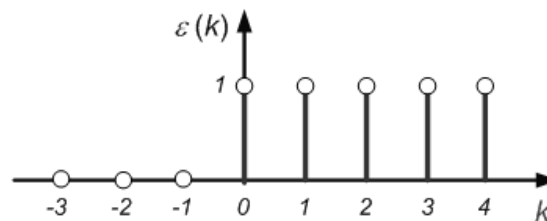


Figura 9.4 – Sinal em degrau unitário.

O sinal em impulso unitário define-se como

$$\delta(k) = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

e a sua representação temporal é apresentada na Figura 9.5

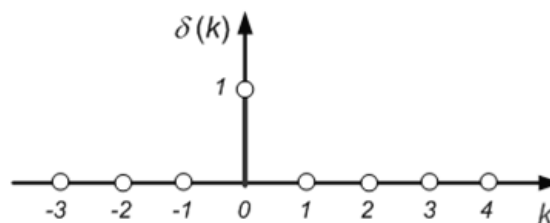


Figura 9.5 – Sinal em impulso unitário.

9.3. Sistema discreto

Um **sistema discreto** é um sistema cujos sinais de entrada e saída são sinais discretos.

A representação do sistema discreto é realizada, no domínio do tempo, através de uma equação recursiva que descreve como a saída no instante k se relaciona com entradas e saídas noutros instantes do tempo.

Por exemplo, a equação dada por

$$y(k) = 0.5 \cdot y(k - 1) + 2 \cdot u(k)$$

diz-nos que o valor do sinal de saída y no instante presente k é igual a metade do valor observado no instante anterior $k - 1$, adicionado de duas vezes o valor presente no mesmo instante k na entrada de excitação u . Estas equações são designadas por **equações às diferenças**, e estão para os sistemas discretos como as equações diferenciais estão para os sistemas contínuos.

O sistema discreto diz-se **dinâmico** se a saída $y(k)$ depender não só da entrada $u(k)$ no mesmo instante de tempo discreto mas também de outros valores passados de u ou de y , por exemplo, $u(k - 1)$, $u(k - 2)$, $y(k - 1)$, $y(k - 2)$.

Exemplo 9.1

O sistema descrito pela equação

$$y(k) = 0.2y(k - 1) + 3u(k - 2)$$

é um sistema dinâmico, pois a saída y no instante actual k depende directamente da saída no instante anterior $k - 1$, e da entrada u dois instantes anteriores do tempo discreto, ou seja, $k - 2$.

■

Exemplo 9.2

Um bloco de ganho

$$y(k) = 5u(k)$$

é uma função estática entre a entrada e a saída (amplificação de ganho constante e igual a 5), sem memória, e logo não é um sistema dinâmico.

■

Exemplo 9.3

Note-se no entanto que o sistema definido por

$$y(k) = 5u(k - 1)$$

é um sistema dinâmico, pois necessita de memória para reter o valor da entrada durante um ciclo de amostragem.

■

A definição de **sistema discreto linear** não é diferente da já apresentada para os sistemas contínuos. Se o sistema verificar os princípios da sobreposição e a escalabilidade de sinais, diz-se linear.

À semelhança também da definição em tempo contínuo, um sistema discreto diz-se **invariante no tempo** se a resposta a uma dada entrada for independente do instante em que é aplicada.

Um sistema diz-se **causal** se a saída $y(k)$ no instante presente não depender de nenhuma entrada futura, por exemplo, $u(k + 1)$.

Exemplo 9.4

Os sistemas representados a seguir são caracterizados quanto à causalidade do modo indicado:

$$y(k) = y(k - 1) + 3u(k - 2) \quad : \text{ sistema causal}$$

$$y(k) = y(k - 1) + 3u(k + 2) \quad : \text{ sistema não causal}$$

$$y(k) = y(k + 1) + 3u(k + 1) \quad : \text{ sistema causal (porquê?)}$$

■

A **causalidade no sentido lato** aceita que a saída dependa de uma entrada no mesmo instante. A **causalidade no sentido estrito** obriga a que haja, pelo menos, o atraso de uma amostra entre a entrada e a saída.

O sistema diz-se **estável no sentido assintótico** se, para um sinal de entrada com energia finita, a saída tender para zero de forma assintótica. Lembra-se que, tender para zero de forma assintótica significa que existem coeficientes $\alpha, a > 0$ tal que a função representativa do sinal tende para zero verificando a condição

$$|y(k)| < \alpha a^k$$

para qualquer valor de k . A verificação da estabilidade será revista mais adiante após introduzir os conceitos de função de transferência.

Exemplo 9.5

O sistema dado pela equação às diferenças

$$y(k) = y(k - 1) + u(k - 1)$$

não é estável no sentido assintótico, visto que se a entrada for um impulso unitário que tem energia finita, a saída será dada por

$$y(k) = \begin{cases} 1 & k > 0 \\ 0 & k \leq 0 \end{cases}$$

que não tende para zero de forma assintótica.

■

9.4. Equações às diferenças

No domínio do tempo discreto, um **sistema linear invariante no tempo** pode ser representado pela respetiva equação às diferenças que estabelece a relação entre sucessivas amostras das entradas e saídas do sistema, como

$$y(k+n) + a_1y(k+n-1) + \dots + a_ny(k) = b_0u(k+m) + b_1u(k+m-1) + \dots + b_mu(k)$$

ou, de forma equivalente,

$$y(k) + a_1y(k-1) + \dots + a_ny(k-n) = b_0u(k-d) + b_1u(k-d-1) + \dots + b_mu(k-n)$$

com $d = n - m$, igual ao atraso entre a entrada e a saída. Note-se que se o valor do atraso d for negativo o sistema não será causal. O valor de n define **a ordem do sistema** representado.

Em simultâneo, as equações às diferenças oferecem um mecanismo direto de simulação do sistema que representam, pois podem ser escritas na seguinte forma recorrente

$$y(k) = -a_1y(k-1) - \dots - a_ny(k-n) + b_0u(k-d) + b_1u(k-d-1) + \dots + b_mu(k-n)$$

permitindo calcular os valores da saída y para instantes sucessivos a partir dos valores calculados nas iterações anteriores e da sequência de entrada que excita o sistema. É necessário, naturalmente, conhecer as condições iniciais do sistema a simular. Esta perspetiva será usada também na realização do controlador.

9.5. Solução das equações às diferenças

A **solução de uma equação às diferenças** pode ser obtida por recorrência de equações. Conhecido o sinal de entrada $u(k)$ e as condições iniciais $y(0)$, $y(-1)$, ..., $y(-n+1)$, a própria equação às diferenças fornece os sucessivos valores da saída.

Exemplo 9.6

Considere-se o sistema discreto representado pela seguinte equação às diferenças de ordem 1

$$y(k) = 0.5y(k-1) + u(k-1)$$

com entrada nula $u(k) = 0$ e condição inicial $y(0) = 1.0$.

Os sucessivos valores da saída y são calculados de acordo com a tabela seguinte:

$k = 0$	$y(0) = 1.0$
$k = 1$	$y(1) = 0.5y(0) + u(0) = 0.5$
$k = 2$	$y(2) = 0.5y(1) + u(1) = 0.25$
$k = 3$	$y(3) = 0.5y(2) + u(2) = 0.125$
$k = 4$	$y(4) = 0.5y(3) + u(3) = 0.0625$

e a resposta é representada na Figura 9.6.

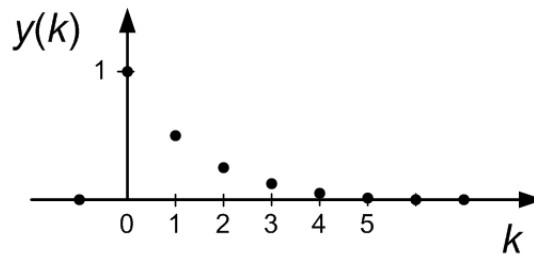


Figura 9.6 – Resposta livre do sistema do Exemplo 9.6.

Exemplo 9.7

Considere-se novamente o sistema discreto representado pela equação às diferenças

$$y(k) = 0.5y(k - 1) + u(k - 1)$$

agora, com entrada unitária em degrau unitário e condição inicial nula $y(0) = 0$.

Os sucessivos valores da saída y são calculados de acordo com a tabela seguinte:

$k = 0$	$y(0) = 0.0$
$k = 1$	$y(1) = 0.5y(0) + u(0) = 1.0$
$k = 2$	$y(2) = 0.5y(1) + u(1) = 1.5$
$k = 3$	$y(3) = 0.5y(2) + u(2) = 1.75$
$k = 4$	$y(4) = 0.5y(3) + u(3) = 1.875$

e a resposta é representada na Figura 9.7.

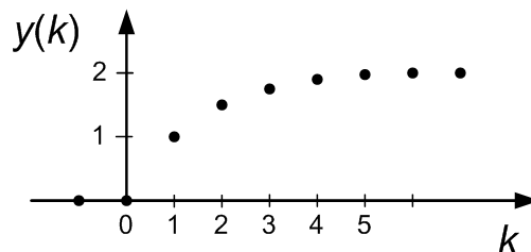


Figura 9.7 – Resposta forçada do Exemplo 9.7.

Esta forma de obtenção da solução apresenta, no entanto, um problema. Se necessitarmos de obter o valor da resposta para $k = 100$, necessitamos de calcular primeiro todos os valores anteriores.

Um método alternativo para resolver as equações às diferenças consiste em procurar e propor funções que, substituídas na equação às diferenças, permitam encontrar o valor da solução. As funções exponenciais de base a , ou $x(k) = a^k$, são candidatas naturais a soluções, dado que

$$x(k - 1) = a^{k-1} = a^{-1} \cdot a^k = a^{-1}x(k)$$

ou seja, as amostras sucessivas destas funções relacionam-se entre si apenas por fatores proporcionais. Note-se que mesmo a função constante é um caso particular desta função exponencial quando $a = 1$.

Exemplo 9.8

Considere-se novamente o sistema discreto representado pela equação às diferenças

$$y(k) = 0.5y(k - 1) + u(k - 1)$$

com entrada em degrau unitário e condição inicial $y(0) = 1$.

Considere-se uma solução genérica do tipo exponencial da forma

$$y(k) = \alpha a^k + \beta$$

e substitua-se na equação às diferenças sob análise. Resulta então (para $k \geq 1$)

$$\alpha a^k + \beta = 0.5 \cdot (\alpha a^{k-1} + \beta) + 1$$

Igualando, nos dois lados da equação, os termos constantes e os termos que dependem da função exponencial obtém-se, respetivamente,

$$\beta = 0.5\beta + 1$$

e

$$\alpha a^k = 0.5\alpha a^{-1}a^k$$

donde se pode concluir que $\beta = 2.0$ e $a = 0.5$. O valor de α , em falta, é obtido da condição inicial

$$y(0) = \alpha a^k + \beta = \alpha \cdot 0.5^0 + 2.0$$

resultando $\alpha = -1$. A solução final é então dada por

$$y(k) = -0.5^k + 2 = 2 \cdot (1 - 0.5^{k-1})$$

para $k \geq 0$.

Uma forma mais abreviada de escrever o resultado anterior é fazer uso do sinal em degrau unitário para ativar a resposta no tempo certo, escrevendo apenas

$$y(k) = 2(1 - 0.5^{k-1}) \cdot \varepsilon(k)$$

■

9.6. Resposta impulsiva

Define-se **resposta impulsiva** $h(k)$ como o sinal de saída dum sistema em resposta a um sinal à entrada em impulso unitário. Como no caso contínuo, o conhecimento da resposta impulsiva caracteriza completamente o comportamento dinâmico do sistema discreto linear e invariante no tempo.

Exemplo 9.9

Considere-se novamente o sistema discreto representado pela equação às diferenças

$$y(k) = 0.5y(k-1) + u(k-1)$$

sem pré-história, ou seja $y(0) = 0$, e excitado pelo impulso unitário. A resposta impulsiva do sistema pode ser calculada usando os métodos apresentados atrás.

O impulso unitário estabelece um valor para a saída no instante $k = 1$ dado por

$$y(1) = 0.5y(0) + u(0) = 1$$

Considerando uma solução do tipo exponencial da forma

$$y(k) = \alpha a^k + \beta$$

a substituição na equação às diferenças em estudo, para $k > 1$, resulta em

$$\alpha a^k + \beta = 0.5 \times (\alpha a^{k-1} + \beta)$$

Igualando, nos dois lados da equação, os termos constantes e os termos que dependem da função exponencial obtém-se, respetivamente,

$$\beta = 0.5\beta$$

e

$$\alpha a^k = 0.5\alpha a^{-1}a^k$$

donde se conclui que $\beta = 0$ e $a = 0.5$. O valor de α obtém-se da condição para $k = 1$, tal que

$$y(1) = \alpha a^1 + \beta = \alpha 0.5^1 + 0 = 1.0$$

resultando em $\alpha = 2.0$.

A resposta impulsiva é assim dada por

$$h(k) = y(k) = 2 \cdot 0.5^k \cdot \varepsilon(k - 1)$$

como se pode observar na Figura 9.8.

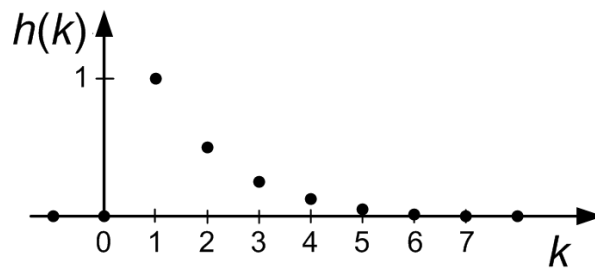


Figura 9.8 – Resposta impulsiva do Exemplo 9.9.

■

9.7. Cálculo da convolução

A utilização da resposta impulsiva e as características de linearidade e invariância no tempo permitem sistematizar, numa fórmula não recorrente, o cálculo da solução de uma equação às diferenças de um sistema discreto excitado com um sinal genérico $u(k)$.

Para compreensão do significado da fórmula são necessários dois passos, como se indica a seguir:

1. Qualquer sinal de excitação $u(k)$ pode ser decomposto na sobreposição de um número infinito de impulsos unitários deslocados no tempo. Ou seja, um sinal de entrada genérico $u(k)$ pode ser representado por

$$u(k) = \sum_{j=-\infty}^{+\infty} \delta(k - j) \cdot u(j)$$

Para um instante particular $k = k_0$ os termos da série são todos nulos, exceto para o elemento ativado pela função δ para o índice $j = k_0$, resultando $u(k_0) = 1 \cdot u(k_0)$.

2. A saída do sistema como reação ao sinal genérico $u(k)$ pode ser reconstruída pela sobreposição das respostas impulsivas de cada um dos impulsos da série presente na equação anterior. Ou seja, cada uma das parcelas da série $u(j)\delta(k - j)$ terá como resposta $u(j)h(k - j)$. Aplicando o princípio da sobreposição resulta

$$y(k) = \sum_{j=-\infty}^{+\infty} h(k - j) \cdot u(j)$$

sendo esta fórmula designada por **produto de convolução** entre o sinal de excitação u e a resposta impulsiva h . No domínio do tempo discreto, o produto de convolução também é representado pelo símbolo $*$, tal como

$$y(k) = h(k) * u(k) = \sum_{j=-\infty}^{+\infty} h(k-j) \cdot u(j) = \sum_{j=-\infty}^{+\infty} h(j) \cdot u(k-j)$$

Exemplo 9.10

Considere-se o sistema discreto representado pela equação às diferenças

$$y(k) = 0.5y(k-1) + u(k-1)$$

da qual se conhece já a resposta impulsiva (ver Exemplo 9.9)

$$h(k) = 2 \cdot 0.5^k \cdot \varepsilon(k-1)$$

Considere-se agora o sinal de excitação dado por

$$u(k) = \varepsilon(k-1) - 3\varepsilon(k-2) + \varepsilon(k-3) + \varepsilon(k-4)$$

que está representado na Figura 9.9.

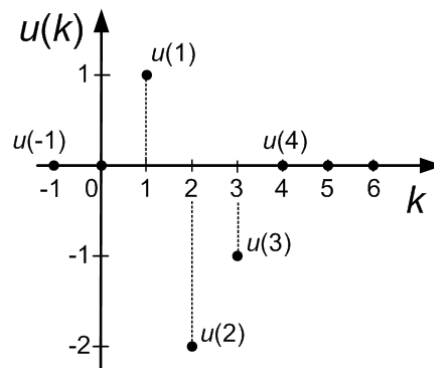


Figura 9.9 – Sinal de entrada do Exemplo 9.10.

Pela observação da Figura 9.9 é possível escrever também o sinal de entrada como

$$\begin{aligned} u(k) &= \dots + 0 \cdot \delta(k+1) + 0 \cdot \delta(k) + \delta(k-1) - 2 \cdot \delta(k-2) - \delta(k-3) + 0 \cdot \delta(k-4) + \dots = \\ &= \sum_{j=-\infty}^{+\infty} u(j) \cdot \delta(k-j) \end{aligned}$$

A saída do sistema $y(k)$ é agora dada pela sobreposição das respostas impulsivas, representada na Figura 9.10,

$$\begin{aligned} y(k) &= \dots + 0 \cdot h(k+1) + 0 \cdot h(k) + h(k-1) - 2 \cdot h(k-2) - h(k-3) + 0 \cdot h(k-4) + \dots = \\ &= \sum_{j=-\infty}^{+\infty} u(j) \cdot h(k-j) \end{aligned}$$

e, como $h(k) = 2 \cdot 0.5^k \cdot \varepsilon(k-1)$, obtém-se

$$y(k) = h(k-1) - 2 \cdot h(k-2) - h(k-3) = \begin{cases} 0 & k \leq 1 \\ 1 \cdot 0.5^{k-1} & k = 2 \\ 1 \cdot 0.5^{k-1} - 2 \cdot 0.5^{k-2} & k = 3 \\ 1 \cdot 0.5^{k-1} - 2 \cdot 0.5^{k-2} - 1 \cdot 0.5^{k-3} & k \geq 4 \end{cases}$$

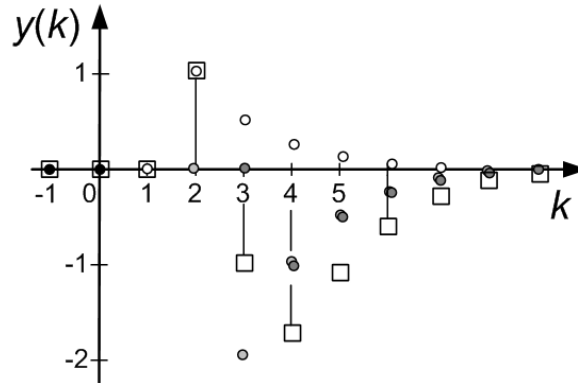


Figura 9.10 – Sinal de saída $y(k)$ do Exemplo 9.10 (símbolo \square).

■

9.8. Transformação em z

No estudo dos sistemas em tempo contínuo, introduziu-se a transformação de Laplace (unilateral direita) como forma alternativa de resolver as equações diferenciais de ordem n genérica. Essa introdução abriu depois o caminho que nos levou à função de transferência e tudo aquilo que viemos a descobrir depois.

A transformação no domínio da variável z , muitas vezes designada apenas por **transformação em z** , realizará o mesmo papel na análise dos sistemas discretos.

A transformação em z é um operador que transforma um sinal $x(k)$ no domínio do tempo discreto k , numa função $X(z)$ no domínio da variável complexa z . A operação é definida por

$$X(z) = \sum_{j=0}^{+\infty} x(j) \cdot z^{-j}$$

A transformação em z aparece representada na literatura em diversas formas equivalentes tais como

$$\text{TZ}\{x(k)\} = \text{Z}\{x(k)\} = X(z)$$

A partir da definição é possível calcular a transformação em z de qualquer sinal, embora isso possa não ser simples. Os exemplos seguintes calculam as transformações em z para os equivalentes discretos do impulso unitário e do degrau unitário.

Exemplo 9.11

Considere-se o sinal em impulso unitário $x(k) = \delta(k)$, a sua transformação em z é dada por

$$X(z) = \sum_{j=0}^{+\infty} x(j) \cdot z^{-j} = \sum_{j=0}^{+\infty} \delta(j) \cdot z^{-j} = \delta(0) \cdot z^{-0} = 1$$

ou seja

$$\mathbf{TZ}\{\delta(k)\} = 1$$

■

Exemplo 9.12

Considere-se o sinal em degrau unitário $x(k) = \varepsilon(k)$, a sua transformação em z é dada por

$$X(z) = \sum_{j=0}^{+\infty} x(j) \cdot z^{-j} = \sum_{j=0}^{+\infty} \varepsilon(j) \cdot z^{-j} = 1 + z^{-1} + z^{-2} + z^{-3} + \dots = \frac{1}{1 - z^{-1}}$$

segundo a regra conhecida para uma série geométrica, neste caso com razão $r = z^{-1}$.

Resulta assim

$$\mathbf{TZ}\{\varepsilon(k)\} = \frac{1}{1 - z^{-1}} = \frac{z}{z - 1}$$

■

As tabelas seguintes apresentam algumas transformações em z para alguns sinais e algumas das propriedades da transformação em z .

Tabela de transformações em z

	Sinal	Transformação em z
1	$\delta(k)$	1
2	$\varepsilon(k)$	$\frac{z}{z - 1}$
3	$a^k \cdot \varepsilon(k)$	$\frac{z}{z - a}$
4	$k \cdot \varepsilon(k)$	$\frac{z}{(z - 1)^2}$
5	$k \cdot a^k \cdot \varepsilon(k)$	$\frac{a \cdot z}{(z - a)^2}$
6	$a^k \cdot \sin \Omega_0 k \cdot \varepsilon(k)$	$\frac{a \cdot \sin \Omega_0 \cdot z}{z^2 - 2 \cdot a \cdot \cos \Omega_0 z + a^2}$
7	$a^k \cdot \cos \Omega_0 k \cdot \varepsilon(k)$	$\frac{(z - a \cdot \cos \Omega_0) \cdot z}{z^2 - 2 \cdot a \cdot \cos \Omega_0 z + a^2}$

Tabela de Propriedades da transformação em z

Propriedades da transformação em z		
P1	Definição	$X(z) = \sum_{j=0}^{+\infty} x(j) \cdot z^{-j}$
P2	Inversa	$x(k) = \text{TZ}^{-1}\{X(z)\} = \frac{1}{2\pi j} \oint_C X(z)z^{k-1} dz$
P3	Linearidade	Se $w(k) = a \cdot x(k) + b \cdot v(k)$, então $W(z) = a \cdot X(z) + b \cdot V(z)$
P4	Avanço no tempo (um passo)	$\text{TZ}\{x(k+1)\varepsilon(k)\} = z \cdot X(z) - z \cdot x(0)$
P5	Avanço no tempo (dois passos)	$\text{TZ}\{x(k+2)\varepsilon(k)\} = z^2 X(z) - z^2 x(0) - zx(1)$
P6	Avanço no tempo (n passos)	$\text{TZ}\{x(k+n)\varepsilon(k)\} = z^n X(z) - z^n x(0) \dots - zx(n-1)$
P7	Atraso (n passos)	$\text{TZ}\{x(k-n)\varepsilon(k-n)\} = z^{-n} X(z)$
P8	Convolução	$\text{TZ}\left\{\sum_{j=-\infty}^{+\infty} h(k-j) \cdot x(j)\right\} = H(z) \cdot X(z)$
P9	Produto por a^k	$\text{TZ}\{a^k \cdot x(k)\} = X\left(\frac{z}{a}\right)$
P10	Produto por k	$\text{TZ}\{k \cdot x(k)\} = -z \cdot \frac{dX(z)}{dz}$
P11	Teorema do valor final	$\lim_{k \rightarrow +\infty} x(k) = \lim_{z \rightarrow 1} (z-1)X(z)$
P12	Teorema do valor inicial	$x(0) = \lim_{z \rightarrow +\infty} X(z)$

Uma vez introduzida a transformação em z , estamos em condições de ganhar generalidade na análise dos sistemas discretos através do conceito de função de transferência de um sistema discreto.

9.9. Função de transferência

Define-se a **função de transferência** de um sistema discreto dinâmico linear e invariante no tempo como a função que relaciona as transformações em z dos sinais de entrada e de saída do sistema, ou seja

$$H(z) = \frac{Y(z)}{U(z)}$$

Analicamente, a função de transferência confunde-se com a transformação em z da resposta impulsiva

$$H(z) = \text{TZ}\{h(k)\}$$

e isto sucede pelo facto da transformação em z do impulso unitário ser igual à unidade. No entanto, a função de transferência representa uma relação entrada-saída, independente do sinal de excitação.

Note-se que para o caso discreto é possível escrever a função de transferência como

$$H(z) = \text{TZ}\{h(k)\} = \sum_{j=0}^{+\infty} h(j)z^{-j} = h(0) + h(1)z^{-1} + h(2)z^{-2} + h(3)z^{-3} + \dots$$

Fazendo uso das propriedades da transformação em z , em especial as relativas ao avanço e atraso de sinais no tempo (propriedades P4 a P6), é possível fazer a transposição direta da equação às diferenças do sistema linear e invariante no tempo para a função de transferência correspondente. Assim, o sistema representado pela equação às diferenças dada por

$$y(k+n) + a_1y(k+n-1) + \dots + a_ny(k) = b_0u(k+m) + b_1u(k+m-1) + \dots + b_mu(k)$$

tem o equivalente no domínio da variável z dado por

$$z^nY(z) + a_1z^{n-1}Y(z) + \dots + a_nY(z) = b_0z^mU(z) + b_1z^{m-1}U(z) + \dots + b_mU(z)$$

que resulta na função de transferência

$$F(z) = \frac{Y(z)}{U(z)} = \frac{b_0z^m + b_1z^{m-1} + \dots + b_m}{z^n + a_1z^{n-1} + \dots + a_n}$$

com $n \geq m$ para garantir a causalidade.

9.10. Código do controlador

A transposição direta entre a função de transferência e a equação às diferenças é necessária, especialmente, no passo de projeto em que após dimensionado o controlador $C(z)$ será

necessário realizá-lo no código da rotina relativa ao controlador. O **código do controlador** é a equação às diferenças escrita numa forma específica.

Exemplo 9.13

Considere-se que foi projetado o controlador com a seguinte estrutura

$$C(z) = \frac{U(z)}{E(z)} = \frac{b_0z - b_1}{z - a_1}$$

onde $\{a_1; b_0; b_1\}$ é o conjunto de parâmetros do controlador; $U(z) = \text{TZ}\{u(k)\}$ é o sinal de actuação a produzir pelo controlador; e $E(z) = \text{TZ}\{e(k)\}$ é o sinal de erro obtido a partir da diferença entre o sinal de referência $r(k)$ (*set-point*) e a saída da instalação $y(k)$.

O código do controlador é obtido fazendo a transposição da função de transferência para a equação às diferenças correspondente

$$u(k + 1) - a_1u(k) = b_0e(k + 1) - b_1e(k)$$

e reescrita para produzir, em cada instante de amostragem, a ação de controlo no instante atual k , ou seja

$$u(k) = a_1u(k - 1) + b_0e(k) - b_1e(k - 1)$$

a partir de informação relativa a valores obtidos no passado $\{u(k - 1); e(k - 1)\}$ ou valores relativos a entradas no presente $\{e(k)\}$.

Note-se que a variável k usada no processo de recorrência é uma variável muda. Na última passagem, para obter o código do controlador, subentende-se uma mudança de variável de k para $k - 1$.

A Figura 9.11 apresenta em formato de pseudo-código uma possível rotina (ou método) de controlo incluída numa aplicação informática tipo SCADA (*supervisory control and data acquisition*) responsável pela realização da lei de controlo. Note-se que os valores relativos a variáveis passadas são guardados em variáveis (*static*) para que possam estar disponíveis em chamadas futuras da rotina. Subentende-se aqui também a existência de outras rotinas *readAD(.)* e *writeDA(.)* responsáveis, respectivamente por lerem e escreverem os valores medidos e a actuar na instalação. Estas rotinas são normalmente fornecidas pelo fabricante das cartas de aquisição de dados para diversas linguagens de programação

■

```

void controller(double rk) {
    double uk, yk, ek;
    static double uk_1, ek_1;
    // rk – reference value
    readAD(portIn, yk);
    ek = rk - yk;
    uk = a1*uk_1 + b0*ek - b1*ek_1;
    writeDA(portOut, uk)
    // update state for next iteration
    uk_1 = uk;
    ek_1 = ek;
    return;
} // end of controller routine

```

Figura 9.11 – Possível pseudo-código relativo ao Exemplo 9.13.

Dentro do nosso objetivo de projetar sistemas de controlo, usaremos a função de transferência para descrever a instalação, que designaremos $F(z)$, e o controlador que resultará do projecto, que designaremos $C(z)$. É importante reflectir neste ponto que, se o controlador, realizado através do computador, tem realidade discreta; a instalação, que existe fora do computador, tem realidade contínua. Qual o significado então de $F(z)$, a instalação discreta? Esta função descreve como a dinâmica da instalação é vista a partir do interior do domínio discreto, i.e. a partir do computador, com o processo de amostragem em curso com período de amostragem T_s . Conclusão inevitável: para a mesma instalação (que no contínuo modelávamos através de uma função de transferência $F(s)$) iremos obter funções de transferência $F(z)$ diferentes para valores diferentes do período de amostragem.

9.11. Seleção do período de amostragem

A seleção do período de amostragem T_s é um dos aspetos mais importantes no projeto de um sistema de controlo por computador porque no tempo que medeia a leitura de duas amostras, a instalação encontra-se em anel aberto (i.e. sem controlo). Evidentemente, teremos que fazer este período suficientemente curto para que isso não coloque a instalação em risco. Mas, mais velocidade requer sistemas mais dispendiosos e que consomem mais recursos, e.g. energia.

9.12. Teorema da amostragem (Nyquist-Shannon)

A recuperação de um sinal analógico a partir das suas amostras só é possível se a frequência de amostragem definida como

$$\omega_s = \frac{2\pi}{T_s}$$

for, pelo menos, duas vezes superior à máxima frequência incluída do sinal, ou seja

$$\omega_s > 2\omega_{\max}$$

Visto na perspectiva oposta, frequência máxima do sinal analógico que permite a sua recuperação a um dado ritmo de amostragem é

$$\omega_{\max} < \omega_N = \frac{\pi}{T_s}$$

onde $\omega_N = \omega_s/2$ é designada por **frequência de Nyquist**.

A Figura 9.12 ilustra através de um exemplo o que sucede quando esta regra não é rejeitada. Considere-se a amostragem a um ritmo de $f_s = 1$ Hz, logo com $T_s = 1$ s, e dois sinais a digitalizar com frequências de $f_1 = 0.40$ Hz e $f_2 = 0.60$ Hz. Pode observar-se na figura em baixo à esquerda que as amostras, tiradas de segundo a segundo, coincidem exactamente para estes dois sinais originalmente diferentes. No entanto, no processo de recuperação das amostras novamente para o contínuo não existe forma de os diferenciar.

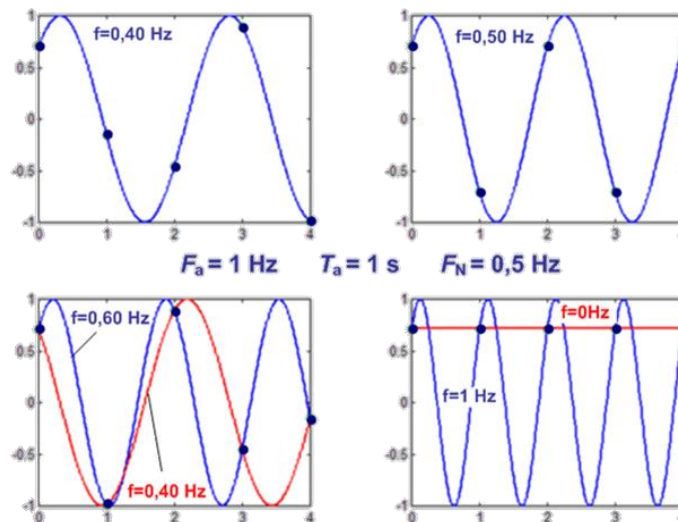


Figura 9.12 – O fenómeno do falseamento (*aliasing*).

Sucede que, para este exemplo, a frequência de Nyquist é igual a $f_N = f_s/2 = 0.5$ Hz. O sinal de frequência $f_1 < f_N$ é recuperado de forma adequada, mas o sinal de frequência $f_2 > f_N$ é recuperado também com a frequência f_1 .

No gráfico em baixo à direita da Figura 9.12, observamos o que sucede quando o sinal tem frequência $f = 2f_N = 1$ Hz. As amostras são todas de valor constante. Na recuperação resulta um sinal de frequência 0 Hz (i.e. um sinal constante).

Em resumo, os sinais analógicos com frequências f_i que estejam acima da frequência de Nyquist f_N serão reconstruídas como sinais com frequências $2f_N - f_i$, correspondendo a um rebatimento em redor da frequência de Nyquist (ver Figura 9.13).

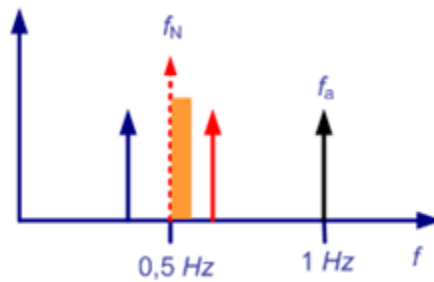


Figura 9.13 – O fenômeno do falseamento (*aliasing*) na representação em frequência.

O sinal a amostrar deve estar limpo de componentes na frequência acima da frequência de Nyquist. De modo a garantir a inexistência de componentes espectrais acima da frequência de Nyquist deve introduzir-se um filtro passa-baixo (como representado na Figura 9.14) antes da amostragem para remover as essas componentes que resultarão, na recuperação, em ruído sobre as frequências abaixo do limiar.

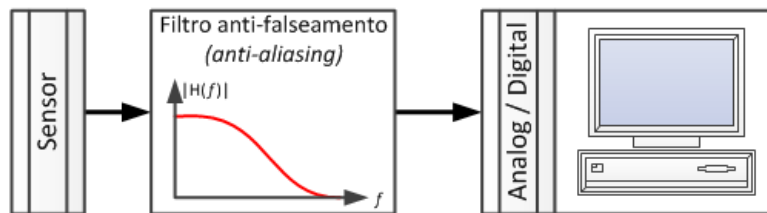


Figura 9.14 – A introdução do filtro *anti-aliasing*.

Existindo o filtro ideal o problema resolvia-se estabelecendo a frequência de corte do filtro igual à frequência de Nyquist, que por sua vez seria a frequência máxima de interesse para o sistema. No entanto, mesmo para um filtro de 2ª ordem (com decaimento a -40 dB/dec.) apenas se consegue uma atenuação de 100 vezes após uma década acima da frequência de corte do filtro.

As condições impostas pelo teorema de Nyquist-Shannon (ideais) não satisfazem o uso prático pelas seguintes razões:

- nenhum sinal físico é de banda limitada;
- realização de um filtro anti-falseamento passa-baixo não é perfeita; e
- a reconstrução de Shannon (rejeita a Teoria de Sinais) é não-causal;
- a reconstrução de Shannon é de cálculo complexo, preferindo-se uma aproximação de reconstrução por retenção, i.e. tomar o valor do sinal no intervalo de amostragem pelo valor da última amostra disponível. O método da reconstrução por retenção é obviamente mais fácil (leia-se económico) de realizar.

Por estes motivos, na prática do controlo digital, a seleção da frequência de amostragem realiza-se pela regra mais conservadora

$$\omega_s = (10 \cdots 20) \times \omega_{\max}$$

ou de forma equivalente,

$$\omega_N = (5 \cdots 10) \times \omega_{\max}$$

onde ω_{\max} representa a largura de banda do sistema de controlo pretendido já em anel fechado. Esta gama de valores adequados, aparentemente larga, permite ao engenheiro de controlo alguma liberdade no estabelecimento do valor de amostragem. De seguida, voltaremos a este assunto, analisando em que situações se devem utilizar os limites inferior ou superior destas regras de dimensionamento.

9.13. Duas abordagens alternativas no projeto do controlador em tempo discreto

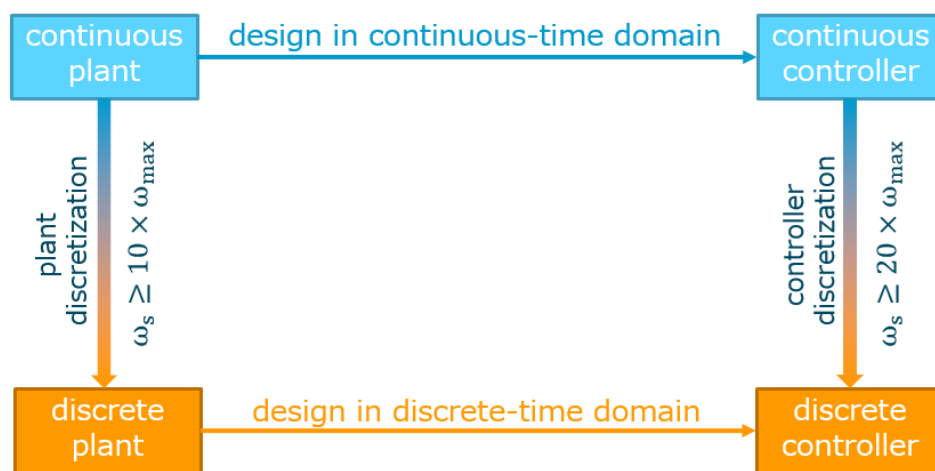


Figura 9.15 – Duas abordagens possíveis à realização do projeto.

No projeto de sistemas de controlo por computador existem duas alternativas principais representadas na Figura 9.15:

1. a substituição por emulação de um controlador contínuo $C(s)$, já existente ou projetado no contínuo para o efeito, pelo seu equivalente discreto $C(z)$; ou
2. o projeto de raiz de um controlador em tempo discreto a partir do equivalente discreto $F(z)$ da instalação contínua descrita por $F(s)$.

Na prática, o projeto resultará melhor quanto mais reduzida for a frequência de amostragem (maior período de amostragem). Em geral, para o mesmo desempenho do sistema de controlo conseguem-se valores mais reduzidos do ritmo de amostragem seguindo a abordagem 2. ou seja, projetar o controlador $C(z)$ de raiz.

Em situações em que o tempo de processamento não é uma limitação, pode ser interessante considerar a abordagem 1. Pense-se num contrato de substituição de umas dezenas de controladores que estão em funcionamento, mas que se pretende fazer evoluir de uma

tecnologia analógica para o digital. Pode ser economicamente muito compensador, especialmente pela interrupção da produção por um período mínimo. Não fica invalidada a aplicação da segunda abordagem num tempo posterior para melhorar o desempenho de cada cadeia de regulação.

10. Análise dinâmica do sistema discreto

No capítulo anterior introduziu-se a função de transferência e referiu-se a sua utilização na perspetiva da modelação da instalação a controlar $F(z)$ e na perspetiva do controlador a projetar $C(z)$. Neste capítulo faz-se a análise das características dinâmicas de um sistema representado por uma função de transferência, independentemente da sua utilização.

10.1. Álgebra de blocos

Esta secção dedicada à **álgebra de blocos** é breve, no sentido em que tudo aquilo que se aplicava à manipulação de blocos representando funções de transferência na variável complexa s , aplica-se sem qualquer modificação aos esquemas de blocos representando funções de transferência na variável complexa z . Aconselha-se assim a revisão da secção 4.3 *Álgebra de blocos*.

10.2. Polos e zeros da função de transferência discreta

A análise do comportamento dinâmico de um sistema discreto através da sua função de transferência $G(z)$ é realizada observando a posição das raízes dos polinómios numerador $N(z)$ e denominador $D(z)$. Assim, seja $G(z)$ a função de transferência definida por

$$G(z) = \frac{N(z)}{D(z)}$$

onde o polinómio numerador, de ordem m , é dado por

$$N(z) = b_0 z^m + b_1 z^{m-1} + \dots + b_{m-1} z + b_m$$

e o polinómio denominador, de ordem n , é dado por

$$D(z) = z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n$$

Se o sistema for causal a ordem do denominador é superior ou igual à ordem do numerador, ou seja, $n \geq m$.

As m raízes do numerador $N(z)$ designam-se por **zeros** da função de transferência e representam-se no plano complexo da variável z por um zero - 'o'.

As n raízes do denominador $D(z)$ designam-se por **polos** da função de transferência e representam-se no plano complexo da variável z por um xis - 'x'.

Tal como registado nos sistemas em tempo contínuo, o lugar destas raízes, polos ou zeros, condiciona o comportamento dinâmico do sistema discreto.

Exemplo 10.1

Considere-se o sistema descrito pela função de transferência

$$G(z) = \frac{3.0z^3 - 7.5z^2 + 5.25z - 1.125}{z^4 - 3.3z^3 + 3.92z^2 - 1.988z + 0.3536}$$

O polinómio numerador tem as raízes (zeros da função de transferência)

$$N(z) = 3.0z^3 - 7.5z^2 + 5.25z - 1.125 = 0 \quad \left| \begin{array}{l} z_1 = +1.5 \\ z_2 = +0.5 \\ z_3 = +0.5 \end{array} \right.$$

e o polinómio denominador tem as raízes (polos da função de transferência)

$$D(z) = z^4 - 3.3z^3 + 3.92z^2 - 1.988z + 0.3536 = 0 \quad \left| \begin{array}{l} p_1 = +1.3 \\ p_2 = +0.8 + 0.2j \\ p_3 = +0.8 - 0.2j \\ p_4 = +0.4 \end{array} \right.$$

Na Figura 10.1 observa-se o lugar destas raízes no diagrama polos-zeros.

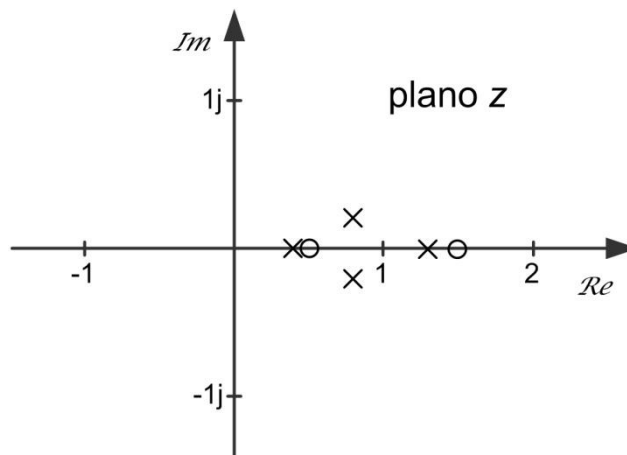


Figura 10.1 – Lugar das raízes da função de transferência do Exemplo 10.1.

■

10.3. Estabilidade assintótica

A **estabilidade** é uma das características mais procuradas de um sistema dinâmico e é muito importante saber concluí-la a partir do diagrama de polos-zeros, também para o caso discreto. Relembra-se que a estabilidade de um sistema linear é uma característica intrínseca do sistema e portanto não depende do sinal externo de excitação.

Um sistema discreto diz-se **estável no sentido assintótico** se, dada a sua resposta impulsiva, existirem valores $\alpha, a > 0$, tal que, qualquer que seja o valor do tempo k

$$|h(k)| < \alpha a^k$$

Significa isto que a resposta impulsiva deve aproximar-se de zero (quando o tempo vai para infinito) mais depressa que uma exponencial decrescente com parâmetros positivos adequados (Figura 4.15). Uma consequência desta definição é que a resposta impulsiva de um sistema estável no sentido assintótico converge obrigatoriamente para zero.

Um sistema dinâmico representado pela sua função de transferência é **estável no sentido assintótico** se e só se todos os seus polos estão posicionados no interior do círculo de raio unitário, ou seja, se e só se

$$|p_i| < 1 \quad \text{para } i = 1, \dots, n$$

A posição dos zeros não influencia a estabilidade do sistema, podendo estar localizados em qualquer posição do plano complexo.

Em oposição, um polo diz-se **instável** quando se apresenta fora do círculo unitário.

Exemplo 10.2

Considere-se novamente o sistema descrito pela função de transferência

$$G(z) = \frac{3.0z^3 - 7.5z^2 + 5.25z - 1.125}{z^4 - 3.3z^3 + 3.92z^2 - 1.988z + 0.3536}$$

Observou-se já no Exemplo 10.1 que os polos da função de transferência são dados por

$$D(z) = z^4 - 3.3z^3 + 3.92z^2 - 1.988z + 0.3536 = 0 \quad \left\{ \begin{array}{l} p_1 = +1.3 \\ p_2 = +0.8 + 0.2j \\ p_3 = +0.8 - 0.2j \\ p_4 = +0.4 \end{array} \right.$$

Como pelo menos um polo, p_1 , está fora do círculo unitário, o sistema não é estável no sentido assintótico (ver Figura 10.2).

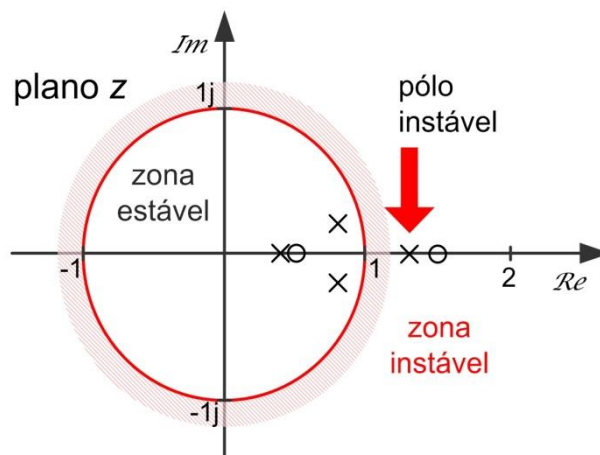


Figura 10.2 – Polo instável do Exemplo 10.2.



Exemplo 10.3

Considere-se o sistema descrito pela função de transferência

$$G(s) = \frac{3z^3 - 7.5z^2 + 5.25z - 1.125}{z^4 - 2.9z^3 + 3.44z^2 - 1.876z + 0.36}$$

Neste caso os polos da função de transferência são dados por

$$D(z) = z^4 - 2.9z^3 + 3.44z^2 - 1.876z + 0.36 = 0 \quad \left| \begin{array}{l} p_1 = +0.9 \\ p_2 = +0.8 + 0.6j \\ p_3 = +0.8 - 0.6j \\ p_4 = +0.4 \end{array} \right.$$

Como os polos p_2 e p_3 têm módulo unitário, o sistema não é estável no sentido assintótico. Nesta situação em que os polos estão sobre a fronteira da estabilidade (e não existe mais nenhum no exterior do círculo unitário) o sistema diz-se **marginalmente estável** (Figura 10.3).

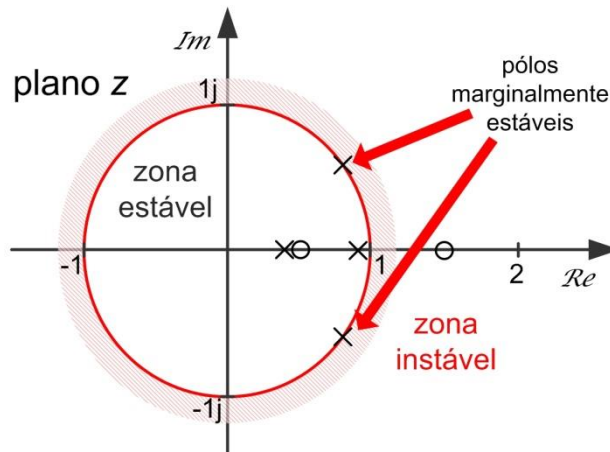


Figura 10.3 – Polos marginalmente estáveis (Exemplo 10.3).

■

10.4. Critério de estabilidade de Jury

À semelhança do critério de Routh-Hurwitz para sistemas contínuos, é conveniente ter uma forma de verificar a estabilidade de um sistema discreto a partir da sua função de transferência sem realizar o cálculo explícito do lugar dos polos. Evidentemente, com as ferramentas computacionais que temos ao nosso dispor hoje em dia, o cálculo do lugar dos polos é imediato e uma ferramenta deste tipo foi claramente mais relevante no passado. No entanto, estes métodos permitem ao engenheiro detetar um sistema instável com algumas verificações de cálculo fácil.

Assim, o **critério de estabilidade de Jury** introduz um conjunto de condições a verificar pelo polinómio característico, i.e. o denominador da função de transferência.

Cada uma das condições é necessária mas não suficientes. Logo, todas as condições têm de ser verificadas para concluir sobre a estabilidade do sistema em causa.

Condições de estabilidade de Jury

Considere-se o polinómio característico de ordem n dado por

$$P(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n$$

com $a_0 > 0$. As condições de estabilidade são

1. $|a_n| < a_0$
2. $P(1) > 0$
3. $P(-1) \cdot (-1)^n > 0$, i.e. $P(-1)$ deve ser positivo para n par e negativo para n ímpar.
4. $(n - 2)$ regras adicionais que resultam da construção de uma tabela.

Iremos introduzir as regras adicionais relativas ao ponto 4. através do exemplo seguinte.

Exemplo 10.4

Caracterize-se o sistema definido pela seguinte função de transferência quanto à estabilidade usando o teste de Jury

$$G(z) = \frac{1}{z^4 - 1.2z^3 + 0.07z^2 + 0.3z - 0.08}$$

A equação característica com $n = 4$ (par!)

$$P(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n$$

tem os coeficientes

$$a_0 = 1 \quad a_1 = -1.2 \quad a_2 = 0.07 \quad a_3 = 0.3 \quad a_4 = -0.08$$

Verifiquemos então as condições de Jury, lembre-se todas necessárias:

1. $|a_4| = 0.08 < a_0 = 1$ - Verifica-se!
2. $P(1) = 1 - 1.2 + 0.07 + 0.3 - 0.08 = 0.09 > 0$ - Verifica-se!
3. $P(-1) = 1 + 1.2 + 0.07 - 0.3 - 0.08 = 1.89 > 0$ - Verifica-se!
4. A verificação das $(4 - 2) = 2$ regras adicionais requerem a construção da tabela seguinte:

Tabela de Jury

z^0	z^1	z^2	z^3	z^4	
$a_4 = -0.08$				$a_0 = 1$	$b_3 = a_4 a_4 - a_0 a_0 = -0.994$
$a_0 = 1$				$a_4 = -0.08$	
$a_4 = -0.08$			$a_1 = -1.2$		$b_2 = a_4 a_3 - a_1 a_0 = 1.176$
$a_0 = 1$			$a_3 = 0.3$		
$a_4 = -0.08$		$a_2 = 0.07$			$b_1 = a_4 a_2 - a_2 a_0 = -0.076$
$a_0 = 1$		$a_2 = 0.07$			
$a_4 = -0.08$	$a_3 = 0.3$				$b_0 = a_4 a_1 - a_3 a_0 = -0.204$
$a_0 = 1$	$a_1 = -1.2$				
$b_3 = -0.994$			$b_0 = -0.204$		$c_2 = b_3 b_3 - b_0 b_0 = 0.946$
$b_0 = -0.204$			$b_3 = -0.994$		
$b_3 = -0.994$		$b_1 = -0.076$			$c_1 = b_3 b_2 - b_1 b_0 = -1.184$
$b_0 = -0.204$		$b_2 = 1.176$			
$b_3 = -0.994$	$b_2 = 1.176$				$c_0 = b_3 b_1 - b_2 b_0 = 0.315$
$b_0 = -0.204$	$b_1 = -0.076$				
$c_2 = 0.946$		$c_0 = 0.315$			

As 2 condições em falta são então:

- $|b_3| = 0.994 > |b_0| = 0.204$ - Verifica-se!
- $|c_2| = 0.946 > |c_0| = 0.315$ - Verifica-se!

Uma vez que se verificam todas as condições, a equação característica tem todas as suas raízes com módulo inferior a 1, ou seja, o sistema sob análise tem todos os polos localizados no interior do círculo unitário e logo é estável. Deixa-se como exercício essa verificação através do cálculo exato da localização dos polos.



Note-se que a tabela só requer ser construída para sistemas de ordem $n > 2$.

Exemplo 10.5

Considere o sistema discreto dado pela função de transferência

$$F(z) = \frac{0.3679z + 0.2642}{(z - 0.3679)(z - 1)}$$

colocado num esquema de retroação como o apresentado de seguida na Figura 10.4.

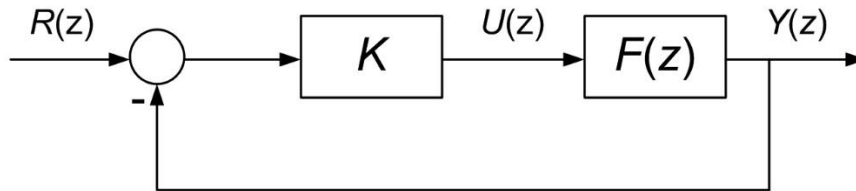


Figura 10.4 – Sistema em anel fechado (Exemplo 10.5).

Podemos agora usar o critério de Jury para estudar a gama de valores do ganho K que resultam num sistema em anel fechado estável.

A equação característica é agora obtida resolvendo o esquema de retroação

$$H(z) = \frac{K \cdot F(z)}{1 + K \cdot F(z)} = \frac{K \cdot (0.3679z + 0.2642)}{(z - 0.3679)(z - 1) + K \cdot (0.3679z + 0.2642)}$$

ou seja,

$$P(z) = z^2 + (0.3679K - 1.3679)z + (0.2642K + 0.3679)$$

Como a ordem do polinómio característico é apenas $n = 2$ (par), basta aplicar as três primeiras condições:

1. $|0.2642K + 0.3679| < 1 \quad \Leftrightarrow \quad -1 < 0.2642K + 0.3679 < 1$
2. $P(1) = 1 + (0.3679K - 1.3679) + (0.2642K + 0.3679) = 0.6321K > 0$
3. $P(-1) = 1 - (0.3679K - 1.3679) + (0.2642K + 0.3679) = -0.1037K + 2.7358 > 0$

ou equivalentemente

1. $-5.18 < K < 2.39$
2. $K > 0$
3. $K < 26.38$

resultando então um intervalo de estabilidade

$$K \in]0; 2.39[$$

■

10.5. Equivalência entre respostas impulsivas contínuas e discretas

Concluimos anteriormente que, para garantir a estabilidade assintótica do sistema descrito por uma função de transferência $G(z)$, os polos devem apresentar-se todos no interior do círculo de raio unitário centrado na origem.

Podemos perguntar agora qual o sistema discreto com resposta impulsiva coincidente com as amostras da resposta impulsiva de um determinado sistema contínuo, por exemplo, com apenas um polo.

Considere-se então o sistema contínuo dado por

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

Sabemos que a sua resposta impulsiva é dada por

$$h(t) = e^{-at} \cdot \varepsilon(t)$$

Se amostrarmos este sinal com um ritmo de amostragem $\omega_s = 2\pi/T_s$, obtemos um sinal discreto dado por

$$h(k) = e^{-akT_s} \cdot \varepsilon(k) = (e^{-aT_s})^k \cdot \varepsilon(k)$$

Finalmente, aplicando a transformação em z inversa, obtemos a função de transferência com $h(k)$ como resposta impulsiva

$$G(z) = \frac{z}{z - e^{-aT_s}}$$

Conclui-se assim que um sistema contínuo com um polo em $p = -a$ tem uma resposta impulsiva coincidente com o sistema discreto com um polo em $p = e^{-aT_s}$.

Exemplo 10.6

Considere o sistema contínuo dado por

$$G(s) = \frac{1}{s + 0.2}$$

com constante de tempo $\tau = 1/a = 5$ s. A resposta impulsiva é então dada por

$$h(t) = e^{-\frac{t}{5}} \cdot \varepsilon(t)$$

como representado na Figura 10.5.

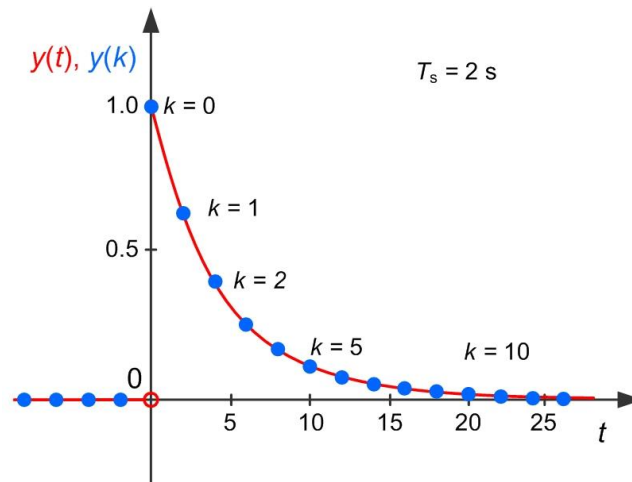


Figura 10.5 – Respostas impulsivas equivalentes com $a = 0.2$ e $T_s = 2$ s.

O sistema equivalente discreto com tempo de amostragem $T_s = 2.0$ s é dado por

$$G(z) = \frac{z}{z - e^{-0.2 \cdot 2.0}} = \frac{z}{z - 0.6703}$$

com resposta impulsiva dada por

$$h(k) = 0.6703^k \cdot \varepsilon(k)$$

também representada na Figura 10.5.

Observa-se que as respostas coincidem nos instantes discretos tornando estes sistemas, contínuo e discreto, equivalentes num certo sentido.

■

Para já registemos que, quando consideramos polos da função de transferência do sistema contínuo sobre o eixo real, existe uma transposição para os polos da função de transferência do sistema discreto que respeita a fórmula

$$z = e^{s \cdot T_s}$$

Na realidade, esta fórmula pode ser estendida a todas as localizações do plano s e correspondente **mapeamento** no plano z .

O próximo capítulo dedica-se ao problema da conversão entre sistemas em tempo contínuo e seus equivalentes em tempo discreto, e vice-versa.

10.6. Respostas de sistemas de primeiro grau

Antes de se fechar este capítulo, interessa estudar como a localização no plano z dos polos e zeros da função de transferência determina o tipo de resposta temporal dos sistemas.

Começemos pelo sistema de 1º grau descrito pela função de transferência

$$G(z) = \frac{z}{z - a}$$

onde o valor do parâmetro a determina a localização do polo, e logo, a estabilidade assintótica do sistema. Quando o parâmetro a , que coincide com a localização do único polo, está no interior do círculo unitário i.e. $-1 < a < 1$, a resposta impulsiva deste sistema

$$h(k) = a^k \varepsilon(k)$$

irá convergir para zero. No entanto, se o parâmetro a for negativo iremos ter uma resposta que irá trocar de sinal a cada amostra. Significa que conseguimos ter uma resposta oscilante mesmo com um sistema de primeiro grau, que não acontecia nos sistemas contínuos.

Exemplo 10.7

Considere o sistema descrito pela função de transferência

$$G(z) = \frac{z}{z - 0.5}$$

com polo em $p = 0.5$, estável.

A Figura 10.6 apresenta a resposta impulsiva convergente para zero.

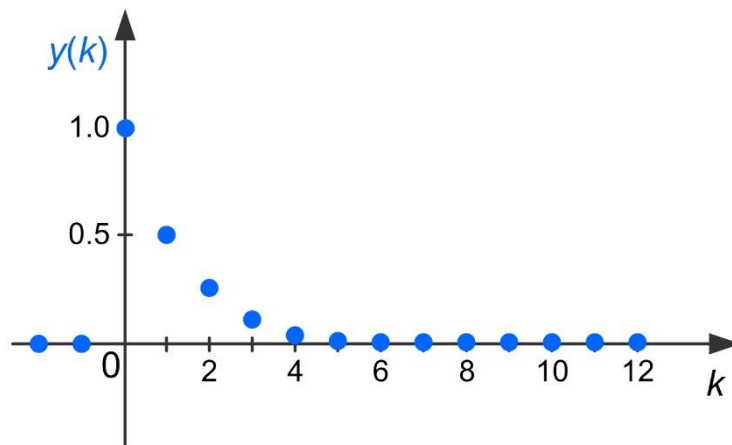


Figura 10.6 - Resposta impulsiva do sistema de 1º grau com $a = 0.5$.



Exemplo 10.8

Considere agora o sistema descrito pela função de transferência

$$G(z) = \frac{z}{z + 0.5}$$

com polo em $p = -0.5$, estável.

A Figura 10.7 apresenta a resposta impulsiva também convergente para zero mas de forma oscilatória.

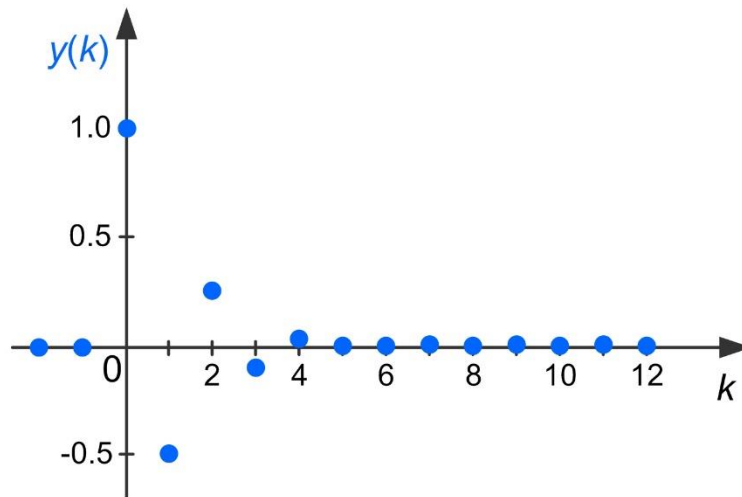


Figura 10.7 - Resposta impulsiva do sistema de 1º grau com $a = -0.5$.

■

Exemplo 10.9

Considere agora o sistema descrito pela função de transferência

$$G(z) = \frac{z}{z - 1}$$

com polo em $p = 1$, não estável.

A Figura 10.8 apresenta a resposta impulsiva que coincide com o degrau unitário, logo não convergente.

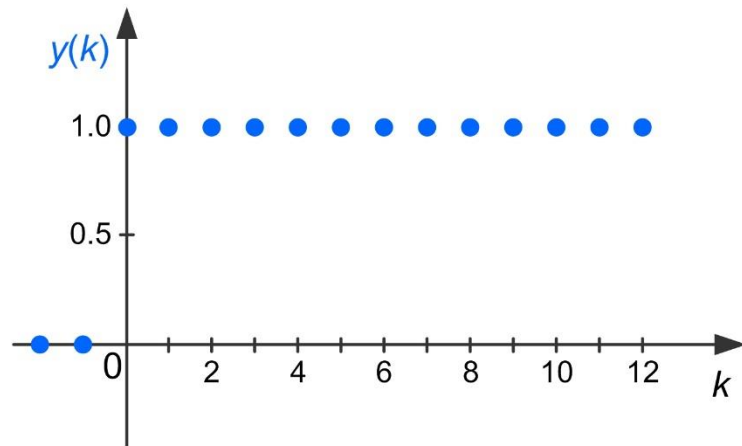


Figura 10.8 - Resposta impulsiva do sistema de 1º grau com $a = 1$.

■

Exemplo 10.10

Considere agora o sistema descrito pela função de transferência

$$G(z) = \frac{z}{z + 1}$$

com polo em $p = -1$, não estável.

A Figura 10.9 apresenta a resposta impulsiva que oscila entre os valores $+1$ e -1 , e logo não convergente.

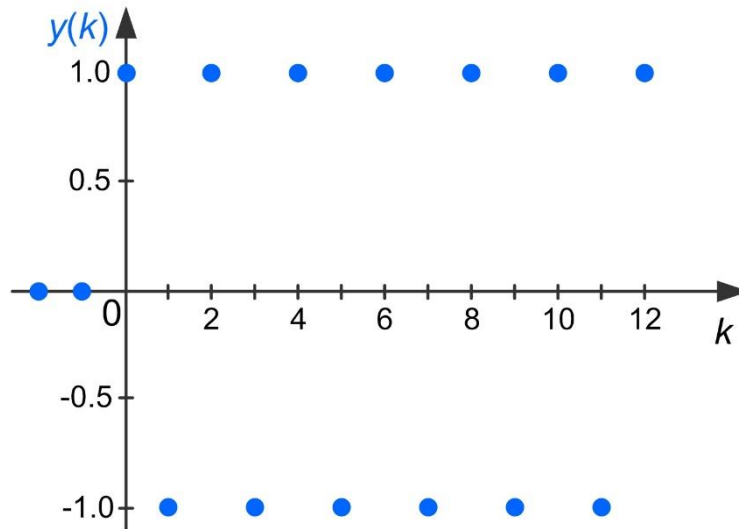


Figura 10.9 - Resposta impulsiva do sistema de 1º grau com $a = -1$.

■

10.7. Respostas de sistemas de segundo grau

Considere-se o sistema de 2º grau descrito pela função de transferência

$$G(z) = \frac{a \cdot \sin \Omega_0 \cdot z}{z^2 - 2 \cdot a \cdot \cos \Omega_0 z + a^2}$$

Cujos polos estão localizados em

$$p = a \cdot \cos \Omega_0 \pm j \cdot a \cdot \sin \Omega_0$$

ou seja, a uma distância a da origem e com um ângulo com o eixo real de Ω_0 . O valor do parâmetro a determina novamente a estabilidade assintótica do sistema. Quando o parâmetro a respeita a condição $-1 < a < 1$, a resposta impulsiva deste sistema

$$h(k) = a^k \cdot \sin \Omega_0 k \cdot \varepsilon(k)$$

irá convergir para zero. Novamente, se o parâmetro a for negativo iremos ter uma resposta que irá trocar de sinal a cada amostra.

Exemplo 10.11

Considere o sistema descrito pela função de transferência

$$G(z) = \frac{0.8 \cdot \sin \frac{\pi}{3} \cdot z}{z^2 - 2 \cdot 0.8 \cdot \cos \frac{\pi}{3} z + 0.8^2} = \frac{0.6928z}{z^2 - 0.8z + 0.64}$$

com polos localizados em $p = 0.4 \pm j \cdot 0.6928$, estável. Note-se que o zero na origem é responsável apenas pela manutenção da diferença no número de polos e zeros a um.

A Figura 10.10 apresenta a resposta impulsiva convergente para zero de forma oscilatória.

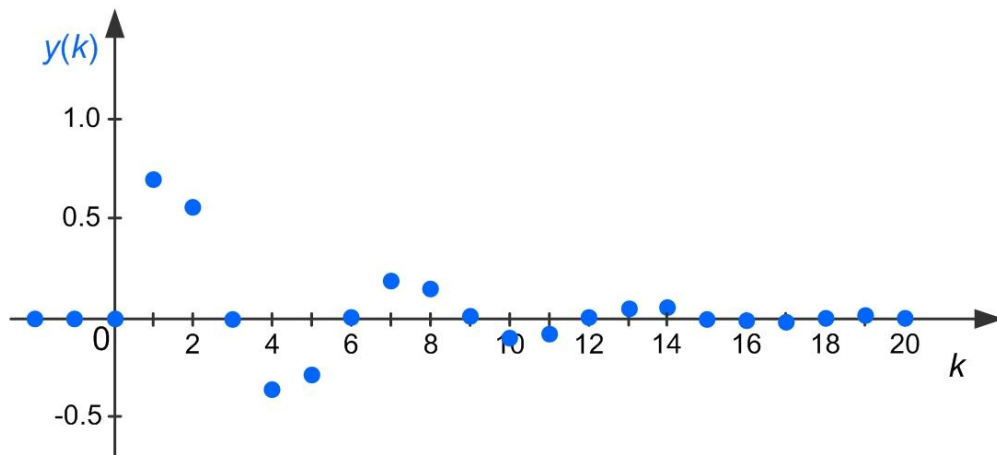


Figura 10.10 - Resposta impulsiva do sistema de 2º grau com $a = 0.8$ e $\Omega_0 = \pi/3$.

■

11. Métodos de discretização

Quando se abordou o problema da seleção do período de amostragem, referiu-se que existem duas alternativas principais no projeto de sistemas de controlo por computador:

1. a substituição por emulação de um controlador contínuo $C(s)$, já existente ou projetado no contínuo para o efeito, pelo seu equivalente discreto $C(z)$; ou
2. o projeto de raiz de um controlador em tempo discreto a partir do equivalente discreto $F(z)$ da instalação contínua descrita por $F(s)$.

Referiu-se ainda que, para o mesmo desempenho do sistema de controlo conseguem-se valores mais reduzidos do ritmo de amostragem seguindo a abordagem 2. ou seja, projetar o controlador $C(z)$ de raiz no domínio do tempo discreto.

Este capítulo dedica-se ao problema da obtenção do equivalente discreto de um sistema contínuo que, nas abordagens descritas acima, pode ser um controlador (abordagem 1.) ou uma instalação (abordagem 2.). No entanto, pela forma particular como cada um destes dois elementos se inserem no sistema de controlo, os **métodos de discretização** adequados serão também diferentes para discretizar uma instalação ou um controlador. Note-se que seguiremos apenas umas das duas abordagens possíveis e nesse sentido não teremos, no mesmo processo de projeto, de realizar simultaneamente a discretização de uma instalação e de um controlador.

11.1. Discretização de controladores contínuos

Considere-se o esquema de blocos da Figura 11.1 com um sistema de controlo no domínio da variável s de Laplace, i.e. no domínio do tempo contínuo.

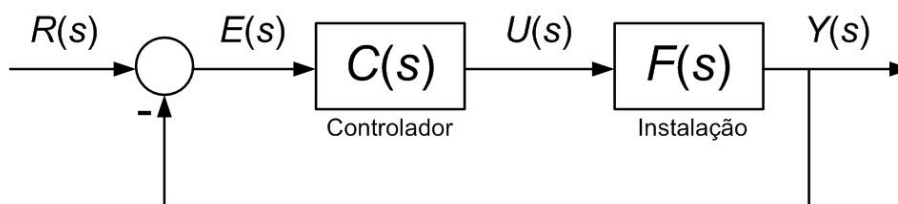


Figura 11.1 – Sistema de controlo em tempo contínuo.

Estamos interessados em substituir o controlador contínuo por um controlador discreto $C(z)$ mantendo o mesmo desempenho do sistema de controlo, em anel fechado, que se observava em tempo contínuo. Ou seja, queremos que o controlador discreto, programado no computador, **emule** o comportamento do controlador contínuo realizado numa qualquer tecnologia analógica, e.g. eletrónica analógica.

Temos então de encontrar o controlador discreto $C(z)$ que colocado na estrutura da Figura 11.2 apresente o mesmo comportamento dinâmico do controlador $C(s)$.

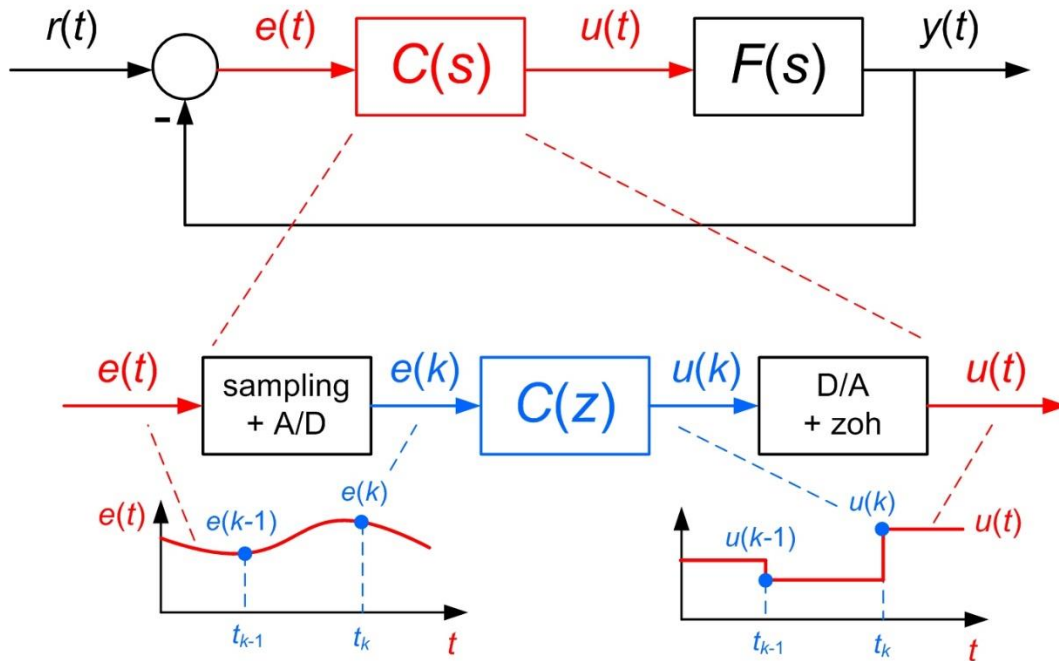


Figura 11.2 – Substituição do controlador contínuo por emulação.

O controlador contínuo é substituído por um conjunto com três partes:

- a aquisição do sinal relativo ao erro por amostragem (*sampling*) do sinal contínuo $e(t)$ e correspondente digitalização i.e. conversão analógico-digital (A/D);
- a realização do controlador discreto que recebe como entrada a amostra digitalizada $e(k)$ para produzir o sinal relativo à ação de controlo $u(k)$; e
- a conversão do sinal de atuação de volta ao tempo contínuo $u(t)$ através do conversor digital-analógico com retenção de ordem zero i.e. com manutenção do valor registado constante até ao instante de amostragem seguinte.

Ao contrário do controlador contínuo, devido ao processo de amostragem, o controlador discreto não tem conhecimento sobre a evolução do sinal do erro $e(t)$ entre instantes de amostragem.

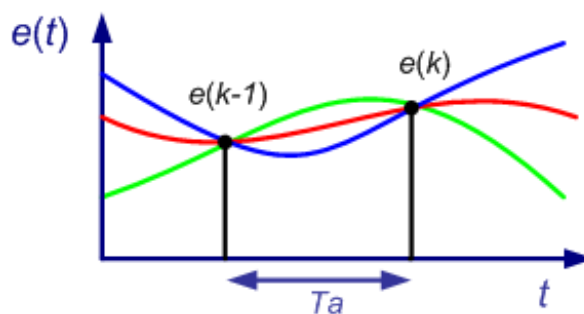


Figura 11.3 – Infinitude de possibilidades de evolução do sinal $e(t)$ entre amostras.

De facto, como se representa na Figura 11.3, existe uma infinidade de possibilidades para a evolução do sinal entre amostras. Assim, o processo de discretização terá que assumir algum tipo de comportamento do sinal entre instantes de amostragem.

O sinal de erro pode assumir diversos comportamentos entre amostras, mas o cálculo da ação de controlo pelo controlador discreto é baseado apenas nas amostras $e(k)$.

Isto significa que a emulação do controlador pelo computador resulta numa aproximação do comportamento existente em tempo contínuo.

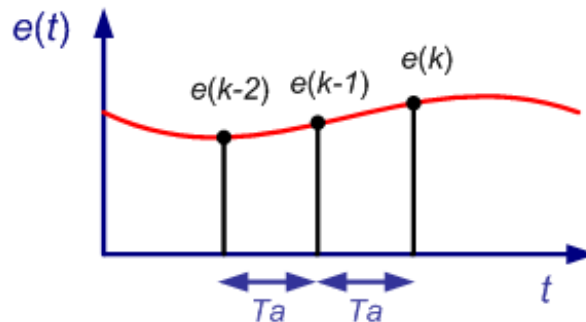


Figura 11.4 – A redução do período de amostragem T_s melhora a aproximação.

A aproximação do controlador contínuo pelo controlador discretizado é tanto melhor quanto menor for o período de amostragem (Figura 11.4). No entanto, recorde-se a existência de vários motivos que impedem a redução do período de amostragem.

Conclui-se assim, que independentemente do método a utilizar, o processo de discretização do controlador é aproximado, i.e. que se observarão diferenças de desempenho entre o sistema de controlo com os controladores contínuo e discreto. Este é o motivo pelo qual a abordagem 1., descrita acima, exige um ritmo de amostragem ainda mais conservador. Note-se que não há forma de condicionar o sinal entre amostras visto que o sinal do erro deriva diretamente da saída da instalação $y(t)$.

11.2. Métodos de discretização por aproximação

Neste documento referir-se-ão três métodos para a discretização de controladores por aproximação:

- método de Euler *backward*;
- método de Euler *forward*;
- método de Tustin ou da bilinear.

Métodos de Euler

À função de transferência em tempo contínuo está sempre associada uma equação diferencial linear ordinária. Os métodos de Euler consistem em aproximar as derivadas no ponto em

análise pelas diferenças entre o valor do sinal em dois instantes de amostragem consecutivos, como representado na Figura 11.5.

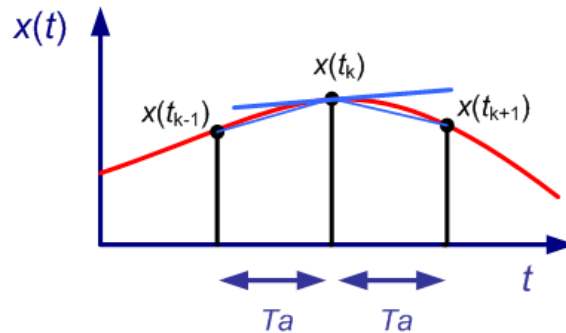


Figura 11.5 – Aproximações à derivada nos métodos de Euler.

O método de Euler *backward* realiza essa aproximação considerando o instante presente e o instante anterior, ou seja

$$\frac{dx(t)}{dt} \approx \frac{x(t_k) - x(t_{k-1})}{T_s}$$

e o método de Euler *forward* realiza essa aproximação considerando o instante presente e o instante posterior, ou seja

$$\frac{dx(t)}{dt} \approx \frac{x(t_{k+1}) - x(t_k)}{T_s}$$

Assim, considere-se um bloco integrador puro com função de transferência

$$C(s) = \frac{U(s)}{E(s)} = \frac{1}{s}$$

com a correspondente equação diferencial dada por

$$\frac{du(t)}{dt} = e(t)$$

No método de Euler *backward* esta operação de derivação é substituída pela equação às diferenças

$$\frac{u(t_k) - u(t_{k-1})}{T_s} = e(t_k)$$

com a correspondente função de transferência dada por

$$C(z) = \frac{U(z)}{E(z)} = \frac{T_s}{1 - z^{-1}} = \frac{T_s \cdot z}{z - 1}$$

e no método de Euler *forward* a operação de derivação é substituída pela equação às diferenças

$$\frac{u(t_{k+1}) - u(t_k)}{T_s} = e(t_k)$$

com a correspondente função de transferência dada por

$$C(z) = \frac{U(z)}{E(z)} = \frac{T_s}{z - 1}$$

A Figura 11.6 apresenta outra interpretação possível dos métodos de Euler *backward* e *forward*. A partir da função de transferência original do integrador em tempo contínuo podemos escrever

$$u(t_k) - u(t_{k-1}) = \int_{t_{k-1}}^{t_k} e(\sigma) d\sigma \approx \begin{cases} T_s e(t_k) & \text{backward} \\ T_s e(t_{k-1}) & \text{forward} \end{cases}$$

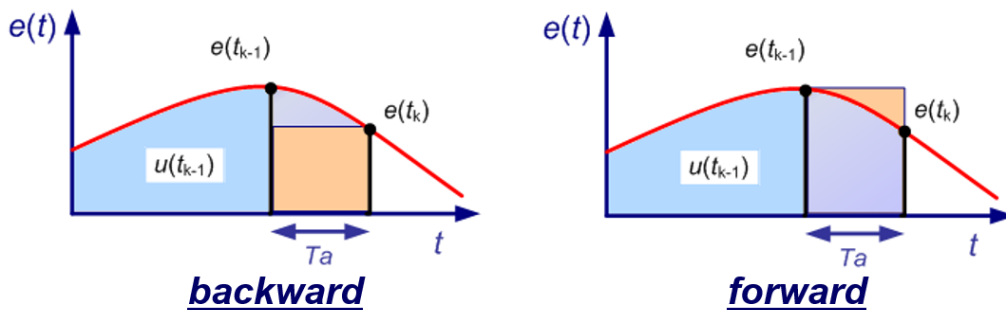


Figura 11.6 – Área de integração nos métodos de Euler.

Os métodos de Euler, sendo os mais simples, exigem ritmos de amostragem elevados.

Método de Tustin ou da bilinear

Na sequência da interpretação anterior, o método de Tustin propõe que a área de integração seja calculada através de um trapézio com área equivalente à média das obtidas pelos dois métodos de Euler,

$$u(t_k) - u(t_{k-1}) = \int_{t_{k-1}}^{t_k} e(\sigma) d\sigma \approx \frac{e(t_k) + e(t_{k-1})}{2} T_s$$

como representado na Figura 11.7. Esta aproximação corresponde a uma função de transferência como equivalente do integrador, dada por

$$C(z) = \frac{U(z)}{E(z)} = \frac{T_s}{1 - z^{-1}} = \frac{T_s}{2} \cdot \frac{z + 1}{z - 1}$$

Este método é também conhecido por aproximação trapezoidal ou método da bilinear.

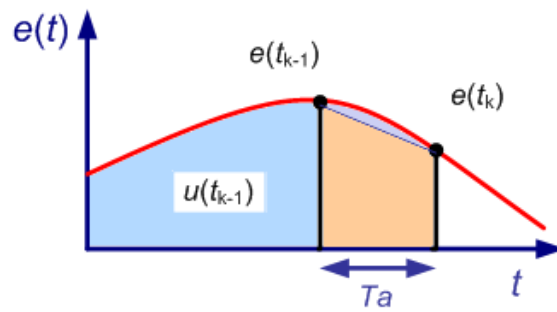


Figura 11.7 – Área de integração no método de Tustin.

O bloco integrador como o bloco dinâmico elementar

Qualquer sistema linear pode ser colocado como o resultado da composição de blocos estáticos (ganhos e somas) e blocos integradores (dinâmicos).

Exemplo 11.1

Considere-se o sistema dado pela função de transferência

$$C(s) = \frac{U(s)}{E(s)} = \frac{s}{s^2 + 2s + 3}$$

É possível reescrever a função de transferência como

$$C(s) = \frac{U(s)}{E(s)} = \frac{s}{s^2 + 2s + 3} = \frac{\frac{1}{s}}{\left(1 + \frac{2}{s}\right) + 3 \cdot \frac{1}{s^2}} = \frac{\frac{1}{s}}{1 + 3 \cdot \frac{1}{s} \cdot \frac{1}{\left(1 + 2 \cdot \frac{1}{s}\right)}}$$

A Figura 11.8 mostra a realização em esquema de blocos relativa à função de transferência como apresentada acima.

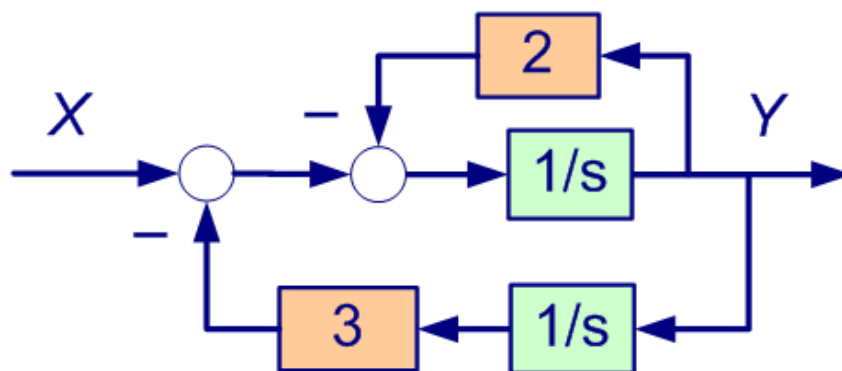


Figura 11.8 – Esquema de blocos relativo ao Exemplo 11.1.

■

Este resultado mostra que é possível isolar a componente dinâmica do sistema na forma de blocos integradores.

Os métodos de Euler e Tustin propõem que agora os blocos integradores contínuos sejam substituídos pelos seus equivalentes discretos segundo a tabela

Método	Processo de substituição
Euler <i>backward</i>	$\frac{1}{s} \rightarrow \frac{T_s \cdot z}{z - 1}$
Euler <i>forward</i>	$\frac{1}{s} \rightarrow \frac{T_s}{z - 1}$
Tustin	$\frac{1}{s} \rightarrow \frac{T_s}{2} \cdot \frac{z + 1}{z - 1}$

Esta substituição é equivalente a substituir na função de transferência (contínua) a variável s pelos seguintes equivalentes:

Método	Processo de substituição
Euler <i>backward</i>	$s \rightarrow \frac{z - 1}{T_s \cdot z}$
Euler <i>forward</i>	$s \rightarrow \frac{z - 1}{T_s}$
Tustin	$s \rightarrow \frac{2}{T_s} \cdot \frac{z - 1}{z + 1}$

Exemplo 11.2

Considere-se o processo de discretização, pelo método de Tustin com $T_s = 2$ s, do controlador contínuo dado por:

$$C(s) = \frac{2}{s + 4}$$

O método aplica-se substituindo a variável s pelo seu equivalente para o caso do método de Tustin, i.e.

$$C(z) = \frac{2}{\left(\frac{2}{T_s} \cdot \frac{z - 1}{z + 1}\right) + 4} = \frac{2}{\left(\frac{z - 1}{z + 1}\right) + 4} = \frac{2z + 2}{5z + 3}$$

■

Mapeamento do domínio s no domínio z dos métodos aproximados

Na Figura 11.9, observa-se o mapeamento dos 3 métodos aproximados entre o domínio da variável s e o domínio da variável z . Observa-se que o método de Tustin é o único que faz um mapeamento perfeito entre as regiões estável no contínuo e a estável no discreto.

No método de Euler *forward* a imagem de um sistema estável pode ser instável no domínio z .

No método de Euler *backward* a imagem de um sistema instável no domínio s pode ser estável no domínio z .

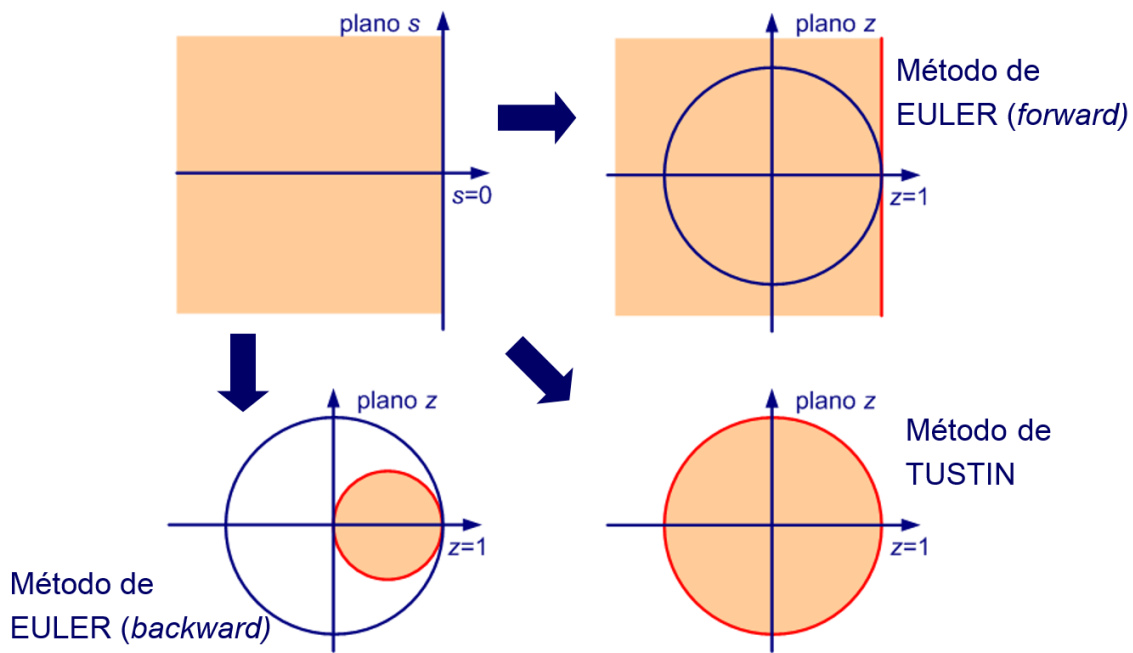


Figura 11.9 – Mapeamento de s para z dos métodos aproximados.

11.3. Discretização de instalações

Considere-se agora o projeto de controladores em tempo discreto desenvolvidos de raiz, ou seja, sem qualquer controlador em tempo contínuo como ponto de partida. Interessamos agora realizar o projeto na perspectiva que a própria instalação existe no domínio do tempo discreto (embora saibamos que isso não é verdade). Esta perspectiva está representada no esquema de blocos da Figura 11.10.

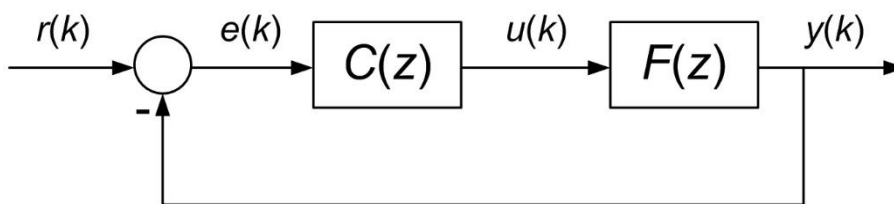


Figura 11.10 – Sistema de controlo em tempo discreto.

Podemos então fazer a pergunta: qual é a função de transferência no domínio z equivalente à instalação contínua vista através do processo de amostragem que nos liga ao mundo contínuo?

Temos então de encontrar a instalação discreta $F(z)$ que colocada na estrutura da Figura 11.11 apresente o mesmo comportamento dinâmico da instalação $F(s)$.

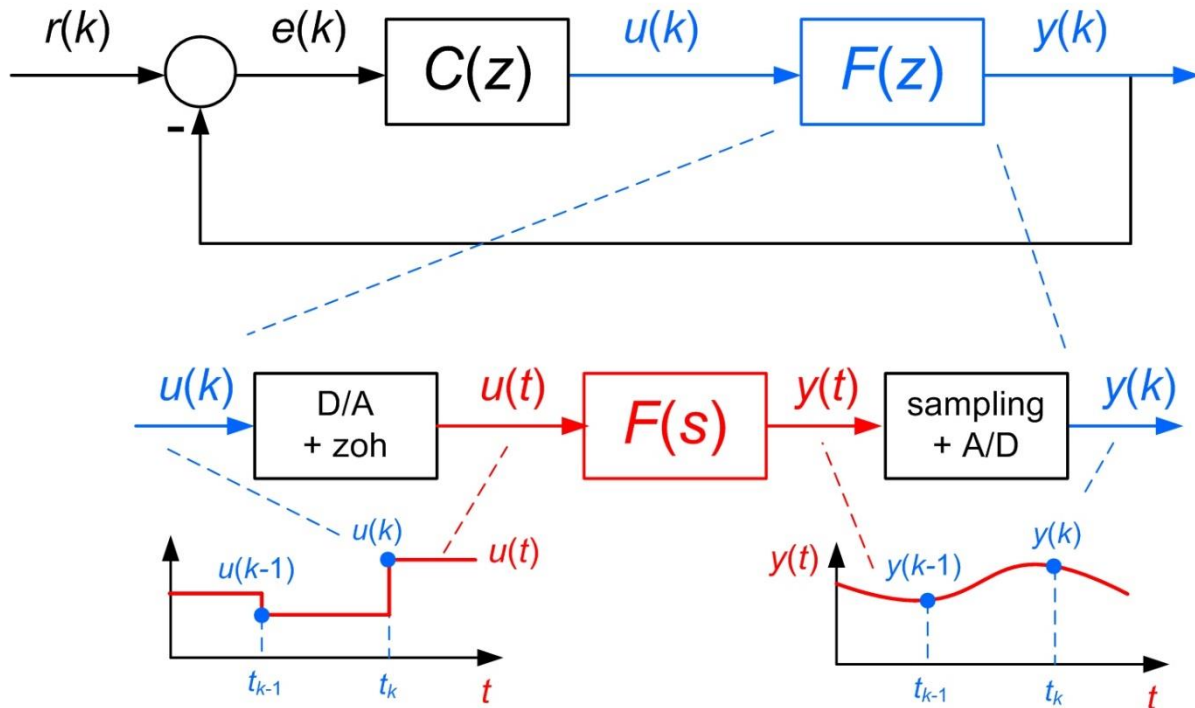


Figura 11.11 – Substituição do controlador contínuo por emulação.

O controlador discreto é o equivalente a um conjunto com três partes:

- a conversão do sinal de atuação $u(k)$ para o tempo contínuo $u(t)$ através do conversor digital-analógico com retenção de ordem zero i.e. com manutenção do valor registado constante até ao instante de amostragem seguinte
- a dinâmica da instalação em tempo contínuo descrita pela função de transferência $F(s)$ que recebe como entrada a excitação em degraus $u(t)$ e reagem com o sinal medido $y(t)$; e
- a aquisição do sinal medido por amostragem (*sampling*) do sinal contínuo $y(t)$ e correspondente digitalização i.e. conversão analógico-digital (A/D);

A principal diferença face ao problema da discretização de controladores contínuos, apresentado nas secções anteriores, é que agora o sinal de entrada do sistema a discretizar, i.e. a instalação $F(s)$, é bem conhecido entre instantes de amostragem. Visto que a excitação é realizada através do mecanismo com retenção de ordem zero – *zero-order hold* – podemos realizar a discretização de forma exata.

11.4. Método de discretização ZOH

O método de discretização ZOH (de *zero-order hold*) determina qual a função de transferência $F(z)$ que é equivalente à função de transferência $F(s)$ quando ambas são excitadas por um degrau. A equivalência é traduzida pela coincidência das amostras discretas com o sinal contínuo nos instantes de amostragem, ou seja

$$y(k) = y(t_k)$$

onde $y(k)$ é o sinal em tempo discreto, saída de $F(z)$ e $y(t_k)$ é o sinal em tempo contínuo nos instantes particulares $t_k = k \cdot T_s$. A Figura 11.12 ilustra o problema de equivalência que procuramos resolver.

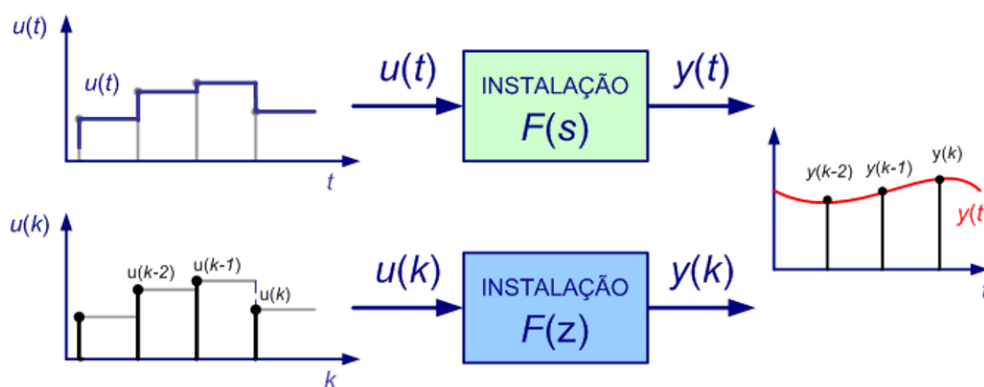


Figura 11.12 – Equivalência ZOH contínuo discreto do bloco instalação.

Como as instalações, contínua e equivalente discreta, são descritas por funções de transferência, são ambas lineares. Logo, basta procurar a equivalência para a excitação com um degrau unitário. A linearidade é então responsável pela generalização a sequências de degraus de amplitudes diversas.

Do lado contínuo podemos escrever

$$y(t_k) = \text{TL}^{-1} \left\{ F(s) \cdot \frac{1}{s} \right\} \Big|_{t=t_k=kT_s}$$

e do lado discreto

$$y(k) = \text{TZ}^{-1} \left\{ F(z) \cdot \frac{1}{1-z^{-1}} \right\}$$

ou seja, em ambos os casos as respetivas funções de transferência multiplicadas pelas transformadas dos degraus, e transportadas para os domínios do tempo.

Igualando,

$$\text{TZ}^{-1} \left\{ F(z) \cdot \frac{1}{1-z^{-1}} \right\} = \text{TL}^{-1} \left\{ F(s) \cdot \frac{1}{s} \right\} \Big|_{t=t_k=kT_s}$$

ou, de forma equivalente,

$$F(z) = (1 - z^{-1}) \cdot \text{TZ} \left\{ \text{TL}^{-1} \left\{ F(s) \cdot \frac{1}{s} \right\} \Big|_{t=kT_s} \right\}$$

sendo esta a fórmula do **método de discretização ZOH**.

Exemplo 11.3

Considere-se o processo de discretização, pelo método de discretização ZOH com $T_s = 0.2$ s, da instalação dada por:

$$F(s) = 10 \frac{s + 2}{s + 4}$$

Aplicando o método por passos, calculamos

$$F(s) \cdot \frac{1}{s} = 10 \frac{s + 2}{s + 4} \cdot \frac{1}{s}$$

aplicamos a transformada de Laplace inversa

$$\text{TL}^{-1} \left\{ F(s) \cdot \frac{1}{s} \right\} = \text{TL}^{-1} \left\{ 10 \frac{s + 2}{s + 4} \cdot \frac{1}{s} \right\} = 5(e^{-4t} + 1)\varepsilon(t)$$

substituímos t por $k \cdot T_s$ e aplicamos a transformação em z

$$\text{TZ} \left\{ \text{TL}^{-1} \left\{ F(s) \cdot \frac{1}{s} \right\} \Big|_{t=kT_s} \right\} = \text{TZ} \{ 5(e^{-4kT_s} + 1)\varepsilon(k) \} = \frac{5z}{z - e^{-4T_s}} + \frac{5z}{z - 1}$$

e finalmente multiplicamos por $(1 - z^{-1})$ e simplificamos a expressão, resultando

$$F(z) = \frac{10z - 7.247}{z - 0.4493}$$

■

Pela forma como está proposto, verifica-se que o método ZOH mapeia os polos existentes no plano s em z através da relação

$$z = e^{s \cdot T_s}$$

Esta relação não é conforme no sentido em que existem pontos diferentes em s que são mapeados para o mesmo ponto em z . Isto sucede quando a componente imaginária dos polos está acima, em módulo, do valor da frequência de Nyquist.

12. Controlo por retroação em tempo discreto

No capítulo 5 apresentaram-se os diversos tipos de sistemas de controlo em tempo contínuo, nomeadamente:

- sistema de comando
- sistema de seguimento
- sistema de regulação

Estas tipologias continuam a fazer sentido em tempo discreto, sendo a única diferença a forma como o sistema de controlo é realizado. As técnicas de projeto do sistema de controlo têm de ser também adaptadas às limitações introduzidas pelo sistema amostrado.

12.1. Especificação do sistema de controlo

A especificação do sistema de controlo para suportar um projeto a partir da função de transferência e focado no posicionamento dos polos segue uma abordagem semelhante ao que vimos em tempo contínuo. Na Figura 12.1 repete-se a resposta temporal genérica onde se identificam os principais parâmetros que caracterizam o comportamento desejado para um sistema de controlo.

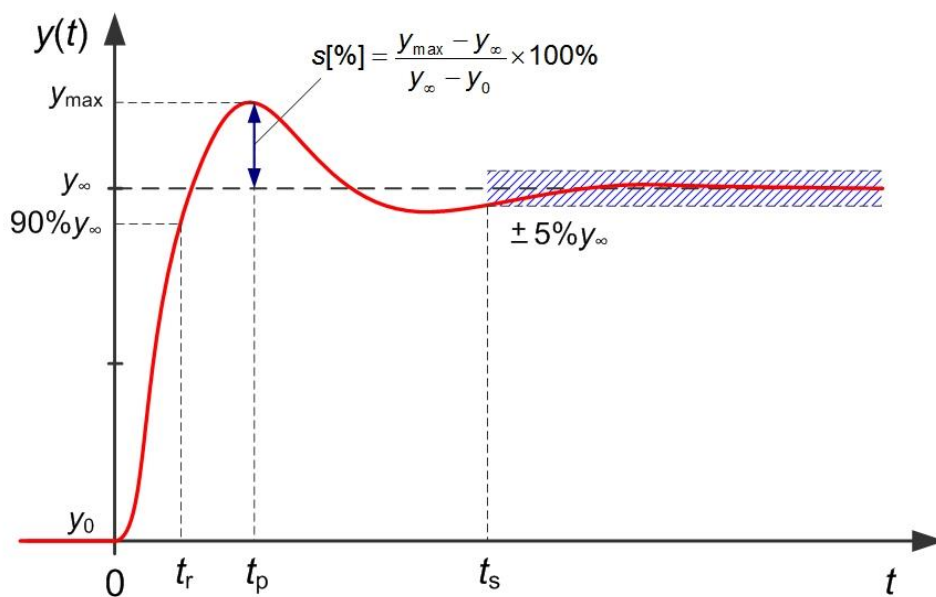


Figura 12.1 – Parâmetros de especificação no domínio do tempo (resposta ao degrau).

Vimos já também que é possível traduzir a maioria dos conjuntos de parâmetros de especificação por uma função de transferência de 2ª ordem sem zeros dada por

$$H^*(s) = \frac{K \omega_n^2}{s^2 + 2D\omega_n s + \omega_n^2}$$

a que chamaremos **modelo de especificação**.

Através da seleção adequada dos parâmetros K , D e ω_n ; representamos o modelo de comportamento dinâmico desejado que exhibe os mesmos parâmetros de especificação. Recomenda-se a leitura da secção 5.2 onde se introduziram uma diversidade de fórmulas relacionando os parâmetros γ_∞ , $s[\%]$, t_r e t_s ; com o conjunto K , D e ω_n .

Faz-se uma nota neste ponto para o leitor não confundir a quantidade ω_n relativa à frequência natural dos polos do sistema de 2ª ordem com a quantidade ω_N relativa à frequência de Nyquist e diretamente relacionada com o período de amostragem.

Uma vez encontrados os parâmetros K , D e ω_n de $H^*(s)$, como especificar o comportamento desejado para o sistema de controlo, ou seja, como determinar $H^*(z)$?

No capítulo anterior introduziu-se o método de discretização ZOH que se identificou como o mais adequado para a discretização de instalações pelo facto que a atuação sobre estas é realizada através da retenção do valor entre amostras, ou melhor dizendo, através de uma sequência de degraus de amplitudes diversas.

Agora, o sistema de controlo $H(z)$ tem como entrada o sinal de referência (ou comando) que na maioria das situações apresenta variações em degrau. Assim, o método adequado para discretizar o modelo de especificação $H^*(s)$ é também o ZOH.

Alternativamente, podemos aplicar diretamente a fórmula que mapeia o plano s em z , respeitada pelo método ZOH (ver capítulo anterior), dada por

$$z = e^{s \cdot T_s}$$

Exemplo 12.1

Considere-se a instalação dada pela função de transferência

$$F(s) = \frac{0.5}{(s + 1.0)(s + 0.5)}$$

com dois polos em $p_1 = -1.0$ e $p_2 = -0.5$; apresentando a resposta ao degrau unitário registada na Figura 12.2. Observa-se no gráfico da resposta que o sistema tem ganho estático unitário e apresenta um tempo subida a 90% de aproximadamente $t_r = 6$ s.

O responsável por esta instalação está interessado em obter um tempo de reação seis vezes mais rápido, i.e. $t_r = 1$ s; e sem sobrelevação assinalável, i.e. $s[\%] \leq 5\%$.

Através das fórmulas dos dois parâmetros podemos escrever

$$t_{r[90\%]} \approx \frac{3.7D}{\omega_n} = 1.0 \text{ s}$$

e

$$s[\%] = 100\% \times e^{\frac{-D\pi}{\sqrt{1-D^2}}} = 5\%$$

que resulta num coeficiente de amortecimento de $D = 0.7$ e frequência natural $\omega_n = 2.6$ rad/s.

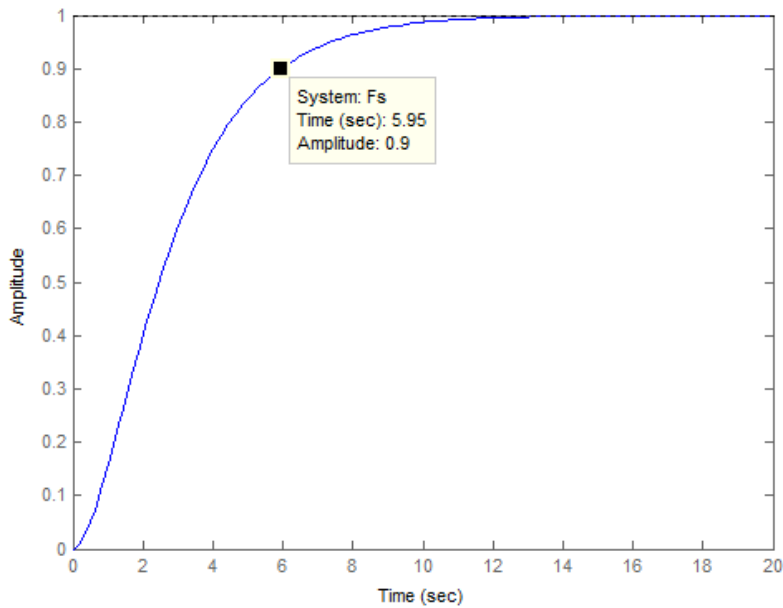


Figura 12.2 – Resposta ao degrau da instalação do Exemplo 12.1.

Assim, o modelo de especificação pode ser construído como ($K = 1$)

$$H^*(s) = \frac{K\omega_n^2}{s^2 + 2D\omega_n s + \omega_n^2} = \frac{6.76}{s^2 + 3.64s + 6.76}$$

com polos em $p_{1,2} = -1.82 \pm j 1.87$.

O tempo de amostragem deve ser selecionado através das regras apresentadas nos capítulos anteriores. A frequência máxima coincide, neste caso, com os polos dominantes que impõem a largura de banda, i.e. $\omega_{max} = 2.6$ rad/s. A frequência de Nyquist deve então estar compreendida no intervalo

$$\omega_N = (5 \cdots 10) \times \omega_{max} = 13.0 \cdots 26.0 \text{ rad/s}$$

ou seja,

$$T_s = \frac{\pi}{\omega_N} = 0.12 \cdots 0.24 \text{ s}$$

Então, aplicando o método ZOH ao modelo de especificação com o valor selecionado de $T_s = 0.15$ s, obtém-se o modelo de especificação no domínio z dado por

$$H^*(z) = \frac{0.06313z + 0.0526}{z^2 - 1.464z + 0.5793}$$

com polos em $p_{1,2} = 0.7318 \pm j 0.2092$.

A Figura 12.3 apresenta a resposta ao degrau do modelo de especificação em tempo discreto onde se pode observar que cumpre no limite as especificações.

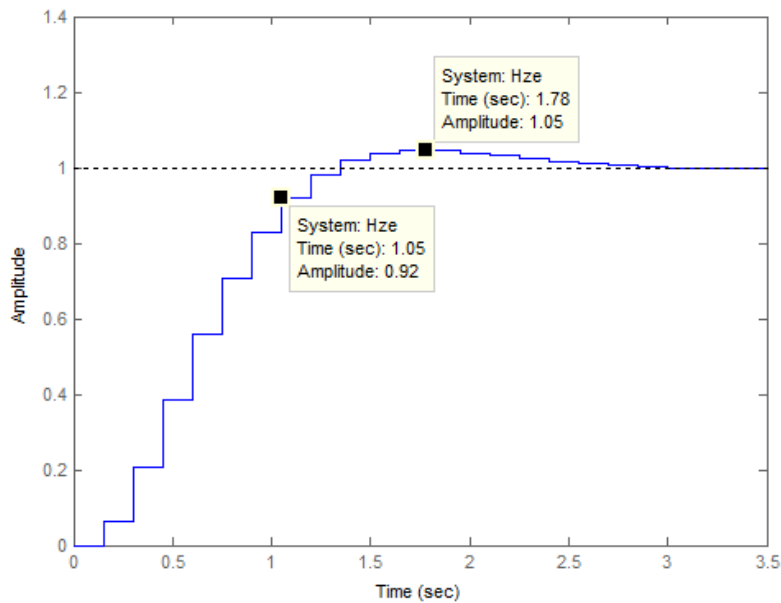


Figura 12.3 – Resposta ao degrau do modelo de especificação do Exemplo 12.1.

12.2. Diagrama de localização de raízes

Antes de fazermos uso da especificação, podemos primeiro colocar a questão relativa ao impacto da retroação sobre o comportamento dinâmico da instalação. Como o comportamento dinâmico depende diretamente da localização das raízes da função de transferência, a questão resume-se a determinar as localizações possíveis dos polos e zeros da função de transferência $H(z)$ do sistema de controlo.

No caso mais simples temos a retroação proporcional representada na Figura 12.4 que resulta num sistema de controlo com função de transferência dada por

$$H(z) = \frac{Y(z)}{R(z)} = \frac{K \cdot F(z)}{1 + K \cdot F(z)}$$

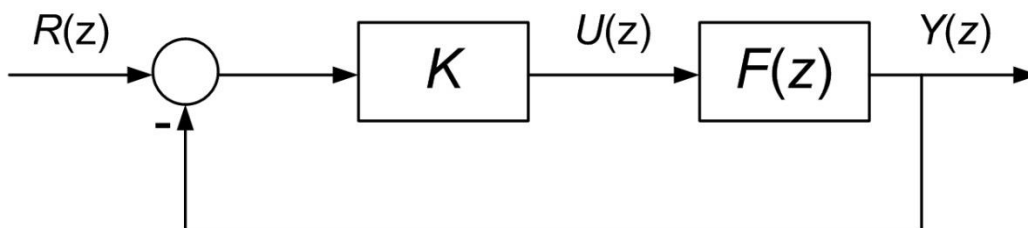


Figura 12.4 – Retroação proporcional no domínio z .

Esta expressão é, à parte da variável s trocada por z , exatamente igual à que obtivemos no capítulo 0 (secção 5.4) para o sistema de controlo no contínuo. Assim, tudo aquilo que foi

apresentado relativo à localização de raízes do anel fechado, nomeadamente a construção e análise do **diagrama de localização de raízes** (ou *root-locus*), aplica-se aos sistemas discretos sem qualquer modificação.

Em conclusão, as regras de realização do diagrama de localização de raízes são as mesmas para os sistemas representados nos domínios s ou z .

Existe apenas a diferença nas regiões de localização dos polos que resultam em sistemas estáveis nos contínuo e no discreto, como foi já referido. O diagrama de localização de raízes no domínio z deve ser analisado quanto à estabilidade contra a região definida pelo círculo unitário centrado na origem.

Exemplo 12.2

Considere-se de novo a instalação dada pela função de transferência

$$F(s) = \frac{0.5}{(s + 1.0)(s + 0.5)}$$

O equivalente discreto obtido pelo método ZOH (use-se novamente $T_s = 0.15$ s) é dado por

$$F(z) = \frac{0.005221z + 0.004844}{z^2 - 1.788z + 0.7985} = \frac{0.005221z + 0.004844}{(z - 0.9277)(z - 0.8607)}$$

com dois polos em $p_1 = 0.9277$ e $p_2 = 0.8607$; e um zero em $z_1 = -0.9277$.

As Figura 12.5 e Figura 12.6 apresentam o diagrama de localização de raízes do sistema de controlo com a instalação $F(z)$ e retroação proporcional.

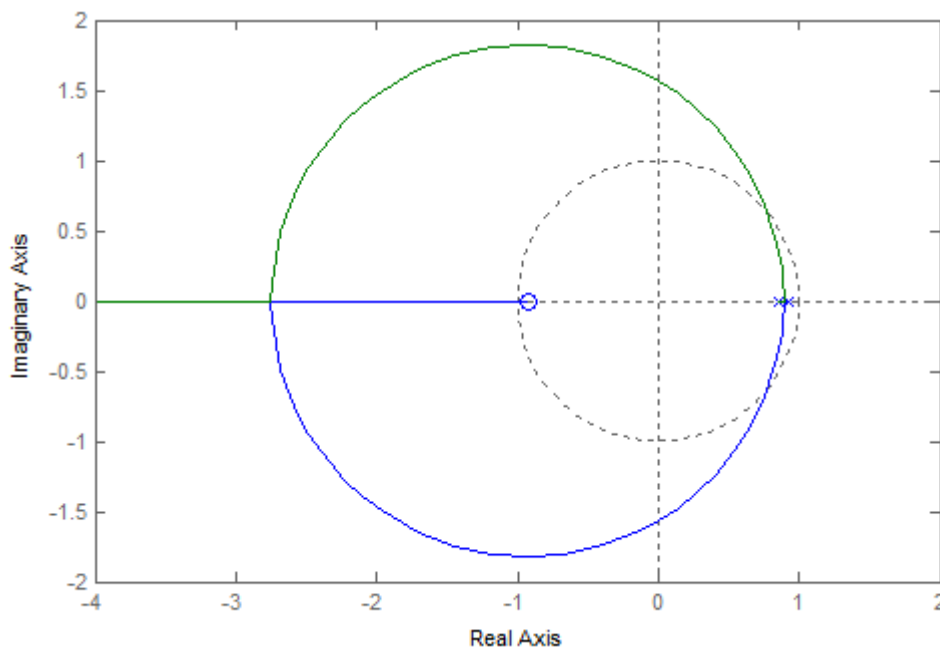


Figura 12.5 – Diagrama de localização de raízes do Exemplo 12.2.

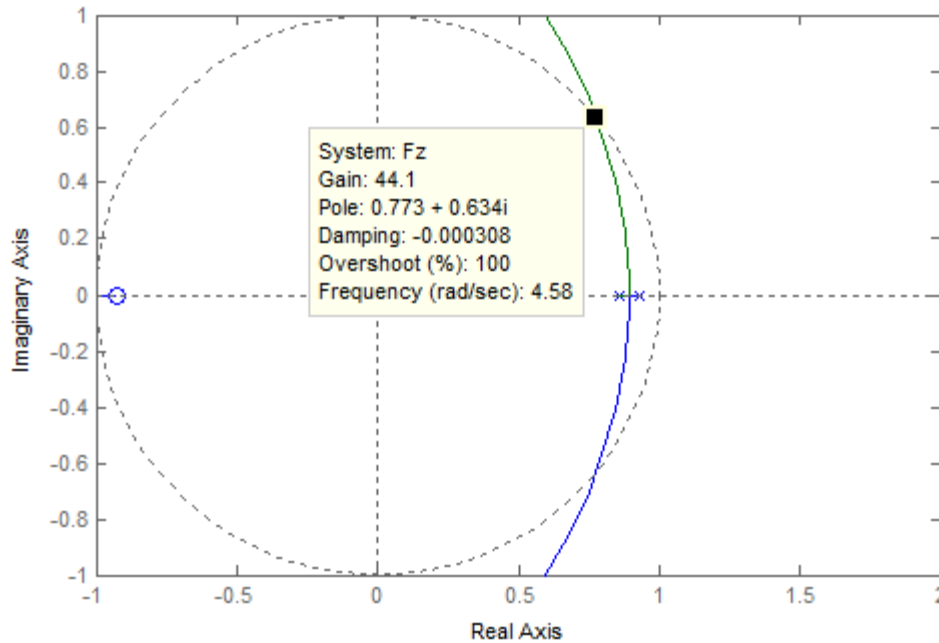


Figura 12.6 – Diagrama de localização de raízes do Exemplo 12.2 (detalhe).

O valor do ganho crítico, a partir do qual o sistema de controlo deixará de ser estável, pode ser encontrado aplicando o método de Jury, já descrito.

Assim, desenvolvendo a função de transferência no anel fechado da Figura 12.4 obtém-se

$$H(z) = \frac{K \cdot F(z)}{1 + K \cdot F(z)} = \frac{(0.005221z + 0.004844)K}{(z^2 - 1.788z + 0.7985) + (0.005221z + 0.004844)K}$$

que resulta num polinómio característico dado por

$$P(z) = z^2 - (1.788 - 0.005221K)z + (0.7985 + 0.004844K)$$

Como o polinómio é de grau 2, apenas é necessário verificar as três condições:

1. $-1 < 0.7985 + 0.004844K < 1$
2. $P(1) = 1 - 1.788 + 0.005221K + 0.7985 + 0.004844K = 0.0105 + 0.010065K > 0$
3. $P(-1) = 1 + 1.788 - 0.005221K + 0.7985 + 0.004844K = 2.5865 - 0.000377K > 0$

que resulta num intervalo de estabilidade para o ganho K de

$$-1.04 < K < 41.6$$

como se ilustra na Figura 12.6.

■

Exemplo 12.3

Considere-se ainda o desenvolvimento realizado no exemplo anterior e verifique-se a possibilidade de cumprir a especificação apresentada no Exemplo 12.1. A Figura 12.7 repete

o diagrama de localização de raízes mas acrescenta a localização desejada para os polos obtidas no Exemplo 12.1.

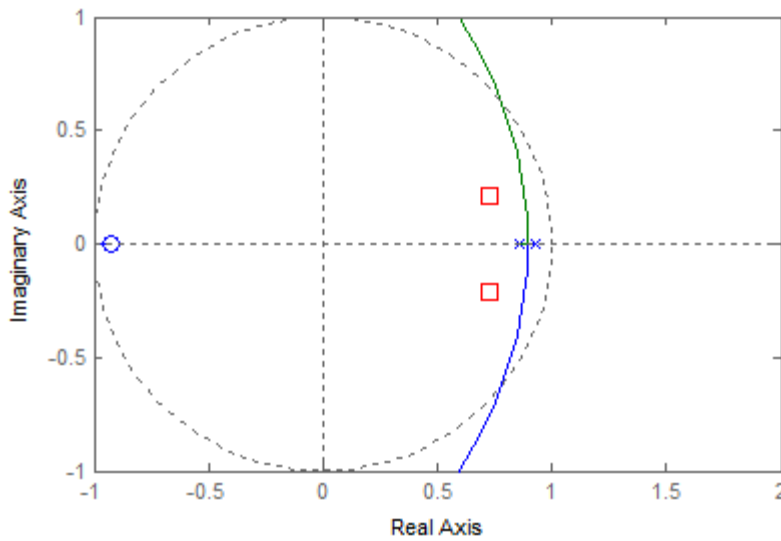


Figura 12.7 – Diagrama de localização de raízes e especificação dos polos.

No gráfico observa-se que o diagrama não atravessa a posição desejada, e logo não existe nenhum valor possível de ganho K que verifique a especificação.

Nesta situação, teremos de aumentar a complexidade do controlador acrescentando dinâmica adicional para além do ganho. Propõe-se assim um controlador com a estrutura

$$C(z) = K \frac{z - b}{z - a}$$

O conjunto do controlador e da instalação terão agora uma dinâmica de grau 3, a menos que haja cancelamentos polo-zero. Sugere-se usar o cancelamento como estratégia para o projeto do controlador.

Faça-se o zero do controlador igual ao polo mais rápido da instalação, ou seja

$$b = 0.8607$$

Temos agora um polinómio característico dado por

$$P(z) = (z - 0.9277)(z - a) + (0.005221z + 0.004844)K$$

que queremos igual ao denominador de $H^*(z)$ do **Exemplo 12.1**

$$P^*(z) = z^2 - 1.464z + 0.5793$$

Este problema tem solução analítica dada por

$$a = 0.58037; K = 8.4412$$

A Figura 12.8 mostra o diagrama de localização de raízes do anel aberto

$$G(z) = \frac{z - 0.8607}{z - 0.58037} \cdot F(z)$$

ou seja, do controlador ainda sem a seleção do valor do ganho K em série com a instalação. O novo diagrama atravessa a localização dos polos especificada e o ganho que corresponde à localização desejada coincide com o valor encontrado analiticamente.

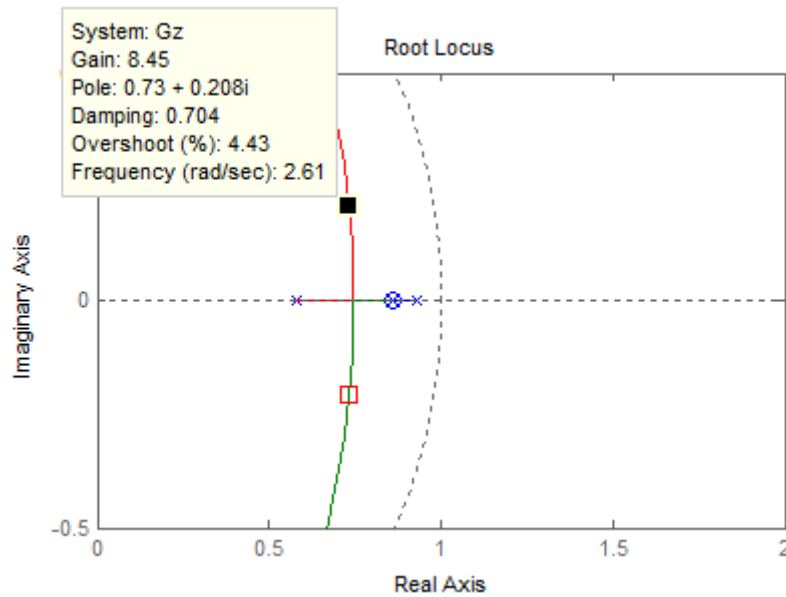


Figura 12.8 – Diagrama de localização de raízes e especificação dos polos.

A Figura 12.9 apresenta a resposta final obtida para o sistema de controlo com uma excitação em degrau no valor da referência.

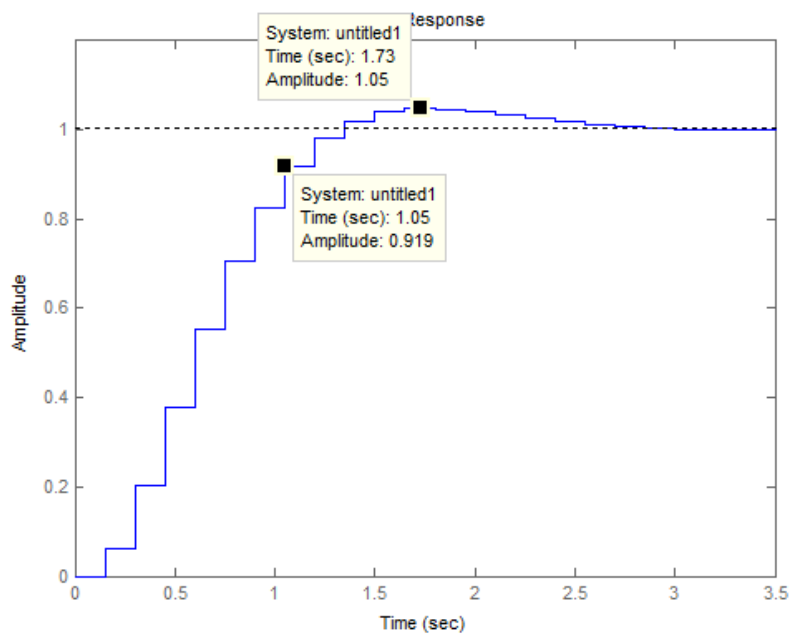


Figura 12.9 – Resposta ao degrau unitário do sistema de controlo (com correção de referência).

■

13. Resposta em frequência em tempo discreto

Foi referido no capítulo 6 que, quando a complexidade das instalações é muito elevada, a obtenção da função de transferência pela via da modelação física do processo pode ser uma tarefa quase irrealizável. Os métodos de resposta em frequência constituem uma alternativa aos métodos baseados na função de transferência de forma explícita e o capítulo presente descreve como podemos tirar partido dessas metodologias de controlo também em tempo discreto.

13.1. Sinal discreto sinusoidal

Antes de avançarmos para o controlo, importa introduzir (ou relembrar) alguns conceitos fundamentais da teoria dos sinais discretos.

Um sinal $x(k)$ diz-se sinusoidal de amplitude A , frequência ω e (desvio de) fase ϕ , se a sua variação no tempo discreto k puder ser representada pela função (Figura 13.1)

$$x(k) = A \sin(\Omega \cdot k + \phi)$$

A fase ϕ define o valor do argumento da função seno quando k é igual a zero e a amplitude A escala o valor máximo do sinal, visto que a saída da função seno varia no intervalo $[-1; +1]$.

A frequência angular Ω diz-nos quantos radianos são percorridos pelo argumento da função seno por cada instante de tempo discreto. É, portanto, medida em radianos, ou rad. Como a função seno é periódica com 2π (equivalente a 360°), define-se o período de tempo discreto N como

$$\Omega \cdot N = 2\pi$$

ou seja

$$\Omega = \frac{2\pi}{N}$$

Ora aqui aparece a primeira diferença face ao tempo contínuo. Em geral, o valor do período N , calculado por

$$N = \frac{2\pi}{\Omega}$$

não é necessariamente um número inteiro, a menos que a frequência Ω seja um divisor de 2π . Significa que o sinal sinusoidal não é necessariamente periódico em tempo discreto, no sentido em que pode não existir nenhum número inteiro N tal que

$$x(k + N) = x(k)$$

Este facto não retira potencial ao estudo da resposta em frequência em tempo discreto e focaremos a nossa atenção na frequência da senoide contínua que está subjacente, mesmo que amostrada de forma assíncrona relativamente ao período.

Na perspetiva que nos interessa analisar, ou seja, na ligação com o domínio do tempo contínuo onde se encontram as instalações a controlar, a frequência angular Ω relaciona-se com a frequência angular ω , em rad/s, através da expressão

$$\Omega = \omega \cdot T_s$$

Uma vez fixado o valor do período de amostragem T_s , é indiferente referirmos a frequência angular Ω , em rad, ou a frequência angular ω , em rad/s. É importante, no entanto, não confundir as duas.

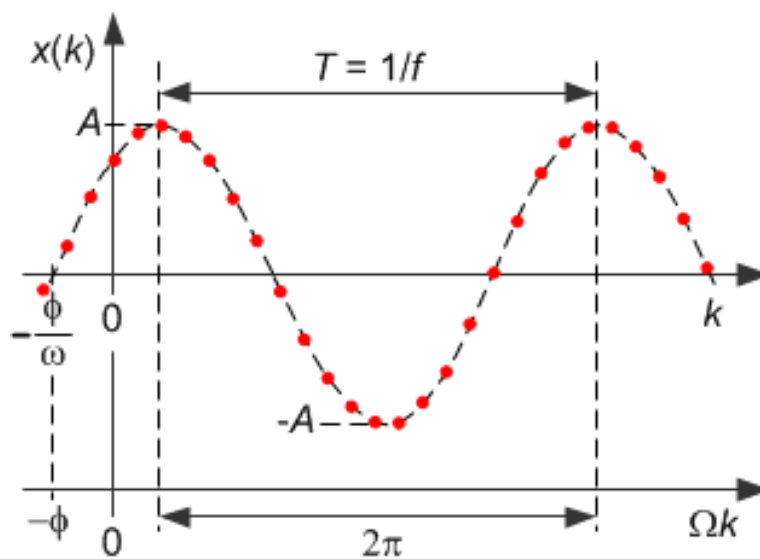


Figura 13.1 – Sinal sinusoidal de frequência angular Ω .

Consultando a tabela das transformações em z obtém-se para a transformação em z do sinal $x(k)$

$$X(z) = A \frac{(z + \sin(\Omega - \phi))z}{z^2 - 2z \cos \Omega + 1}$$

13.2. Resposta no tempo discreto a um sinal sinusoidal

Considere-se agora um sistema $G(z)$ excitado com um sinal sinusoidal $u(k)$ de frequência Ω e amplitude U_0 . Se os n polos p_i de $G(z)$ forem todos estáveis e diferentes entre si, a saída

$$Y(z) = G(z)U(z)$$

admite a expansão em frações simples

$$Y(z) = \frac{r_1 z}{z - p_1} + \frac{r_2 z}{z - p_2} + \dots + \frac{r_n z}{z - p_n} + \frac{r_0}{z - e^{j\Omega}} + \frac{r_0^*}{z - e^{-j\Omega}}$$

onde r_i são os resíduos, sendo r_0^* é o conjugado de r_0 . Se aplicarmos a transformação em z inversa a esta expressão obtém-se

$$y(k) = r_1 p_1^k + r_2 p_2^k + \dots + r_n p_n^k + 2|r_0| \sin(\Omega k + \phi)$$

Na Figura 13.2, representando a resposta $y(k)$, podem identificar-se duas componentes sobrepostas. A primeira, constituída pelos n termos exponenciais, é transitória extinguindo-se à velocidade do polo mais lento. Passado este transitório permanece o termo – designado por regime permanente - dado por

$$y_p(k) = 2|r_0| \sin(\Omega k + \phi)$$

Resta saber quanto valerão as quantidades $|r_0|$ e ϕ .

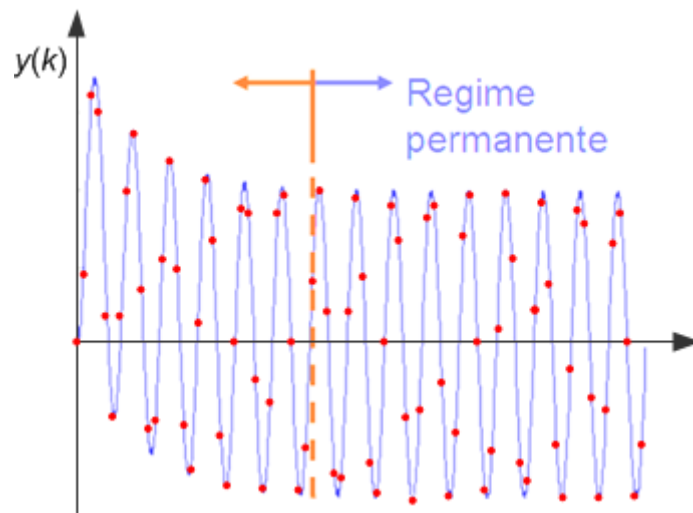


Figura 13.2 – Resposta a excitação sinusoidal em regime permanente.

O cálculo dos resíduos r_0 e r_0^* , complexos conjugados, é feito a partir das expressões

$$r_0 = \lim_{z \rightarrow e^{j\Omega}} (z - e^{j\Omega})Y(z) = \frac{U_0 e^{j\Omega}}{2j} \lim_{z \rightarrow e^{j\Omega}} G(z) = + \frac{U_0 e^{j\Omega}}{2j} G(e^{j\Omega})$$

$$r_0^* = \lim_{z \rightarrow e^{-j\Omega}} (z - e^{-j\Omega})Y(z) = + \frac{U_0 e^{-j\Omega}}{2j} \lim_{z \rightarrow e^{-j\Omega}} G(z) = - \frac{U_0 e^{-j\Omega}}{2j} G(e^{-j\Omega})$$

que substituído na expressão de $Y_p(z)$ resulta

$$Y_p(z) = \frac{r_0}{z - e^{j\Omega}} + \frac{r_0^*}{z - e^{-j\Omega}} = \frac{U_0}{2j} \cdot \left\{ \frac{G(e^{j\Omega}) \cdot e^{j\Omega}}{z - e^{j\Omega}} - \frac{G(e^{-j\Omega}) \cdot e^{-j\Omega}}{z - e^{-j\Omega}} \right\}$$

com transformação em z inversa dada por (demonstre como exercício)

$$y_p(k) = U_0 \cdot |G(e^{j\Omega})| \cdot \sin(\Omega k + \arg\{G(e^{j\Omega})\})$$

Assim conclui-se que, se um sistema discreto linear invariante no tempo, estável, descrito pela sua função de transferência $G(z)$ for excitado por um sinal sinusoidal

$$u(k) = U_0 \sin(\Omega k) \varepsilon(k)$$

responde, após a extinção de um transitório, com um regime permanente

$$y_p(k) = U_0 \cdot M_G(\Omega) \sin(\Omega k + \phi_G(\Omega)) \varepsilon(k)$$

também sinusoidal de igual frequência Ω mas com um ganho de amplitude dado por

$$M_G(\Omega) = |G(e^{j\Omega})|$$

e desvio de fase dado por

$$\phi_G(\Omega) = \arg\{G(e^{j\Omega})\}$$

A quantidade complexa $G(e^{j\Omega})$ é designada por **resposta em frequência**.

Como no caso contínuo, sendo uma função complexa, é possível representar $G(e^{j\Omega})$ graficamente de diversas formas, tais como:

- Representação direta de $G(e^{j\Omega})$ num plano complexo para todos os valores possíveis da frequência Ω . Esta representação designa-se por diagrama polar e é a base para o traçado do **diagrama de Nyquist**.
- Representação num diagrama cartesiano de $M_G(\Omega)$ e $\phi_G(\Omega)$ para todos os valores possíveis da frequência Ω . Esta representação designa-se por diagrama de Nichols e permite-nos tirar conclusões semelhantes àquelas que tiraremos com o diagrama de Nyquist.
- Representação em separado das quantidades $M_G(\Omega)$ e $\phi_G(\Omega)$ em função da frequência Ω . Esta representação é designada por **diagrama de Bode**. Porque a variável $\Omega = \omega \cdot T_s$ aparece explicitamente, estes diagramas contêm mais informação que as duas representações referidas antes.

O exemplo seguinte introduz estas representações diagramáticas possíveis, entre outras.

Exemplo 13.1

Considere-se a instalação de 2ª ordem com a função de transferência

$$F(z) = \frac{0.2z}{(z - 0.8)^2}$$

A resposta em frequência da instalação é obtida substituindo a variável s por $j\omega$, ou seja

$$F(e^{j\Omega}) = \frac{0.2 \cdot e^{j\Omega}}{(e^{j\Omega} - 0.8)^2} = \frac{0.2 \cdot e^{j\Omega}}{((\cos \Omega - 0.8) + j \sin \Omega)^2}$$

com ganho

$$M_F(\Omega) = |F(e^{j\Omega})| = \frac{0.2}{\cos^2 \Omega + 0.64 - 1.6 \cos \Omega - \sin^2 \Omega}$$

e fase

$$\phi_F(\Omega) = \arg\{F(e^{j\Omega})\} = \Omega - 2 \operatorname{atan}\left(\frac{\sin \Omega}{\cos \Omega - 0.8}\right)$$

O estudo destas duas funções da frequência angular Ω pode fazer-se pela substituição de certos valores notáveis. Iniciemos esta análise, pelo limite na baixa frequência ($\Omega \rightarrow 0$), observando que

$$M_F(e^{j0}) = 5$$

$$\phi_F(e^{j0}) = 0^\circ$$

e na alta frequência ($\Omega \rightarrow \pi$)

$$M_F(e^{j\pi}) = 0.0617$$

$$\phi_F(e^{j\pi}) = -180^\circ$$

Um ponto notável intermédio ocorre com $\Omega = 0.22$ rad resultando para a fase o valor de

$$\phi_F(e^{j0.22}) = -90^\circ$$

e para o ganho

$$M_F(e^{j0.22}) = 2.52$$

Na Figura 13.3 observa-se o registo da resposta em frequência no plano complexo.

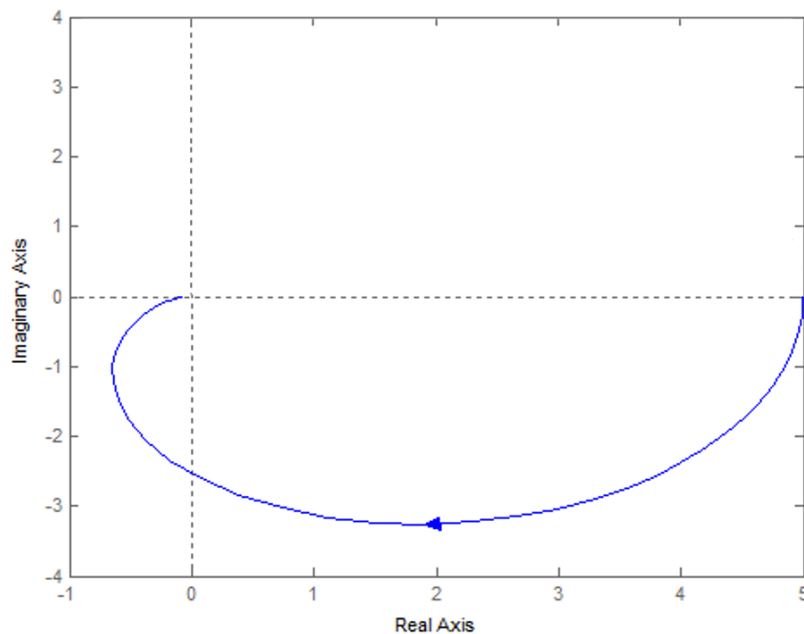


Figura 13.3 – Diagrama polar da resposta em frequência do Exemplo 13.1.

■

13.3. Critério de Nyquist

Verifiquemos agora como se transporta o critério de Nyquist para os sistemas em tempo discreto. Considere-se o sistema de controlo em anel fechado da Figura 13.4.

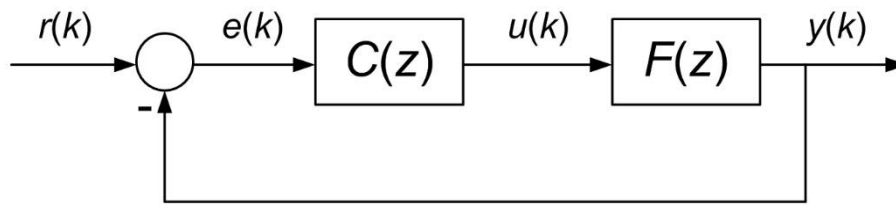


Figura 13.4 – Esquema de controlo de seguimento por retroação.

Se o anel aberto for dado por

$$G_0(z) = F(z)C(z) = \frac{N_0(z)}{D_0(z)}$$

a função de transferência do anel fechado será dada por

$$H(z) = \frac{G_0(z)}{1 + G_0(z)} = \frac{\frac{N_0(z)}{D_0(z)}}{1 + \frac{N_0(z)}{D_0(z)}} = \frac{N_0(z)}{D_0(z) + N_0(z)}$$

Assim, os polos do anel aberto são dados pelas raízes de $D_0(z)$ e os polos do anel fechado são dados pelas raízes de $D_0(z) + N_0(z)$.

Considere-se agora a função

$$1 + G_0(z) = 1 + \frac{N_0(z)}{D_0(z)} = \frac{D_0(z) + N_0(z)}{D_0(z)}$$

e o contorno de Nyquist que envolve todo o exterior do círculo unitário, ou seja a região instável do plano z , no sentido dos ponteiros do relógio.

A imagem do contorno de Nyquist através da função

$$1 + G_0(z) = \frac{D_0(z) + N_0(z)}{D_0(z)}$$

terá um número de envolvimentos em redor da origem igual à diferença entre os polos instáveis (i.e. no interior do contorno de Nyquist) do anel fechado e os polos instáveis do anel aberto (ver Figura 13.5).

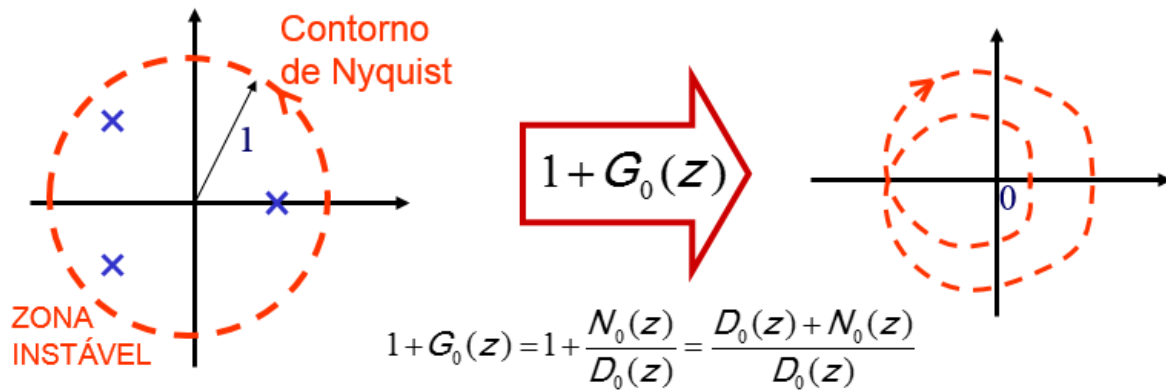


Figura 13.5 – Teorema dos Envolvimentos aplicado ao Contorno de Nyquist.

Sendo mais fácil analisar a imagem da função $G_0(z)$ (o anel aberto), a contagem de envoltimentos, em redor da origem, da função $1 + G_0(z)$ é equivalente à contagem de envoltimentos da função $G_0(z)$ em redor do ponto -1 , como se ilustra na Figura 13.6.

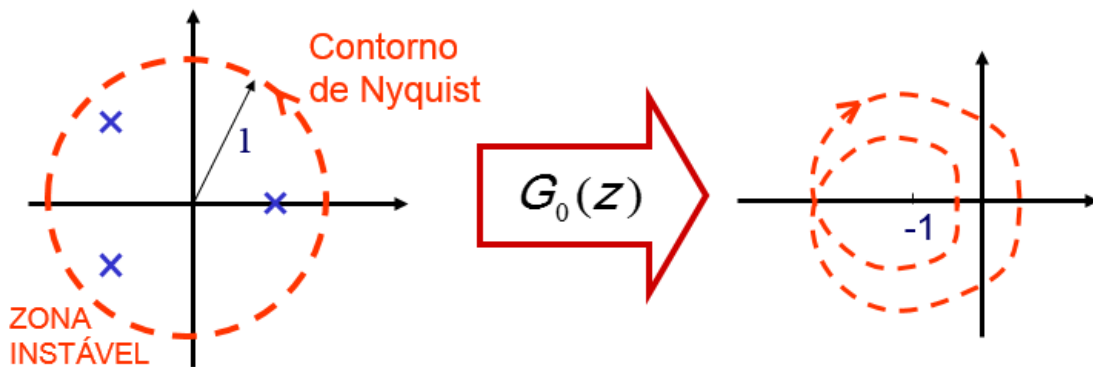


Figura 13.6 – Diagrama de Nyquist como a imagem do contorno de Nyquist através de $G_0(z)$.

À semelhança do caso contínuo, o critério de Nyquist enuncia-se da seguinte forma. A imagem do contorno de Nyquist através da função $G_0(z)$ terá um número de envoltimentos em redor do ponto (-1) igual à diferença entre os polos instáveis do anel fechado e os polos instáveis do anel aberto, que equivale a escrever

$$N = Z - P$$

onde

- P – é o número de polos instáveis do anel aberto;
- Z – é o número de polos instáveis do anel fechado e
- N – é o número de envoltimentos em redor do ponto -1 a origem no mesmo sentido do contorno de Nyquist.

O diagrama que se constrói através da imagem do contorno por $G_0(z)$ designa-se por diagrama de Nyquist de $G_0(z)$.

No caso mais geral, em que o anel aberto do sistema $G_0(z)$ é estável (i.e. $P = 0$), o critério de Nyquist resume-se a

$$Z = N$$

ou seja, para que o sistema em anel fechado resulte estável ($Z = 0$), o diagrama de Nyquist do anel aberto $G_0(z)$ não pode envolver o ponto -1 (i.e. $N = 0$). Se envolver o ponto -1 , o número N de voltas corresponde ao número de polos instáveis do sistema em anel fechado. Resulta disto, que, tal como no caso contínuo, para um anel aberto estável e no objetivo de ter um sistema em anel fechado estável, a circulação do ponto -1 deve ser evitada.

Considere-se agora o caso, menos frequente, em que o anel aberto do sistema $G_0(z)$ é instável (i.e. $P > 0$). Para que, novamente, o sistema em anel fechado resulte estável ($Z = 0$), o diagrama de Nyquist do anel aberto $G_0(z)$ tem de envolver o ponto -1 , em sentido contrário ao do contorno de Nyquist, num número de voltas igual a

$$N = -P$$

(negativo porque no sentido contrário). Assim, para um anel aberto instável e no objetivo de ter um sistema em anel fechado, a circulação do ponto -1 tem de ocorrer no sentido contrário ao dos ponteiros do relógio P vezes.

Exemplo 13.2

Considere-se novamente o sistema em anel aberto dado por

$$G_0(z) = \frac{0.2z}{(z - 0.8)^2}$$

que é estável, pois o seu polo de multiplicidade 2 vale 0.8, e portanto $P = 0$. Se traçarmos o diagrama de Nyquist deste sistema, ou seja, o diagrama polar de

$$G(e^{j\Omega}) = \frac{0.2e^{j\Omega}}{(e^{j\Omega} - 0.8)^2}$$

para valores de $\Omega \in [-\pi, +\pi]$, i.e. o contorno de Nyquist, obtém-se o resultado da Figura 13.7.

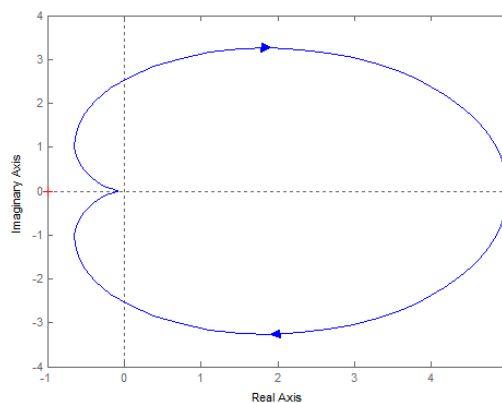


Figura 13.7 – Diagrama de Nyquist de $G_0(z)$ do Exemplo 13.2.

Este diagrama de Nyquist não realiza qualquer envolvimento em redor do ponto -1 , ou seja, $N = 0$. Assim,

$$Z = N + P = 0$$

o anel fechado não terá qualquer polo instável. De modo a verificar este resultado, podemos calcular diretamente o anel fechado

$$H(z) = \frac{G_0(z)}{1 + G_0(z)} = \frac{\frac{0.2z}{(z - 0.8)^2}}{1 + \frac{0.2z}{(z - 0.8)^2}} = \frac{0.2z}{z^2 - 1.4z + 0.64}$$

que, com dois polos em $0.70 \pm j0.39$ no interior do círculo unitário, é estável. ■

13.4. Diagrama de Nyquist e a resposta em frequência

Podemos, nesta altura, perguntar qual a relação entre o diagrama de Nyquist e a resposta em frequência, também introduzida neste capítulo. Repare-se que o contorno fechado de Nyquist é constituído pelas seguintes partes:

- o semicírculo de raio unitário centrado na origem entre $z = e^{j0} = 1$ e $z = e^{j\pi} = -1$;
- e o semicírculo de raio unitário centrado na origem entre $z = e^{-j\pi} = -1$ e $z = e^{j0} = 1$.

Note-se que uma circulação no sentido contrário ao dos ponteiros do relógio do círculo unitário é equivalente a um envolvimento da região complementar do círculo unitário no sentido contrário. Portanto, o sentido do contorno de Nyquist é o dos ponteiros do relógio.

O diagrama de Nyquist é a imagem destes dois troços através de $G_0(z)$. Mas, a imagem do primeiro troço é o diagrama polar da resposta em frequência $G(e^{j\Omega})$. A imagem do segundo troço é a resposta em frequência para frequências negativas $G_0(e^{-j\Omega})$, que resultará no diagrama polar da resposta em frequência rebatido em torno do eixo real.

13.5. Diagrama de Bode

O diagrama de Bode representa, explicitamente em dois diagramas paralelos, as funções de ganho ($M_G(\Omega) = |G(e^{j\Omega})|$) e fase ($\phi_G(\Omega) = \arg\{G(e^{j\Omega})\}$) em função da frequência angular Ω . Este diagrama pode ser representado para o anel aberto do sistema, estando intimamente ligado ao diagrama de Nyquist, mas também para o anel fechado, fornecendo informações de carácter diferente.

A novidade face ao caso contínuo é que o diagrama de Bode fica confinado à gama de frequências entre $\Omega = 0$ e $\Omega = \pi$, melhor dizendo, entre as frequência $\omega = 0$ e $\omega = \omega_N$.

13.6. Construção aproximada do diagrama de Bode – método das assíntotas

O traçado do diagrama de Bode através do método das assíntotas, tal como foi apresentado para o caso contínuo, deixou de ser possível com a gama de frequências limitada e com a dependência na variável Ω através da função exponencial $e^{j\Omega}$.

No entanto, podemos aplicar uma transformação na variável Ω que nos permita recuperar todo o potencial dessa metodologia.

Viu-se anteriormente que a transformação bilinear (usada na conversão contínuo-discreto de Tustin) mapeava o semi-plano complexo definido em s no círculo unitário em z (ver Figura 13.8) através da fórmula

$$z = \frac{1 + \frac{sT_s}{2}}{1 - \frac{sT_s}{2}}$$

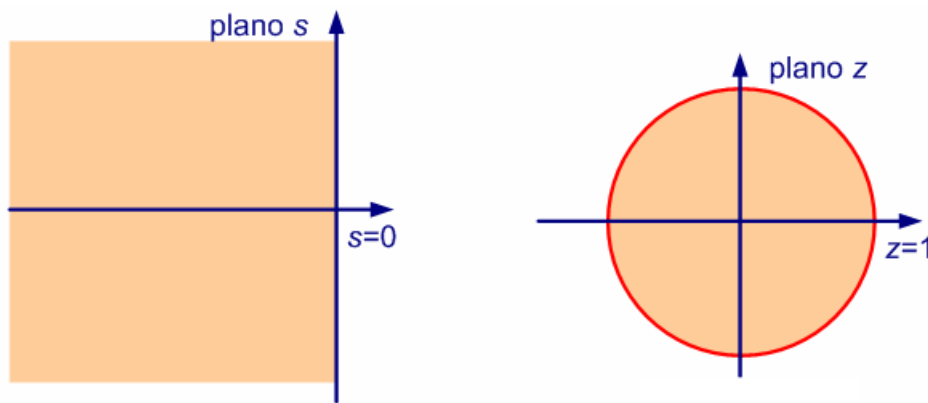


Figura 13.8 – Mapeamento realizado pela transformação bilinear.

A metodologia a introduzir tira partido deste resultado para realizar o traçado das respostas em frequência recuperando o traçado do diagrama de Bode assímptótico do caso contínuo.

Vamos criar um espaço complexo “contínuo”, designado por domínio da variável de Laplace w com características semelhantes ao espaço em da variável de Laplace s , mas diferente. De seguida, transportamos os sistemas a analisar em tempo discreto para este novo espaço através de um processo de conversão discreto-contínuo usando o método da bilinear (Figura 13.9). Como esta transformação é conforme, podemos ir e voltar retornando ao ponto inicial.

Assim, as transformações de ida e de volta são dadas por

$$w = \frac{2}{T_s} \cdot \frac{z - 1}{z + 1} \quad \longleftrightarrow \quad z = \frac{1 + \frac{wT_s}{2}}{1 - \frac{wT_s}{2}}$$

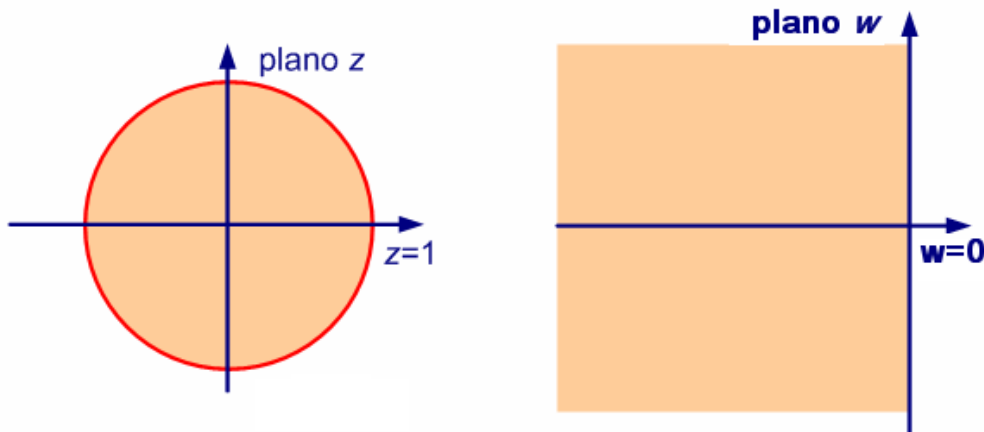


Figura 13.9 – Mapeamento realizado pela transformação bilinear para o domínio da variável w .

Podemos agora realizar o traçado do diagrama de Bode da função de transferência $G(w)$ imagem da função $G(z)$, usando tudo o que sabíamos do método das assíntotas.

Fazendo um paralelo, o traçado do diagrama de Bode realizava-se obtendo a resposta em frequência por substituição da variável de Laplace s por $j\omega$. Agora, o traçado do diagrama de Bode realiza-se obtendo a resposta em frequência por substituição da variável de Laplace w por jv , definindo uma nova frequência v (representada pela letra grega ν com correspondente latino na letra N) com unidades de rad/s.

Se o parâmetro T_s , usado nas fórmulas de conversão bilineares, coincidir com o tempo de amostragem usado originalmente para discretizar a instalação $F(s)$, podemos estabelecer uma relação entre as frequências ω e v .

No processo de discretização da instalação através do método ZOH observámos a relação

$$\Omega = \omega T_s$$

Agora, no segundo salto do discreto para o novo contínuo em w temos

$$e^{j\Omega} = \frac{1 + j\frac{\nu T_s}{2}}{1 - j\frac{\nu T_s}{2}} = \frac{\left(1 + j\frac{\nu T_s}{2}\right)^2}{1 + \left(\frac{\nu T_s}{2}\right)^2} = e^{j2 \operatorname{atan}\frac{\nu T_s}{2}}$$

ou seja, estabelece-se uma relação entre as frequências Ω e v dada por

$$\Omega = 2 \operatorname{atan}\frac{\nu T_s}{2}$$

Finalmente, eliminando a frequência Ω nas duas expressões obtêm-se as relações

$$\omega = \frac{2}{T_s} \cdot \operatorname{atan}\frac{\nu T_s}{2} \quad \longleftrightarrow \quad \nu = \frac{2}{T_s} \cdot \tan\frac{\omega T_s}{2}$$

A Figura 13.10 apresenta a curva definida por estas relações.

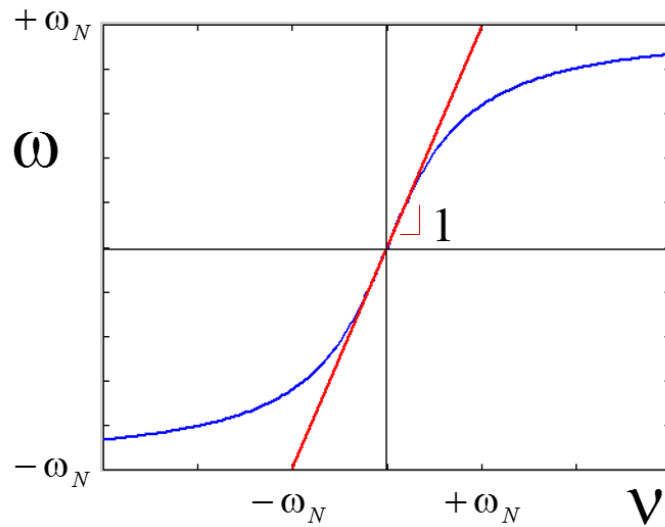


Figura 13.10 – Relação entre as frequências ω e v .

Como a função tangente apresenta uma inclinação unitária em redor da origem, as frequências ω e v aproximam-se nas baixas frequências. No sentido oposto, quando a frequência ω se aproxima do limite imposto pela frequência de Nyquist ω_N , a frequência v é “esticada” até infinito. É esta característica da transformação que nos abre espaço para podermos aplicar tudo o que aprendemos em tempo contínuo usando o método das assíntotas.

Note-se ainda que a curva da Figura 13.10 apresenta-se em escalas logarítmicas como desenhado na Figura 13.11

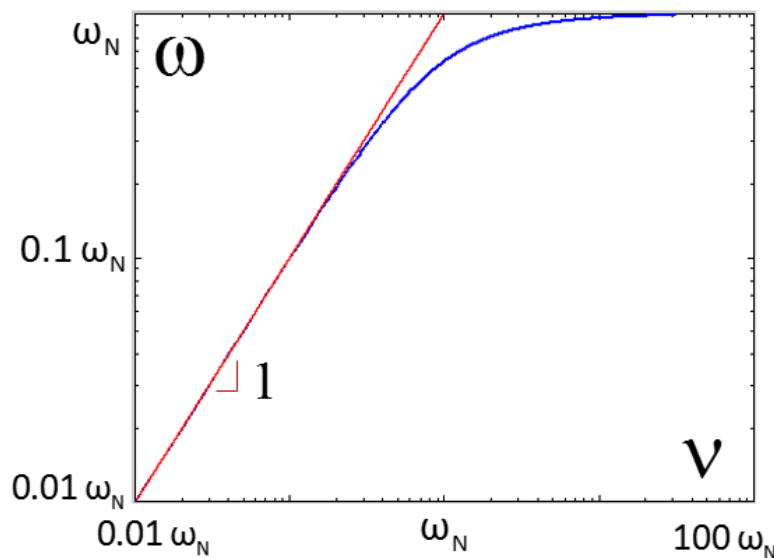


Figura 13.11 – Relação entre as frequências ω e v , em escalas logarítmicas.

Observa-se aqui que só para frequências próximas de $0.2\omega_N$ é que se começa a observar diferença. No entanto, o valor de $0.2\omega_N$ é também aquele que temos considerado como limite da frequência máxima de trabalho ω_{max} . Ou seja, na zona normal de trabalho, as duas frequências têm um andamento coincidente.

13.7. Projeto do controlador através da resposta em frequência

A metodologia de projeto do controlador usando a resposta em frequência é a seguinte:

1. Obter a função $F(z)$ através da conversão contínuo-discreto aplicada à função de transferência $F(s)$ da instalação contínua.
2. Aplicar a transformação bilinear à função de transferência $F(z)$ para obter a função de transferência da instalação no domínio da nova variável de Laplace w , $F(w)$.
3. Realizar o projeto do novo controlador $C(w)$ como se realizava em tempo contínuo.
4. Aplicar a transformação bilinear inversa à função de transferência $C(w)$ para obter o controlador discreto $C(z)$.

14. Sistemas discretos em espaço de estados

O método do espaço de estados baseia-se na descrição da instalação a controlar através de um conjunto de n equações lineares de 1ª ordem, na representação de um sistema dinâmico linear de ordem n , introduzindo o conceito de **estado do sistema**.

O estado de um sistema dinâmico é o conjunto mínimo de variáveis, ditas de estado, x_1, x_2, \dots, x_n , tal que o seu conhecimento num dado instante de tempo discreto k_0 é memória suficiente para determinar a evolução futura do sistema para $k > k_0$.

O exemplo seguinte serve para motivar a introdução desta representação.

Exemplo 14.1

Considere-se o sistema dado pela cascata de funções de transferência da Figura 14.1.

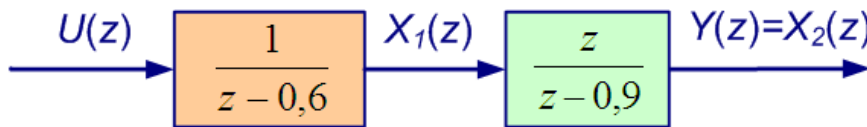


Figura 14.1 – Esquema de blocos do sistema do **Exemplo 14.1**.

O sistema é assim dado por

$$Y(z) = \frac{z}{z - 0.9} X_1(z) = \frac{z}{z - 0.9} \cdot \frac{1}{z - 0.6} U(z)$$

definindo-se um variável interna x_1 , entre os dois blocos, e uma variável também interna x_2 que, neste exemplo, coincide com a saída y .

A representação em espaço de estados é definida no domínio do tempo através da descrição de como o instante futuro de cada variável de estado depende do valor atual de todo o estado e das entradas.

Seguindo na mesma figura da entrada u para a saída y , o primeiro bloco pode ser convertido na sua equação às diferenças correspondente dada por

$$x_1(k + 1) = 0.6 \cdot x_1(k) + u(k)$$

O segundo bloco pode, por seu lado, ser descrito pela equação às diferenças

$$x_2(k + 1) = 0.9 \cdot x_2(k) + x_1(k + 1)$$

ou, por substituição da equação anterior nesta última,

$$x_2(k + 1) = 0.6 \cdot x_1(k) + 0.9 \cdot x_2(k) + u(k)$$

É assim possível escrever um sistema de equações onde ficam definidas a forma recorrente como as duas variáveis de estado x_1 e x_2 evoluem no tempo k , como

$$\begin{cases} x_1(k+1) = 0.6 \cdot x_1(k) + u(k) \\ x_2(k+1) = 0.6 \cdot x_1(k) + 0.9 \cdot x_2(k) + u(k) \end{cases}$$

Repetindo, o sistema descreve de forma recorrente como o valor do estado no instante seguinte $k + 1$ depende do valor do estado no instante atual k , bem como da entrada do sistema no instante atual k .

Para terminar a descrição do sistema devemos adicionar de que forma a saída do sistema depende do estado ou da entrada, que para este exemplo se resume a

$$y(k) = x_2(k)$$

■

As equações apresentadas no exemplo anterior realizam uma boa descrição do sistema discreto mas não são facilmente tratáveis se quisermos desenvolver um conjunto de ferramentas de projeto em seu redor, como fizemos com a função de transferência. A análise da estabilidade de um sistema descrito por uma função de transferência $F(z)$ realiza-se obtendo as raízes do polinómio denominador e verificando se estão no interior da região de estabilidade. Devemos procurar algo que estructure da mesma forma a descrição do sistema e nos possibilite obter resultados equivalentes ou melhores. Isso é possível introduzindo a notação matricial na descrição em espaço de estados.

14.1. Vetor de estado e modelo de estado

As variáveis de estado aparecem normalmente agregadas na forma matricial, definindo o **vetor de estado** definido como

$$\mathbf{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{bmatrix}$$

ou, numa forma transposta mais compacta,

$$\mathbf{x}(k) = [x_1(k) \quad x_2(k) \quad \cdots \quad x_n(k)]^T$$

Usando esta notação, a realização em espaço de estados é descrita como

$$\mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k)$$

$$y(k) = \mathbf{C} \cdot \mathbf{x}(k) + D \cdot u(k)$$

no qual as variáveis vetoriais ou matriciais são representadas em fonte **negrita-ereta**, enquanto as variáveis escalares são, como habitualmente, representadas em fonte *itálica*.

A primeira equação, designada por **equação da dinâmica**, diz-nos de uma forma agregada que o estado \mathbf{x} no instante seguinte $k + 1$ depende do estado \mathbf{x} no instante atual k através da matriz \mathbf{A} . Esta matriz \mathbf{A} é forçosamente quadrada e de dimensão $(n \times n)$. Adicionalmente, o estado seguinte depende da entrada u no instante atual k através da matriz \mathbf{B} de dimensão $(n \times 1)$.

A segunda equação, designada por **equação de saída**, diz-nos que a saída y , em cada instante k , depende de uma forma direta do valor do estado \mathbf{x} através da matriz \mathbf{C} de dimensão $(1 \times n)$; e, eventualmente, da entrada u através da “matriz” \mathbf{D} , também no instante k . Note-se que, ao contrário da primeira equação, a segunda é absolutamente desprovida de dinâmica, i.e. sem memória.

Esta notação está já preparada para sistemas que tenham múltiplas entradas e múltiplas saídas, sendo uma das mais-valias do método.

Assim, em geral para um sistema com n_i sinais de entrada e n_o sinais de saída, \mathbf{u} e \mathbf{y} passam a vetores e a realização em espaço de estado toma a forma

$$\mathbf{x}(k + 1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot \mathbf{u}(k)$$

$$\mathbf{y}(k) = \mathbf{C} \cdot \mathbf{x}(k) + \mathbf{D} \cdot \mathbf{u}(k)$$

no qual

- \mathbf{A} é a matriz da dinâmica $(n \times n)$;
- \mathbf{B} é a matriz de entrada $(n \times n_i)$;
- \mathbf{C} é a matriz da saída $(n_o \times n)$;
- \mathbf{D} é a matriz de transmissão $(n_o \times n_i)$;

No entanto, neste texto introdutório trataremos apenas o caso mais simples do sistema com uma entrada e uma saída, i.e. $n_i = n_o = 1$. A Figura 14.2 apresenta uma representação esquemática do modelo de estado, também designado por **esquema de simulação**, que mostra de uma forma direta como fluem os sinais no modelo. O bloco central é constituído simplesmente por um banco de atrasos de uma amostra representados pelo termo z^{-1} multiplicado pela matriz identidade \mathbf{I} .

O exemplo seguinte mostra como podemos estruturar as equações diferenciais encontradas no exemplo anterior no formato matricial do modelo de estado.

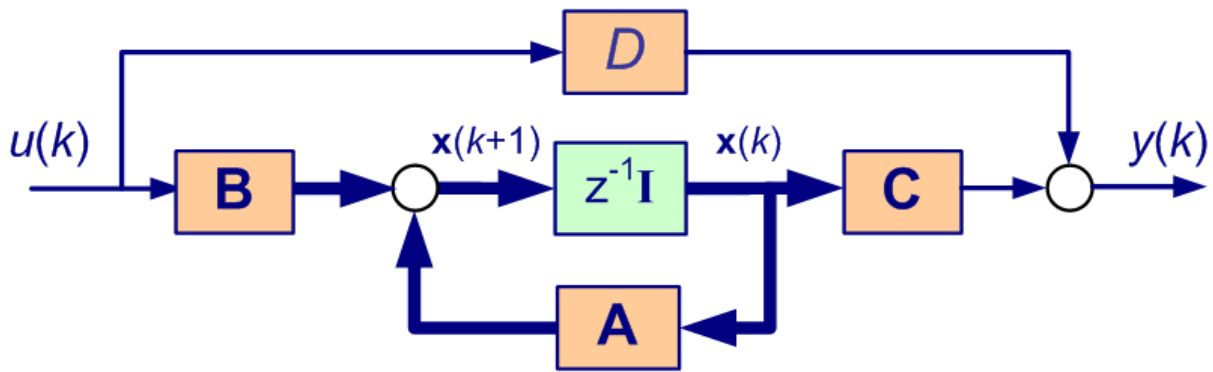


Figura 14.2 – Esquema de simulação do modelo de estado.

Exemplo 14.2

Se continuarmos o **Exemplo 14.1** a partir das equações obtidas

$$\begin{cases} x_1(k + 1) = 0.6 \cdot x_1(k) + u(k) \\ x_2(k + 1) = 0.6 \cdot x_1(k) + 0.9 \cdot x_2(k) + u(k) \end{cases}$$

e

$$y(k) = x_2(k)$$

podemos perguntar agora: qual é o modelo de estado correspondente?

Arrumando os elementos nas diversas entradas das matrizes com dimensão adequada, obtém-se

$$\begin{bmatrix} x_1(k + 1) \\ x_2(k + 1) \end{bmatrix} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot u(k)$$

$$y(k) = [0 \quad 1] \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + [0] \cdot u(k)$$

ficando definidas as matrizes

$$\mathbf{A} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{C} = [0 \quad 1] \quad \mathbf{D} = [0]$$

e o esquema de simulação da Figura 14.3.

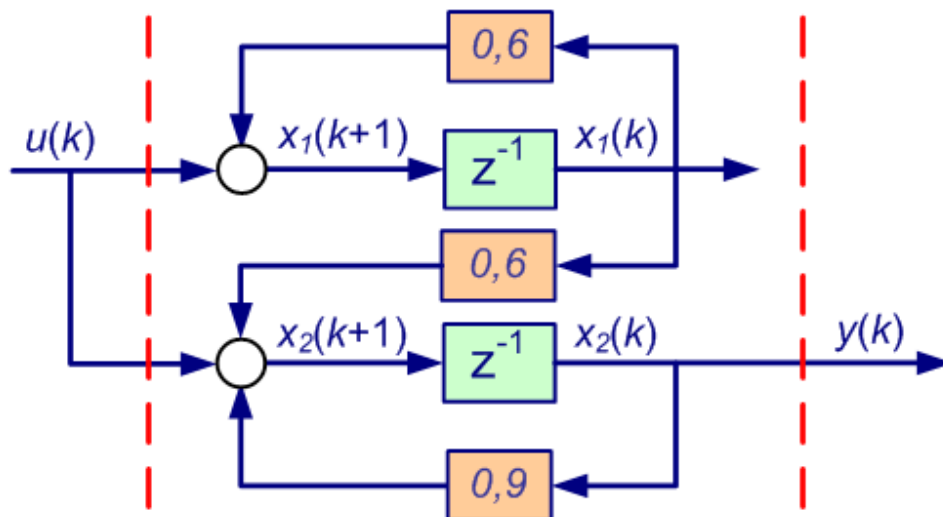


Figura 14.3 – Esquema de simulação do modelo de estado do Exemplo 14.2.

14.2. Matriz de transição de estado

Pensemos agora num sistema representado em modelo de estado mas sem excitação à entrada, i.e. $u(k) \equiv 0$. A equação da dinâmica resume-se assim a

$$\mathbf{x}(k + 1) = \mathbf{A} \cdot \mathbf{x}(k)$$

que pode ser iterada a partir de uma determinada condição inicial $\mathbf{x}(0) = \mathbf{x}_0$.

Deste modo, para cada instante do tempo teremos

$$\mathbf{x}(1) = \mathbf{A} \cdot \mathbf{x}(0) = \mathbf{A} \cdot \mathbf{x}_0$$

$$\mathbf{x}(2) = \mathbf{A} \cdot \mathbf{x}(1) = \mathbf{A} \cdot \mathbf{A} \cdot \mathbf{x}_0$$

$$\mathbf{x}(3) = \mathbf{A} \cdot \mathbf{x}(2) = \mathbf{A}^3 \cdot \mathbf{x}_0$$

⋮

$$\mathbf{x}(k) = \mathbf{A} \cdot \mathbf{x}(k - 1) = \mathbf{A}^k \cdot \mathbf{x}_0$$

ou seja, a solução do sistema autónomo é dada por

$$\mathbf{x}(k) = \mathbf{\Phi}(k) \cdot \mathbf{x}(0)$$

onde $\mathbf{\Phi}(k) = \mathbf{A}^k$ é a matriz de transição entre o estado inicial $\mathbf{x}(0)$ e o estado em qualquer instante futuro k . Note-se que sendo \mathbf{A} uma matriz, o cálculo de \mathbf{A}^k não é trivial.

14.3. Solução geral da equação de estado

A partir da definição de matriz de transição estamos em condições para determinar a solução das equações de estado para o caso mais geral com a entrada u não nula. Note-se que em cada instante do tempo discreto, a entrada u reposiciona o estado numa nova condição inicial da evolução futura.

A solução geral da equação de estado

$$\mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k)$$

é dada por

$$\mathbf{x}(k) = \Phi(k) \cdot \mathbf{x}(0) + \sum_{i=0}^{k-1} \Phi(k-i-1) \cdot \mathbf{B} \cdot u(i)$$

que resulta da sobreposição da resposta da equação homogénea com a resposta forçada pela entrada u . A resposta forçada é, por sua vez, resultante da sobreposição dos efeitos da entrada em cada instante do passado do sistema.

Fica como exercício a demonstração deste resultado.

14.4. Relação entre modelo de estado e função de transferência

Se aplicarmos a transformação em z às equações de estado é possível convertê-lo para o formalismo da função de transferência. Considere-se então o modelo de estado

$$\mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k)$$

$$\mathbf{y}(k) = \mathbf{C} \cdot \mathbf{x}(k) + D \cdot u(k)$$

e apliquemos a transformação em z à primeira equação (com condição inicial nula). Obtém-se assim

$$z \cdot \mathbf{X}(z) = \mathbf{A} \cdot \mathbf{X}(z) + \mathbf{B} \cdot U(z)$$

ou

$$(z \cdot \mathbf{I} - \mathbf{A}) \cdot \mathbf{X}(z) = \mathbf{B} \cdot U(z)$$

ou ainda

$$\mathbf{X}(z) = (z \cdot \mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{B} \cdot U(z)$$

Com a segunda equação do modelo de estado escrevemos

$$Y(z) = \mathbf{C} \cdot \mathbf{X}(z) + D \cdot U(z)$$

ou, a partir do resultado obtido antes para $\mathbf{X}(z)$,

$$Y(z) = \mathbf{C} \cdot (z \cdot \mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{B} \cdot U(z) + D \cdot U(z)$$

Concluindo, é possível obter a função de transferência a partir da realização de estado através da fórmula compacta

$$F(z) = \mathbf{C} \cdot (z \cdot \mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{B} + D$$

Note-se que uma forma de calcular a inversa de uma matriz é usar a relação

$$\mathbf{M}^{-1} = \frac{\text{adj } \mathbf{M}}{\det \mathbf{M}}$$

onde $\text{adj } \mathbf{M}$ é a matriz adjunta de \mathbf{M} ; e $\det \mathbf{M} = |\mathbf{M}|$ é o seu determinante.

Por exemplo, para uma matriz genérica de dimensão 2×2 a aplicação desta fórmula resulta em

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}{a \cdot d - b \cdot c}$$

que, de certa forma, é fácil de memorizar (os elementos na diagonal principal trocam de lugar e os restantes trocam de sinal, dividindo pelo determinante).

Aplicando este método à fórmula do cálculo da função de transferência resulta que

$$F(z) = \mathbf{C} \cdot (\mathbf{z} \cdot \mathbf{I} - \mathbf{A})^{-1} \cdot \mathbf{B} + D = \frac{\mathbf{C} \cdot \text{adj} (\mathbf{z} \cdot \mathbf{I} - \mathbf{A}) \cdot \mathbf{B} + D \cdot |\mathbf{z} \cdot \mathbf{I} - \mathbf{A}|}{|\mathbf{z} \cdot \mathbf{I} - \mathbf{A}|}$$

concluindo-se que o denominador da função de transferência, cujas raízes são os polos, podem ser encontrados nas soluções de

$$|\mathbf{z} \cdot \mathbf{I} - \mathbf{A}| = 0$$

Voltaremos a esta equação quando analisarmos a estabilidade em espaço de estados.

14.5. Cálculo da matriz de transição

Disse-se atrás que o cálculo da matriz de transição $\Phi(k) = \mathbf{A}^k$ não era trivial. Os desenvolvimentos para a obtenção da função de transferência a partir do modelo de estado sugerem uma das formas para a calcular.

Se aplicarmos a transformação em z à equação homogénea

$$\mathbf{x}(k + 1) = \mathbf{A} \cdot \mathbf{x}(k)$$

obtém-se (agora com a condição inicial livre)

$$z \cdot \mathbf{X}(z) - z \cdot \mathbf{x}(0) = \mathbf{A} \cdot \mathbf{X}(z)$$

ou

$$\mathbf{X}(z) = (\mathbf{z} \cdot \mathbf{I} - \mathbf{A})^{-1} z \cdot \mathbf{x}(0)$$

e aplicando a transformação em z inversa resulta

$$\mathbf{x}(k) = \mathcal{TZ}^{-1}\{(\mathbf{z} \cdot \mathbf{I} - \mathbf{A})^{-1} \cdot z\} \cdot \mathbf{x}(0)$$

que pela própria definição de matriz de transição como a matriz que nos transfere do estado inicial para um estado futuro k , resulta que

$$\Phi(k) = \mathbf{A}^k = \mathbf{TZ}^{-1}\{(z \cdot \mathbf{I} - \mathbf{A})^{-1} \cdot z\}$$

14.6. Conversão da função de transferência em modelo de estado

Vimos como converter o modelo de estado em função de transferência. Note-se que, nessa passagem, a referência explícita ao estado desaparece. A função de transferência é um modelo entrada/saída, por vezes também referido como modelo caixa preta ou *black-box* (no sentido de ser opaca àquilo que se opera no seu interior). O modelo de estado, em oposição, é um modelo que descreve as operações internas do sistema fazendo emergir o conceito de estado. O modelo de estado diz-se assim um modelo caixa branca ou *white-box*.

A parte menos fácil de compreender é que enquanto a função de transferência de um determinado sistema é única (à parte de fatores comuns ao numerador e denominador que se cancelam); por outro lado, existem infinitas realizações possíveis em modelo de estado. Vejamos um exemplo de duas realizações diferentes da mesma função de transferência.

Exemplo 14.3

Considere-se de novo o sistema do **Exemplo 14.1** com função de transferência dada por

$$F(z) = \frac{z}{(z - 0.9)(z - 0.6)}$$

e para o qual encontrámos o modelo de estado na forma

$$\begin{cases} x_1(k + 1) = 0.6 \cdot x_1(k) + u(k) \\ x_2(k + 1) = 0.6 \cdot x_1(k) + 0.9 \cdot x_2(k) + u(k) \end{cases}$$

e

$$y(t) = x_2(k)$$

Consideremos a possibilidade de substituir as duas ($n = 2$) variáveis de estado x_1 e x_2 por outras assim definidas

$$w_1 = x_1 + x_2$$

$$w_2 = x_1 - x_2$$

Portanto,

$$\begin{cases} w_1(k + 1) = 1.2 \cdot x_1(k) + 0.9 \cdot x_2(k) + 2 \cdot u(k) \\ w_2(k + 1) = 0.0 \cdot x_1(k) - 0.9 \cdot x_2(k) + 0 \cdot u(k) \end{cases}$$

mas como

$$x_1 = \frac{w_1 + w_2}{2}$$

$$x_2 = \frac{w_1 - w_2}{2}$$

obtém-se

$$\begin{cases} w_1(k+1) = +1.05 \cdot w_1(k) + 0.15 \cdot w_2(k) + 2 \cdot u(k) \\ w_2(k+1) = -0.45 \cdot w_1(k) + 0.45 \cdot w_2(k) \end{cases}$$

com a equação de saída dada por

$$y(k) = 0.5 \cdot w_1(k) - 0.5 \cdot w_2(k)$$

Ou seja, temos duas realizações em espaço de estados que resultam na mesma função de transferência, são elas:

$$\begin{aligned} \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} &= \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot u(k) \\ y(k) &= [0 \quad 1] \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + [0] \cdot u(k) \end{aligned}$$

com o vetor de estado $\mathbf{x}(k) = [x_1(k) \quad x_2(k)]^T$; e

$$\begin{aligned} \begin{bmatrix} w_1(k+1) \\ w_2(k+1) \end{bmatrix} &= \begin{bmatrix} 1.05 & 0.15 \\ -0.45 & 0.45 \end{bmatrix} \cdot \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} \cdot u(k) \\ y(k) &= [0.5 \quad -0.5] \cdot \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} + [0] \cdot u(k) \end{aligned}$$

com o vetor de estado $\mathbf{w}(k) = [w_1(k) \quad w_2(k)]^T$.

O leitor é convidado a aplicar a fórmula da conversão de modelo de estado para função de transferência à segunda realização para verificar que se obtém a função de transferência inicial.

■

Quer isto dizer que enquanto converter modelos de estado em funções de transferência é um processo convergente: o resultado final é sempre o mesmo; converter funções de transferência em modelos de estado é um processo divergente pois existem infinitas soluções, todas equivalentes, para o mesmo ponto de partida. Dessas soluções, existem um número finito de realizações, ditas canónicas, que permitem a transposição direta da função de transferência para uma realização de estado.

Iremos apresentar aqui apenas uma delas, a forma canónica da controlabilidade, mas convida-se o leitor a procurar outras na literatura se assim for o seu interesse.

Considere-se então uma função de transferência genérica dada por

$$F(z) = \frac{b_0 z^n + b_1 z^{n-1} + \dots + b_{n-1} z + b_n}{z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n}$$

O modelo de estado equivalente, na forma canónica da controlabilidade é dado por

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_{n-1}(k+1) \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{bmatrix} \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u(k)$$

$$y(k) = [(b_n - a_n b_0) \quad (b_{n-1} - a_{n-1} b_0) \quad \dots \quad (b_1 - a_1 b_0)] \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_{n-1}(k) \\ x_n(k) \end{bmatrix} + [b_0] u(k)$$

Exemplo 14.4

Considere-se de novo o sistema do **Exemplo 14.1** com função de transferência dada por

$$F(z) = \frac{z}{(z - 0.9)(z - 0.6)} = \frac{z}{z^2 - 1.5z + 0.54}$$

A realização de estado obtida diretamente através da forma canónica da controlabilidade é dada por

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -0.54 & +1.5 \end{bmatrix} \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot u(k)$$

$$y(k) = [0 \quad 1] \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + [0] \cdot u(k)$$

convidando-se mais uma vez o leitor a verificar que mais este modelo de estado converte para a mesma função de transferência.

14.7. Valores próprios e estabilidade

A estabilidade de um sistema representado por um modelo de estado terá de ser obrigatoriamente analisada através das características da matriz da dinâmica **A**, dado o seu papel no modelo.

Foi já referido em outras partes deste texto que a estabilidade de um sistema linear só depende das suas características intrínsecas, independentemente do sinal de excitação à entrada. Se o sistema não for estável, basta uma breve injeção de energia para que este já não retorne ao seu ponto de equilíbrio. A equação da dinâmica permite observar que é a matriz **A** que caracteriza como a evolução do estado depende do seu valor atual. Então, em que condições é que o vetor do estado **x(k)** na equação recorrente

$$\mathbf{x}(k + 1) = \mathbf{A} \cdot \mathbf{x}(k)$$

converge ou diverge do seu ponto de equilíbrio, i.e. $\mathbf{x} = \mathbf{0}_n$?

Recordemos o conceito de valores e vetores próprios de uma matriz.

Definem-se os n valores próprios λ_i da matriz \mathbf{A} como as n raízes da equação

$$|\lambda \mathbf{I} - \mathbf{A}| = 0$$

e definem-se os vetores próprios \mathbf{v}_i da matriz \mathbf{A} como os n vetores não nulos, cada um associado a um valor próprio λ_i , tal que

$$\mathbf{A} \cdot \mathbf{v}_i = \lambda_i \cdot \mathbf{v}_i$$

Que poderemos concluir daqui?

Se os n vetores próprios forem todos independentes então constituem uma base do espaço em \mathbb{R}^n . Logo, qualquer vetor pode ser representado por uma combinação linear dos vetores próprios, inclusivamente, o vetor de estado inicial, i.e.

$$\mathbf{x}(0) = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$$

Iterando agora no sistema

$$\mathbf{x}(1) = \mathbf{A} \cdot \mathbf{x}(0) = \mathbf{A} \cdot \{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2\} = \alpha_1 \cdot \mathbf{A} \cdot \mathbf{v}_1 + \alpha_2 \cdot \mathbf{A} \cdot \mathbf{v}_2$$

ou ainda

$$\mathbf{x}(1) = \mathbf{A} \cdot \mathbf{x}(0) = \mathbf{A} \cdot \{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2\} = \alpha_1 \cdot \lambda_1 \cdot \mathbf{v}_1 + \alpha_2 \cdot \lambda_2 \cdot \mathbf{v}_2$$

As iterações futuras resultam na forma

$$\mathbf{x}(k) = \mathbf{A}^k \cdot \mathbf{x}(0) = \mathbf{A}^k \cdot \{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2\} = \alpha_1 \cdot \lambda_1^k \cdot \mathbf{v}_1 + \alpha_2 \cdot \lambda_2^k \cdot \mathbf{v}_2$$

que convergirão para zero se e só se todos os valores próprios forem, em valor absoluto, menores que 1.

Um sistema dinâmico representado pelo seu modelo em espaço de estados

$$\mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k)$$

é assintoticamente estável se todos os valores próprios da matriz da dinâmica estiverem no interior do círculo unitário.

Assim, um sistema dinâmico representado pelo seu modelo de estado é **estável no sentido assintótico** se e só se todos os seus valores próprios estão posicionados no interior do círculo de raio unitário, ou seja, se e só se

$$|\lambda_i| < 1 \quad \text{para } i = 1, \dots, n$$

Em oposição, um valor próprio diz-se **instável** quando se apresenta fora do círculo unitário.

É importante realçar neste ponto que os polos da função de transferência coincidem com os valores próprios da matriz da dinâmica. O leitor é convidado a comparar as equações apresentadas neste capítulo para a determinação dos polos e dos valores próprios.

A Figura 14.4 ilustra a região de estabilidade do sistema representado pelo modelo de estado, relativamente à posição dos seus valores próprios.

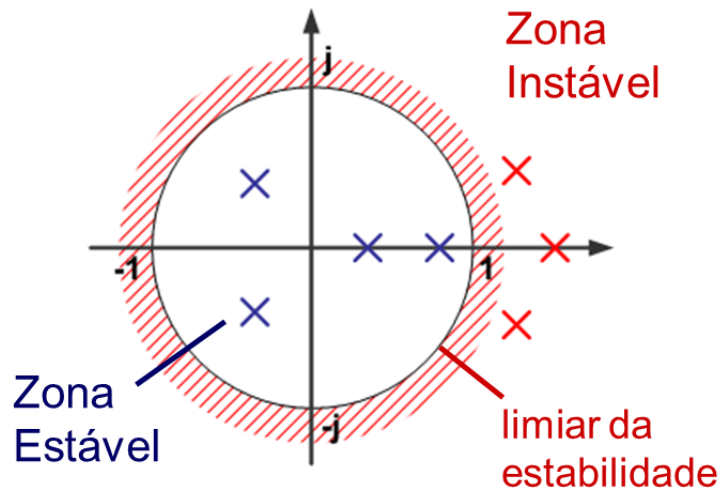


Figura 14.4 – Zona de estabilidade relativa à posição dos valores próprios em modelo de estado.

Exemplo 14.5

Considere-se de novo a instalação que admite a representação de estados dada por

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot u(k)$$

$$y(k) = [0 \quad 1] \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + [0] \cdot u(k)$$

Calculando os valores próprios iniciamos com

$$\det\{\lambda \mathbf{I} - \mathbf{A}\} = \det\left\{\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix}\right\} = \det\left\{\begin{bmatrix} \lambda - 0.6 & 0 \\ -0.6 & \lambda - 0.9 \end{bmatrix}\right\} = (\lambda - 0.6)(\lambda - 0.9) = 0$$

com raízes

$$\lambda_1 = 0.6 \text{ e } \lambda_2 = 0.9$$

ambas no interior do círculo unitário, e logo a instalação é estável. ■

14.8. Vetores próprios e desacoplamento dos modos dinâmicos

Os vetores próprios têm informação direcional sobre a relação entre as variáveis de estado que pode ser aproveitada no projeto dos sistemas de controlo. Vejamos primeiro um exemplo sobre a forma de os determinar.

Exemplo 14.6

Continuando o exemplo anterior, calculem-se agora os vetores próprios.

Associado ao primeiro valor próprio $\lambda_1 = 0.6$, temos

$$\begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = 0.6 \cdot \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix}$$

que resulta no sistema indeterminado

$$\begin{cases} 0.6 \cdot v_{11} = 0.6 \cdot v_{11} \\ 0.6 \cdot v_{11} = -0.3 \cdot v_{12} \end{cases}$$

que nos diz que são vetores próprios todos os vetores com uma direção tal que a segunda coordenada é o simétrico do dobro da primeira, em particular, é vetor próprio o vetor dado por

$$\mathbf{v}_1 = \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Associado ao segundo valor próprio $\lambda_2 = 0.9$, temos

$$\begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = 0.9 \cdot \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix}$$

que resulta no sistema indeterminado

$$\begin{cases} 0.6 \cdot v_{21} = 0.9 \cdot v_{21} \\ 0.6 \cdot v_{21} + 0.9 \cdot v_{22} = 0.9 \cdot v_{22} \end{cases}$$

ou seja, a componente v_{21} é forçosamente nula e a componente v_{22} é qualquer, em particular, é vetor próprio o vetor dado por

$$\mathbf{v}_2 = \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Note-se que qualquer vetor não nulo que tenha a mesma direção de \mathbf{v}_1 e \mathbf{v}_2 são também vetores próprios associados a λ_1 e λ_2 , respetivamente. Na realidade, faz mais sentido referir direções próprias da matriz da dinâmica \mathbf{A} .

■

Defina-se a **matriz modal** \mathbf{M} como a matriz quadrada de dimensão $n \times n$ cujas colunas são os vetores próprios da matriz \mathbf{A} . Na condição que os valores próprios da matriz \mathbf{A} sejam todos diferentes e reais, observa-se que o produto

$$\mathbf{\Lambda} = \mathbf{M}^{-1} \cdot \mathbf{A} \cdot \mathbf{M}$$

é uma matriz diagonal com os valores próprios na diagonal principal. Assim, nas condições enunciadas, a matriz $\mathbf{\Lambda}$ resultante é dada por

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \lambda_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \lambda_n \end{bmatrix}$$

Se usarmos a matriz modal para definir uma nova realização de estado tal que o novo vetor de estado seja dado por

$$\mathbf{z}(k) = \mathbf{M}^{-1} \cdot \mathbf{x}(k)$$

multiplicando (à esquerda) a equação da dinâmica do modelo de estado pela matriz \mathbf{M}^{-1} e introduzindo o elemento $\mathbf{M} \cdot \mathbf{M}^{-1} = \mathbf{I}$, obtemos uma nova realização de estado que é dada por

$$\mathbf{M}^{-1} \cdot \mathbf{x}(k+1) = \mathbf{M}^{-1} \cdot \mathbf{A} \cdot (\mathbf{M} \cdot \mathbf{M}^{-1}) \cdot \mathbf{x}(k) + \mathbf{M}^{-1} \cdot \mathbf{B} \cdot u(k)$$

$$y(k) = \mathbf{C} \cdot (\mathbf{M} \cdot \mathbf{M}^{-1}) \cdot \mathbf{x}(k) + D \cdot u(k)$$

ou

$$\mathbf{z}(k+1) = \mathbf{\Lambda} \cdot \mathbf{z}(k) + \mathbf{B}_M \cdot u(k)$$

$$y(k) = \mathbf{C}_M \cdot \mathbf{z}(k) + D \cdot u(k)$$

onde a matriz $\mathbf{\Lambda} = \mathbf{M}^{-1} \cdot \mathbf{A} \cdot \mathbf{M}$ é a nova matriz da dinâmica (com os mesmos valores próprios); e $\mathbf{B}_M = \mathbf{M}^{-1} \cdot \mathbf{B} = [b_{M,1} \ \cdots \ b_{M,n}]^T$ e $\mathbf{C}_M = \mathbf{C} \cdot \mathbf{M} = [c_{M,1} \ \cdots \ c_{M,n}]$ são as novas matrizes de entrada e saída.

Após a transformação através da matriz modal, que tem informação sobre as direções próprias da dinâmica do sistema, o sistema resultante caracteriza-se por ter os modos desacoplados. Como consequência, a evolução de cada uma das novas variáveis de estado é independente das restantes, i.e.

$$z_i(k+1) = \lambda_i \cdot z_i(k) + b_{M,i} \cdot u(k)$$

Exemplo 14.7

Continuando na instalação do **Exemplo 14.6**, a matriz modal (cujas colunas são vetores próprios da matriz \mathbf{A}) é dada por

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$$

com inversa (a verificar pelo leitor)

$$\mathbf{M}^{-1} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$$

Aplicando a transformação ao modelo de estado obtém-se a nova matriz da dinâmica

$$\mathbf{\Lambda} = \mathbf{M}^{-1} \cdot \mathbf{A} \cdot \mathbf{M} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.9 \end{bmatrix}$$

na forma diagonal no qual os elementos da diagonal principal coincidem com os dois valores próprios $\lambda_1 = 0.6$ e $\lambda_2 = 0.9$.

Podemos ainda calcular as novas matrizes de entrada e saída como

$$\mathbf{B}_M = \mathbf{M}^{-1} \cdot \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

e

$$\mathbf{C}_M = \mathbf{C} \cdot \mathbf{M} = [0 \quad 1] \cdot \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} = [-2 \quad 1]$$

Assim, a nova realização de estados é dada por

$$\begin{bmatrix} z_1(k+1) \\ z_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} z_1(k) \\ z_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} \cdot u(k)$$

$$y(k) = [-2 \quad 1] \cdot \begin{bmatrix} z_1(k) \\ z_2(k) \end{bmatrix} + [0] \cdot u(k)$$

representada na Figura 14.5. No esquema da figura pode observar-se a separação entre os modos dinâmicos associados aos dois valores próprios, como se tratasse de dois subsistemas que correm em paralelo sem se influenciarem mutuamente.

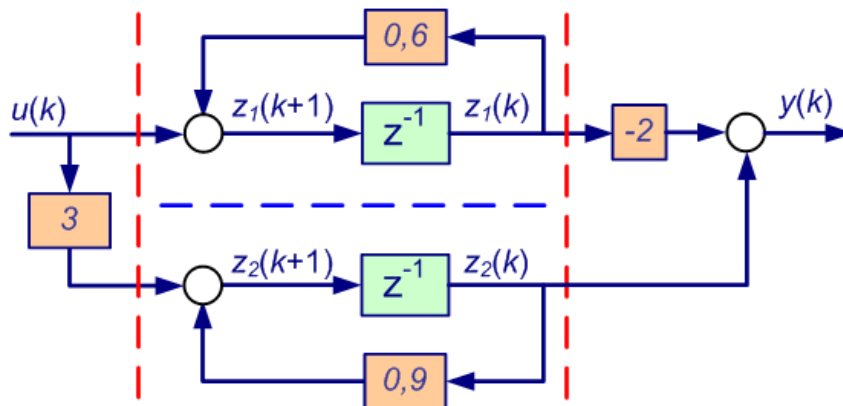


Figura 14.5 – Esquema de simulação do modelo de estado do Exemplo 14.2.

■

Deixa-se como exercício ao leitor o cálculo dos novos vetores próprios da realização na forma diagonal e respetiva interpretação do resultado.

Uma propriedade da forma diagonal é que o cálculo da matriz de transição passou a ser trivial porque

$$\Phi(k) = \Lambda^k = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}^k = \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix}$$

Esta propriedade sugere uma forma alternativa de cálculo da matriz de transição (nas outras formas não trivial) como se apresenta na Figura 14.6.

1. Calculam-se os valores e vetores próprios.

2. Se os valores próprios forem todos reais e diferentes constrói-se a matriz modal \mathbf{M} e calcula-se a sua inversa \mathbf{M}^{-1} .
3. Finalmente, calcula-se a matriz de transição na realização original como

$$\Phi(k) = \mathbf{M} \cdot \Lambda^k \cdot \mathbf{M}^{-1}$$

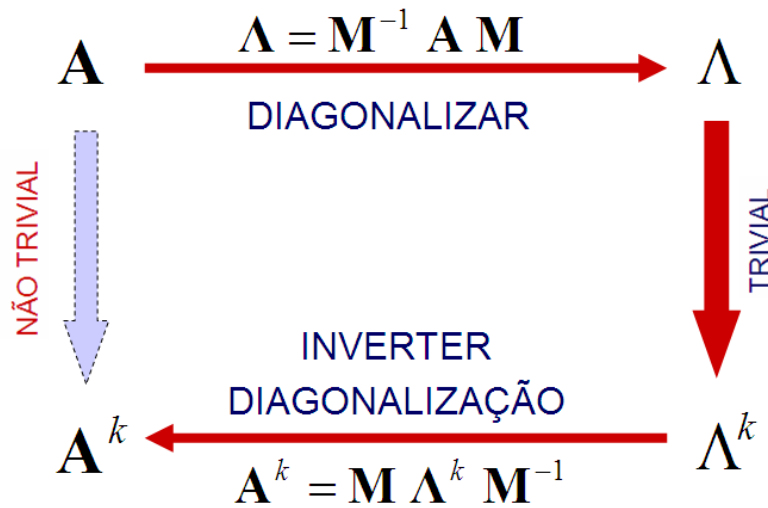


Figura 14.6 – Forma alternativa de cálculo da matriz de transição.

Antes de fechar este capítulo, importa referir que este assunto não se encerra aqui e é possível generalizar estes resultados mesmo para sistemas cuja matriz da dinâmica apresente valores próprios de multiplicidade superior à unidade ou complexos conjugados. Convida-se o leitor a procurar na bibliografia dedicada ao tema esta generalização.

15. Controlo em espaço de estados

A motivação para se introduzir a formulação em espaço de estados está relacionada com o desenvolvimento de alternativas para o projeto do controlo. Nesse sentido, é expectável que as novas ferramentas tirem partido do conhecimento acrescido sobre a dinâmica interna da instalação a controlar e, em particular, do acesso ao estado.

Interessa agora questionar se é possível condicionar o estado do sistema e em que extensão. Será possível, escolhendo convenientemente a sequência de entrada $u(k)$, levar o estado de um ponto inicial para um dado estado final \mathbf{x}_f ?

Refletindo sobre a estrutura do modelo de estado o que se pode concluir é que a resposta a essa questão deverá depender dos valores das matrizes **A** e **B**.

15.1. Controlabilidade completa

O sistema

$$\mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k)$$

$$y(k) = \mathbf{C} \cdot \mathbf{x}(k) + D \cdot u(k)$$

diz-se **completamente controlável** se, para $\mathbf{x}(0) = \mathbf{0}_n$ e qualquer valor do estado \mathbf{x}_f , existir um instante de tempo finito k_f e um sinal de entrada $u(k)$ definido no intervalo $k \in [0; k_f]$ tal que

$$\mathbf{x}(k_f) = \mathbf{x}_f$$

Esta definição diz-nos que o sistema é controlável se tivermos o poder de levar o conjunto de variáveis de estado para valores pré-determinados numa sequência finita de ações de controlo. Vejamos o seguinte exemplo.

Exemplo 15.1

Considere novamente a instalação introduzida no capítulo anterior no **Exemplo 14.1** dado por

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot u(k)$$

$$y(k) = [0 \quad 1] \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + [0] \cdot u(k)$$

Considere-se ainda que queremos levar o estado do sistema a partir de condições iniciais nulas para um ponto final dado por e.g.

$$\mathbf{x}_f = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Iterando a equação da dinâmica do modelo de estado temos

$$\begin{bmatrix} x_1(1) \\ x_2(1) \end{bmatrix} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot u(0) = \begin{bmatrix} u(0) \\ u(0) \end{bmatrix}$$

Não existe nenhum valor a atribuir a $u(0)$ que nos faça atingir o estado desejado \mathbf{x}_f num passo. Fazemos outro então. Iterando mais uma vez obtemos

$$\begin{bmatrix} x_1(2) \\ x_2(2) \end{bmatrix} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} u(0) \\ u(0) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot u(1) = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \begin{bmatrix} u(0) \\ u(0) \end{bmatrix} + \begin{bmatrix} u(1) \\ u(1) \end{bmatrix}$$

que igualando ao estado desejado \mathbf{x}_f resulta na sequência solução

$$u(0) = -1.1$$

$$u(1) = +2.7$$

■

No exemplo anterior mostrámos que era possível atingir um determinado estado final, mas precisaríamos agora de mostrar o mesmo para todos os valores possíveis de \mathbf{x}_f para poder dizer que o sistema é completamente controlável. Precisamos de uma forma mais expedita de demonstrar esta característica.

Note-se que no exemplo anterior iterámos o sistema à procura de ortogonalidade suficiente nas entradas subsequentes para podermos atingir um dado ponto no espaço em \mathbb{R}^2 . Vendo de uma perspetiva mais geral fizemos

$$\mathbf{x}(1) = \mathbf{A} \cdot \mathbf{x}(0) + \mathbf{B} \cdot u(0) = \mathbf{B} \cdot u(0)$$

seguido de

$$\mathbf{x}(2) = \mathbf{A} \cdot \mathbf{x}(1) + \mathbf{B} \cdot u(1) = \mathbf{A} \cdot \mathbf{B} \cdot u(0) + \mathbf{B} \cdot u(1)$$

Ao fim de dois passos temos a soma de dois vetores: $\mathbf{A} \cdot \mathbf{B}$ escalado por $u(0)$ e \mathbf{B} escalado por $u(1)$. A capacidade de atingir qualquer ponto no espaço \mathbb{R}^2 depende da existência de alguma ortogonalidade entre as direções do vetor $\mathbf{A} \cdot \mathbf{B}$ e do vetor \mathbf{B} . Se os vetores forem colineares por mais passos que demos não sairemos da mesma reta.

Então estamos em condições intuitivas para aceitar o seguinte resultado geral.

O sistema descrito por uma realização de estados é completamente controlável se e só se a **matriz de controlabilidade** definida por

$$\mathbf{M}_c = [\mathbf{B} \quad \mathbf{A}\mathbf{B} \quad \mathbf{A}^2\mathbf{B} \quad \dots \quad \mathbf{A}^{n-1}\mathbf{B}]$$

tiver característica i.e. o número de linhas ou colunas linearmente independentes, igual a n .

A forma mais expedita de verificar que a matriz tem característica completa é verificar que o seu determinante é diferente de zero.

Exemplo 15.2

Continuando o exemplo anterior com matrizes **A** e **B** dadas por

$$\mathbf{A} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

a matriz de controlabilidade é calculada como

$$\mathbf{M}_c = [\mathbf{B} \quad \mathbf{AB}] = \begin{bmatrix} 1 & 0.6 \\ 1 & 1.5 \end{bmatrix}$$

com determinante

$$|\mathbf{M}_c| = \begin{vmatrix} 1 & 0.6 \\ 1 & 1.5 \end{vmatrix} = 0.9$$

não nulo, e logo este sistema é completamente controlável. ■

Note-se que o problema de encontrar a sequência de controlo que nos leva do estado inicial ao estado final não é a mais interessante no projeto do controlo em tempo real que é o objeto do nosso estudo. Interessa-nos sim obter uma lei de controlo que resulte numa retroação permanente das variáveis da instalação que possam ser medidas através de sensores para atingir um comportamento dinâmico a especificar.

15.2. Controlo por retroação linear de variáveis de estado

Dada uma realização de estado

$$\mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k)$$

$$y(k) = \mathbf{C} \cdot \mathbf{x}(k) + D \cdot u(k)$$

com polinómio característico

$$p(\lambda) = |\lambda \mathbf{I} - \mathbf{A}| = \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n$$

e respetivo conjunto de valores próprios

$$\{\lambda_1, \lambda_2, \dots, \lambda_n\}$$

será possível através de uma lei de retroação linear de variáveis de estado

$$u(k) = -\mathbf{K} \cdot \mathbf{x}(k) + u_c(k)$$

reposicionar os valores próprios do sistema de controlo como se apresenta na Figura 15.1.

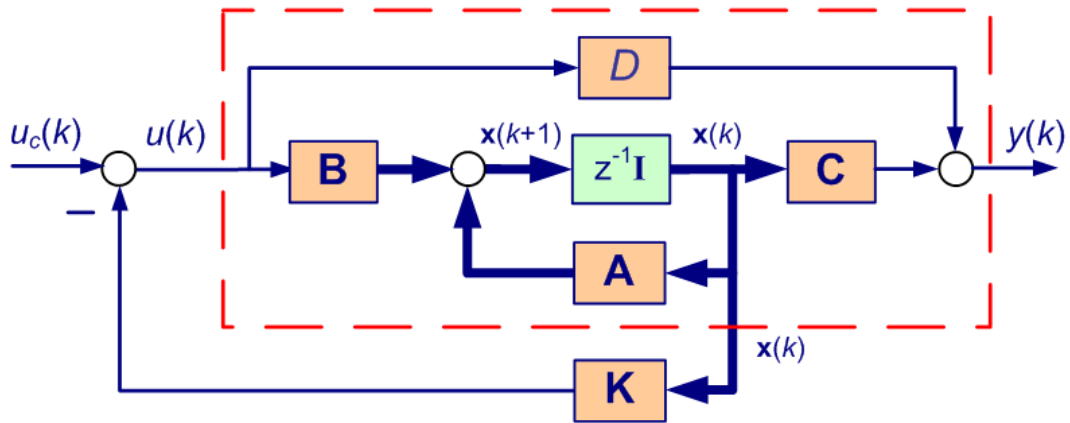


Figura 15.1 – Retroação linear de variáveis de estado.

A pergunta pode colocar-se na forma: será possível, através de uma lei de controlo da forma

$$u(k) = -\mathbf{K} \cdot \mathbf{x}(k) + u_c(k)$$

e por escolha adequada do vetor de ganhos de retroação \mathbf{K} , posicionar os valores próprios do sistema em anel fechado em valores específicos, e.g. nas raízes de um polinómio

$$\alpha(\lambda) = \lambda^n + \alpha_1 \lambda^{n-1} + \dots + \alpha_{n-1} \lambda + \alpha_n$$

qualquer?

Note-se que o papel da entrada de comando $u_c(k)$ colocada na lei de controlo serve apenas para manter algum acesso ao sistema a partir do exterior.

Existem aqui dois problemas: saber se é possível e determinar o valor do vetor de ganhos de retroação \mathbf{K} .

Substituindo a lei de controlo no modelo de estado obtemos

$$\mathbf{x}(k + 1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot \{-\mathbf{K} \cdot \mathbf{x}(k) + u_c(k)\}$$

$$y(k) = \mathbf{C} \cdot \mathbf{x}(k) + D \cdot \{-\mathbf{K} \cdot \mathbf{x}(k) + u_c(k)\}$$

que desenvolvendo resulta

$$\mathbf{x}(k + 1) = \{\mathbf{A} - \mathbf{B} \cdot \mathbf{K}\} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u_c(k)$$

$$y(k) = \{\mathbf{C} - D \cdot \mathbf{K}\} \cdot \mathbf{x}(k) + D \cdot u_c(k)$$

ou seja, um modelo de estado diferente e cuja matriz da dinâmica depende do vetor de ganhos \mathbf{K} a dimensionar.

A questão que se coloca é então relativa à seleção \mathbf{K} tal que

$$|\lambda \mathbf{I} - \mathbf{A} + \mathbf{B} \cdot \mathbf{K}| = \alpha(\lambda) = \lambda^n + \alpha_1 \lambda^{n-1} + \dots + \alpha_{n-1} \lambda + \alpha_n$$

Os valores próprios podem ser posicionados arbitrariamente se e só se o sistema for completamente controlável, i.e. se a sua matriz da controlabilidade tiver característica completa.

A determinação do valor do ganho \mathbf{K} que realiza o reposicionamento pode ser realizada através da **fórmula da Ackermann** com a seguinte redação.

Considere-se o modelo de estado e um polinómio $\alpha(\lambda)$ de ordem n , cujas raízes coincidem com o valor desejado dos valores próprios para o sistema de controlo (com retroação de estado). Então, o vetor de ganhos pode ser calculado a partir da fórmula compacta

$$\mathbf{K} = [0 \quad \dots \quad 0 \quad 1] \cdot \mathbf{M}_c^{-1} \cdot \alpha(\mathbf{A})$$

onde \mathbf{M}_c é a matriz de controlabilidade; e

$$\alpha(\mathbf{A}) = \mathbf{A}^n + \alpha_1 \mathbf{A}^{n-1} + \dots + \alpha_{n-1} \mathbf{A} + \alpha_n \mathbf{I}_n$$

Exemplo 15.3

Continuemos o exemplo deste capítulo para o qual já se demonstrou ser completamente controlável. Coloque-se o problema de determinar a lei de controlo por retroação de estado de modo a posicionar os valores próprios em

$$0.8 \pm j0.2$$

O primeiro passo de verificação da controlabilidade ficou feito no **Exemplo 15.2** concluindo que é possível realizar o controlador. Porque a característica da matriz de controlabilidade é completa é possível calcular a sua inversa dada por

$$\mathbf{M}_c^{-1} = \begin{bmatrix} 1 & 0.6 \\ 1 & 1.5 \end{bmatrix}^{-1} = \begin{bmatrix} 1.7 & -0.7 \\ -1.1 & 1.1 \end{bmatrix}$$

De seguida, desenvolvemos o polinómio $\alpha(\lambda)$ com raízes no local especificado, ou seja

$$\alpha(\lambda) = (\lambda - 0.8 - j0.2)(\lambda - 0.8 + j0.2) = \lambda^2 - 1.6\lambda + 0.68$$

e usamo-lo para calcular

$$\alpha(\mathbf{A}) = \mathbf{A}^2 - 1.6 \cdot \mathbf{A} + 0.68 \cdot \mathbf{I}_2 = \begin{bmatrix} 0.08 & 0 \\ -0.06 & 0.05 \end{bmatrix}$$

Finalmente, o vetor de ganhos é calculado por

$$\mathbf{K} = [0 \quad 1] \cdot \mathbf{M}_c^{-1} \cdot \alpha(\mathbf{A}) = [0 \quad 1] \cdot \begin{bmatrix} 1.7 & -0.7 \\ -1.1 & 1.1 \end{bmatrix} \cdot \begin{bmatrix} 0.08 & 0 \\ -0.06 & 0.05 \end{bmatrix} = [-0.16 \quad 0.056]$$

O controlador é então dado por

$$\mathbf{u}(k) = -\mathbf{K} \cdot \mathbf{x}(k) + \mathbf{u}_c(k) = -[-0.16 \quad 0.056] \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \mathbf{u}_c(k)$$

ou

$$\mathbf{u}(k) = 0.16 \cdot x_1(k) - 0.056 \cdot x_2(k) + \mathbf{u}_c(k)$$

■

Neste ponto impõem-se alguns comentários aos resultados que observámos acima.

Da fórmula de Ackermann, é óbvia a necessidade da matriz de controlabilidade ser invertível. Para sistemas com uma controlabilidade “fraca” os ganhos resultantes podem ter valores anormalmente elevados. Na implementação prática, deve verificar-se a consistência dos valores teóricos. Por exemplo, os ganhos podem ser de tal modo elevados que uma pequena variação da variável de estado produza imediatamente a saturação do atuador.

A retroação de estado exige acesso ao estado. As diversas variáveis de estado têm de ser medidas através de sensores para que possam ser usadas no cálculo do controlo. Para uma implementação prática, deve ter-se acesso através de n sensores às n variáveis de estado. No entanto, existem diversas razões pelas quais este acesso ao estado pode não ser possível de realizar, nomeadamente, se as variáveis de estado da nossa realização:

- não são variáveis com significado físico;
- são variáveis com significado físico mas não existe sensor para as medir;
- são variáveis com significado físico mas de acesso impossível na instalação;
- são variáveis com significado físico e existem os sensores mas não é económico.

entre outros possíveis motivos.

Precisamos de uma forma de aproveitar estes resultados mesmo para os casos em que só temos acesso parcial ao estado. Este assunto será desenvolvido no capítulo seguinte através da introdução dos observadores assintóticos.

16. Controlo com observadores de estado

Viu-se já que o projeto de controladores em espaço de estados pode ser realizado de forma muito expedita a partir do local desejado para o posicionamento dos valores próprios do sistema de controlo. Verifica-se a controlabilidade da instalação e calcula-se o vetor de ganhos do controlador através de uma fórmula compacta – a fórmula de Ackermann. Se o modelo representativo da instalação, na forma de quatro matrizes $\{A, B, C, D\}$, for de boa qualidade é garantido o posicionamento dos valores próprios no lugar geométrico desejado.

Esta facilidade e qualidade do projeto tem um custo que nem sempre é possível cobrir. Para realizarmos a retroação de estado necessitamos de ter acesso ao estado através de sensores que disponibilizem os sinais para retroação. Como fazer quando estes sinais não estão disponíveis?

Se queremos fazer controlo por retroação (que é o tema deste documento) algum sinal terá que ser medido através dum sensor. Frequentemente está disponível apenas a variável física a controlar que designamos pela saída da instalação (e até agora representada pela letra y). Portanto, estamos na situação em que a saída do sistema y é medida, mas as variáveis de estado x_i não estão todas disponíveis para o controlo.

16.1. Observador assintótico

A solução para este problema tira partido da capacidade de processamento do computador para simular o comportamento da instalação em paralelo com a mesma (ver Figura 16.1) e disponibilizar uma estimativa do estado, representada pelo símbolo $\hat{x}(k)$.

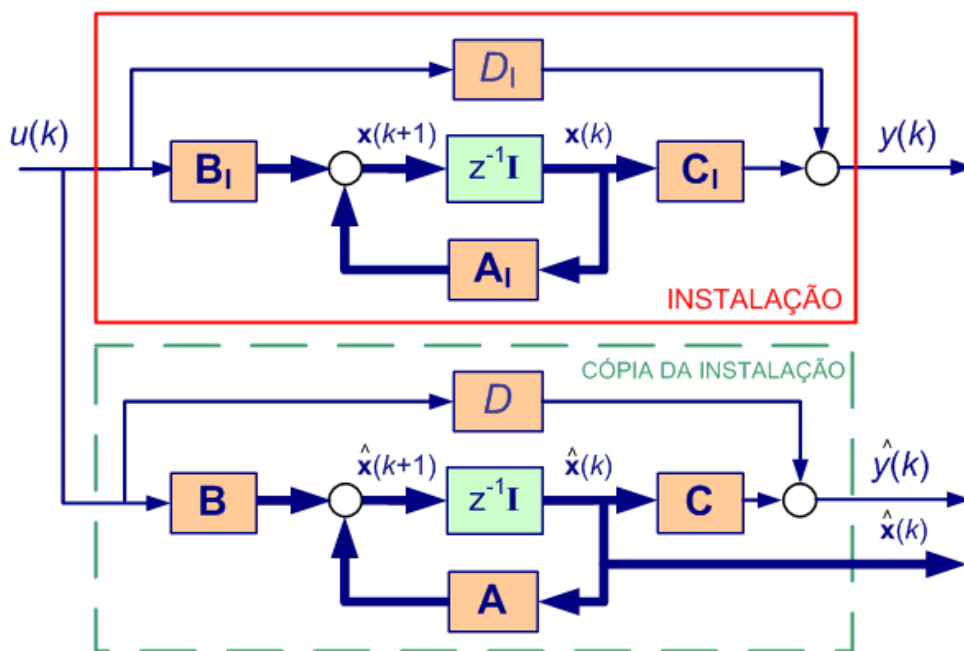


Figura 16.1 – Simulação paralela para gerar estimativa do estado.

Se o modelo $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ fosse um modelo completo e perfeito do comportamento da instalação podíamos usar o estado estimado $\hat{\mathbf{x}}(k)$ do simulador para realizar controlo, mesmo sem ter acesso ao estado real $\mathbf{x}(k)$.

No entanto, o esquema da figura, tal como está representado, está condenado ao insucesso. Os motivos são:

- A aproximação apresentada pressupõe que é conhecido com exatidão o modelo de estado da instalação $\{\mathbf{A}_I, \mathbf{B}_I, \mathbf{C}_I, \mathbf{D}_I\}$. O modelo da instalação é conhecido sempre com um certo grau de incerteza associado.
- A aproximação apresentada pressupõe que existe um modelo de estado da instalação $\{\mathbf{A}_I, \mathbf{B}_I, \mathbf{C}_I, \mathbf{D}_I\}$. Se a instalação for não linear o modelo de estado só é válido em redor do ponto onde foi obtida a linearização.
- A aproximação apresentada pressupõe que conhecemos com exatidão as condições iniciais $\mathbf{x}(0)$ do estado da instalação. Adicionalmente, as perturbações que afetam a instalação não estão disponíveis para afetar o modelo em simulação.
- Na realidade, o conjunto de incerteza associado ao processo de modelação resultará numa divergência entre o comportamento da instalação e o comportamento do simulador em paralelo.
- Finalmente, se a estimativa do estado for usada para retroação, estaremos perante um controlador puramente antecipativo (*feedforward*), ou seja, sem qualquer tipo de retroação a partir da saída da instalação.

Necessitamos de algo que utilize a melhor informação provinda da instalação, i.e. a saída $y(k)$, e compara-la com o seu equivalente no simulador, i.e. a saída estimada $\hat{y}(k)$.

O esquema do **observador de estado assintótico** propõe que as saídas sejam comparadas e utilizadas para corrigir a estimativa do estado.

Assim, o observador de estado é uma réplica da instalação, excitado pela mesma entrada, com o estado atualizado segundo o modelo de estado

$$\hat{\mathbf{x}}(k+1) = \mathbf{A} \cdot \hat{\mathbf{x}}(k) + \mathbf{B} \cdot u(k) + \mathbf{L} \cdot (y(k) - \hat{y}(k))$$

$$\hat{y}(k) = \mathbf{C} \cdot \hat{\mathbf{x}}(k) + \mathbf{D} \cdot u(k)$$

O termo $\mathbf{L} \cdot (y(k) - \hat{y}(k))$ será nulo se a saída estimada for igual à saída da instalação, não executando qualquer correção à evolução do estado estimado $\hat{\mathbf{x}}(k+1)$. Se existir um desvio, a matriz \mathbf{L} de dimensão $(n \times 1)$ indica em que direção a estimativa deve ser corrigida.

A Figura 16.2 apresenta a estrutura final do sistema de controlo com a instalação, o observador e a retroação do estado.

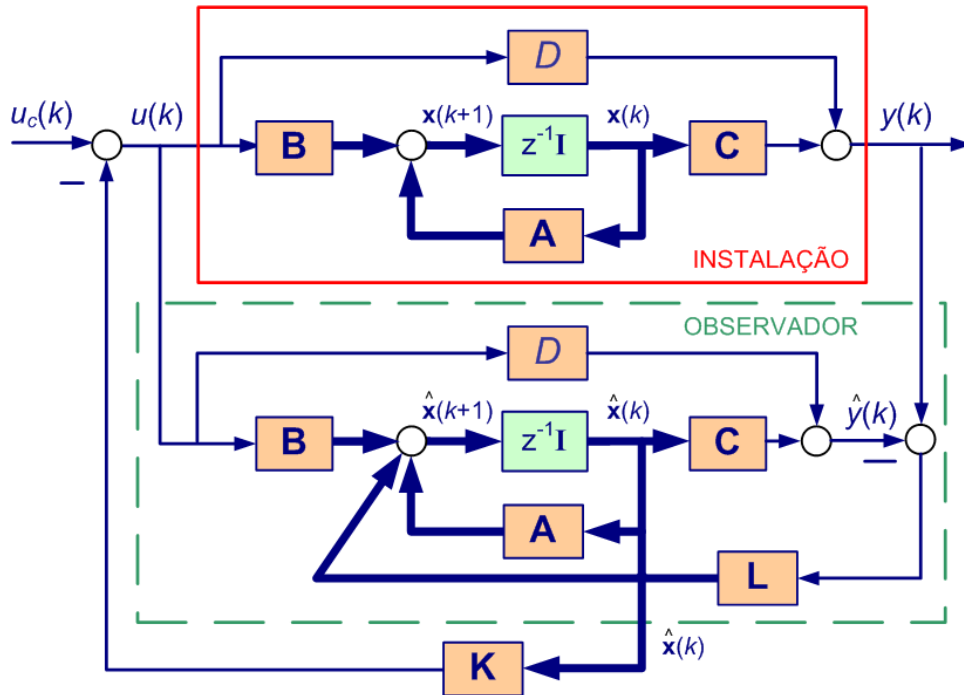


Figura 16.2 – Simulação paralela para gerar estimativa do estado.

Substituindo a estimativa da saída $\hat{y}(k)$ na equação da dinâmica do observador obtém-se

$$\hat{\mathbf{x}}(k + 1) = (\mathbf{A} - \mathbf{L} \cdot \mathbf{C}) \cdot \hat{\mathbf{x}}(k) + (\mathbf{B} - \mathbf{L} \cdot \mathbf{D}) \cdot u(k) + \mathbf{L} \cdot y(k)$$

Ou seja, o observador é um sistema dinâmico com valores próprios nas soluções de

$$|\lambda \mathbf{I} - \mathbf{A} + \mathbf{L} \cdot \mathbf{C}| = 0$$

Mais importante que a dinâmica do observador isolado, é a dinâmica do conjunto observador-instalação, nomeadamente, o erro de estimativa do estado

$$\tilde{\mathbf{x}}(k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k)$$

é governado pela dinâmica

$$\tilde{\mathbf{x}}(k + 1) = (\mathbf{A} - \mathbf{L} \cdot \mathbf{C}) \cdot \tilde{\mathbf{x}}(k)$$

Se os valores próprios da matriz da dinâmica do erro $(\mathbf{A} - \mathbf{L} \cdot \mathbf{C})$ forem estáveis, i.e. no interior do círculo unitário, então o erro entre o estado real e o estado estimado converge assintoticamente para zero. Por este motivo o observador é designado de assintótico.

Surgem as questões sobre a possibilidade de realizar tal observador; e de onde e como posicionar os valores próprios do observador.

A realização do observador assintótico só é possível se a instalação for observável.

16.2. Observabilidade completa

O sistema

$$\mathbf{x}(k + 1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k)$$

$$y(k) = \mathbf{C} \cdot \mathbf{x}(k) + D \cdot u(k)$$

diz-se **completamente observável** se existir um instante de tempo $k_1 > 0$ tal que o conhecimento da saída $y(k)$ no intervalo $k \in [0; k_1]$, for suficiente para determinar o estado inicial $\mathbf{x}(0)$.

Esta definição diz-nos que o sistema é observável se conseguirmos estimar o estado num instante do tempo a partir de uma sequência finita de amostras da saída $y(k)$. Vejamos o seguinte exemplo.

Exemplo 16.1

Considere novamente a instalação introduzida no capítulo anterior no **Exemplo 14.1** com entrada identicamente nula $\mathbf{u}(k) \equiv \mathbf{0}$, dada por

$$\begin{bmatrix} x_1(k + 1) \\ x_2(k + 1) \end{bmatrix} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$

$$y(k) = [0 \quad 1] \cdot \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$

e o conhecimento da sequência de saída $y(0) = 2$ e $y(1) = 1$.

Iterando as equações do modelo de estado temos

$$y(0) = [0 \quad 1] \cdot \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}$$

Ficamos a saber que o estado inicial $x_2(0) = 2$ mas o valor de $x_1(0)$ é indeterminado.

Façamos outro então. Iterando mais uma vez obtemos

$$y(1) = [0 \quad 1] \cdot \begin{bmatrix} x_1(1) \\ x_2(1) \end{bmatrix} = [0 \quad 1] \cdot \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \cdot \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}$$

que resulta no conhecimento $y(1) = 0.6 \cdot x_1(0) + 0.9 \cdot x_2(0)$ que resulta em

$$x_1(0) = \frac{1}{0.6}(1 - 0.9 \cdot 2) = -1.3$$

Portanto, o resultado da estimação do estado inicial é

$$x_1(0) = +2.0$$

$$x_2(0) = -1.3$$

■

No exemplo anterior mostrámos que era possível estimar um determinado estado inicial, mas precisaríamos agora de mostrar o mesmo para todas as sequências de saída para poder dizer que o sistema é completamente observável. Precisamos de uma forma mais expedita de demonstrar esta característica.

À semelhança com o problema da controlabilidade, a observabilidade está dependente da existência de ortogonalidade suficiente na forma como o estado é projetado na saída.

No exemplo anterior em espaço em \mathbb{R}^2 , temos

$$y(0) = \mathbf{C} \cdot \mathbf{x}(0)$$

$$y(1) = \mathbf{C} \cdot \mathbf{A} \cdot \mathbf{x}(0)$$

ou

$$\begin{bmatrix} y(0) \\ y(1) \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C} \cdot \mathbf{A} \end{bmatrix} \cdot \mathbf{x}(0)$$

que tem solução $\mathbf{x}(0)$ apenas se existir alguma ortogonalidade entre as direções do vetor \mathbf{C} e do vetor $\mathbf{C} \cdot \mathbf{A}$.

Então estamos em condições intuitivas para aceitar o seguinte resultado geral.

O sistema descrito por uma realização de estados é completamente observável se e só se a **matriz de observabilidade** definida por

$$\mathbf{M}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \mathbf{C}\mathbf{A}^2 \\ \vdots \\ \mathbf{C}\mathbf{A}^{n-1} \end{bmatrix}$$

tiver característica i.e. o número de linhas ou colunas linearmente independentes, igual a n .

A forma mais expedita de verificar que a matriz tem característica completa é verificar que o seu determinante é diferente de zero.

Exemplo 16.2

Continuando o exemplo anterior com matrizes \mathbf{A} e \mathbf{C} dadas por

$$\mathbf{A} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix}$$

$$\mathbf{C} = [0 \quad 1]$$

a matriz de controlabilidade é calculada como

$$\mathbf{M}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0.6 & 0.9 \end{bmatrix}$$

com determinante

$$|\mathbf{M}_o| = \begin{vmatrix} 0 & 1 \\ 0.6 & 0.9 \end{vmatrix} = -0.6$$

não nulo, e logo este sistema é completamente observável.

■

16.3. Projeto do observador assintótico

Coloca-se agora a questão de como dimensionar o valor do vetor de ganhos do observador \mathbf{L} tal que

$$|\lambda \mathbf{I} - \mathbf{A} + \mathbf{L} \cdot \mathbf{C}| = \beta(\lambda) = \lambda^n + \beta_1 \lambda^{n-1} + \dots + \beta_{n-1} \lambda + \beta_n$$

Os valores próprios do observador podem ser posicionados arbitrariamente se e só se o sistema for completamente observável, i.e. se a sua matriz da observabilidade tiver característica completa.

A determinação do valor do ganho \mathbf{L} que realiza o reposicionamento pode ser realizada através da **fórmula da Ackermann** com a seguinte redação.

Considere-se o modelo de estado e um polinómio $\beta(\lambda)$ de ordem n , cujas raízes coincidem com o valor desejado dos valores próprios para o observador. Então, o vetor de ganhos pode ser calculado a partir da fórmula compacta

$$\mathbf{L} = \beta(\mathbf{A}) \cdot \mathbf{M}_o^{-1} \cdot \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

onde \mathbf{M}_o é a matriz de observabilidade; e

$$\beta(\mathbf{A}) = \mathbf{A}^n + \beta_1 \mathbf{A}^{n-1} + \dots + \beta_{n-1} \mathbf{A} + \beta_n \mathbf{I}_n$$

Exemplo 16.3

Continuemos o exemplo deste capítulo para no qual para instalação em causa já se demonstrou ser completamente observável. Coloque-se o problema de determinar o ganho do observador de modo a posicionar ambos os valores próprios em **0.14**.

O primeiro passo de verificação da observabilidade ficou feito no **Exemplo 16.2** concluindo que é possível realizar o observador. Porque a característica da matriz de observabilidade é completa é possível calcular a sua inversa dada por

$$\mathbf{M}_o^{-1} = \begin{bmatrix} 0 & 1 \\ 0.6 & 0.9 \end{bmatrix}^{-1} = \begin{bmatrix} -1.5 & 1.7 \\ 1 & 0 \end{bmatrix}$$

De seguida, desenvolvemos o polinómio $\beta(\lambda)$ com raízes no local especificado, ou seja

$$\beta(\lambda) = (\lambda - 0.14)^2 = \lambda^2 - 0.28\lambda + 0.0196$$

e usamo-lo para calcular

$$\beta(\) = A^2 - 0.28 \cdot A + 0.0196 \cdot I_2 = \begin{bmatrix} 0.21 & 0 \\ 0.73 & 0.58 \end{bmatrix}$$

Finalmente, o vetor de ganhos é calculado por

$$L = \beta(A) \cdot M_0^{-1} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.21 & 0 \\ 0.73 & 0.58 \end{bmatrix} \cdot \begin{bmatrix} -1.5 & 1.7 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.35 \\ 1.2 \end{bmatrix}$$

A Figura 16.3 mostra a estrutura final do controlador com o observador dimensionado neste exemplo e o controlador dimensionado em exemplos anteriores para a mesma instalação. Convida-se o leitor a interpretar os diversos blocos presentes no esquema.

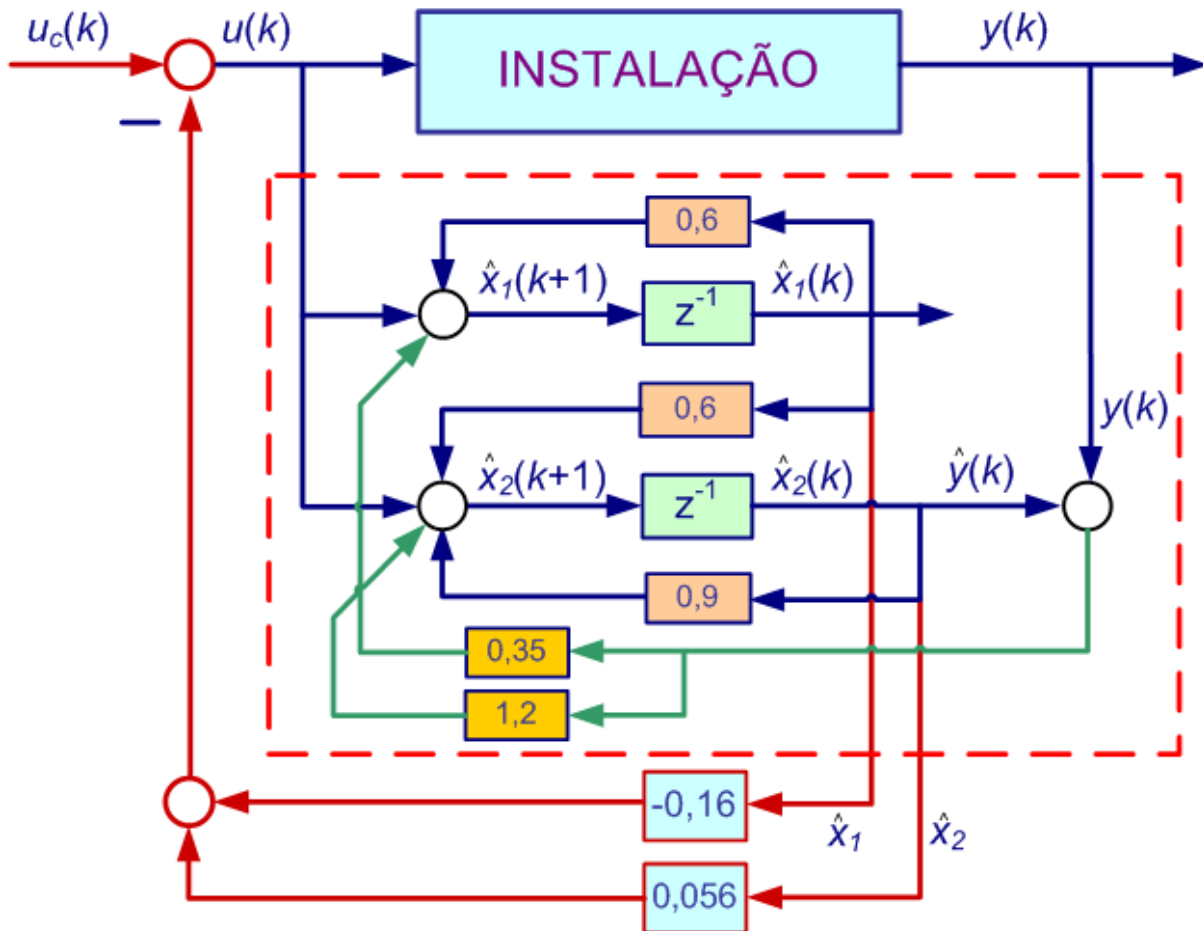


Figura 16.3 – Estrutura final do controlador com o observador do Exemplo 16.3.

■

No exemplo anterior selecionou-se uma dinâmica para o observador com valores próprios em 0.14 sem mais comentários adicionais. Como devemos especificar esta dinâmica?

No mínimo, será desejável que a dinâmica do observador, i.e. a dinâmica do erro de estimação, convirja para zero com constantes de tempo inferiores às constantes de tempo da instalação (controlada ou não). Em tempo contínuo, usa-se a regra de garantir uma relação entre constantes de tempo de, pelo menos, dez vezes. O valor usado no exemplo derivou dessa prática. Como seria esperado, quanto mais rápida for a dinâmica solicitada maior o valor

absoluto dos ganhos resultantes. Isto pode levar a que certas variáveis intermédias atinjam valores mais elevados que a ordem de grandeza das variáveis de estado.

No entanto, em controlo em tempo discreto, a implementação do observador é realizada inteiramente na forma programada, à partida sem qualquer restrição nos valores máximos das variáveis intermédias. Assim, faz sentido seleccionar para a dinâmica do observador um conjunto de valores próprios todos na origem, ou seja, fazer $\beta(\lambda) = \lambda^n$.

16.4. Teorema da separação dos modos

Falta justificar um passo que, de certa forma, demos depressa de mais no último exemplo. Dimensionámos o controlador sem pensar na existência do observador. E dimensionámos o observador sem levar em consideração que iríamos usar o estado observado para retroacção. No final juntámos ambos – controlador e observador – e esperamos que funcione.

O que nos permite ter tal expectativa?

A resposta reside no teorema da separação dos modos que apresentamos de seguida.

A instalação é descrita pelo modelo

$$\mathbf{x}(k + 1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k)$$

$$\mathbf{y}(k) = \mathbf{C} \cdot \mathbf{x}(k) + D \cdot u(k)$$

O controlador realiza

$$u(k) = -\mathbf{K} \cdot \hat{\mathbf{x}}(k) + u_c(k)$$

e o observador, na perspectiva do erro de estimação $\tilde{\mathbf{x}}(k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k)$, tem a dinâmica descrita por

$$\tilde{\mathbf{x}}(k + 1) = (\mathbf{A} - \mathbf{L} \cdot \mathbf{C}) \cdot \tilde{\mathbf{x}}(k)$$

A junção destas componentes resulta num modelo de estado global dado por

$$\begin{bmatrix} \mathbf{x}(k + 1) \\ \tilde{\mathbf{x}}(k + 1) \end{bmatrix} = \begin{bmatrix} (\mathbf{A} - \mathbf{B} \cdot \mathbf{K}) & \mathbf{B} \cdot \mathbf{K} \\ \mathbf{0}_{n \times n} & (\mathbf{A} - \mathbf{L} \cdot \mathbf{C}) \end{bmatrix} \begin{bmatrix} \mathbf{x}(k) \\ \tilde{\mathbf{x}}(k) \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0}_{n \times 1} \end{bmatrix} u_c(k)$$

Existe uma propriedade da álgebra relativa ao cálculo do determinante de matrizes compostas por outras matrizes dada por

$$\begin{vmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{0} & \mathbf{R} \end{vmatrix} = |\mathbf{P}| \times |\mathbf{R}|$$

Aplicando esta propriedade à determinação dos valores próprios do sistema global obtém-se

$$\left| \lambda \cdot \mathbf{I}_{2n} - \begin{bmatrix} (\mathbf{A} - \mathbf{B} \cdot \mathbf{K}) & \mathbf{B} \cdot \mathbf{K} \\ \mathbf{0}_{n \times n} & (\mathbf{A} - \mathbf{L} \cdot \mathbf{C}) \end{bmatrix} \right| = |\lambda \cdot \mathbf{I}_n - \mathbf{A} + \mathbf{B} \cdot \mathbf{K}| \times |\lambda \cdot \mathbf{I}_n - \mathbf{A} + \mathbf{L} \cdot \mathbf{C}| = 0$$

Ou seja, o teorema da separação dos modos estabelece que os vetores de ganhos do controlador K e do observador L podem ser projectados em separado, que a dinâmica resultante é a sobreposição das duas.

16.5. Função de transferência do controlador

No final do processo do projeto do controlo recorrendo ao observador para estimar o estado temos um controlador que inclui o observador e a retroação do estado (estimado) como se representa na Figura 16.4.

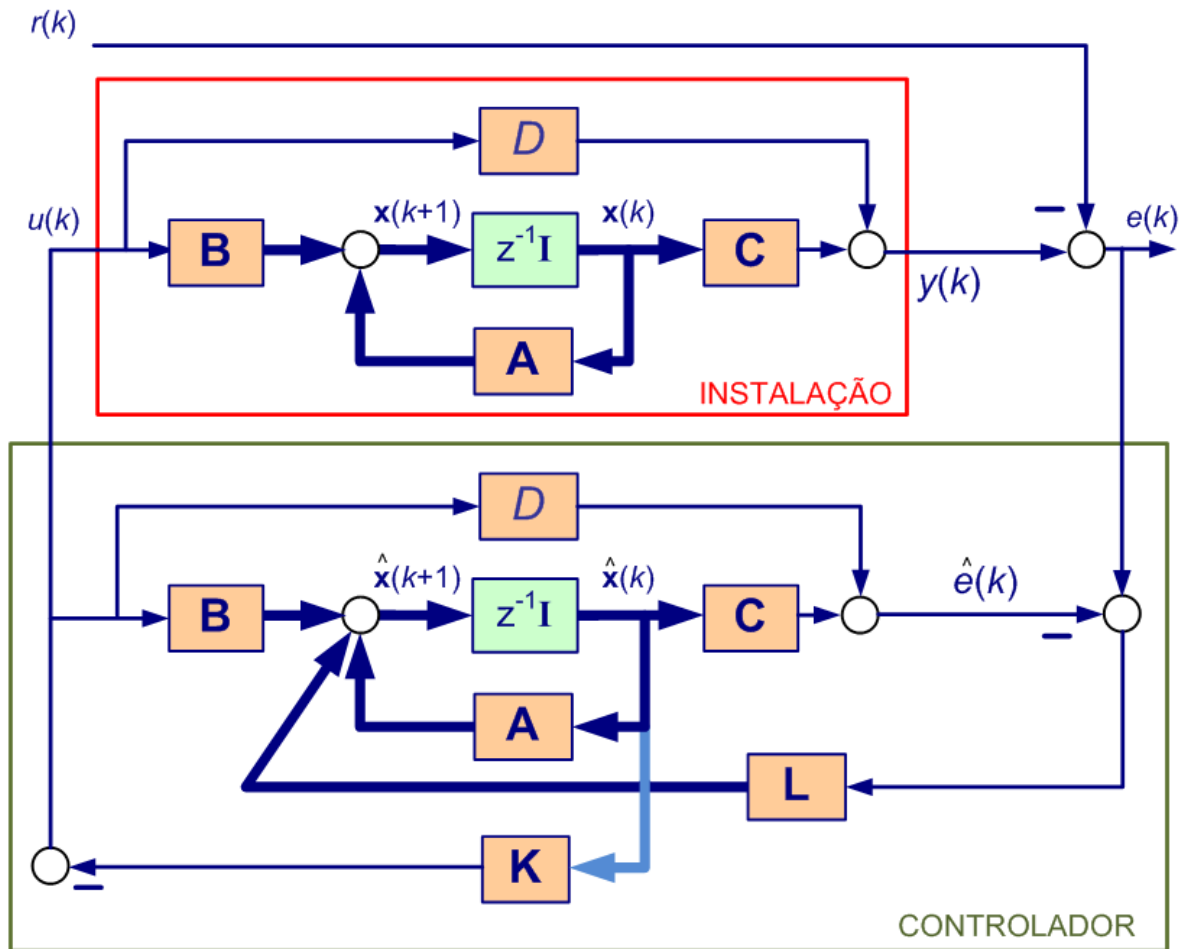


Figura 16.4 – Esquema global do sistema de controlo.

Todos os elementos que não pertencem à instalação pertencem ao controlador. Podemos colocar a questão de qual o modelo de estado só do controlador e, eventualmente, qual a função de transferência correspondente. Um motivo natural seria a facilidade de implementação do controlador. Em vez de programar todos os elementos do observador um por um, podemos, a partir da função de transferência, programar diretamente a equação às diferenças com a lei de controlo.

Observando a figura, deixa-se como exercício ao leitor, demonstrar que o modelo de estado do controlador é definido por

$$\hat{\mathbf{x}}(k+1) = (\mathbf{A} - \mathbf{B} \cdot \mathbf{K} - \mathbf{L} \cdot \mathbf{C} - \mathbf{L} \cdot \mathbf{D} \cdot \mathbf{K}) \cdot \hat{\mathbf{x}}(k) + \mathbf{L} \cdot e(k)$$

$$u(k) = -\mathbf{K} \cdot \hat{\mathbf{x}}(k)$$

com função de transferência, obtida a partir da fórmula correspondente, dada por

$$C(z) = -\mathbf{K} \cdot (z \cdot \mathbf{I} - \mathbf{A} + \mathbf{B} \cdot \mathbf{K} + \mathbf{L} \cdot \mathbf{C} + \mathbf{L} \cdot \mathbf{D} \cdot \mathbf{K})^{-1} \cdot \mathbf{L}$$

ou

$$C(z) = \mathbf{K} \cdot (z \cdot \mathbf{I} - \mathbf{A} + \mathbf{B} \cdot \mathbf{K} + \mathbf{L} \cdot \mathbf{C} + \mathbf{L} \cdot \mathbf{D} \cdot \mathbf{K})^{-1} \cdot \mathbf{L}$$

se o erro de seguimento for definido da forma habitual $e(k) = r(k) - y(k)$.

Exemplo 16.4

Na linha dos exemplos que temos seguido, pode calcular-se agora o controlador correspondente ao sistema de controlo da Figura 16.3.

Reúna-se o conjunto de elementos relativos à instalação, observador e controlador:

$$\mathbf{A} = \begin{bmatrix} 0.6 & 0 \\ 0.6 & 0.9 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{C} = [0 \quad 1] \quad \mathbf{D} = [0]$$

$$\mathbf{K} = [-0.16 \quad 0.056] \quad \mathbf{L} = \begin{bmatrix} 0.35 \\ 1.2 \end{bmatrix}$$

que resulta no controlador

$$C(z) = \frac{0.0112z + 0.02184}{z^2 - 0.404z + 0.038}$$

■