

DE GRUYTER

GRADUATE

*Paul Șerban Agachi, Mircea Vasile Cristea,  
Alexandra Ana Csavdári, Botond Szilágyi*

# ADVANCED PROCESS ENGINEERING CONTROL

Copyright 2017, De Gruyter. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

ERSCO  
AN: 145  
Enginee  
Account

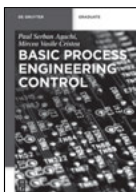
DE  
GRUYTER

ng : eBook Collect  
Paul Șerban  
Control  
141

host) - printed on 2017-03-27 08 AM Vi  
Mircea Vasile Cristea, Alexandra Ana Csavdári, Botond Szilágyi

Agachi, Cristea, Csavdári, Szilágyi  
**Advanced Process Engineering Control**  
De Gruyter Graduate

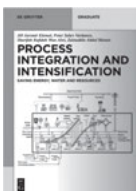
## Also of interest



*Basic Process Engineering Control.*

Agachi, Cristea, 2014

ISBN 978-3-11-028981-7, e-ISBN 978-3-11-028982-4



*Process Integration and Intensification.*

*Saving Energy, Water and Resources*

Klemeš, Varbanov, Wan Alwi, Manan, 2014

ISBN 978-3-11-030664-4, e-ISBN 978-3-11-030685-9

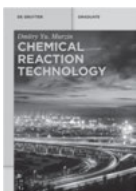


*Process Technology.*

*An Introduction*

De Haan, 2015

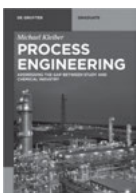
ISBN 978-3-11-033671-9, e-ISBN 978-3-11-033672-6



*Chemical Reaction Technology.*

Murzin, 2015

ISBN 978-3-11-033643-6, e-ISBN 978-3-11-033644-3



*Process Engineering.*

*Addressing the Gap between Studies and Chemical Industry*

Kleiber, 2016

ISBN 978-3-11-031209-6, e-ISBN 978-3-11-031211-9



*Reviews in Chemical Engineering.*

Dan Luss, Neima Brauner (Editors-in-Chief)

ISSN 0167-8299, e-ISSN 2191-0235

Paul Şerban Agachi, Mircea Vasile Cristea,  
Alexandra Ana Csavdári, Botond Szilágyi

# Advanced Process Engineering Control

---

**DE GRUYTER**

## Authors

### **Prof. Dr. Paul Șerban Agachi**

Babeș-Bolyai University  
Faculty of Chemistry and Chemical Engineering  
Arany Janos Street 11  
400028 Cluj Napoca  
Romania  
serban.agachi@ubbcluj.ro

Botswana International University of  
Science and Technology  
Private Bag 16, Palapye  
Botswana  
agachip@biust.ac.bw

### **Prof. Dr. Mircea Vasile Cristea**

University Babeș-Bolyai  
Faculty of Chemistry and Chemical Engineering  
Arany Janos Street 11  
400028 Cluj-Napoca  
Romania  
mcristea@chem.ubbcluj.ro

### **Assoc. Prof. Dr. Alexandra Ana Csavdári**

Babeș-Bolyai University  
Faculty of Chemistry and Chemical Engineering  
Arany Janos Street 11  
400028 Cluj-Napoca  
Romania  
acsavdari@chem.ubbcluj.ro

### **Dr. Botond Szilágyi**

Babeș-Bolyai University  
Faculty of Chemistry and Chemical Engineering  
Arany Janos Street 11  
400028 Cluj-Napoca  
Romania  
botiszilagyi@yahoo.com

ISBN 978-3-11-030662-0

e-ISBN (PDF) 978-3-11-030663-7

e-ISBN (EPUB) 978-3-11-038816-9

### **Library of Congress Cataloging-in-Publication Data**

A CIP catalog record for this book has been applied for at the Library of Congress.

### **Bibliographic information published by the Deutsche Nationalbibliothek**

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2017 Walter de Gruyter GmbH, Berlin/Boston

Typesetting: Compuscript Ltd., Shannon, Ireland

Printing and binding: CPI books GmbH, Leck

Cover image: Janaka Dharmasena/Hemera/Thinkstock

♻️ Printed on acid-free paper

Printed in Germany

[www.degruyter.com](http://www.degruyter.com)

## Preface

The present work, *Advanced Process Engineering Control*, is intended to be the continuation of the authors' *Basic Process Engineering Control* published by DeGruyter in 2014. It presents the main and conventional type control loops in process industries. Titles containing the concept of process engineering were deliberately chosen to suggest the inclusion, within the same approach, of processes other than the traditional ones. These come from outside the traditional fields of chemistry and petrochemistry: the sphere of pharmaceuticals, wastewater management, water purification, water reserve management, construction material industry, food processing, household or automotive industries.

During use and development of automatic control systems, control analysis and control system design for process industries have followed the traditional unit operation approach. It means that all control loops are established individually for each unit or piece of equipment in the plant and that the final plantwide control system represents the sum of the individual parts. The disadvantage of this method is the difficulty in stabilizing potential conflicts among individual loops. One very handy method of avoiding these interactions is the different tuning of control loops: those controlling the most important parameters are tuned tight and the others loose. Despite any process complexity, the unit operation approach provided reasonable results and remains in use on a large scale for designing control systems. Consequently, the book follows this traditional approach but provides updates to new industrial achievements.

Modern process plants are designed for flexible production and maximization of energy and material savings, especially within the frame of globalization and strong competition among manufacturers. Additionally, according to the fourth paradigm of process /chemical engineering, the processes have to fulfill tight environmental constraints. Industrial plants become more complex and have therefore strong interactions between process units. As a consequence, the failure of one unit might have a negative effect both on overall productivity and on environmental performances. This situation raises important control issues. A significant example is that of the thermally integrated plants, a concept born during the global energy crisis that started in 1973. Energy recovery became a priority for the industry and at the same time a scientific challenge. The necessity of redesigning industrial processes in terms of energetic efficiency was identified. Moreover, it was discovered that energy saving can be achieved using retrofit and recovery of extra energy from all secondary sources of a process. These aspects posed complex control problems because of the weak process controllability (effects of all disturbances are collected at the end of the process and reintroduced as enhanced disturbances at the input).

The emergence and continuous development of advanced control techniques provided solutions for plantwide control at any level of process complexity and in the

above-mentioned conditions. According to Willis and Tham, a definition of the advanced process control can be formulated as “a systematically studied approach for the choice of pertinent techniques and their integration into a co-operative management and control system that will significantly enhance plant operation and profitability”.

Applied on complex chemical processes, advanced control is able to improve product yield, reduce energy consumption, increase plant capacity, improve product quality and consistency, enhance process safety, and reduce environmental impact. The benefits of the advanced control implementation are noticeable in the overall operating costs of a plant. These can decrease by 2% to 6%. Another benefit is the reduction of process variability. As a consequence, a plant can be operated at its designed capacity.

The present book is structured into two parts. Part I, entitled *Advanced Process Control* comprises chapters 1–7 and defines as advanced control any control system that surpasses simple and conventional loops. This could mean either smarter control configuration (cascade, feedforward, ratio, inferential, digital, or multivariable control) or improved regulator features (fuzzy, model predictive or optimal control). Approaches for the design of plantwide control systems are also presented. Part II, entitled *Applied Process Engineering Control* includes chapters 8–14 and refers to control solutions for the so-called unit operations: reaction and separation processes (distillation, absorption-desorption, extraction, evaporation, drying, and crystallization). The reader can check her or his level of comprehension by solving the problems and exercises proposed in Chapter 15. These cover the entire list of discussed topics.

The authors hope that by including many industrial examples and applications as well as their own and other researcher’s experience accumulated over many years within the Group of Computer Aided Process Control, the present work will be useful for all interested parties in process engineering and process control: students in electrical, chemical, or process engineering; specialists in chemical, petrochemical or automation companies; professionals of water or natural gas management; etc.

The idea of this book series describing the main aspects of modern process engineering as applied to (not only) chemical industry belongs to Prof. Dr. Paul Șerban Agachi. He initiated the manuscripts, developed their structure, and coordinated the authors. More than 20 years ago, he recognized the ever-increasing importance of the subject and founded the *Group of Computer Aided Process Engineering* at the Faculty of Chemistry and Chemical Engineering of the *Babeș-Bolyai University* in Cluj-Napoca, Romania. Many professionals emerged from it, and the three younger authors of the present work have also started their carriers here. Although writers have exchanged ideas and discussed all topics of this book, work was distributed in agreement with individual strengths, experience, and competencies: Prof. Dr. Paul Șerban Agachi was in charge of chapters 1, 4, 12 and 13; Prof. Dr. Mircea Vasile Cristea shared his experience in chapters 2, 3, and 5–7; Assoc. Prof. Dr. Alexandra Csavdári wrote chapters 8 and 11, and coauthored with

young Eng. Botond Szilágyi chapters 9, 10, and 14 with assistance from Ş. Agachi. The list of problems and exercises in Chapter 15 is the result of a joint effort.

Finally, within the framework of this laborious enterprise, the authors gratefully acknowledge graduate students Maria Gherman, Abhilash Nair, Zsolt Tasnadi-Asztalos, László Zsolt Szabó and Hoa Pham Tai (engineers working in the *Group of Computer Aided Process Engineering*) for their dedicated and valuable help.

Cluj-Napoca, September 2016

*The authors*





# Contents

**Preface — v**

**Part I: Advanced Process Control — 1**

**1 Complex and nonconventional control systems — 3**

- 1.1 Cascade control systems — 3
- 1.1.1 Processes in series — 3
- 1.1.2 Processes in parallel — 10
- 1.2 Feedforward control systems — 15
- 1.3 Ratio control systems — 24
- 1.4 Inferential control systems — 27
- 1.5 Selective control systems — 28

**References — 30**

**2 Model predictive control — 32**

- 2.1 Introduction — 32
- 2.2 MPC history — 32
- 2.3 Basics of MPC control strategy — 34
- 2.4 Types of MPC process models — 42
- 2.4.1 Impulse and step response models — 43
- 2.4.2 State-space models — 49
- 2.4.3 Time series models — 49
- 2.5 Predictions for MPC — 50
- 2.6 Optimization for MPC — 60
- 2.7 MPC tuning — 64
- 2.8 MPC stability — 66
- 2.9 Nonlinear MPC — 68

**References — 71**

**3 Fuzzy control — 75**

- 3.1 Introduction — 75
- 3.2 Fuzzy sets — 75
- 3.3 Typical membership functions of the fuzzy sets — 77
- 3.4 Operations with fuzzy sets — 81
- 3.5 Fuzzy logic — 83

**References — 91**

**4 Optimal control systems — 92**

- 4.1 Steady-state optimal control — 92
  - 4.2 Dynamic optimal control of batch processes — 102
  - 4.3 Dynamic optimal control of continuous processes — 111
- References — 118

**5 Multivariable control — 119**

- 5.1 Introduction — 119
  - 5.2 Multiloop control — 120
    - 5.2.1 Interaction among control loops — 120
    - 5.2.2 Pairing the control loops — 126
    - 5.2.3 Tuning the multiloop controllers — 128
    - 5.2.4 Decoupling interaction for multiloop control — 129
  - 5.3 Multivariable centralized control — 133
- References — 134

**6 Plantwide control — 136**

- 6.1 Introduction — 136
  - 6.2 Premises of plantwide control — 137
  - 6.3 Designing the plantwide control strategy — 139
- References — 143

**7 Linear discrete systems and Z transform — 145**

- 7.1 Introduction — 145
  - 7.2 Discrete systems described by input-output relationship — 147
    - 7.2.1 Sampling the continuous signals — 147
    - 7.2.2 Reconstruction of the continuous signals from their discrete values — 153
    - 7.2.3 Analytical description of the discrete systems — 156
    - 7.2.4 Z transform — 160
    - 7.2.5 Z transform of several simple functions — 162
    - 7.2.6 Inverse of the Z transform — 163
    - 7.2.7 Z transfer function — 166
    - 7.2.8 Z transfer function of the sampled system — 168
    - 7.2.9 Z transfer function of the interconnected systems — 169
  - 7.3 Discrete PID controller — 171
  - 7.4 Other forms of the discrete controllers — 173
- References — 175

**Part II: Applied Process Engineering Control — 177****8 Reaction unit control — 179**

- 8.1 Introduction — 179
- 8.2 Basic concepts of ideal continuous and batch units — 179
- 8.3 Temperature control — 182
  - 8.3.1 Into thermal instability — 182
  - 8.3.2 Out of thermal instability — 184
  - 8.3.3 Temperature control in practice – continuous units — 188
  - 8.3.4 Temperature control in practice – batch units — 195
- 8.4 Pressure control — 200
- 8.5 Liquid level control — 202
- 8.6 pH control — 202
  - 8.6.1 pH and titration curves — 202
  - 8.6.2 pH regulator characteristics — 206
  - 8.6.3 Aspects of pH control in practice — 208
- 8.7 End-point detection and product-quality control — 210
  - 8.7.1 Some analyzer types — 210
  - 8.7.2 End-point detection reliability issues — 211
- 8.8 Control structure design for reaction units — 212
  - 8.8.1 Principles of control structure design — 212
  - 8.8.2 Control structure design for homogeneous ideal units — 218
  - 8.8.3 Control structure design for some heterogeneous units — 222
- References — 230

**9 Control of distillation processes — 233**

- 9.1 Economic constraints of distillation — 233
- 9.2 The recovery factor — 234
- 9.3 Lowering energy demand of distillation units — 237
- 9.4 General control of continuous distillation columns — 239
  - 9.4.1 Mass and energy balance imposed control issues — 239
  - 9.4.2 Control solutions — 249
- 9.5 Control issues of continuous distillation column dynamics — 254
- 9.6 Control issues of batch distillation columns — 259
- References — 260

**10 Control of absorption processes — 262**

- References — 269

**11 Control of extraction processes — 270**

References — 278

**12 Control of evaporation processes — 279**

References — 286

**13 Control of drying processes — 287**

13.1 Batch drying control — 289

13.1.1 Conventional batch drying control — 289

13.1.2 Advanced batch drying control — 292

13.2 Continuous adiabatic drying — 299

References — 302

**14 Control of crystallization processes — 303**

14.1 The process of crystallization — 303

14.2 Crystal size distribution control — 308

14.2.1 Model-free crystal size distribution control — 309

14.2.2 Model-based crystal size distribution control — 313

References — 316

**15 Problems and exercises — 318**

15.1 Advanced process control — 318

15.2 Applied process engineering control — 322

**Index — 325**

---

## Part I: Advanced Process Control

The first part of the book, entitled *Advanced Process Control*, refers, in the view of the authors, to all levels of complexity of automatic control systems beyond the simple feedback control loop, the backbone of the industrial process control.

Advanced control emerges from the human need to make groups of phenomena behave in a desired way, meeting productivity, quality, efficiency, safety, and environmentally friendly objectives. Usually, this assembly of physical, chemical, and biological phenomena shows complicated behavior and asks for complex control systems.

Within this context, the authors have treated first the nonconventional systems, such as cascade, feed-forward, ratio, and inferential control. The following chapters deal with the most important and significant control techniques for process industries: model predictive control, multivariable control including decoupling, fuzzy control, optimal control, plantwide control as well as the method of applying control algorithms in practice.

Model predictive control is undoubtedly the most industrially applied advanced control algorithm. Its roots are located in the practice-proven native applications, followed by the theoretical developments based on the open-loop optimal control and receding horizon control. Model predictive control makes possible the intelligent embedment of the process behavior described by the model while taking advantage of its prediction capability and coupling it with solving the constrained optimization to find the best control solution.

Fuzzy logic, with its extension to fuzzy control, is able to provide human consistency to the way control action is conceived. As human experts are able to solve complex control problems on the basis of the rule based approach, the fuzzy controllers may achieve the same task, being able to embed in the controller the human expertise in a systematic way.

Multivariable control strives to consider the complex system as a whole and accordingly aims to simultaneously achieve all desired control objectives, despite the possibly interacting effects of the multiple controls. Decentralized and centralized control are the two facets of the multivariable control approach featuring particular motivations from the practice applying perspective.

Optimal control is practically devoted to optimize one objective function such as the quality control criterion (e.g. integral absolute error IAE or integral of time weighted absolute error ITAE), the economic benefit, the return of investment, the duration of a batch process, or the conversion rate of a raw material. The optimal control is approached here both from the point of view of the process and of the control engineer. The steady-state control is practically an optimization of the process that uses the control elements as instruments. This approach is extended to batch and continuous processes as well.

Plantwide control is still an emerging field of research and development. It strives to find the harmony of coupling the regulatory layer of the control hierarchy with the supervisory/advanced control layers. However, a set of guidelines for the development of a systematic plantwide control approach are available and they are useful for the control system design.

The theoretical concepts of part I are the fundamentals for the control applications presented in part II.

# 1 Complex and nonconventional control systems

We consider all *automatic control systems* (ACS) which exceed as complexity, the traditional feedback control loops, as being complex and nonconventional or advanced. As Agachi and Cristea [1] have stated, the controllability of a process can be improved by using more complex systems as cascade, feed-forward, predictive, multivariable, adaptive, or optimal control systems. The first two, at which we add two special configurations of ratio and inferential control, are classified in this book as complex and nonconventional.

## 1.1 Cascade control systems

### 1.1.1 Processes in series

This type of control system is the next in complexity after the simple feedback control. The better performances of the cascade reside in capturing the effects of the disturbances inside the process and not at its end (Fig. 1.1). In this way, the controller can intervene earlier, being more efficient in cancelling the effects of the disturbance.

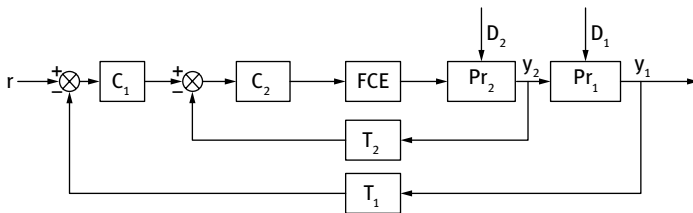


Fig. 1.1: Cascade ACS.

The process can be “decomposed” in two subprocesses, one faster in response to the intervention of the controller ( $Pr_2$ ) and the other slower ( $Pr_1$ ). As a matter of fact, the process is a whole, but the system captures an *intermediate* controlled variable by means of the inner loop transducer ( $T_2$ ).

The cascade consists of two nested loops, the “fast” inner one with the index<sub>2</sub> and the “slow” outer one with the index<sub>1</sub>. The temperature control of a continuous stirred tank reactor (CSTR) is given here as an example of applying the cascade control (Fig. 1.2). The inner loop is formed of the “faster” process, the heat transfer in the jacket ( $Pr_2$ ), the resistance thermometer in the jacket ( $T_2$ ), the controller  $C_2$ , and the final control element, which operates for both nested loops, manipulating the cooling or heating agent flow. The outer loop is formed of the “slower” process, the heat transfer in the reactor ( $Pr_1$ ), the resistance thermometer in the jacket ( $T_1$ ), the controller  $C_1$ , which gives the setpoint to be followed to  $C_2$ .



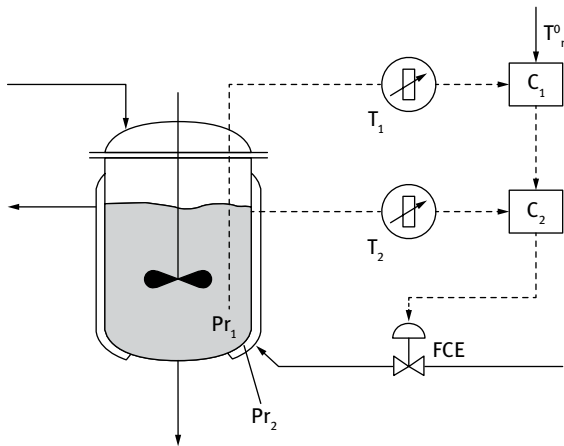


Fig. 1.2: Cascade ACS for the temperature of a non-isothermal CSTR.

The configuration is most efficient when the frequent disturbances are in the cooling/heating agent circuit because the ACS does not wait for the inner temperature of the reactor to change, but intervenes after the slightest change of temperature in the jacket.

Another example of applying successfully the cascade control is represented in Fig. 1.3.

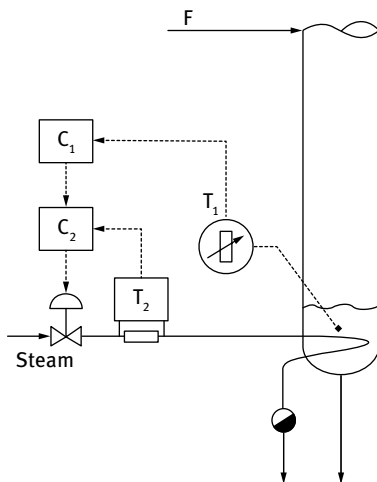


Fig. 1.3: Temperature cascade control in the bottom of a distillation column.

In this case, the steam flow to the reboiler is considered the inner process variable, subjected to the “faster” changes than the temperature in the bottom. The inner loop manages the steam flow, whereas the outer loop manages the temperature.

Generally, the cascade control is most efficient when disturbances of the type  $D_2$  are most frequent and important as magnitude. Because of the interaction between

$Pr_2$  and  $Pr_1$  is not only one way, from left to right interaction, the efficiency of the cascade should still be considered when type  $D_1$  disturbances occur. The efficiency is the highest when

$$3 \leq \frac{T_1}{T_2} \leq 10 \quad (1.1)$$

where  $T_1$  and  $T_2$  are the time constants of the two subprocesses.

### Structure of the controllers and tuning the controllers' parameters

The main goal of the cascade ACS is the tight control of the output  $y_1$  and less accurate control of the intermediate variable  $y_2$ . This is why the inner controller  $C_2$ , can have a simple P structure [2]. The external controller  $C_1$ , function of the slowness of the  $Pr_1$  can have a PI or PID structure (PID is designed for very slow processes, as heat or component transfer) [2]. In the case of PID structure of the external controller, the presence of the D element in the control algorithm of  $C_2$  has to be avoided since it amplifies any sudden small change of disturbance signals.

Tuning the controllers' parameters has to take into consideration the strong interaction between the two controllers,  $C_1$  imposing the setpoint value for  $C_2$ .

Experimentally, tuning has the following steps:

1. One of the experimental tuning methods is chosen (e.g. Ziegler-Nichols or Cohen-Coon) [3];
2. With both controllers on the manual mode of operation, the inner controller is fixed at the beginning to  $PB2_{max}$  (P structure); the chosen method is applied and  $PB2_{opt}$  is found;
3. With the  $C_2$  controller on automatic mode of operation, the inner loop becomes a dynamic element in the outer loop (Fig. 1.4);
4. The procedure of tuning the parameters of  $C_1$  is repeated for the “new” control loop with the inner loop as dynamic element of the outer loop.

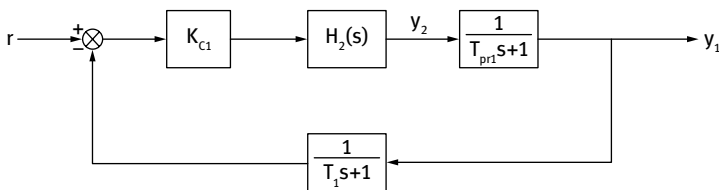


Fig. 1.4: The “new” structure of the cascade reduced to feedback control loop.

### Example 1.1

Tuning the parameters of the controllers in a series cascade system.

The control system in Fig. 1.5 has two controllers, one, the internal, with an already chosen gain,  $K_{c2} = 12$ , and the other, which has to be chosen properly.



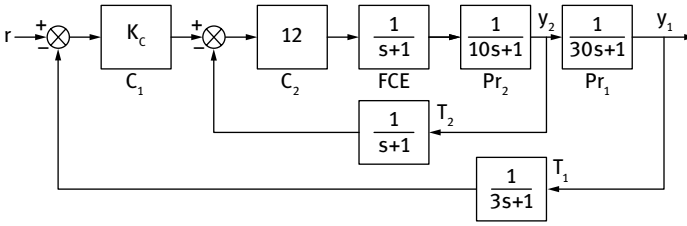


Fig. 1.5: Block diagram of the cascade control system corresponding to the CSTR in Fig. 1.2.

The obtained values of the  $C_1$  controller should be compared with those of a simple feedback loop controller, controlling the same process. Analysis has to be made to compare the values obtained in both cases and explain why the cascade is superior. All constants are expressed in minutes. The feedback control loop of  $y_1$  is presented in Fig. 1.6.

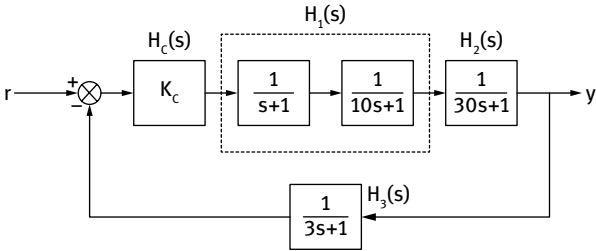


Fig. 1.6: Block diagram of the feedback control system controlling the heat transfer process in the CSTR from Fig. 1.2.

According to chapter 12.4 of [2], the crossover frequency condition for tuning the controller is

$$\frac{360}{2\pi} (-\tan^{-1}(1\omega) - \tan^{-1}(10\omega) - \tan^{-1}(30\omega) - \tan^{-1}(3\omega)) = -180^\circ \quad (1.2)$$

and

$$\omega_{osc} = 0.18 \text{ rad} \cdot \text{min}^{-1} \quad (1.3)$$

meaning one oscillation at every 35 minutes.

From relationship 11.3 [3],

$$K_{C1} \times \frac{1}{\sqrt{1 + 1 \times 0.18^2}} \cdot \frac{1}{\sqrt{1 + 10^2 \times 0.18^2}} \cdot \frac{1}{\sqrt{1 + 30^2 \times 0.18^2}} \cdot \frac{1}{\sqrt{1 + 3^2 \times 0.18^2}} = 0.5 \quad (1.4)$$

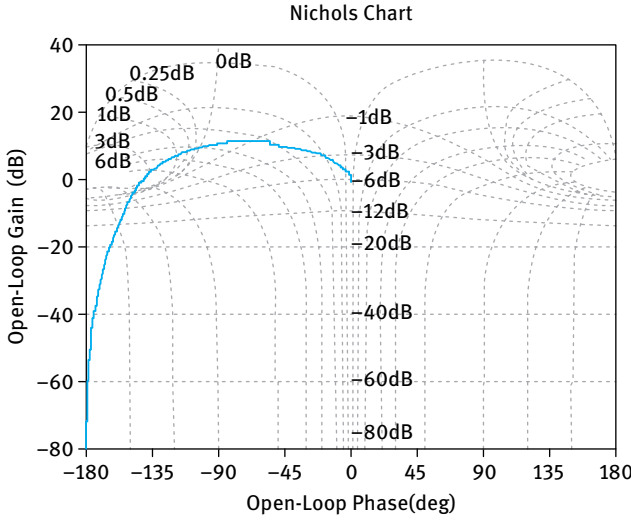
it results that

$$K_{Copt} = 6.5 \quad (1.5)$$

According to Fig. 1.4, the “simplified” control system has as  $T_{pr1} = 30$  min and  $T_{T1} = 3$  min and an additional element with the transfer function

$$H_1(s) = \frac{12 \cdot \frac{1}{s+1} \cdot \frac{1}{10s+1}}{1 + 12 \cdot \frac{1}{s+1} \cdot \frac{1}{10s+1} \cdot \frac{1}{s+1}} \tag{1.6}$$

To determine the parameters of the  $C_1$  controller, we may use the Black-Nichols diagrams (Fig. 1.7) [4].



**Fig. 1.7:** Black-Nichols diagrams used to determine the closed-loop gain and the phase angle when we know both open-loop values.

Using the open-loop phase angle

$$\varphi_{ol}(\omega) = 0 - \tan^{-1}(1 \cdot \omega) - \tan^{-1}(10 \cdot \omega) - \tan^{-1}(1 \cdot \omega) \tag{1.7}$$

and the open loop module

$$M_{ol}(\omega) = 12 \cdot \frac{1}{\sqrt{1+1 \cdot \omega^2}} \cdot \frac{1}{\sqrt{1+10^2 \cdot \omega^2}} \cdot \frac{1}{\sqrt{1+1 \cdot \omega^2}} \tag{1.8}$$

we obtain the following closed-loop values  $\varphi_{cl}$  and  $M_{cl}$  from the Black-Nichols plots:

$\omega$ [rad · min <sup>-1</sup> ]	$\varphi_{ol}$ [deg]	$M_{ol}$	$\varphi_{cl}$ [deg]	$M_{cl}$	$\varphi_{tot}$ [deg]
0.1	-56	8.40	-5	0.93	-93.0
0.2	-86	5.16	-11	0.97	-122.5
0.3	-104	3.48	-19	1.03	-144.6
0.4	-119	2.10	-28	1.17	-160.6
0.5	-12	1.87	-32	1.32	-174.5
0.6	-142	1.38	-40	1.60	-197.0

meaning that the crossover frequency for the cascade is  $\omega_{osc} = 0.53 \text{ rad} \cdot \text{min}^{-1}$  or an oscillation at every 12 minutes.

The alternative calculus can be done by processing directly the transfer function from eq. (1.5).

The frequency function is

$$\begin{aligned}
 H_1(j\omega) &= \frac{12 \cdot \frac{1}{j\omega + 1} \cdot \frac{1}{10j\omega + 1}}{1 + 12 \cdot \frac{1}{j\omega + 1} \cdot \frac{1}{10j\omega + 1} \cdot \frac{1}{j\omega + 1}} \\
 &= \frac{12 - 1440\omega - 3158\omega^2 + 5040\omega^3 - 40068\omega^4 - 37320\omega^6 - 12000\omega^8}{1 - 264\omega + 50\omega^2 - 1964\omega^3 - 1059\omega^4 - 82460\omega^5 - 19220\omega^6 + 6200\omega^7} + \\
 &\quad j \frac{-12 - 120\omega - 18371\omega^2 - 4584\omega^3 + 17532\omega^4 - 36120\omega^5 - 1200\omega^6 - 39720\omega^7}{1 - 264\omega + 50\omega^2 - 1964\omega^3 - 1059\omega^4 - 82460\omega^5 - 19220\omega^6 + 6200\omega^7}
 \end{aligned} \tag{1.9}$$

and from it  $M_{cl}(\omega)$ ,  $\varphi_{cl}(\omega)$  for different values of  $\omega$  is obtained directly, but in a much more laborious manner.

From equation 11.3,  $K_{c1}$  can be calculated:

$$K_{c1} \cdot M_{c1}(\omega = 0.53) \cdot \frac{1}{\sqrt{1 + 30^2 \times 0.53^2}} \cdot \frac{1}{\sqrt{1 + 3^2 \times 0.53^2}} = 0.5 \tag{1.10}$$

resulting  $K_{c1} = 10.6$

If we compare the results of the controllers parameter tuning in both cases (feedback versus cascade), we obtain

ACS feedback		ACS cascade	
$\omega_{osc} [\text{rad} \cdot \text{min}^{-1}]$	$K_c$	$\omega_{osc} [\text{rad} \cdot \text{min}^{-1}]$	$K_{c1}$
0.16	6.5	0.53	10.6

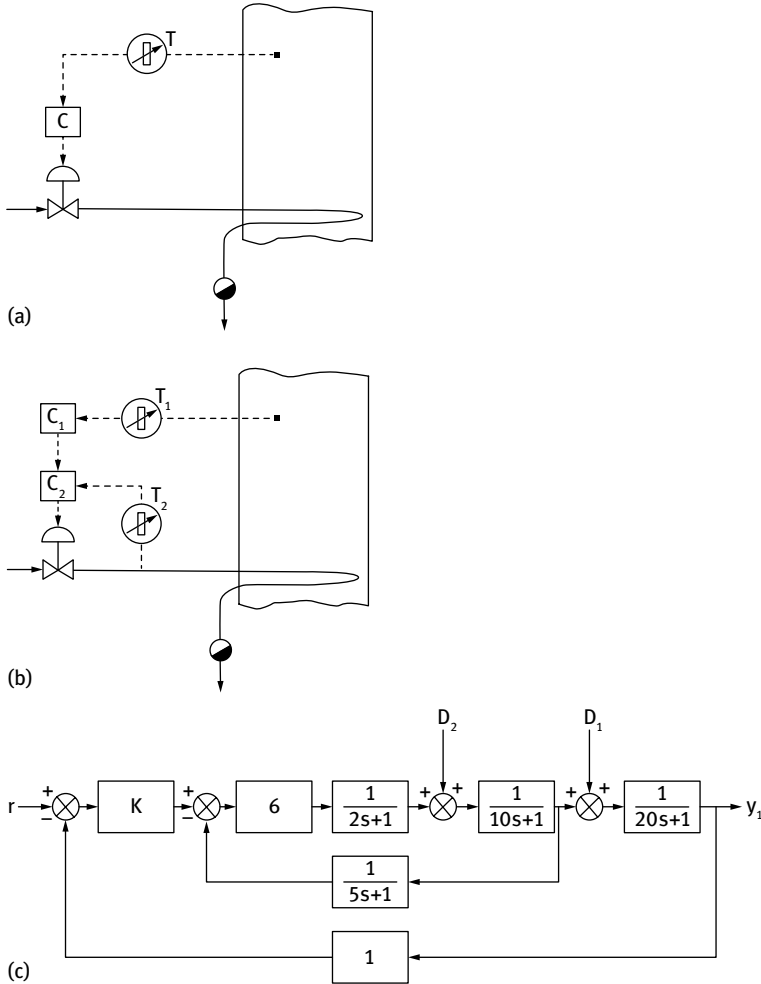
results that are in the favor of the cascade:

- the crossover frequency is pushed for the cascade in the range of higher frequencies, meaning that the chance of a “normal” disturbance (of lower frequency) to destabilize the loop is lower; thus, the loop is more stable;
- the response of the system is three times faster in its intervention to eliminate the disturbance effect ( $P_{osc\,casc} < P_{osc\,fb}$ ); the result is also a shorter settling time in the case of the cascade ACS;
- the gain of the controller is higher, with the result of a smaller offset of the controlled variable.

The practical and measurable economic results are visible from Example 1.2 [4].

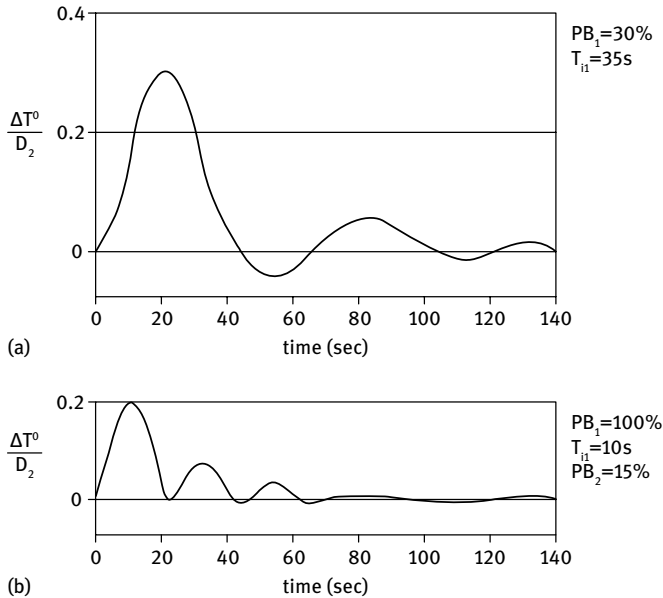
**Example 1.2**

The temperature control of a dryer is made both with a feedback control and cascade control systems (Fig. 1.8). Data of the process and elements of the ACS are given in Fig. 1.8 c.



**Fig. 1.8:** Temperature control of a dryer. (a) Feedback temperature control; (b) cascade temperature control; (c) block diagram of the control system with the data of the components of the system (time constants are measured in seconds).

Considering a load disturbance of 10% (the variation of the steam quality), the response of both systems is shown in Fig. 1.9 with the tuning of both controllers,  $PB = 30\%$  and  $T_i = 35\text{s}$  (a) and  $PB_1 = 100\%$ ,  $T_{i1} = 10\text{s}$  and  $PB_2 = 15\%$ .



**Fig. 1.9:** The response of feedback and cascade control systems at a load disturbance (type  $D_2$ ). (a) response of the cascade; (b) response of the feedback control.

One may observe the following facts:

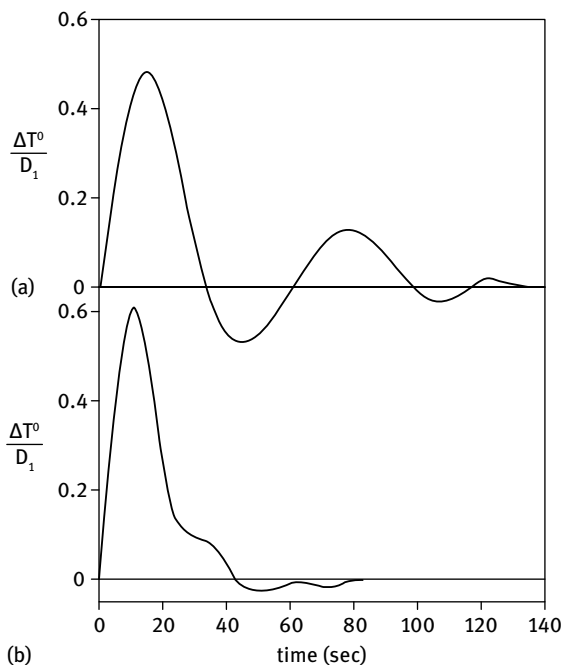
- the overshoot is smaller in the case of the cascade (20% versus 30%);
- the settling time is much smaller, decreasing from 150s to 90s; these two first facts show an increased stability of the system;
- the ratio of the areas under the response curve, signifying the energy consumption in both cases of control, is  $S_b = \frac{1}{4} S_a$ , implying a very important economy of consumed energy.

Even when the disturbance is not of the type  $D_2$ , but  $D_1$  (Fig. 1.10), the cascade is still very economically efficient ( $S_b = \frac{1}{2} S_a$ ). The disturbance could be a higher humidity of the material introduced in the dryer.

It should be noted that the performance of the cascade control in the latter case is worse since the ACS does not profit on the fast response of the internal loop, but only on the inverse interaction between the dryer and its heater.

### 1.1.2 Processes in parallel

The first who defined the concept was Luyben in 1973 [5]. Luyben referred to the systems for which the manipulated variable influences two processes in parallel,



**Fig. 1.10:** The response of the temperature control systems (feedback and cascade) in the case of a disturbance of type  $D_1$ . (a) feedback control; (b) cascade control.

with two output variables (Fig. 1.11). One example is the influence of both the top tray temperature and the composition at the top of a distillation column, *via* its reflux. The primary loop of the cascade is that controlling the overhead composition and the secondary one is controlling the temperature of the top tray. The manipulative variable (reflux flow) affects overhead composition and tray temperature through two parallel process transfer functions,  $H_{pr1}$  and  $H_{pr2}$ . According to the original notations, these correspond to  $G_M$  and  $G_S$  respectively.

Another example was given by Rao *et al.* [8] for a liquefied petroleum gas (GPL) splitter model; its the structure is given in Fig. 1.12.

$Pr_1$  and  $Pr_2$  are the two processes in parallel, the primary and the secondary one; the secondary loop is much faster than the primary one.  $H_{C1}$  and  $H_{C2}$  are the controllers of the loops.  $H_{D1}$  and  $H_{D2}$  are the transfer functions of the channels of the disturbance in the process.

The cascade control is especially beneficial when the secondary loop is much faster than the primary one, as in the series cascade systems.

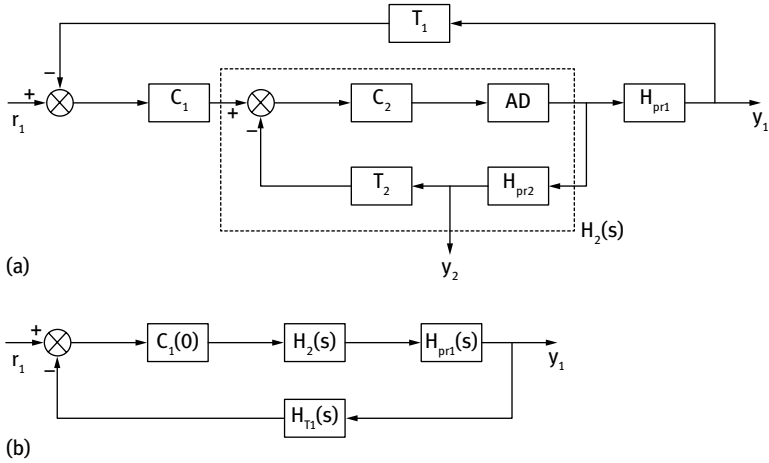
Tuning the controllers starts with the  $C_2$  controller by using the inner open-loop transfer function.





with the consequent  $\varphi_{ol}(\omega)$  and  $M_{ol}(\omega)$  from [4]. Because of simplicity reasons, the value of the actuator transfer function  $H_{AD}(s)$  as well as those of the transducers' transfer functions  $H_{T1}(s)$  and  $H_{T2}(s)$  in Fig. 1.12. are considered to be equal to 1.

The controller  $C_1$  is tuned, considering that function  $H_{pr2}(s)$  is placed on the reaction of loop 2 (Fig. 1.13).



**Fig. 1.13:** The block diagram of the tuning of parallel cascade; (a) the “restructured” control system; (b) The equivalent system for tuning the  $C_1$  controller.

The transfer function of the external loop is:

$$H_1(s) = \frac{H_{C1} \cdot \frac{H_{C2}(s) \cdot H_{AD}(s) \cdot H_{pr2}(s) \cdot H_{T2}(s)}{1 + H_{C2}(s) \cdot H_{AD}(s) \cdot H_{pr2}(s) \cdot H_{T2}(s)} \cdot H_{pr1}(s)}{1 + H_{C1} \cdot \frac{H_{C2}(s) \cdot H_{AD}(s) \cdot H_{pr2}(s) \cdot H_{T2}(s)}{1 + H_{C2}(s) \cdot H_{AD}(s) \cdot H_{pr2}(s) \cdot H_{T2}(s)} \cdot H_{pr1}(s) \cdot H_{T2}(s)} \quad (1.12)$$

**Example 1.3**



Tuning the parameters in a parallel cascade system.

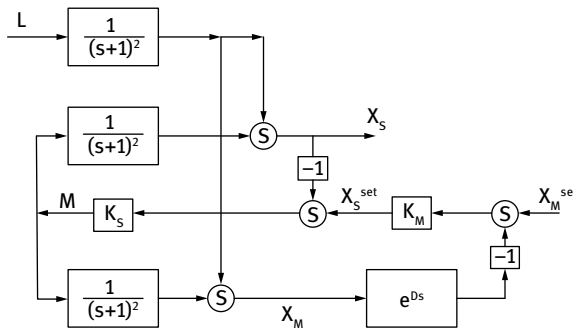
Let us take the example given by Luyben in [5]. He defined the master loop (loop 1 in previous notation) that of composition and slave loop (loop 2 in previous notation) that of temperature with the consequent notations. Because of clarity reasons, we decided to keep Luyben’s notations.

Considering the system in Fig. 1.11, Luyben particularized the transfer functions of the process as (Fig. 1.14)

$$H_{pr1}(s) = H_{prM}(s) = \frac{1}{(T_M s + 1)^2} \text{ and } H_{pr2}(s) = H_{prS}(s) = \frac{1}{(T_S s + 1)^2} \quad (1.13)$$

by giving only one value for  $T_S = 1$  min and several values for  $T_M = 0.5, 1, 2, 4$  min, respectively;  $T_S$  and  $T_M$  stand for the time constants of the process transfer functions of slave and master loops, respectively.

For the master loop, additional dead time was considered  $D = 0.5, 1, 2$  min to notice the difference in calculating the gain of the master controller,  $G_{CM}$ . The controllers are considered proportional and  $G_{CS}$  was set to 1 to obtain a damping ratio of 0.707.



**Fig. 1.14:** Particular block diagram for the alternatives feedback and parallel cascade system in Fig. 1.11. In the figure,  $T_S = T_M = 1$  min. Meaning of notations is:  $L$  – load disturbance;  $D$  – deadtime;  $M$  – manipulative variable;  $X_M^{set}$  and  $X_S^{set}$  – setpoints of master and slave controllers;  $K_M$  and  $K_S$  – gains of master and slave controllers.

The tuning of the master controller is given in Tab. 1.1. Luyben used the root locus method for tuning [6, 7]. But other tuning method, as presented in [3], gives the same results.

**Tab. 1.1:** The results of the master P controller tuning for different ACSs (feedback, series, and parallel cascade) for different time constants and dead times.  $K_S = 1$  and the damping coefficient of the slave loop is 0.7.

T [min]	$K_M$		$K_M$		Additional dead time $\tau$ [min]	$K_M$ ACS	
	ACS feedback		ACS cascade			feedback	cascade parallel
	series	parallel	series	parallel			
0.5	0.7	3	0.6	2.5	0.5	2.500	2.500
1.0	0.7	3	1	2	1	1.350	1.560
2.0	0.7	3	2	4	2	0.862	1.136
4.0	0.7	3	3	13	T = 1 min		

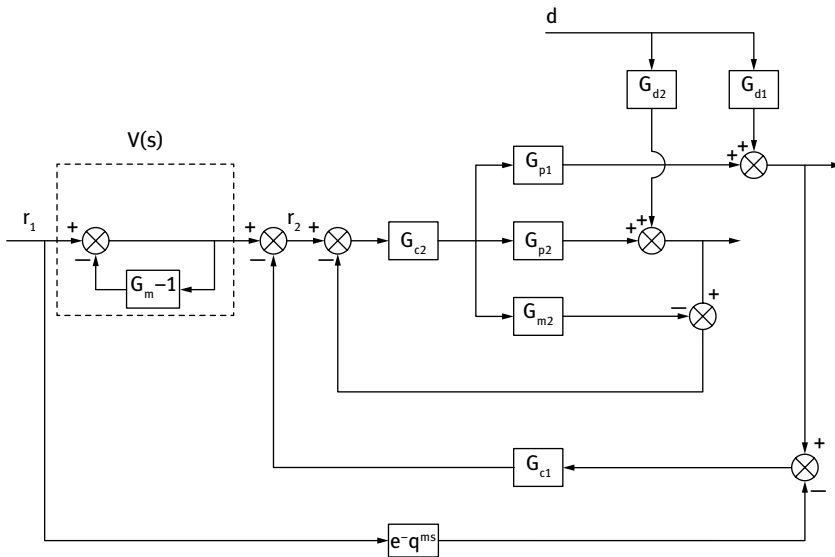
The supplementary dead time can belong to a gas chromatograph in the composition loop, meaning that the transducer  $T_1$  ( $T_M$  in Luyben’s version) has the transfer function  $e^{-Ds}$ , where  $D$  takes different values (from 0.5, 1 or 2 minutes) to see the influence of this element on the tuned values of the master controller: the larger  $D$  is, the smaller the controller gain is, due to the reduced controllability of the process plus transducer.

More recently there have been attempts to revise the tuning techniques by proposing procedures of autotuning that are extensively discussed in [8].

The results reported in [8–11] are encouraging the use of parallel cascade control, with amendments to the original thinking of Luyben as adding a dead-time compensator for the dead time existing in the outer loop (Fig. 1.1 and 1.15). The proposed control structure [11] uses the outer loop controller  $G_{c1}$  in the feedback path. As a result, although  $G_{c1}$  is meant to reject the load disturbance, it contributes to the stabilization of the process in the outer loop.

Actually, as Luyben affirms, one cannot appreciate truly the performances of different types of cascade control but only in comparison with the non-cascade situation – see Fig.1.16. It is obvious that the performance of the cascade control loop is better (shorter settling time and smaller steady-state error). In the cascade system, the manipulative variable begins immediately its intervention when load disturbance occurs, whereas in the feedback system action starts later with the time elapse of the dead time.

If cascade control is not efficient enough, a more complex control system should be used.

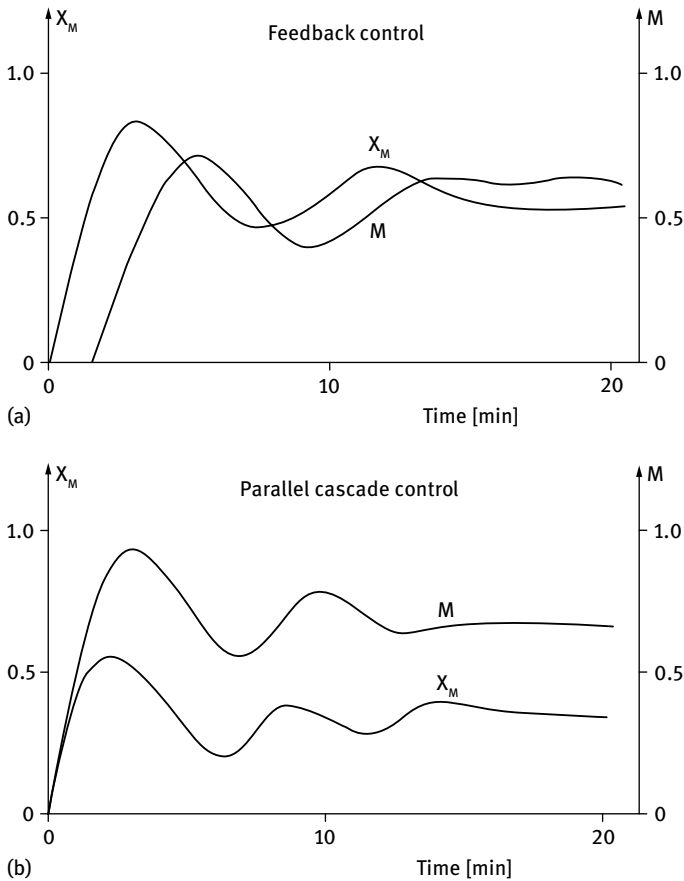


**Fig. 1.15:** Parallel cascade with dead time compensator [11].  $G_{p1}$  and  $G_{p2}$  stand for transfer functions of the primary and secondary process;  $G_{m1}$  and  $G_{m2}$  for those of the primary and secondary process models, and  $G_{d1}$  and  $G_{d2}$  for the disturbance paths in the primary and secondary loops, respectively.

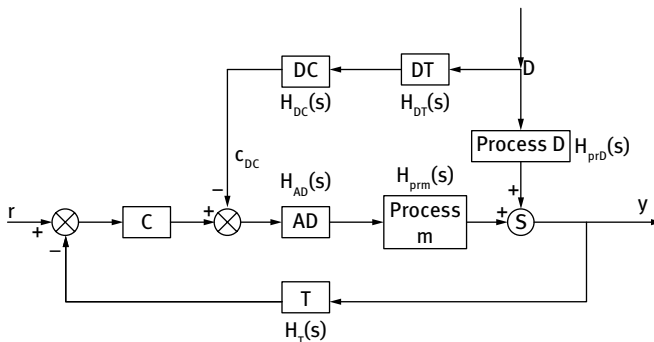
## 1.2 Feedforward control systems

In [1], when the controllability of the process is discussed, for a poor controllability, a superior organization of the control system is needed: either a feedforward or a model predictive structure. In this section, we propose for examination the feedforward control.

Feedforward control (Fig. 1.17) is one of the most advanced control structures used on a large scale in industry.



**Fig. 1.16:** The comparative results of controlling the output of a 2 min dead time for the primary process: (a) feedback control; (b) parallel cascade control.  $X_M$  stands for the controller output and  $M$  for the manipulative variable [5].



**Fig. 1.17:** Feedforward ACS block diagram.

It captures the most important disturbance of the process, measures and processes it in a disturbance controller (*DC*) in such a way that when the disturbance action is propagated through process *D* path to the exit, it is totally annulled by the counteraction on the path “*DT-DC-AD-process m*” (Fig. 1.18).

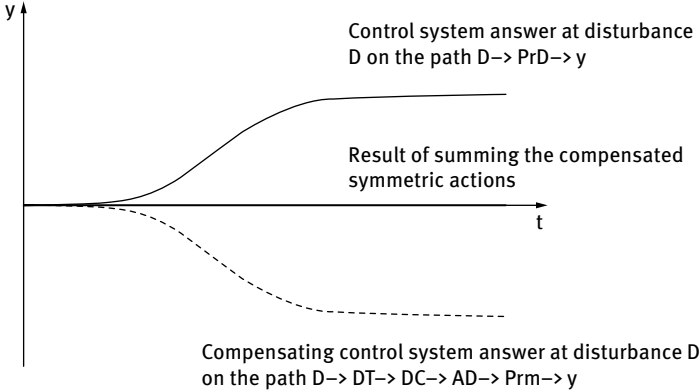


Fig. 1.18: The way the disturbance effect in the process is annulled by the action of the DC.

The *DC* has to be so “intelligent” that its action should be perfectly coordinated in time and magnitude with the action of the main disturbance *D* in the process to completely counteract it. Thus, it has to contain all the information related to the gains and delays in all elements involved: process *D*, process *m*, *DT*, *AD*. As observed from Fig. 1.16, the result of the feedforward control is the sum of the symmetric actions on both process *D* and “*DT-DC-AD-process m*” paths, meaning that the disturbance does not disturb actually at all the process. To obtain such a result, one must know the dynamic behavior of all elements, meaning that their models are known. Because *D* is not the only disturbance on the process, but the main one, the effect of the other disturbances is rejected through a regular feedback ACS containing the regular controller (*C*), transducer (*T*), and actuating device (*AD*).

The synthesis of the *DC* is realized based on the desired behavior of the system expressed mathematically by (1.14).

$$\frac{\Delta y(t)}{\Delta D(t)} = 0 \text{ or } \frac{Y(s)}{D(s)} = 0 \tag{1.14}$$

This means that the effect of the disturbance on the process output controlled variable is 0.

Thus,

$$Y(s) = [C(s) - CD(s)] \cdot H_{AD}(s) \cdot H_{pr m}(s) + D(s) \cdot H_{pr D}(s) \tag{1.15}$$

with

$$C(s) = H_C(s)[R(s) - H_T(s) \cdot Y(s)] \text{ and } CD(s) = H_{DC}(s) \cdot H_{DT}(s) \cdot D(s) \tag{1.16}$$

Equation (1.14) becomes

$$Y(s) = R(s) \frac{C(s) \cdot H_{AD}(s) \cdot H_{prm}(s)}{1 + H_T(s) \cdot H_C(s) \cdot H_{AD}(s) \cdot H_{prm}(s)} + D(s) \frac{H_{prD}(s) - H_{DC}(s) \cdot H_{DT}(s) \cdot H_{AD}(s) \cdot H_{prm}(s)}{1 + H_T(s) \cdot H_C(s) \cdot H_{AD}(s) \cdot H_{prm}(s)} \quad (1.17)$$

To have no effect of the disturbance on the process, the numerator of the second term in eq. (1.17) has to be 0,

$$H_{prD}(s) - H_{DC}(s) \cdot H_{DT}(s) \cdot H_{AD}(s) \cdot H_{prm}(s) = 0 \quad (1.18)$$

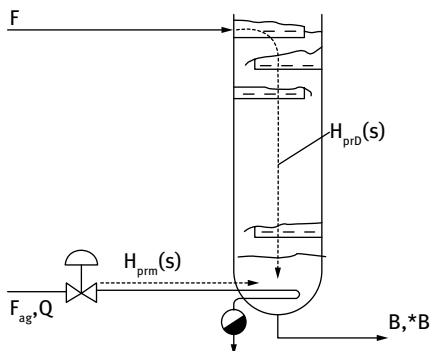
and from which the DC transfer function can be calculated:

$$H_{DC}(s) = \frac{H_{prD}(s)}{H_{DT}(s) \cdot H_{AD}(s) \cdot H_{prm}(s)} \quad (1.19)$$

It is remarkable that the *DC* contains all the information needed to compensate perfectly the effect of the disturbance in the process; all gains and delays on the propagation paths are embedded in its transfer function.

### **i** Example 1.3

Consider a distillation column subjected to the main disturbance, the feed flow. This is very often true in a train of columns where the feed flows cannot be controlled. The controlled variable is the light component concentration (or the temperature) in the bottom of the column. Figure 1.19 describes the technological process of the column.



**Fig. 1.19:** Distillation process subjected to feedforward control.

According to the mathematical model of the binary distillation column described in [12], each tray can be described as having a capacitive behavior. Since we have  $k$  trays on the concentration section of the column, the transfer function of transfer path process  $D$  is

$$H_{prD}(s) = \frac{K_{eD}e^{-\tau_{eD}s}}{1 + T_{eD}s} = \sum_1^k \frac{K_i}{1 + T_i s} \quad (1.20)$$

where  $K_{eD}$ ,  $T_{eD}$ , and  $\tau_{eD}$  are the equivalent gain, time constant, and dead time, respectively, on the transfer path  $F \rightarrow x_B$ ;  $K_i$  and  $T_i$  are the gain and the time constant for the tray  $i$  (series of trays each with capacitive behavior), respectively.

On the path  $m \rightarrow x_B$ , the process has practically only one element, the bottom of the column (including the reboiler or the steam serpentine). It exhibits a capacitive behavior and has a transfer function of

$$H_{pr m}(s) = \frac{K_B}{1 + T_B s} \quad (1.21)$$

where  $K_B$  and  $T_B$  are the gain and the time constant of the bottom, respectively.

The other elements of the feedforward control are the disturbance transducer (flow transducer), the actuating device (control steam valve) and the additional transducer (temperature transducer). All their delays (in comparison with the delays in the column) are insignificant; therefore, one can approximate their behavior as being proportional with the gains  $K_{DT}$ ,  $K_{AD}$  and  $K_T$ , respectively.

Thus, the transfer function of the feedforward controller is for this case study

$$H_{DC}(s) = \frac{\frac{K_{eD}e^{-\tau_{eD}s}}{1 + T_{eD}s}}{K_{DT} \cdot K_{AD} \cdot \frac{K_B}{1 + T_B s}} = \frac{K_{eD}e^{-\tau_{eD}s} \cdot (1 + T_B s)}{K_{DT} \cdot K_{AD} \cdot K_B \cdot (1 + T_{eD}s)} \quad (1.22)$$

and the corresponding feedforward control system is presented in Fig. 1.20.

Equation (1.22) can be very easily transformed in a practical feedforward control algorithm by using the Z-transform [13]

$$c_{DC}(nT^*) = \frac{T_{eD}}{\beta} \cdot c_{DC}[(n-1)T^*] + \frac{K\alpha}{\beta} \cdot x_{DT}[(n-2)T^*] - \frac{KT_B}{\beta} \cdot x_{DT}[(n-3)T^*] \quad (1.23)$$

where  $T^*$  is the sampling time chosen so as  $\tau_{eD} = kT^*$ ;  $K = K_{DT} \cdot K_{AD} \cdot K_B$ ;  $\alpha = T_B + T^*$  and  $\beta = T_{eD} + T^*$ ;  $x_{DT}$  is the signal of the disturbance transmitter at different sampling times.

In some works of Niesenfeld [14, 15], the authors mentioned an experiment worth to be cited: they measured the performances of a feedforward control system applied to a distillation column, exactly in the format mentioned in Example 1.3. The results are exposed in Tab. 1.2.



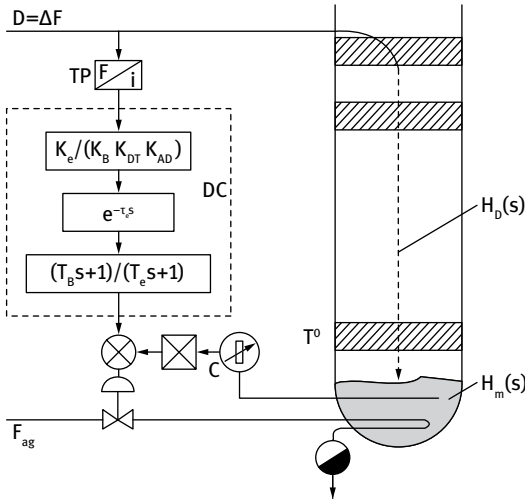


Fig. 1.20: Feedforward control system of the distillation column from Fig. 1.17.

Tab. 1.2: The results of the feedforward control of the distillation column in [14, 15].

Disturbance $\Delta F$ [%] from $F_{max}$	ACS structure	Oscillations damping time [min]
60–70%	PID	10
	feedforward	no oscillations
70–80%	PID	maintained oscillations
	feedforward	3

It is worth mentioning that due to the increased stability of the process, a 15% saving in energy consumption was estimated.

Another experiment carried out in 1987 at the *Brazi Refinery* in Romania by a group of professors from the *Petroleum and Gas Institute*, Ploiești, Romania [16], reported 20% energy savings at a benzene – toluene column with 90 trays and a 2 m diameter. The usually reported consumption of steam was in the average of 4 t/h.

To give an idea on the significance of a 20% savings, let us make a cost estimation. At the average price of 70 USD/Gcal and a total operating timespan of 7200 hours/year, with a latent heat of 525 kcal/kg, the total amount of consumption for the mentioned column would be

$$Q_{year} = 7200 \text{ h} \cdot 4000 \text{ kg/h} \cdot 525 \text{ kcal/kg} = 15,120 \text{ Gcal/year}$$

The monetary savings per year are of  $MS = 0.2 \cdot 15,120 \text{ Gcal/year} \cdot 70 \text{ USD/Gcal} = 211,680 \text{ USD}$

The investment for such an industrial application can be estimated at a total of 123,000 USD by including:

- human resources (2 process engineers to work on process models and computer algorithms) ~ 8,000 USD/ month = 96,000 USD/year;
- a process computer ~10,000 USD;

- control equipment (controller, transducers, control valve) ~ 15,000 USD;
- electric and air fittings, rack, manpower ~2,000 USD.

This means that the investment can be recovered in about 7 months.

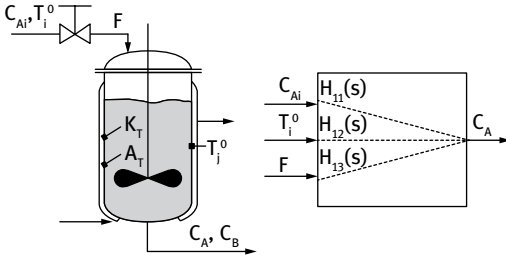
Furthermore, a process computer, the major infrastructure additional investment, can manage not only one but several applications.

There are situations when more important disturbances are influencing significantly the process and they have to be considered in the feedforward control [17–21], together with more advanced techniques as the model predictive control.

**Example 1.4**



Consider a CSTR (Fig. 1.21) in which a first-order exothermic reaction  $A \rightarrow B$ , takes place.  $k$  and  $\Delta H_r$  are the rate constant and heat of reaction, respectively. The process is disturbed essentially by  $C_{Ai}$  and  $T_i^\circ$ .



**Fig. 1.21:** CSTR subjected to two measurable main disturbances.  $T_j^\circ$  is the jacket temperature;  $K_T$  and  $A_T$  are the heat transfer characteristics.

The output controlled variable is  $C_A$  and the manipulative variable is  $F$ . The difference from the previous example of feedforward control is that there are two important disturbances instead of one, which have as consequence the existence of two disturbance controllers,  $DC_1$  for  $C_{Ai}$  and  $DC_2$  for  $T_i^\circ$ . The two process transfer paths for the two disturbances are described by the following equations:

$$V \frac{dC_A}{dt} = F(C_{Ai} - C_A) - VkC_A \tag{1.24}$$

$$V\rho c_p \frac{dT^\circ}{dt} = F\rho c_p(T_i^\circ - T^\circ) - \Delta H_r V k C_A - K_T A_T (T^\circ - T_j^\circ) \tag{1.25}$$

$$k = k_0 e^{-\frac{E}{RT^\circ}} \tag{1.26}$$

The linearized equations, by following the procedure from [22] to obtain linear transfer functions, are presented in the equations (1.27)–(1.28):

$$\frac{dC_A}{dt} = a_{11}C_A + a_{12}T^\circ + a_{13}C_{Ai} + a_{15}F \tag{1.27}$$

$$\frac{dT^\circ}{dt} = a_{21}C_A + a_{22}T^\circ + a_{24}T_i^\circ + a_{25}F + a_{26}T_j^\circ \quad (1.28)$$

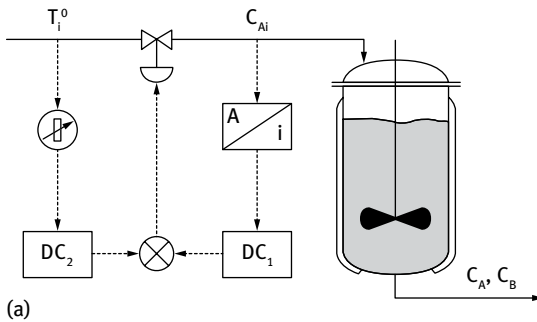
In these orders 1–6 of the correlation coefficients' indexes are allocated to the variables  $C_A$ ,  $T^\circ$ ,  $C_{Ai}$ ,  $T_i^\circ$ ,  $F$ , and  $T_j^\circ$ , respectively.

$$a_{11} = -\frac{F_n}{V} - k_n; a_{12} = -\frac{C_{An}Ek_n}{RT_n^{\circ 2}}; a_{13} = \frac{F_n}{V}; a_{15} = \frac{C_{Ai} - C_{An}}{V}$$

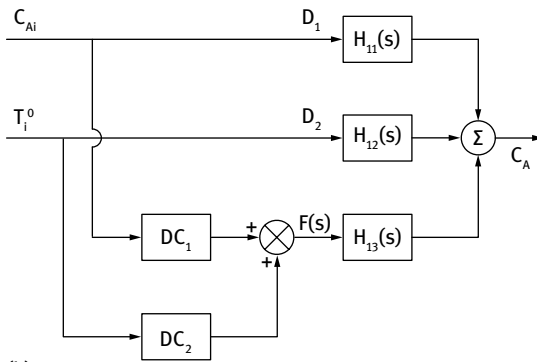
$$a_{21} = -\frac{\Delta H_r k_n E}{RT_n^{\circ 2} \rho c_p} - \frac{F_n}{V} - \frac{K_T A_T}{V \rho c_p}; a_{24} = \frac{F_n}{V}; a_{25} = \frac{T_i^\circ - T_n^\circ}{V}; a_{26} = \frac{K_T A_T}{V \rho c_p}$$

The variables with the index  $n$  are at steady state and have steady-state values. Density, specific heat, and volume are considered constant.

The block diagram of the feedforward control system is presented in Fig. 1.22.



(a)



(b)

**Fig. 1.22:** Block diagram of the CSTR feedforward control of both main disturbances:  $C_{Ai}$  and  $T_i^\circ$ . (a) feed forward scheme without feedback control; (b) block diagram of the feed forward CSTR control system with two identified disturbances.

Thus, the transfer functions which describe the system are:

$$C_A(s) = H_{11}(s)C_{Ai}(s) + H_{12}(s)T_i(s) + H_{13}(s)F(s) + H_{14}(s)T_j(s) \quad (1.29)$$

$$T(s) = H_{21}(s)C_{Ai}(s) + H_{22}(s)T_i(s) + H_{23}(s)F(s) + H_{24}(s)T_j(s) \quad (1.30)$$

and from eqs. (1.27)–(1.28) the following is obtained:

$$sC_A(s) = a_{11}C_A(s) + a_{12}T(s) + a_{13}(s)C_{Ai}(s) + a_{15}(s)F(s) \quad (1.31)$$

$$sT(s) = a_{21}C_A(s) + a_{22}T(s) + a_{24}T_i(s) + a_{23}F(s) + a_{24}T_j(s) \quad (1.32)$$

Finally

$$C_A(s) =$$

$$\begin{aligned} & \frac{a_{13}(s-a_{22})}{s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21}} C_{Ai}(s) + \frac{a_{12}a_{24}}{s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21}} T_i(s) + \\ & \frac{a_{12}a_{23} + a_{15}(s-a_{22})}{s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21}} F(s) + \frac{a_{12}a_{26}}{s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21}} T_j(s) \end{aligned} \quad (1.33)$$

and

$$T(s) =$$

$$\begin{aligned} & \frac{a_{13}a_{21}}{s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21}} C_{Ai}(s) + \frac{a_{24}(s - a_{11})}{s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21}} T_i(s) + \\ & \frac{a_{15}a_{21} + a_{25}(s - a_{11})}{s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21}} F(s) + \frac{a_{26}(s - a_{11})}{s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21}} T_j(s) \end{aligned} \quad (1.34)$$

The condition for invariability of the output concentration is  $C_A(s) = 0$ , thus

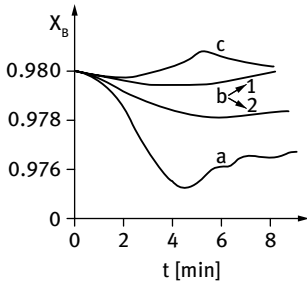
$$C_A(s) = H_{11}(s)C_{Ai}(s) + H_{12}(s)T_i(s) + H_{13}(s)F(s) = 0 \quad (1.35)$$

and then, from (1.19):

$$H_{DC1}(s) = \frac{H_{11}(s)}{H_{13}(s)} \text{ and } H_{DC2}(s) = \frac{H_{12}(s)}{H_{13}(s)} \text{ or}$$

$$H_{DC1}(s) = \frac{a_{13}(s - a_{22})}{a_{12}a_{23} + a_{15}(s - a_{22})} \quad \text{and} \quad H_{DC2}(s) = \frac{a_{12}a_{24}}{a_{12}a_{23} + a_{15}(s - a_{22})} \quad (1.36)$$

The efficiency of the feedforward control at a change of input flow on the bottom concentration is shown in Fig. 1.23.



**Fig. 1.23:** The efficiency of the feedforward control at an input change in the molar feed flow of a distillation column. (a) feedback control; (b) feedforward control only (1 – for process 1; 2 – for process 2); (c) combined feedforward and feedback.

One important drawback is that linearization which decreases the degree of accuracy. Research has been done in the direction of nonlinear feedforward control.

To summarize, the feedforward control is a very successful control strategy in the case of poor controllability processes. However, it has some drawbacks:

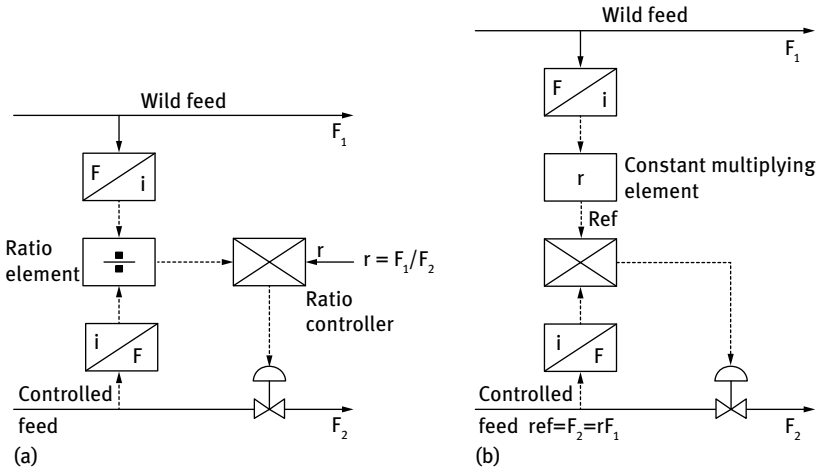
- the main disturbances should be measurable and captured;
- it is quite difficult to elaborate good dynamic models for the entire process; in addition, in case of linear feedforward control, linearization reduces the accuracy of the model;
- high level of competency is requested from process engineers.

For a long time, the interest in development of feedforward control was somehow tempered and determined especially by the fact in many situations the ideal algorithms elaborated from process models could not be implemented in practice. Recent developments including internal model control and model predictive control, or “neuromorphic” control by using artificial neural networks, and robust nonlinear controllers are nowadays presented in the literature [23–25].

### 1.3 Ratio control systems

In many situations of process engineering, one has to keep constant the ratio between two flows of material, especially when one of the flows cannot be controlled from the beginning: this is the situation of keeping constant the ratio between reflux and distillate in a distillation column, or the ratio of flows of the reactants in a CSTR, or the ratio of flows of primary and secondary solvents in a liquid-liquid extraction process, or the air-gas ratio in a burner. The general block diagram of the ratio control systems is given in Fig. 1.24.

The noncontrolled, independent flow is sometimes named “wild feed” and the controlled, dependent flow, following closely the noncontrolled one, is called “controlled feed”. In version (a) of the ratio control system, the two flows are measured with flow transducers and the values of their signal are divided inside a ratio relay; the controller becomes a ratio controller, with the setpoint equal to the desired ratio between the two flows. The controller acts by modifying the controlled feed stream in such a way that the ratio between the wild and the controlled streams is kept at the setpoint.



**Fig. 1.24:** Ratio control systems in two possible versions: (a) ratio control based on variable ratio error; (b) ratio control based on error of the flow subjected to ratio constraint.

$$r = \frac{F_B}{F_A} = r_{set} \tag{1.37}$$

In this case, the control system reacts slowly when the value of the controlled stream is relatively large and acts faster when the value of the controlled stream is relatively small.

In version (b) of the ratio control system, the wild feed is multiplied with a constant  $r$  in the multiplication element and the result is fed as the setpoint to the controller; the result is that the controlled stream is kept by the flow control system to the setpoint (variable) value that varies with flow  $A$ .

$$F_B = F_{Bset} = rF_A \tag{1.38}$$

In this version, the action of the control system is independent from the relative amount of the controlled feed stream.

**Example 1.5**

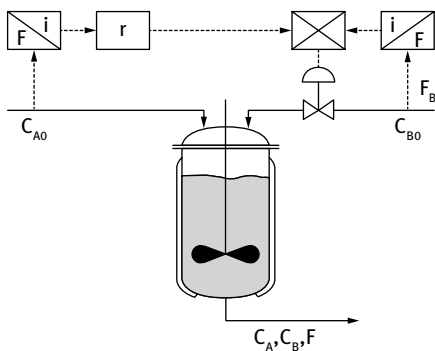
One good example of the application of ratio control systems is that of a reversible reaction  $aA + bB \rightleftharpoons cC$  taking place in a CSTR (see Fig. 1.25.); to maximize the pro-



duction of  $C$ , one has to move the equilibrium of the reaction to the right, keeping an excess of  $B$  over  $A$ . It means that whenever the flow of  $A$  increases, the ratio control system automatically increases the flow of  $B$ , and keeps constant a certain excess ratio value (Fig. 1.22). In manual control operation mode of a CSTR, the highest excess of  $B$  over  $A$  is often used by fixing the controlled flow at maximum. This is feasible, yet the disadvantages are the following:

1. In the case of economically natural recycling of  $B$ , stage following the CSTR is the separation whose costs depends of the molar fraction of  $B$  in the mixture  $B$ – $C$ . The higher the fraction of  $B$ , the higher the separation costs;
2. In the case of not recycling  $B$  due to its very low price, either there is no separation stage and  $B$  severely impurifies  $C$  because of its high quantity, or there exists a separation stage with the adjacent costs so that a solution has to be found for depositing and further using  $B$ .

The advantages of the automatic ratio control are thus obvious.



**Fig. 1.25:** CSTR with ratio control.  $F_A$  and  $F_B$  are the volumetric reactants flow rates;  $C_{Ai}$  and  $C_{Bi}$  are the input molar concentrations of the reactants.

From chemical kinetics information [26], the excess factor that has to be respected is

$$\gamma = \frac{aC_{B0}}{bC_{A0}} \quad (1.39)$$

where  $C_{A0}$  and  $C_{B0}$  are the initial concentrations of both reactants after mixing their flows:

$$C_{A0} = \frac{F_A \cdot C_{Ai}}{F_A + F_B} \text{ and } C_{B0} = \frac{F_B \cdot C_{Bi}}{F_A + F_B} \quad (1.40)$$

Keeping in mind that the ratio control system operates with volumetric flows and keeps constant their ratio, the value of setpoint ratio  $r$  calculated from eqs. (1.38)–(1.40) is:

$$r = \frac{F_B}{F_A} = \gamma \frac{bC_{Ai}}{aC_{Bi}} \quad (1.41)$$

By supposing that in the CSTR the following reaction takes place:  $FeCl_3 + 3H_2O \Leftrightarrow Fe(OH)_3 + 3HCl$ , with the mass concentration of the ferric chloride solution of 20%, by knowing the molecular weights of feed solution chemicals  $FeCl_3$  and  $H_2O$  (162,2 and 18 kg/kmol respectively), their densities (1135 and 1000 kg/m<sup>3</sup> respectively), and finally by using equations (1.38)–(1.40), the molar concentrations of feed solutions can be calculated. These are 36.77 kmol/m<sup>3</sup> and 55.6 kmol/m<sup>3</sup>. The values yield  $\gamma = 0.22$ ,  $r = 0.43$ , and can be fixed on the ratio relay.

The approximate variation of the flows in the ratio control system is given in Fig. 1.23. The tuning of the controller parameters is based on the process response curve method [4] to  $K_{p\ opt} = 0.35 K_{pr} \frac{T_{pr}}{\tau_{pr}}$  and  $T_{i\ opt} = 1.2\tau_{pr}$ .

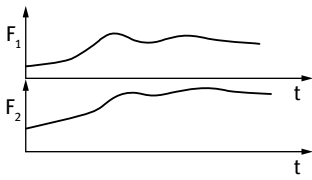


Fig. 1.26: Evolution of flows in the ratio control system. The controlled flow  $F_2$  follows with a small delay the wild feed  $F_1$ .

### 1.4 Inferential control systems

Sometimes it is useful to not control directly one variable, but a calculated one. For example, it is better to control the mass flow instead of volumetric flow of liquids or gases, especially when large variations of temperature occur.

It is known that the mass flow can be calculated as  $F_m = F_v \rho = F_v \frac{pM_g}{RT^\circ}$ , where the volumetric flow  $F_v$  is measured with an orifice plate transducer. If one limits the measurement at the flow transducer, the mass content remains unknown. This is why in inferential control, measurements of pressure, temperature, and additional computing elements are added (Fig. 1.27).

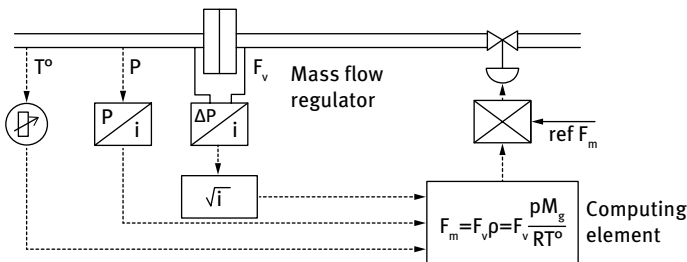


Fig. 1.27: Inferential control of the mass flow of a gas.



An other interesting example is the cascade control of heat content of the reboiler in the bottom of a distillation column (Fig. 1.28).

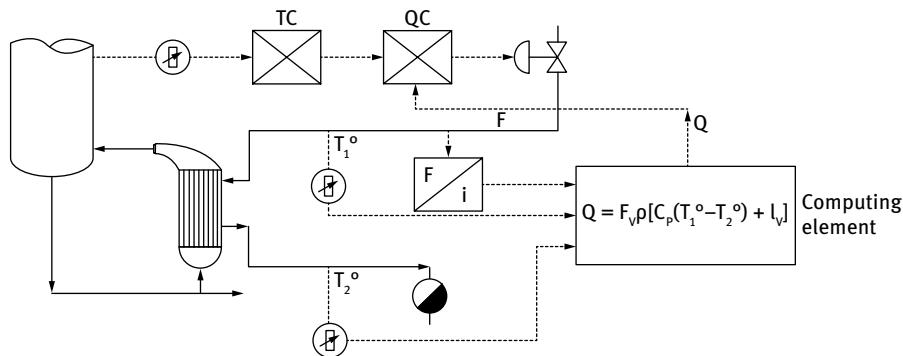


Fig. 1.28: Bottom temperature cascade control with the inner controlled variable  $Q$ .

The controlled variable of the inner loop of the cascade is the heat flow depending on the input and output temperatures of the steam in and out streams from the reboiler ( $T_1^o$  and  $T_2^o$ ). Thus, the controlled variable in the inner loop is  $Q = F\rho[c_p(T_1^o - T_2^o) + l_v]$ , where  $c_p$ , the specific heat, and  $l_v$ , the latent heat of vaporization, are considered constant in the usual range of temperature variation.

Thus, the final controlled parameter is the bottom temperature,  $T^o$ , and the intermediate controlled one is  $Q$ .

## 1.5 Selective control systems

There are situations when the appropriate functioning of a control loop depends not only on the controlled variable, but also on other variables. Usually, the level in the bottom of a separation column is controlled through the bottom flow  $B$ . The temperature profile in the stripping section can be kept constant through the heat flow into the reboiler. In start-up or shut-down situations or at strong disturbances, the level at the bottom may decrease even when the bottom control valve is completely closed. This happens because the vapor flow is higher than the reflux flow from the first tray. If there is no attempt to reduce the heat entering the reboiler, then the column could work with an empty bottom, which can in return destroy it. If an operator observes the event, he/she shifts the operating mode of the level control loop from AUTOMATIC to MANUAL mode and closes the steam valve of the reboiler. This operation requires the operator to have high professional expertise and qualification. A selective control system (Fig. 1.29) can solve this problem.

The selective system has in its composition a selector switch that can choose which control signal is predominant ( $T_B^o$  or  $L_B$ ) in manipulating the reboiler's control valve. If the bottom is nearly empty (supposing a level transducer with the output

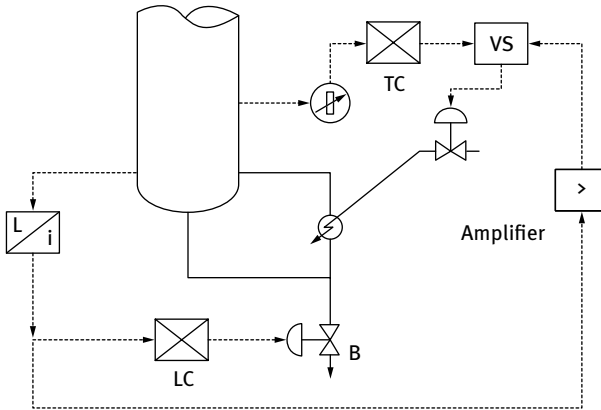
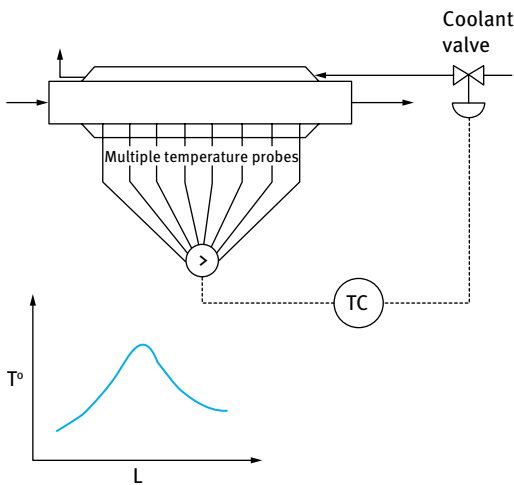


Fig. 1.29: Selective control system for the temperature and level of the bottom of the column.

signal 4–20 mA) at a transducer signal of 5 mA, the value selector VS takes the decision and conveys the smallest signal received: either the temperature signal or the level signal. This way, the reboiler’s steam valve is closed at the higher than normal temperature, and if this is not the case, at the minimum level signal. It has to be mentioned that in this case the bottom flow is reduced to 0.

Similar selective control systems are designed to locate the “hot spot” in plug flow reactors and to select which is the important signal to be considered in the control loop (Fig. 1.30). The position of the “hot spot” is not always the same; it depends on other operating conditions as well.



Temperature profile (qualitative representation along the length of a tubular reactor)

Fig. 1.30: Selective control loop of the highest temperature of a tubular reactor.

The temperatures measured along the reactor enter in a selector switch and the highest value is transmitted to the controller in order to be controlled.

## References

- [1] Agachi, P.S., Cristea, V.M., *Basic Process Engineering Control*, Walter de Gruyter GmbH, Berlin/Boston, 142, 2014.
- [2] Agachi, P.S., Cristea, V.M., *Basic Process Engineering Control*, Walter de Gruyter GmbH, Berlin/Boston, chapter 12, 2014.
- [3] Agachi, P.S., Cristea, V.M., *Basic Process Engineering Control*, Walter de Gruyter GmbH, Berlin/Boston, chapter 11, 2014.
- [4] Agachi, S., *Automatizarea proceselor chimice (Control of Chemical Processes)*, Casa Cartii de Stiinta, Cluj - Napoca, 185, 1994.
- [5] Luyben, W., *Parallel cascade control*, *Industrial & Engineering Chemistry Fundamentals*, 12, 463–467, 1973.
- [6] Lumkes J.H. Jr., *Control Strategies for Dynamic Systems. Design and Implementation*, CRC Press, 207, 2002.
- [7] Baker, G., *PID Tuning of plants with time delay, using root locus*, Master Theses and Graduate Research, Masteral Thesis Paper 4036, San Jose Scholar Works, San Jose University, [http://scholarworks.sjsu.edu/etd\\_theses?utm\\_source=scholarworks.sjsu.edu%2Fetd\\_theses%2F4036&utm\\_medium=PDF&utm\\_campaign=PDFCoverPages](http://scholarworks.sjsu.edu/etd_theses?utm_source=scholarworks.sjsu.edu%2Fetd_theses%2F4036&utm_medium=PDF&utm_campaign=PDFCoverPages), 2011.
- [8] Rao, A.S., et al., *Enhancing the performance of parallel cascade control using Smith predictor*, *ISA Transactions*, 48, 220–227, 2009.
- [9] Bharati, M., et al, *Autotuning of parallel cascade Control using setpoint relay*, *International Journal of Computer Applications Special Issue on “Evolutionary Computation for Optimization Techniques”*, ECOT, 2, 57–61, 2010.
- [10] Vilanova, R., Visioli, A., *PID Control in the Third Millennium: Lessons Learned and New Approaches*, Springer, the Netherlands, chapter 8, 2012.
- [11] Padhan, D.G., Majhi, S., *Synthesis of PID tuning for a new parallel cascade control structure*, Fr.A2, IFAC Conference on Advances in PID Control, PID'12, Brescia (Italy), March 28–30, 2012.
- [12] Agachi, P.S., Cristea, V.M., *Basic Process Engineering Control*, Walter de Gruyter, Berlin/Boston, 87, 2014.
- [13] Agachi, S., *Automatizarea Proceselor Chimice (Chemical Process Control)*, Casa Cartii de Stiinta, Cluj Napoca, 330, 1994.
- [14] Nisenfeld, A. and Miyasak, R., *Applications of feedforward control to distillation columns*, *Automatica*, 9, 319–327, 1973.
- [15] Nisenfeld, A.E., *Reflux or distillate. Which to control?*, *Chemical Engineering*, 78(21), 169–171, 1969.
- [16] Marinoiu, V., Paraschiv, N., Patrascioiu, C., *Conducerea cu calculatorul a unei coloane de distilare (Distillation column computer control)*, Contract Report, No. 18, CP Brazi, 1987.
- [17] Congalidis, J., Richards, J., Harmon Ray, W., *Feedforward and feedback control of a solution copolymerization reactor*, *AIChE Journal*, 35(6), 891–907, 1989.
- [18] Brauner, N., and Lavie, R., *Feedforward quality control through inventory manipulation in periodically perturbed mixed accumulators*, *Computers & Chemical Engineering*, 14(9), 1025–1029, 1990.
- [19] Rovaglio, M., Ranzi, E., Biardi, G., Fontana, M., Domenichini, D., *Rigorous dynamics and feedforward control design for distillation processes*, *AIChE Journal* 36(4), 576–586, 1990.

- [20] Sheffield, R. E., *Integrate process and control system design*, Chemical Engineering Progress, 88(10), 30–35, 1992.
- [21] Biao, H., *Feedforward plus feedback controller performance assessment of MIMO systems*, Control Systems Technology, IEEE Transactions, 8(3), 580–587, 2002.
- [22] Agachi, P.S., Cristea, V.M., *Basic Process Engineering Control*, Walter de Gruyter, Berlin/Boston, 346, 2014.
- [23] Nandong, J., A unified design for feedback-feedforward control system to improve regulatory control performance, *International Journal of Control, Automation and Systems*, 13(1), 91–98, 2015.
- [24] Cus, F., Zuperl, U., Balic, J., *Combined feedforward and feedback control of end milling system*, Journal of Achievements in Materials and Manufacturing Engineering, 45(1), 79–88, 2011.
- [25] Rusli, E., Drews, T.O., Ma, D., Alkire, R., Braatz, R., *Robust nonlinear feedback–feedforward control of a coupled kinetic Monte Carlo–finite difference simulation*, Journal of Process Control, 16, 409–417, 2006.
- [26] Agachi, P.S., Cristea, V. M., *Basic Process Engineering Control*, Walter de Gruyter, Berlin/ Boston, Table 3.2, 75, 2014.

## 2 Model predictive control

### 2.1 Introduction

Currently, the most prominent representative of the model based control strategies is considered to be the model predictive control (MPC) algorithm. This consideration relies on the success MPC has demonstrated in industrial applications, associated to the interest of control researchers for developing its design along the last four decades. Emerged from practice-driven needs for high-performance control, the MPC design theory brings out a natural relationship between the behavior of the process to be controlled, revealed by the model, and the design of the controller. MPC may be considered a member of the internal model control family to whom it shares the direct use of the process model when building the controller, but adding new capabilities such as prediction, optimization, and handling constraints.

MPC strategy relies on the use of the process model to make predictions of the future behavior of the process, as result of both the past but known and the future but unknown inputs, to find the best future input sequence aimed to optimize a control performance index having associated constraints on inputs, states and outputs. The constrained optimization feature of the MPC control strategy is most appreciated, as it allows the controller to overcome the traditional stabilizing capability by involving optimization in the control design while incorporating the constraints always accompanying the control problem. The multivariable approach that MPC methodology is able to cope in a straightforward way is also appreciated for the control of processes with interactions between input and output variables.

The MPC applications cover a very large area of fields with thousands of successful reported industrial implementations [1–2]. The literature shows that almost all the practice and research studies where control is involved may benefit of the MPC approach [2]. Some of them are: refining and production of petrochemicals [3–5], scheduling semiconductor production [6], PVC reactor [7], polyethylene reactor [8], water gas-shift reactor [9], thermal regenerator [10], autoclave composite processing [11], drying [12–13], heat exchanger network [14], drainage and irrigation channels [15–16], waste water treatment [17–18], solar air conditioning and desalination plants [19–20], cruise control [21], flight control [22–23], robotics [24], or medical applications [25]. The trend in the fields' expansion of MPC is continuously widening with applications for large-scale systems, fast dynamic systems, and low-cost systems [26].

### 2.2 MPC history

The MPC history may be considered to have its roots in the linear quadratic regulator (LQR) developed by Kalman in the early 1960s, which described the

proportional controller emerging from the minimization of a performance function that penalizes the squared deviation of the states and inputs from the origin, working on the basis of the state feedback and computed as solution of a Riccati equation [27–28]. But the LQR approach lacked in the ability of handling constraints.

More than five decades ago, the work of Zahed and Whalen [29] revealed the potential of coupling the minimum time optimal control with linear programming, and Propoi [30] showed the benefits of the receding horizon approach. They were the first two ingredients for the birth of the originally called *open-loop optimal control*. However it had to pass some time, until the 1970s, when the potential of this control approach was proven by industrial applications. The results presented by Richalet showed that model predictive heuristic control (MPHC) was a powerful control strategy, as its software implementation IDCOM was successfully used for identification and control [31]. The main characteristics of MPHC were the use of linear impulse response models for making predictions used in a quadratic performance function considered on a finite prediction horizon, to which input and output constraints were associated. The computation of the optimal inputs was based on a heuristic iterative algorithm. Soon, Cutler et al. [4] and Prett et al. [32] elaborated the new control version by the so-called *dynamic matrix control* (DMC) and successfully applied it for the control of the challenging fluid catalytic cracking process. IDCOM and DMC are considered the first generation of the new born industrial MPC technology [1].

The first generation of industrial MPC technology lacked in systematic handling of constraints but this problem was addressed by Cutler et al. [33] who posed the DMC as a quadratic program (QP) with constraints directly included in the optimization problem and the optimal inputs emerging as solution to a QP. The new quadratic dynamic matrix control (QDMC), broadly described by Garcia et al. [34], refined the concept of hard and soft constraints and may be considered the second generation of the industrial MPC technology [1].

The second generation of industrial MPC technology could not handle one important problem occurring in QDMC applications, i.e. tackling the situation of getting an infeasible solution and recovering from infeasibility. The new developed IDCOM-M algorithm uses a controllability supervisor to identify ill condition regions and has a mechanism for their avoidance, operates with two quadratic objective functions (one for output and one for input), makes control for a single future point (denoted as coincidence point, on the reference trajectory), and distinguishes between hard constraints by ranking them in the order of priority [35–36]. This is considered as a representative of the third generation of the industrial MPC technology [1].

It is a recognized fact that at the beginning of the MPC technology development the MPC industrial applications have preceded the MPC theoretical design, but the order has been reversed in the last decades. During the 1980s to the 1990s, the MPC

theory mostly focused on the quantitative analysis of the industrial MPC algorithm performance. Since the 1990s, the MPC qualitative synthesis theory has predominated over the industrial applications. A large interest was shown and effort made by the scientific community during the last decades to develop sound theoretical theory for dealing with different MPC design aspects emerged from process requirements, such as stability, tuning, optimization, robustness, uncertainty, etc., with the aim of building MPC controllers able to guarantee both general and specific plant performance [26]. In the future, these theoretical developments need to be validated by the industrial applications. Practical solutions are expected to match the two aspects of the MPC development directions, such as reducing the computational burden, simplifying the theoretical approach, and rendering physical meaning to the complex MPC theoretical research results. The costs for maintenance and training should be kept as low as possible and the need for low-cost MPC controllers should be also satisfied [26].

The MPC technology is far from reaching its maximum potential and further developments are expected both from theory and practice. The approach of MPC control for large-scale, fast dynamic, and low-cost systems is a current and challenging area of research for large communities of researchers and practitioners. One of the major research directions remains the way that process nonlinearity may be handled both in MPC applications and the development of an efficient but comprehensive nonlinear model predictive control (NMPC) theoretical framework. Adaptive, robust (tube-based), distributed, embedded, cooperative, economic, and stochastic (scenario based) MPC are just several hot topics of research waiting for new ideas and solutions.

### 2.3 Basics of MPC control strategy

In the following a tutorial of linear MPC is presented in a simple but intuitive way. At its root the MPC strategy may be considered to have the same basic structure as the traditional feedback-feedforward control loop, as it is presented comparatively in Fig. 2.1.

As shown in Fig. 2.1, there is no difference between the arrangements of the elements building the two control loops, i.e. the process, controller, measuring devices, and final control element have the same role and position in the feedback (plain line) and feedforward (dotted line) control configurations. The meaning of the signals in the loops remains the same too, i.e. the setpoint (reference), output (controlled variable), measurements (both for the output and for the disturbance), control (input) variable and manipulated variable carry the same information. What makes the difference stands in the way the control (input) variable is computed. The classical controller is replaced by an optimization algorithm that runs online.

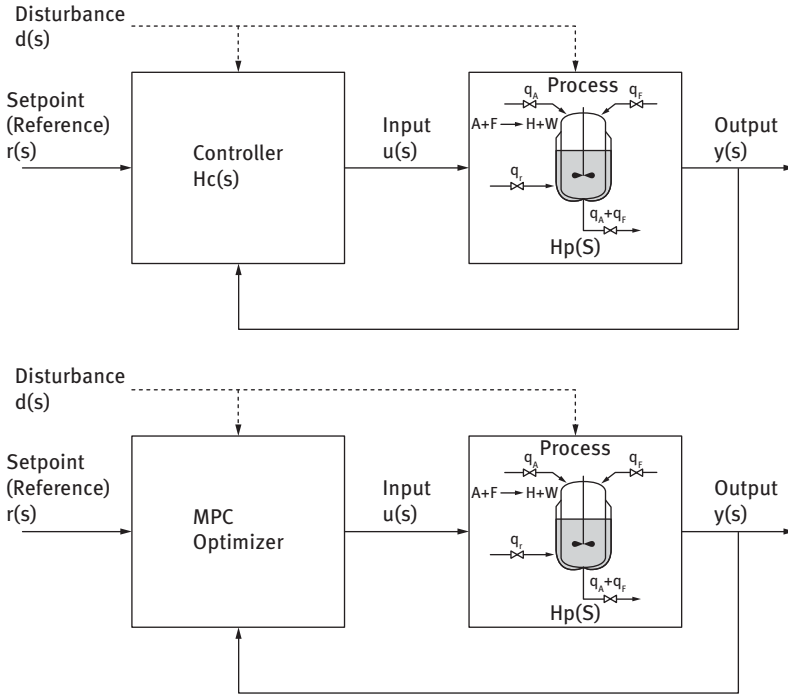


Fig. 2.1: (a) Traditional and (b) MPC control structures.

For the MPC control loop, the control variable is generated by the MPC controller as a result of solving an optimization problem. The performance index of the optimization problem consists of a function that evaluates the square difference between the desired reference and the predicted output variable, over a future time interval denoted as prediction horizon. As the performance index depends on the future predicted variable, it is implicitly also depending on the future values of the control variable. This makes the independent (unknown) variable of the performance index to be, in fact, the future control values, which will be computed by solving the minimization problem.

A process model is needed for making predictions on the future output variable evolution, as they depend on the future changes of the input variable. Mathematically, the process model includes in the performance index the relationship between the future values of the controlled variable and the future values of the input variable (the unknown of the optimization problem).

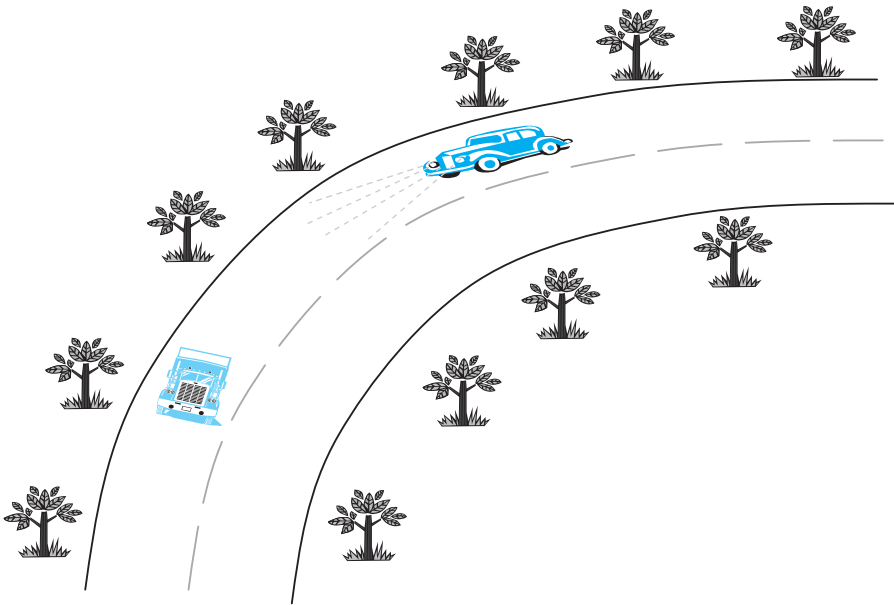
But what emerges from solving this minimization problem is only an open-loop optimal control. There is an important weakness of the open-loop optimal control, i.e. no feedback is included yet. This is introduced by the so-called receding horizon approach which will be further described. As the MPC algorithm is implemented on a computer, the discrete time approach is usually the one that is assumed for the



MPC control strategy. The discrete approach introduces the so-called future input sequence having values changing only at the sampling time moments and remaining constant in between (i.e. zero order hold, ZOH, assumption).

### **i** Example 2.1

To better understand the MPC control approach, we may consider a trivial but intuitive tutoring example. It consists of the way a car (or bike) is controlled in traffic. The analogy of controlling a process and driving a car on a road will reveal the natural way that MPC control strategy is conceived. Let us consider the task of driving a car on the road from a starting point to a destination point, as it is schematically presented in Fig. 2.2.



**Fig. 2.2:** Analogy between driving a car and the MPC control.

The driver (controller) has to manipulate the steering wheel, the accelerator (throttle pedal) and the brakes to stay in lane, obey speed limitations, avoid overacceleration or overbraking, etc. The position of the car on the lane may be considered as controlled variable and the middle of the lane trajectory as reference (setpoint) for the car system.

Reaching the destination point on the best possible trajectory, i.e. keeping the position of the car as much as possible on the middle of the lane all along the starting-destination travel, may be planned before leaving the starting point. This

means planning the best control actions for the steering wheel, accelerator and brakes for all future moments of the planned road trip. Finding these optimal control actions is done by the driver who knows not only the desired trajectory but has also knowledge about the behavior of his car when manipulates the steering wheel, accelerator, and brakes. This means the driver has a mental model of his car. Only on the basis of this model the driver is able to make the best plan aiming to keep the minimum distance between the car position and the reference trajectory. This best plan, done before leaving the departure point, represents the optimal open-loop control of the car. It is worthy to mention that it relies on the model of the car and predicts its future position all along the time of the trip (time horizon) while computing the best future sequence for the control actions.

As it may be observed from the driving example, the control problem may be formulated on the basis of a performance index to be minimized, i.e. the deviation (possibly square) distance between the current (and future) position of the car and the reference trajectory, while conforming to constraints, such as the obeying speed limitations and avoiding over acceleration or over braking.

Unfortunately, although this optimal open-loop control could be considered of interest, it shows to have major problems. These problems are related to the unknown events that may appear during the trip, making ineffective the optimal plan. Such unpredictable events are other vehicles moving on the same lane with different speed or stopping (as the truck is in Fig. 2.1), meeting wet track with low adhesion, people or animals crossing the road in front of the car, etc. All these events may be considered unknown (unmeasured) disturbances. Such disturbances cannot be handled by the open-loop optimal control hypothetical plan made by the driver before starting the trip.

There is still a solution to these problems, while keeping the prediction and optimization benefits introduced by the open-loop optimal control strategy. The driver should make his plan on a shorter time interval and apply it. Then, the driver should resume building a new plan upon reaching the end of this time interval, taking into account at that moment the unexpected event that appeared in the meantime. The new plan should be applied during a new time interval. And the succession of building new plans based on current information and their application during their subsequent time interval should be repeated all along the trip, up to its end. This manner of driving a car is in fact very close to the way the driver acts by making a plan for his best future driving actions on a short period of time, considering all available information about the environment, applying it, and repeating this planning and execution again.

The general theoretical approach for coping with the open-loop optimal control weakness is to split the time intended for controlling the system in small time intervals and compute an open-loop optimal control on each of these time periods. Every new open-loop optimal control plan should be performed when the system reaches the end of the previous time interval and using in control the information on the outputs (states) of the system at this moment (possibly affected by unmeasured

disturbances). What results is a succession of optimal plans computed and applied all over the period of control. Resuming the design of a new control plan is denoted as the receding horizon control. It is a necessary ingredient of the MPC strategy as it brings feedback into the control approach by considering, when planning, the current values of the outputs or states. This is obviously needed to counteract the effect of unmeasured disturbances. As the discrete time approach is commonly used for implementing MPC, the moments of building new optimal open-loop control plans are set at *each of the sampling time moments*. It is important to note that length of a new open-loop optimal control period of time spans over multiple sampling time intervals. Although the plan is made over a large future time period, only the first of the computed optimal control variable is sent (applied) to the controlled process. The optimization is done again at the very next sampling moment and a new open-loop optimal control plan is rebuilt. This feature of the MPC control approach is presented in Fig. 2.3.

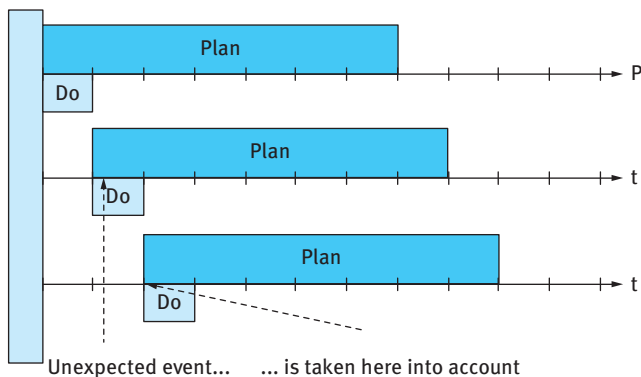


Fig. 2.3: Receding horizon feature of MPC.

As presented in Fig. 2.3, an unexpected event is taken into consideration in a new optimization plan at the very next sampling time moment following its occurrence, when this feedback information is used for the new prediction.

The future time period (multiple of the sampling time) over which the output is predicted when performing the open-loop control optimization, i.e. future plan duration with respect to the output, is denoted as the *prediction (output) horizon*. This prediction horizon is shifted one sampling time-step ahead into the future at each sampling period when the optimization is resumed. This is also shown in Fig. 2.3.

At first sight, it might be considered the same length for the future time period on which the optimal control variable sequence is computed. This is usually not the case for practical MPC implementations when this future time period associated to

the optimal control variable is considered to have a smaller length, i.e. the control variable is allowed to change only for a shorter future period of time and then remains constant for the rest of the time up to the prediction horizon end. This approach is due to reasons of sparing the computation resources, since the number of unknown of the optimization problem is reduced and the computation load too. The future time period (multiple of the sampling time) on which the open-loop control variable is computed when performing the open-loop control optimization, i.e. future plan duration with respect to the control (input) variable, is denoted as the *input horizon*.

Recalling the driving car example previously presented, the prediction horizon may be illustrated as the corresponding time needed by the car during night traffic to cover the distance its headlights are revealing, i.e. how long (in time) the driver sees in front of the road to make his open-loop optimal control plan.

In a nutshell, it may be concluded that MPC control strategy makes a repeated open-loop optimization of a performance index with respect to the future control actions, on the basis of the predictions made for the output variable, using the model of the process. The family of the MPC methodology has several versions, such as: IDCOM [31], IDCOM-M [35], DMC [4], MAC [3], QDMC [34], GPC [37–38], SPC [39], or UPC [40].

The representation of the MPC horizons for an arbitrary current moment  $k$  and considering the input variable emerging from a ZOH is presented in Fig. 2.4.

An important element of the MPC strategy is the *model* of the process. As presented before, the prediction made at any current (sampling) moment, used in the optimization performance index, relies on the model. Implication of the *prediction* in the control algorithm allows the MPC strategy to consider the future behavior of the controlled process, based on current information available, and take early counteracting measures, before the output is substantially diverged from the reference trajectory.

It may be considered that, taking into account the future reference changes, MPC prediction is allowing “feedforward” control with respect to the setpoint future evolution. This “look ahead” or “preview” capability is bringing efficiency to the algorithm. Coming back to the driving car example, the control based on traditional feedback control (such as PID) would mean driving the car only by looking in the lateral rear windows (or rear-view mirror). Control based on looking ahead ability would mean driving the car looking forward through the front window and making control based on knowing the future road reference trajectory.

Nevertheless, the classical feedforward control made on the basis of knowing the measured disturbances is also possible in the MPC framework. This is also based on a model of the process which in this case must consider the disturbance-output path. It is sufficient to include in the output prediction of the optimization index the measured disturbance effect, associated to the input (control variable) effect, to obtain a combined feedback-feedforward MPC controller.

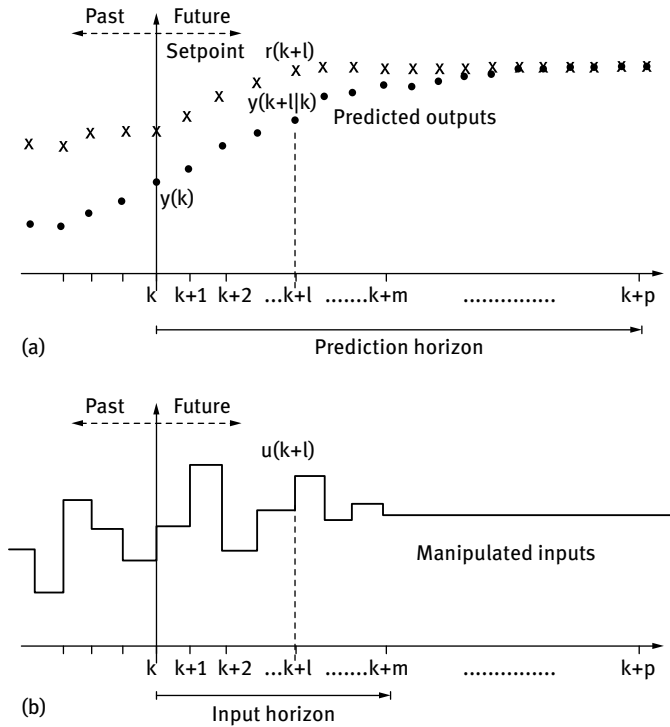


Fig. 2.4: Significance of the prediction and output horizons.

As the control problem is discrete, the optimization problem is performed in a discrete framework. A typical form of the MPC optimization problem, with its associated performance index, for the input-output process model description has the following form:

$$\min_{\Delta u(k) \dots \Delta u(k+m-1)} \{J(u, k)\} = \min_{\Delta u(k) \dots \Delta u(k+m-1)} \left\{ \sum_{l=1}^p ||Ywt_l[r(k+l) - y(k+l|k)]||^2 + \sum_{l=1}^m ||Uwt_l[\Delta u(k+l-1)]||^2 \right\} \quad (2.1)$$

where

- $J(u, k)$  is the performance index (cost function),
- $r(k)$  is the reference trajectory vector,
- $y(k)$  is the process output vector,
- $u(k)$  is the process input (control signal) vector, (i.e. manipulated variable),
- $\Delta u(k)$  is the process input (control) incremental signal (i.e. manipulated variable move)  $\Delta u(k) = u(k) - u(k-1)$ ,
- $p$  is the prediction horizon,

- $m$  is the input (control) horizon,
- $Ywt_l$  is the output (error) weighting factor for the moment  $(k + l)T$ ,
- $Uwt_l$  is the input (control) weighting factor for the moment  $(k + l)T$ ,
- $y(k + l, |k)$  is the prediction of  $y(k + l)$  for the generic moment “ $k + l$ ”, based on information available at time “ $k$ ”,
- $\|y(k)\|^2 = y(k)^2$  for scalar case ( $\|y(k)\|^2 = y^T(k) \cdot y(k)$  for vector case),
- $\|\Delta u(k)\|^2 = u(k)^2$  for scalar case ( $\|\Delta u(k)\|^2 = \Delta u^T(k) \cdot \Delta u(k)$  for vector case).

As presented in eq. (2.1), the performance index consists of two terms [41–42]. The first one penalizes the reference tracking error vector, multiplied by the output weighting factor  $Ywt_l$  and the second one that penalizes the input (control) move vector, multiplied by the input weighting factor  $Uwt_l$ .

Another form of the input-output MPC optimization formulation is framed in the generalized model predictive control (GPC) with a simplified performance index [37–38]:

$$\min_{\Delta u(k) \dots \Delta u(k+m-1)} \{J(u, k)\} = \min_{\Delta u(k) \dots \Delta u(k+m-1)} \left\{ \sum_{l=p_1}^{p_2} \|[r(k+l) - P(q)y(k+l|k)]\|^2 + \sum_{l=1}^m \|\lambda^2 [\Delta u(k+l-1)]\|^2 \right\} \quad (2.2)$$

where

- $Ywt_l = 1$  and  $Uwt_l = \lambda^2$ ,
- $p_1$  is the minimum prediction (output) horizon,
- $p_2$  is the prediction (output) horizon,
- $P(q)$  in a polynomial with the desired closed-loop poles,  $P(q) = a_r q^{-r} + a_{r-1} q^{-r+1} + \dots + a_1 q^{-1} + 1$  (in the simplest case  $P(q) = 1$ ).

Despite the fact that the input-output MPC formulation of the performance index is more intuitive and directly involves from practice emerged signals, the input-state-output model is very frequently used due to its comprehensive approach. The space-based MPC optimization may be formulated as [43–44]:

$$\min_{\Delta u(k) \dots \Delta u(k+m-1)} \{J(u, k)\} = \min_{\Delta u(k) \dots \Delta u(k+m-1)} \left\{ \sum_{l=1}^p \mathbf{Q} \|\mathbf{x}(k+l|k)\|^2 + \sum_{l=1}^m \mathbf{R} \|\Delta \mathbf{u}(k+l-1)\|^2 \right\} \quad (2.3)$$

where

- $\mathbf{x}(k)$  is the vector of the states,
- $\Delta\mathbf{u}(k)$  is the vector of the inputs move,
- $\mathbf{Q}$  is the state weighting matrix,
- $\mathbf{R}$  is the input (control) weighting matrix.

Most of the MPC performance index formulations are based on the square norm  $\|\cdot\|^2$ , as presented before. The 1-norm (based on the sum of the vector absolute values) or the infinity-norm (based on the maximum absolute value of vector's elements) are rarely used.

Constraints are always associated to the MPC optimization problem defined by eqs. (2.1) to (2.3). Solving constrained optimization problems is the most important gain of MPC. The capability to take into account constraints in a systematic way during the design and implementation of the controller makes it the most popular advanced control algorithm. Simple constraints, such as the final control element saturation, or more complex constraints on the states or on the controlled outputs, such as those coming out from equipment, technological, safety, environmental or economic reasons are putting challenges on the MPC based control. What results is the best control sequence of the control variable changes that brings the predicted output as close as possible to the reference, while satisfying all input, states and output constraints. For linear models with linear constraints and the performance index as presented in eq. (2.1), the solution of the optimization problem can be found using QP algorithms. If a 1-norm or infinity-norm performance index is used, linear programming algorithms will provide the solution. Both algorithms are convex and show convergence.

MPC may be successfully used for the control of processes showing large time constants, large pure time delay, nonminimum phase or even instability, and having the same systematic approach both for SISO and MIMO systems. It is also valued the MPC ability to work with unequal (excess) number of manipulated vs. the number of controlled variables. As the model is explicitly used in the controller design MPC has inherent robustness properties to the model-process mismatch and, to a certain extent, may also cope with structural modifications of the process, as long as they are revealed by the prediction model.

## 2.4 Types of MPC process models

The optimal input (control) sequence  $\{u(k), u(k+1), \dots, u(k+l), \dots, u(k+m-1)\}$  is minimizing the square error (first term) of the optimization index, possibly associated to the input variable change (second term) and constraints. The error term is computed as difference between the future reference and the predicted sequences. The predicted sequence is calculated on the basis on the process model. The explicit use of the model for designing the MPC controller differentiates the model based controller design, compared to the traditional PID design. Different types of models may be used

for making predictions [39]. A typical linear, discrete, and time invariant model of the process has the form:

$$y(k) = H_p^*(q)u(k) + d(k) = H_p(q)\Delta u(k) + H_{pD}(q)\Delta d(k) = H_p(q)\Delta u(k) + H_{pD}(q)e(k) \quad (2.4)$$

where  $y(k)$  denotes the output controlled variable,  $u(k)$  is the input (control or manipulated) variable,  $d(k)$  is the disturbance variable,  $\Delta u(k)$  is the input variable incremental change ( $\Delta u(k) = u(k) - u(k-1)$ ),  $\Delta d(k)$  is the disturbance variable incremental change ( $\Delta d(k) = d(k) - d(k-1)$ ) considered to be the white noise signal with zero mean  $e(k)$ ,  $H_p(q)$  is the discrete transfer function of the process on the input-output path,  $H_{pD}(q)$  is the discrete transfer function of the process on the disturbance-output path, and  $q$  is the forward shift operator. The output-input increment model assumes that the disturbance is not changing. As a result, in the prediction, it will not be necessary to include the model of the disturbance estimate, as it is already included in the incremental variables.

#### 2.4.1 Impulse and step response models

The most intuitive models emerged from direct process identification are the step and impulse response models. They have been used as first MPC models at the beginning of the MPC applications development, but were later extended to the more elaborate first principle models.

For the SISO case, the impulse response  $h_i$  and the unit step response  $s_i$  of a linear time invariant process are related by the well-known equations:

$$h_i = s_i - s_{i-1}, \quad s_i = \sum_{j=1}^i h_j \quad (2.5)$$

The discrete transfer functions are defined by

$$H_p^*(q) = \frac{y(q)}{u(q)} = \sum_{i=1}^{\infty} h_i q^{-i} \quad (2.6)$$

$$H_p(q) = \frac{y(q)}{\Delta u(q)} = \sum_{i=1}^{\infty} s_i q^{-i} \quad (2.7)$$

and, as result,

$$H_p(q) = \frac{y(q)}{\Delta u(q)} = \frac{y(q)}{(1-q^{-1})u(q)} = H_p^*(q) \frac{1}{(1-q^{-1})} = \sum_{i=1}^{\infty} h_i \frac{q^{-i}}{(1-q^{-1})} \quad (2.8)$$



From practical reasons, truncated impulse and step response models (up to a multiple of  $n$  sampling times,  $nT$ , i.e. the time process needs to settle and reach almost the steady state) are used instead of the infinite response models. They have the following form in the discrete time formulation:

$$y(k) = \sum_{i=1}^n h_i u(k-i) + d(k) \tag{2.9}$$

$$y(k) = \sum_{i=1}^{n-1} s_i \Delta u(k-i) + s_n u(k-n) + d(k) \tag{2.10}$$

For the MIMO case, when the system has a  $v$ -dimensional input vector  $\mathbf{u}(k) = [u_1(k) \ u_2(k) \ \dots \ u_v(k)]^T$  and a  $w$ -dimensional output vector  $\mathbf{y}(k) = [y_1(k) \ y_2(k) \ \dots \ y_w(k)]^T$ , the truncated response models have the forms:

$$\mathbf{y}(k) = \sum_{i=1}^n \mathbf{H}_i \mathbf{u}(k-i) + \mathbf{d}(k) \tag{2.11}$$

$$\mathbf{y}(k) = \sum_{i=1}^{n-1} \mathbf{S}_i \Delta \mathbf{u}(k-i) + \mathbf{S}_n \mathbf{u}(k-n) + \mathbf{d}(k) \tag{2.12}$$

where the impulse response matrix is

$$\mathbf{H}_i = \begin{bmatrix} H_{11}(i) & H_{12}(i) & \dots & H_{1v}(i) \\ H_{21}(i) & H_{22}(i) & \dots & H_{2v}(i) \\ \vdots & \vdots & \vdots & \vdots \\ H_{w1}(i) & H_{w2}(i) & \dots & H_{wv}(i) \end{bmatrix} \tag{2.13}$$

For this MIMO case, the step response matrix is related to the the impulse response matrix by

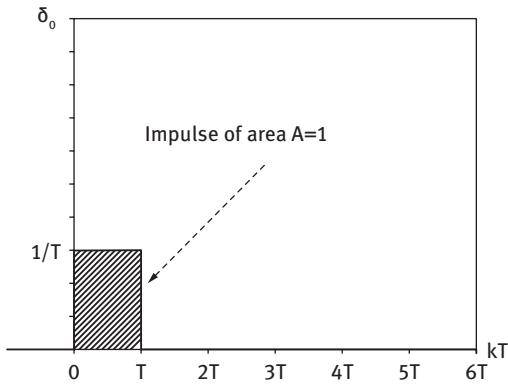
$$\mathbf{S}_i = \sum_{l=1}^i \mathbf{H}_l \tag{2.14}$$

or

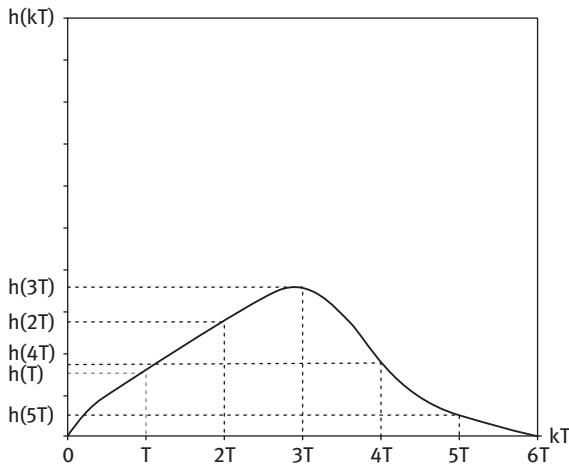
$$\mathbf{H}_i = \mathbf{S}_i - \mathbf{S}_{i-1} \tag{2.15}$$

**Example 2.2**

The unit impulse and step response models may be obtained from experimental identification data. Consider the unit impulse input and its corresponding truncated unit impulse response presented in Figs. 2.5 and 2.6:



**Fig. 2.5:** Unit impulse input signal.



**Fig. 2.6:** Truncated unit impulse response signal.

and unit step input with its corresponding truncated step response presented in Figs. 2.7 and 2.8:

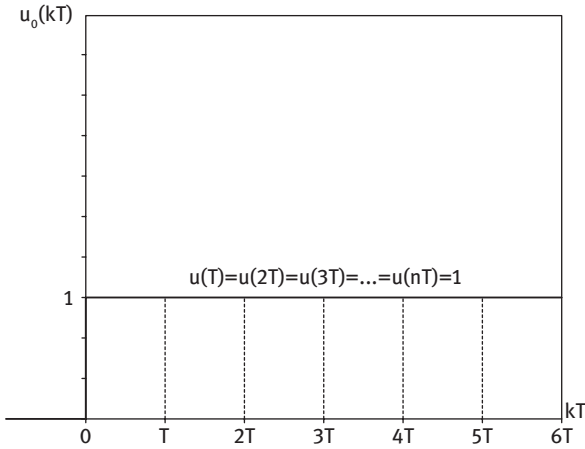


Fig. 2.7: Unit step input signal.

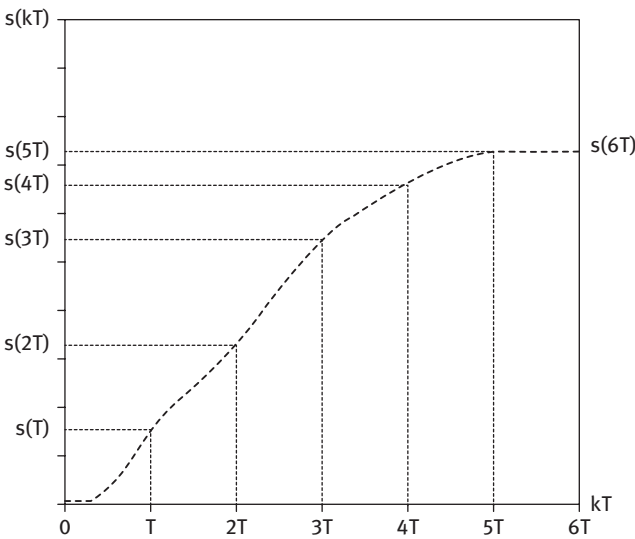
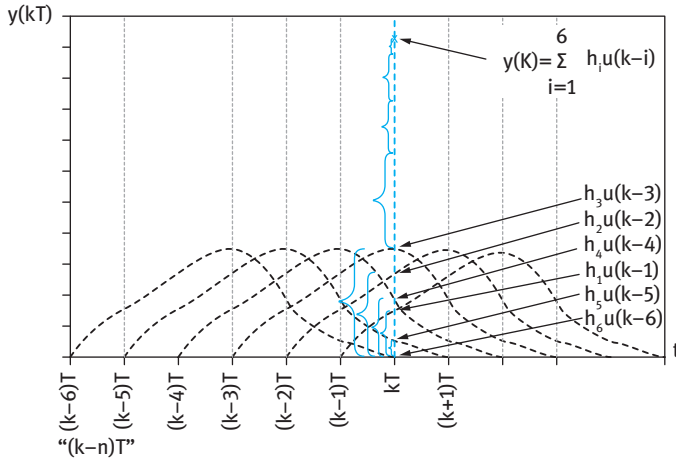


Fig. 2.8: Truncated unit step response signal.

Based on the homogeneity and additivity properties of the linear systems, implicit in the convolution model presented in eq. (2.9) and the unit impulse response parameters presented in Fig. 2.6, it is possible to compute the response of the system to an arbitrary input. A trivial but tutorial example is the computation of the system's response to the unit step input signal, using information emerging

from the unit impulse truncated response. The graph representation of this computation is given in Fig. 2.9 for the input step signal of the same form as presented in Fig. 2.7.



**Fig. 2.9:** Computation of the unit input step response signal based on the truncated impulse response model.

Figure 2.9 shows the computation by summation of the step response  $y(k)$ , at the generic time moment  $t = kT$ , using the truncated impulse response model presented in eq. (2.9).

### Example 2.3

In this example, it is shown the way the output (response) of the system can be determined on the basis of the the truncated step model presented in eq. (2.10) and the unit step response parameters presented in Fig. 2.8. In this example, the unit ramp input signal is considered and the response of the system is computed. The input ramp signal and its equivalents (obtained by decomposition) are presented in Fig. 2.10. The graph representation of the response computation is given in Fig. 2.11.

Figure 2.11 shows the computation of the step response  $y(k)$ , at the generic time moment  $t = kT$ , using the truncated step response model presented in eq. (2.10). Both examples 2.2 and 2.3 assume that no disturbance is present,  $d(k) = 0$ .

Examples 2.2 and 2.3 reveal the natural way impulse or step response models may be used for computing the process output for an arbitrary input, using the truncated

**i**

impulse or step response equations. The latter will be used for making the MPC predictions.

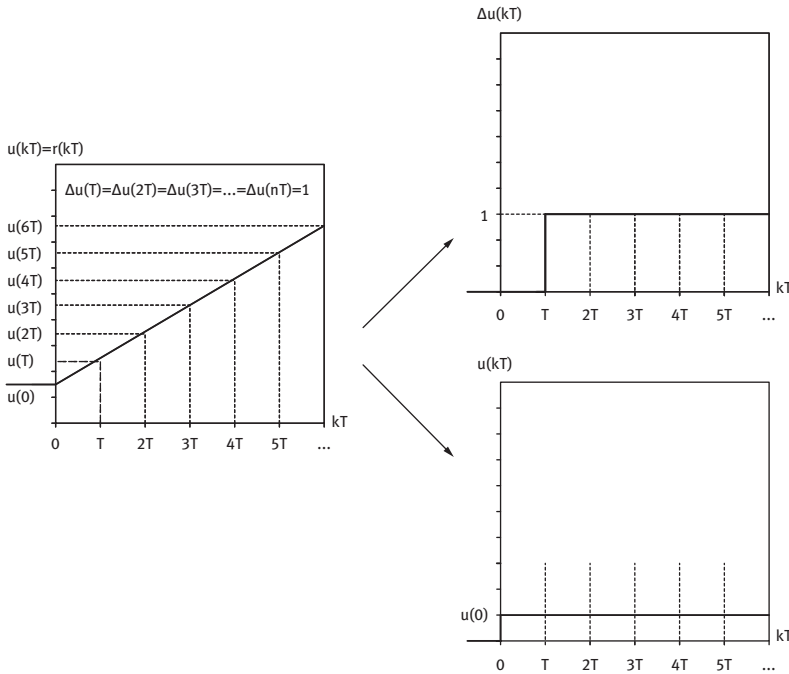


Fig. 2.10: Unit ramp input signal and its equivalent signals.

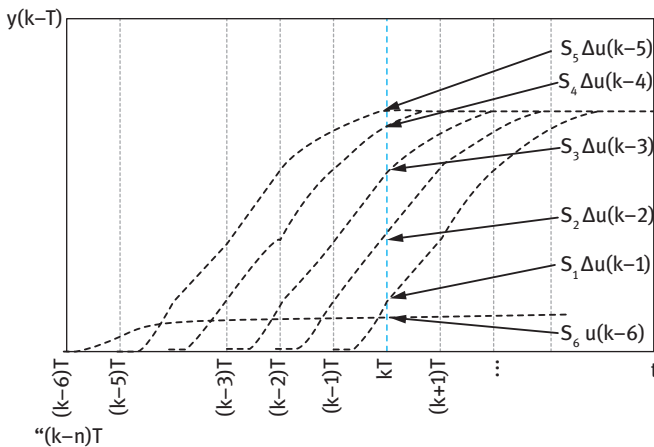


Fig. 2.11: Computation of the unit input ramp response signal based on the truncated unit step response model.

### 2.4.2 State-space models

The time-invariant state-space description of the system offers a comprehensive representation of the dynamic and steady-state behavior, since the internal system's variables (states) are also revealed and can be controlled. The state-space realization shows to be appropriate both for SISO and MIMO system description, using the same mathematical form. The general discrete form of the linear time invariant state-space model has the following representation:

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\Delta\mathbf{u}(k) + \mathbf{K}\mathbf{e}(k) \\ \mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{e}(k) \end{cases} \quad (2.16)$$

where the matrices  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , and  $\mathbf{K}$  may be directly computed from their continuous corresponding ones, i.e. the matrices of the continuous linear time invariant state-space description.

Nevertheless, the input-output models may be also converted in the state space form, the latter being also preferred due to its inherent numerical computations reliability. The relationship between the input-output description and the state-space formulation may be directly obtained from eqs. (2.16) using the Z transform. Using the forward shift operator, the relationship between the two types of models may be described by

$$\mathbf{H}_p(q) = \mathbf{C}(q\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \quad (2.17)$$

$$\mathbf{H}_{pD}(q) = \mathbf{C}(q\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{K} + \mathbf{I} \quad (2.18)$$

It is worth to mention that due to the fact that usually not all of the states are directly measurable a state space estimator is necessary for constructing the initial (current moment) state from the available measured outputs.

### 2.4.3 Time series models

The CARIMA (Controlled Auto-Regressive Integrated Moving Average) type of models has been also used to make MPC prediction [37–39]. Also known as transfer function models, CARIMA models have gained large acceptance especially for SISO systems. The development of the MIMO CARIMA models is straightforward, but becomes cumbersome for applications to large systems.

Its largest popularity among other transfer function models relies on the capability of including uncertainty in a simple way, being suitable to reveal the slowly changing disturbances that exhibit nonzero steady state.

The discrete equation describing the model is

$$A(q)y(k) = B(q)u(k) + C(q)\frac{\xi(q)}{\Delta(q)} \quad (2.19)$$

In eq. (2.19), the shift operator  $q^{-1}$  was used and  $\Delta(q) = 1 - q^{-1}$ .  $A(q)$ ,  $B(q)$ ,  $C(q)$  are polynomials in the time shift-operator  $q^{-1}$ . The factor  $\xi(q)$  considers the uncertainty on the input-output path.  $\xi(q)$  is assumed to have zero mean and finite variance.

Based on the above equation, the following form of the difference equation can be derived:

$$A(q)\Delta(q)y(k) = B(q)\Delta(q)u(k) + C(q)\xi(q) \quad (2.20)$$

and as a result, the transfer functions of eq. (2.4) get the following forms:

$$H_p(q) = \frac{B(q)}{A(q)\Delta(q)} \quad (2.21)$$

$$H_{pD}(q) = \frac{C(q)}{A(q)\Delta(q)} \quad (2.22)$$

## 2.5 Predictions for MPC

At each current moment of time  $kT$ , the output of the process is considered over the future time horizon  $pT$  by making a set of  $p$ -step ahead predictions of the output:  $\{y(k+1|k), y(k+2|k), \dots, y(k+l|k), \dots, y(k+p|k)\}$ . The predictions are based on information available at current time step  $kT$  and depend on both past and future values (with respect to current time) of the input variable  $u(k)$  or on its increments  $\Delta u(k)$ . However, the effect of past input variable on the predicted output may be directly computed. This computation is based on knowing all the past *input (control) variable values* and the *model* of the process. The effect of the future input variable sequence  $\{u(k), u(k+1), \dots, u(k+l), \dots, u(k+m-1)\}$  or sequence of input variable move  $\{\Delta u(k), \Delta u(k+1), \dots, \Delta u(k+l), \dots, \Delta u(k+m-1)\}$  on the predicted output is not yet known. It will be the task of minimizing the performance index to find this latter sequence, as the performance function penalizes the square error (square difference between the reference trajectory and the predicted output).

From the cause-effect viewpoint, the prediction has two components. One component is due to the action of the past and future input variable moves  $H_p(q)\Delta u(k)$  and the second is due to the action of the disturbances  $d(k) = H_{pD}(q)e(k)$  [39]. The assumption that  $e(k)$  is the zero-mean white noise will be considered further and the best prediction for future values of  $e(k+l) = \Delta d(k+l) = d(k+l) - d(k+l-1)$  will be  $e(k+l|k) = \Delta d(k+l|k) = 0$ , for  $l > 0$ .

The development of the prediction equations for SISO case will be presented both for the input-output and the state-space model approaches, but the extension

to the MIMO case is straightforward [39]. A simple form of the prediction equation may be obtained if the prediction  $p$  and the control horizon  $m$  are considered equal,  $p = m = N$ .

The vector of predicted values of the output variable based on information available at current time step  $kT$  is defined as

$$\hat{\mathbf{y}}(k) = [y(k+1|k) \ y(k+2|k) \ \dots \ y(k+l|k) \ \dots \ y(k+N|k)]^T \quad (2.23)$$

and the vector of future values of the control variable move:

$$\hat{\mathbf{u}}(k) = [\Delta u(k) \ \Delta u(k+1) \ \dots \ \Delta u(k+l) \ \dots \ \Delta u(k+N-1)]^T \quad (2.24)$$

A prediction equation may be developed with the same form for both the input-output and the state-space model approaches. Its formulation is

$$\hat{\mathbf{y}}(k) = \mathbf{M}\hat{\mathbf{u}}(k) + \hat{\mathbf{y}}_0(k) \quad (2.25)$$

where  $\mathbf{M}$  is the so-called *predictor matrix* and describes the effect of future unknown input moves on the predicted output and  $\hat{\mathbf{y}}_0$  is the so-called *free run output* and describes the effect of past input moves (before the current time step,  $kT$ ) on the predicted output.

When *step response models are used for prediction*, the predicted output variable emerges from the step truncated response model presented in eq. (2.10), where the generic moment  $lT$  is considered all along the prediction horizon ( $0 < l < N$ ). The prediction is described, for the SISO case, by [45]:

$$y(k+l|k) = \sum_{i=1}^l s_i \Delta u(k+l-i) + \sum_{i=l+1}^{n-1} s_i \Delta u(k+l-i) + s_n u(k+l-n) + d(k+l|k) \quad (2.26)$$

In eq. (2.26), the first term considers the effects of the sequence of future control moves,  $\{\Delta u(k), \Delta u(k+1), \dots, \Delta u(k+l-1)\}$ , on the predicted output variable  $y(k+l|k)$ . The second term of eq. (2.26) considers the effect of the the past control moves,  $\{\Delta u(k-1), \Delta u(k-2), \dots, \Delta u(k-n+1)\}$ , on the predicted output variable  $y(k+l|k)$ . The third term in eq. (2.26) considers the furthestmost input that affects the predicted output  $y(k+l|k)$ , according to the length of the truncation time,  $nT$ . The last term considers the prediction of future assumed disturbances on the same predicted output variable  $y(k+l|k)$ . As presented before, it is also assumed that  $e(k+l|k) = \Delta d(k+l|k) = 0$  and the computation of the disturbance effect is given by:

$$d(k+l|k) = d(k|k) = y_m(k) - y(k|k) \quad (2.27)$$



where  $y_m(k)$  is the measured value of the output at the current time  $kT$  and  $y(k|k)$  is the predicted value of the output, at the same moment of time, provided by the model:

$$y(k|k) = \sum_{i=1}^{n-1} s_i \Delta u(k-i) + s_n u(k-n) \tag{2.28}$$

This last term in eq (2.26) is the one bringing feedback in the MPC control algorithm.

For the MIMO case the prediction is described, in a similar way, by [2]:

$$\mathbf{y}(k+l|k) = \sum_{i=1}^l \mathbf{S}_i \Delta \mathbf{u}(k+l-i) + \sum_{i=l+1}^{n-1} \mathbf{S}_i \Delta \mathbf{u}(k+l-i) + \mathbf{S}_n \mathbf{u}(k+l-n) + \mathbf{d}(k+l|k) \tag{2.29}$$

and

$$\mathbf{d}(k+l|k) = \mathbf{d}(k|k) = \mathbf{y}_m(k) - \mathbf{y}(k|k) = \mathbf{y}_m(k) - \sum_{i=1}^{n-1} \mathbf{S}_i \Delta \mathbf{u}(k-i) + \mathbf{S}_n \mathbf{u}(k-n) \tag{2.30}$$

Assuming the length of the prediction horizon be larger than the step response truncation time,  $N \geq n$ , and  $s_{n+i} = s_n$  for  $0 \leq i \leq N-n$ , the predictor matrix of eq. (2.25) gets the following form (SISO case):

$$\mathbf{M} = \begin{bmatrix} s_1 & 0 & 0 & \cdots & 0 \\ s_2 & s_1 & 0 & \cdots & 0 \\ s_3 & s_2 & s_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_N & s_{N-1} & s_{N-2} & \cdots & s_1 \end{bmatrix} \tag{2.31}$$

As a result, the free run output  $\hat{\mathbf{y}}_0$  of eq. (2.25) may be described by

$$\hat{\mathbf{y}}_0(k) = \begin{bmatrix} s_2 & s_3 & \cdots & s_N \\ s_3 & s_4 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ s_N & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \Delta u(k-1) \\ \Delta u(k-2) \\ \vdots \\ \Delta u(k-N+1) \end{bmatrix} + \begin{bmatrix} s_N \cdot u(k-N+1) \\ s_N \cdot u(k-N+2) \\ \vdots \\ s_N \cdot u(k-1) \end{bmatrix} + \begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix} (y_m(k) - y(k|k)) \tag{2.32}$$

**i** **Example 2.4**

Based on the truncated step model presented in eq. (2.10) and the unit step response parameters presented in Fig. 2.8, the computation of the predictions for MPC and an intuitive graphic representation of this computation are presented in the following.

The example considers the step response model truncated to  $n = 6$  sampling time steps. The input for which prediction will be computed is presented in Fig. 2.12.

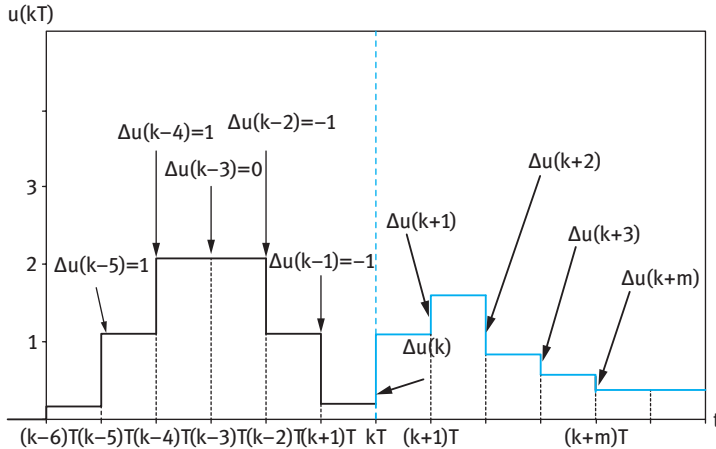


Fig. 2.12: Input signal with arbitrary form used for MPC prediction.

Figure 2.12 shows the arbitrary input signal for which the prediction will be performed, revealing both its past and the future values (with respect to the current time  $kT$ ). The increment (input move) values are also presented.

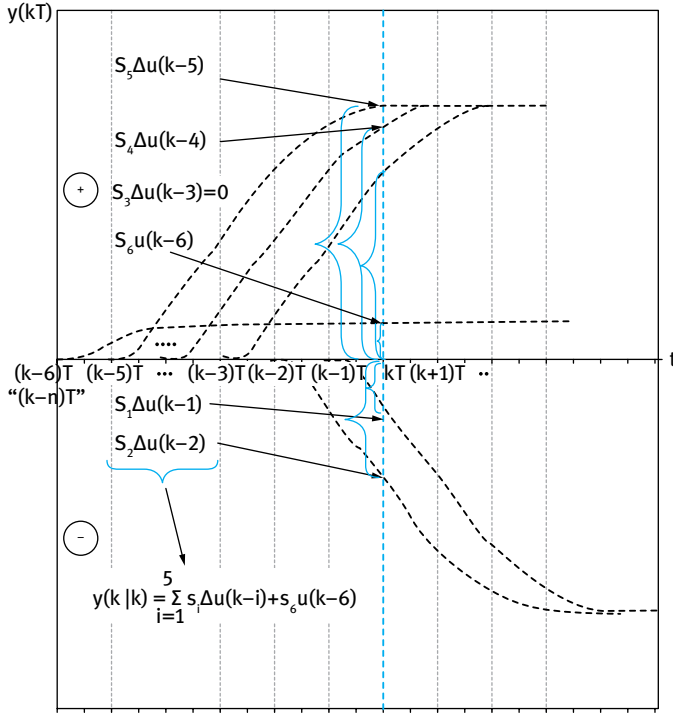
The predicted output for the current moment of time is computed with eq. (2.26), for  $l = 0$ , or with eq. (2.28) and has the following analytical form:

$$\begin{aligned}
 y(k|k) &= \sum_{i=1}^0 s_i \Delta u(k-i) + \sum_{i=1}^5 s_i \Delta u(k-i) + s_6 u(k-6) = s_1 \Delta u(k-1) + s_2 \Delta u(k-2) + \\
 &+ s_3 \Delta u(k-3) + s_4 \Delta u(k-4) + s_5 \Delta u(k-5) + s_6 u(k-6) = \\
 &= s_1 \cdot (-1) + s_2 \cdot (-1) + s_3 \cdot 0 + s_4 \cdot (1) + s_5 \cdot (1) + s_6 u(k-6) = \quad (2.33) \\
 &= -s_1 - s_2 + 0 + s_4 + s_5 + s_6 u(k-6)
 \end{aligned}$$

As expected, the prediction of the output only depends on past known inputs.

Figure 2.13 shows the graphical computation of the predicted response  $y(k|k)$ , at the current time moment  $t = kT$ . The segments corresponding to the components of the prediction, which should be graphically added to build  $y(k|k)$ , are individually revealed in the same figure. They are considered with plus sign above the abscissa and with negative sign below it.

The computation of the prediction based on summation of step response oriented segments is intuitive and simple.



**Fig. 2.13:** Graphical representation of the components that build by summation the prediction for  $y(k|k)$ , based on the truncated step response model.

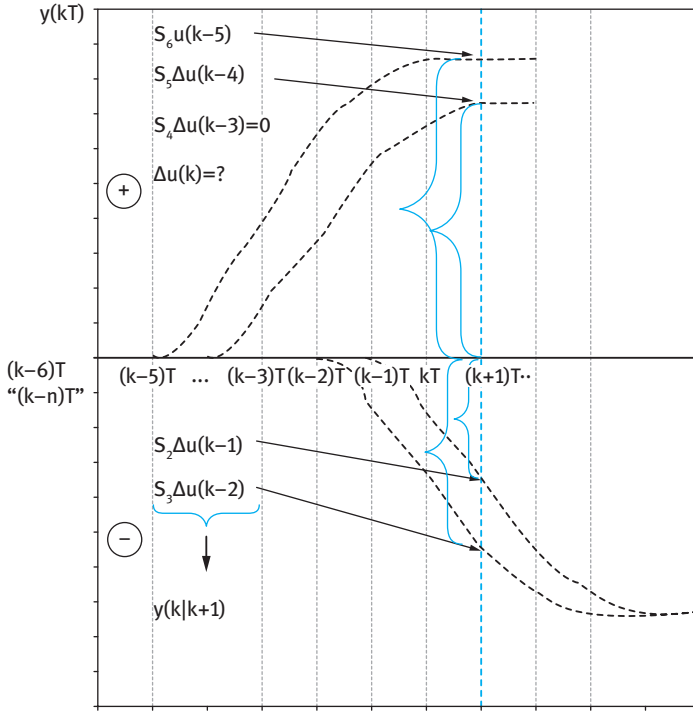
Continuing the same computation approach, prediction for the following sampling time moment,  $y(k + 1|k)$ ,  $l = 1$ , with the same eq. (2.26), has the following analytical form:

$$\begin{aligned}
 y(k + 1|k) &= \sum_{i=1}^1 s_i \Delta u(k + 1 - i) + \sum_{i=2}^5 s_i \Delta u(k + 1 - i) + s_6 u(k + 1 - 6) = \\
 &= s_1 \Delta u(k) + s_2 \Delta u(k - 1) + \\
 &+ s_3 \Delta u(k - 2) + s_4 \Delta u(k - 3) + s_5 \Delta u(k - 4) + s_6 u(k - 5) = \\
 &= s_1 \Delta u(k) + s_2 \cdot (-1) + s_3 \cdot (-1) + s_4 \cdot 0 + s_5 \cdot (1) + s_6 u(k - 5) = \quad (2.34) \\
 &= s_1 \Delta u(k) - s_2 - s_3 + 0 + s_5 + s_6 u(k - 5)
 \end{aligned}$$

Prediction of the output for the future moment of time  $(k + 1)T$  depends on the known past inputs and the future (unknown) input variable move  $\Delta u(k)$ .

Figure 2.14 shows the graphical computation of the predicted response  $y(k + 1|k)$ , for the time moment  $(k + 1)T$ , on the basis of information known at the current time moment  $t = kT$ . The segments corresponding to the components of the prediction,

which should be graphically added to build  $y(k + 1|k)$ , are individually revealed in the Fig. 2.14. Using the same representation, they are considered with plus sign above the abscissa and with negative sign below it.



**Fig. 2.14:** Graphical representation of the components that build by summation the prediction for  $y(k + 1|k)$ , based on the truncated step response model.

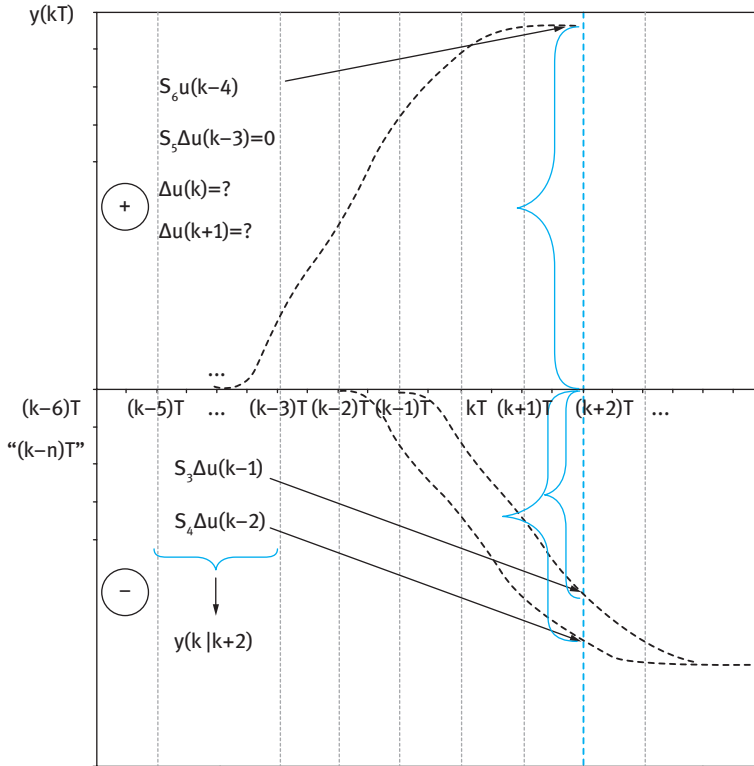
As noticed, the contribution of the input variable move  $\Delta u(k)$  is not represented in Fig. 2.14 because it is not known at current moment  $kT$ . It will be computed at the optimization step of MPC algorithm.

Based on the same procedure, the computation of the prediction for  $y(k + 2|k)$ ,  $l = 2$ , has the following analytical form:

$$\begin{aligned}
 y(k + 2|k) &= \sum_{i=1}^2 s_i \Delta u(k + 2 - i) + \sum_{i=3}^5 s_i \Delta u(k + 2 - i) + s_6 u(k + 2 - 6) = \\
 &= s_1 \Delta u(k + 1) + s_2 \Delta u(k) + \\
 &+ s_3 \Delta u(k - 1) + s_4 \Delta u(k - 2) + s_5 \Delta u(k - 3) + s_6 u(k - 4) = \\
 &= s_1 \Delta u(k + 1) + s_2 \Delta u(k) + s_3 \cdot (-1) + s_4 \cdot (-1) + s_5 \cdot 0 + s_6 u(k - 4) = (2.35) \\
 &= s_1 \Delta u(k + 1) + s_2 \Delta u(k) - s_3 - s_4 + s_5 \cdot 0 + s_6 u(k - 4)
 \end{aligned}$$

Prediction of the output for the future moment of time  $(k + 2)T$ , depends on the known past inputs and the future (unknown) input variable moves  $\Delta u(k)$  and  $\Delta u(k + 1)$ .

Figure 2.15 shows the graphical computation of the predicted response  $y(k + 2|k)$ , using the same approach of adding oriented segments.



**Fig. 2.15:** Graphical representation of the components that build by summation the prediction for  $y(k + 2|k)$ , based on the truncated step response model.

Considering  $m = p = N = n = 6$ , the output predictions for all future time moments,  $(k + l)T$ ,  $l = 1, \dots, 6$ , depend on known past input moves  $\{\Delta u(k - 1), \Delta u(k - 2), \dots, \Delta u(k - 5)\}$ , on past input  $u(k - 6)$  and future and present (**unknown**) input variable moves  $\{\Delta u(k), \Delta u(k + 1), \dots, \Delta u(k + 5)\}$ . They may be computed in a similar manner as presented before. At the very core of the prediction computation stays the convolution property of the linear time invariant systems.

When state-space models are used for prediction, the predicted state and output variables, both for SISO and MIMO case, emerge from the state-space model presented in eq. (2.16). They have the following form:

$$\mathbf{x}(k + 1|k) = \mathbf{A}\mathbf{x}(k|k) + \mathbf{B}u(k) + \mathbf{K}e(k) \quad (2.36)$$

$$\mathbf{y}(k+1|k) = \mathbf{C}\mathbf{x}(k+1|k) + \mathbf{e}(k+1|k) \quad (2.37)$$

where the prediction of the disturbance will be considered  $\mathbf{e}(k+l|k) = 0$ .

Based on eq. (2.36), the state prediction  $\mathbf{x}(k+l|k)$  for future time steps,  $l > 0$ , may be obtained as

$$\mathbf{x}(k+l|k) = \mathbf{A}^k \mathbf{x}(k) + \sum_{i=1}^l \mathbf{A}^{i-1} \mathbf{B} \Delta \mathbf{u}(k+l-i) + \mathbf{A}^{k-1} \mathbf{K} \mathbf{e}(kT) \quad (2.38)$$

and the predicted output  $\mathbf{y}(k+l|k)$  as

$$\mathbf{y}(k+l|k) = \mathbf{C}\mathbf{x}(k+l|k) = \mathbf{C} \left[ \mathbf{A}^k \mathbf{x}(k) + \sum_{i=1}^l \mathbf{A}^{i-1} \mathbf{B} \Delta \mathbf{u}(k+l-i) + \mathbf{A}^{k-1} \mathbf{K} \mathbf{e}(kT) \right] \quad (2.39)$$

The corresponding predictor matrix  $\mathbf{M}$  and the free-run term  $\hat{\mathbf{y}}_0$  from eq. (2.25) may be obtained as

$$\mathbf{M} = \begin{bmatrix} \mathbf{CB} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CAB} & \mathbf{CB} & \cdots & \mathbf{0} \\ \vdots & & & \ddots & \vdots \\ \mathbf{CA}^{N-1}\mathbf{B} & \mathbf{CA}^{N-2}\mathbf{B} & \mathbf{CA}^{N-3}\mathbf{B} & \cdots & \mathbf{CB} \end{bmatrix} \quad (2.40)$$

$$\mathbf{y}_0(k) = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^N \end{bmatrix} \mathbf{x}(k|k-1) + \begin{bmatrix} \mathbf{CK} \\ \mathbf{CAK} \\ \vdots \\ \mathbf{CA}^{N-1}\mathbf{K} \end{bmatrix} [\mathbf{y}_m(k) - \mathbf{C}\mathbf{x}(k|k-1)] \quad (2.41)$$

It may be observed that the relationship between the step response model  $\mathbf{S}_i$  (of the model  $\mathbf{H}^*_p(q)$ ) and the  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  matrices is given by

$$\mathbf{S}_i = \mathbf{CA}^{i-1}\mathbf{B} \quad (2.42)$$

For the measured disturbances, feedforward MPC may be directly performed on the basis of predictions for the measured disturbance effect on the predicted output. Both truncated step/impulse response models and state space models may be used. They have forms similar with those already presented for the manipulated (input) variable.

A particular but commonly used form of the state-space formulation may be built using the step response parameters of the process  $s_1, \dots, s_n$  to the input

$u$  and the step response parameters of the process  $s_1^d, \dots, s_n^d$  to the measured disturbance  $d$  [2, 46].

$$\mathbf{Y}(k) = \mathbf{M}^* \mathbf{Y}(k-1) + \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix} \Delta u(k-1) + \begin{bmatrix} S_1^d \\ S_2^d \\ \vdots \\ S_n^d \end{bmatrix} \Delta d(k-1) \quad (2.43)$$

where

$$\mathbf{Y}(k) = [y(k) \ y(k+1) \ \dots \ y(k+n-1)]^T \quad (2.44)$$

and

$$\mathbf{M}^* = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \quad (2.45)$$

Based on this model, the output prediction, made at moment  $kT$  for the future  $N$  sampling time steps  $\mathbf{Y}(k+1|k)$ , is

$$\mathbf{Y}(k+1|k) = \mathbf{M}^{**} \mathbf{y}(k|k) + \mathbf{S}^u \Delta \mathbf{U}(k) + \mathbf{S}^d \Delta \mathbf{d}(k) \quad (2.46)$$

where  $\mathbf{S}^u = \mathbf{M}$ , eq. (2.31), is the matrix containing the step response parameters to the input variable  $u$  and  $\mathbf{S}^d$  consists of the step response parameters to the measured disturbance variable  $d$ :

$$\mathbf{S}^d = \begin{bmatrix} s_1^d & 0 & 0 & \dots & 0 \\ s_2^d & s_1^d & 0 & \dots & 0 \\ s_3^d & s_2^d & s_1^d & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_N^d & s_{N-1}^d & s_{N-2}^d & \dots & s_1^d \end{bmatrix} \quad (2.47)$$

The input move vector is defined by

$$\Delta \mathbf{U}(k) = [\Delta u(k) \ \Delta u(k+1) \ \dots \ \Delta u(k+N-1)]^T \quad (2.48)$$

and the matrix  $M^{**}$  is given by

$$M^{**} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \vdots & 1 & \cdots & 0 \end{bmatrix} \quad (2.49)$$

For the state-space-based prediction, it may be necessary to use a state observer to compute the states that are not directly measurable. This computation is based on the control and measured output variables, associated to the process model.

When CARIMA time series models are used for prediction, the predicted output variable emerges from solving two Diophantine equations [37–39]:

$$C(q) = E_j(q) A(q) \Delta(q) + q^j F_j(q) \quad (2.50)$$

$$E_j(q)B(q) = G_j(q)C(q) + q^j L_j(q) \quad (2.51)$$

The predictor matrix becomes

$$M = \begin{bmatrix} s_1^* & 0 & 0 & \cdots & 0 \\ s_2^* & s_1^* & 0 & \cdots & 0 \\ s_3^* & s_2^* & s_1^* & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_N^* & s_{N-1}^* & s_{N-2}^* & \cdots & s_1^* \end{bmatrix} \quad (2.52)$$

where  $s_i^*$  are the step response values of  $H_p^*(q)$ . They are given by the coefficients of the  $G_j(q)$  polynomial.

The free run output of eq. (2.25) becomes

$$\hat{y}_0(k) = \begin{bmatrix} L_1 u^*(q) \\ L_2 u^*(q) \\ L_3 u^*(q) \\ \vdots \\ L_N u^*(q) \end{bmatrix} + \begin{bmatrix} F_1 y^*(q) \\ F_2 y^*(q) \\ F_3 y^*(q) \\ \vdots \\ F_N y^*(q) \end{bmatrix} \quad (2.53)$$



where  $u^*(q)$  and  $y^*(q)$  are

$$u^*(q) = C^{-1}(q)\Delta u(k-1) \quad (2.54)$$

$$y^*(q) = C^{-1}(q) y(k) \quad (2.55)$$

## 2.6 Optimization for MPC

As presented in eqs. (2.1)–(2.3), the MPC algorithm relies on solving a receded optimization problem for computing, at each time step, the open-loop optimal control sequence, from which only the first control move is sent to the process input. Usually, the optimization index penalizes the sum of the square error (difference between the future reference trajectory and the predicted output) and the square input variable move, subject to constraints. The performance index is quadratic in the unknown input move sequence  $\{\Delta u(k), \Delta u(k+1), \dots, \Delta u(k+l), \dots, \Delta u(k+m-1)\}$ .

For the MIMO case, with the prediction equations (2.29)–(2.30), the optimization problem has the following formulation:

$$\min_{\Delta u(k) \dots \Delta u(k+m-1)} \{J(\mathbf{u}, k)\} = \min_{\Delta u(k) \dots \Delta u(k+m-1)} \left\{ \sum_{l=1}^p \mathbf{Y}_{wt} \|\mathbf{y}(k+l|k) - \mathbf{r}(k+l)\|^2 + \sum_{l=1}^m \mathbf{U}_{wt} \|\Delta \mathbf{u}(k+l-1)\|^2 \right\} \quad (2.56)$$

Although the *quadratic* performance index in the state space formulation is preferred and extensively applied, the linear performance index may be also used. Despite the fact that when using the linear performance index the optimization problem is solved in an easier manner, compared to the quadratic performance index, the first one shows some weaknesses, making it less attractive. Among them, the possibility of obtaining not-unique solutions, showing slow control action far away from origin and a lot of action close to origin, difficulty of tuning, and always activating some constraints may be mentioned. The use of quadratic performance index is harder to solve but leads to unique solution while showing lot of control action far away from origin and smooth action close to origin.

When the quadratic optimization MPC problem is not constrained, an analytical optimal solution may be obtained (good resemblance to the classical LQG regulation). However, when constraints are associated to the optimization problems (2.1–2.3), the optimal solution may be obtained using QP, sequential programming, or nonlinear programming (NLP) numerical solution techniques, as no analytical solution exists. QP refers to an optimization problem having quadratic performance index and linear constraints.

The constraints are always present in control applications, having equipment, technological, safety, environmental, or economic origin. MPC is highly appreciated for its capability to handle constraints in a systematic way, compared to traditional control, such as PID control. In the latter case, handling constraints is made by detuning the controller, with negative effect on the control performance, and by ad hoc solutions of using lead-lag and high-low logic blocks [1]. Process requirements may ask for the control task to keep a process variable as close as possible but below a specified constrained value. It is, for example, the case of the reactor riser or regenerator temperature control for the fluid catalytic cracking unit, aimed to keep the temperature in the very close vicinity of the desired temperature value (where optimal economic results are obtained), but below it and with no overshoot [5, 14, 49].

Constraints act on input (control), output, state, and input move variables, as presented by:

$$\begin{aligned} \mathbf{u}_{min} &\leq \mathbf{u}(k) \leq \mathbf{u}_{max} \\ \mathbf{y}_{min} &\leq \mathbf{y}(k) \leq \mathbf{y}_{max} \\ \mathbf{x}_{min} &\leq \mathbf{x}(k) \leq \mathbf{x}_{max} \\ \Delta \mathbf{u}_{min} &\leq \Delta \mathbf{u}(k) \leq \Delta \mathbf{u}_{max} \end{aligned} \quad (2.57)$$

for any moment of time  $kT$ .

Based on the output prediction equation (2.25) and the state prediction equation (2.38), the constraints on predicted inputs and states may be reformulated as constraints on the future control variables  $\hat{\mathbf{u}}(k)$ :

$$\hat{\mathbf{u}}_{min} \leq \mathbf{C}_u \hat{\mathbf{u}} \leq \hat{\mathbf{u}}_{max} \quad (2.58)$$

There are cases when the control variable has to be kept close to an optimal value, while it is used to counteract the effect of disturbances and finally be brought to this optimal value. This situation may be met when an optimization layer, working on the top of the MPC control layer, sends optimum values for the control variables. Typical circumstances may arise when a higher number of control variables are used, compared to a smaller number of controlled outputs. Handling this situation is possible by adding in the the performance index a supplementary term. This term penalizes the deviation of the control variable from its reference value, as presented by the following formulation of the optimization problem:

$$\begin{aligned} \min_{\Delta u(k) \dots \Delta u(k+m-1)} \{J(\mathbf{u}, k)\} &= \min_{\Delta u(k) \dots \Delta u(k+m-1)} \left\{ \sum_{l=1}^p \|Ywt_l[r(k+l) - y(k+l|k)]\|^2 + \right. \\ &\left. + \sum_{l=1}^m \|Uwt_l \Delta u(k+l-1)\|^2 + \sum_{l=1}^m \|Uwt_l^{ref} [u(k+l-1) - u^{ref}(k+l-1)]\|^2 \right\} \end{aligned} \quad (2.59)$$

where  $u^{ref}$  is the reference input variable and  $Uwt_l^{ref}$  is the input reference weighting factor.

The incentive of this approach over constrained optimization consists of putting less restriction on  $\Delta u(k)$  for obtaining the optimal control move. This will also enlarge the feasible region of the optimization problem provided that there are enough degrees of freedom to bring the value of the performance index to zero, i.e. the steady-state outputs be brought to their desired reference values.

Situations may arise when the constrained optimization problem does not have a feasible solution. In such cases, it might be suitable to allow the constraints to be violated, but in the less possible extent. This approach is called constraints softening and may be achieved by the help of a new variable  $\delta > 0$ , denoted as *slack variable*. This variable allows the enlargement of the hard constraints limits, as presented in the following inequalities:

$$\hat{\mathbf{u}}_{min} - \mathbf{U}_{min} \delta \leq \mathbf{C}_u \hat{\mathbf{u}} \leq \hat{\mathbf{u}}_{max} + \mathbf{U}_{max} \delta \quad (2.60)$$

where  $\mathbf{U}_{min}$  and  $\mathbf{U}_{max}$  are the minimum and maximum relaxation nonnegative constant vectors, respectively.

The slack variable  $\delta$  also becomes an unknown variable of the optimization problem and its optimal value will minimize the constraints violation. In this case, the form of the optimization problem for MPC gains a new term and gets the formulation:

$$\min_{\Delta u(k) \dots \Delta u(k+m-1), \delta} \{J(u, k)\} = \min_{\Delta u(k) \dots \Delta u(k+m-1)} \left\{ \sum_{l=1}^p \|\mathbf{Y}wt_l[r(k+l) - y(k+l|k)]\|^2 + \right. \\ \left. + \sum_{l=1}^m \|\mathbf{U}wt_l \Delta u(k+l-1)\|^2 + \sum_{l=1}^m \|\mathbf{U}wt_l^{ref}[u(k+l-1) - u^{ref}(k+l-1)]\|^2 + \rho_\delta \delta \right\} \quad (2.61)$$

where  $\rho_\delta$  is the weighting factor for penalizing the violation of constraints.

The analytical solution of the MPC unconstrained optimization problem may be obtained as solution of a least squares problem [37–39]. This solution can be demonstrated for the optimization problem described in eq. (2.2), considering  $P(q) = 1$ ,  $p_1 = 1$ ,  $p_2 = N$ ,  $m < N$ , and introducing the notation  $\hat{\mathbf{r}}(k)$  for the future reference vector:

$$\hat{\mathbf{r}}(k) = [r(k+1) \dots r(k+l) \dots r(k+N)]^T \quad (2.62)$$

The vector of future unknown values of the control variable move is described by

$$\hat{\mathbf{u}}(k) = [\Delta u(k) \Delta u(k+1) \dots \Delta u(k+l) \dots \Delta u(k+m-1)]^T \quad (2.63)$$

The MPC minimization problem may be sequentially formulated as:

$$\begin{aligned}
 \min_{\Delta u(k) \dots \Delta u(k+m-1)} \{J(u, k)\} &= \min_{\Delta u(k) \dots \Delta u(k+m-1)} \left\{ \sum_{l=1}^N \|\hat{r}(k+l) - y(k+l)\|^2 + \right. \\
 &+ \left. \sum_{l=1}^N \|\lambda^2 \Delta u(k+l-1)\|^2 \right\} = \min_{\Delta u(k) \dots \Delta u(k+m-1)} \{ \|\hat{r}(k) - \hat{y}(k)\|^2 + \lambda^2 \|\hat{u}(k)\|^2 \} = \\
 &= \min_{\Delta u(k) \dots \Delta u(k+m-1)} \{ [\hat{r}(k) - \hat{y}(k)]^T [\hat{r}(k) - \hat{y}(k)] + \lambda^2 \hat{u}^T(k) \hat{u}(k) \} = \\
 &= \min_{\Delta u(k) \dots \Delta u(k+m-1)} \{ \hat{u}^T(k) \hat{\mathbf{A}}^T \hat{\mathbf{A}} \hat{u}(k) + 2\hat{\mathbf{b}} \hat{\mathbf{A}} \hat{u}(k) + \hat{\mathbf{b}}^T \hat{\mathbf{b}} \}
 \end{aligned} \tag{2.64}$$

where the last form of the performance index was obtained by replacing  $\hat{y}(k)$  with its equal term presented by the general prediction equation (2.25) and  $\hat{\mathbf{A}}$  matrix and  $\hat{\mathbf{b}}$  vector have the following form:

$$\hat{\mathbf{A}} = \begin{bmatrix} -\mathbf{M} \\ \lambda \mathbf{I} \end{bmatrix} \tag{2.65}$$

and

$$\hat{\mathbf{b}} = \begin{bmatrix} \hat{r}(k) - \hat{y}_0(k) \\ \mathbf{0} \end{bmatrix} \tag{2.66}$$

The solution of this minimization problem in the least squares sense, having  $\hat{u}(k)$  vector as unknown, is given by

$$\hat{u}(k) = -(\hat{\mathbf{A}}^T \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}}^T \hat{\mathbf{b}} = (\mathbf{M}^T \mathbf{M} + \lambda^2 \mathbf{I})^{-1} \mathbf{M}^T (\hat{r}(k) - \hat{y}_0(k)) \tag{2.67}$$

### Example 2.5

The derivation of the solution presented in eq. (2.67) emerges from solving a linear system of equations having the form:



$$\mathbf{Ax} + \mathbf{b} = \mathbf{0} \tag{2.68}$$

where is  $\mathbf{x}$  the vector of unknown ( $\dim(\mathbf{x}) = m \times 1$ ),  $\mathbf{A}$  is a matrix of constant values ( $\dim(\mathbf{A}) = N \times m$ ,  $\text{rank}(\mathbf{A}) = m$ ), and  $\mathbf{b}$  is a vector of constant free terms ( $\dim(\mathbf{b}) = N \times 1$ ,  $m < N$ ). As the number of unknown of the linear system is higher than the number of equations, the eq. (2.68) may only be solved in the least squares sense, i.e. by minimizing the square of the residue  $\boldsymbol{\rho} = \mathbf{b} + \mathbf{Ax}$ :

$$\min_{\mathbf{x}} \{ \boldsymbol{\rho}^T \boldsymbol{\rho} \} = \min_{\mathbf{x}} \{ (\mathbf{Ax} + \mathbf{b})^T (\mathbf{Ax} + \mathbf{b}) \} \tag{2.69}$$

The least squares solution may be found by computing the derivative of the minimization index from eq. (2.69) and making it equal to zero. The solution is obtained successively in the following:

$$\begin{aligned}\frac{d(\boldsymbol{\rho}^T \boldsymbol{\rho})}{d\mathbf{x}} &= 2 \left( \frac{d\boldsymbol{\rho}}{d\mathbf{x}} \right)^T \boldsymbol{\rho} = 2\mathbf{A}^T(\mathbf{A}\mathbf{x} + \mathbf{b}) \\ 2\mathbf{A}^T(\mathbf{A}\mathbf{x} + \mathbf{b}) &= \mathbf{0} \\ \mathbf{A}^T\mathbf{A}\mathbf{x} &= -\mathbf{A}^T\mathbf{b} \\ \mathbf{x} &= -(\mathbf{A}^T\mathbf{A})^{-1} \mathbf{A}^T\mathbf{b}\end{aligned}\quad (2.70)$$

Note that the second derivative of the residue is positive:

$$\frac{d^2(\boldsymbol{\rho}^T \boldsymbol{\rho})}{d\mathbf{x}^2} = 2 \mathbf{A} \mathbf{A}^T \geq \mathbf{0} \quad (2.71)$$

which makes the optimal value to be a minimum.

For the MIMO case of the optimization problem presented in eq. (2.56), i.e. the optimal control sequence becomes:

$$\hat{\mathbf{u}}(k) = (\mathbf{M}^T \mathbf{Y}_{wt}^T \mathbf{Y}_{wt} \mathbf{M} + \mathbf{U}_{wt}^T \mathbf{U}_{wt})^{-1} \mathbf{M}^T \mathbf{Y}_{wt}^T \mathbf{Y}_{wt} (\hat{\mathbf{r}}(k) - \hat{\mathbf{y}}_0(k)) \quad (2.72)$$

The forms of the optimal MPC control moves, presented in eqs. (2.67) and (2.72), may be reduced to the following well-known form:

$$\hat{\mathbf{u}}(k) = \mathbf{K}_{MPC} \mathbf{E}(k+1|k) \quad (2.73)$$

where  $\mathbf{K}_{MPC}$  is the controller gain matrix and  $\mathbf{E}(k+1|k)$  the vector of future predicted errors for zero future manipulated variable moves (future error produced by the *free run output*).

Due to the receding horizon control approach, only the first of the control variable move:

$$\Delta \mathbf{u}(k) = [1 \ 0 \ 0 \ \dots \ 0] \hat{\mathbf{u}}(k) \quad (2.74)$$

is extracted from the optimal control vector and sent to the controlled process. The resulting explicit control law is linear and time invariant, provided the weighting factors are also time invariant.

## 2.7 MPC tuning

Tuning the MPC controller is not straightforward, but some general guidelines may be depicted to provide desired reference tracking, no steady-state offset and disturbance rejection [2, 39]. MPC has good inherent robustness to model-pant

mismatch. Meanwhile, there are no simple mathematical conditions for ensuring stability.

There are several MPC tuning parameters [40, 47]. They are: sampling time  $T$ , model horizon  $n$ , prediction horizon  $p$  (including both  $p_1$  and  $p_2 = n$  parameters of the MPC formulation described in eq. (2.2)), control horizon  $m$ , penalty weighting factor  $\lambda$  or matrices  $\mathbf{Y}_{wt}$  and  $\mathbf{U}_{wt}$ , and filters.

The sampling period  $T$ : must be chosen such as to describe the dynamic behavior without losing relevant quick process variable changes, but at the same time to avoid overloading the numerical computation implied by finding the solution of the constrained optimization problem [48]. It is the Shannon theorem that states the compromise for selecting the best value of the sampling time, i.e. half of the period corresponding to the highest considerable angular frequency  $\omega_s$  present in the harmonic decomposition of the sampled signal  $T \leq \pi/\omega_s$ . This consideration may be reduced to the very simple rule  $T = 0.1(\tau_d + T_d)$ , where  $\tau_d$  and  $T_d$  are the pure time delay (dead time) and the dominant time constant of the open-loop system.

Model horizon  $n$ : This should be selected such as  $nT$  exceeds 95% of the open-loop system settling time. It may be also recommended the value for  $n$  such as the last considered value of the impulse response parameter  $h_n$  is of the order of magnitude comparable to the measurement error for the output variable. Typical values may vary between 20 and 70, but it may also depend on available computation resources.

Prediction horizon  $p$ : The product of the prediction horizon and sampling time  $pT$  should cover the time necessary for the closed-loop system to achieve steady state. Typical values vary between 20 and 30. Short prediction horizon produces large control variable moves ending in instability. Long prediction horizon produces less control variable moves and slower response. There is a critical minimum horizon length to achieve a stable closed-loop system; the setting  $p = n + m$  is suggested. The minimum horizon  $p_1 T$  of the performance index (2.2) should be chosen one sampling step larger than the pure time delay  $p_1 = \tau_d + 1$ .

Control horizon  $m$ : This should be chosen about one-fourth to one-third of the prediction horizon. Typical values vary between 1 and 4. Short control horizon performs good control due to smaller control variable moves and reduced computation resources. When  $m$  is increased the control variable moves become larger, with less robustness and increased computational load, having also an increased degree of freedom in computing the control moves. For the case of using the CARIMA prediction models (or equivalents), the control horizon may be chosen equal to the order of polynomial  $A(q)$ .

Penalty weighting factor  $\lambda$ : This should be chosen as small as possible (but positive). For  $\lambda = 0$  the control action is not penalized. Increasing  $\lambda$  will make the control action less aggressive.

Penalty weighting matrices  $\mathbf{Y}_{wt}$  and  $\mathbf{U}_{wt}$ : These are usually diagonal, positive definite matrices. The values of the diagonal elements in the  $\mathbf{Y}_{wt}$  matrix are measures of the importance of the control effort assigned for each of the controlled variables

(MIMO case). As one element of  $\mathbf{Y}_{wt}$  is increased, the deviation from reference (set-point) value of the corresponding controlled variable is decreased. Adding in the performance index the term for penalizing the movement of the control (manipulated) variables, weighted by  $\mathbf{U}_{wt}$ , reduces the excessive manipulated variable move. For this reason  $\mathbf{U}_{wt}$  is also called as “move suppression factor”. Increasing the value of one element from  $\mathbf{U}_{wt}$  decreases the corresponding manipulated variable change, with effect on degradation of the control performance but increasing robustness. The relative magnitudes of  $\mathbf{Y}_{wt}$  and  $\mathbf{U}_{wt}$  have to be considered, as the magnitudes for the controlled and manipulated variables may be of different order of magnitude. The weighting matrices  $\mathbf{Y}_{wt}$  and  $\mathbf{U}_{wt}$  may be considered nondiagonal, as interactions between controlled variables are important, but tuning their values is not straightforward. The weighting matrices  $\mathbf{Y}_{wt}$  and  $\mathbf{U}_{wt}$  may be considered time-varying when control performance has to be dynamically changed, but choosing their appropriate time-dependent functions is difficult. A possible first choice for choosing the values of the weighting matrices is selecting each of their diagonal elements as the inverse of the maximum allowed change in the corresponding output or input variables. These initial guess values should be refined by an iterative simulation procedure [49].

**Feedback filter:** Introducing a suitable filter on the feedback signal may provide good disturbance rejection. The  $P(q)$  polynomial of the performance index formulation (2.2) may be also used to obtain desired closed-loop control performance by choosing the location of its poles.

Scaled dynamic sensitivity of the MPC performance index with respect to the tuning parameters may reveal their importance on the control performance, suggesting tuning quantitative measures [50].

## 2.8 MPC stability

Development of the necessary and sufficient conditions for ensuring linear MPC stability is not a trivial task. Most of the developed theoretical conditions for guaranteeing stability are sufficient conditions. One of the most important properties MPC algorithm must also satisfy is feasibility, i.e. capability to find solution for the optimization problem.

During the last two decades, researchers have focused in their studies on two important ways for guaranteeing stability and feasibility, with significant developments [51]. The first one addresses the development of conditions imposed on the terminal performance function and on supplementary constraints for the terminal states or outputs [39]. *Terminal* values refer to values at the end of the prediction or control horizon. The second direction has been aimed to find conditions for guaranteeing stability by using sufficiently large (infinite) prediction and control horizons.

Infinite control and prediction horizons,  $m = p = \infty$ , make unconstrained MPC algorithm be similar to the optimal LQ solution. But handling the constraints on an infinite horizon formulation is extremely difficult (or even impossible), especially from the computation effort point of view. One potential approach allows the prediction horizon be infinite but makes the control horizon be finite. This way, constraints handling becomes a finite-dimensional problem. Conditions may be developed to obtain the closed-loop stability for monotonic performance index.

The terminal point equality constraint adds a supplementary constraint to the constrained MPC by imposing the condition that the controlled output reaches the reference value or the controlled state reaches the origin (for the state-space MPC performance index formulation) at the end of the prediction horizon:

$$\begin{aligned} \mathbf{y}(k+p+j) &= \mathbf{r}(k+p), \quad j = 1, \dots, N \\ \mathbf{x}(k+p+j) &= \mathbf{0}, \quad j = 1, \dots, N \end{aligned} \quad (2.75)$$

Such sufficient conditions for guaranteeing asymptotic stability may exhibit feasibility problems if the control horizon is also small, due to the limited degrees of freedom. One solution is to replace the constraint on the terminal point with the constraint of keeping the terminal outputs or states inside a terminal region  $\Omega$ , where stability is fulfilled.

$$\begin{aligned} \mathbf{y}(k+p+j) &\in \Omega_y, \quad j = 1, \dots, N \\ \mathbf{x}(k+p+j) &\in \Omega_x, \quad j = 1, \dots, N \end{aligned} \quad (2.76)$$

A terminal penalty term,  $E(\mathbf{x}(k+p))$ , may also be used for obtaining stability by adding it to the optimization performance index:

$$\begin{aligned} \min_{\Delta \mathbf{u}(k) \dots \Delta \mathbf{u}(k+m-1)} \{J(\mathbf{u}, k)\} &= \min_{\Delta \mathbf{u}(k) \dots \Delta \mathbf{u}(k+m-1)} \left\{ \sum_{l=1}^p \mathbf{Q} \|\mathbf{x}(k+l|k)\|^2 + \right. \\ &\quad \left. + \sum_{l=1}^m \mathbf{R} \|\Delta \mathbf{u}(k+l-1)\|^2 + E(\mathbf{x}(k+p)) \right\} \end{aligned} \quad (2.77)$$

Contraction constraints can also offer guarantees for MPC algorithm stability. Such a contraction constrained has the form:

$$\|\mathbf{x}(k_0+p|k)\| \leq \varepsilon \|\mathbf{x}(k_0)\| \quad (2.78)$$

where  $\varepsilon$  is a less than one but positive contraction factor [52].



Recursive feasibility and stability conditions have been developed by researchers [53–54], but their practical application is still lacking due to the mathematical complexity [26].

## 2.9 Nonlinear MPC

During the last two decades, a large effort has been devoted both by academia and industry to develop and apply the Nonlinear Model Predictive Control NMPC techniques [54–62]. Significant results have been obtained but this MPC research of top interest is yet a challenging subject. Although the linear MPC does have well-developed methodologies for obtaining the desired control performance while satisfying the stability requirement, the NMPC techniques do not have such general and mature construction.

Essentially, the NMPC techniques are striving to compute the control law as solution of a receding open-loop optimization problem using the predictions of the output or state variable based on the nonlinear model of the plant. Founded on linearized process models, the linear MPC can handle control of nonlinear processes, especially when the goal is to keep the constrained process variables close to the desired nominal steady state points. In this case the use of a linear model for predictions, associated to a quadratic performance index, result in a convex QP optimization problem which can be solved online with available numerical algorithms. The need for the NMPC becomes necessary when the control task asks for tracking a changing reference which implies the covering of a pronounced nonlinearity domain. Typical case is the batch, start-up or shut down operation. Another set of control instances asking for NMPC is the disturbance rejection control for processes with strong nonlinearities and important changes of the disturbances.

There are different types of models that may be used in the NMPC framework. The most valued are the mechanistic (first principle) models, but statistical black box models are also much appreciated. The first category of models takes advantage of their sound physical meaning and capacity to extrapolate beyond the available process measured data or any particular known operating point. The second category allows the identification of the model from process measurements and this may be performed with less time and human resources compared to the former category. But black box models lose the physical interpretation and make extrapolation questionable. However, the NMPC applications are still dependable in a large extent to the models emerged from process identification.

From the time variable point of view, both continuous and discrete time models are used in NMPC. Due to the only on digital computer implementation of the NMPC applications and despite to the difficulty of transforming the continuous time nonlinear models in their corresponding discrete time form (unlike the continuous time linear models case), the discrete nonlinear model formulation is preferred. From the continuous or discrete time type of the signals the theoretical approach for NMPC has developed several methodologies for the association between the model used and the control signal obtained. They

are: continuous-time model with continuous control signal, continuous-time model with discrete-time control signal and discrete-time model with discrete-time control signal formulations [39].

The mathematical formulation of the NMPC optimization problem with associated constraints, for the discrete-time model with discrete-time control signal is presented in the following:

$$\min_{\mathbf{u}(k) \dots \mathbf{u}(k+m-1)} \{J(\mathbf{u}, k)\} = \min_{\mathbf{u}(k) \dots \mathbf{u}(k+m-1)} \left\{ \sum_{l=1}^p \|\mathbf{Y}_{wt}[\mathbf{r}(k+l) - \mathbf{y}(k+l|k)]\|^2 + \sum_{l=1}^m \|\mathbf{U}_{wt}\mathbf{u}(k+l-1)\|^2 \right\} \quad (2.79)$$

subject to

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}, \mathbf{u}, k) \quad (2.80)$$

$$\mathbf{y}(k) = \mathbf{g}(\mathbf{x}, k) \quad (2.81)$$

$$\mathbf{g}_1(\mathbf{x}) = \mathbf{0} \quad (2.82)$$

$$\mathbf{x}(k) = \mathbf{x}_{est}(k) \quad (2.83)$$

$$\mathbf{u}_{min}(k+l) \leq \mathbf{u}(k+l) \leq \mathbf{u}_{max}(k+l) \quad (2.84)$$

$$\mathbf{u}(k+l-1) - \Delta\mathbf{u}_{max} \leq \mathbf{u}(k+l) \leq \mathbf{u}(k+l-1) + \Delta\mathbf{u}_{max} \quad (2.85)$$

$$\mathbf{u}(k+l) = \mathbf{u}(k+m-1), l = m-1, \dots, p \quad (2.86)$$

$$\mathbf{x}_{min}(k+l) \leq \mathbf{x}(k+l) \leq \mathbf{x}_{max}(k+l) \quad (2.87)$$

$$\mathbf{y}_{min}(k+l) \leq \mathbf{y}(k+l) \leq \mathbf{y}_{max}(k+l) \quad (2.88)$$

Equations (2.80) and (2.81) describe the nonlinear discrete state-space model, where both  $\mathbf{f}$  and  $\mathbf{g}$  functions are nonlinear. Equation (2.82) reveals the nonlinear algebraic equations of the model. Equation (2.83) shows the use of the estimated state in the discrete state-space model, as initial state. The inequality constraints (2.84) and (2.85) correspond to the constraints on the control variable and the control variable moves. The additional constraint on the control variable (2.86) enforces the condition to keep the control variable values unchanged for the time interval between the end of the input horizon and the end of the prediction horizon. Constraints from eqs. (2.87) and (2.88) correspond to the constraints on the state and the output variables.

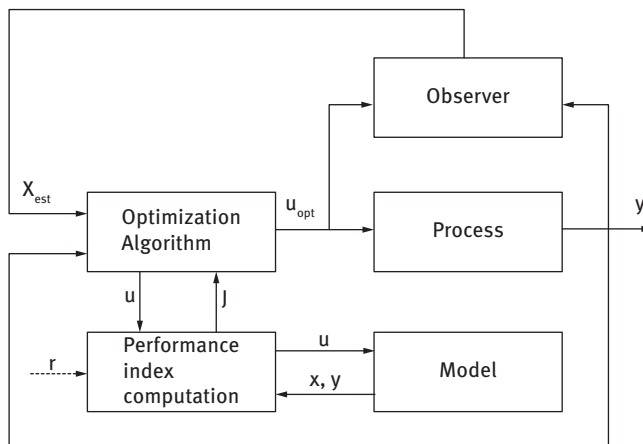
For the case of NMPC presented in eqs. (2.79) to (2.88), where discrete time models are used, solving the constrained optimization is simplified compared to NMPC using

continuous models due to the fact that solution of the optimization problem might be performed without the need in numerical integration for finding the states associated to the next moment of time. Alternation of the optimal control variables computation, followed by substitution, makes the algorithm of finding the solution to the constrained optimization less demanding. Anyway, even for the case of using the continuous models the solving of the model differential equations is also based on numerical approach, i.e. the discretization is inherently performed but the computational effort is increased.

Solving the nonlinear programming (NLP) problem described by eqs. (2.79) to (2.88) is generally non convex and requires good NLP software for finding the online solution. There are some approaches for finding solution to the NLP of NMPC [2].

One simple approach is to use a *linearized form of the model* equations around an operating point and successively update the linearizing as the operating point is changing or on a time scheduling basis. If the operating point cannot be directly found by measurements, it may be the estimated from available measurements. This approach transforms NPMPC into linear MPC and for the latter the QP may be used for obtaining the solution if a quadratic performance index is considered. An improved successive linearizing approach consists of using the linearized model for computing only the component of the predictions due to future control moves and the nonlinear model for computing the prediction component due the past (known) control moves.

Another approach for solving the NPMPC optimization problem is the *sequential algorithm* [2, 63, 64]. This is a two-step method where sequential model solution and optimization are performed. It is supposed that a continuous model is used for prediction and it consists of a set of ordinary differential equations (ODEs). The control inputs are computed by the NLP solver and subsequently the ODE numerical solver computes the solution of the model equations. A schematic representation of this algorithm is presented in Fig. 2.16.



**Fig. 2.16:** Schematic representation of the sequential NMPC algorithm.

A sequence of input control is first considered and the system of model ODEs is solved to get the states and outputs. With these computed values, the performance index is evaluated and the optimization problem is solved by establishing a new input and improved control sequence. This succession of optimization and model solution is performed up to the finding of the optimal control input sequence.

The *simultaneous algorithm* is another solving approach to the NPMC optimization problem [2, 65]. It is schematically presented in Fig. 2.17.

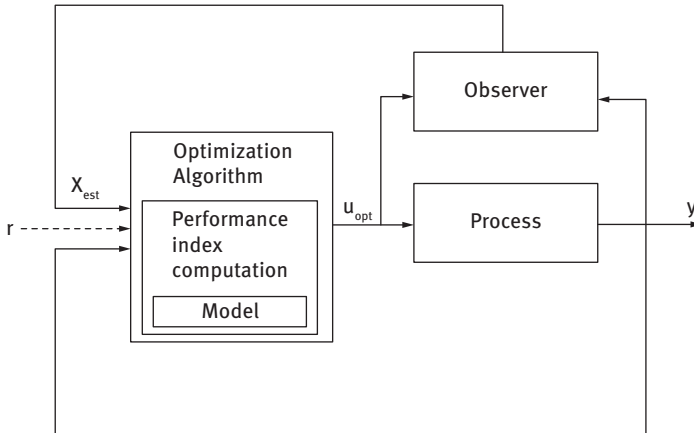


Fig. 2.17: Schematic representation of the simultaneous NMPC algorithm.

In this approach, the model equations, in discretized form, are considered as constraints to the optimization problem. An NLP solver is used to solve the constrained optimization problem.

Further research is devoted today to extend the MPC theory and applications. Some of the directions for these developments are MPC for large-scale systems, either with distributed or hierarchical configurations, MPC for rapid dynamic systems, MPC for uncertain systems, stochastic MPC, adaptive MPC, embedded MPC, low cost MPC, economic MPC, output MPC, and nonlinear MPC.

## References

- [1] Qin, S.J., Badgwell, T.A., *A survey of industrial model predictive control technology*, Control Engineering Practice, 11(7), 733–764, 2003.
- [2] Agachi, P.S., Nagy Z.K., Cristea, V.M., Imre-Lucaci A., *Model Based Control – Case Studies in Process Engineering*, Wiley-VCH, Weinheim, 2006.
- [3] Prett, D.M., Gillette, R.D., *Optimization and Constrained Multivariable Control of a Catalytic Cracking Unit*, AIChE National Meeting, Houston, TX, 1979.
- [4] Cutler, C.R., Ramaker, B.L., *Dynamic matrix control—a computer control algorithm*, in Proceedings of the joint automatic control conference, 1980.

- [5] Cristea, M.V., Agachi, S.P., Marinoiu, M.V., *Simulation and model predictive control of a UOP fluid catalytic cracking unit*, Chemical Engineering and Processing, 42, 67–91, 2003.
- [6] Wang, W.L., Rivera, D.E., Kempf, K.G., *Model predictive control strategies for supply chain management in semiconductor manufacturing*, International Journal of Production, Economics, 107(1), 56–77, 2007.
- [7] Nagy, Z., Agachi, S., *Model predictive control of a PVC batch reactor*, Computers and Chemical Engineering, 21(6), 571–591, 1997.
- [8] Van Brempt, W., Backx, T., Ludlage, J., Van Overschee, P., De Moor, B., Tousain, R., *A high performance model predictive controller: application on a polyethylene gas phase reactor*, Control Engineering Practice, 9, 829–835, 2001.
- [9] Wright, G.T., Edgar, T.F., *Nonlinear model predictive control of a fixed-bed water-gas shift reactor: an experimental study*, Computers and Chemical Engineering, 18, 83–102, 1994.
- [10] Muske, K.R., Howse, J.W., Hansen, G.A., Cagliostro, D.J., *Model-based control of a thermal regenerator. Part 1: dynamic model*, Computers and Chemical Engineering, 24, 2519–2531, 2000.
- [11] Dufour P., Michaud D.J., Toure Y., Dhurjati P.S., *A partial differential equation model predictive control strategy, application to autoclave composite processing*, Computers and Chemical Engineering, 28(4): 545–556, 2004.
- [12] Dufour P., Toure Y., Blanc D., Laurent P., *On nonlinear distributed parameter model predictive control strategy: On-line calculation time reduction and application to an experimental drying process*, Computers and Chemical Engineering, 27(11), 1533–1542, 2003.
- [13] Cristea, V.M., Roman, R., Agachi, S.P., *Neural Networks Based Model Predictive Control of the Drying Process*, ESCAPE 13, Computer Aided Chemical Engineering 14, 1–4 June, Lappeenranta, Finland, 389–394, 2003.
- [14] Iancu, M., Cristea, M.V., Agachi, P.S., *Retrofit design of heat exchanger network of a fluid catalytic cracking plant and control based on MPC*, Computers and Chemical Engineering, 49, 205–216, 2013.
- [15] van Overloop P.J., Weijs S., Dijkstra S., *Multiple model predictive control on a drainage canal system*, Control Engineering Practice, 16(5), 531–540, 2008.
- [16] Gomez M., Rodellar J., Mantecon J.A., *Predictive control method for decentralized operation of irrigation canals*, Applied Mathematical Modelling, 26(11), 1039–1056, 2002.
- [17] Shen, W., Chen, X., Corriou, J.P., *Application of model predictive control to the BSM1 benchmark of wastewater treatment process*, Computers and Chemical Engineering, 32, 2849–2856, 2008.
- [18] Ostace, G.S., Cristea, V.M., Agachi, P.S., *Cost reduction of the wastewater treatment plant operation by MPC based on modified ASM1 with two-step nitrification/denitrification model*, Computers and Chemical Engineering, 15(11), 2469–2479, 2011.
- [19] Garcia-Gabin, W., Zambrano, D., Camacho, E.F., *Sliding mode predictive control of a solar air conditioning plant*, Control Engineering Practice, 17(6), 652–663, 2009.
- [20] Roca, L., Guzman, J.L., Normey-Rico, J.E., Berenguel, M., Yebra, L., *Robust constrained predictive feedback linearization controller in a solar desalination plant collector field*, Control Engineering Practice, 17(9), 1076–1088, 2009.
- [21] Asadi, B., Vahidi, A., *Predictive cruise control: utilizing upcoming traf c signal information for improving fuel economy and reducing trip time*, IEEE Transactions on Control Systems Technology, 19(3), 707–714, 2011.
- [22] Keviczky, T., Balas, G.J., *Receding horizon control of an F-16 aircraft: a comparative study*, Control Engineering Practice, 14(9), 1023–1033, 2006.
- [23] Alexis, K., Nikolakopoulos, G., Tzes, A., *Switching model predictive attitude control for a quadrotor helicopter subject to atmospheric disturbances*, Control Engineering Practice, 19(10), 1195–1207, 2011.

- [24] From, P.J., Gravdahl, J.T., Lillehagen, T., Abbeel, P., *Motion planning and control of robotic manipulators on seaborne platforms*, Control Engineering Practice, 19(8), 809–819, 2011.
- [25] Percival, M.W., Wang, Y., Grosman, B., Dassau, E., Zisser, H., Jovanovic, L., Doyle, F.J. III., *Development of a multiparametric model predictive control algorithm for insulin delivery in type 1 diabetes mellitus using clinical parameters*, Journal of Process Control, 21(3), 391–404, 2011.
- [26] Xi, Y.-G., Li, D.-W., Lin, S., *Model predictive control – status and challenges*, Acta Automatica Sinica, 39(3), 222–236, 2013.
- [27] Kalman, R.E., *Contributions to the theory of optimal control*, Bulletin de la Societe Mathematique de Mexicana, 5, 102–119, 1960.
- [28] Kalman, R.E., *A new approach to linear filtering and prediction problems*, Transactions of ASME, Journal of Basic Engineering, 87, 35–45, 1960.
- [29] Zadeh, L.A., Whalen, B.H., *On optimal control and linear programming*, IRE Transactions on Automatic Control, (7)4, 45, 1962.
- [30] Propoi, A.I., *Use of LP methods for synthesizing sampled-data automatic systems*, Automatic Remote Control, 24, 837, 1963.
- [31] Richalet, J.A., Rault, A., Testud, J.L., Papon, J., *Model predictive heuristic control: applications to an industrial process*, Automatica, 14, 413–428, 1978.
- [32] Prett, D.M., Gillette, R.D., *Optimization and constrained multivariable control of a catalytic cracking unit*, AIChE National Meeting, Houston, TX, 1979.
- [33] Cutler, C., Morshedi, A., Haydel, J., *An industrial perspective on advanced control*. In AIChE annual meeting, Washington, DC, October 1983.
- [34] Garcia, C.E., Morshedi, A.M., *Quadratic programming solution of dynamic matrix control (QDMC)*, Chemical Engineering Communications, 46, 73–87, 1986.
- [35] Grosdidier, P., Froisy, B., Hammann, M., *The IDCOM-M controller*, in T.J. McAvoy, Y. Arkun, E. Zafiriou (Eds.), Proceedings of the 1988 IFAC workshop on model based process control, 31–36, Pergamon Press, Oxford, 1988.
- [36] Froisy, J.B., Matsko, T., *IDCOM-M application to the Shell fundamental control problem*, AIChE Annual Meeting, November 1990.
- [37] Clarke, D.W., Mohtadi, C., Tuffs, P.S., *Generalized predictive control – part 1. The basic algorithm*, Automatica, 23(2), 137–148, 1987.
- [38] Clarke, D.W., Mohtadi, C., Tuffs P.S., *Generalized predictive control – part 2. Extensions and interpretations*, Automatica, 23(2), 149–160, 1987.
- [39] van den Boom, T.J.J., *Model Based Predictive Control*, LernModul 6. Swiss Society for Automatic Control, 1997.
- [40] Soeterboek, A.R.M., *Predictive Control-A unified approach*, Prentice Hall, New York, 1992.
- [41] Seborg, D.E., Edgar, T.F., Mellichamp, D.A., *Process Dynamic and Control*, John Wiley & Sons, 649–669, 1989.
- [42] Munske, K.R., Rawlings, J.B., *Model predictive control with linear models*, AIChE Journal, 39(2), 262–287, 1993.
- [43] Lee, J.H., Morari, M., Garcia, C.E., *State-space interpretation of model predictive control*, Automatica, 30(4), 707–717, 1994.
- [44] Balchen, J.G., Ljungquist, D., Strand, S., *State space predictive control*, Chemical Engineering Science, 1992, 47(4), 787–807, 1992.
- [45] Garcia, C.E., Prett, D.M., Morari, M., *Model predictive control: theory and practice – a survey*, Automatica, 25(3), 335–348, 1989.
- [46] Brosilow, C., Joseph, B., *Techniques of Model-Based Control*, Prentice Hall, New York, 2002.
- [47] Clarke, D.W., Mohtadi, C., *Properties of Generalized*, Automatica, 25(6), 859–875, 1989.

- [48] Zafiriou, E., Morari, M., *Design of robust digital controllers and sampling-time selection for SISO systems*, International Journal of Control, 44(3), 711–735, 1986.
- [49] Kalra, L., Georgakis, C., *Effect of process nonlinearity on the performance of linear model predictive controllers for the environmentally safe operation of a fluid catalytic cracking unit*, Industrial and Engineering Chemistry Research, 33, 3063–3069, 1994.
- [50] Cristea, M.V., Agachi, S.P., *Model predictive control of inferred variables and dynamic sensitivity analysis applied to MPC tuning*, Buletinul Universitatii “Petrol-Gaze” Ploiesti, 52(1), 52–57, 2000.
- [51] Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O.M., *Constrained model predictive control: stability and optimality*. Automatica, 36(6), 789–814, 2000.
- [52] Zheng, A., Morari, M., *Global Stabilization of Linear Discrete-Time Systems with Bounded Controls – A Model Predictive Control Approach*, ACC, 1994.
- [53] Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O.M., *Constrained model predictive control: stability and optimality*, Automatica, 36, 789–814, 2000.
- [54] Mayne, D.Q., *Model predictive control: recent developments and future promise*, Automatica, 50, 2967–2986, 2014.
- [55] Qin, S.J., Badgwell, T.A., In F. Allgower, A. Zheng (Eds.), *An overview of nonlinear model predictive control applications*, Birkhauser, Switzerland, 2000.
- [56] Bauer, M., Craig, I.K., *Economic assessment of advanced process control – a survey and framework*, Journal of Process Control, 18, 2–18, 2008.
- [57] Findeisen, R., Allgöwer, F., Biegler L.T., *Assessment and Future Directions of Nonlinear Model Predictive Control*, Springer, Berlin, 2007.
- [58] Magni, L., Raimondo, D.M., Allgöwer, F., *Nonlinear Model Predictive Control: Towards New Challenging Applications*, Springer-Verlag, Berlin, 2009.
- [59] Klatt, K.U., Marquadt, W., *Perspectives for process systems engineering | personal views from academia and industry*. Computers and Chemical Engineering, 33(3), 536–550, 2009.
- [60] Manenti, F., *Considerations on nonlinear model predictive control techniques*, Computers and Chemical Engineering, 35, 2491–2509, 2011.
- [61] Allgöwer, F., Findeisen, R., Nagy, Z.K., *Nonlinear model predictive control: From theory to application*, Journal of the Chinese Institute of Chemical Engineers, 35, 299–315, 2004.
- [62] Zavala, V.M., Biegler, L.T., *The advanced-step NMPC controller: optimality, stability and robustness*, Automatica, 45, 86–93, 2009.
- [63] Jang, S., Joseph, B., Mukai, H., *Control of constrained multivariable nonlinear processes using a two-stage approach*, Industrial and Engineering Chemistry Research, 26, 2106–2114, 1987.
- [64] Bequette, B.W., *Nonlinear predictive control using multi-rate sampling*, Canadian Journal of Chemical Engineering, 69, 136–143, 1991.
- [65] Patwardhan, A.A., Rawlings, J.B., Edgar, T.F., *Nonlinear model predictive control*, Chemical Engineering Communications, 87, 123–141, 1990.

## 3 Fuzzy control

### 3.1 Introduction

The control system is aimed to generate and send its control decisions to the controlled process such as its steady-state and dynamic behavior conforms to desired process performance. It is obvious that process characteristics are governing the design of the controller. Consequently, control performance is directly related to the capability of describing the process behavior and to the way process information is embedded in the controller. Neither the model creation nor the controller design tasks is trivial. Among the important reasons hindering these endeavors are, process complexity, incomplete knowledge, uncertainty of the description, or stochastic behavior of the process. The knowledge on the process behavior may have different forms, i.e. models built on analytical (first principle) or statistical basis. The controller design based on analytical models is preferred, but as complexity increases, the approach may become stiff, consumes more computing resources, becomes very sensitive to disturbances, becomes less robust, or accumulates errors. Statistical models may be an alternative for such cases because they may be built on heuristic methods. Accounting for the tolerance to incomplete determination and uncertainty may lead to the design of a controller able to fulfil the control task in a similar way humans do, i.e. not necessarily on the basis of very precise (mathematical) knowledge or representation (e.g. on ordinary or partially differential equations), but on a less precise evaluation that involves the use of some acting rules. Finally, this approximate approach may result in efficient control, able to overcome the problems mentioned before and to partially render the controller design a human way of acting. The fuzzy controller is a representative of this class. The fuzzy approach, for both modeling and controller design, is based on the fuzzy logic. Its main elements are presented in the next sections.

### 3.2 Fuzzy sets

In classical (Boolean) set theory, a large(r) set  $X$ , having a subset  $F$ ,  $F \subset X$ , has elements  $x \in X$  that either belong to the set  $F$ ,  $x \in F$ , or do not belong to it,  $x \notin F$ . The classical set  $F$  is denoted as a crisp set. As opposed to the crisp set  $F$ , a fuzzy set  $F$  has elements that may belong to the set  $F$ , in a more or less extent. The membership degree of an element to a fuzzy set is defined by a real value in the interval  $[0, 1]$  and shows how much truth is in the statement “ $x$  belongs to the set  $F$ ”. The membership value of 1 denotes total belonging and the membership value of 0 shows not at all belonging to the set  $F$ . It was in 1965 when Lotfi Zadeh introduced the notion of fuzzy set and opened the new horizon of the fuzzy logic research and applications [1].



**i** Example 3.1

Consider a set of metal balls that have been coated with white, black, and white-black combinations of paints. This is the large set  $X$ , of both classical and fuzzy approach. Consider the set  $F$  to be the set of gray balls.

The *classical (crisp) set* of gray balls consists of balls (elements) having been coated with a combination of 50% white paint and 50% black paint. All other balls, coated with other combinations of white-black paints or exclusively coated white or black paint, do not belong to the crisp set  $F$ . This approach of building the set  $F$  is a very accurate one, as it considers only balls coated with exactly equal white-black paint mixture. Other crisp sets may be defined on the large set  $X$  of balls, such as white set of balls  $F_W$ , black set of balls  $F_B$ , etc.

The *fuzzy set* of gray balls  $F$  consists of balls (elements) having been coated not only with a combination of 50% white paint and 50% black paint but also with different percentage of black-and-white paint. According to this approach, balls having been coated with varying proportion of white-black paint (e.g. combination of 80% white paint and 20% black paint or combination of 20% white paint and 80% black paint), also belong to the fuzzy set  $F$ . Furthermore, it may be stated that balls from this fuzzy set belong to the set  $F$  more or less, according to how much the combination of their white-black paints comes close to the 50% white paint and 50% black paint combination. For example,

- for a ball coated with a combination of 50% white paint and 50% black paint, it may be assigned total membership to the fuzzy set  $F$  (with associated membership value of 1),
- for a ball coated with a combination of 75% white paint and 25% black paint, it may be assigned partial membership to the fuzzy set  $F$  (with associated membership value of 0.5),
- for a ball coated with a combination of 25% white paint and 75% black paint, it may be assigned partial membership to the fuzzy set  $F$  (with associated membership value of 0.5),
- for a ball exclusively coated with white paint (100%) or black paint (100%), it may be assigned membership to the fuzzy set  $F$  with associated membership value of 0.

Other fuzzy sets may be defined on the large set  $X$  of balls, such as white set of balls  $F_W$  or black set of balls  $F_B$ . Either of them contains the balls coated with different combinations of white-black paints but have associated different membership values, proportional to the closeness of their color to the white or black paint.

**Definition:** A fuzzy set  $F$  (on a large set  $X$ , named universe of discourse) is represented by the set of pairs [2]:

$$F = \{(x, m_F(x)) | x \in X\} \quad (3.1)$$

where  $m_F$  is a characteristic function, named membership function of the fuzzy set  $F$ , defined by

$$m_F : X \rightarrow [0, 1] \quad (3.2)$$

As a result of the fuzzy set definition, the membership function assigns to every element  $x$  from  $F$  a real value  $m_F(x)$  which belongs to the interval  $[0, 1]$ . The value  $m_F(x)$  is denoted as the membership degree of the element  $x$  from the fuzzy set  $F$  and shows how much the element  $x$  belongs to the set  $F$ . The value of 0 denotes no membership (false) and the value of 1 shows total membership (true). Even though the fuzzy set is a time-invariant structure, it may be successfully used for the design of the fuzzy controller.

Note that for a crisp set  $F$ , such a characteristic function could be also associated but it would take values only from the set  $\{0, 1\}$ , i.e. if the element  $x$  belongs to the classical set  $F$  it will take the value of 1 and if the element  $x$  does not belong to the classical set  $F$  it will take the value of 0.

### 3.3 Typical membership functions of the fuzzy sets

For the description of the membership functions, a set of typical function forms are commonly used in applications, although there is no restriction about the form of their choice [3, 5].

Figure 3.1 shows the triangular membership function.

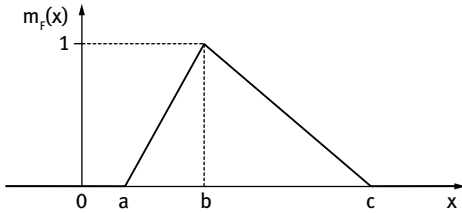


Fig. 3.1: Triangular membership function.

Its analytical form is

$$m_F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b \leq x \leq c \\ 0, & x > c \end{cases} \quad (3.3)$$

Figure 3.2 shows the trapezoidal membership function.

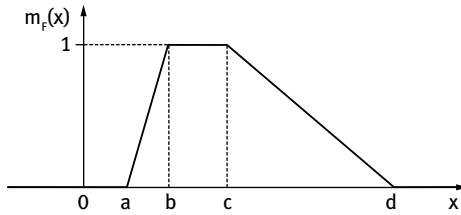


Fig. 3.2: Trapezoidal membership function.

Its analytical form is

$$m_F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ \frac{d-x}{d-c}, & c \leq x \leq d \\ 0, & x > d \end{cases} \quad (3.4)$$

Figure 3.3 shows saturation membership functions:

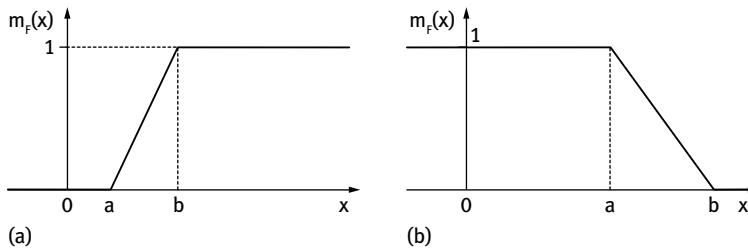


Fig. 3.3: High a) and Low b) Saturation membership functions.

Their analytical forms are

$$m_F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases} \quad (3.5)$$

$$m_F(x) = \begin{cases} 1, & x < a \\ 1 - \frac{x-a}{b-a}, & a \leq x \leq b \\ 0, & x > b \end{cases} \quad (3.6)$$

For the triangular, trapezoidal, and saturation membership functions, the parameters  $a$ ,  $b$ ,  $c$ , and  $d$  are constants, but they may be adjusted according to the needs. These membership functions are piecewise continuous functions.

A particular membership function form is the singleton membership function. Its shape, presented in Fig. 3.4, emerges from a Dirac impulse function with a finite width  $\epsilon$ .

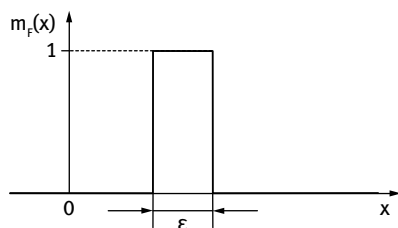


Fig. 3.4: Singleton membership function.

Smooth continuous membership functions are also used, such as sigmoidal saturation functions (including their difference) or Gauss bell-shaped function. The latter is presented in Fig. 3.5.

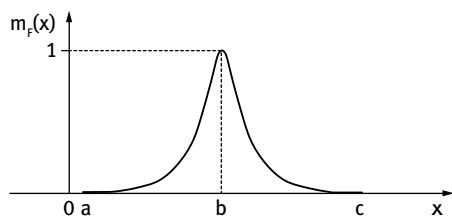


Fig. 3.5: Bell-shaped membership function.

Its equation is of the form

$$m_F(x) = e^{-\frac{(x-b)^2}{2(c-a)^2}} \quad (3.7)$$

for which, the restriction to the interval  $[a, c]$  is usually considered.

### Example 3.2

Consider the system of a mixing tank in which two inlet flows are entering, one of cold water and one of hot water. The tank outlet flow of mixed water feeds the downstream process with water at the desired temperature of  $T^o = 25^\circ\text{C}$ . The temperature in the tank is controlled either by the inlet flow of hot water or by the inlet flow of cold water. The process variable of first importance is the temperature in the tank. It may be considered that the tank temperature can be classified in three categories (ranges): *cold water* with temperature in the interval  $[18, 23]^\circ\text{C}$ , *warm water* with temperature in the



interval  $[23, 27]$  °C, and *hot water* with temperature in the interval  $[27, 32]$  °C. It is possible to associate, both crisp and fuzzy sets to the three water temperature categories. The cold, warm, and hot temperature crisp sets are presented in Fig. 3.6.

According to the crisp set approach, a water temperature of  $T^\circ = 23.3$  °C belongs to the warm water set and does not belong to the cold water set or to the hot water set.

One possible choice of fuzzy sets associated to the three water temperature categories are presented in Fig. 3.7.

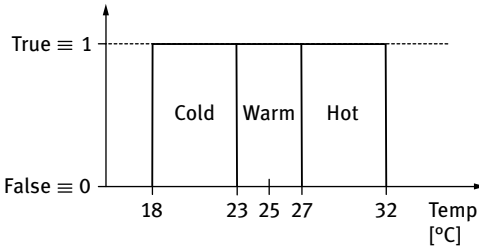


Fig. 3.6: Crisp water temperature sets.

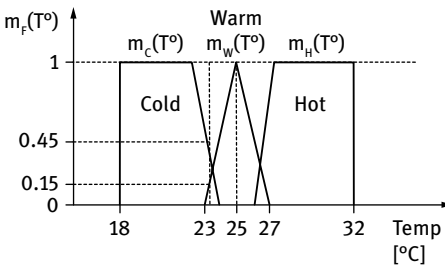


Fig. 3.7: Cold, Warm and Hot Water fuzzy water temperature sets.

The fuzzy sets cold and hot water have saturation forms (low, respectively high) while the warm water fuzzy set is of triangular form. All of them have as universe of discourse the temperature interval  $[18, 32]$  °C. The membership functions associated to the three fuzzy sets  $m_c(T^\circ)$ ,  $m_w(T^\circ)$ , and  $m_h(T^\circ)$  show a gradual transition for the membership of the elements belonging to adjacent sets. For example, the temperature element of sets  $T^\circ = 23.3$  °C belongs to cold water fuzzy set in the extent (trueness) described by the membership value  $m_c(T^\circ) = 0.45$ , while the same element belongs to warm water fuzzy set in the extent (trueness) described by the membership value  $m_w(T^\circ) = 0.15$ . This shows that the water temperature  $T^\circ = 23.3$  °C is considered to be cold in a larger extent (trueness) than warm. According to the way the fuzzy sets are defined, the same temperature element has a membership value of  $m_h(T^\circ) = 0$ , showing that it may not be considered as a hot temperature.

The triangular form of the warm water fuzzy set was selected for revealing a gradual measure of proximity to the temperature value of  $T^\circ = 25$  °C, which is of highest importance for the considered mixing process and its associated control system task of keeping the temperature to this setpoint value.

The control philosophy is simple and may be concentrated in the following linguistic statements: if the water temperature is cold, then increase the hot inlet flow; if the water temperature is hot, then decrease the hot inlet flow; if the water temperature is warm, then do not change the hot inlet flow (or gently increase or decrease the hot inlet flow, as the water temperature is slightly below or slightly above the desired value of  $T^o = 25^\circ\text{C}$ ).

As noticed from this example, for the temperature variable three fuzzy sets were defined. First, they are described by their names, denoted by linguistic variables: *cold*, *warm*, and *hot*. Second, the membership functions associated to these sets show the extent a certain temperature (element) may be considered to belong to each of the sets.

### 3.4 Operations with fuzzy sets

Operations with classical sets may be extended to the fuzzy sets. The latter will involve the membership functions associated to the fuzzy sets when operations with fuzzy sets are performed.

For two fuzzy sets  $M$  and  $N$ , defined on the universe of discourse  $F$  and having the membership functions  $m_M(x)$  and  $m_N(x)$ , the union  $M \cup N$ , intersection  $M \cap N$ , and complement  $C_M$  operations for  $x \in F$  are described by [3, 4]

$$m_{M \cup N}(x) = m_M(x) \vee m_N(x) = \max(m_M(x), m_N(x)) \quad (3.8)$$

$$m_{M \cap N}(x) = m_M(x) \wedge m_N(x) = \min(m_M(x), m_N(x)) \quad (3.9)$$

$$m_{C_M}(x) = 1 - m_M(x) \quad (3.10)$$

Other properties of the fuzzy set operations may be also developed. They are:

- the empty set  $\emptyset \subseteq F$ , which has

$$m_{\emptyset}(x) = 0 \quad (3.11)$$

- the total set, which is characterized by

$$m_F(x) = 1 \quad (3.12)$$

- the two fuzzy sets are equal,  $M = N$ , if and only if their membership functions are identical:

$$m_M(x) = m_N(x) \quad (3.13)$$

Most of the operations with the classical sets are also met for the fuzzy sets, except the excluded middle and the law of contraction, as presented in the following:

$$M \cup C_M \neq F \quad (3.14)$$

$$M \cap C_M \neq \emptyset \quad (3.15)$$

On the other side, two fuzzy specific operations may be performed:

- the algebraic product of two fuzzy sets,  $M \cdot N$ , where the membership function of the product set is defined by:

$$m_{M \cdot N}(x) = m_M(x) \cdot m_N(x) \quad (3.16)$$

- the algebraic sum of two fuzzy sets  $M + N$ , where the membership function of the sum set is:

$$m_{M+N}(x) = m_M(x) + m_N(x) - m_M(x) \cdot m_N(x) \quad (3.17)$$

The union, intersection, and complement operations with the temperature fuzzy set for cold water  $C$ , and warm water  $W$ , from Example 3.2, are presented in Figs. 3.8 to 3.10.

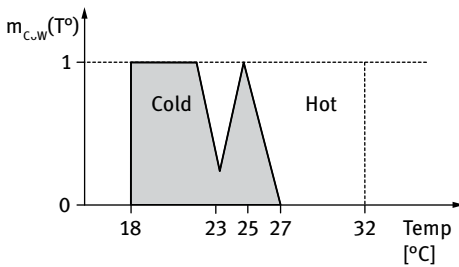


Fig. 3.8: Fuzzy union operation between cold and warm water temperature fuzzy sets.

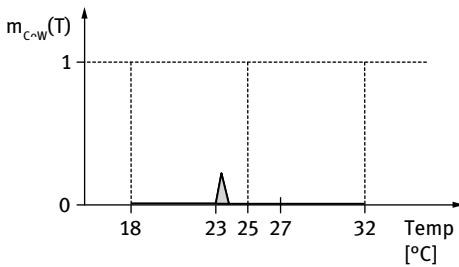


Fig. 3.9: Fuzzy intersection operation between cold and warm water temperature fuzzy sets.

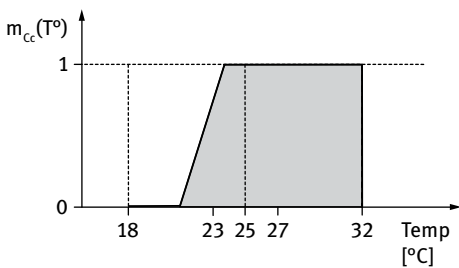


Fig. 3.10: Fuzzy complement operation of the cold water temperature fuzzy set.

Note that the results of the operations with fuzzy sets are fuzzy sets, too.

### 3.5 Fuzzy logic

Having its origin in the human way of reasoning, fuzzy logic is actually its formalization. Fuzzy logic is a generalization of the Boolean logic where the discrete approach with the only two true-false (1 and 0) values is replaced by a continuous approach. Its mathematical fundamentals may be found in the polyvalent logic of Lukasiewicz. A linguistic statement, i.e. a proposition  $Pp$ , may only have the *True* (1) or the *False* (0) evaluation (result) when classical predicate logic is considered [3, 4]. In fuzzy logic, a proposition  $Pp$  may have intermediate evaluations between *True* and *False*, having assigned a fuzzy set. Its membership function reveals the degree of truth of the proposition by a numerical value in the interval  $[0, 1]$ .

For one of the propositions mentioned in Example 3.2 related to the control philosophy, i.e. “if the water temperature is *cold* then ...”, the degree of trueness may be well described by the fuzzy set *cold water*. Water tank with a temperature of  $T^\circ = 19^\circ\text{C}$  may be very well considered as cold water (membership value of  $m_c(T^\circ) = 1$ ), while water tank with a temperature of  $T^\circ = 23^\circ\text{C}$  may be only “partially” considered as cold water (membership value of  $m_c(T^\circ) \approx 0.5$ ).

As presented in the previous paragraphs, we associate to the deterministic variables their fuzzy correspondents, i.e. the fuzzy variables. These fuzzy variables are linguistic variables. A linguistic variable uses words to characterize its values, as it is formulated in natural language. These words are described by fuzzy sets on the universe of discourse. Each word of the linguistic variable has associated a fuzzy set to show the membership measure of one *deterministic value* (element), from the universe of discourse, of belonging to the category described by that particular word. Equivalently, the fuzzy set associated to that word shows the degree of trueness of the linguistic statement formulated by the word of the linguistic variable. To conclude, the deterministic variable has as values scalars (or vectors) expressed by numbers, while the fuzzy variable has as values different linguistic attributes expressed by words (formalized as fuzzy sets).

Usually, the linguistic variable uses two words to describe it. The first one (primary term) has the same name with the deterministic variable it originates from and the second one (linguistic class) denotes the category of the primary term using adjectives or adverbs.

For Example 3.2, the deterministic variable “water temperature” has scalar values in the interval  $[18, 32]$  °C. To this deterministic variable, it was associated the fuzzy variable “water temperature”. The latter has the following values, described in words: “cold water temperature”, “warm water temperature”, and “hot water temperature”. The adjectives “cold”, “warm”, and “hot” denote the category of the primary term “temperature”. Each of the linguistic variables “cold water temperature”, “warm water temperature”, and “hot water temperature” have associated a fuzzy set. The linguistic statement “the water temperature is cold” is formalized as a fuzzy set. For a particular deterministic value of the variable “water temperature”, e.g. for  $T^\circ = 23.3^\circ\text{C}$ , the previously mentioned statement is true with a degree of trueness described by the membership value of the associated fuzzy set “cold water”, i.e.  $m_c(23.3^\circ\text{C}) = 0.45$ .



The transformation of a deterministic variable into its corresponding fuzzy variable (the latter having associated linguistic values) is denoted by *fuzzification*. The number of the linguistic (fuzzy) variables associated to a deterministic variable depends on the particular problem, but usually, a number of five linguistic values are sufficient. Typically, they are Very Small, Small, Mean, Large and Very Large.

For a fuzzy control system, the error deterministic variable (difference between the setpoint and the controlled variable) may have the corresponding fuzzy linguistic variables: Negative Large (NL), Negative Small (NS), Zero (Z), Positive Small (PS) and Positive Large (PL), as shown in Fig. 3.11.

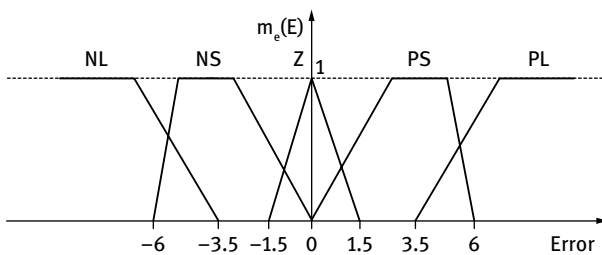


Fig. 3.11: Five fuzzy sets defined for the error variable of the fuzzy controller.

The human knowledge, represented in a linguistic way, is usually formulated in the form of *if-then* rules. It is in fact a transformation of the knowledge into a nonlinear (cause-effect) mapping. The fuzzy logic is based on information represented by fuzzy sets and performs fuzzy *if-then* rules on these linguistic variables. The fuzzy *if-then* rule is a conditional statement involving two propositions:

$$\mathbf{if} \{ \text{fuzzy proposition } Pa \} \mathbf{then} \{ \text{fuzzy proposition } Pc \} \quad (3.18)$$

The  $Pa$  proposition is denoted as *antecedent* and the  $Pc$  proposition as *consequent*. The linguistic formulated *if-then* rule makes a correlation between the premise/cause  $Pa$  and the consequent/effect  $Pc$ , building a relationship between them. According to the degree of trueness of the antecedent statement, the degree of trueness of the consequent statement will be satisfied. The mathematical correspondent for building this relationship is the *implication*. Although not similar, the fuzzy implication corresponds to the crisp composition of functions.

The fuzzy implication starts from the trueness assessment of one element  $x$  to be member of the fuzzy set  $P$ , and as a consequence, it follows by establishing the trueness of the elements  $y$  (of  $Q$ ) to belong to the fuzzy set  $Q$ . The result of a fuzzy implication is also a fuzzy set  $Q^*$ . The fuzzy set  $Q^*$  has the same linguistic attributes as the fuzzy variable  $Q$ . The fuzzy implication is usually referred as *fuzzy inference* [3, 4, 6].

Recalling the control philosophy presented in Example 3.2, the following fuzzy sets may be considered:

- (a)  $P$ : The water temperature is *cold*,
- (b)  $Q$ : Hot inlet flow is *increased*.

which are involved in the control rule “if the water temperature is *cold* then *increase* the hot inlet flow”. In this example there are three fuzzy sets for the fuzzy variable “Water temperature”, according to the linguistic attributes: *Cold*, *Warm*, and *Hot*. Additionally, there are three fuzzy sets for the fuzzy variable “Hot inlet flow”, according to the linguistic attributes: *Increase*, *Decrease* and *No change*. In this particular example  $x$  is the *water temperature* and  $y$  is the *hot inlet flow*.

As a result, the fuzzy implication  $P \rightarrow Q$  described by the rule “if the water temperature is *cold* then *increase* the hot inlet flow” is a new fuzzy set  $Q^*$  showing how much “hot inlet flow is *increased*”.

The membership function of the fuzzy set  $Q^*$  (fuzzy inference)  $m_{P \rightarrow Q}(x, y)$  is computed on the basis of the membership functions of the fuzzy sets  $P$ ,  $m_P(x)$ , and  $Q$ ,  $m_Q(y)$ . There are different ways of defining this fuzzy implication [3, 4]:

1. Mamdani inference:

$$m_{P \rightarrow Q}(x, y) = \min(m_P(x), m_Q(y)) \quad (3.19)$$

2. Tagaki-Sugeno-Kang (TSK) inference:

$$m_{P \rightarrow Q}(x, y) = g(x) \quad (3.20)$$

where  $g(x)$  is a function of the crisp input variable  $x$  (or variables). For the case of the *first-order Sugeno inference*, the function  $g(x_1, x_2, \dots, x_n)$  has a linear dependence on the input variables  $x_1, x_2, \dots, x_n$ :

$$m_{P \rightarrow Q}(x, y) = g(x) = p_0 + p_1x_1 + p_2x_2 + \dots + p_nx_n \quad (3.21)$$

3. Boolean inference:

$$m_{P \rightarrow Q}(x, y) = \max(1 - m_P(x), m_Q(y)) \quad (3.22)$$

4. Zadeh I inference:

$$m_{P \rightarrow Q}(x, y) = \min(1, 1 - m_P(x) + m_Q(y)) \quad (3.23)$$

5. Zadeh II inference:

$$m_{P \rightarrow Q}(x, y) = \min[\max(m_P(x), m_Q(y)), 1 - m_P(x)] \quad (3.24)$$

In applications, the most used are the Mamdani and Tagaki-Sugeno-Kang fuzzy logic inference methods.

The antecedent part of the fuzzy rule may have a single statement or a compound statement that merges several statements (premises). Each statement involves a

different fuzzy variable. They are connected in the antecedent proposition by the words *and* and *or*. The trueness of the antecedent is evaluated on the basis of the operations with fuzzy sets corresponding to the individual statements, i.e. union and/or intersection operators (as presented in the 3.4 paragraph).

To illustrate the combination of fuzzy variables, consider for Example 3.2, as additional input, the *rate of change* of the temperature in the tank. This crisp variable has a corresponding fuzzy variable *temperature rate of change* with three linguistic attributes: *Positive*, *Zero*, and *Negative*. The following *if-then* rule may be added for the control strategy of the temperature in the tank:

$$\begin{aligned} & \text{if } \{ \{ \text{the water temperature is Cold} \} \text{ and } \{ \text{the water temperature change is Positive} \} \} \\ & \text{then } \{ \text{the hot inlet flow is No change} \} \end{aligned} \quad (3.25)$$

The antecedent of this fuzzy rule implies the computation of the membership function as a compound statement, i.e. the intersection (*and* operator) of the fuzzy sets “the water temperature is cold” and “the water temperature change is positive”. This means the truth of the antecedent is computed as the *min* value of the involved fuzzy sets’ membership functions:

$$m_{C \cap \Delta+}(x) = m_C(x) \wedge m_{\Delta+}(x) = \min(m_C(x), m_{\Delta+}(x)) \quad (3.26)$$

where the membership functions for the fuzzy sets “the water temperature is cold” and “the water temperature change is positive” have been denoted by  $m_C(x)$  and respectively  $m_{\Delta+}(x)$ .

Every fuzzy system consists in a set of rules. As more than one fuzzy rule becomes active for a given set of antecedents, an *aggregation* process for the active rules is needed. There are two types of aggregation mechanisms. The first one is denoted by “and” aggregation and is determined by intersection of the individual rule consequents. The second one is named “or” aggregation and is determined by union of the individual rule consequents.

The last component of the fuzzy logic technique is the *defuzzifier*. This element transforms the result of the implications (aggregated consequent of the active set of rules), which are also fuzzy sets, into crisp (representative) values for the fuzzy variables (sets). This transformation is usually based on the *computation of the centroid C* (center-of-gravity) of the aggregated fuzzy set  $F$ , according to the following formula [4]:

$$C = \frac{\int_F x m_F(x) dx}{\int_F m_F(x) dx} \quad (3.27)$$

where the integral is computed over the universe of discourse.

It may be noticed that fuzzy logic may be successfully used for modeling applications, especially where first principle models are inoperable or difficult to develop due to process complexity, but for which human expertise or human knowledge is available under the form of linguistic assessments. Furthermore, fuzzy logic may be used to design control systems able to better cope with process uncertainty or incomplete description, compared to the analytical (equation based) approach, and be successful in processing information affected by noise.

The general structure of the fuzzy system that uses the fuzzy logic is presented in Fig. 3.12.

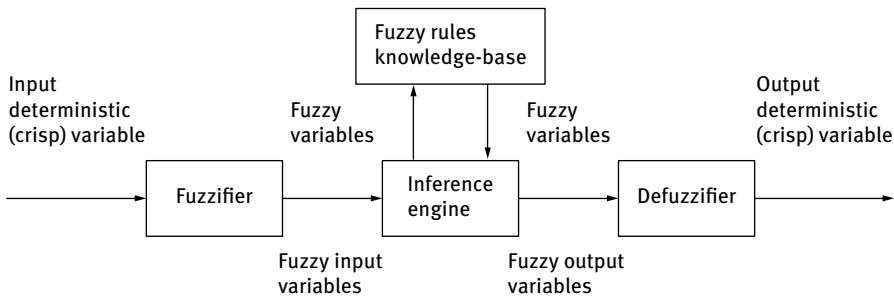


Fig. 3.12: General structure of a fuzzy system.

The general steps for developing a fuzzy system (model or controller) are [3]:

1. description of the knowledge base of the problem,
2. selection of the input-output (state) variables,
3. fuzzyfication of the crisp input and output variables,
4. building the fuzzy inference rules and setting the mechanism of aggregation of the rules,
5. defuzzyfication of the fuzzy output variables,
6. setting mechanisms for making fuzzy system adaptive.

### Example 3.3

Consider the example of a fuzzy controller designed to maintain the temperature in the mixing tank at the setpoint temperature value of  $T^{\circ} = 25^{\circ}\text{C}$ , as presented in Example 3.2. The controller has two crisp inputs: the error  $e$  (difference between the setpoint and the current temperature) and the differentiated error  $de/dt$ . The controller output is the mixing tank inlet flow rate of cold water.

The associated fuzzy variables for the considered crisp inputs and the fuzzy membership functions are presented in Fig. 3.13(a) and 3.13(b).



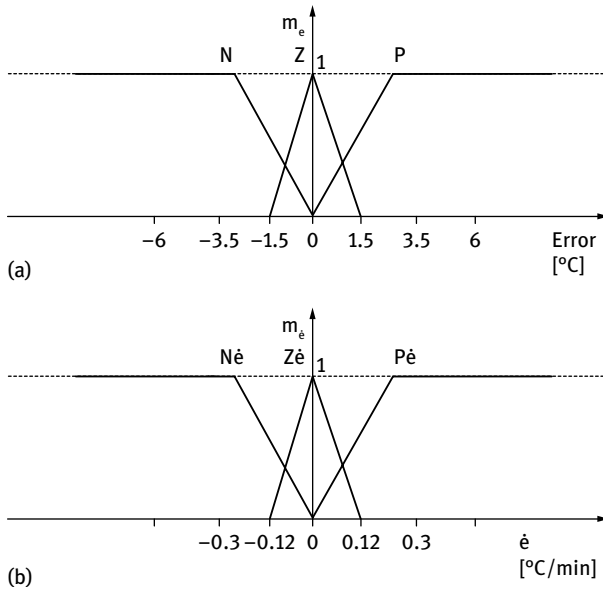


Fig. 3.13: Input fuzzy membership functions of the fuzzy controller (a) error and (b) differentiated error.

For the error variable, three linguistic variables have been considered:  $N$ -negative (high water temperature),  $Z$ -zero (water temperature close to the setpoint value), and  $P$ -positive (low water temperature). For the differentiated error variable (error rate) three linguistic variables have been also considered:  $N\dot{e}$ , negative (rising tank water temperature),  $Z\dot{e}$ , zero (lack of tank water temperature change), and  $P\dot{e}$ , positive (falling tank water temperature).

The associated fuzzy variable for the considered crisp output is presented in Fig. 3.14.

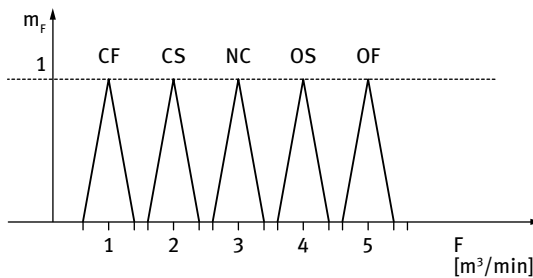


Fig. 3.14: Output fuzzy membership functions of the fuzzy controller.

For the controller output variable, i.e. the mixing tank inlet flow rate of *cold water*, have been considered five linguistic fuzzy variables: *CF*, close fast; *CS*, close slow; *NC*, no change; *OS*, open slow; *OF*, open fast.

Five fuzzy inference rules have been designed for the operation of the controller. They are

1. if (error  $e$  is negative  $N$ ) then (cold water flow rate  $F$  is open fast  $OF$ ),
2. if (error  $e$  is positive  $P$ ) then (cold water flow rate  $F$  is close fast  $CF$ ),
3. if (error  $e$  is zero  $Z$ ) and (error rate  $\dot{e}$  is positive  $Pe$ ) then (cold water flow rate  $F$  is close slow  $CS$ ),
4. if (error  $e$  is zero  $Z$ ) and (error rate  $\dot{e}$  is negative  $Ne$ ) then (cold water flow rate  $F$  is open slow  $OS$ ),
5. if (error  $e$  is zero  $Z$ ) and (error rate  $\dot{e}$  is zero  $Ze$ ) then (cold water flow rate  $F$  is no change  $NC$ ).

Mamdani inference mechanism is chosen for the fuzzy implication operation. A graphical representation for the computation of the fuzzy controller output is presented in Fig. 3.15. The inlet values are the error, with a value of  $e = -0.7^\circ\text{C}$ , and the differentiated error, with a value of  $\dot{e} = 0.10^\circ\text{C}/\text{min}$ . For rules no. 3, 4 and 5 the minimum *min* function has been used, as the antecedent parts are joined by the *and* operator. For each rule, the Mamdani inference is used to compute the membership function of the consequent. Furthermore, the resulting membership functions of the rules are aggregated using the *or* rule. As a result, the aggregated membership function is obtained  $m_{F \text{ aggregated}}$ .

The defuzzification is performed using the following centroid method:

$$C = \frac{\sum_{i=1}^n x_i m_{F_i}(x)}{\sum_{i=1}^n m_{F_i}(x)}, \quad (3.28)$$

where  $n$  is the number of segments (areas) in which the aggregated area has been divided in,  $x_i$  is the position of each segment, and  $m_{F_i}(x)$  is the membership value (area) corresponding to each segment. The centroid is situated at the value of  $C = 3.1 \text{ m}^3/\text{min}$ .

The design of the fuzzy controller presented in the Example 3.3 shows the procedure that resulted in a PD fuzzy controller. The design procedure may be used for building other types of fuzzy controllers.

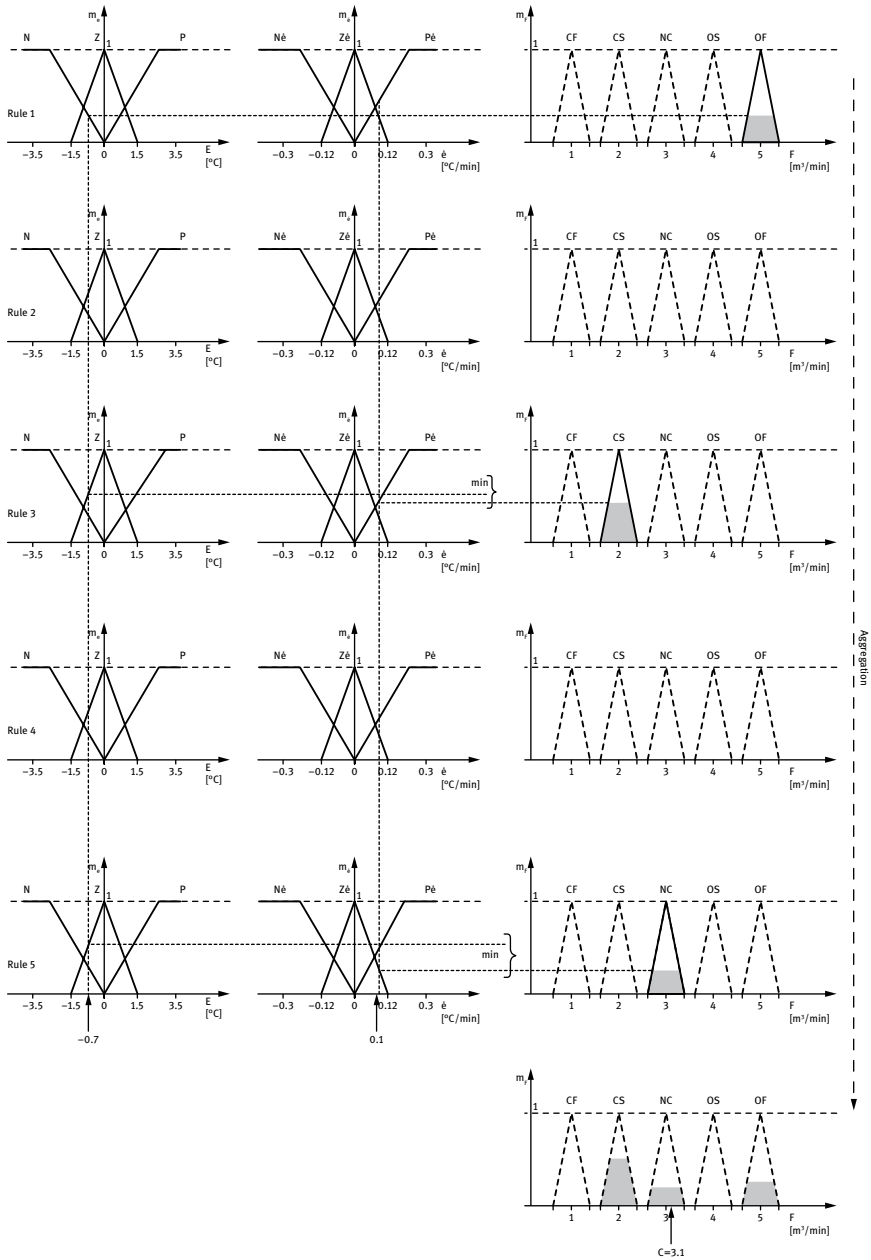


Fig. 3.15: PD fuzzy controller design.

## References

- [1] Zadeh, L.A., *Fuzzy sets*, Information and Control, 8, 338–353, 1965.
- [2] Palm, R., Driankov D., Hellendorn H., *Model Based Fuzzy Control*, Springer-Verlag, Berlin, New York, Heidelberg, 1997.
- [3] Sofron, E., Bizon, N., Ionita, S., Raducu, R., *Sisteme de Contro fuzzy, Modelare si Proiectare Asistate de Calculator*, ALL Educational, Bucuresti, 1998.
- [4] Sablani, S.S., Rahman, M.S., Datta, A.K., Mujumdar, A.S., Editors, *Handbook of Food and Bioprocess Modeling Techniques*, Taylor and Francis Group, CRC Press, 2007.
- [5] Klirk, J.G., Yuan, B., *Fuzzy Sets and Fuzzy Logic – Theory and Applications*, Prentice-Hall PTR, New Jersey, 1995.
- [6] Siler, W., Buckley, J.J., *Fuzzy Expert Systems and Fuzzy Reasoning*, John Wiley & Sons, New Jersey, Wiley-Interscience, 2005.



## 4 Optimal control systems

It seems that optimal control was not very much in attention in recent years since model-based control techniques, which have embedded an optimization procedure (minimization of the error), were applied extensively in industry (see Chapter 2 of the present book). It remains an important issue, since the minimization of costs *via* control became more and more important.

Optimal control of a process can refer at three distinct aspects:

- Optimal control in steady state, based on a scope or objective function (benefit, yield, pollution index, etc.) that has to be optimized, and thus, the optimal set-points of the controlled parameters are calculated [1];
- Optimal control in dynamic state of the batch processes, this being referred mainly at the minimization of the duration of the batch, or maximization of some characteristics as reaction conversion, respecting the constraints of quality [2, 3, 4];
- Optimal control in dynamic state of the continuous processes [5, 6], where the problem is to optimize a control quality criterion (IAE, ISE, ITAE, ITSE [7, 8], etc.).

Of course, in approaching optimal control, optimization techniques, which are not the subject of the present text, have to be known.

### 4.1 Steady-state optimal control

This type of control implies the existence of a steady-state model of the controlled process, having as an output an economic or technical variable and as input or state variables the key parameters of the process. The scope function is thus realized and subjected to optimization [9].

In the case of optimized control of the sequential processes, methods of identifying optimal policies are used [10].

#### Pontryagin's maximum principle

The method "maximum principle" was elaborated in 1956 by Pontryagin, a renowned Russian mathematician. This method is an extension of the variational calculus (or calculus of variations) upon the optimization problems described by ordinary differential equations. In the case of this type of problems, one single variable is considered, as time or space. The condition imposed to the variable is that it has to have a fixed value at the beginning and the end of the evolution of the process. Thus, a value  $t_0 = 0$  and another one  $t_f$  are imposed corresponding to the time domain on which the optimization is done, or a  $z_0 = 0$  and another  $z_f = l$  corresponding to the

space, where  $l$  is the length of the plug flow reactor, for example. Thus, considering a dynamic system (Fig. 4.1):

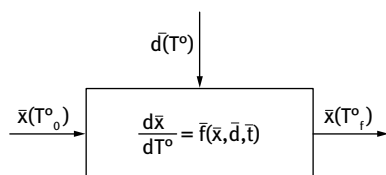


Fig. 4.1: The fundamental problem.

- whose state is characterized by the state vector  $\bar{x}(t)$ ;
- evolution in time is characterized by a set of decision variables that are also functions of time and that form the control/decision vector  $\bar{d}(t)$ ; these decision variables can vary in an admissible domain  $D$ ;
- the model of the system can be represented by a system of differential equations (4.1) such as:

$$\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, x_2, \dots, x_n, d_1, d_2, \dots, d_m, t) \\ \frac{dx_2}{dt} &= f_2(x_1, x_2, \dots, x_n, d_1, d_2, \dots, d_m, t) \\ &\dots \\ \frac{dx_n}{dt} &= f_n(x_1, x_2, \dots, x_n, d_1, d_2, \dots, d_m, t) \end{aligned} \quad (4.1)$$

The maximum principle proposes finding the decision vector  $\bar{d}^*(t)$  meaning the optimal solution among the admissible decision vectors  $\bar{d}(t)$  which maximizes the objective/scope function:

$$f_{ob} = \int_{t_0}^t F(\bar{x}(t), \bar{d}(t)) dt \stackrel{\Delta}{=} \max \quad (4.2)$$

it is subject to local constraints of the form:

$$\frac{dx_i}{dt} = f_i(\bar{x}(t), \bar{d}(t)), \text{ for } i = 1, 2, \dots, n, \quad (4.3)$$

with the initial conditions  $x_i = x_{i,0}$  at  $t = t_0$  and  $i = 1, 2, \dots, n$ .

According to the maximum principle, for the function 4.2 to be maximum, the decision vectors  $\bar{d}(t)$  have to be chosen in such a way that the Hamiltonian

$$H = F(\bar{x}(t), \bar{d}(t)) + \sum_{i=1}^n \lambda_i(t) f_i(\bar{x}(t), \bar{d}(t)) \quad (4.4)$$

is also at maximum;  $\lambda_i(t)$  are the Lagrange multipliers.

A way to solve the problem is to determine the extremum of the Hamiltonian as it follows:

1. values for the functions  $\bar{d}(t)$  are proposed for the whole domain of existence of  $t \in [t_0, t_f]$ ;
2. the differential system (4.3) is solved numerically, where the functions  $\bar{d}(t)$  have the previously established values and thus functions  $\bar{x}(t)$  are obtained;
3. the equation

$$\frac{d\lambda_i(t)}{dt} = -\sum_{i=1}^n \frac{\partial f_i}{\partial x_i} \lambda_i(t) - \frac{\partial F}{\partial x_i} \text{ for } i = 1, 2, \dots, n \quad (4.5)$$

is solved numerically for  $\lambda_i = 0$  at  $t = t_f$ ;

4. the approximation done in the first stage is improved by optimizing the Hamiltonian (eq. (4.4)) by using the gradient method:

$$d_i(t)^{(k+1)} = d_i(t)^{(k)} + s_g \frac{\partial H^{(k)}}{\partial d_i(t)} \quad (4.6)$$

The cycle is retaken with operation 2. The optimization is finished when functions  $\bar{d}(t)$  do not change anymore.

$s_g$  is the step done on the Hamiltonian direction.

### **i** Example 4.1

Optimal control of the methanol process using Pontryagin's maximum principle: The method has been applied to control the temperature profile at the methanol plant at C.C. Victoria, Romania [11]. The calculated increase in production was 33.8%, but in reality it reached 27%.

Methanol mass production makes optimal control possible and tempting because a fairly small 1% increase of the production brings important economic benefits. The case study is for a methanol reactor with a production of 210,000 t/year. The process to which we refer to is based on the low-pressure technology, the methanol synthesis being carried out in a multilayer reactor on a Co-ZnO-Al<sub>2</sub>O<sub>3</sub> catalyst with external cooling after each layer (Fig. 4.2).

The reactor design led to the characteristics presented in Tab. 4.1.  $\xi_1$  and  $\xi_2$  are the chemical conversions referring to the following reactions:



It is to be mentioned that a major design hypothesis was that of an adiabatic operation with the layer input temperature of 523 K.

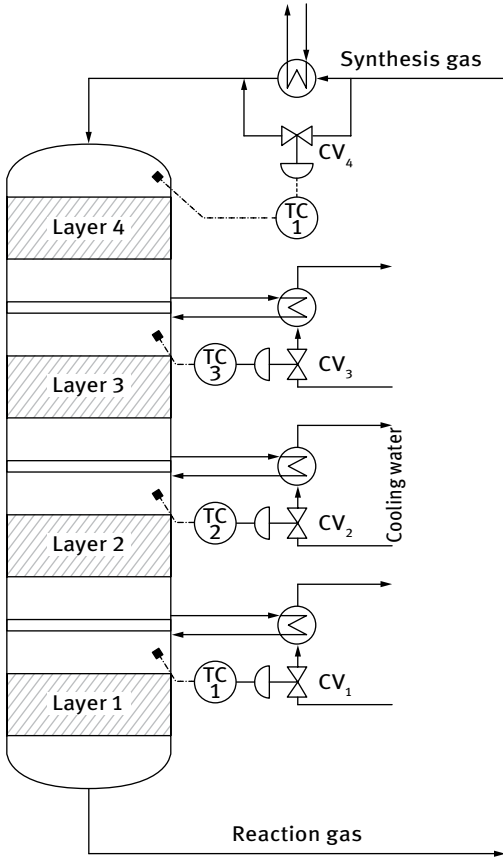


Fig. 4.2: Methanol synthesis reactor discussed in Example 4.1.

Tab. 4.1: Characteristics of the methanol synthesis reactor in Example 4.1.

Layer	Layer thickness [m]	Temperature [K]		Output conversion	
		input	output	$\xi_1$	$\xi_2$
1	0.3750	523	543.109	0.12421	0.000445
2	0.7050	523	543.136	0.25632	0.115788
3	1.0875	523	543.039	0.39172	0.201761
4	0.5925	523	543.983	0.44079	0.221330

### Thermodynamics of the process

Thermodynamic data referring to reactions (4.7) and (4.8) are presented in Tab. 4.2.

**Tab. 4.2:** Thermodynamic data of the methanol synthesis process described in Example 4.1.

Reaction	$\Delta_r H_{298}^0$ [kcal/mol]	$\Delta_r S_{298}^0$ [cal/(mol*K)]	$\Delta_r G_{298}^0$ [kcal/mol]
$\text{CO} + 2 \text{H}_2 \rightarrow \text{CH}_3\text{OH}$	-21.68	-52.31	-6.0916
$\text{CO}_2 + \text{H}_2 \rightarrow \text{CO} + \text{H}_2\text{O}$	8.64	10.02	5.8540

The equilibrium constants of both reactions are given by eqs. (4.9) and (4.10), respectively.

$$K_{p1} = \frac{K_{f1}}{K_{y1}} = \frac{p_{\text{CH}_3\text{OH}}}{p_{\text{CO}} p_{\text{H}_2}^2} \quad (4.10)$$

$$K_{p2} = \frac{K_{f2}}{K_{y2}} = \frac{p_{\text{CO}_2} p_{\text{H}_2\text{O}}}{p_{\text{CO}} p_{\text{H}_2}} \quad (4.11)$$

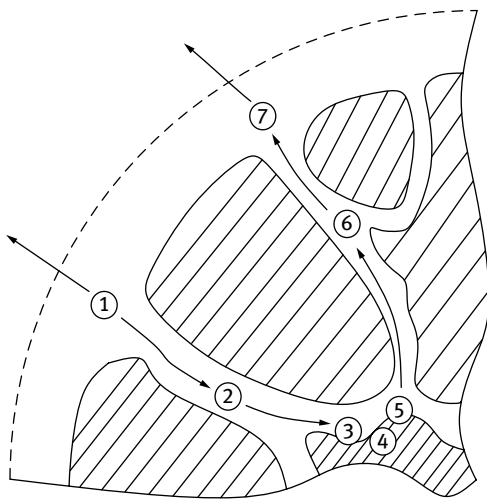
$K_{y_i}$  are the fugacity constants,  $p_{\text{CH}_3\text{OH}} \dots$  are the partial pressures of the components of both reactions, and

$$K_{f_i} = e^{-\frac{\Delta G_{T,i}^0}{RT}}, \quad i = 1, 2 \quad (4.12)$$

By means of a computer program, the equilibrium constants, the fugacity and  $K_{f_i}$  constants were determined within the temperature range of 0°C–300°C as well as pressure range of 10–100 bar.

### Kinetics of the process

The catalytic process develops in the catalyst's pores (Fig. 4.3) by following the sequence:

**Fig. 4.3:** The stages of the catalytic process described in Example 4.1.

1. reactant transport through the limit diffusion layer to the surface of the catalyst granule;
2. diffusion of the reactant in the catalyst's pores;
3. activated adsorption of the reactant R on the catalyst surface (chemisorption);
4. reaction on the catalyst centers on its surface, when reactant R transforms into product P;
5. desorption of product P from the catalyst;
6. diffusion of product through the catalyst's pores;
7. diffusion of the product away from the catalyst granule.

Natta et al. [12] arrived at the conclusion that the rate determining step is the chemical reaction between the species adsorbed on the catalyst, with an activation energy of 17.1 kcal/mol.

In Tab. 4.3, the different reported expressions for the reaction rate of the process are given.

**Tab. 4.3:** Rate of reaction forms proposed in the literature.

Authors	Reaction rate expression
Atroschenko, Zazorin [32]	$r_1 = k_1 p_{H_2} \sqrt[4]{\frac{p_{CO}}{p_{CH_3OH}}} - k_2 \sqrt[4]{\frac{p_{CH_3OH}}{p_{CO}}}$ $f_{CO} p_{CO} f_{H_2}^2 p_{H_2}^2 - f_{CH_3OH} \frac{p_{CH_3OH}}{K_{cg}}$
Natta, Pasquon [43]	$r_1 = \frac{f_{CO} p_{CO} f_{H_2}^2 p_{H_2}^2 - f_{CH_3OH} \frac{p_{CH_3OH}}{K_{cg}}}{A(1 - B f_{CO} p_{CO} + C f_{H_2}^2 p_{H_2}^2 + D f_{CH_3OH} p_{CH_3OH} + E f_{CO_2} p_{CO_2})}$
Uchida, Ogino [58]	$r_1 = k_1 \left[ \sqrt[4]{p_{CO} p_{H_2}^2} - \frac{p_{CH_3OH}}{k_{p1} \sqrt[3]{p_{CO} p_{H_2}^2}} \right]$
Cappelli [11]	$r_1 = k_1 \left( \frac{p_{H_2} \sqrt{p_{CO}}}{\sqrt[3]{p_{CH_3OH}^2}} - \frac{1}{k_{p1} p_{H_2} \sqrt{p_{CO}}} \sqrt[3]{p_{CH_3OH}} \right)$
Siminiceanu [48]	$r_2 = \frac{k_2 (p_{CO_2} p_{H_2} k_{p_2} - p_{CO} p_{H_2O})}{K' p_{CO} + p_{CO_2}}$
Experimentally determined on ICI catalyzer	$r_1 = k_1 \left( \frac{p_{H_2} \sqrt{p_{CO}}}{\sqrt[3]{p_{CH_3OH}^2}} - \frac{1}{k_{p1}} \frac{\sqrt[3]{p_{CH_3OH}}}{p_{H_2} \sqrt{p_{CO}}} \right)$ $k_1 = 9.24 \times 10^6 e^{-11423.1/T}$

The terms in Tab. 4.3 are detailed in [11].

### Mathematical model of the process

A quasi-homogeneous, unidimensional model was proposed by taking into consideration the following realities and simplifying assumptions:

- there are few data about the structural and textural characteristics of the catalyst;
- assumption of plug-flow behavior of the gaseous phase is supported by the uniformity of its velocity profile in the cross-sectional area of the catalyst;

- important temperature gradients can be neglected due to adiabatic operation;
- pressure drop on the reactor is neglected in comparison with the operating pressure;
- variation of gas velocity depending on temperature is considered;
- variation of gas velocity due to the mole number change is considered;
- average molar heat of the mixture is recalculated depending on composition and temperature.

Consequently, the mathematical model is composed of the relations of conversion and temperature:

$$\frac{dp_{CO}}{dt} = -r_1 + r_2 \quad (4.13)$$

where  $r_1$  and  $r_2$  are rates of consumption and formation of CO in reactions (4.7) and (4.8), respectively. They are presented in Tab. 4.3 (Siminiceanu and ICI catalyst [11]).

The rates can be expressed through:

$$-r_1 = \frac{dp_{CO}}{d\xi_1} \cdot \frac{d\xi_1}{dt} \quad (4.14)$$

$$r_2 = \frac{dp_{CO}}{d\xi_2} \cdot \frac{d\xi_2}{dt} \quad (4.15)$$

Noting with  $v_0$  the initial fictive gas mixture velocity in the reactor and with  $v$  the current one, one may pass to the space derivatives:

$$\frac{d}{dt} = \frac{d}{dz} \cdot \frac{dz}{dt} \text{ where } \frac{dz}{dt} = v = v_0(1 + \varepsilon\xi_1) \frac{T^\circ}{T_0} \quad (4.16)$$

Equations (4.14) and (4.15) become:

$$-r_1 = v_0(1 + \varepsilon\xi_1) \frac{T^\circ}{T_0} \frac{dp_{CO}}{d\xi_1} \cdot \frac{d\xi_1}{dz} \text{ and} \quad (4.17)$$

$$r_2 = v_0(1 + \varepsilon\xi_1) \frac{T^\circ}{T_0} \frac{dp_{CO}}{d\xi_2} \cdot \frac{d\xi_2}{dz} \quad (4.18)$$

Considering that:

$$p_{CO} = x_{CO}p$$

and

$$x_{CO} = x_{CO}^0 \frac{1 - \xi_1 + \beta\xi_2}{1 + \varepsilon\xi_1}$$

relations (4.17) and (4.18) become [4, 5]:

$$-r_1 = x_{CO} p v_0 (1 + \varepsilon \xi_1) \frac{T^\circ}{T_0} \frac{1 - \xi_1 + \beta \xi_2}{1 + \varepsilon \xi_1} \cdot \frac{d\xi_1}{dz} \quad (4.19)$$

$$r_2 = \beta p v_0 x_{CO} \frac{T^\circ}{T_0} \cdot \frac{d\xi_2}{dz} \quad (4.20)$$

and

$$\begin{aligned} x_{H_2} &= x_{CO}^0 \frac{\alpha - 2\xi_1 - \beta \xi_2}{1 + \varepsilon \xi_1} \\ x_{CH_3OH} &= x_{CO}^0 \frac{\gamma + \xi_1}{1 + \varepsilon \xi_1} \\ x_{H_2O} &= x_{CO}^0 \frac{1 - \xi_2}{1 + \varepsilon \xi_1} \\ x_{CH_3OH} &= x_{CO}^0 \frac{\delta + \xi_2}{1 + \varepsilon \xi_1} \end{aligned} \quad (4.21)$$

where

$$\alpha = \frac{x_{H_2}^0}{x_{CO}^0}; \beta = \frac{x_{CO_2}^0}{x_{CO}^0}; \gamma = \frac{x_{CH_3OH}^0}{x_{CO}^0}; \delta = \frac{x_{H_2O}^0}{x_{CO}^0}; \varepsilon = -2x_{CO}^0$$

When replacing the molar fractions (4.21), together with the expressions (4.19) and (4.20) in the rate expressions [4, 5], the following is obtained:

$$r_1 = k_1 \left( \frac{p x_{CO}^0}{1 + \varepsilon \xi_1} \right)^{0.84} \left[ \frac{(1 - \xi_1 + \beta \xi_2)^{0.8} (\alpha - 2\xi_1 - \beta \xi_2)}{(\gamma + \xi_1)^{0.66}} - \frac{1}{K_{p1} (p x_{CO})^2} \frac{(\gamma + \xi_1)^{0.34} (1 + \varepsilon \xi_1)^2}{(\alpha - 2\varepsilon_1 - \beta \varepsilon_2) \sqrt{1 - \xi_1 + \beta \xi_2}} \right] \quad (4.22)$$

$$r_2 = k_2 \frac{p x_{CO}^0}{1 + \varepsilon \xi_1} \frac{K_{p2} (\alpha - 2\xi_1 - \beta \xi_2) (1 - \xi_2) - (1 - \xi_1 + \beta \xi_2) (\delta + \varepsilon_2)}{k' (1 - \xi_1 + \beta \xi_2) + \beta (1 - \xi_2)} \quad (4.23)$$

Finally, from eqs. (4.19), (4.20), (4.22), and (4.23), conversion-space function is obtained

$$\begin{aligned} \frac{d\xi_1}{dz} &= \frac{k_1}{v_0} \frac{T_0^\circ}{T^\circ} \frac{1}{1 + \varepsilon + \beta \varepsilon \xi_1} \left[ \frac{(\alpha - 2\xi_1 - \beta \xi_2) \sqrt{1 - \xi_1 + \beta \xi_2}}{(\gamma + \xi_1)^{0.66}} - \frac{1}{(p x_{CO}^0)^2} \frac{(\gamma + \xi_1)^{0.34} (1 + \varepsilon \xi_1)^2}{K_{p1} (\alpha - 2\varepsilon_1 - \beta \varepsilon_2) \sqrt{1 - \xi_1 + \beta \xi_2}} \right] \\ &\quad \left( \frac{1 + \varepsilon \xi_1}{p x_{CO}^0} \right)^{0.16} \end{aligned} \quad (4.24)$$

$$\frac{d\xi_2}{dz} = \frac{k_2}{v_0} \frac{T_0^\circ}{T^\circ} \frac{K_{p2} (\alpha - 2\xi_1 - \beta \xi_2) (1 - \xi_2) - (1 - \xi_1 + \beta \xi_2) (\delta + \varepsilon_2)}{k' (1 - \xi_1 + \beta \xi_2) + \beta (1 - \xi_2)} \quad (4.25)$$



Because both rate and equilibrium constants depend on temperature, the equation of heat balance (since the reactor operates under adiabatic regime) has to complete the model:

$$\frac{dT^\circ}{dz} = \frac{1}{1 + \varepsilon\xi_1} \left( x_{CO}^0 \frac{\Delta_r H_1}{c_p} \frac{d\xi_1}{dz} - x_{CO_2}^0 \frac{\Delta_r H_2}{c_p} \frac{d\xi_2}{dz} \right) \quad (4.26)$$

### Optimal temperature profile in the methanol synthesis reactor

The maximization of the methanol synthesis is targeted and the control policy strives for a temperature profile along the reactor,  $T^\circ(z)$ .

The objective function is

$$f_{ob} = \xi_1 = \int_0^L \frac{d\xi_1}{dz} dz \quad (4.27)$$

By comparing (4.27) with (4.2), one may observe that  $F = \frac{d\xi_1}{dz}$  and that  $\xi_1$  and  $\xi_2$  are corresponding to the dependent variables  $\bar{x}(z)$ .

The local constraints (eq. (4.3)) are:

$$\begin{aligned} \frac{d\xi_1}{dz} &= f_1(\xi_1, \xi_2, T^\circ) \\ \frac{d\xi_2}{dz} &= f_2(\xi_1, \xi_2, T^\circ) \end{aligned} \quad (4.28)$$

while the initial conditions are:

$$\xi_1 = 0, \xi_2 = 0 \text{ at } z = z_0 = 0$$

Equation (4.5) can be written, under the condition  $F = \frac{d\xi_1}{dz} = f_1$ , as

$$\begin{aligned} \frac{d\lambda_1}{dz} &= -\frac{\partial f_1}{\partial \xi_1} \lambda_1 - \frac{\partial f_2}{\partial \xi_1} \lambda_2 - \frac{\partial F}{\partial \xi_1} = -(1 + \lambda_1) \frac{\partial f_1}{\partial \xi_1} - \lambda_2 \frac{\partial f_2}{\partial \xi_1} \\ \frac{d\lambda_2}{dz} &= -\frac{\partial f_1}{\partial \xi_2} \lambda_1 - \frac{\partial f_2}{\partial \xi_2} \lambda_2 - \frac{\partial F}{\partial \xi_2} = -(1 + \lambda_1) \frac{\partial f_1}{\partial \xi_2} - \lambda_2 \frac{\partial f_2}{\partial \xi_2} \end{aligned} \quad (4.29)$$

with the final conditions  $\lambda_1 = 0, \lambda_2 = 0$  at  $z = L$

The Hamiltonian H becomes:

$$H = F + \lambda_1 f_1 + \lambda_2 f_2 = (1 + \lambda_1) f_1 + \lambda_2 f_2 \quad (4.30)$$

and its derivative related to the control policy is:

$$\frac{\partial H}{\partial T^\circ} = (1 + \lambda_1) \frac{\partial f_1}{\partial T^\circ} + \lambda_2 \frac{\partial f_2}{\partial T^\circ} \quad (4.31)$$

Equation (4.31) is used to correct the temperature profile:

$$T^{\circ}(z)^{(k+1)} = T^{\circ}(z)^k + S_g \left( \frac{\partial H}{\partial T^{\circ}} \right)^{(k)} \quad (4.32)$$

The computing program for calculating the optimum temperature profile has the following sequences:

1. Initially, an isothermal regime for the whole reactor is supposed, with the input temperature  $T_0^{\circ}$ ; thus, for  $k = 1$ ,  $T_i^{(1)} = T_0^{\circ}$ ,  $i = 1, 2, \dots, n$ ;
2. The equation system (4.26)–(4.28) is integrated by using the Euler method, and by considering the conditions from (4.28);
3. The system (4.29) is integrated from  $z = L$  to  $z = 0$  with the final conditions  $\xi_{1,n} = 0$ ,  $\xi_{2,n} = 0$ :

$$\lambda_{i,j} = \lambda_{i,j+1} + \Delta z \left[ -(1 + \lambda_{i,j+1}) \left( \frac{\partial f_1}{\partial \xi_i} \right)_{j+1} - (\lambda_{2,j+1}) \left( \frac{\partial f_2}{\partial \xi_i} \right)_{j+1} \right]$$

where

$$i = 1, 2 \text{ and } j = n - 1, n - 2, \dots, 1$$

4. Temperatures are recalculated

$$T_i^{\circ(k+1)} = T_i^{\circ(k)} + S_g \left( \frac{\partial H}{\partial T_i^{\circ}} \right)^{(k)}$$

where  $i = 1, 2, \dots, n$

One passes to a new iteration from point 2.

**Tab. 4.4:** Optimal theoretic temperature profile and theoretical optimal performances in the methanol reactor of Example 4.1.

$z$ [m]	$\xi_1$	$\xi_2$	$\lambda_1$	$\lambda_2$	$T$ [K]
0.000	0.000	0.000	1.338	-1.338	573.660
0.276	0.517	0.487	1.001	-0.820	546.150
0.552	0.546	0.447	0.775	-0.554	542.940
0.828	0.570	0.430	0.605	-0.386	540.990
1.104	0.591	0.420	0.469	-0.271	539.480
1.380	0.610	0.414	0.357	-0.189	538.170
1.656	0.628	0.410	0.263	-0.129	536.980
1.932	0.644	0.408	0.183	-0.084	535.860
2.208	0.658	0.406	0.114	-0.049	534.810
2.484	0.672	0.405	0.053	-0.022	533.820
2.760	0.685	0.404	0.000	0.000	532.910

The results of the iterations and the optimum temperature profile which ensure  $\xi_1 = \max$  are given in Tab. 4.4.

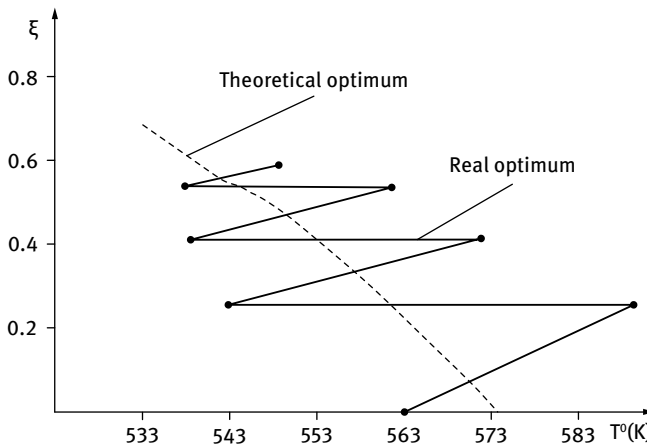
The optimum temperature profile from Tab. 4.4 cannot be practically obtained because its tuning can be carried out only between the catalyst layers by using the

external heat exchangers – see Fig. 4.2. The input temperature in each next layer has to be controlled in such a way that all input and output temperatures have values around the optimal profile (see Tab. 4.5 and Fig. 4.4).

**Tab. 4.5:** Optimal real temperature profile and the consequent performances in the methanol reactor of Example 4.1.

Layer	Layer thickness [m]	Temperature [K]		Output conversion	
		input	output	$\xi_1$	$\xi_2$
1	0.3750	563	586.36	0.2562	0.4161
2	0.7050	543	571.71	0.4115	0.4160
3	1.0875	538	561.51	0.5369	0.4139
4	0.5925	538	548.40	0.5893	0.4035

It is important to observe that the real performance ( $\xi_1 = 0.589$ ) is inferior to the theoretical performance obtained after applying the optimization procedure ( $\xi_1 = 0.685$ ), but with 33.8% higher than the performance obtained in the traditional adiabatic operation,  $\xi_1 = 0.440$  (Tab. 4.1). The saw tooth temperature profile is obtained by fixing each controller to the appropriate setpoint value:  $T_{1set}^\circ = 563$  K,  $T_{2set}^\circ = 543$  K,  $T_{3set}^\circ = 538$  K,  $T_{4set}^\circ = 538$  K.



**Fig. 4.4:** The quasi-optimal saw tooth profile of temperatures in the methanol reactor.

## 4.2 Dynamic optimal control of batch processes

In the given situation, there are two ways of approaching the problem:

- To obtain a *maximum* of quantity  $Q$ , in a given time period  $T$ .
- To obtain a certain product quantity  $Q$ , in a *minimum* operational time.

In the first of these situations, a maximization of a function of type

$$I = \int_0^T Q dt \text{ is subjected to the constraint} \quad (4.33)$$

$$\bar{y} = \frac{\int_0^T y_Q dt}{\int_0^T Q dt} = y^*, \text{ where } \bar{y} \text{ is the average specification of the quality of the product (it can}$$

be a mass or a molar concentration or fraction).

The second situation requires a minimization of a time function

$$t_T^* = \min \int_0^{t_T} dt, \text{ or the dividing of the process in } N \text{ discrete stages,} \quad (4.34)$$

$$t_T^* = \min \sum_{i=1}^N t_i, \text{ where } t_i \text{ is the duration of one stage.}$$

#### Example 4.2

One has to determine the optimal operation policy of a batch distillation, by manipulating the reflux flow to obtain a certain given quantity in a minimum of time [13].

A technique of dynamic programming was applied [10].

The process of batch distillation of one fraction resides in the following operations:

1. Loading the batch to be separated into the bottom of the column;
2. Bringing the temperature to the boiling point;
3. Separation of the lightest fraction based on the difference of volatility, until the composition reaches a certain minimum value.

Corresponding to the continuous decrease of the light component molar fraction in the bottom,  $x_B$ , the molar fraction of the light component in the distillate,  $x_D$ , decreases as well (when the reflux ratio is kept constant).

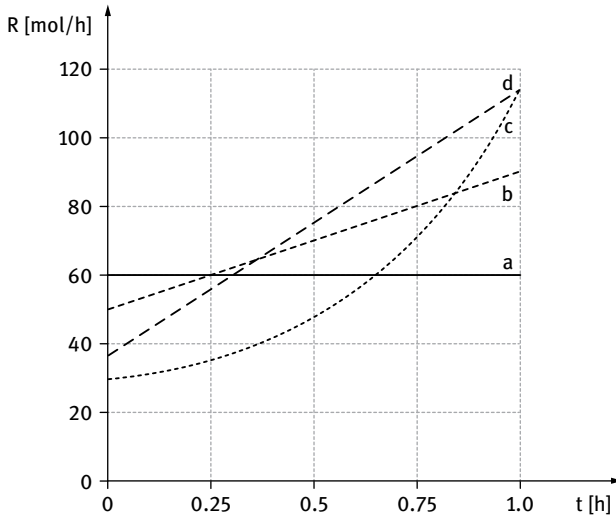
After writing the mass balance for the column, the continuous decrease of the batch quantity in the bottom,  $M_B$ , is expressed through eqs. (4.35) and (4.36):

$$\frac{dM_B(t)}{dt} = -Q_D(t) \quad (4.35)$$

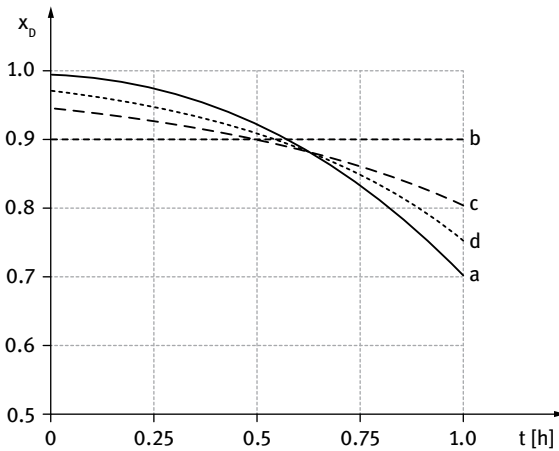
$$\frac{dx_B(t)}{dt} = \frac{Q_D(t)}{M_B(t)} \{x_B(t) - x_D[x_B(t), Q_D(t)]\} \quad (4.36)$$

where  $Q_D(t)$  and  $M_B(t)$  are the quantities separated at time  $t$  at the top and the bottom of the load, respectively.





(a)



(b)

**Fig. 4.5:** Batch distillation operation: operation with constant reflux ratio a; operation with constant molar fraction b; optimal operation with maximum distillate quantity in a given time c; optimal operation with a certain quantity Q in a minimum of time d.

If  $t_T$  is the total batch distillation time and  $\bar{x}_D$  the average desired molar fraction of the collected distillate, the mass balance below describes the process:

$$\bar{x}_D \int_0^{t_T} Q_D(t) dt = \int_0^{t_T} Q_D(t) \cdot x_D [x_B(t), Q_D(t)] dt \quad (4.37)$$

The vapor flow produced in the column reboiler is considered constant, a fact which ensures an adequate thermal regime for the bottom of the column.

As observed in Fig. 4.5, the column can be operated with either a constant reflux ratio (the molar fraction at the top is decreasing in time) or with a progressive reflux ratio increase (with maintaining an average  $\bar{x}_D$  constant; this means the continuous decrease of the distillate flow rate  $D$ ). This happens simultaneously with the decrease of the light component concentration in the bottom.

The goal of control action is to minimize the duration of the distillation process of the prescribed distillate quantity, by also achieving its specified purity. In this situation, the optimization of the batch distillation process becomes a problem of allocation [10, 13]. It is solved in the following sequences:

1. The considered function

$$f(x_1, x_2, \dots, x_i, \dots, x_N) = g_1(x_1) + g_2(x_2) + \dots + g_i(x_i) + \dots + g_N(x_N) \quad (4.38)$$

is subjected to the constraints:

$$\sum_{i=1}^N x_i = y, \quad x_i \geq 0$$

where  $y$  is the total quantity of resources and the functions  $g_i(x_i)$  can have any form. The variables  $x_i$  are considered resources to be allocated optimally to  $N$  activities, each activity being characterized through its own objective function  $g_i(x_i)$ . The objective function of the system is a sum of functions that characterizes each separate activity. The problem has a finite optimum only in the case of the limited quantity of resources. Thus, the optimal value of the function (4.38) is:

$$f^*(y) = \text{opt } f(x_1, x_2, \dots, x_N) \quad (4.39)$$

2. The properties to be considered are:
  - a.  $g_i(0) = 0$  for any  $i = \overline{1, N}$  when  $f_i^*(0) = 0$
  - b.  $f_i^*(y) = g_i(y)$ ,  $y \geq 0$ ; thus, a recurrence equation between  $f_i^*$  and  $f_{i-1}^*$  can be found:

$$f_i(y) = g_i(x_i) + f_{i-1}(y - x_i)$$

- c. The optimum choice of  $x_i$  is the one optimizing  $f_i(y)$ ,

$$f_i^*(y) = \text{opt}_{0 \leq x_i \leq y} [g_i(x_i) + f_{i-1}^*(y - x_i)] \text{ for } i = 2, 3, \dots, N, y \geq 0 \quad (4.40)$$

Consequently, by using eq. (4.34), the value of  $t_i$ , that is the duration of one distillation stage can be expressed. It is defined as the interval in which the bottom molar fraction decreases from  $x_{B, i-1}$  to  $x_{B, i}$ .

By taking into account the known quantity and quality of the load in the bottom of the column at the beginning of the batch distillation process,  $M_{B,0}$  and  $x_{B,0}$ , the final (after  $N$  stages) molar fraction of the volatile component in the bottom,  $x_{B,N}$ , as well as the average molar fraction of the collected distillate,  $\bar{x}_D$ , the following equations (4.41– 4.44) can be written:

$$M_{B,0} - M_{B,N} = P_N \quad (4.41)$$

$$M_{B,0}x_{B,0} - M_{B,N}x_{B,N} = P_N \bar{x}_D \quad (4.42)$$

where the total quantity of distillate is

$$P_N = \sum_{i=1}^N D_i, \text{ with } D_i \text{ as the quantity extracted at each stage.} \quad (4.43)$$

The problem can be thus seen as a problem of allocation where the activities are the stages 1, 2, ...,  $N$ , characterized by the variation  $\Delta x_w$  of the concentration of the volatile component in the bottom:

$$\Delta x_B = \frac{x_{B,0} - x_{B,N}}{N} \quad (4.44)$$

The resources, the quantities of distillate  $D_i$ . The functions  $g_i$  from the allocation problems are the time periods  $t_i$ .

To determine the optimal function, the mass balance equations for stage  $i$  are written as follows:

$$M_{B,i-1} - M_{B,i} = D_i \quad (4.45)$$

$$M_{B,i-1}x_{B,i-1} - M_{B,i}x_{B,i} = D_i x_{D,i} \quad (4.46)$$

But:

$$M_{B,i} = M_{B,0} - \sum_{j=1}^i D_j \text{ and} \quad (4.47)$$

$$x_{B,i-1} = x_{B,0} - (i-1)\Delta x_B, \text{ or} \quad (4.48)$$

$$x_{B,i} = x_{B,0} - i\Delta x_B \quad (4.49)$$

From eqs. (4.45)–(4.49):

$$x_{D,i} = x_{B,0} + \left( \frac{M_{B,0} - \sum_{j=1}^i D_j}{D_i} - i + 1 \right) \Delta x_B \quad (4.50)$$

Thus,

$$x_{D,i} = x_{D,i} \left( D_i, \sum_{j=1}^i D_j \right) \text{ and} \quad (4.51)$$

$$x_{B,av} = x_{B,i-1} - \frac{\Delta x_B}{2} \quad (4.52)$$

When having the average concentrations in the bottom and at the top, the average value of the reflux in stage  $i$  can be determined. To reduce the volume of calculations, the Fenske-Underwood-Gilliland method [14] and not the iterative one (tray to tray) is used. With this method, coefficients

$$A = \frac{R - R_{\min}}{R + 1} \text{ and } B = \frac{S - S_{\min}}{S + 1} \quad (4.53)$$

are calculated.

The minimum reflux  $R_{\min}$  and minimum number of theoretical transfer units  $S_{\min}$  are calculated with this method, where the feed molar fraction  $x_F$  is replaced with  $x_{B, \text{fav}}$ . To determine the reflux  $R$  corresponding to a given  $S$ , the function  $A = f(B)$  is obtained through the least square method:

$$A = 0.96 - 2.09B + 0.21B^2 + 2.82B^3 - 2.90B^4$$

for  $0 \leq B \leq 0.51$  and

$$A = 0.86 - 0.95B - 2.48B^2 + 3.00B^3$$

for  $B > 0.51$

If  $S < S_{\min}$  and  $B < 0$ , respectively, the fractionation is impossible even at infinite reflux. Finally,  $R_i(x_{D_i})$  is calculated:

$$R_i = R_i \left( D_i, \sum_{j=1}^i D_j \right) \quad (4.54)$$

The mass balance equation for the vapors produced in the bottom of the column during one stage is:

$$Q_V t_i = D_i (R_i + 1) \quad (4.55)$$

where  $Q_V$  is the constant, known vapor flow rate produced in the reboiler.

From previous relations (4.39) and (4.40),

$$t_i = g_i \left( D_i, \sum_{j=1}^i D_j \right) \quad (4.56)$$

and the desired optimal value [7] is:

$$t_T^* = \min_{D_1, D_2, \dots, D_N} [g_1(D_1) + g_1(D_2, D_1 + D_2) + \dots + g_N(D_N, D_1 + D_2 + \dots + D_N)] \quad (4.57)$$

To calculate the optimal policy  $f_i^*(P)$ , one has to consider the allocations  $D_i$ .

For  $x_{D_i} < 1$ ,

$$D_i > \frac{(M_{B,0} - P)\Delta x_B}{1 - x_{B,0} + (i-1)\Delta x_B} > 0 \text{ and for } i = 1, D_i = P, \text{ and thus}$$

$$D_1 > \frac{M_{B,0}\Delta x_B}{1 - x_{B,0}} > 0$$

Imposing  $D_1 > 0$ ,  $D_2 > 0$ , and  $D_{i-1} > 0$ , it results that  $D_i < P$ . Let us apply this principle in the case of batch fractionation:



$$f_i^*(P) = \min_{D_{i\min} < D_i < P} [g_i(D_i, P) + f_{i-1}^*(P - D_i)] \tag{4.58}$$

with  $i = 2, 3, \dots, N$  and  $f_1^*(P) = g_1(D_1)$

Because the objective function depends for each stage  $i$  on the allocation  $D_i$  and the total quantity of resources  $P$ , the solution will be modified, and the functions  $g_i$  can not be separately calculated (they depend on  $P$ ). Dividing the domain  $[0, P_N]$  in  $N$  intervals of dimension  $\Delta P$ , and because  $D_i$  take discrete values in the nodes of the network, the results of the minimization procedure

(eq. (4.58)) need  $1 + 2 + \dots + (M - 1) = \frac{(M - 1)M}{2}$  evaluations of the function

$f_i(P)$  in each stage, instead of  $N + 1$  as in the usual procedure. Because in any stage  $i$ ,  $D_{i+1}, D_{i+2}, \dots, D_N$  cannot be zero but at most equal to  $\Delta P$ , the highest rational value for  $P$  is  $[M - (N - i)]\Delta P$ . Consequently, for a series of values  $j(1 \leq j \leq N)$ , a fake very high value will be allocated to the functions  $f_i^*(j\Delta P)$  and similarly also the conventional value of the corresponding decision  $D_i^*(j\Delta P)$ .

Finally, the optimal fractionation policy will be obtained by swiping the *optimal decisions matrix*, starting from the element in the row  $j = M$ , corresponding to  $P_N$  and the column  $i = N$ , corresponding to  $x_{B,N}$ .

The matrix resolution procedure is the following:

1. The domain  $0 \leq y \leq y_N$  in which the resources  $y$  can take values is divided in  $M$  intervals of magnitude  $\Delta$ . This way, only the discrete values of the network  $R$  will be considered:  $R = R\{0, \Delta, 2\Delta, \dots, j\Delta, \dots, M\Delta\}$ ; it will be admitted that each term of the sequence of functions  $f_i^*(y)$  will be evaluated only in the nodes of this network and all allocation variables will be discretized in these points to avoid interpolation calculations.
2. For  $i = 1, f_1^*(y) = g_1(y)$  will be determined immediately. Further,

$$f_1^*(j\Delta) = g_1(j\Delta) \tag{4.59}$$

where  $j = 0, 1, 2, \dots, M$ , the total resources quantity takes the discrete values of the  $R$  network.

3. For  $i = 2$ , eq. (4.40) becomes:

$$f_2^*(y) = \text{opt}_{0 \leq x_1 \leq y} [g_2(x_2) + f_1^*(y - x_2)], \text{ or, in the discrete form,}$$

$$f_2^*(j\Delta) = \text{opt}_{0 \leq k \leq j} \{g_2(k\Delta) + f_1^*[(j - k)\Delta]\}, j = 0, 1, 2, \dots, M \tag{4.60}$$

The optimization process evolves through direct evaluation (exhaustive search):

- the string of values  $g_2(k\Delta)$  for  $k = 0, 1, 2, \dots, M$  is calculated;
  - because  $f_1^*[(j - k)\Delta]$  has been previously calculated, for each of the  $M + 1$  data sets out of the  $j + 1$  values of  $f_2^*(j\Delta)$ , the optimal (maximum or minimum) value is chosen (eq. (4.68)).
4. The procedure at point 3 is repeated for the following activities for  $i = 3, 4, \dots, N$ . Finally, the optimal decisions matrix  $k_i^*(j)$  summarized in Tab. 4.6 is obtained.

Tab. 4.6: The matrix of optimal decisions for Example 4.2.

Stage	1	2	...	<i>i</i>	...	<i>N</i> -1	<i>N</i>
State							
0		$k_1^*(0)$	–	$k_i^*(0)$	–	$k_{N-1}^*(0)$	$k_N^*(0)$
1		$k_2^*(1)$	–	$k_i^*(1)$	–	$k_{N-1}^*(1)$	$k_N^*(1)$
2		$k_2^*(2)$	–	$k_i^*(2)$	–	$k_{N-1}^*(2)$	$k_N^*(2)$
–		–	–	–	–	–	–
–		–	–	–	–	–	–
–		–	–	–	–	–	–
<i>j</i>		$k_2^*(j)$	–	$k_i^*(j)$	–	$k_{N-1}^*(j)$	$k_N^*(j)$
–		–	–	–	–	–	–
–		–	–	–	–	–	–
–		–	–	–	–	–	–
<i>M</i> -1		$k_2^*(M-1)$	–	$k_i^*(M-1)$	–	$k_{N-1}^*(M-1)$	$k_N^*(M-1)$
<i>M</i>		$k_2^*(M)$	–	$k_i^*(M)$	–	$k_{N-1}^*(M)$	$k_N^*(M)$

- To obtain the optimal allocation policy for a process with a total number of *i* activities ( $i \leq N$ ) and  $y = j\Delta$  ( $y \leq y_M, j \leq M$ ), the following procedure is used:
  - in column *i* and row *j* from the optimal decisions matrix, the solution  $k_i^*(j)$  is found, corresponding to  $x_i^*(y)$ ;
  - this value is deducted from the total quantity of resources, the available resources for the other *i* - 1 activities, being obtained;
  - on the line corresponding to row *j* -  $k_i^*(j)$  and column *i* - 1 of the optimal decisions matrix, the optimal allocation for the activity *i* - 1,  $k_{i-1}^*[j - k_i^*(j)]$  is found;
  - the procedure is repeated for the other decreasing values of *i* until all values  $k_i^*$  of  $\Delta = x_i^*$  are found. This is the desired optimal allocation policy.

The advantage of the formulation of the optimal fractionation policy as a resource allocation policy is that it allows the establishment of the final optimal bottom composition and determines its degree of separation that is economically appropriate to operate.

The optimal decisions matrix allows solving any allocation problem with maximum *M* activities and maximum quantity of resources  $y_M$ . In the present situation, this means that for any final concentration of the residue larger or equal to  $x_{B,N}$  and for any quantity less or equal to  $P_N$ , the distillation does not advance until the final concentration  $x_{B,N}$  and stops at  $x_{B,i}(x_{B,i} \geq x_{B,N})$ .

By keeping in mind the distillate's desired purity, the following equations can be written:

$$M_{B,0} - M_{B,i} = P_i, \text{ the mass balance equation} \tag{4.61}$$

$$M_{B,0}x_{B,0} - M_{B,i}x_{B,i} = P_i\bar{x}_D, \text{ the component balance equation} \tag{4.62}$$

from where

$$P_i = M_{B,0} \frac{x_{B,0} - x_{B,i}}{\bar{x}_D - x_{B,i}} \quad (4.63)$$

with  $P_i \leq P_N$  because  $x_{B,i} \geq x_{B,N}$  and  $\bar{x}_D \geq x_{B,0}$ .

The optimal policy will be obtained by starting from the element of the optimal decisions matrix from column  $i$  and row  $j$  where  $j = \frac{P_i}{\Delta P}$ . Because for any  $x_{B,i} \geq x_{B,N}$  an optimal distillation policy can be obtained, that one corresponding to the most convenient value  $x_{B,i}^*$  will be chosen. Finally, the following can be determined:

- the optimal policy  $D_i^*$ ,  $i = 1, 2, \dots, i^*$  ( $i^*$  corresponds to  $x_{w,i}$ )
- $R_i^*$ , the corresponding optimal policy of the reflux.

The numerical application is run with the following initial data:

$$M_{B,0} = 500 \text{ mol}; x_{B,0} = 0.2; x_{B,N} = x_{\min} = 0.05; x_{B,\max} = 0.12; \bar{x}_D = 0.8; \alpha = 2.5;$$

$$Q_V = 1000 \text{ mol/h}; M = 15; N = 200; \Delta P = \frac{P_N}{N} = \frac{100}{200} = 0.5 \text{ mol}; \Delta x_B = 0.01.$$

A code for calculating the optimal policy was written and the result is presented in Tab. 4.7.

**Tab. 4.7:** The optimal operation policy of the batch distillation in Example 4.2.

Stage no. (i)	Total time (h)	Composition in the bottom ( $x_B$ , mol%)	Total distilled quantity ( $P$ , mol)	Values on interval				
				Duration of the interval (h)	Distilled quantity ( $P_i$ , mol)	Distillate flow rate ( $D$ , mol/h)	Reflux rate	Distillate composition ( $x_D$ , mol%)
1	0.030	0.1200	7.500	0.030	7.5	251.890	2.97	0.8567
2	0.060	0.1800	15.000	0.030	7.5	250.978	2.98	0.8367
3	0.091	0.1700	22.000	0.031	7.0	223.159	3.43	0.8625
4	0.123	0.1600	29.000	0.032	7.0	221.111	3.52	0.8229
5	0.155	0.1500	36.000	0.032	7.0	217.711	3.59	0.8029
6	0.188	0.1400	43.000	0.033	7.0	213.059	3.69	0.8931
7	0.223	0.1300	49.500	0.035	6.5	184.912	4.41	0.8131
8	0.259	0.1200	56.000	0.036	6.5	179.078	4.58	0.7931
9	0.297	0.1100	62.500	0.036	6.5	171.953	4.82	0.7731
10	0.337	0.1000	69.000	0.040	6.5	163.561	5.11	0.7531
11	0.379	0.0900	75.500	0.042	6.5	153.896	5.50	0.7875
12	0.426	0.0800	81.500	0.047	6.0	126.877	6.88	0.7675
13	0.478	0.0700	87.500	0.052	6.0	115.988	7.62	0.7475
14	0.536	0.0600	93.500	0.058	6.0	103.775	8.64	0.7044
15	0.600	0.0500	100.000	0.064	6.5	101.705	8.83	0.6754

The optimal operation shows a nonlinear decrease of the distillate flow  $D$ , 2.5 times from the beginning, and a 3-fold nonlinear increase of the reflux rate. The result is a quasi-constant composition  $x_D$ . The reduction of the batch time, as compared to the classical operation (linear decrease of  $D$ ), is of 5%. In the volatile essential oils industry, a batch can be of 24-hour duration, and 5% means 1.2 hours per batch with the corresponding steam consumption. The result are process intensification and reduction of energy consumption.

### 4.3 Dynamic optimal control of continuous processes

This optimal control procedure targets to stabilize the behavior of an automatic control system (ACS), of minimizing the steady-state error, as well as the transient time, with the ultimate goal of increasing the economic efficiency of the control action.

The approach is based on the minimization of the *quadratic performance index* as  $J_{IAE}$  or  $J_{ITAE}$ . The system is described through the state equation:

$$\dot{\bar{x}} = A\bar{x} + B\bar{u} + C\bar{d} \quad (4.64)$$

where  $\bar{x}$  is the *state vector* (dimension  $nx1$ ),  $\bar{u}$  is the *control vector* (dimension  $mx1$ ), and  $\bar{d}$  is the *disturbance vector* (dimension  $px1$ ).

Such a system is derived from the mathematical model of a process [15], but with the condition that the equations are either linear or linearized.

The control vector that minimizes the performance index is:

$$J = \int_0^{\infty} P(\bar{x}, \bar{u}) dt \quad (4.65)$$

where  $P(\bar{x}, \bar{u})$  is a quadratic function of  $\bar{x}$  and  $\bar{u}$ .

It can be shown [16] that for such a performance index, where the limits are 0 and  $\infty$ , the resulting control law is a function of the state vector  $\bar{x}$ :

$$\bar{u}(t) = K\bar{x}(t) \quad (4.66)$$

where  $K$  is a matrix of dimensions  $m \times n$ .

The performance index to be minimized is quadratic,

$$J = \int_0^{\infty} (\bar{x}^T Q \bar{x} + \bar{u}^T R \bar{u}) dt \quad (4.67)$$

where  $Q$  and  $R$  are symmetric matrices, positive-definite, called *weight* matrices. These matrices determine the importance of the states of the system and the energy penalties *via* control actions, in the cost function. The vector  $\bar{u}$  is not supposed to be subjected to constraints.

If matrix  $K$  can be determined in such a way that the index  $J$  from (4.67) is minimized, then the vector  $\bar{u}(t)$  is optimal for any initial state  $\bar{x}(0)$ .

By substituting (4.66) in (4.64) and considering the disturbance as being 0,

$$\dot{\bar{x}} = (A - BK)\bar{x} \quad (4.68)$$

and then replacing eq. (4.66) in eq. (4.67), the performance index becomes:

$$J = \int_0^{\infty} \bar{x}^T (Q + K^T RK) \bar{x} dt \quad (4.69)$$

By considering that:

$$\bar{x}^T (Q + K^T RK) = -\frac{d}{dt} \bar{x}^T S \bar{x} \quad (4.70)$$

where  $S$  is a positive-definite symmetrical matrix, the following can be written:

$$\bar{x}^T (Q + K^T RK) \bar{x} = -\bar{x}^T S \dot{\bar{x}} - \bar{x}^T S \dot{\bar{x}} = \bar{x}^T [(A - BK)^T S + S(A - BK)] \bar{x} \quad (4.71)$$

Because this equation has to be true for any  $\bar{x}$ ,

$$(A - BK)^T S + S(A - BK) = -(Q + K^T RK) \quad (4.72)$$

If  $(A - BK)$  is a stable matrix, there is a positive-definite matrix  $S$  satisfying the equation (4.72). Because all *eigenvalues* of the matrix  $(A - BK)$  have real negative parts,

$$\bar{x}(\infty) \rightarrow 0 \text{ and } J = \bar{x}^T S \bar{x}(0).$$

$R$  is a symmetric positive-definite matrix,  $R = T^T T$ , where  $T$  is a nonsingular matrix, so that eq. (4.71) can be written in another way:

$$\begin{aligned} (A^T - K^T B)S + S(A - BK) + Q + K^T T^T T K &= 0 \text{ or} \\ A^T S + SA + [TK - (T^T)^{-1} B^T S]^T [TK - (T^T)^{-1} B^T S] - SBR^{-1} B^T S + Q &= 0 \end{aligned} \quad (4.73)$$

Minimization of  $J$  relative to  $K$  requires the minimization of

$$\bar{x}^T [TK - (T^T)^{-1} B^T S]^T [TK - (T^T)^{-1} B^T S] \bar{x} \text{ relative to } K. \quad (4.74)$$

Because this value is nonnegative, the minimum is located where the expression (4.74) is equal to 0 and  $TK = (T^T)^{-1} B^T S$ .

The compensating matrix

$$K = T^{-1} (T^T)^{-1} B^T S = R^{-1} B^T S \quad (4.75)$$

Matrix  $S$  from (4.75) has to satisfy eq. (4.72) or the reduced Riccati equation

$$A^T S + SA - SBR^{-1} B^T S + Q = 0 \quad (4.76)$$

- This way, the synthesis procedure of the optimal control is reduced to:
- finding the solution of the reduced matrix Riccati equation relative to  $S$ ;
  - substitution of the matrix  $S$  in equation (4.76).

**Example 4.3**

Considering the system described through

$$\dot{\bar{x}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \bar{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \bar{u} \text{ and using control law (4.66)}$$

$$\bar{u}(t) = -K\bar{x}(t)$$

$K$  has to be determined to minimize the performance index:

$$J = \int_0^{\infty} (\bar{x}^T Q \bar{x} + \bar{u}^T \bar{u}) dt \quad (4.77)$$

$Q$  has the form:

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \alpha \geq 0 \text{ and } R = [1] \quad (4.78)$$

The Riccati equation which has to be solved is (4.76) and thus:

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} + \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} [1] [0 \ 1] \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and further} \quad (4.79)$$

$$\begin{bmatrix} 0 & 0 \\ S_{11} & S_{12} \end{bmatrix} + \begin{bmatrix} 0 & S_{11} \\ 0 & S_{21} \end{bmatrix} - \begin{bmatrix} S_{12}^2 & S_{12}S_{22} \\ S_{12}S_{22} & S_{22}^2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

resulting the equations ( $S$  is symmetric):

$$\begin{aligned} 1 - S_{12}^2 &= 0 \\ S_{11} - S_{12}S_{22} &= 0 \\ \alpha + 2S_{12}S_{22}^2 &= 0 \end{aligned} \quad (4.80)$$

From here,

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} \sqrt{\alpha+2} & 1 \\ 1 & \sqrt{\alpha+2} \end{bmatrix} \text{ and the optimal control matrix}$$

$K = R^{-1}B^T S = [1 \ \sqrt{\alpha+2}]$  and the optimal control law:

$$u = -K\bar{x} = -x_1 - \sqrt{\alpha+2} x_2 \quad (4.81)$$

**i** Example 4.4 [17]

Optimal control of a binary distillation column (Fig. 4.6). The column has the control configuration given in the figure below. An optimal control law of the distillate composition has to be written by using the distillate flow ( $D$ ) as control variable when the feed flow ( $F$ ) suffers disturbances.

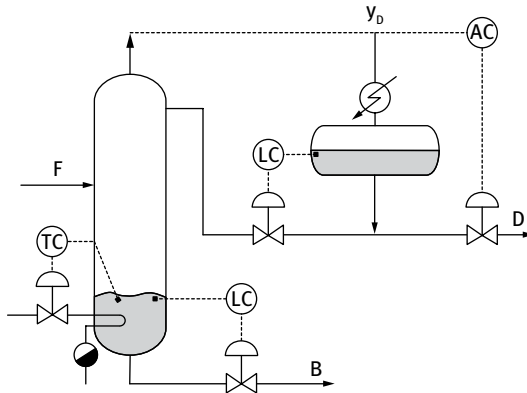


Fig. 4.6: Binary continuous distillation column with the corresponding control loops.

The block scheme of the distillation process is presented in Fig. 4.7.

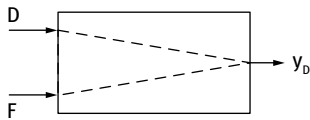


Fig. 4.7: The block scheme of the distillation process, having input  $D$  (distillate flow) as control action,  $F$  (feed flow) as disturbance, and  $y_D$  (distillate concentration) as output.

The influence of the control action ( $D$ ) and the disturbance ( $F$ ) on the distillate composition,  $y_D$ , is given by the *transfer function*:

$$Y_D(s) = G_1(s)D(s) + G_2(s)F(s) = D_e(s) + F_e(s) \tag{4.82}$$

where  $G_1(s)$  and  $G_2(s)$  are the transfer functions corresponding to the transfer paths  $D \rightarrow y_D$  and  $F \rightarrow y_D$  and  $D_e$  and  $F_e$  are equivalent transfer functions.

For  $G_1$ , the expression found in [17] is:

$$G_1(s) = \frac{K_1}{T_1s + 1} = -\frac{0.031}{0.55s + 1}. \tag{4.83}$$

Hence,

$$Y_D(s) = -\frac{0.031}{0.55s + 1} D(s) + F_e(s) \text{ and when passing to time domain then } \tag{4.84}$$

$$0.55 \frac{dy_D(t)}{dt} + y_D(t) = -0.031D(t) + d_1(t) \tag{4.85}$$

where  $d_1(t)$  is the expression in time of the effect of the feed disturbance on the distillate composition.

When trying to put the equation in the form of the state equation (4.64), the result is:

$$\frac{dy_D(t)}{dt} = -\frac{1}{0.55}y_D(t) - \frac{0.031}{0.55} D(t) - \frac{1}{0.55} d_1(t) \tag{4.86}$$

By noting the variables  $x_1 = y_D - y_{D,ref}$ ,  $u = D - D_{ref}$ ,  $d = d_1 - d_{1,ref}$ , equation above becomes the state equation (4.64):

$$\frac{dx_1}{dt} = ax_1(t) + bu(t) + cd(t) \text{ with } a = -\frac{1}{0.55}; b = 0.031a; c = a$$

At  $t = 0$ , the feed disturbance is considered to occur (step disturbance  $d(t) = d_0$ ), the previous deviations of the variables being considered 0:  $x_1(0) = 0$  and  $d(0) = 0$ .

We note with  $x_2$  the state variable,  $x_2 = bu(t) + cd_0$  and  $u_1 = u(t)$ .

The previous state equation becomes thus a system:

$$\begin{aligned} \frac{dx_1}{dt} &= ax_1(t) + x_2(t) \\ \frac{dx_2(t)}{dt} &= bu_1(t) \end{aligned} \tag{4.87}$$

The proposed performance index is:

$$J = \frac{1}{2} \int_0^{t_f} (x_1^2 + ru^2)dt = \frac{1}{2} \int_0^{t_f} [(y_D - y_{D,ref})^2 + rD^2]dt \tag{4.88}$$

in which the first term represents the deviation of the distillate composition from the steady-state value while the second stands for the cost of the distillate flow change.  $r$  is the penalty for the change and has to be  $\leq 1$ .

$u_1$  is sought after. By applying the optimal control law (4.66) and relationship (4.75), then:

$$\bar{u}_1(t) = K \bar{x}_1(t) = -R^{-1}B^T S \bar{x}_1(t) \tag{4.89}$$

where  $[R] = [r]$ ;  $[B]^T = \begin{bmatrix} 0 & -\frac{0.031}{0.55} \end{bmatrix}$ ;  $[S] = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$  where  $S_{ij}$  are calculated from the Riccati equation (4.76).

$$\begin{aligned} &\begin{bmatrix} -\frac{1}{0.55} & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} + \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} -\frac{1}{0.55} & 1 \\ 0 & 0 \end{bmatrix} - \\ &\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{0.031}{0.55} \end{bmatrix} [r] \begin{bmatrix} 0 & -\frac{0.031}{0.55} \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ equivalent with} \end{aligned}$$



$$\begin{aligned}
-\frac{2}{0.55} S_{11} + \frac{0.031^2}{0.55^2} S_{11} S_{21} r &= -1 \\
S_{11} - \frac{1}{0.55} S_{12} - \frac{0.031^2}{0.55^2} S_{11} S_{21} r &= 0 \\
S_{11} - \frac{1}{0.55} S_{21} - \frac{0.031^2}{0.55^2} S_{21} S_{22} r &= 0 \\
S_{12} + S_{22} - \frac{0.031^2}{0.55^2} S_{22}^2 r &= 0
\end{aligned} \tag{4.90}$$

The optimal solution is then:

$$\begin{aligned}
u_1(t) &= \begin{bmatrix} -\frac{1}{r} \\ 0 \end{bmatrix} \begin{bmatrix} 0 & -\frac{0.031}{0.55} \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{0.031}{0.55} (S_{21}x_1 + S_{22}x_2) \text{ or} \\
u(t) &= \frac{0.056}{r} \left[ S_{21}x_1 + S_{22} \left( x_1 + \frac{1}{0.55} x_1 \right) \right] = \frac{0.056}{r} \left( S_{21} + \frac{1}{0.55} S_{22} \right) x_1 + \frac{0.056}{r} S_{22} x_1
\end{aligned} \tag{4.91}$$

Integrating,

$$u(t) = \frac{0.056}{r} S_{22} x_1(t) + \frac{0.056}{r} \left( S_{21} + \frac{1}{0.55} S_{22} \right) \int_0^{t_f} x_1(t) dt + u_0 \tag{4.92}$$

we find  $S_{22}$  by solving the equation:

$$\begin{aligned}
4q^4 S_{22}^6 - 4(pq^3 + 2q^3) S_{22}^5 - (p^2 q^2 - 8pq^2 + 4q^2 - 4pq^3) S_{22}^4 - (2p^2 q + 4pq - 4p^2 q^2 + 8pq^2) \\
\times S_{22}^3 + (p^2 - 4p^2 q - p^3 q) S_{22}^2 - (p^3 - q) S_{22} - \frac{1}{2} p = 0,
\end{aligned} \tag{4.93}$$

$S_{21}$  is calculated from equation:

$$S_{21} = -\frac{1 + pS_{11}}{qS_{11}} \tag{4.94}$$

and  $S_{11}$  from:

$$S_{11} = S_{22} (qS_{22} - 1) (2S_{22}q - p) \tag{4.95}$$

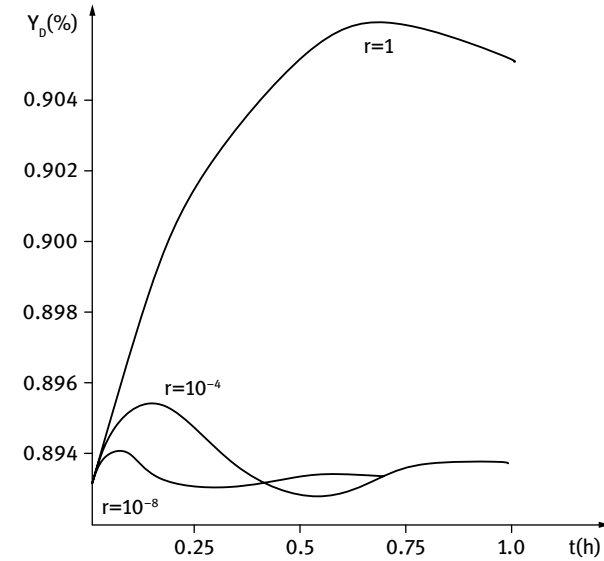
where  $p = -\frac{2}{0.55}$  and  $q = \frac{0.031^2}{0.55^2} r$ .

Consequently, from eqs. (4.92)–(4.95), the optimal control law is:

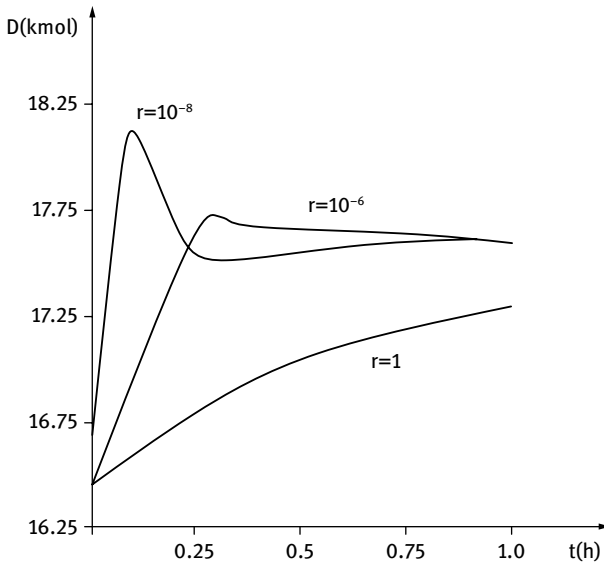
$$D(t) = \frac{0.056}{r} - S_{22} (y_D - y_{D,ref}) + \frac{0.056}{r} \left( S_{21} + \frac{1}{0.55} S_{22} \right) \int_0^{t_f} (y_D - y_{D,ref}) dt + D_0 \tag{4.96}$$

where  $D_0$  is the distillate flow at  $y_D = y_{D,ref}$

It can be observed that a PI-type control function is obtained, but which is depending on the penalty imposed to the second term (Fig. 4.8).



(a)



(b)

**Fig. 4.8:** The optimal control of the distillate purity ( $y_D$ ): (a) through the distillate flow ( $D$ ) and (b) depends on the penalty term.

## References

- [1] Kirk, D.E., *Optimal Control Theory – An Introduction*, Prentice Hall, 4, 2004.
- [2] Zhang, G.P., Rohani, S., *On line optimal control of a seeded batch cooling crystallizer*, Chemical Engineering Science, 58(9), 1887, 2003.
- [3] Sheikhzadeh, M., Trifkovic, M., Rohani, S., *Real-time optimal control of an anti-solvent isothermal crystallization process*, Chemical Engineering Science, 63(3), 829, 2008.
- [4] Smets, I., Claes, J., November, E., Bastin, G., Van Impe, J., *Optimal adaptive control of (bio)chemical reactors: past, present and future*, Journal of Process Control, 14, 795, 2004.
- [5] Manon, Ph., Valentin-Roubinet, C., Gilles, G., *Optimal control of hybrid dynamical systems: application in process engineering*, Control Engineering Practice, 10(2), 133, 2002.
- [6] Engell, S., *Feedback control for optimal process operation*, Journal of Process Control, 17(3), 203, 2007.
- [7] Agachi, P.S., Cristea, V. M., *Basic Process Engineering Control*, Walter de Gruyter, Berlin/ Boston, 297, 2014.
- [8] Yi, C., *Learning PID Tuning III: Performance Index Optimization, a Tool and Tutorial to Perform Optimal PID Tuning*, File Exchange, Matlab Central, <http://www.mathworks.com/matlab-central/fileexchange/18674-learning-pid-tuning-iii--performance-index-optimization/content/html/optimalpidtuning.html>.
- [9] Imre Lucaci, A., Agachi P.S., *Optimizarea proceselor din industria chimica (Optimization of the Chemical Industry Processes)*, Editura Tehnica, Bucuresti, Ch. 5, 2002.
- [10] Imre Lucaci, A., Agachi P.S., *Optimizarea proceselor din industria chimica (Optimization of the Chemical Industry Processes)*, Editura Tehnica, Bucuresti, Ch. 7, 2002.
- [11] Agachi, S.P., Vass, E., *Optimization of a methanol reactor using Pontryagin's maximum principle*, Revista De Chimie, 6, 513–521, 1995.
- [12] Natta, G., Pino, P., Mazzanti, G., Pasquon, J., *Kinetics of methanol synthesis reaction*, Chim. e Ind. (Milano) 35(6), 705, 1953.
- [13] Agachi, S., *Automatizarea proceselor chimice (Chemical Process Control)*, Ed. Casa Cartii de Stiinta, Cluj, 337, 1994.
- [14] Seader, J.D., *Perry's Chemical Engineer's Handbook*, 6th Ed., McGraw-Hill, New York, NY, 13–37 – 13–39, 1984.
- [15] Agachi, P.S., Cristea, V.M., *Basic Process Engineering Control*, Walter de Gruyter, Berlin/Boston, 63, 2014.
- [16] Singh, M., *Applied Industrial Control: An Introduction (International Series of Systems and Control)*, Elsevier, prg.5.6, 1980.
- [17] Bozga, G., *Conducerea cu calculatoare a proceselor chimice (Chemical process computer control)*, Institutul Politehnic București, 224, 1989.

# 5 Multivariable control

## 5.1 Introduction

Usually, the control of a plant aims to keep more than one of its outputs at desired setpoints, either constant or changing, and to counteract the action of disturbances, to provide both safe and efficient operation and to conform to environmental requirements. The single variable control is in the large majority of cases not sufficient to fulfil the multiple requests demanded by the control tasks and the need for the control of multiple output variables is obvious. Multivariable control is the solution for multiple-input-multiple-output (MIMO) processes to fulfil their associated multiple control objectives. Multivariable control involves multiple manipulated variables and the design of the control system becomes complex, sometimes asking for the support of advanced control methodologies.

Although the general control concepts of the single loop control are also valid for the multivariable control, there are some additional and particular aspects derived from the multiple-variables characteristic. Some of these aspects are presented in the following.

The starting point of designing any multivariable control system is to clearly define the control objectives. From the control objectives arises the need for specifying which of the the process outputs to be controlled. Choosing the controlled outputs has to rely first on the analysis of the possibility to either directly measure the intended to control outputs or to infer their values from other available measured outputs.

Following the selection of the controlled outputs, the selection of the manipulated variables (inputs) is of first role importance. This category of process input variables are the handles used by the controllers to act on the controlled outputs. Additionally, the analysis must be carried on to identify other relevant process inputs affecting the controlled outputs, i.e. the disturbances. Having both qualitative and quantitative knowledge on the disturbances means offering valuable information for the control system design. Particularly, the identification of the measured disturbances offers the opportunity to perform feedforward control and efficiently counteract part of the undesired effects of the disturbances. Comprehensive information on the disturbances is useful for the multivariable control system design, as it may reveal both the source of disturbances and the amplitude of their effect on the different controlled variables.

The major problems for the multivariable control design originate in the existence of the interaction among inputs and outputs of the MIMO systems, i.e. a manipulated variable may influence several controlled variables. This feature, not considered in the case of the single variable control design, may affect not only the multivariable control performance but even the stability of the controlled process.

The interaction among the process' variables that are controlled may produce upsets in all the other controlled (or noncontrolled) variables. Design methods for the multivariable control should be developed to either reduce the effects of the interactions or eliminate them.

There are two fundamental approaches for the multivariable control. They are the *multiloop control*, also known as *decentralized control*, and the *centralized (coordinated) control* [1]. The multiloop control uses multiple single variable controllers where each manipulated variable is assigned to an output variable (and establishes a control loop), with the aim of controlling the output. The centralized control takes into account all measured outputs and simultaneously generates all manipulated variables in such a way that process interacting effects are counteracted when the manipulated variables computation is performed.

Subsequent to the selection of the sets of controlled and manipulated variables, the multiloop control design implies the specification of the control loops configuration. This task consists in deciding the *pairing* of the controlled variables and the manipulated variables. This pairing is of most importance for reducing the interaction effects among the multiple control loops and, as a result, for the attainable control performance desired for the multivariable control system [1–3].

For the processes where the number of controlled and the number of manipulated variables are not the same, the multivariable control design has to cope with this circumstance and provide specific control methods.

## 5.2 Multiloop control

The most commonly applied approach of the multivariable control is the multiloop control strategy, although from theoretical point of view, the centralized control presents obvious incentives. However, the latter also may have to comply with possible implementation impediments [1].

The multiloop control uses several single-loop controllers, each of them devoted to control one process output variable. Besides historical reasons of precedency, the incentives of this control approach rely on its simplicity, both with respect to the use of unsophisticated architecture and to the exploitation of common control laws, such as the PID algorithm. Furthermore, the multiloop control is easily understood and accepted by the operators especially when their direct intervention is needed for coping with instrumentation malfunction or significant process upsets.

### 5.2.1 Interaction among control loops

The key aspect in designing the multiloop control system is the interaction among the manipulated and controlled variables, i.e. the way each input (manipulated) variable influences all of the output variables of the process. As a result, any

single-loop controller can no longer be considered decoupled with respect to the other single-loop controllers. The presence of the process interaction induces interaction among the individual control loops. Example 5.1 presents a simple but intuitive process with interaction among two control loops.

### Example 5.1

Consider a mixing process of two liquid streams presented in Fig. 5.1.

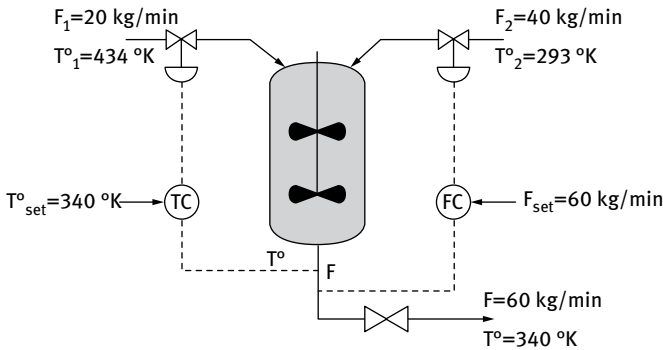


Fig. 5.1: Interacting flow and temperature control loops in the mixing tank.

The streams entering the mixing tank have the inlet mass flows  $F_1$  and  $F_2$ , ( $F_1 < F_2$ ), and the inlet temperature  $T_1^\circ$  and  $T_2^\circ$ , ( $T_1^\circ > T_2^\circ$ ). The outlet mass flow  $F$  has the outlet temperature  $T^\circ$ . It is assumed that mixing is perfect and the densities of the input flows are equal. The specific heats of the inlet flows are equal, too. The mixing process goal is to keep both the outlet mass flow and outlet temperature at desired setpoints, in order to provide smooth operation for the downstream unit. This is performed by the outlet flow  $FC$  and outlet temperature  $TC$  controllers presented in Fig. 5.1 [3].

Consider the system to be initially in steady state. Suppose that a change in the setpoint  $F_{set}$  of the mass flow controller is made to increase the desired outlet flow. This task is accomplished by the flow controller  $FC$  by increasing the inlet flow  $F_2$ . As the inlet flow  $F_2$  is changed (increases) by  $FC$ , the temperature in the mixing tank  $T^\circ$  changes also (decreases). This change is due to the different (lower) temperature of the  $F_2$  stream, compared to the temperature of the  $F_1$  stream. Consequently, as the temperature in the mixing tank is changed from its initial value (equal to the setpoint value), the temperature controller starts acting by changing (increasing) the inlet flow  $F_1$ . In conclusion, the setpoint change in the flow control loop produces the need for control intervention provided by the temperature control loop. The same interacting behavior may be observed if a change of the outlet temperature setpoint  $T^\circ_{set}$  is initiated, resulting in the need for control action in the flow control loop. The two control loops are also showing interacting effects if a disturbance is acting on any of the inlet flows or temperatures.

The interaction among the control loops may destabilize the process and, as the number of control loops is increasing, this undesired effect may be intensified. Therefore, a new task is coming to the control design. This task is the loop pairing and consists in specifying for each controlled output the corresponding manipulated variable to be used for its control, in such a way that the loop interaction to be minimal [1, 2, 8].

In the following, the case of an interacting control system for a multivariable process with two manipulated variables and two controlled outputs will be considered. The  $2 \times 2$  open-loop system is presented in Fig. 5.2 [1, 2].

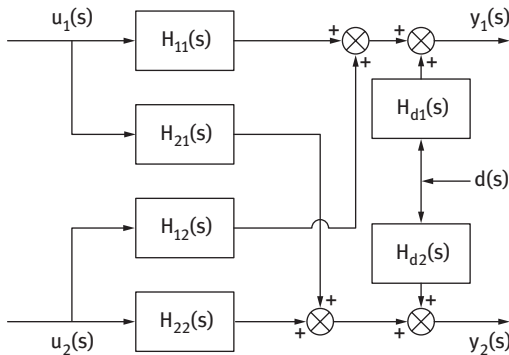


Fig. 5.2: Diagram of the  $2 \times 2$  open-loop system.

The transfer matrix of the  $2 \times 2$  open-loop system may be described by

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \begin{bmatrix} H_{11}(s) & H_{12}(s) \\ H_{21}(s) & H_{22}(s) \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} + \begin{bmatrix} H_{d1}(s) \\ H_{d2}(s) \end{bmatrix} d(s), \quad (5.1)$$

where  $H_{ij}(s)$  is the transfer function relating output  $y_i(s)$  to input  $u_j(s)$  and  $H_{di}(s)$  is the transfer function relating output  $y_i(s)$  to disturbance  $d(s)$ .

For the multivariable process to be steady-state controllable, i.e. the controlled outputs to be brought at steady state to the desired setpoints in the presence of disturbances, the inverse of the steady-state gain matrix should exist. This is equivalent to the condition that the determinant of the steady-state gain matrix  $\mathbf{H}_s$  is nonzero [1].

### **i** Example 5.2

The transfer matrix for the mixing process described in Example 5.1 can be obtained from the mass and heat balance equations:

$$\frac{dm}{dt} = F_1 + F_2 - F \quad (5.2)$$

$$\frac{d}{dt}(mc_p T^\circ) = F_1 c_p T_1^\circ + F_2 c_p T_2^\circ - F c_p T^\circ \quad (5.3)$$

Considering that the mass of liquid in the mixing tank ( $m = 600$  kg) is not changing,  $dm/dt = 0$ , eqs. (5.2) and (5.3) become:

$$F = F_1 + F_2 \quad (5.4)$$

$$m \frac{dT^\circ}{dt} = F_1 T_1^\circ + F_2 T_2^\circ - F T^\circ \quad (5.5)$$

Linearization of eq. (5.5) around the steady-state values of variables presented in Fig. 5.1 is leading to the following linear forms, using the deviation variables  $\Delta T^\circ = T^\circ - T_s^\circ$ ,  $\Delta F_1 = F_1 - F_{1s}$ , and  $\Delta F_2 = F_2 - F_{2s}$  (subscript *s* has been used for *steady state*).

$$m \frac{d\Delta T^\circ}{dt} = T_1^\circ \Delta F_1 + T_2^\circ \Delta F_2 - T_s^\circ \Delta F_1 - T_s^\circ \Delta F_2 - (F_{1s} + F_{2s}) \Delta T^\circ \quad (5.6)$$

and

$$\frac{m}{(F_{1s} + F_{2s})} \frac{d\Delta T^\circ}{dt} + \Delta T^\circ = \left( \frac{F_{2s}(T_1^\circ - T_2^\circ)}{(F_{1s} + F_{2s})^2} \right) \Delta F_1 + \left( \frac{F_{1s}(T_2^\circ - T_1^\circ)}{(F_{1s} + F_{2s})^2} \right) \Delta F_2 \quad (5.7)$$

The matrix transfer function of the system becomes

$$\mathbf{H}(s) = \begin{bmatrix} H_{11}(s) & H_{12}(s) \\ H_{21}(s) & H_{22}(s) \end{bmatrix} = \begin{bmatrix} \frac{1}{F_{2s}(T_1^\circ - T_2^\circ)} & \frac{1}{F_{1s}(T_2^\circ - T_1^\circ)} \\ \frac{m}{(F_{1s} + F_{2s})^2} & \frac{m}{(F_{1s} + F_{2s})^2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1,567 & -0.783 \end{bmatrix} \begin{matrix} 1 \\ 10s + 1 \\ 10s + 1 \end{matrix} \quad (5.8)$$

revealing a time constant of  $T_p = 10$  [min] for the temperature change.

Finally, the static (steady-state) gain matrix of the open-loop process has the following form:

$$\mathbf{H}_s = \begin{bmatrix} 1 & 1 \\ 1,567 & -0.783 \end{bmatrix} \quad (5.9)$$

As  $\det(\mathbf{H}_s) = -2.349 \neq 0$ , it may be concluded that the process is controllable with the chosen pairing of the manipulated and controlled variables.

The selection of the manipulated and controlled variables for building the multiloop control configuration aims to find the multivariable control structure that presents the weakest interaction among the individual control loops. For *n*-manipulated and *n*-controlled variables, the number of possible pairing is equal to *n!* alternatives [1, 2].

The multiloop control system must cope both with circumstances when all loops work in control mode and with situations when only part of the individual loops are operating in automatic mode and the others are in manual mode.

The interaction effect on the control performance is investigated by considering two cases. The first one assesses the closed-loop operation and the transfer function when only one loop is closed (in automatic mode) and all the other loops are in open loop (in manual mode). The second evaluation is performed for the case when all control loops are closed. For simplicity of presentation, this investigation is done for the



$2 \times 2$  loop control system presented in Fig. 5.3, but the extension to a larger dimension of the MIMO control is straightforward. For the same reason, the transfer function of the final control elements and measuring devices are considered equal to unity.

Fig. 5.3 shows the first case when *one loop is closed* and the other is open.

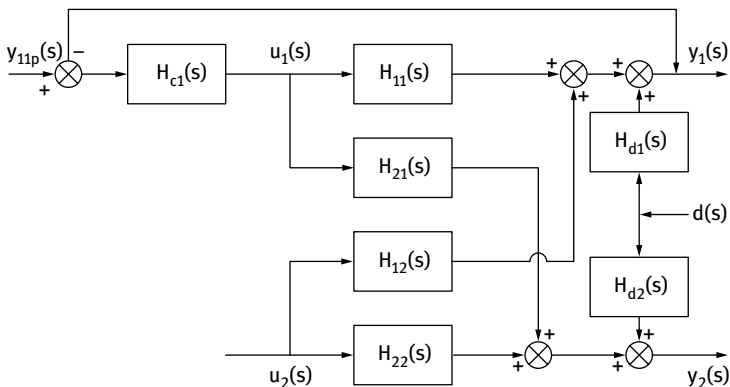


Fig. 5.3: Interaction assessment for the case when only one loop is closed.

It may be noticed that both outputs are influenced by the setpoint of the closed loop, according to the following transfer functions:

$$y_1(s) = \frac{H_{c1}(s)H_{11}(s)}{1 + H_{c1}(s)H_{11}(s)} y_{1sp}(s) + H_{d1}(s)d(s) \quad (5.10)$$

$$y_2(s) = \frac{H_{c1}(s)H_{21}(s)}{1 + H_{c1}(s)H_{11}(s)} y_{1sp}(s) + H_{d2}(s)d(s) \quad (5.11)$$

This is due to the transfer path provided by  $H_{21}(s)$ .

The second case, when *all control loops are closed*, is presented in Fig. 5.4.

When both loops are closed, the change of the loop 1 setpoint  $y_{1sp}$  is producing a change of the controlled output  $y_1$  due to the action of the loop 1 control system. This is the direct effect of the manipulated variable  $u_1$  change on output  $y_1$  and it is represented in Fig. 5.4 by the dash-dotted line. At the same time, the change of the manipulated variable  $u_1$  is producing a change on output  $y_2$  due to the  $H_{21}$  transfer path. As a result,  $y_2$  deviates from its desired setpoint  $y_{2sp}$  and loop 2 starts to operate for counteracting the developed offset. Meanwhile, the action of the manipulated variable  $u_2$  of loop 2 produces a second, indirect effect on output  $y_1$ , due to the  $H_{12}$  transfer path. This indirect effect is represented in Fig. 5.4 by the dashed line. Moreover, due to this new change of output  $y_1$ , the control action is again activated in loop 1 and this produces a new change of  $y_2$ . This interaction between the control loops induces oscillating behavior on both controlled variables, producing an extended settling time or, in the extreme case, instability [7]. Similar behavior may be initiated by the change of the loop 2 setpoint or by the action of disturbances.

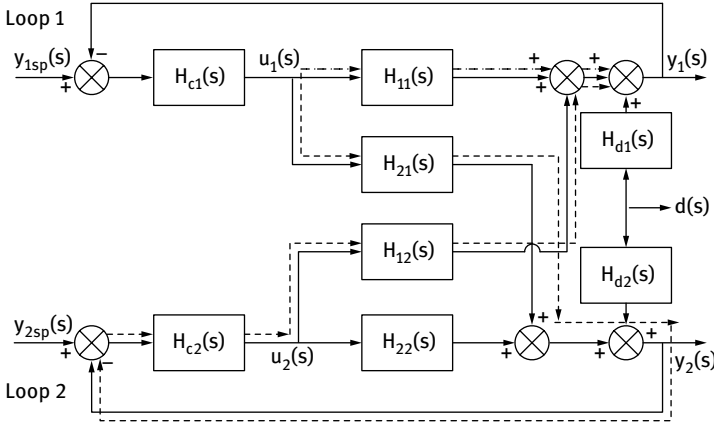


Fig. 5.4: Interaction assessment for the case when all loops are closed.

The closed-loop transfer functions of the controlled system, when both control loops are closed, may be described by:

$$y_1(s) = G_{11}(s)y_{1sp}(s) + G_{12}(s)y_{2sp}(s) + H_{d1}(s)d(s) \tag{5.12}$$

$$y_2(s) = G_{21}(s)y_{1sp}(s) + G_{22}(s)y_{2sp}(s) + H_{d2}(s)d(s) \tag{5.13}$$

with

$$G_{11}(s) = \frac{H_{11}H_{c1} + (H_{11}H_{22} - H_{12}H_{21})H_{c1}H_{c2}}{(1 + H_{11}H_{c1})(1 + H_{22}H_{c2}) - H_{12}H_{21}H_{c1}H_{c2}} \tag{5.14}$$

$$G_{12}(s) = \frac{H_{12}H_{c2}}{(1 + H_{11}H_{c1})(1 + H_{22}H_{c2}) - H_{12}H_{21}H_{c1}H_{c2}} \tag{5.15}$$

$$G_{21}(s) = \frac{H_{21}H_{c1}}{(1 + H_{11}H_{c1})(1 + H_{22}H_{c2}) - H_{12}H_{21}H_{c1}H_{c2}} \tag{5.16}$$

$$G_{22}(s) = \frac{H_{22}H_{c2} + (H_{11}H_{22} - H_{12}H_{21})H_{c1}H_{c2}}{(1 + H_{11}H_{c1})(1 + H_{22}H_{c2}) - H_{12}H_{21}H_{c1}H_{c2}} \tag{5.17}$$

The transfer functions developed for the two investigated cases show that to ensure stability for both operating scenarios, i.e. when either of the loops is closed and the other one is open loop, and when both loops are closed, it is necessary that simultaneously the roots of the following characteristic equations should have negative real parts:

$$\begin{aligned} 1 + H_{c1}(s)H_{11}(s) &= 0, \\ 1 + H_{c2}(s)H_{22}(s) &= 0, \\ (1 + H_{11}H_{c1})(1 + H_{22}H_{c2}) - H_{12}H_{21}H_{c1}H_{c2} &= 0 \end{aligned} \tag{5.18}$$

The conclusion is that due to the interacting effects the controllers should be tuned simultaneously to provide the desired control performance and stability.

### 5.2.2 Pairing the control loops

The interacting effects of the multiloop control system is dependent on the form of the MIMO system transfer matrix  $\mathbf{H}(s)$ . A diagonal form of  $\mathbf{H}(s)$  is revealing the lack of interaction. The selection of the pairing among manipulated variables and controlled variables for finding the multiloop control structure with minimum interaction between the loops is based on the Relative Gain Array (RGA). RGA, also called Bristol array, was first introduced by Bristol [5], and it measures the interaction in the process [6]. The elements of the RGA matrix, denoted by  $\Lambda$ , consist in the ratio of the open loop to the close loop gain for each input-output path of the process. The generic element of the RGA matrix  $\lambda_{ij}$ , associated to the transfer path starting from input  $u_j$  and finishing to output  $y_i$ , is described by either of the following forms of the ratio:

$$\lambda_{ij} = \frac{\left(\frac{\Delta y_i}{\Delta u_j}\right)_{u_i = \text{const.}, l \neq j}}{\left(\frac{\Delta y_i}{\Delta u_j}\right)_{y_p = \text{const.}, p \neq i}} \quad \text{or} \quad \lambda_{ij} = \frac{\left(\frac{\partial y_i}{\partial u_j}\right)_{u_i = \text{const.}, l \neq j}}{\left(\frac{\partial y_i}{\partial u_j}\right)_{y_p = \text{const.}, p \neq i}} \quad (5.19)$$

The numerator of  $\lambda_{ij}$ , i.e. the open-loop gain, is the corresponding element of the static gain matrix  $\mathbf{H}_s$  and shows the gain between input  $u_j$  and output  $y_i$  when only one loop (the one involving input  $u_j$  and output  $y_i$ ) is closed and the others are open. The denominator of  $\lambda_{ij}$ , i.e. the closed loop gain, is the steady-state gain between input  $u_j$  and output  $y_i$  when all control loops are closed [1, 6].

The RGA matrix can be determined either by experiments or by computation methods. The first methodology implies two experiments for obtaining the numerator and denominator of each  $\lambda_{ij}$ , and the second one computes the Hadamard product between the static gain matrix  $\mathbf{H}_s$  and its inverse transposed matrix  $(\mathbf{H}_s^{-1})^T$ :

$$\Lambda = \mathbf{H}_s * [\mathbf{H}_s^{-1}]^T \quad (5.20)$$

where the symbol  $*$  denotes the element by element matrix multiplication [3].

The RGA matrix is scale independent and the sum of its elements on each row or column are equal to unity.

The values of the RGA matrix may be interpreted for the selection of the multiple-loop control pairing, according to the following remarks [1]:

1. When  $\lambda_{ij} = 0$ , there is no direct steady-state open-loop relationship between input  $u_j$  and output  $y_i$ . It means this pairing is not desirable and in the majority of cases it is not acceptable. Nevertheless, when all loops are closed the controller paired with the value of  $\lambda_{ij} = 0$  will operate and in very particular cases this pairing might be acceptable (due to its indirect effect).

2. When  $\lambda_{ij} = 1$ , the change of input  $u_j$  directly influences only output  $y_i$  and not by other indirect paths (of the other closed loops), i.e. the loop paired according to  $\lambda_{ij} = 1$  is decoupled with respect to the other loops. It means this pairing is the most desirable to be applied for the control of output  $y_i$ . Nevertheless, it should be noticed that changes in input (manipulated)  $u_j$  used in this loop may also influence other controlled outputs.
3. When  $\lambda_{ij} < 0$ , the direct effect of input  $u_j$  on output  $y_i$  has a different sign compared to the indirect effect generated by the manipulated variables of the other closed loops. This may change the feedback of the control loop from negative to positive, with undesired consequences on the stability. This pairing is not acceptable (unless special measures for ensuring stability are introduced).
4. When  $0 < \lambda_{ij} < 1$ , the direct effect of input  $u_j$  on output  $y_i$  is smaller compared to the sum of direct and indirect effect generated by the manipulated variables of the other closed loops. As  $\lambda_{ij}$  approaches zero, this indirect effect is larger and the interaction effect is larger too. This pairing is acceptable especially when  $\lambda_{ij}$  approaches to 1.
5. When  $\lambda_{ij} > 1$ , the direct effect of input  $u_j$  on output  $y_i$  is larger compared to the sum of direct and indirect effect generated by the manipulated variables of the other closed loops. As  $\lambda_{ij}$  approaches 1, the indirect effect is smaller and the interaction effect is smaller too. This pairing is acceptable especially when  $\lambda_{ij}$  approaches to 1.
6. When  $\lambda_{ij} = \infty$ , the sum of the direct and indirect effects of input  $u_j$  on output  $y_i$  has a zero value and this means output  $y_i$  may not be controlled by input  $u_j$ . This pairing is not acceptable.

### Example 5.3

The computation of the RGA matrix for the mixing process described in Example 5.1 results in the following form of  $\mathbf{A}$ :

$$\mathbf{A} = \mathbf{H}_s * [\mathbf{H}_s^{-1}]^T = \begin{bmatrix} 1 & 1 \\ 1,567 & -0.783 \end{bmatrix} * \begin{bmatrix} 0.333 & 0.667 \\ 0.425 & -0.425 \end{bmatrix} = \begin{bmatrix} 0.333 & 0.667 \\ 0.667 & 0.333 \end{bmatrix} \begin{matrix} u_1 = F_1 & u_2 = F_2 \\ y_1 = F \\ y_2 = T^\circ \end{matrix} \quad (5.21)$$

According to the elements of  $\mathbf{A}$  from eq. (5.21) the most favorable pairing is the following. The mixer output mass flow  $F$  should be controlled by cold inlet flow stream  $F_2$  ( $\lambda_{12} = 0.667$ ) and the mixer output temperature  $T^\circ$  should be controlled by hot inlet flow stream  $F_1$  ( $\lambda_{21} = 0.667$ ). This pairing is already presented in Fig. 5.1. The appropriateness of the pairing suggested by the RGA matrix may be also confirmed by the phenomenological and practical assessment relying on the observation that the high value of the cold inlet flow stream  $F_2 = 40$  [kg/min] is able to influence output flow  $F$  in a greater extent than the cold inlet flow  $F_1 = 20$  [kg/min] would.

RGA is a measure of interaction for steady state, and this limitation should be considered as the transient regime is not revealed by it. Furthermore, the analytical computation of the RGA matrix is based on linearized models of the process. The accuracy of the linearized models may be dramatically affected when the operating point of the process changes significantly or the nonlinearity is pronounced. During



the last decades, important research effort has been devoted for developing a theoretical frame for using the frequency-dependent RGA in interaction assessment, but the results do not yet have the power of insight and generality characteristic to the steady-state RGA [9].

Measuring the interaction in a multiloop control system may be performed on the basis of the singular value decomposition of the static gain matrix  $\mathbf{H}_s$  [1, 14]. This decomposition reveals the directions in the controlled variables space that make the manipulated variables have the largest and the smallest change. The *condition number* is the ratio between the largest and the smallest changes in these identified directions, i.e. ratio between the maximum and the minimum singular values. As the condition number has a high value, it shows high interaction being present in the multiloop control system and it offers a measure for the pairing the control loops. Correlations between RGA and condition number may be performed when analysing pairing [10, 12].

One of the problems that emerged from the practice of multi-loop control is to cope with situations when part of the control loops are in automatic mode while the others are not in closed loop mode due to operation in manual mode or are out of service because of failures. Pairing of the control loops should consider and ground its design on the assessment of the control performance for such operation situations. Maintaining the integrity of the system means preserving the stability of the multiloop control system in case of operating with several control loops open and without the need of changing the signs of the feedback controllers [1, 13]. Assessment of the integral stabilizability may be performed by the *Niederlinski index* (NI) test, assuming that multiloop controllers are fitted with integral action. The NI can be computed by the following formula:

$$NI = \frac{\det(\mathbf{H}_s^*)}{\prod_{k=1}^n (\mathbf{H}_s^*)_{kk}} \quad (5.22)$$

where  $\mathbf{H}_s^*$  is the static gain matrix of the process, emerged from the  $\mathbf{H}_s$  gain matrix by arranging the loop pairing on the main diagonal of the matrix.

If the NI value is negative, the loop pairing design should not be accepted. Designs with positive NI may be considered as candidates for pairing the control loops.

### 5.2.3 Tuning the multiloop controllers

Tuning the multiloop controllers does not have a very well defined theoretical approach and clearly formulated methodology. However, there are general guidelines that may be used in association to the interaction analysis and the assessment of the disturbance effects on the multiloop control performance [11].

First, the tuning should take into consideration the objectives of the control problem and, directly emerged from that, generate a hierarchy of importance among

the controlled variables. To the most important of them, the multiloop control pairing and tuning of the associated control loops should provide tighter control performance (setpoint tracking, disturbance rejection, and fast response), while to the others, the performance might be less strict [1].

A first choice of tuning can be the experimental tuning made on the real process by repeated adjustments of the controllers' tuning parameters and by assessment the control performance or made on a dynamic simulator of the process. Results obtained from the latter greatly depend on the model accuracy and might be less precise, but may be suitable when producing experiments on the process is not convenient. Usually, for this tuning approach the starting values of the tuning parameters are those specific to the single-loop tunings (without taking into account the interaction), but adjusted for obtaining stability. In most of the cases, the tuning may proceed by detuning these initial tuning parameters' values for coping with the interaction between the control loops.

The best tuning parameters may be also obtained by minimizing one or several of the control performance indices (e.g. ISE, IAE, ITAE), on the basis of a dynamic simulator and by the use of efficient optimization algorithms. Again, the obtained best tuning parameters are dependent on the involved dynamic model.

The multiloop controllers should include the integral mode (effect) to eliminate the steady-state offset. RGA provides the basic information for the pairing design and is the main instrument used for this purpose. Niederlinski index and condition number are complementing the RGA information. The tuning is performed after the loop pairing step. Controlled variables having assigned higher importance will be paired with manipulated variables demonstrating fast effect. Additionally, an analysis may be performed to assess the effect of disturbances on the controlled variables. This may be done by the relative disturbance gain [8].

#### 5.2.4 Decoupling interaction for multiloop control

The RGA offers information for the design of the loop pairing such as the interaction between the multiloop control loops is minimal. However, it is possible that even for this best multiloop design and proper tuning, the control performance to be still poor. RGA proposes a “palliative” solution to treat the interacting problem, but a “surgical” one still exists. This is called decoupling and consists in the design of special elements, called decouplers, whose role is to reduce or even eliminate loops interaction by transforming the closed-loop transfer matrix in a new one, but with a diagonal form [1]. As a result, the control loops would operate such as they are independent and noninteracting. This solution introduces new control paths from each manipulated variable to all other controlled outputs, such as the added effect of loops interaction and the one generated by

the decoupler should compensate each other. The process inherent interactions, described by  $H_{ij}(s)$ , cannot be removed but the decouplers introduce new effects (paths) of the manipulated variables  $u_j$  on the controlled outputs  $y_i (i \neq j)$  in such a way to produce compensation for this interaction. As a result, the control loops will not interact anymore. The decoupler is designed according to the following relationship:

$$D_{ij}(s) = -\frac{H_{ij}(s)}{H_{ii}(s)} \tag{5.23}$$

The  $2 \times 2$  system with the associated multiloop controllers presented in Fig. 5.4 will be further used for showing the decoupler design. Consider a disturbance or setpoint change in loop 1. The manipulated variable  $u_1$  will change to bring the controlled output to the setpoint  $y_{1sp}$ . This change of  $u_1$  will produce a change on controlled output  $y_2$ , due to process interaction. To keep  $y_2$  unchanged, the decoupler  $D_{21}(s)$  will be introduced such as the process effect of  $u_1$  to  $y_2$  to be eliminated. The transfer function of this decoupler is

$$D_{21}(s) = -\frac{H_{21}(s)}{H_{22}(s)} \tag{5.24}$$

The control loops with  $D_{21}$  decoupler is presented in Fig. 5.5.

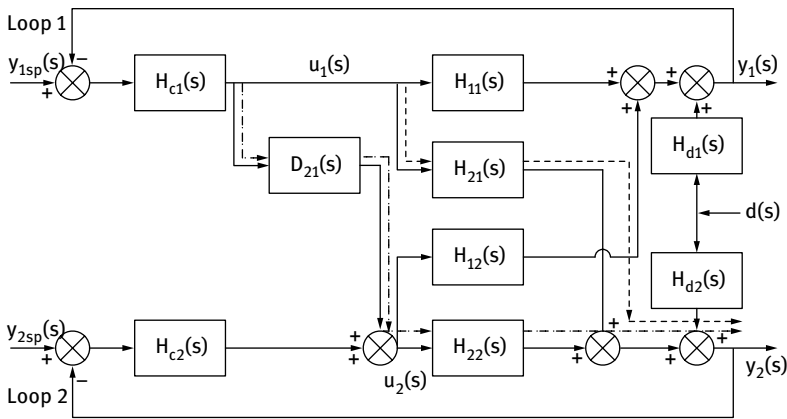


Fig. 5.5: Decoupling effect of Loop 1 on Loop 2.

As presented in Fig. 5.5, the effect of  $u_1$  on  $y_2$  (dotted line) is compensated by the effect introduced by the decoupler  $D_{21}$  path (dash-dotted line). The transfer function of the decoupler  $D_{21}(s)$  is obvious. This decoupler prevents the transmission of any change from loop 1 to loop 2 (one-way decoupling).

The same approach may be used to make the decoupling of loop 2 to loop 1, by designing the decoupler  $D_{12}$ :

$$D_{12}(s) = -\frac{H_{12}(s)}{H_{11}(s)} \tag{5.25}$$

The  $2 \times 2$  system with the associated multiloop controllers and decouplers is presented in Fig. 5.6.

It may be observed that the decoupling effect is very similar to the feedforward control and the decouplers are feedforward elements. The decoupler acts immediately as the manipulated variable changes and produces at every moment of time an effect that is equal but with opposed sign to the process interacting effect.

Depending on the time constants and dead time of the process transfer functions  $H_{ij}(s)$ , the design of the decoupler may be physically unrealizable. Partial decoupling or steady-state decoupling may be used in such cases.

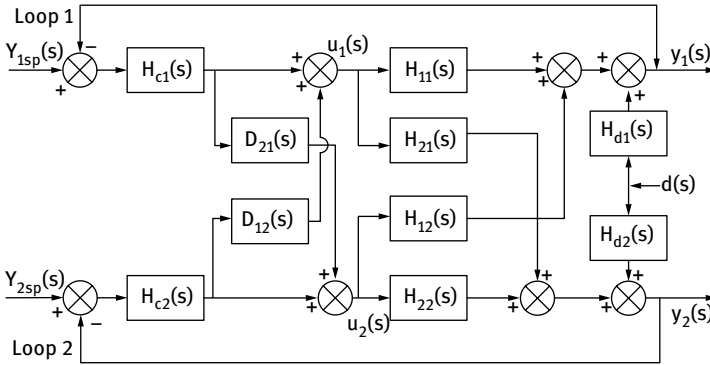


Fig. 5.6: Complete decoupling of the multiloop control for the  $2 \times 2$  system.

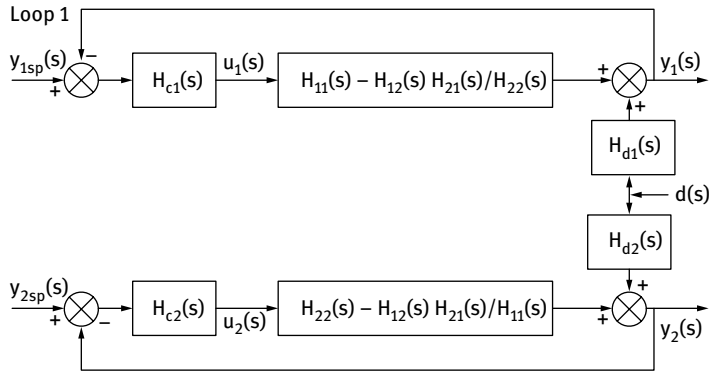
The multiloop control system with complete decoupling has the following closed-loop transfer function dependence on the setpoints and disturbance:

$$y_1(s) = \frac{\left(H_{11} - \frac{H_{12}H_{21}}{H_{22}}\right)H_{c1}}{1 + \left(H_{11} - \frac{H_{12}H_{21}}{H_{22}}\right)H_{c1}} y_{1sp}(s) + H_{d1}(s)d(s) \tag{5.26}$$

$$y_2(s) = \frac{\left(H_{22} - \frac{H_{12}H_{21}}{H_{11}}\right)H_{c2}}{1 + \left(H_{22} - \frac{H_{12}H_{21}}{H_{11}}\right)H_{c2}} y_{2sp}(s) + H_{d2}(s)d(s) \tag{5.27}$$

The block diagram representation of the decoupled closed loops is presented in Fig. 5.7.





**Fig. 5.7:** Equivalent block diagram showing the complete decoupling of the multiloop control system for the  $2 \times 2$  system.

Although the dynamics of the process with decouplers is different, compared to the original process, tuning of the decoupled control loops may be performed in an independent way and both control performance and stability may be straightforwardly achieved.

The complete decoupling is very sensitive to the accuracy of describing the process and decouplers transfer functions. Due to this reason complete decoupling is almost impossible for the processes featuring strong nonlinearity and without considering an adaptive strategy to cope with changes in the linearized model or in the nonstationary parameters of the process. This situation is especially cumbersome for the processes where the pairing was made according to large relative gain values ( $\lambda_{ij} > 1$ ).

The sensitivity of the decoupling to the accuracy of the model description is very much reduced when only part of the decouplers are implemented in the multiloop control structure. Partial decoupling, also denoted as one-way decoupling, may bring important benefits especially if it is applied for the most important controlled variables.

Even though complete dynamic decoupling is desired from the theoretical perspective of effectiveness, the steady-state decoupling based on the static gains of the decouplers transfer functions may bring benefits compared to the lack of any decoupling.

### **i** Example 5.4

The computation of the decouplers for the multiloop control of the mixing process described in Example 5.1 results in the following form of the decouplers:

$$D_{12}(s) = -\frac{H_{12}(s)}{H_{11}(s)} = -\frac{1}{1} = -1 \quad (5.28)$$

$$D_{21}(s) = -\frac{H_{21}(s)}{H_{22}(s)} = -\frac{1.567}{-0.783} = -2 \quad (5.29)$$

It may be noticed that, in this particular case, the dynamic and static decouplers have the same form. This is because both manipulated variables have equal dynamic effects on each of the controlled outputs (in fact, the first controlled variable is instantaneously affected by both manipulated variables and its dynamics is lacking).

The form of the equivalent transfer functions for the decoupled control loops are

$$y_1(s) = \frac{3H_{c1}}{1 + 3H_{c1}} y_{1sp}(s) \quad (5.30)$$

$$y_2(s) = -\frac{2.35 H_{c2}}{10s + 1 - 2.35 H_{c2}} y_{2sp}(s) \quad (5.31)$$

### 5.3 Multivariable centralized control

Multivariable centralized control aims to develop the control law for the action of all manipulated variables of a MIMO system on the basis of the inputs and outputs of the process and considering all process interactions. The emerged control law considers the system as a whole and not divided in decentralized control loops. The way the multivariable controller outputs are computed takes into account, at each instance of time, the effect of all interacting effects of the manipulated variables on all controlled outputs. The multivariable controller is described by a transfer matrix  $\mathbf{H}_c(s)$  having a form that, in the general case, it is not diagonal.

The most fundamental concept of *perfect control* for a MIMO process, i.e. the design of a controller that should be able to keep the process outputs equal to the desired input values (setpoints) at any moment of time (featuring no time lag or time delay), would imply a controller whose transfer matrix is equal to the inverse of the process transfer matrix. But such a MIMO controller should also operate for removing the effect of the measured and unmeasured disturbances. Due to the action of unmeasured disturbances and due to the process transfer matrix lack of accuracy of representing the real process, the controller should have to rely on feedback information on the outputs deviation from their desired setpoints. Based on feedback, the controller should generate the manipulated variables with the aim of keeping outputs to the desired setpoints. Unfortunately, although it is mathematically possible to design such a perfect controller, its transfer matrix form would not be physically realizable due to the fundamental lack of conforming to the cause-effect principle or equivalently, due to the need of a pure anticipative behavior of the controller. However, it is still possible to design a multivariable controller to have only the physically realizable part of the process inverse

transfer matrix (e.g. at least a controller equal to the inverse of the process gain matrix), but the control performance will be degraded accordingly.

The design of such a multivariable controller does not yet have a general, very well defined and practical methodology (although for the SISO case the internal model control has developed a design framework). Nevertheless, there are some methodologies that design such a multivariable centralized controller having an explicit or an implicit form of the emerged control law.

The most renowned and validated by practice implementations of centralized multivariable control is the model predictive control (MPC)-based design [15]. MPC is founded on the computation of its control algorithm using different forms of the multivariable process model that are used for predicting the future behaviour of the controlled variables. Taking into consideration the model-based predictions, a control performance index is optimized with respect to the manipulated variables while considering the future prediction horizon. On the basis of measured information describing the process outputs, feedback is introduced by the receding horizon approach and by making corrections to the process predictions at every sampling moment. MPC has straightforward means for distributing both the control importance among the controlled variables and the control effort among the manipulated variables. They are performed by means of the weighting matrices within the optimization index. The success of the MPC applications is owing not only to its capability of coping with constraints in a systematic way but also due to its multivariable centralized approach. The MPC principle is described in Chapter 2 of the book.

## References

- [1] Marlin, T.E., *Process Control – Designing Processes and Control Systems for Dynamic Performance*, McGraw-Hill, Inc., 2000.
- [2] Stephanopoulos, G., *Chemical Process Control – An Introduction to Theory and Practice*, Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1984.
- [3] Agachi, S.P., Cristea M.V., *Basic Process Engineering Control*, Walter de Gruyter GmbH, Berlin/Boston, 2014.
- [4] Romagnoli, J.A., Palazoglu, A., *Introduction in Process Control*, CRC Press, Taylor and Francis Group, Boca Raton, FL, 2006.
- [5] Bristol, E., *On a New Measure of Interaction for Multivariable Process Control*, IEEE Transaction on Automatic Control, AC-11, 133–134, 1996.
- [6] Shinskey, F.G., *Process Control Systems*, McGraw-Hill, New York, 1988.
- [7] McAvoy, T., *Connection between Relative Gain and Control Loop Stability and Design*, AIChE Journal, 27, 4, 613–619, 1981.
- [8] McAvoy, T., *Interaction Analysis*, Instrument Society of America, Research Triangle Park, NC, 1983.
- [9] Skogestad, S., Lundstrom, P., Morari M., *Selecting the Best Distillation Control Configuration*, AIChE Journal, 36, 753–764, 1990.
- [10] Ortega, J.M., *Matrix Theory, A Second Course*, Plenum Press, New York, 1987.

- [11] Skogestad, S., Morari M., *Effect of Disturbance Directions on Closed-Loop Performance*, *IEC Research*, 26, 2029–2035, 1987.
- [12] Grosdidier, P., Morari, M., Holt, B., *Closed Loop Properties from Steady State Gain Information*, *Ind. Eng. Chem. Fund.*, 24, 221–235, 1985.
- [13] Chiu, M., Arkun, Y., *Decentralized Control Structure Selection Based On Integrity Considerations*, *IEC Res.*, 29-369–373, 1990.
- [14] Campo, P., Morari, M., *Achievable Closed-loop Properties of Systems under Decentralized Control: Conditions Involving Steady-State Gain*, *IEEE Trans. Automatic Control*, 39, 932–943, 1994.
- [15] Agachi, P.S., Nagy Z.K., Cristea, V.M., Imre-Lucaci A., *Model Based Control – Case Studies in Process Engineering*, Wiley-VCH, Weinheim, 2006.

## 6 Plantwide control

### 6.1 Introduction

Process engineering concepts of automatic control are applied in a wide range of industries such as production of chemicals, pharmaceuticals, petroleum, plastics, pulp and paper, automotive, food, textiles, glass, metals, building materials, electricity, and electric and electronic devices. The challenges of the continuously changing market demands and conditions, associated with the need to spare natural resources, and the requirements to save energy while preventing environmental pollution are asking for safe, flexible, economically efficient, and environmentally friendly solutions provided by automatic control. The control systems represent efficient tools to accomplish such demanding tasks, and as intelligent control solutions are conceived, they may also represent a cost-effective and environmentally friendly way of responding to these challenges.

The generic structure of a comprehensive control system consists in a hierarchical architecture of layers, as presented in Fig. 6.1 [1].

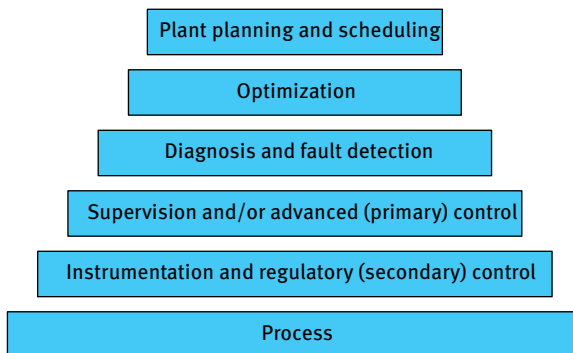


Fig. 6.1: Hierarchical structure of the whole plant control system.

The *instrumentation and regulatory control* layer performs not only the measurement and data acquisition from the plant but also the regulatory control at the most basic level of the control hierarchy. The layer interacts with the higher layers by sending and receiving data. This is the core of a distributed control system (DCS). The control systems at this level usually have single loops, with typically PID structure, and are intended to control the secondary variables of the process such as inventory, level, pressure, flow, and temperature. Its modularity allows the quick and prompt intervention for maintaining the secondary variables at their desired setpoints despite the action of disturbances.

The *supervision and/or advanced layer* computes the reference values for the regulatory layer or directly controls process variables on the basis of a multivariable model of the process. Primary controlled variables may be identified at this level, and they are subject to control, in association with or by the means of the regulatory layer. Beside fulfillment of control performance-based objectives, at this layer, economic objective functions might be considered for the computation of the reference values that are further sent to the underlying regulatory layer. Model predictive control is the most applied advanced control algorithm for this layer, although it may be used in other layers, too.

The *diagnosis and fault detection layer* aims to make the operation of the process, as much as possible, not depending on the faults and hence enhances the safe and reliable work of the subsystems belonging to all other layers. Models of the process may be also used for the identification of the causes producing malfunction of either process or instrumentation equipment, to offer early information and to manually or automatically mitigate their future negative consequences. This layer is not always present in the comprehensive control hierarchy, but its beneficial outcomes are highly appreciated.

The *optimization layer* uses a model of the process to compute the best manipulated variables or references for the supervisory/advanced lower layer. The computation is performed by optimization of an economic objective function or of the process throughput. Usually, at this layer, a steady-state model of the process is involved in the computation of the optimal values for the manipulated variables representing references for the supervisory/advanced lower layer and for the degrees of freedom remained from the lower levels. At this level, dynamic models may be also used for optimization, but their sampling time is larger than the ones associated with supervisory/advanced and regulatory layers.

The objective of the *plant planning and scheduling layer* is to generate the production planning for an extended period compared to the sampling time of the lower layers. It takes into consideration the market demands, sets the way the resources are allocated, and schedules the plan of operation. The economic objective type is dominating at this level and the decisions are mainly made for maximizing the profit. Optimization algorithms can be also used at this level to find the optimal solutions to be sent as targets to the lower layer.

## 6.2 Premises of plantwide control

Plantwide control consists in the control system aimed to operate the whole plant with the task of achieving the desired objectives, despite the action of disturbances and conforming to imposed constraints. Plantwide control may also imply the specification and placement of the measured variables, specification of the

manipulated variables, and the decomposition of the control task into smaller problems to be solved based on the overall control approach [2].

The systematic design of the plantwide control system is becoming more and more important due to the increasing complexity of the processes to be controlled and the ever-growing economic, energetic, and environmental constraints to be fulfilled. Despite the stress exerted by these driving forces, the plantwide control in a large number of industrial applications is very much based on heuristic approaches, making use of experts' knowledge. The research performed during the last two decades in this field has investigated several approaches [3]. They are considering either heuristic- or optimization-based methods and struggle for proposing an algorithmic approach for the selection of the most favorable plantwide control structure [4–9].

Traditional plantwide control practice starts from control solutions specific to different unit operations intended to make the plant operation reliable, while getting good understanding and acceptance by the operators [10]. This approach usually begins with fixing the production throughput at the plant input and is followed by the sequential design of the control system along the unit operations contributing to the successive processing stages. But this approach does not guarantee the attainment of the optimal control structure or the best operation of the plantwide control system. Nevertheless, some systematic approaches have been proposed during the last decades, and they have been successfully tested by simulation and industrial applications [3, 8].

It is obvious that the fundamental objective of any plantwide control system is to maintain the mass, energy, and momentum balance of the whole process [3]. Generally, it may be considered that momentum balance can be obtained with moderate effort, as flow and pressure measurements are available and control valves, pumps, and compressor are efficient manipulated variables. Energy balance may also be achieved by the support of the low-cost temperature measurements and by the manipulation of the cooling or heating utilities. The most difficult task is to maintain the balance of the material entering and leaving the whole system. This not only considers the overall mass balance but also the mass balance for almost every component. As concentration measurements are usually not available, due to their instruments' cost and complexity, and the independent manipulation of each chemical species flow is also problematic, the mission of preserving the components' mass balance is the most challenging task. This balance should prevent the components' accumulation or consumption in the process and becomes an important objective for the control system. Such duty of the control system is aggravated by the existence of the recycle streams that may negatively affect the process controllability, and in the extreme case, losing stability. It is also well recognized that for processes containing reduced inventory units, such as buffer tanks, the sequential plantwide design of the control system methodology based on unit operations may show poor control performance.

It is noteworthy to mention that plantwide control design may have to meet two different requirements. The first one is specific to the case of the newly designed

process and has the objective of maximizing the efficiency of the operation for the nominal throughput of the plant. This situation may also be encountered in cases of upgrading the old control system. This case is driven by the economic target of minimizing the utility, energy, raw material, and environment protection costs.

The second requirement is characterized by the objective of maximizing the throughput of the process. It may be driven by the market conditions of favorable prices for the products and asks for obtaining profit by maximizing the throughput. The maximized throughput approach implies the appearance of a bottleneck on the process production flow. The localization of the throughput manipulator may affect the structure of the control system. It may be also concluded that control system design has to be performed such as it will show a robust response to most of the possible modifications in the operating conditions.

Operators' acceptance of the plantwide control strategy is an aspect worthy to consider, as very complex control systems require trained personnel to operate in the challenging industrial environment. Changing the traditional way of operation to the plantwide control solutions implies the formation of qualified human resources by special operators training programs and the use of operating training systems.

### 6.3 Designing the plantwide control strategy

It is obvious that plantwide control design should be directly related to the process design. Traditionally, the control design was performed after the process design was finished. The result of this sequencing may affect the controllability of the process and reduce the achievable control performance. One major problem of this historical approach for the control system design may come out from the fact that process design is performed almost exclusively on the basis of the steady-state description of the process. Unfortunately, this does not reveal possible problems arising during transient operating regimes. Specifically, the process design implies the setting of the nominal steady-state operating point that might be not attainable in a simple manner during the dynamic operating mode. Additionally, the effect of the disturbances, of new imposed constraints, or of process load changes may affect the controllability of the process and the efficiency of the control system design may be degraded.

As a result of these circumstances, it becomes obvious that both process and control system design should be done by an integrated approach. This implies a joint design approach that relies on merging the process knowledge with the control theory competency and possibly associated with understanding economic indicators and their assessment methods.

The plantwide control system design, in its most favorable theoretical approach, should have the capability of bringing the process operation at the optimum operating point (or in its neighborhood). The optimal control, for a centralized multivariable approach, may emerge from running plantwide real-time optimization to obtain the best control solution and by the use of comprehensive models [11].



Meanwhile, the plantwide control system could be designed to make it simple, possibly with single loops control structure, and have the benefit of the self-regulating or self-optimizing property of the process subsystems that facilitates the accomplishment of the optimal operation in a natural way.

Although a comprehensive solution to the plantwide control design problem is not yet matured, one systematic approach for designing the control system is proposed in [4, 8]. This plantwide control design approach is considered to provide answers to the two main groups of design tasks: design of the regulatory control layer and design of the economics driven control layer.

The first task objective is to identify the suitable secondary controlled variables and their associated control variables (pairing), including control of the mass and energy inventories. The proposed design of the regulatory control layer is performed such as the remaining degrees of freedom (primary control variables) are straightforwardly used for the design of the plantwide economic optimization strategy.

The second design task objective aims to integrate in the plantwide control design the economic optimization outcomes while identifying the primary controlled variables. This part of the design is centered on economic objectives, with their associated constraints (such as the environmental or safety requirements), and is based on steady-state models. The main challenges for the economic optimization integration in the plantwide control system design are the specification of the primary controlled variables and the best localization of the throughput manipulator.

The control structure design methodology consists in two parts with seven steps [4]. The *top-down part* deals with economic optimization group of tasks and the *bottom-up part* with the regulatory group of tasks, to which are added the supervision and the dynamic optimization steps.

The *top-down part* of the control system design consists in four steps. They are [4]:

1. Defining the economic objectives and their associated constraints. Based on the steady-state models of the process an optimization economic index  $I$  is defined. This economic optimization objective consists in either maximizing the profit or maximizing the production rate (throughput). Typical disturbances are included in the steady models such as feed flows and composition, kinetic parameters or potential changes of market parameters.
2. Selecting the set of control variables (degrees of freedom) and solving the economic optimization problem. Optimization is done with respect to the control variables,  $u$ , and in the presence of typical disturbances,  $d$ . Steps 1 and 2 analyze the optimal operation of the process and are prerequisites for the selection of the most important variables to control or to maintain constant, from an economic point of view.
3. Choosing the primary controlled variables. This is the most difficult part of the top-down part and it has to specify which variables should be controlled to get an optimal (or at least close-to-optimal) economic operation. The most fundamental method of selecting the primary controlled variables is to choose them equal to the gradient of

the optimization index function, i.e.  $I_u = dl/du$ . This approach is based on the observation that by keeping  $I_u = 0$ , independent of disturbances, it makes the economic optimization to be fulfilled (as necessary condition for optimality). However, it may be difficult to develop an analytical expression for  $I$  or to have available process measurements for all variables of  $I_u$ . The selection of the primary controlled variables may be done on the basis of process insight. Two categories of variables are important to be considered as good candidates for primary controlled variables. They are the so-called *active constraints variables* and the *self-optimizing variables*.

The *active constraints variables* are characterized by the fact that the optimum is very much dependent on these variables and therefore they should be strictly maintained at desired setpoints. A small deviation of the active constraints variables from their optimal values has large effect on the deviation from optimality. Their desired values are usually maximum or minimum values of process variables such as composition, temperature, flow. Very good control performance for the control of active constraints variables should be accomplished.

The *self-optimizing variables* are characterized by the fact that when they are maintained constant they drive the process close to the optimal operation, despite the action of the disturbances. Small deviations of the self-optimizing variables from their optimal values do not have large effect on the deviation from optimality. Therefore, in most cases, a very good control performance for the control of self-optimizing variables is not necessary. To be chosen as primary controlled variables, the self-optimizing variables have to be measurable variables or at least to have the possibility of being inferred from other measured variables [7]. The self-optimizing variables must be considered in relation to remaining unconstrained degrees of freedom (of the regulatory or economic optimization layers).

Besides the physical insight approach, a set of four systematic methods have been developed for discovering appropriate self-optimizing variables [4]:

- (a) The first method assesses different sets of candidates for the controlled (self-optimizing) variables, each of them considered to be maintained at constant values. An evaluation is made for each candidate set with respect to the deviation of the optimization index function, computed with the constant values of the self-optimizing variables, from the optimal value of the performance function, in the presence of disturbances. The set presenting the most reduced deviation is selected.
- (b) Another method selects the self-optimizing variables to be equal to the gradient of the optimization index with respect to the degrees of freedom (as already mentioned).
- (c) If the previous method is not feasible due to lack of available measurements of the variables involved in  $I_u$ , a nullspace method may be employed to find optimal measurement combinations to be used in  $I_u$  [4, 12, 13].
- (d) The last method selects the self-optimizing variables on the basis of the variables presenting large scaled gain, computed as the ratio between the process gain and the span of the controlled variable [4, 14].

4. Setting the location of the throughput manipulator. Setting the production rate and the position of the throughput manipulator has direct impact on the maximization of the economic objective. At the same time, the structure of the inventory control loops and their pairing are affected by the throughput manipulator placement. Although the traditional location of the throughput manipulator is set at the inlet feed position, placing it in other position may bring benefits, such is the location at the bottleneck unit of the process.

The *bottom-up part* of the control system design consists in the following three steps [4]:

5. Selection of the *secondary controlled variables* used in the regulatory control layer, whose intended task is to stabilize the process. Pairing of the secondary controlled variables with control (manipulated) variables should be also accomplished at this step. Inventory control will be addressed, too. The main task of the regulatory layer is to reduce the deviation (by providing short transient time) of the secondary controlled variables from their associated setpoints.

A systematic procedure for the selection of the secondary controlled variables is presented in [15]. This approach involves three stages in the design, aiming to select the secondary controlled variables in such a way to provide efficient indirect control of the primary controlled variables, considering both the servo and the regulatory control performance. The design approach is based on minimizing the scaled integral absolute error (IAE) of the the primary controlled variables. IAE is used as a measure for indicating the most favorable secondary controlled variables selection intended to indirectly control the primary controlled variables. The three stages of the selection methodology are [15]

- (a) The first stage generates an initial candidate set of secondary controlled variables, identifies disturbances, and formulates a subset selection constraint to find if a control variable may be used for control of a candidate controlled variable. Additionally, the input to secondary and primary controlled output models and disturbance models are developed during this stage. The IAEs of the cascaded primary control loops are computed for all possible pairings of the triplet controlled-secondary controlled-primary controlled variables. IAEs are scaled to account for the economic importance of the primary controlled variable. The generation of the candidate set of secondary controlled variables can be done either by considering a larger set of candidate variables or by reducing the set using process insight. The selection of relevant disturbances can be done only on process insight. The subset selection constraint formulates a matrix of only logical 1 and logical 0 values that shows if the pairing of each candidate control variable with each controlled variable is acceptable from the servo and regulatory performance. The subset selection constraint is then used for pruning (of supernodes) in the branch and bound algorithm.
- (b) The second stage consists in the selection of the set of secondary controlled variables on the basis of the summed and scaled IAEs of the primary

- controlled variables, considering a cascaded structure where the secondary controlled variables are used as control variables for the control of the primary ones. Constrained minimization of the summed and scaled IAEs of the primary control loops is used for the selection of the optimal set of secondary controlled variables. The constraints of this mixed integer optimization problem take into account the interactions between control loops, at the regulatory and supervisory control layers, and the servo control performance at the primary control loop. RGAs are computed for interaction assessment for both the regulatory control layer and for the supervisory control layer [16, 17].
- (c) The third stage of the methodology consists in the performance evaluation of the control system emerged from the selection of the secondary controlled variables in case of off-design operating conditions. As the design is based on linear models and nominal operating conditions, the evaluation of the secondary controlled variables control performance has to be tested on the nonlinear model or the process. Particularly, gain switching may occur due to operation changes in the process and they should be avoided for the selected pairing. The final selection of the secondary controlled variables is made based on this last stage.
6. Design of the supervisory control layer. Economic considerations may be taken into account at this control layer, such as whether good control performance at this level can comply with economic objective targets. The design must decide which of the controlled variables of the secondary control layer have to be controlled by the supervisory layer by means of imposed setpoints. The control variables of the supervisory control layer are usually the secondary controlled variables of the regulatory layer. At this control layer, both decentralized and centralized control may be used.
  7. Design of the optimization layer. At this control layer the design is dominated by economic objectives accomplishment. It is founded on a centralized control approach and it uses steady-state models or dynamic models with large sampling time, compared to the one of the subordinated layers.

Plantwide control design may be needed both for new plants and for upgrading older ones. The latter is driven by the economic context of the market, desired specifications of the products and the constraints on the production feed.

## References

- [1] Zhu, Y., *Multivariable System Identification for Process Control*, Pergamon, Elsevier, Oxford, UK, 2001.
- [2] Dimian, A.C., Bildea, S., Kiss, A., *Integrated design and simulation of chemical processes*, Computer Aided Chemical Engineering, 35, second edition, Elsevier, Amsterdam, 2014.
- [3] Larsson, T., Skogestad, S., *Plantwide control: A review and a new design procedure. Modeling, Identification and Control*, 21, 209–240, 2000.

- [4] Downs, J.J., Skogestad, S., *An industrial and academic perspective on plantwide control*, Annual Reviews in Control, 35, 99–110, 2011.
- [5] Kookos, I.K., Perkins, J.D., *An algorithmic method for the selection of multivariable process control structures*, Journal of Process Control, 12, 85–99, 2002.
- [6] Luyben, W.L., Tyreus, B.D., Luyben, M.L., *Plantwide process control*, McGraw-Hill, New York 1998.
- [7] Skogestad, S., *Plantwide control: The search for the self-optimizing control structure*, Journal of Process Control, 10, 487–507, 2000.
- [8] Skogestad, S., *Control structure design for complete chemical plants*, Computers and Chemical Engineering, 28(1–2), 219–234, 2004.
- [9] Zheng, A., Mahajanam, R.V., Douglas, J.M., *Hierarchical procedure for plantwide control system synthesis*, AIChE Journal, 45(6), 1255–1265, 1999.
- [10] Agachi, S.P., Cristea M.V., *Basic Process Engineering Control*, Walter de Gruyter, Berlin/Boston, 2014.
- [11] Agachi, P.S., Nagy Z.K., Cristea, V.M., Imre-Lucaci A., *Model Based Control – Case Studies in Process Engineering*, Wiley-VCH Verlag, Weinheim, 2006.
- [12] Alstad, V., Skogestad, S., *Null space method for selecting optimal measurement combinations as controlled variables*, Industrial and Engineering Chemistry Research, 46(3), 846–853, 2007.
- [13] Alstad, V., Skogestad, S., Hori, E.S., *Optimal measurement combinations as controlled variables*, Journal of Process Control, 19, 138–148, 2009.
- [14] Halvorsen, I.J., Skogestad, S., Morud, J.C., Alstad, V., *Optimal selection of controlled variables*, Industrial and Engineering Chemistry Research, 42(14), 3273–3284, 2003.
- [15] Dustin Jones, D., Bhattacharyya, D., Turton, R., Zitney, S.E., *Plant-wide control system design: Secondary controlled variable selection*, Computers and Chemical Engineering, 71, 253–262, 2014.
- [16] Marlin, T.E., *Process Control – Designing Processes and Control Systems for Dynamic Performance*, McGraw-Hill, New York, 2000.
- [17] Stephanopoulos, G., *Chemical Process Control – An Introduction to Theory and Practice*, Prentice-Hall, Englewood Cliffs, NJ, 1984.

# 7 Linear discrete systems and Z transform

## 7.1 Introduction

The continuous time systems deal with systems having continuous (with respect to time) input  $\mathbf{u}(t)$ , output  $\mathbf{y}(t)$  and state  $\mathbf{x}(t)$  variables and having the mathematical representation based on functions with the time as the independent variable [1]. These independent variables have a continuous set of values for time,  $t \in \mathbf{R}$ . Unlike the continuous systems, the discrete systems (also named sampled systems) have the input  $\mathbf{u}(n \cdot T)$ , output  $\mathbf{y}(n \cdot T)$  and state  $\mathbf{x}(n \cdot T)$  variables represented by discrete functions with respect to time, functions that are only defined at particular moments of time  $t \in \mathbf{Z}$ . These time moments are usually chosen equally spaced  $t = n \cdot T$ ,  $n = 0, 1, 2, \dots$ , and consequently, are multiples of the time interval  $T$ , named the *sampling time*. A discrete time function  $f(t)$  is defined as [1]:

$$f(t) = \begin{cases} \text{exists,} & \forall t = nT, \quad n \in N \\ 0 \text{ or undefined,} & \forall t \neq nT \end{cases} \quad (7.1)$$

Although the need for describing the discrete systems appeared before the spectacular development of the digital control systems using the computer, it may be stated that the perspective of these applications represented the main motivation for the progress of the discrete system theory. The explanation is that every digital system is working on the basis of a sequential mode of operation. This sequential mode puts in a successive arrangement all operations that have to be performed [2–4]. Thus, a digital control system based on a computing system is not able to generate a control action at each moment of time. The arithmetic and logic unit of the computer has to perform several operations that may not be simultaneous, such as reading data from the process, analogue-to-digital (A/D) conversion, processing the acquired data on the basis of a computation algorithm, generation of the control variable, and digital-to-analogue conversion (D/A) [3]. The digital control system using the computer is presented in Fig. 7.1 [1].

The output of the process is a continuous signal. This continuous signal is transformed in a digital form using the A/D converter. The A/D conversion is performed at discrete moments of time  $t = n \cdot T$ ,  $n = 0, 1, 2, \dots$ , also named sampling moments. The clock of the system sets the sampling moments. The sampling is generally defined as the operation of extracting a small part, named sample, from the whole. In the context of the computer control, sampling is the transformation of a continuous time signal into a succession of numeric values representing the values of the continuous signal at the time moments  $y(n \cdot T)$ . The sampled signals are discrete signals obtained from the continuous signals by the sampling operation. The succession of numeric values obtained after the sampling operation is processed by the computing algorithm that generates a new succession of numeric values  $u(n \cdot T)$  representing the control action (commands). These commands are determined only at

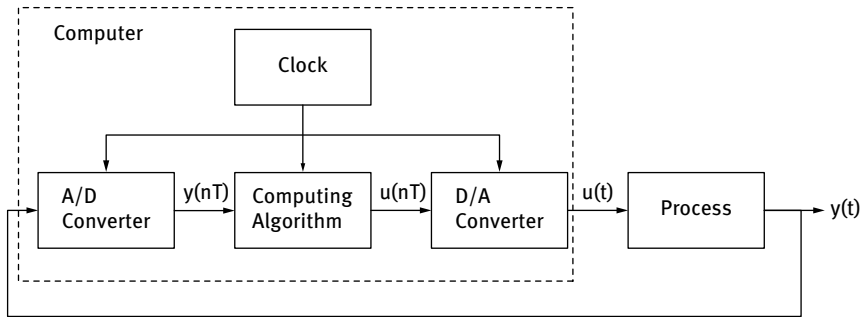


Fig. 7.1: Schematic representation of the process controlled by a computer system.

discrete moments of time and sent to the D/A converter. The D/A converter transforms the succession of the commands' values into a continuous signal  $u(t)$  that is further sent to the controlled process. Usually, this transformation into the continuous signal consists in keeping the computed value at a sample moment constant in time until the next one. All operations, A/D conversion, computation of the command, and D/A conversion, are happening at discrete and successive moments of time, marked by the clock of the computing system. Consequently, the computing system operates with discrete signals having values changing only at the sampling moments of time.

In the control system using the computer, both continuous and discrete signals are present, leading to the following consequences:

- before being “read” by the computer, it is necessary to transform the continuous signals  $y(t)$  into discrete signals  $y(n \cdot T)$ ,
- in order to have an efficient action upon the process it is necessary to transform the discrete signals produced by the computer  $u(n \cdot T)$  into continuous signals  $u(t)$ ,
- the mathematical instruments used for describing the behavior of the continuous linear systems (differential equations or transfer functions) are not convenient for describing the dynamic behavior of the control system based on the computer. Therefore, it is necessary to find specific ways for describing the discrete systems.

Beside the computer-based systems there are also other systems being implicitly discrete. The radar localization is an example of such a system that processes information on the spatial position of an object only after the electromagnetic reflected wave is again generated and processed (after a full rotation of the electromagnetic wave beam). As a result, it produces a signal with new values only at equally spaced and discrete intervals of time. An important category of analytical instruments, such as the mass spectrometers or the gas chromatographs, are implicit discrete systems because they work sequentially. Usually, these instruments require disjoint (successive) time intervals for extracting the sample, for performing the analysis, and for preparing the new analysis [1].

## 7.2 Discrete systems described by input-output relationship

### 7.2.1 Sampling the continuous signals

Consider the way that a continuous signal may be transformed in a discrete signal, operation commonly named as sampling the continuous signal [1, 4, 5, 16]. Consider a continuous signal  $y(t)$  transmitted over a transmission line that is interrupted by a switch, named sampling element (sampler), as it is schematically presented in Fig. 7.2.

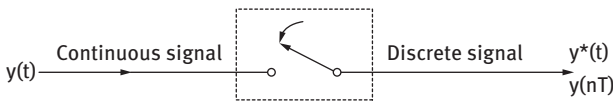


Fig. 7.2: Basic representation of the sampling element.

The sampling element (switch) is closing every  $T$  seconds and remains closed for a very short (approaching zero) time interval.

Fig. 7.3(a) presents the continuous signal  $y(t)$  and Fig. 7.3(b) the discrete signal obtained after the sampling operation, denoted with  $y^*(t)$ .

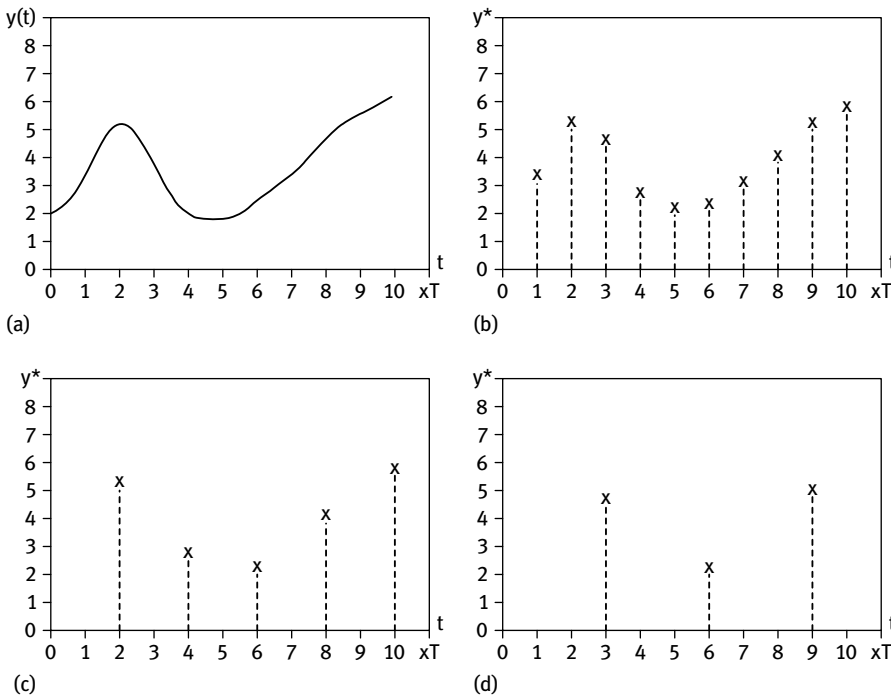


Fig. 7.3: Representation of (a) continuous signal, (b) sampled signal with the reference sampling period, (c) sampled signal with double sampling period, (d) sampled signal with triple sampling period.



The sampled signal consists of a series of (point-wise defined) values that are equal to the values of the continuous signal at each moment of time, the latter being multiple of the sampling period  $t = n \cdot T, n = 0, 1, 2, \dots$ . Fig. 7.3(c) and Fig. 7.3(d) present the signals obtained after sampling the same continuous signal with double and triple sampling period (compared with the reference sampling period presented in Fig. 7.3(b)).

On this basis, the following remarks may be formulated [1]:

- i) As the sampling period is getting smaller and is approaching zero  $T \rightarrow 0$ , the sampled signal  $y^*$  is approaching the continuous signal  $y(t)$ , but this situation requires a very large number of sampled values.
- ii) As the sampling time is getting larger, less sampled values are produced but the sampled signal deteriorates meaning that the reconstruction of the original signal from the discrete values may become imperfect or even impossible.

Therefore, there will be analyzed the conditions for which the reconstruction of a continuous signal is possible from the sampled values and the way of choosing the sampling period. To answer this question, the following particular cases are considered.

Consider the case of the continuous first-order system having the step response presented in Fig. 7.4 [1].

Analyzing the response, it may be stated that the sampling operation has to be performed with a sampling period smaller than the time constant of the system (process)  $T < T_p$  to capture the dynamics of this signal. Practical considerations suggest the selection the sampling time between  $T = 0.1 \cdot T_p$  and  $T = 0.2 \cdot T_p$  to obtain sampled values able to reflect, in a convenient way, the output variable dynamics of the first-order system.

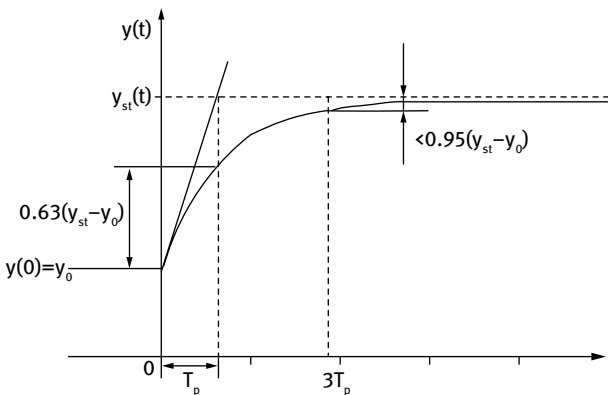


Fig. 7.4: Step response of the first-order system.

For the case of the first-order system with dead time  $\tau_m$ , having the step response presented in Fig. 7.5, the selection of the sampling time is performed as follows [1]:

- if the dead time is of the same order of magnitude with the time constant of the process  $\tau_m \approx T_p$ , the sampling time has to be about one tenth of the time constant  $T = 0.1 \cdot T_p$  or of the dead time  $T = 0.1 \cdot \tau_m$ , whichever is smaller,
- if the dead time is much larger than the time constant of the system (process)  $\tau_m \gg T_p$ , the sampling time is chosen about one tenth of the dead time value,  $T = 0.1 \cdot \tau_m$ ,
- if the dead time is much smaller than the time constant of the system  $\tau_m \ll T_p$ , the sampling time is chosen about one tenth of the time constant of the system  $T = 0.1 \cdot T_p$ .

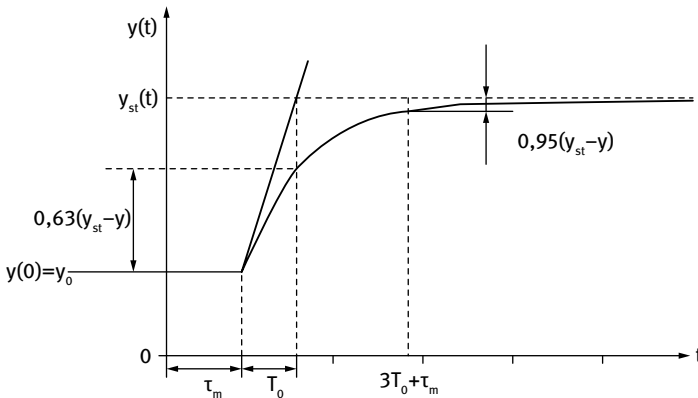
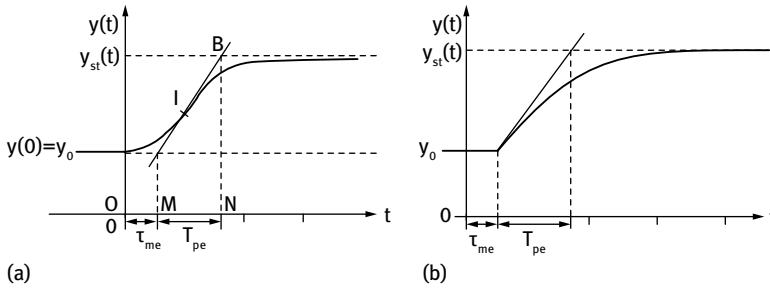


Fig. 7.5: Step response of the first-order system with dead time.

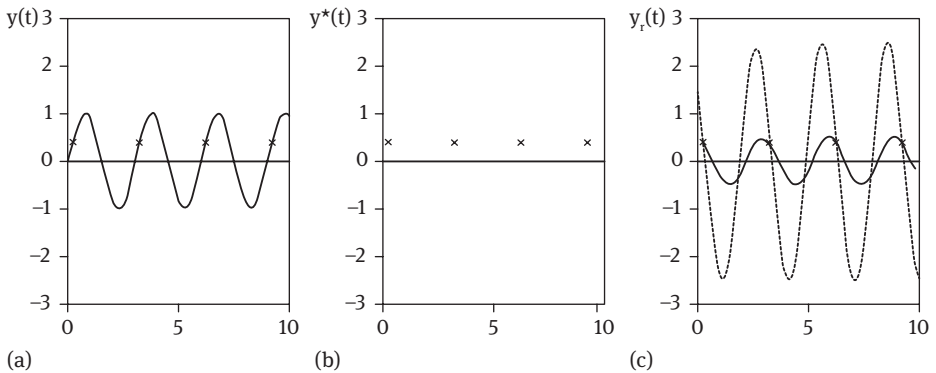
For the case of the overdamped second-order system the selection of the sampling period may be reduced to the case of the equivalent first-order system having an equivalent time constant  $T_{pe}$  and an equivalent dead time  $\tau_{me}$ , as presented in Fig. 7.6 [1, 6, 15].

The equivalence is accomplished graphically by drawing the tangent line, in the inflexion point  $I$ , to the step response plot of the second-order system and obtaining the segments  $OM = \tau_{me}$  and  $MN = T_{pe}$ . These segments are determined by the projections  $M$  and  $N$ , on the abscissa axis, of the tangent line intersection points  $A$  and  $B$  with the asymptotes of the initial and final steady state values. Figure 7.6(b) presents the step response of the first-order system with dead time that is equivalent to the step response of the second-order system presented in Fig. 7.6(a).



**Fig. 7.6:** (a) Step response of the second-order system. (b) Equivalent (first order with dead time) step response of the second-order system.

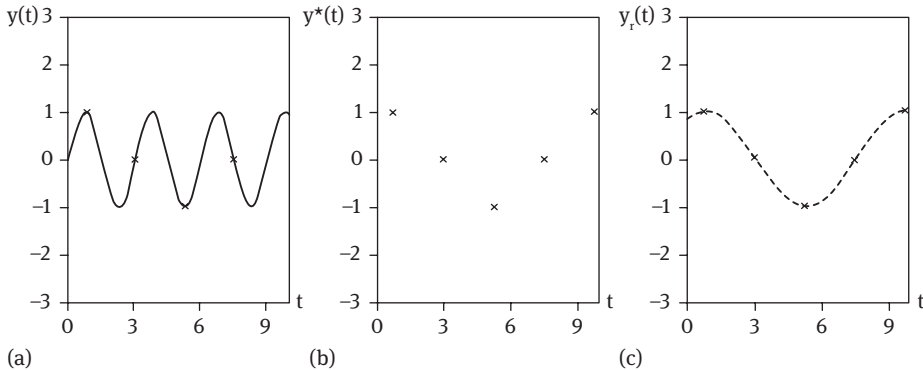
For the case of the underdamped second-order system or for the case of the linear systems subject to the sinusoidal shaped input signal the sampling period is chosen less than the half of the period corresponding to the sine signal. The plots presented in Fig. 7.7 suggest this choice [1, 4].



**Fig. 7.7:** (a) Sinusoidal signal subject to the sampling operation with the sampling period  $T$  equal to the period of the signal, (b) Values of the sampled signal  $y^*(t)$ , (c) Reconstructed signals from the sampled values.

From Fig. 7.7, it may be noticed that for the case of the sampling operation performed with a sampling period equal to the period of the sinusoidal signal  $T = 2\pi/\omega$  the sampled signal  $y^*(t)$  from Fig. 7.7(b) is obtained. This signal has values that do not reflect the shape of the continuous sinusoidal signal  $y(t)$  that was subject to the sampling operation. If the reconstruction of the continuous signal is desired from the values of the sampled signal, different reconstructed signals  $y_r(t)$  may be obtained, as presented in Fig. 7.7(c). Obviously, they are different than the original signal subjected to the sampling operation. It is concluded that the sampling period has to be chosen less than or equal to the period of the sinusoidal signal.

Figure 7.8 presents the case of the sampling with a period less than (with a  $\frac{3}{4}$  factor) the sinusoidal signal period  $T = 0.75(2\pi/\omega)$ . For this case again, the reconstructed signal  $y_r(t)$  from the values of the sampled is not satisfying the desired quality, as presented in Fig. 7.8(c).



**Fig. 7.8:** (a) Sinusoidal signal sampled with the sampling period equal to three-quarters of the period of the continuous signal. (b) Values of the sampled signal  $y^*(t)$ . (c) Reconstructed signal from the sampled values.

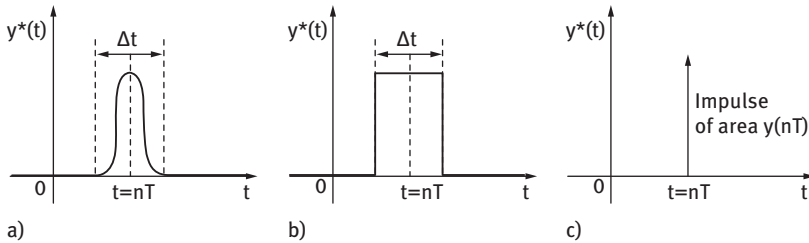
The theorem of Shannon offers an exact quantitative measure for choosing the sampling period. It is based on the fact that a periodic signal may be represented as an infinite sum of periodic sinusoidal component functions (Fourier series) [1, 4].

**Theorem of Shannon:** The sampling period of a continuous signal has to be chosen less than or equal to the half of the period corresponding to the highest frequency  $\omega_{max}$  contained in the continuous signal, i.e.  $T \leq \pi/\omega_{max}$ .

In the following, a way for describing quantitatively the sampling operation will be investigated using a mathematically coherent model [1]. To accomplish this task the starting point is the sampling element represented as a switch situated in closed position for a very short time interval  $\Delta t$ , around the sampling moments  $t = n \cdot T$ ,  $n = 0, 1, 2, \dots$ . The signal obtained as a result of this *physical sampling operation*, for a generic sampling moment  $t = n \cdot T$ , is an impulse of the form presented in Fig. 7.9(a). If the real impulse signal is considered as having ideal (sharp) flanks, the form of the impulse in the Fig. 7.9(b) is obtained.

To develop a compact mathematical description, it is assumed that the ideal sampling element acts instantly, being in the closed position only for a very short time interval  $\Delta t \rightarrow 0$ . To keep the same “power” (area) of the impulse, when  $\Delta t \rightarrow 0$ , it is necessary that the height of the impulse to become infinite. This impulse having a duration approaching zero and a height approaching infinity has an *area equal to the amplitude of the continuous signal* at the corresponding sampling moment. But this impulse having a duration approaching zero, an infinite amplitude,

and a finite area of the impulse is the Dirac function. The representation of this Dirac impulse, for the generic time moment  $t = n \cdot T$ , is presented in Fig. 7.9(c) [1, 4].



**Fig. 7.9:** (a) Real impulse signal (b) Real impulse signal having ideal flanks (c) Ideal impulse signal.

The mathematical form of the impulse presented in Fig. 7.9(c) is:

$$y^*(n \cdot T) = y(n \cdot T) \cdot \delta(t - n \cdot T), \quad (7.2)$$

where the notation  $\delta(t - n \cdot T)$  has been used for the shifted Dirac impulse, at the moment  $t = n \cdot T$ .

Consequently, the *ideal sampling element* provides a series of Dirac impulses at the sampling moments  $t = n \cdot T$ ,  $n = 0, 1, 2, \dots$ , with each Dirac impulse having the area equal to the value of the continuous signal at the corresponding moment of time. Therefore, the sampled signal  $y^*(t)$  may be mathematically represented as a sum of Dirac impulses weighted by the value of the continuous signal from the sampling moments:

$$\begin{aligned} y^*(T) &= y^*(0) + y^*(T) + y^*(2T) + \dots = \\ &= y(0) \delta(t) + y(T) \delta(t - T) + y(2T) \delta(t - 2T) + \dots \end{aligned} \quad (7.3)$$

or

$$y^*(t) = \sum_{n=0}^{\infty} y(nT) \delta(t - nT) \quad (7.4)$$

This ideal way of representing the sampling signal obtained from a continuous signal is based on the assumption that at each sampling moment the area (“power”) of the impulse is equal to the amplitude of the continuous signal. The notation “\*”, in superscript, is usually used for denoting the sampled signal. Between the sampling moments, the value (“power”) of the sampled signal is equal to zero.

For the series of discrete values, the discrete signal (sampled signal) preserves the character of *function with respect to time* (with continuous argument). Following this

fact, the Laplace transform may be applied for both members of eq. (7.4). The Laplace transform of the sampled signal  $Y^*(s)$  may be described by [1, 4]:

$$\begin{aligned} Y^*(s) &= L(y^*(t)) = L\left(\sum_{n=0}^{\infty} y(nT) \delta(t - nT)\right) \\ &= \sum_{n=0}^{\infty} y(nT) L(\delta(t - nT)) = \sum_{n=0}^{\infty} y(nT) e^{-nTs} L(\delta(t)) \quad (7.5) \\ &= \sum_{n=0}^{\infty} y(nT) e^{-nTs} \end{aligned}$$

### 7.2.2 Reconstruction of the continuous signals from their discrete values

The reconstruction of the continuous signals from the discrete signals is important because the large majority of the real processes are continuous and their input variables are also continuous [1]. For the case of the process controlled by the computer, the commands generated by the computer are provided periodically and have discrete-time values. They are represented by a succession of nonzero values just for certain moments of time, marked by the clock of the computer system (between the sampling moments the values are equal to zero or, more generally, indefinite). This succession of impulses is not able to act efficiently in the process because such command would make the control valve to open for a very short time interval and then to close until the next sampling moment, in a periodic operating mode. It is therefore necessary to reconstruct a continuous signal from the discrete time values [4, 5].

Consider a discrete time signal produced by a discrete system (e.g. by a computer) consisting in a succession of values represented by the series of impulses:

$$u^*(0) = u(0) \delta(t), u^*(T) = u(T) \delta(t - T), u^*(2T) = u(2T) \delta(t - 2T), \dots \quad (7.6)$$

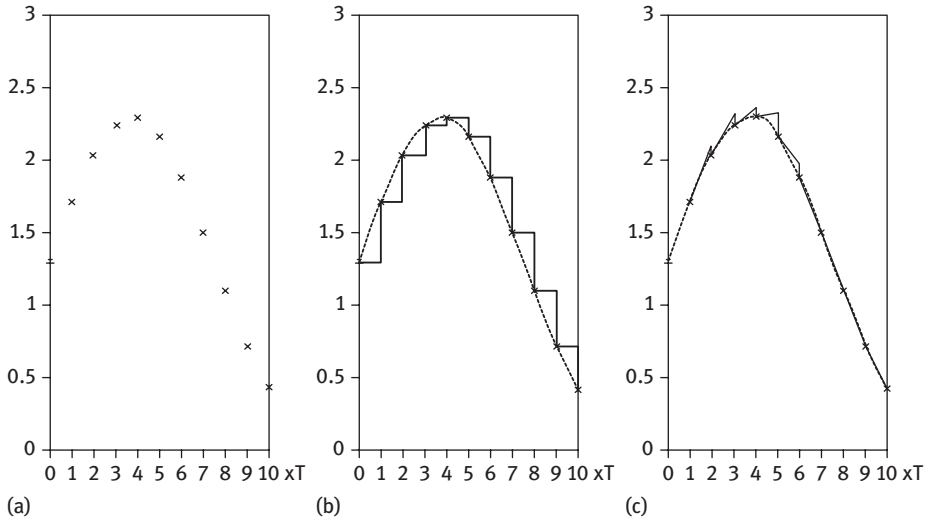
This discrete signal is represented in Fig. 7.10(a) [1]:

The simplest way for reconstructing the continuous signal is to maintain constant the discrete value from a sampling moment up to the next sampling moment:

$$u(t) = u(nT) \text{ for } nT \leq t < (n+1)T, n = 0, 1, 2, \dots \quad (7.7)$$

This way of reconstructing the continuous signal, presented in Fig. 7.10(b), may be assigned to a system that transforms the discrete values in continuous variables according to eq. (7.7). This system is denoted as the *Zero Order Hold* (ZOH) element.

There are also other ways for reconstructing the continuous signal from the discrete values. If two successive discrete values are considered, for example,  $u[(n-1)T]$



**Fig. 7.10:** (a) Discrete signal. (b) Reconstructed signal using the zero-order hold ZOH element. (c) Reconstructed signal using the first-order hold FOH element.

and  $u(nT)$ , a linear extrapolation based on these two values may be considered to determine the values of the signal over the next sampling interval. The following continuous signal is obtained:

$$u(t) = u(nT) + \frac{u(nT) - u[(n-1)T]}{T}(t - nT),$$

for  $nT \leq t < (n+1)T, n = 1, 2, \dots$  (7.8)

This way of reconstructing the continuous signal, presented in Fig. 7.10(c), may be assigned to a system that transforms the discrete values in continuous variables according to eq. (7.8). This system is denoted as the *First Order Hold* (FOH) element.

The mathematical basis for determining the equation for the hold element, irrespective of its order, is the Taylor series development of the function  $u(t)$  [1, 2, 4]. The development is performed around the discrete value  $u(nT)$ :

$$u(t) = u(nT) + \left(\frac{du}{dt}\right)_{t=nT}(t - nT) + \frac{1}{2!}\left(\frac{d^2u}{dt^2}\right)_{t=nT}(t - nT)^2 + \dots +$$
 (7.9)

If only the first term from the Taylor series is taken into consideration, the zero-order hold element is obtained. If two terms of the Taylor series are considered, the first-order hold element is obtained.

The first-order hold element needs at least two values for beginning the reconstruction of the continuous signal, but the zero-order hold element needs just one

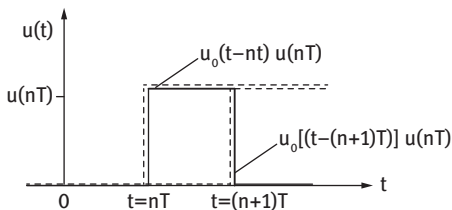
value for the reconstruction. Second-, third-, and higher-order hold elements may be developed for reconstructing the continuous signal, but they need three, four, or more values for beginning it. As the order of the hold element is increasing, the computation effort is growing, bringing poor improvement in the precision of the continuous signal reconstruction. If the sampling period is small, the differences between the hold elements are relatively insignificant. These are the reasons for using, with very good results, the zero-order hold element in the practical applications. Increasing the precision of the reconstructed signal is merely performed by decreasing the sampling period than by increasing the order of the hold element.

In general, it may be considered that when the sampling period is too large, the reconstruction of the continuous signal from its discrete values with a first-order element is better for the case of slow varying signals. The reconstruction of the continuous signal from its discrete values is unsatisfactory both with a first- or zero-order hold element for rapid varying signals. For the case of the rapid varying signals, the reconstruction of the continuous signal using the first-order element may lead to peaks and may determine excessive large values of the command.

The output of the zero-order hold element may be considered as a succession of rectangular impulses having the duration equal to the sampling period. Each of these impulses may be considered as being formed from the difference of two-step signals shifted with a sampling period of time.

$$u(t) = u(nT)[u_0(t - nT) - u_0(t - (n + 1)T)]. \quad (7.10)$$

where  $u_0(t)$  is the unit step signal. This impulse signal, for the generic time moment  $t = nT$ , obtained from the difference of the two step signals is presented in Fig. 7.11 [1, 2, 4]:



**Fig. 7.11:** Impulse signal obtained from the difference of two step signals.

Applying the Laplace transform for both members of eq. (7.10) leads to the transfer function of the zero-order hold element  $H_0(s)$  [1, 2, 4]:

$$U(s) = u(nT) \left[ e^{-snT} \frac{1}{s} - e^{-s(n+1)T} \frac{1}{s} \right]$$



$$\begin{aligned}
 U(s) &= u(nT) e^{-snT} \left[ \frac{1 - e^{-sT}}{s} \right] \\
 U(s) &= L(u(nT) \delta(t - nT)) \frac{1 - e^{-sT}}{s} \\
 \frac{U(s)}{L(u(nT) \delta(t - nT))} &= \frac{1 - e^{-sT}}{s} = H_0(s)
 \end{aligned} \tag{7.11}$$

In the above equations, it has been considered that the input of the zero-order hold element is the impulse  $u(nT) \cdot \delta(t - nT)$ , having the Laplace transform  $u(nT) \cdot e^{-nTs}$ , and the output is the signal  $u(t)$  with the Laplace transform  $U(s)$ .

### 7.2.3 Analytical description of the discrete systems

The discrete system has the input and output variables represented by sequences of numeric values for the input  $u(n \cdot T)$  and the output  $y(n \cdot T)$ , respectively,  $n = 0, 1, 2, \dots$ . The discrete system is characterized by the relationship between these sequences of numeric values. The mathematical form of this relationship is represented by the *difference equations* also named recurrent equations. For a discrete system, the difference equation stating the input and output relationships has the general form [1]:

$$F(y(n), y(n+1), \dots, y(n+N), u(n), u(n+1), \dots, u(n+M), n) = 0, \tag{7.12}$$

where  $F: \mathbf{C}^{N+M+2} \times T \rightarrow \mathbf{C}$  is, in general, a time-dependent nonlinear function. The relationship (7.12) is true for every integer  $n \in T = \{n_0, n_0 + 1, n_0 + 2, \dots\}$  or  $n \in \mathbf{Z}$ .  $M$  and  $N$  are integer and positive constants.

In the following, the *linear and discrete time-invariant* systems will be investigated. They are described by [1]:

$$\begin{aligned}
 q_N y(n+N) + q_{N-1} y(n+N-1) + \dots + q_1 y(n+1) + q_0 y(n) &= \\
 = p_M u(n+M) + p_{M-1} u(n+M-1) + \dots + p_1 u(n+1) + p_0 u(n)
 \end{aligned} \tag{7.13}$$

where  $q_i, i = 0, 1, \dots, N$  and  $p_j, j = 0, 1, \dots, M$  are real or complex constant coefficients. For non-anticipating discrete systems, the condition  $N > M$  is fulfilled.

#### **i** Example 7.1

An example of a discrete system is a digital computer system that receives as input the sequence of values of a physical variable  $u = \{u(0), u(1), u(2), \dots\}$  and transforms it

into the sequence of numeric values  $y = \{y(0), y(1), y(2), \dots\}$ , according to the recurrent algorithm [1, 4]:

$$y(n+1) = \alpha y(n) + (1-\alpha) u(n+1), \quad (7.14)$$

where  $n = 0, 1, 2, \dots$  and  $\alpha$  is a positive constant,  $0 < \alpha < 1$ .

This digital system is named *exponential smoothing filter*. It may be noticed that the output  $y(n+1)$  at any moment of time  $(n+1)T$  is determined as a weighted mean of the current input  $u(n+1)$  and of the output  $y(n)$  from the previous moment of time  $nT$ . As the constant  $\alpha$  is approaching the value of 1, the output from the previous moment is given higher importance and the output from the current moment of time is smoother.

If the equation describing the discrete system is repeatedly applied, the following equation is obtained:

$$y(n) = \alpha^n y(0) + (1-\alpha) \sum_{k=0}^{n-1} \alpha^k u(n-k). \quad (7.15)$$

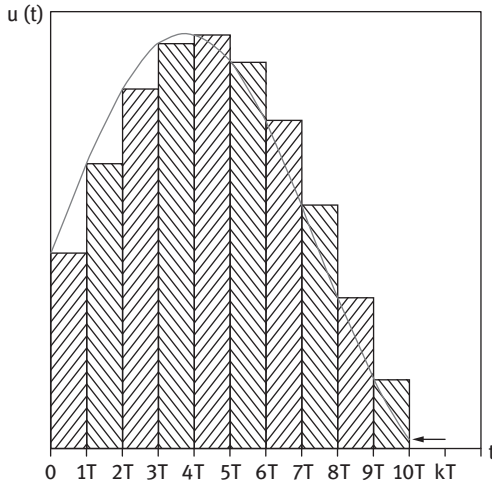
This equation presents the fact that the output  $y(n)$  is equal with the weighted sum of all inputs from the previous moments of time, starting with  $u(1)$  and of the initial output  $y(0)$ . For  $|\alpha| < 1$ , the effect on the output of the initial value  $y(0)$  is decreasing asymptotically as  $n$  is increasing. The weighting coefficients  $\alpha^k$  are exponentially decreasing as they are weighting values situated further in the past.

The question is how can the continuous system description (model) be transformed into a discrete description (model)? It is known that algebraic, differential, or integral equations can represent continuous systems. Operations equivalent to the basic operations met in the description of the continuous systems may be also developed [1, 2, 14].

1. The *algebraic* relations of the continuous models, between values of the input and output at the same moment of time, are transformed into relationships of the same form for the discrete models. For example, the simple relationship  $y(t) = k \cdot u(t)$  of a continuous time gain element model is transformed into the discrete equivalent relationship  $y(nT) = k \cdot u(nT)$ .
2. The *integration* operation, from the continuous models, is transformed into a *sum* for the discrete models. The integral  $\int u(t) dt$  has the significance of the area under the plot of the function  $u(t)$  and may be approximated with a sum of areas corresponding to the rectangles obtained by sampling the time. These rectangles have one side equal to the sampling period and the other side equal to the value of the function at the corresponding sampling moment.

$$\int_0^t u(t) dt \cong T \sum_{k=0}^n u(kT), \quad (7.16)$$

These rectangles are represented in Fig. 7.12 [1].



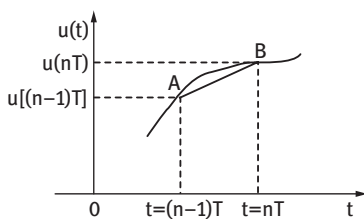
**Fig. 7.12:** Approximation of the integral from the continuous models with the sum for the discrete models.

The way of transforming the integral of a function from the continuous case into its discrete correspondent is not unique, depending on the different numerical methods for approximating the integral: rectangle, trapezoid, etc.

3. The *derivative*  $du/dt$  from the continuous models is transformed into the difference equations. The simplest numerical method for approximating the derivative is given by

$$\frac{du}{dt} \cong \frac{u(nT) - u[(n-1)T]}{T}. \tag{7.17}$$

The slope of the straight line  $AB$  from Fig. 7.13 is the approximation of the derivative [1].



**Fig. 7.13:** Approximation of the derivative from the continuous models with the difference equation for the discrete models.

The transformation of the derivative into its discrete model correspondent is not unique, depending on the different numeric methods for approximating the

derivative: Euler, Runge-Kutta, etc. There are methods for the numeric approximation of the derivative based on the right (forward) difference  $dy/dt = (1/T)(y[(n + 1)T] - y(nT))$  or on the left (backward) difference  $dy/dt = (1/T)(y(nT) - y[(n - 1)T])$ . High order derivatives are approximated by high order difference equations. Usually, for terms of the form  $y(nT)$  the notation  $y_n$  is used.

For obtaining the discrete time model, starting from the continuous time model, the following steps have to be performed [1, 2, 14]:

1. Start with the continuous model described by algebraic, differential and integral equations,
2. Approximate the derivatives of any order by finite differences,
3. Approximate the integrals by a method of numeric integration,
4. Values of the algebraic terms are directly introduced with their values from the corresponding sampling moments.

### Example 7.2

The discrete model of the continuous first-order system is determined as follows [1, 2, 14]. The nonlinear first order continuous system is described by

**i**

$$\frac{dy}{dt} = f(u, y), \quad (7.18)$$

where  $f$  is a nonlinear function.

Approximating the derivative by the first-order forward difference  $dy/dt = (1/T) \cdot (y_{n+1} - y_n)$ , the following difference equation is obtained:

$$y_{n+1} = y_n + Tf(u_n, y_n), \quad (7.19)$$

showing the way that output at the next moment of time  $y_{n+1}$  depends on the values of the input  $u_n$  and output  $y_n$  from the current moment.

For the case of the linear system, as it is the process of liquid accumulation in a tank, the continuous linear system is described by the differential equation:

$$T_p \frac{dy}{dt} + y = ku, \quad (7.20)$$

and the discrete time form of the model, based on the same first-order forward difference, becomes

$$y_{n+1} = \left(1 - \frac{T}{T_p}\right) y_n + \frac{kT}{T_p} u_n. \quad (7.21)$$

**i** Example 7.3

The discrete time model of the linear second-order system, consisting in the process of liquid accumulation in a series of two tanks, will be determined [1, 2, 14]. The second-order system is described by the second-order differential equation:

$$T_1 T_2 \frac{d^2 y}{dt^2} + (T_1 + T_2) \frac{dy}{dt} + y = ku. \quad (7.22)$$

The second-order derivative is approximated by the second-order forward difference:

$$\begin{aligned} \frac{d^2 y}{dt^2} &= \frac{d}{dt} \left( \frac{dy}{dt} \right) \approx \frac{d}{dt} \left[ \frac{y_{n+1} - y_n}{T} \right] \approx \frac{1}{T} \left[ \frac{y_{n+2} - y_{n+1}}{T} - \frac{y_{n+1} - y_n}{T} \right] \\ &= \frac{1}{T^2} (y_{n+2} - 2y_{n+1} + y_n) \end{aligned} \quad (7.23)$$

Replacing the first- and second-order derivatives with their approximations, the following discrete time form is obtained:

$$\begin{aligned} T_1 T_2 \left[ \frac{1}{T^2} (y_{n+2} - 2y_{n+1} + y_n) \right] + (T_1 + T_2) \frac{1}{T} (y_{n+1} - y_n) + y_n &= ku_n \\ y_{n+2} &= \frac{1}{T_1 T_2} \{ [2T_1 T_2 - (T_1 + T_2)T] y_{n+1} + [(T_1 + T_2)T - T_1 T_2 - T^2] y_n + kT^2 u_n \} \end{aligned} \quad (7.24)$$

Determination of the derivatives on the basis of numeric approximations, for the case of continuous physical (real) signal, may bring practical difficulties due to the noise superposed on the pure signal. The process noise has low amplitude but high frequency (rapid change in the time unit) and may generate components with unacceptable high amplitude when the derivation operation is used. To overcome this inconvenience, digital filters (Example 7.1) are used to smooth the measured signal [1, 2, 14].

The transformation of the continuous model presenting dead time into its discrete correspondent may be performed directly if the dead time is approximated with a multiple (integer) of the sampling period  $\tau_m = k \cdot T$ . Consider a first-order system with dead time described by the differential equation:

$$T_p \frac{dy}{dt} + y = k_p u(t - \tau_m). \quad (7.25)$$

Its equivalent discrete form is

$$y_{n+1} = \left( 1 - \frac{T}{T_p} \right) y_n + \frac{k_p T}{T_p} u_{n-k}. \quad (7.26)$$

### 7.2.4 Z transform

According to the concepts presented in the previous paragraphs, it may be concluded that difference equations represent a simple and natural way for the mathematical description of the discrete time systems behavior. This is similar to the

description of continuous time systems by the differential equations [1, 4, 12]. The *Z transform* offers an elegant and convenient method for the analysis and synthesis of the linear discrete systems in the same way that for the continuous systems is used the Laplace transform [6]. The incentives brought by the use of the *Z transform* consist in the simplicity of the representation and directly offering an instrument for qualitative and quantitative analysis of linear discrete systems [7].

The following row of values is obtained by the sampling of a continuous signal  $y(t)$ , using the sampling time  $T$ :

$$y(0), y(T), y(2T), y(3T), \dots \quad (7.27)$$

**Definition:** The *Z transform* of the row of sampled values is defined by the sum (series):

$$Z(y(0), y(T), y(2T), \dots) = \sum_{n=0}^{\infty} y(nT)z^{-n} \quad (7.28)$$

The usual appellation for *Z transform* of the row of sampled values is of *Z transform of the function*  $y(t)$ . It is denoted by  $Y(z)$  and its notation is:

$$Z(y(t)) = Y(z) = \sum_{n=0}^{\infty} y(nT)z^{-n} \quad (7.29)$$

The *Z transform* exists only for those values of the independent variable  $z$  for which the infinite series, given by eq. (7.28) or (7.29), has finite values (the series is convergent) [1, 4, 12].

This way of defining the *Z transform* makes the *Z transform* of a sequence of numeric values originating from different continuous functions, but having the same values at the sampling moments, to have the same *Z transform*.

The *Z transform* associates a signal (function) from the so-called *Z domain* (having the  $z$  independent variable) to a discrete time signal (function).

The *Z transform* of a time function is depending on the value chosen for the sampling time  $T$ .

According to the concepts presented when the sampling of the continuous signal has been investigated, for a sequence of sampled values it may be associated a time function  $y^*(t)$  described by a sum of shifted Dirac impulses multiplied by the values of the sampled function corresponding to the sampling moments. The integral (area) of each of the terms from the sum is equal to the value of the function at the corresponding sampling moment.

$$y^*(t) = \sum_{n=0}^{\infty} y(nT) \delta(t - nT). \quad (7.30)$$

Applying the Laplace transform for this continuous time function leads to

$$Y^*(s) = \sum_{n=0}^{\infty} y(nT) e^{-nTs}. \quad (7.31)$$

If the notation  $z = e^{Ts}$  is used, the Z transform of the sequence (7.27) is

$$Y^*(s) \Big|_{z=e^{Ts}} = \sum_{n=0}^{\infty} y(nT)z^{-n} = Y(z). \quad (7.32)$$

It may be concluded that the Z transform of the row of sampled values is a particular form of the Laplace transform. The variables of the two transforms,  $s$  and  $z$ , are linked by the relationship  $z = e^{Ts}$ .

### 7.2.5 Z transform of several simple functions

The Z transform of the following simple functions is important for analyzing the behavior and the synthesis of discrete systems [1, 4, 12]. They are presented in Tab. 7.1.

**Tab. 7.1:** Z transform of usual functions [4].

Time signal	Laplace transform	Z transform
$\delta(t)$	1	1
$u_0(t)$	$\frac{1}{s}$	$\frac{1}{1-z^{-1}}$
$a \cdot t$	$\frac{a}{s^2}$	$\frac{a \cdot T \cdot z^{-1}}{(1-z^{-1})^2}$
$e^{-at}$	$\frac{1}{s+a}$	$\frac{1}{1-e^{-aT} \cdot z^{-1}}$
$t \cdot e^{-at}$	$\frac{1}{(s+a)^2}$	$\frac{T \cdot e^{-aT} \cdot z^{-1}}{(1-e^{-aT} \cdot z^{-1})^2}$
$\sin(\omega t)$	$\frac{\omega}{(s^2 + \omega^2)}$	$\frac{z^{-1} \cdot \sin \omega T}{1 - 2 \cdot z^{-1} \cos \omega T + z^{-2}}$
$\cos(\omega t)$	$\frac{s}{(s^2 + \omega^2)}$	$\frac{1 - z^{-1} \cdot \cos \omega T}{1 - 2 \cdot z^{-1} \cdot \cos \omega T + z^{-2}}$
$1 - e^{-at}$	$\frac{a}{s(s+a)}$	$\frac{(1 - e^{-aT}) \cdot z^{-1}}{(1 - z^{-1})(1 - e^{-aT} \cdot z^{-1})}$
$e^{-at} \cdot \sin(\omega t)$	$\frac{\omega}{(s+a)^2 + \omega^2}$	$\frac{z^{-1} \cdot e^{-aT} \cdot \sin \omega T}{1 - 2 \cdot z^{-1} \cdot e^{-aT} \cdot \cos \omega T + e^{-2aT} \cdot z^{-2}}$
$e^{-at} \cdot \cos(\omega t)$	$\frac{s+a}{(s+a)^2 + \omega^2}$	$\frac{1 - z^{-1} \cdot e^{-aT} \cdot \cos \omega T}{1 - 2 \cdot z^{-1} \cdot e^{-aT} \cdot \cos \omega T + e^{-2aT} \cdot z^{-2}}$

It may be noticed that the Z transform of the presented functions has the form of the ratio between two polynomials with the independent variable  $z^{-1}$  or  $z$ .

### 7.2.6 Inverse of the Z transform

The determination of the inverse Z transform is similar, as methodology, with the determination of the inverse Laplace transform. In the case of the inverse Z transform of a function  $G(z)$  the values of the function  $g(t)$  at the multiples of the sampling time,  $t = n \cdot T$ ,  $n = 0, 1, 2, \dots$ , are computed [1, 4, 12]:

$$Z^{-1}(G(z)) = \{g(0), g(T), g(2T), \dots\}. \quad (7.33)$$

The original (continuous time) function  $g(t)$  may not be recovered from the image  $G(z)$  by the inverse Z transform, but only the values  $g(nT)$ ,  $n = 0, 1, 2, \dots$ . This fact is natural because it is not expected to get from the inverse Z transform more information than it was embedded by the direct Z transform. As a consequence, the sequence of values obtained by the inverse Z transform may correspond to different (continuous time) functions but which share the same values at the sampling moments. It is not possible to compute the sampling time  $T$  by the inverse Z transform.

From practical reasons, finding the inverse Z transform of a function having the form of the ratio between two polynomials in the  $z$  independent variable is desired:

$$G(z) = \frac{p_n z^n + p_{n-1} z^{n-1} + \dots + p_0}{z^n + q_{n-1} z^{n-1} + \dots + q_0} = \frac{P(z)}{Q(z)}. \quad (7.34)$$

A first way to determine the inverse Z transform of the function  $G(z)$  is the direct division of the polynomials  $P(z)$  and  $Q(z)$ . The order of the polynomial  $Q(z)$  has to be always higher than or equal to the order of the polynomial  $P(z)$ , for the case of causal systems. If the following form is obtained after the direct division:

$$G(z) = \frac{P(z)}{Q(z)} = g_0 + g_1 z^{-1} + g_2 z^{-2} + \dots, \quad (7.35)$$

and considering the definition of the Z transform it may be concluded that

$$g(0) = g_0, g(T) = g_1, g(2T) = g_2, \dots \quad (7.36)$$

This means that the row of the successive quotients obtained by the direct division of the polynomials  $P(z)$  and  $Q(z)$  represent the values of the sequence and they form the inverse Z transform of the function  $G(z)$ .

#### Example 7.4

Consider the Z transform of a sequence of values, with the following form [1]:

$$G(z) = \frac{z}{2z^2 - 3z + 1}. \quad (7.37)$$





The inverse Z transform may be determined by the division of the polynomials  $P(z) = z$  and  $Q(z) = 2z^2 - 3z + 1$ :

$$\begin{array}{r}
 z \qquad \qquad \qquad | \ 2z^2 - 3z + 1 \\
 \underline{-z + 1.5 - 0.5z^{-1}} \quad | 0.5z^{-1} + 0.75z^{-2} + 0.875z^{-3} + 0.9375z^{-4} \dots \\
 1.5 - 0.5z^{-1} \\
 \underline{-1.5 + 2.25z^{-1} - 0.75z^{-2}} \\
 1.75z^{-1} - 0.75z^{-2} \\
 \underline{-1.75z^{-1} + 2.625z^{-2} - 0.875z^{-3}} \\
 1.875z^{-2} - 0.875z^{-3} \\
 \underline{-1.875z^{-2} + 2.8125z^{-3} - 0.9375z^{-4}} \\
 1.9375z^{-3} - 0.9375z^{-4} \\
 \text{etc.}
 \end{array}$$

From this division, the form of the function  $G(z)$  is

$$G(z) = 0 + 0.5z^{-1} + 0.75z^{-2} + 0.875z^{-3} + 0.9375z^{-4} + \dots \quad (7.38)$$

The sequence of values at the sampling moments is:

$$g(0) = 0, g(T) = 0.5, g(2T) = 0.75, g(3T) = 0.875, g(4T) = 0.9375, \dots \quad (7.39)$$

The second way to determine the inverse Z transform is the partial fraction expansion of the function  $G(z)$  in simple fractions having well-known inverse Z transform (according to direct Z transform of the set of functions presented in Tab. 7.1). This procedure is similar to the one used for determining the inverse Laplace transform [1].

There are two equivalent ways for obtaining the partial fraction expansion:

- Taking into account that all Z transforms of the functions in Tab. 7.1 have the  $z$  factor at the numerator, it is simple to determine first the partial fraction expansion of the function  $G(z)/z$ . Subsequently, for the partial fraction expansion of the function  $G(z)$  it is applied the inverse Z transform (distributed for each term of the expansion).
- The function  $G(z)$  is transformed into a ratio of two polynomials  $P'(z^{-1})$  and  $Q'(z^{-1})$ , having  $z^{-1}$  independent variable:

$$G(z) = \frac{P(z)}{Q(z)} = \frac{p_n + p_{n-1}z^{-1} + p_{n-2}z^{-2} \dots + p_0z^{-n}}{1 + q_{n-1}z^{-1} + q_{n-2}z^{-2} + \dots + q_0z^{-n}} = \frac{P'(z^{-1})}{Q'(z^{-1})}. \quad (7.40)$$

The partial fraction expansion of this ratio of polynomials is performed obtaining simple fractions with well-known (from Tab. 7.1) inverse Z transforms:

$$G(z) = \frac{P'(z^{-1})}{Q'(z^{-1})} = \frac{c_1}{v_1(z^{-1})} + \frac{c_2}{v_2(z^{-1})} + \dots + \frac{c_n}{v_n(z^{-1})}. \quad (7.41)$$

Coefficients  $c_1, c_2, \dots, c_n$  may be determined according to a simple methodology. The polynomials  $v_1(z^{-1}), v_2(z^{-1}), \dots, v_n(z^{-1})$  are usually of low order and have known inverse Z transform.

Applying the inverse Z transform to the members of eq. (7.41), the function  $f(t)$  is determined as a sum of elementary time functions  $r_1(t), r_2(t), \dots, r_n(t)$ :

$$\begin{aligned} g(t) = Z^{-1}(G(z)) &= Z^{-1}\left(\frac{c_1}{v_1(z^{-1})}\right) + Z^{-1}\left(\frac{c_2}{v_2(z^{-2})}\right) + \dots + Z^{-1}\left(\frac{c_n}{v_n(z^{-n})}\right) \\ &= r_1(t) + r_2(t) + \dots + r_n(t) \end{aligned} \quad (7.42)$$

It has to be mentioned that the values of the function  $g(t)$ , from expression (7.42), are only meaningful for the sampling moments:

$$\begin{aligned} g(0) &= r_1(0) + r_2(0) + \dots + r_n(0) \\ g(T) &= r_1(T) + r_2(T) + \dots + r_n(T) \\ g(2T) &= r_1(2T) + r_2(2T) + \dots + r_n(2T) \\ &\vdots \end{aligned} \quad (7.43)$$

### Example 7.5

Consider the Z transform of a sequence of values, with the form [1]:



$$G(z) = \frac{z}{2z^2 - 3z + 1} = \frac{z^{-1}}{z^{-2} - 3z^{-1} + 2}. \quad (7.44)$$

The roots of the polynomial  $Q'(z^{-1})$ , from the denominator of the function  $G(z^{-1})$ , are equal to  $z^{-1} = 1$  and  $z^{-1} = 2$ . The partial fraction expansion of the function  $G(z^{-1})$  is:

$$G(z) = \frac{z^{-1}}{z^{-2} - 3z^{-1} + 2} = \frac{1}{1 - z^{-1}} - \frac{1}{1 - \frac{1}{2}z^{-1}}. \quad (7.45)$$

Applying the inverse Z transform, the following are obtained:

- for the term  $1/(1 - z^{-1})$ , the unit step function  $u_0(t)$  is obtained
- for the term  $-1/(1 - 0.5z^{-1})$ , the original (discrete time) function is an exponential having for the sampling moments the form  $-e^{n \cdot \ln(0.5)}$ ,  $n = 0, 1, 2, \dots$ , is has been accounted that  $e^{aT} = 0.5$  and equivalently  $aT = \ln(0.5)$ .

Consequently, the sequence of values  $g(nT)$  at the sampling moments, for the function  $g(t)$ , may be determined by

$$g(nT) = u_0(nT) - e^{n \ln(0.5)} = 1 - e^{n \ln(0.5)}, \quad n = 0, 1, 2, \dots \quad (7.46)$$

and the sequence of values is

$$\begin{aligned} g(0) &= 1 - 1 = 0, \\ g(T) &= 1 - e^{\ln(0.5)} = 1 - 0.5 = 0.5, \\ g(2T) &= 1 - e^{2 \ln(0.5)} = 1 - (0.5)^2 = 0.75, \\ g(3T) &= 1 - e^{3 \ln(0.5)} = 1 - (0.5)^3 = 0.875, \\ &\vdots \end{aligned} \quad (7.47)$$

It may be noticed that irrespective of the applied methods, the result of the inverse Z transform from Examples 7.4 and 7.5 are the identical.

### 7.2.7 Z transfer function

Resume the difference equation (7.13) describing the discrete system behavior. Taking into account the property of the causal systems  $N \geq M$ , it is considered that  $N = M = n$  and,  $q_{N=n} = 1$ . The difference equation may be reformulated with the following equivalent forms [1]:

$$\begin{aligned} y(k+n) + q_{n-1}y(k+n-1) + \dots + q_0y(k) &= \\ = p_nu(k+n) + p_{n-1}u(k+n-1) + \dots + p_0u(k) & \end{aligned} \quad (7.48)$$

$$\begin{aligned} y(k) + q_{n-1}y(k-1) + \dots + q_0y(k-n) &= \\ = p_nu(k) + p_{n-1}u(k-1) + \dots + p_0u(k-n) & \end{aligned} \quad (7.49)$$

Applying the Z transform for eq. (7.49) and taking into account the Z transform of a shifted function, the following forms are obtained [8]:

$$\begin{aligned} Y(z) + q_{n-1}z^{-1}Y(z) + \dots + q_0z^{-n}Y(z) &= \\ = p_nU(z) + p_{n-1}z^{-1}U(z) + \dots + p_0z^{-n}U(z) & \end{aligned} \quad (7.50)$$

$$Y(z) = \frac{p_n + p_{n-1}z^{-1} + \dots + p_0z^{-n}}{1 + q_{n-1}z^{-1} + \dots + q_0z^{-n}} U(z) = H(z)U(z)$$

or, in a more compact form:

$$Y(z) = H(z)U(z) \quad (7.51)$$

The notation  $H(z)$  has been used for the  $Z$  transfer function of the discrete system, defined by the expression:

$$H(z) = \frac{p_n + p_{n-1}z^{-1} + \dots + p_0z^{-n}}{1 + q_{n-1}z^{-1} + \dots + q_0z^{-n}} = \frac{p_n z^n + p_{n-1}z^{n-1} + \dots + p_0}{z^n + q_{n-1}z^{n-1} + \dots + q_0} = \frac{Y(z)}{U(z)} \quad (7.52)$$

The  $Z$  transfer function  $H(z)$  describes the relationship between the input and output of a discrete system, in the  $z$  domain. Zero value is assumed for the input and output at time moments previous to the present (initial) moment [9].

Similar to the continuous case, the roots of the polynomial  $P(z)$  from the numerator of the  $Z$  transfer function are denoted as *zeros* of the system and the roots of the polynomial  $Q(z)$  from the denominator of the transfer function are denoted as *poles* of the system. To specify the poles and the zeros it is preferable to use the representation based on the  $z$  variable instead of the  $z^{-1}$  variable. The same poles are specified by both of the representations, but only the first of them clearly indicates the existence of a pole in the origin.

### Example 7.6

To sustain the aforementioned statement, it is considered, for example, the  $Z$  transfer function  $H(z) = z/(z - 1)$ , clearly indicating the presence of a zero in the origin [1, 4, 12]. The equivalent form  $H(z) = 1/(1 - z^{-1})$  does not directly reveal this zero.

Similar to the case of the linear continuous systems, the effect superposition principle and the cause-effect proportionality principle have direct consequences for the way of determining the discrete response of a linear discrete system to any given discrete input. The response (output) of a linear discrete system to a certain discrete input may be determined by the summation of the discrete Dirac impulse responses weighted by the input function values from the sampling moments (the discrete input function may be approximated by this sum of Dirac impulses weighted by the input function values at the sampling moments). Knowing the response to an input represented by the discrete Dirac impulse, the response to any given input may be determined by the help of the *discrete convolution* (having the form of a sum):

$$y(k) = \sum_{i=0}^k h(i)u(k-i). \quad (7.53)$$

By the term *discrete Dirac impulse* it has been denoted the succession of values 1, 0, 0, 0, ..., from the sampling moments 0,  $T$ ,  $2T$ ,  $3T$ , .... The discrete Dirac impulse response, having the notation  $h(iT)$  or simply  $h(i)$ , is represented by the succession of values  $h(0)$ ,  $h(T)$ ,  $h(2T)$ , ....

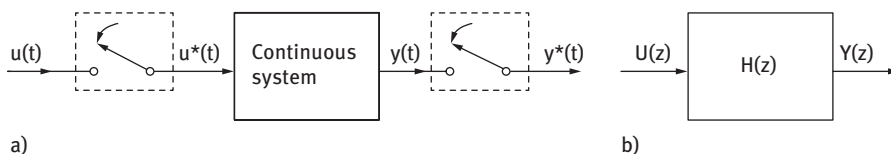
Analyzing the expressions from eqs. (7.51) and (7.53), it may be noticed that, similar to the continuous systems case, for the discrete systems there is also a direct

**i**

correspondence between the *multiplication operation of the Z transform images* (from the  $z$  domain) and the *discrete convolution operation* of the sequences of values of the sampled signals (in the discrete time domain). The discrete Dirac impulse response  $h(iT)$   $i = 1, 2, 3, \dots$  is the inverse Z transform of the Z transfer function  $H(z)$ . In the following, the discrete Dirac impulse will be simply denoted as Dirac impulse and the discrete Dirac impulse response as Dirac impulse response.

### 7.2.8 Z transfer function of the sampled system

For the digital (discrete) control systems it is often met the case when a continuous system is controlled by a digital computing system. For a rigorous analysis it is necessary to describe the sampled system, presented in Fig. 7.14(a), by the  $z$  domain representation according to Fig. 7.14(b) [1, 9].



**Fig. 7.14:** (a) Sampled continuous system and (b) Sampled continuous system using Z transform representation.

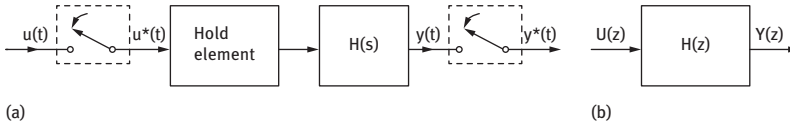
The input function  $u(t)$  of the continuous system is sampled  $u^*(t) = \sum(u(kT) \cdot \delta(t - kT))$  but the output is continuous.

If the differential equation describing the continuous system is known, the Z transfer function  $H(z)$  may be determined by the help of the Z transform. Knowing the transfer function of the continuous system  $H(s)$ , the Z transfer function may be determined by the help of the  $\zeta$  transform defined by [9]:

$$H(z) = Z \left[ L^{-1}(H(s)) \Big|_{t=kT} \right] = \zeta(H(s)). \quad (7.54)$$

It may be mentioned that the inverse  $\zeta$  transform  $H(s) = \zeta^{-1}(H(z))$  is not unique due to the same reason of not being unique the inverse Z transform.

For practical applications the control of a continuous system using a sequence of Dirac impulses  $u^*(t)$  is not possible. The control action must have a “continuous” characteristic, usually given by the zero-order hold element. Its place is between the sampling element and the process, as presented in Fig. 7.15(a) [1, 9].



**Fig. 7.15:** (a) Discretized continuous system with hold element. (b) Discretized continuous system represented using the Z transform.

Knowing the transfer function of the zero-order hold element  $H_0(s) = (1 - e^{-sT})/s$  and taking into account the translation property of the Laplace transform, the determination of the Z transfer function for the system presented in Fig. 7.15(b) leads to

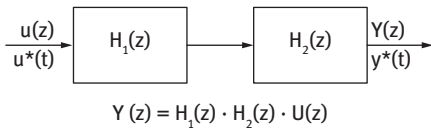
$$\begin{aligned}
 H(z) &= \zeta \left( \frac{1 - e^{-sT}}{s} H(s) \right) = Z \left( L^{-1} \left( (1 - e^{-sT}) \frac{H(s)}{s} \right) \Big|_{t=kT} \right) \\
 &= Z(g(kT) - g(kT - T)) \\
 H(z) &= (1 - z^{-1}) \zeta \left( \frac{H(s)}{s} \right), \tag{7.55}
 \end{aligned}$$

The notation  $g(t)$  has been used for the inverse Laplace transform of the function  $H(s)/s$ , with  $g(t) = \zeta^{-1}(H(s)/s)$ .

**7.2.9 Z transfer function of the interconnected systems**

A few simple cases of discrete control systems, having sampling elements, will be presented in the following [1, 9]. The results may be obtained using the convolution property. It may be noticed that the  $\zeta$  transform presents the property  $\zeta(H_1(z) \cdot H_2(z)) \neq \zeta(H_1(z)) \cdot \zeta(H_2(z))$ .

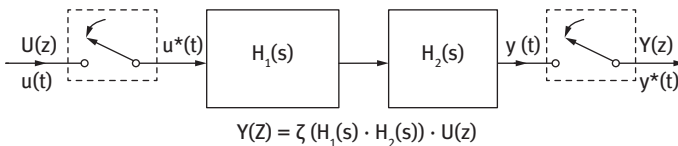
1. Series interconnected discrete systems, Fig. 7.16:



**Fig. 7.16:** Series interconnected discrete systems.

$$Y(z) = H_1(z) \cdot H_2(z) \cdot U(z)$$

2. Series interconnected continuous systems without intermediate sampling element, Fig. 7.17 [1]:



$$Y(z) = \zeta(H_1(s) \cdot H_2(s)) \cdot U(z)$$

**Fig. 7.17:** Series interconnected continuous systems without intermediate sampling element.

3. Series interconnected continuous systems with intermediate sampling element, Fig. 7.18 [1, 9]:

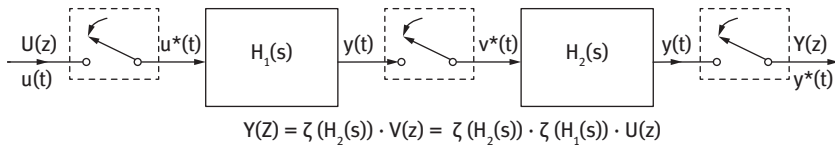


Fig. 7.18: Series interconnected continuous systems with intermediate sampling element.

4. Continuous system before the sampling element, Fig. 7.19 [1, 9]:

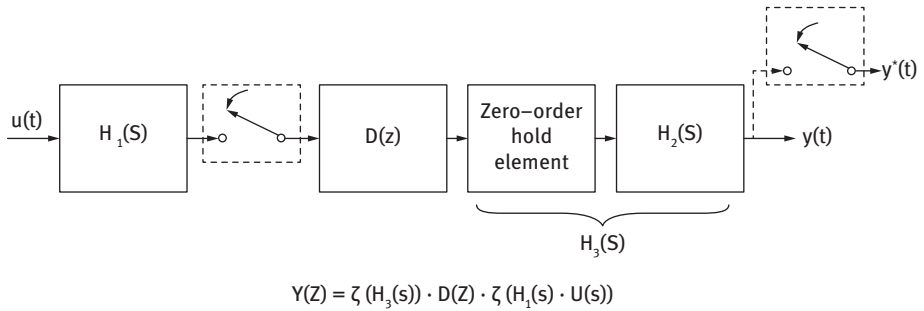


Fig. 7.19: Series interconnected systems with continuous system placed before the sampling element.

5. Discrete feedback control system, Fig. 7.20 [1, 9]:

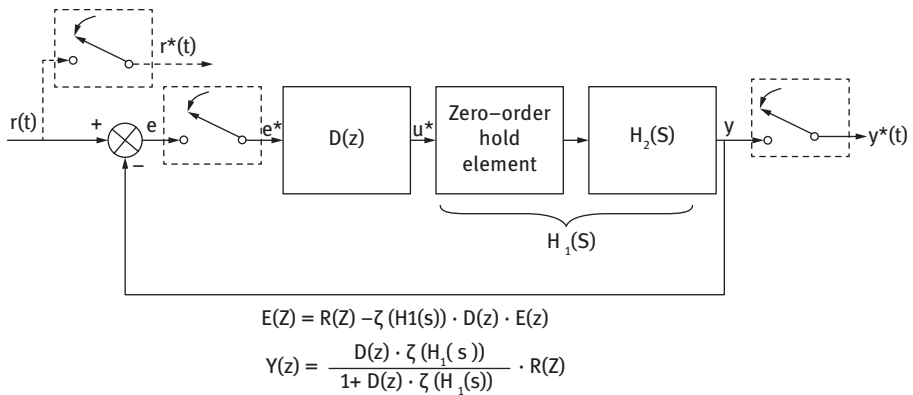


Fig. 7.20: Discrete feedback control system.

### 7.3 Discrete PID controller

The discrete PID controller algorithm emerges from the discretization of the continuous form of the PID control law. As it works on a discrete time bases, the control variable computation has as input the sampled value of the error (or controlled variable) and the computed control variable for a sampled time moment is kept constant until the next sampling time by the zero-order hold element. Generally, the discrete PID tuning and desired performance may be assessed and achieved based on similar considerations used for its continuous correspondent [10]. The discrete PID control law may be described by [4, 11, 12]:

$$u(kT) = K_c[e(kT) + \frac{T}{T_I} \sum_{i=0}^k e(iT) + \frac{T_D}{T}(e(kT) - e((k-1)T))] + u_0 \quad (7.56)$$

where the discrete integral control mode has been obtained by rectangular approximation of the continuous integral of the error and the discrete derivative control mode by the backward difference approximation of the continuous derivative of the error. The initialization term  $u_0$  is the initial value of the manipulated variable, when the error should be zero.

A commonly used form of the PID controller law does not take into account the derivative effect of the setpoint variable (inherent in the error) and changes to the following discrete PID control law [4, 11]:

$$u(kT) = K_c[e(kT) + \frac{T}{T_I} \sum_{i=0}^k e(iT) - \frac{T_D}{T}(y(kT) - y((k-1)T))] + u_0 \quad (7.57)$$

where  $y(kT)$  is the sampled controlled variable. The (7.56) and (7.57) forms of the discrete PID are denoted as *position form* of the algorithm, since they compute the full control variable for every sampling moments.

The summation implied by the integral mode does not have to be performed at every sampling moment as the sum of errors up to the previous sampling moment:

$$S((k-1)T) = \sum_{i=0}^{k-1} e(iT) \quad (7.58)$$

may be saved (stored) and used in the following form of the discrete PID:

$$u(kT) = K_c[e(kT) + \frac{T}{T_I}(S((k-1)T) + e(kT)) - \frac{T_D}{T}(y(kT) - y((k-1)T))] + u_0 \quad (7.59)$$

In the the *velocity form* of the discrete PID algorithm the control variable change is computed, according to the following relationship [4, 11]:

$$\begin{aligned} \Delta u(kT) &= u(kT) - u((k-1)T) \\ &= K_c[e(kT) - e((k-1)T) + \frac{T}{T_I} e(kT) - \frac{T_D}{T}(y(kT) - 2y((k-1)T) + y((k-2)T))] \\ &= K_c[(1 + \frac{T}{T_I})e(kT) - e((k-1)T) - \frac{T_D}{T}(y(kT) - 2y((k-1)T) + y((k-2)T))] \end{aligned} \quad (7.60)$$



The (7.60) form of the velocity algorithm can be obtained by subtracting the position forms corresponding to two consecutive sampling moments, i.e.  $u(kT)$  and  $u((k-1)T)$ .

The velocity form of the discrete PID algorithm has appreciated incentives compared to the discrete position PID as it does not need initialization and is not showing the integral windup undesired behavior.

Due to sampling, an additional time delay (dead time) is appearing in the control loop. This dead time is equal to the sampling time and has negative effect on the control performance. As a result, it may be expected that discrete controller's performance is not as good as its continuous counterpart. This aspect is important especially when the sampling time is large compared to the feedback dynamics. Some authors suggest the choosing of the sample time for the discrete controller such as to be less than 5% of the sum between the feedback dead time and the largest time constant [11].

Nevertheless, tuning the discrete PID controllers may be performed on the bases of the same methods used for tuning the continuous PID controllers (Ziegler-Nichols, Cohen-Coon, minimizing the integral square error or integral absolute error, etc.) [13]. Some authors propose that tuning of the PID discrete controller should be made in a similar way of the continuous PID tuning but considering an additional dead time equal to half of the sampling time [12].

Stability of the discrete systems may be analyzed on the system's Z transfer function, which has the form of a ratio between polynomials in the z variable,  $H(z) = P(z)/Q(z)$ . The discrete system is stable if all roots of the denominator polynomial  $Q(z)$ , denoted as *poles* of the system, are situated inside the unit circle of the complex plane. If at least one of the roots of the polynomial  $Q(z)$  is situated outside the unit circle the system shows instability. Roots of  $Q(z)$  that have real parts near the value  $-1$  (but inside the unit circle) show a strongly oscillating behavior, denoted as *ringing* [4, 11].

The discrete controller may produce the ringing undesired effect if it is not properly designed or tuned. Controller ringing is produced by the changes of the controller output, with high amplitude and different sign (slowly damping oscillations of the control variable), producing degradation of the control system performance. Basically, the reduction of the controller gain and increasing the integral time can reduce the controller ringing of the discrete PID. For the general case of the discrete controllers, ringing may be avoided by designing the poles of the controller such as to be placed inside the positive region of the unit circle and to have absolute values in the [0.4 0.6] interval [11].

Particular attention must be given to the integral mode of the PID controller, when the control valve reaches saturation, i.e. the valve totally open or totally closed position, and the error has not been yet eliminated. This situation makes the control system to be unable of further acting efficiently for obtaining zero offset. However, as time passes, the effect of the control variable on the controlled variable is increasing in time and the error is reduced. At some moment, this delayed reduction of the error determines the control valve to leave its completely open or completely closed position and the control signal will re-enter in the active domain, i.e. between the 0% and the 100% positions

of the valve opening and finally, the error will be eliminated. The problem generated by the integral mode is that during the period of saturation of the control valve, as error is still present, the integral of the error produces the continuous increase or continuous decrease of the control variable value. As a result, the control variable reaches very large or very small values, phenomenon called *integral (or reset) windup*. Bringing back the control variable value in the active region of the control valve may imply a long period of time and this causes the control system to be inefficient in the meanwhile.

Integral windup can be generated by improper design of the control system equipment or by unexpectedly large disturbances. Avoiding the integral windup can be achieved by appropriate equipment design or by special *anti-reset windup* adaptations brought to the traditional PID controller structure. The latter implies the use of a saturation element. Such an anti-reset windup structure, for a PI controller, is presented in Fig. 7.21 (for the continuous case) [11]. As long as signal  $u$  does not reach the saturation limits, the PI controller works as a traditional one. But when saturation is reached the controller output is limited to  $u_{lim}$ .

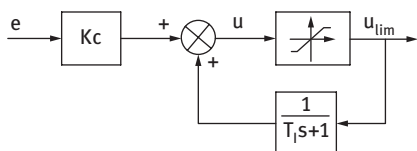


Fig. 7.21: PI controller with anti-reset windup.

Velocity form of the discrete PID controller does not accumulate the integral error and it avoids the integral windup by its design.

## 7.4 Other forms of the discrete controllers

The discrete approach of designing controllers allows the computation of new forms of the controllers that may differ from the classical discrete PID, while satisfying desired control performance [4, 11, 12].

Starting from the Z transfer function of the discrete feedback control loop presented in Fig. 7.20 it is possible to design the controller Z transfer function  $D(z)$  such as for a given form of the Z transform of the setpoint change  $R(z)$  a desired closed loop Z transform response  $Y(z)$  is desired (imposed). The controller satisfying this desired behavior of the closed control loop is described by the following Z transfer function:

$$D(z) = \frac{\frac{Y(z)}{R(z)}}{\zeta(H_1(z)) \left(1 - \frac{Y(z)}{R(z)}\right)} = \frac{\frac{Y(z)}{R(z)}}{H_1(z) \left(1 - \frac{Y(z)}{R(z)}\right)} = \frac{Y(z)}{H_1(z)(R(z) - Y(z))} \quad (7.61)$$

where  $R(z)$ ,  $Y(z)$ , and  $H_1(z)$  are known.

This newly designed discrete controller has the form of a ratio of two polynomials in  $z^{-1}$ :

$$D(z) = \frac{U(z)}{E(z)} = \frac{p_0 + p_1 z^{-1} + \dots + p_n z^{-n}}{1 + q_1 z^{-1} + \dots + q_m z^{-m}} \quad (7.62)$$

This is equivalent to

$$U(z)(1 + q_1 z^{-1} + \dots + q_m z^{-m}) = E(z)(p_0 + p_1 z^{-1} + \dots + p_n z^{-n}) \quad (7.63)$$

Making the inverse Z transform of eq. (7.63), the following algorithm form for computing the control variable  $u(kT)$  at the sampling moment  $t = kT$  may be developed:

$$u(kT) = p_0 e(kT) + p_1 e((k-1)T) + \dots + p_n e((k-n)T) - \dots - q_1 u((k-1)T) - \dots - q_m u((k-m)T) \quad (7.64)$$

Care should be taken that the discrete controller to be physically realizable, i.e. the computation of the control variable  $u(kT)$  at the sampling moment  $t = kT$  to depend only on past or utmost present values of the error and of the control variable. This is equivalent to the condition that powers of  $z$  in the polynomial  $P(z)$  to be strictly negative.

Two of the most renowned results of such new discrete controller designs are the deadbeat and the Dahlin controllers [4].

The *deadbeat controller* is designed requiring the controller (for the sampling moments) to respond to the step change of the setpoint by producing the same step change on the controlled variable but delayed by one sampling period. Usually, this design builds controllers that are very aggressive and show strongly oscillating behavior. This design may be also applied to processes having dead time, but physically realizable forms may be only obtained if the desired step response of the controlled variable cumulates one sampling time period to which is added an integer number of sampling time periods exceeding the process dead time (in which are also included the dead time of the control element and measuring device).

The *Dahlin controller* is designed requiring the controller (for the sampling moments) to respond to the step change of the setpoint by producing a first order with dead time behavior of the controlled variable. This design makes the control to be less aggressive and the the oscillatory behavior is diminished. Physically realizable controllers may be only obtained if the desired response of the controlled variable cumulates one sampling time period to which is added an integer number of sampling time periods exceeding the process dead time (considering also included the dead time of the control element and measuring device).

## References

- [1] Cristea, M.V., Agachi, S.P., *Elemente de Teoria Sistemelor*, Editura Risoprint, Cluj-Napoca, 2002.
- [2] Åström, K.J., Wittenmark, B., *Computer-Controlled Systems Theory and Design*. Prentice Hall, Englewood Cliffs, NJ, 1997.
- [3] Miclea, L., *Fiabilitatea și diagnoza sistemelor digitale*. Universitatea Tehnică Press, Cluj-Napoca, 1998.
- [4] Stephanopoulos, G., *Chemical Process Control An Introduction to Theory and Practice*. Prentice Hall, Englewood Cliffs, NJ, 1984.
- [5] Edward, A.L., Varaiya, P., *Structure and Interpretation of Signals and Systems*. University of California, Berkeley, CA, 2000.
- [6] Schroer, W., *Systetheorie. Vorlesungsmanuskript*. Ulm Universität, 1998.
- [7] Corovei, I., Pop, V., *Transformări integrale, calcul operațional*. Universitatea Tehnică Press, Cluj-Napoca, 1993.
- [8] Mihoc, D., Ceapău, M., Iliescu, S. St., Bornagiu, I., *Teoria și elementele sistemelor de reglare automată*. Editura Didactică și Pedagogică, București, 1980.
- [9] Kottmann, M., Kraus, F., Schaufelberger, W., *Signal und Systemtheorie II*, Skript zur Vorlesung, ETH Institut für Automatik, Zürich, 2001.
- [10] Feștilă, C., *Reglatoare automate-îndrumător de proiectare*. Institutul Politehnic, Cluj-Napoca, 1990.
- [11] Marlin, T.E., *Process Control- Designing Processes and Control Systems for Dynamic Performance*, McGraw-Hill, 2000.
- [12] Franklin, G., Powell, J., Workman, M., *Digital Control of Dynamic Systems*, Addison-Wesley, Reading, MA, 1990.
- [13] Seborg, D.E., Edgar, T.F., Mellichamp, D.A., Doyle III, J.F., *Process Dynamics and Control*, 3rd Edition, John Wiley & Sons, New York, 2003.
- [14] Gerald, C.F., Wheatley, P.O., *Applied Numerical Analysis*, 7th Edition, Addison Wesley, New York, 2003.
- [15] Agachi, P.S., Cristea, M.V., *Basic Process Engineering Control*, Walter de Gruyter, Berlin/Boston, 2014.
- [16] Agachi, S., *Automatizarea Proceselor Chimice (Chemical Process Control)*, Casa Cartii de Stiinta, Cluj Napoca, 320, 1994.



---

## Part II: Applied Process Engineering Control

The reason for the addition of this second part in the present book is that the control of multivariable processes is a complex enterprise, even if handled with independent control loops. Industrial processes can be grouped mainly as belonging either to the *synthesis* or the *separation* stage; each is complex and has multiple inputs and outputs. The synthesis stage occurs in reactors of various types (such as CSTR, BSTR, PFR, and a multitude of other kind of reactors). The separation stage is chosen in agreement with characteristics of the products to be separated. In addition, we would like to draw attention upon the importance of utilities' control (steam, water, electricity, etc.). These are extremely important from economical point of view.

This part, *Applied Process Engineering Control*, applies the theoretical concepts of the first part. The control of any process, either a synthesis or a separation, is discussed from the point of view of the intrinsic needs of the process. Thus, control strategy is proposed in agreement with these requirements.

Consequently, the principles of the control system design used in this chapter are the following:

1. Study of process to be controlled as well as of correlation between controlled and all possible manipulated variables;
2. Design of steady-state mathematical model;
3. Quantifying goal variables (e.g. chemical conversion, material humidity, waste composition) as well as input and state variables;
4. If necessary, sensitivity and relative gain array analysis of the process to eliminate the unimportant correlations between output and input/state variables and simplify the control scheme;
5. Design of control scheme.

## 8 Reaction unit control

### 8.1 Introduction

This chapter deals with basic control strategies of chemical synthesis units. Since temperature affects rates of both physical and chemical processes, its efficient control has to be omnipresent. Consequently, this topic is given special attention; thermal instability and ways to avoid it are described in detail. Yet, pressure and liquid level control as well as ways of coping with the logarithmic dependence of pH on the acid or base concentration are also discussed. Process end-point detection is coupled with aspects of product quality control. The principles of designing the control structure for some homogeneous and heterogeneous, continuous or batch units are presented and electrochemical reactors are also mentioned within this context. The picture is completed with pertinent examples.

The *reactor* is the core of most chemical production facilities, other operations are related to its necessities and performance. Therefore, its smooth and stable operation under well-defined parameters ensures both *profitability* and *safety*. The goal of each control strategy is to minimize as much as possible, if not entirely eliminate, the effect of disturbing parameters on the desired quality of the process in terms of product distribution, energy consumption, exploitation safety, etc. There is no general “recipe” to achieve all that; smart solutions differ from one process to the other but rely on the same basic principles. An appropriately optimized design for both batch and continuous chemical reactors can improve plant productivity up to 25% [1]. Such results may be achieved using adequate individual control loops and strategies to program the main parameters but also to provide sequencing and record-keeping functions.

Success or failure of a certain control system strongly relates to the reactor design. An incorrect layout could result in an unstable reactor, regardless of the employed control strategy and even in the absence of a significant disturbance. For example, strong oscillations in temperature or pressure could compromise product or catalyst, damage equipment or injury personnel. Hence, the choice of adequate control should be in conjunction with the inherent characteristics of each reactor type.

### 8.2 Basic concepts of ideal continuous and batch units

The very basic unit types [1–4] dealt with in chemical reaction engineering and design are *continuous stirred tank reactors* (CSTRs), *tubular plug flow reactors* (PFRs) and *batch stirred tank reactors* (BSTRs). In terms of flow and thermal conditions, they are referred to as *ideal* and may work in *dynamic* or *steady-state*. Description of industrial units, either homogeneous or heterogeneous, derives from these basic models.

Uniform inlet and outlet flows are assumed for continuous units CSTR and PFR. The average amount of time spent by a molecule in the unit is called *residence time*.



It is defined by  $\frac{V}{F}$ , the ratio between the reactor volume  $V$  ( $\text{m}^3$ ) and the total volumetric flow  $F$  ( $\text{m}^3/\text{s}$ ).

A CSTR assumes ideal mixing; therefore, parameters are uniform within the entire volume of the reaction mixture. Each point in space is characterized by the same value of residence time, temperature, and conversion. These are also the attributes of the outlet flow. A PFR assumes total radial but no axial (flow direction) mixing; the reaction mixture advances like a *plug* along the longitudinal axis of the tube. Thus, reactor parameters are distributed along it, and in each point residence time, temperature and conversion have different values. In other words, such a reactor functions like a cascade of infinitely small stirred tank reactors.

A BSTR is operated in cycles of one load each. The time amount necessary for the load to reach the desired end-point of the chemical process is called *reaction time*  $t_{\text{reaction}}$ . Parameters will vary as it advances. This is the reason for their more difficult control (see section 8.7). Yet, since ideal mixing is assumed in such units, the temperature and conversion values will be uniformly distributed at a certain moment through the entire volume of the reaction mass.

Two main types of *batch* units are operated in chemical industry: the pure-batch and the *fed-batch* (for example neutralizing processes). In a *pure-batch* unit, reactants are charged all at once at the beginning of the process. The fed-batch is loaded only partially at the beginning of a cycle and some reactants are added at a well-defined rate during the advance of the reaction.

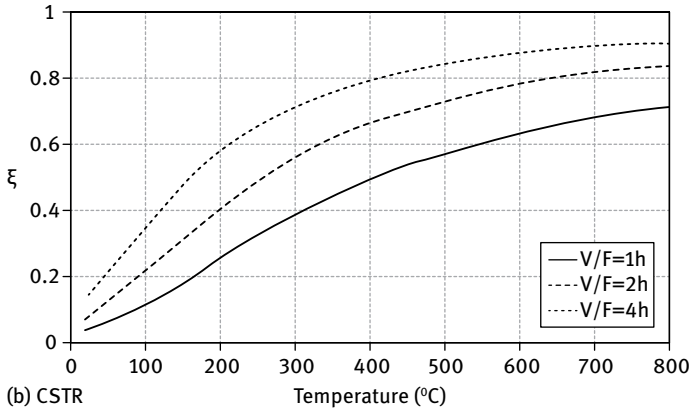
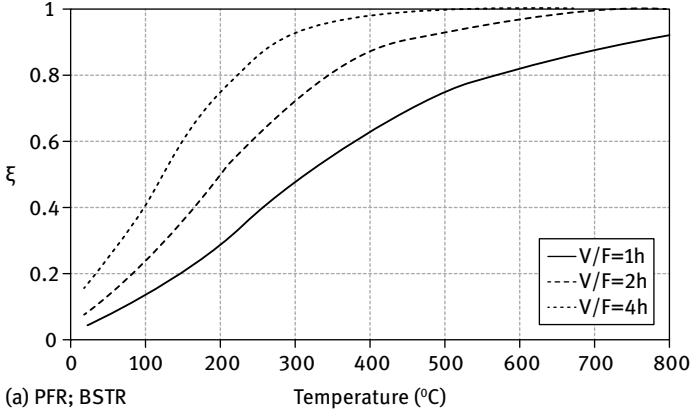
Primordial performance criteria for a reactor are residence time  $\frac{V}{F}$  or reaction time  $t_{\text{reaction}}$ , conversion  $\xi$ , and selectivity  $\Phi$  [1, 3–6]. Conversion is the only advance variable of a chemical process that is defined with respect to the limiting reactant (species  $A$  in equations below), whereas selectivity relates to the desired product (species  $P$ ). The *limiting reactant* is the species present in the smallest stoichiometric amount in the mixture at reaction initiation, while all others are in *excess*.

Definitions (8.1) and (8.2) assign both  $\xi$  and  $\Phi$  values between 0 and 1.  $n_{A0}$  and  $n_{P0}$  stand for the influent amount (expressed in moles) of  $A$  and  $P$ , whereas  $n_A$  and  $n_P$  for the corresponding effluent values. If the total volume of the reaction mass does not vary significantly during reaction course, the molar amounts can be replaced with the corresponding molar concentrations.

$$\xi = \frac{n_{A0} - n_A}{n_{A0}} \quad (8.1)$$

$$\Phi = \frac{n_P - n_{P0}}{n_{A0} - n_A} \quad (8.2)$$

For an isolated first-order reaction  $A \xrightarrow{k} P$ , eqs. (8.3)–(8.5) describe conversion vs temperature  $\xi = f(T^\circ)$  patterns for various unit types [1, 3–4].  $k$  stands for the rate coefficient ( $\text{s}^{-1}$ ) and, as a rule [5–6], depends exponentially on temperature – see eq. (8.6).



**Fig. 8.1:** Conversion vs temperature patterns for: (a) steady-state PFR and BSTR; (b) steady-state CSTR. Kinetic data are  $k_0 = 11.87 \text{ h}^{-1}$ ,  $E_a = 14,000 \text{ J/mol}$ .

$$\xi_{CSTR} = \frac{k \frac{V}{F}}{1 + k \frac{V}{F}} \quad (8.3)$$

$$\xi_{PFR} = 1 - \exp\left(-k \frac{V}{F}\right) \quad (8.4)$$

$$\xi_{BSTR} = 1 - \exp\left(-k t_{\text{reaction}}\right) \quad (8.5)$$

$$k = k_0 \exp\left(-\frac{E_a}{RT^{\circ}}\right) \quad (8.6)$$

where

- $k_0$  stands for the pre-exponential coefficient ( $\text{s}^{-1}$ , same units as  $k$ ),
- $E_a$  for the activation energy (J/mol),

- $R = 3.314 \text{ J/mol} \cdot \text{K}$  is the universal gas constant,
- and  $T^\circ$  stands for the reaction temperature (K).

The above relationships as well as Fig. 8.1 demonstrate that these profiles are non-linear and vary from one reactor design to the other. For the same process as well as identical time and temperature values, PFR and BSTR (Fig. 8.1(a)) are kinetically superior to CSTR (Fig. 8.1(b)) since the corresponding  $\xi$  is higher.

Optimizing  $\xi$  and  $\Phi$  ensures minimum raw material consumption and best product distribution. However, other constraints and limitations also have to be considered: avoiding local overheating of catalyst, ensuring desired particle sizes or molecular weights for a product, batch-to-batch uniformity, etc. Under real operating conditions, residence times are not uniform and their distribution pattern also affects  $\xi$  and  $\Phi$  values [3–4]. Catalyzed or uncatalyzed heterogeneous reaction systems that involve fluid diffusion through various media, fluid adsorption/desorption to a solid surface, or gas absorption/dissolution into a liquid are influenced in addition by the nature of flow regime, diffusion characteristics, or thermodynamics and kinetics of the interphase contact processes.

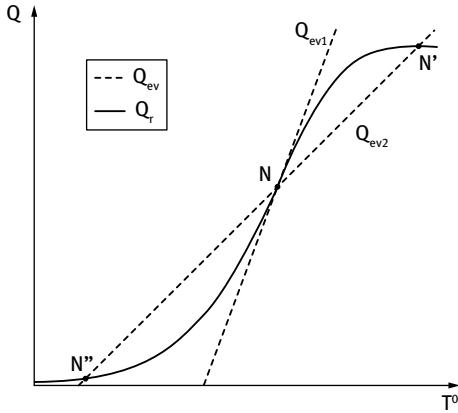
### 8.3 Temperature control

Reaction temperature is almost always one of the controlled variables. Some examples of extreme temperature sensitive goals are lowering side product accumulation rates, keeping narrow polymer molecular weight distributions or protecting integrity of glass-lined reactors. Therefore, some systems demand high performance temperature regulation within  $0.2^\circ\text{C}$ – $0.3^\circ\text{C}$  [1].

Because runaway *exothermic* reactions can destabilize an operating reactor, special attention is given to *thermal stability*. *Endothermic* reactions are self-regulating; if external heat-delivery suffers, the process slows down or stops. Heat transfer is also self-regulating; its rate will increase by improvement of any parameter describing heat exchange.

#### 8.3.1 Into thermal instability

*Thermal instability* of a unit in function may occur only for exothermic processes when the slope of the heat-release  $Q_r$  curve exceeds that of the heat transfer  $Q_{ev}$  (see Fig. 8.2). Such a situation can start *runaway* behavior: an uncontrolled increase of temperature speeds up the process, which in return releases more heat and further elevates temperature. Thus, the process responds to a temperature disturbance with a positive feedback. A controller compensates it with a negative feedback by increasing the heat transfer rate. The stronger feedback decides whether the unit is thermally stable or not.



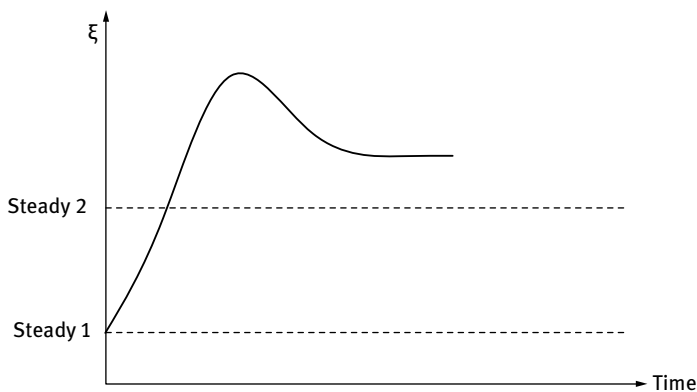
**Fig. 8.2:** Reaction generated  $Q_r$  and evaporated  $Q_{ev}$  heat rate vs temperature patterns. Qualitative representation. (Adapted after [3, 4])

Thermal stability is ensured when:

$$\left( \frac{dQ_{ev}}{dT^o} \right) > \left( \frac{dQ_r}{dT^o} \right) \quad (8.7)$$

Figure 8.3 illustrates the dynamic variation of conversion from steady-state 1 to steady-state 2. It is obvious that the instantaneous variation  $\left( \frac{d\xi}{dT^o} \right)$  of conversion with temperature is higher for the dynamic mode than that for the steady-state (since  $\xi < 1$ ). The relationship between them is described in the literature [7–9]:

$$\left( \frac{d\xi}{dT^o} \right)_{dynamic} = \left[ \frac{1}{1-\xi} \right] \left( \frac{d\xi}{dT^o} \right)_{steady} \quad (8.8)$$



**Fig. 8.3:** Dynamic variation of conversion between two steady states. Qualitative representation. (Adapted after [8])

A singular first-order reaction  $A \xrightarrow{k} P$  is characterized by the rate law:

$$r = k C_A = k_0 \exp\left(-\frac{E_a}{RT^0}\right) C_A \quad (8.9)$$

where  $r$  stands for reaction rate ( $\text{kmol}/\text{m}^3 \cdot \text{s}$ ),  $k$  for the first-order rate coefficient ( $\text{s}^{-1}$ ), and  $C_A$  for the momentary molar concentration ( $\text{kmol}/\text{m}^3$ ) of reactant species  $A$ . The significance of  $E_a$ ,  $R$ , and  $T^0$  are the same as given for eq. (8.6). Using eqs. (8.1)–(8.6), the statement in eq. (8.8) will verify for both a CSTR and a PFR.

– For a continuous stirred tank reactor:

$$\left(\frac{d\xi}{dT^0}\right)_{dynamic} = \left(\frac{E_a}{R(T^0)^2}\right) \xi > \left(\frac{d\xi}{dT^0}\right)_{steady} = \left(\frac{E_a}{R(T^0)^2}\right) \xi(1-\xi) \quad (8.10)$$

– For a plug flow reactor:

$$\left(\frac{d\xi}{dT^0}\right)_{dynamic} = \left(\frac{E_a}{R(T^0)^2}\right) \ln(1-\xi) > \left(\frac{d\xi}{dT^0}\right)_{steady} = \left(\frac{E_a}{R(T^0)^2}\right) (1-\xi) \ln(1-\xi) \quad (8.11)$$

Depending on  $\xi$ , the value of  $\left(\frac{d\xi}{dT^0}\right)$  at the same temperature  $T^0$  can be higher for the dynamic state. Thus, thermal stability condition (8.7) is not satisfied and the unit becomes unstable. As a consequence, correct design requires checking the validity of constraint (8.7) for both dynamic and steady states at setpoint temperature. Otherwise, a reactor apparently thermally stable under steady-state becomes unstable when operated dynamically – see relationship (8.8).

### 8.3.2 Out of thermal instability

Thermal instability can be kept under control, which requires simultaneous solutions for both mass and energy balances. For the exothermic first-order reaction  $A \xrightarrow{k} P$  occurring in a CSTR, the mass balance for reactant  $A$  is

$$V k C_A = F(C_{A0} - C_A) = F \xi C_{A0} \quad (8.12)$$

The corresponding reaction *generated heat rate* is

$$Q_r = -\Delta H_r V k C_A = -\Delta H_r F \xi C_{A0} \quad (8.13)$$

where  $C_{A0}$  is the initial molar concentration of reactant  $A$  ( $\text{kmol}/\text{m}^3$ ) and  $\Delta H_r$  is the exothermic reaction heat ( $\text{kJ}/\text{kmol}$ ). The *evacuated heat rate* depends on the geometry of the reactor (by means of  $A_T$ ) and the characteristics of the material from which it was built and coated (by means of  $K_T$ );  $\rho$  is the reaction mixture density ( $\text{kg}/\text{m}^3$ ),  $C_p$  for a mean reaction mixture heat capacity ( $\text{J}/\text{kg} \cdot \text{K}$ ),  $A_T$  for the heat transfer surface ( $\text{m}^2$ ),

$K_T$  for a mean heat transfer coefficient between the reaction mixture and the coolant ( $\text{J/s} \cdot \text{m}^2 \cdot \text{K}$ ), and  $T_{ag}^{\circ}$  for the coolant average temperature.

$$Q_{ev} = F \rho C_p T^{\circ} + K_T A_T (T^{\circ} - T_{ag}^{\circ}) \quad (8.14)$$

Relationships (8.12)–(8.14) are written under the assumption that no volume variation occurs due to reaction advance. In other words, the density of the entire mixture remains constant regardless of  $\xi$  (such a dependence is described by linear functions in the literature [5–6]). Similarly, the value of  $C_p$  is considered to characterize the entire mixture, to be temperature independent and therefore constant. Thus,  $Q_{ev}$  in eq. (8.14) depends linearly on  $T^{\circ}$  (see Fig. 8.2).

Meanwhile, because of the nonlinear temperature dependence of terms  $k$  and  $\Delta H_r$  [5–6], the generated heat  $Q_r$  exhibits nonlinearity when plotted against temperature. This fact is illustrated by Fig. 8.2. The thermal stability requirement (8.7) is fulfilled in setpoint **N** for the line indicating  $Q_{ev1}$ . If the temperature varies either way around its value in **N**, the process will slip back and recover to the parameters of **N**. Hence, the reactor is thermally stable. Meanwhile, the line corresponding to  $Q_{ev2}$  intersects the curve of  $Q_r$  in more than one point, among which **N** is unstable and both **N'** and **N''** are stable. For example, if the temperature rises over the setpoint value in **N** where  $Q_r > Q_{ev}$ , the disturbance will result in a positive feedback of the process and the temperature will further increase till point **N'**. Beyond it  $Q_r < Q_{ev}$ , and the system recovers to the operating conditions of **N'**. Thus, it describes a thermally stable state. The same judgment can be applied to point **N''**, proving its stability as opposed to **N**. The instability of setpoint **N** can be eased by enhancing the heat transfer through adjustment of  $A_T$ ,  $K_T$ , and  $T_{ag}^{\circ}$  values, respectively.

Analysis of Fig. 8.2 leads to the conclusion that the thermally stable unit has an adequate heat transfer surface. Therefore, it requires for an unsophisticated temperature control loop, as the one exemplified in Fig. 8.4 for a CSTR. This may involve an ordinary negative response proportional-integral controller [1–2, 7–9]. Meanwhile, the thermally unstable reactor, demands a fast and strong enough negative feedback in the control loop to adequately compensate the positive response of the chemical process. Hence, the control system has to fulfill some *constraints* to stabilize this open-loop instability.

*Quantitative assessment* of thermal stability or instability of functioning units is based on checking the heat balance for the setpoint temperature  $T^{\circ}$ . For the first-order reaction discussed above, thermal equilibrium means

$$F \rho C_p T_{in}^{\circ} - \Delta H_r F \xi C_{A0} = F \rho C_p T^{\circ} + K_T A_T (T^{\circ} - T_{ag}^{\circ}) \quad (8.15)$$

The left side of eq. (8.15) stands for the heat-inlet rate, whereas the right side for the outlet rate.  $T_{in}^{\circ}$  stands for the influent reactant temperature. Thermal stability, that is satisfaction of condition (8.7), translates into eq. (8.16), where  $\left(\frac{d\xi}{dT^{\circ}}\right)$  is assessed for both dynamic and steady states.

$$F \rho C_p + K_T A_T > \left| \Delta H_r F C_{A_0} \left( \frac{d\xi}{dT^o} \right) \right| \quad (8.16)$$

Figure 8.1 shows that the conversion  $\xi$  depends nonlinearly on the temperature  $T^o$ . Therefore,  $\left( \frac{d\xi}{dT^o} \right)$  is approximated by a linearization with respect to the setpoint described by  $\bar{\xi}$  and  $\bar{T}^o$  (see Example 8.1 where calculus is carried out for both dynamic and steady states):

$$\bar{\xi} = \xi + \left( \frac{d\xi}{dT^o} \right)_{\bar{T}^o} (\bar{T}^o - T^o) \quad (8.17)$$

The above data serve the computation of time constant  $T_{rm}$  (seconds) and amplification factor  $K_{rm}$  (dimensionless) of reaction mass [1, 7–8]. When these values are positive, the reactor is thermally stable and its temperature control is conventional (see Fig. 8.4).

$$T_{rm} = \frac{V \rho C_p}{F \rho C_p + K_T A_T + \Delta H_r F C_{A_0} \left( \frac{d\xi}{dT^o} \right)_{steady}} \quad (8.18)$$

$$K_{rm} = \frac{K_T A_T}{F \rho C_p + K_T A_T + \Delta H_r F C_{A_0} \left( \frac{d\xi}{dT^o} \right)_{steady}} \quad (8.19)$$

If the reverse is true, and the total heat-outlet rate is lower than the inlet rate, values of  $T_{rm}$  and  $K_{rm}$  become negative. As a result, the process amplification  $\left( \frac{K_{rm}}{\sqrt{1 + T_{rm}^2 \omega^2}} \right)$  is also negative. The phase shift  $\varphi_{rm}$  (rad) is a function of oscillating frequency  $\omega$  (rad/s) and varies between  $-180^\circ$  and  $90^\circ$ .

$$\varphi_{rm} = -180^\circ - \arctg(T_{rm} \omega) = -180^\circ + \arctg(|T_{rm}| \omega) \quad (8.20)$$

The term ( $-180^\circ$ ) indicates a negative gain under steady-state and the sign (+) in front of the  $\arctg$  function implies a negative time constant.

A chemical reactor can function without reaction mass temperature regulation if the control of the coolant temperature is satisfactory [9] (see Fig. 8.4). However, there are two constraints to be imposed on a control system to achieve thermal stability of an unstable reactor:

- First constraint – Imposing lower proportional band  $PB$  values, in agreement with relationship (8.21);  $\omega_{osc}$  stands for the oscillation frequency of the control system:

$$PB < 200 \left( \frac{K_{rm}}{\sqrt{1 + T_{rm}^2 \omega_{osc}^2}} \right) \quad (8.21)$$

- Second constraint – Exclusion of integral time  $T_i$  because of the negative phase shift it generates. Therefore, eq. (8.20) will modify because of regulator phase shift  $\varphi_R$ . The introduction of  $T_i$  brings the system closer to its stability limit described by  $\varphi_{total} = -180^\circ$ .

$$\varphi_{total} = -180^\circ = \varphi_{rm} + \varphi_R = -180^\circ + \arctg(|T_{rm}| \omega) - \arctg\left(\frac{1}{T_i \omega}\right) \quad (8.22)$$

### Example 8.1



Let us consider the described first-order reaction  $A \xrightarrow{k} P$  occurring in liquid phase in a CSTR for which the  $\xi = f(T^\circ)$  dependence corresponds to that in Fig. 8.1(b). The setpoint is given by the ratio  $\frac{\bar{V}}{F} = 2 \text{ h}$  and  $\bar{T}^\circ = 200^\circ\text{C}$ . Inlet flow is preheated to  $T_{in}^\circ = \bar{T}^\circ$ . Other technical data are  $K_T = 2.10^3 \text{ kJ/m}^2 \cdot \text{h} \cdot ^\circ\text{C}$ ;  $A_T = 4 \text{ m}^2$ ;  $F = 5 \text{ m}^3/\text{h}$ ;  $\Delta H_r = -15 \text{ kJ/kmol}$ ;  $C_{A0} = 10 \text{ kmol/m}^3$ ;  $\rho = 10^3 \text{ kg/m}^3$ ;  $C_p = 1 \text{ kJ/kg} \cdot ^\circ\text{C}$ .

Data in Fig. 8.1(b) indicate that the setpoint corresponds to  $\bar{\xi} = 0.4$ . The coolant temperature that ensures thermal stability can be calculated from the energy balance (8.15) by compelling that the heat-outlet flow has to be higher than that of the inlet. Replacement of data yields:  $8 * (200 - T_{ag}^\circ) \geq 300 \text{ kJ/h}$ , where  $T_{ag}^\circ$  is expressed in  $^\circ\text{C}$ . The result is  $T_{ag}^\circ \leq 162.5^\circ\text{C}$ .

A stable reactor, in either dynamic or steady-state, has to check eq. (8.7). Figure 8.1(b) provides data for the calculus of steady-state  $\left(\frac{d\xi}{dT^\circ}\right)$  by employing linearization (8.17) between the setpoint  $\bar{T}^\circ = 200^\circ\text{C}$ ;  $\bar{\xi} = 0.4$  and point  $\bar{T}^\circ = 100^\circ\text{C}$ ;  $\bar{\xi} = 0.2$ . The dynamic value can be further obtained using eq. (8.8).

$$\left(\frac{d\xi}{dT^\circ}\right)_{steady} = \left(\frac{\bar{\xi} - \xi}{\bar{T}^\circ - T^\circ}\right) = \frac{0.4 - 0.2}{200 - 100} \frac{1}{^\circ\text{C}} = 0.0020 \frac{1}{^\circ\text{C}}$$

$$\left(\frac{d\xi}{dT^\circ}\right)_{dynamic} = \left[\frac{1}{1 - \xi}\right] \left(\frac{d\xi}{dT^\circ}\right)_{steady} = \frac{1}{1 - 0.4} \left(0.0020 \frac{1}{^\circ\text{C}}\right) = 0.0033 \frac{1}{^\circ\text{C}}$$

Both values verify thermal stability requirement (8.16): for steady-state  $13 > 1.5 \text{ (kJ/h} \cdot ^\circ\text{C)}$  and for dynamic state  $13 > 2.5 \text{ (kJ/h} \cdot ^\circ\text{C)}$ , respectively. Thus, both operation modes are stable under the specified conditions.

If the setpoint is switched to  $\frac{\bar{V}}{F} = 4 \text{ h}$  at same  $T^\circ = 200^\circ\text{C}$ , then according to Fig. 8.1(b),  $\bar{\xi} = 0.575$ . The same judgment leads to  $T_{ag}^\circ \leq 146.1^\circ\text{C}$ . Estimation of  $\left(\frac{d\xi}{dT^\circ}\right)$  is calculated with respect to  $\bar{T}^\circ = 100^\circ\text{C}$  and  $\bar{\xi} = 0.2$ :

$$\left(\frac{d\xi}{dT^\circ}\right)_{steady} = \frac{0.575 - 0.342}{200 - 100} \frac{1}{^\circ\text{C}} = 0.0023 \frac{1}{^\circ\text{C}}$$



and 
$$\left(\frac{d\xi}{dT^\circ}\right)_{dynamic} = \frac{1}{1-0.575} \left(0.0023 \frac{1}{^\circ\text{C}}\right) = 0.0055 \frac{1}{^\circ\text{C}}$$

Again, both dynamic and steady-state are stable.

Yet, if the reaction heat increases eightfold ( $\Delta H_r = -120 \text{ kJ/kmol}$ ), then at  $200^\circ\text{C}$ , the first setpoint  $\bar{V}/\bar{F} = 2 \text{ h}$ ;  $\bar{\xi} = 0.4$  is stable under steady-state operation but unstable under dynamic conditions. The second setpoint  $\bar{V}/\bar{F} = 4 \text{ h}$ ;  $\bar{\xi} = 0.575$  is unstable in both operation modes. This example proves the importance of thermal stability check calculus during reactor design.

### **i** Example 8.2

Let us consider a process for which the time constant and amplification factor are  $T_{rm} = -2.5 \text{ h}$  and  $K_{rm} = -4$ , respectively [8]. The dead times of equipment [7] add up to  $\tau_e = 0.1 \text{ h}$ . The goal is to determine the parameters of a regulator able to thermally stabilize the unit. The total phase shift is written by applying (8.20) and (8.22):

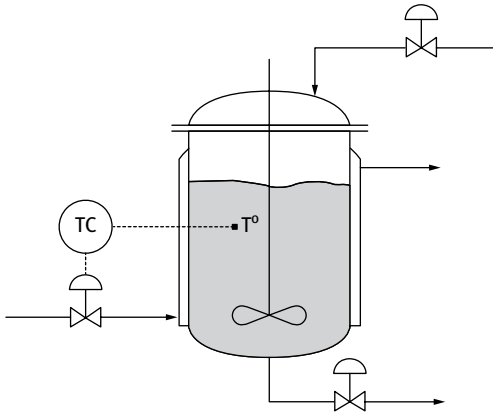
$$\varphi_{total} = -180^\circ = -180^\circ + \arctg |T_{rm}| - 57.3 t_e \omega$$

It yields  $\omega = 15.5 \frac{rad}{h} = 2.46 \frac{1}{h}$  and  $\left(\frac{K_{rm}}{\sqrt{1 + T_{rm}^2 \omega_{osc}^2}}\right) = -0.103$ . The latter corresponds to the limit value of the proportional band of 64%. Yet, the real *PB* has to obey constraint (8.21) and is therefore 32%. The controller gain [7] is  $K_C = \frac{100}{PB} = 1.56$ .

### 8.3.3 Temperature control in practice – continuous units

Thermal stability can be accomplished in many ways. One option is to ensure a difference of approximately  $30^\circ\text{C}$  between the coolant and the bulk reaction mixture to satisfy condition (8.7). Depending on the desired temperature or on whether the employment of a single media is either insufficient or uneconomical, the adequate heat transfer agent can be mono- or multiple-media: for example either water or a mixture of steam and water (also see section 8.3.4, Fig. 8.10).

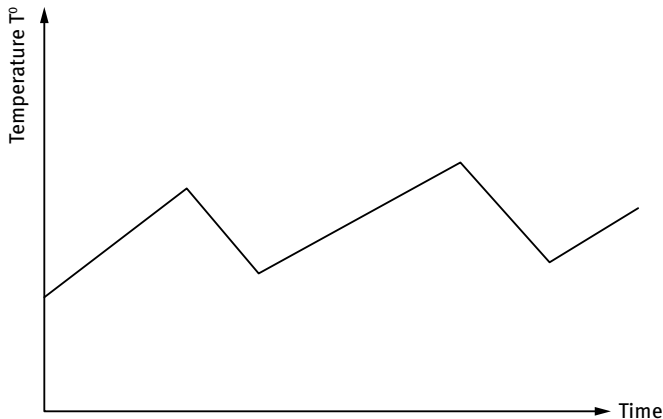
Figure 8.4 depicts a simple control scheme based on manipulating the inlet flow rate of the reactor jacket as a function of controlled reaction mixture temperature; for continuous units the jacket inlet/outlet valve is open (reactor influent and effluent valves are also continuously open). Its disadvantages derive from the possible changes in coolant flow rate and temperature, leading to variations of the jacket dead-time. This influences the controller parameters by means of eqs. (8.18)–(8.19) and (8.20)–(8.22).



**Fig. 8.4:** Temperature regulation for a CSTR by manipulating coolant inlet flow rate. (Adapted after [1, 8])

Another handy option is to add to the reaction mixture some inert ingredient that boils at the desired (setpoint) temperature. This way, the process will be kept almost isotherm whereas production costs might increase because of bigger unit volumes, product separation, inert purification and recirculation, etc. A gas phase example is the hydrocracking unit, where temperature can be controlled by manipulating the hydrogen feed rate because it both dilutes the reactants and cools the mixture (also see Fig. 8.9).

When the equivalent dead time  $\tau_e$  of the controller does not exceed 33%–35% of the reactor's time constant  $T_e$  [7–9], the process can be thermally stabilized, whereas if  $\tau_e/T_e > 0.35$ , the ability of a single control loop to regulate the temperature in the unit



**Fig. 8.5:** Regulated temperature profile (saw teeth) for an unstable reactor. Qualitative representation. (Adapted after [8])

decreases. If it approaches 1, the reactor is uncontrollable. As a result, the temperature oscillates and its profile will resemble the shape of saw teeth as depicted in Fig. 8.5.

In such cases one suitable solution is a *cascade control* (see also chapter 1). This includes *master* and *slave* loops. Because the controlled variable (reaction core temperature) has a slow response to changes in the manipulated variable (coolant flow), it is allowed to adjust the setpoint of a secondary loop whose response to coolant changes is rapid. Thus, the master reaction temperature controller varies the setpoint of the slave jacket temperature loop. The addition of the second loop reduces the oscillation period of the main one; in other words, it corrects for all outside disturbances without allowing them to affect the reaction temperature (compensates the disadvantages of control loop presented in Fig. 8.4). An example of such a configuration is depicted in Fig. 8.6.

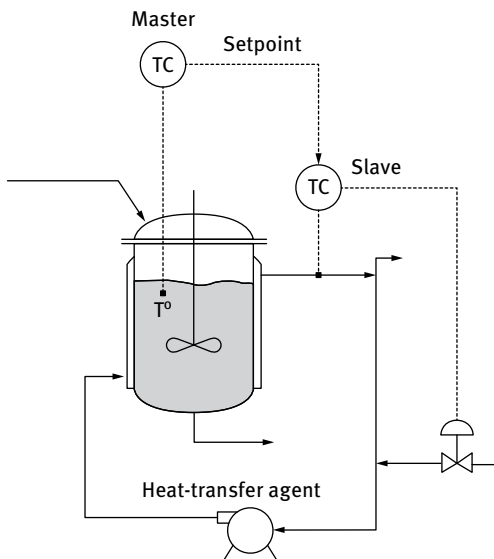


Fig. 8.6: Example of cascade temperature control. (Adapted after [1])

Cascades are stable and function adequately only if the slave is faster than the master loop; the time constant of the slave should not exceed 10% of that of the master [1–2, 9]. Otherwise, it is not able to respond in time to the signal variation detected by the master and the overall loop performance of the cascade will decay.

The slave control loop regulates either the *jacket outlet* or *inlet* temperature. The first option is usually preferred because the jacket time constant and dynamic response, as well as the nonlinearity between jacket-outlet temperature and heat-transfer agent, will be included in the slave loop. As a result, the main nonlinear term within the master loop is removed and the reaction temperature will depend linearly on the

jacket's outlet. If the master's loop oscillation period is cut in half when connected to a slave, the *PB* might change from 30% to 15% and the derivative and integral settings of an interacting controller could also be diminished to the half. This would mean a fourfold overall loop performance improvement [1–2, 9].

Chemical industry often involves processes for which the jacket temperature must be limited and tightly controlled. Such cases are crystallization units or safety systems for thermal shock protection of glass-lined synthesis units (because thermal expansion coefficients for lining and reactor material differ). Thus, the slave control loop is set to regulate the jacket *inlet* temperature. The disadvantage is that the non-linear dynamics of the jacket will be included in the master loop.

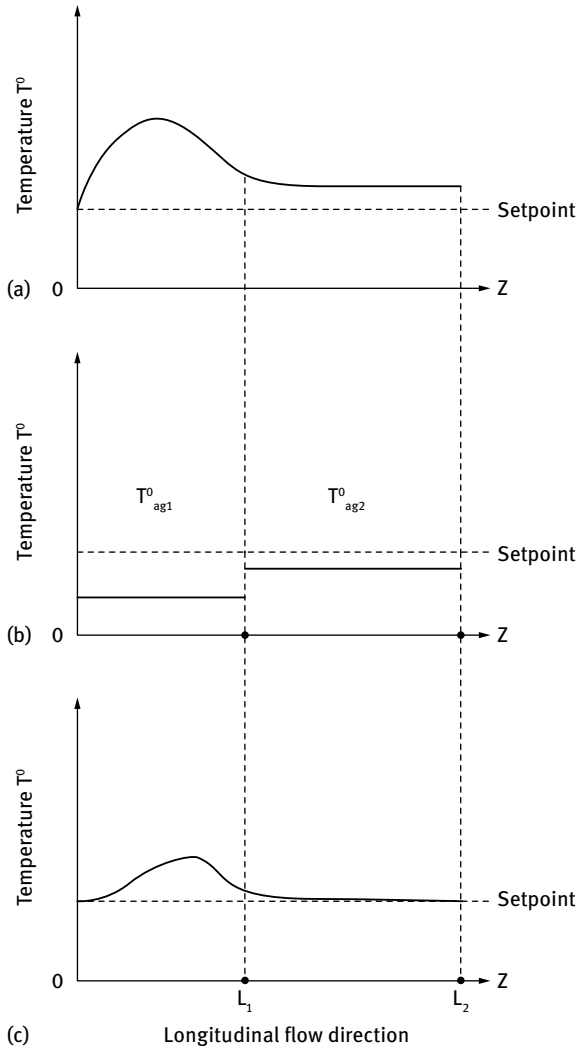
*Model-based controllers* for temperature are recommended when either reaction mechanisms are complex (because of autocatalysis or oscillations) or the process is highly exothermic (danger of runaway behavior). The *predictor* block of such a controller estimates for instance the heat release rate, and feeds forward this value to the slave loop to bias its setpoint. The total effect will be corrected by the *corrector* block and applied on the jacket temperature setpoint. The disadvantage is related to the complexity and accuracy of the model because it has to contain thermodynamic, kinetic, mass, energy, and momentum transfer models of both chemical processes and employed equipment.

According to various descriptions above, a cascade usually controls the reaction temperature by regulating the coolant temperature to match the setpoint. However, in fast and stable processes, a reversed configuration could perform better [9].

The performance of a conventional temperature control loop must be enhanced for chemical processes in which reaction temperature is correlated with reactor pressure. A good example is a reaction for which pressure is due mainly to the vapor pressure of a reactant or a product. In such cases, pressure changes represent almost instantaneous responses to temperature changes or *vice versa* (see Example 8.10). Another example is a gas phase reaction occurring at both high temperature and pressure (oxidation, hydrogenation, etc.) or a high-pressure polymerization. In such cases, partial pressures and temperature are inter-connected within the rate law as well. For these kinds of processes (both temperature and pressure sensitive), a better temperature regulation is carried out using *pressure-compensated temperature control* loops [10].

In tubular plug flow reactors the parameters are distributed on the longitudinal flow direction; residence time, conversion and temperature vary as reaction mixtures advance on the length  $L$  of the tube. Fig. 8.7 shows qualitative temperature profiles vs axial position  $OZ$  for *adiabatic* and *isotherm* chemical processes in profiles (a) and (c) respectively. Profile (b) describes the heat transfer agent that ensures *quasi-isotherm* conditions of (c). The terms “adiabatic” and “isotherm” refer to reactor operation mode, whereas the process itself is exothermic.

To satisfy the demands of an isotherm operation mode, the energy balance in eq. (8.15) is written for infinitely small volumes  $dV$  which correspond to  $dZ$  advances of the plug flow and have  $dA_{T,Z}$  heat transfer surfaces. The reaction mixture enters

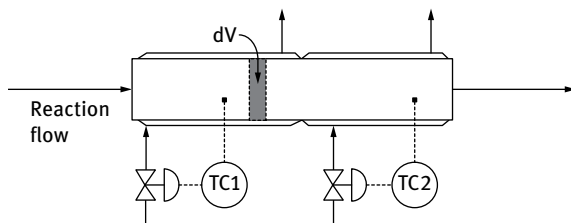


**Fig. 8.7:** Temperature profiles vs axial position  $OZ$  for (a) adiabatic process; (b) quasi-isotherm process; (c) heat transfer agent ensuring conditions in (b). Qualitative representation. (Adapted after [8])

these micro units having  $T_{in}^{\circ}$ ,  $Z$  temperature. It operates and leaves at  $T_Z^{\circ}$ , by reaching conversion  $\xi_Z$ .

$$F \rho C_p (T_{in,Z}^{\circ} - T_Z^{\circ}) - (\Delta H_r)_Z F \xi_Z C_{Ao} = K_T dA_{T,Z} (T_Z^{\circ} - T_{ag}^{\circ}) \quad (8.23)$$

Integration between process imposed limits (tube length  $L$  and  $\frac{V}{F}$  residence time) yields  $T_{opt}^{\circ} = f(Z)$  profiles like those in Fig. 8.7 (that is optimum  $T_Z^{\circ}$  values at various reactor  $Z$  coordinates). From (8.23), values of  $T_{ag}^{\circ}$  are calculated for a desired number of



**Fig. 8.8:** Example of temperature control with separately operating jackets on different portions of a PFR. (Adapted after [8])

tube portions. As depicted by Fig. 8.8, a quasi-isotherm operation mode such as that in Fig. 8.7(c) requires two distinctly operating jackets, each at another  $T_{ag}^{\circ}$  (see Fig. 8.7(b)).

Heterogeneous multi-bed reactors are often modeled as PFR reactors. Practical ways to accomplish desired temperature profiles are *interstage heat exchange* (see Fig. 4.2) [11] or *cold stream injection* (see Fig. 8.9) [1–2, 8, 11].

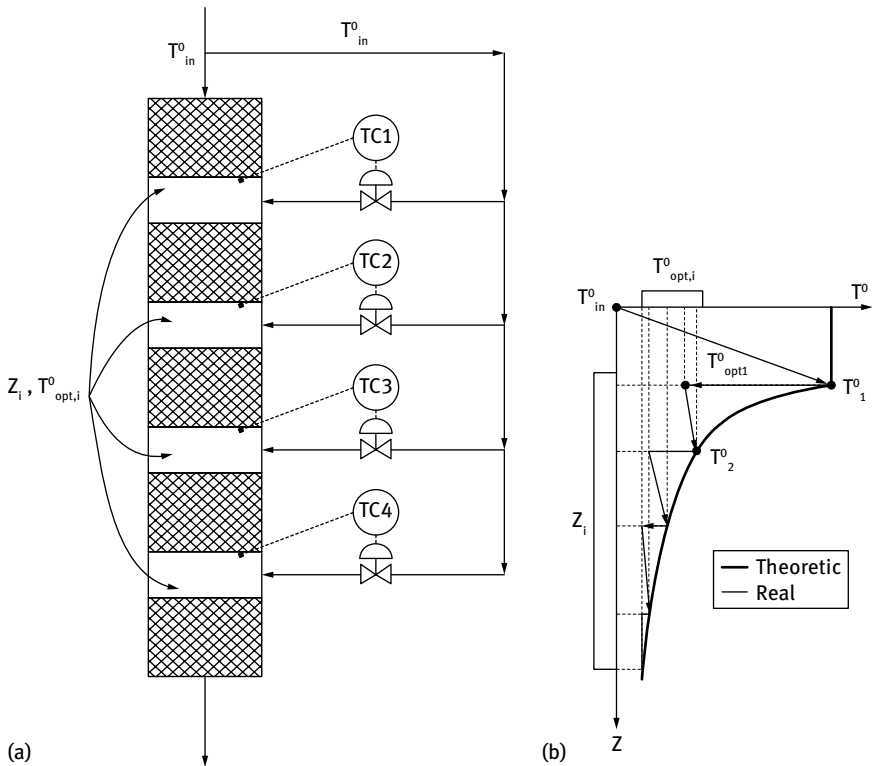
Interstage cooling or heating does not dilute the flow, hence does not require additional beds. Cooling is used for exothermic (sulfur dioxide oxidation, some hydrodealkylation reactions), whereas heating for endothermic equilibrium reactions (dehydrogenation of ethylbenzene to styrene).

Inter-bed injection may be carried out using either a cold gas/liquid phase inert or a cooled reaction mixture. The first option dilutes the main stream and the depth of the bed has to be adapted accordingly to maintain the same conversion. The second option lowers the conversion and additional beds are required to reach its desired value at reactor outlet.

In the industrial ammonia and methanol synthesis (see Example 8.9), hydrogenation of benzene or for some hydrocracking reactions, a fresh and cooled gas is injected between the individual catalyst beds. Its flow rate is designed to satisfy the energy balance in the injection point. Fig. 8.9(a) illustrates such a control strategy. The disadvantage of a costly multiloop configuration is compensated by the possibility of ensuring the *optimum temperature profile* on the longitude of the reactor, a profile that may be neither isotherm nor adiabatic. By using the optimizing principle of Pontryagin [12–13], the setpoint of each individual control loop in Fig. 8.9(a) is chosen to regulate the injected flow in a manner that ensures certain optimum  $T_{opt}^{\circ}$  for the main gas flow after each catalyst bed. In other words, effluent temperature  $T_1^{\circ}$  after the first bed is combined with a fresh and cool stream ( $T_{in}^{\circ}$ ) in order to yield the influent  $T_{opt1}^{\circ}$  in the second bed (see also Fig. 8.9(b)), and so on.

The individual values  $T_{opt1}^{\circ}$  (after the first bed),  $T_{opt2}^{\circ}$  (after the second bed), etc., should ideally lay on the optimum profile  $T_{opt}^{\circ} = f(Z) = f\left(\frac{V}{F}\right)_Z$  depicted in Fig. 8.9(b).

The points  $Z_i$  along the axis of the PFR correspond to various residence time values. The real, (achieved) regulated temperature profile is not as smooth as the theoretical one (see



**Fig. 8.9:** Heterogeneous multi-bed reactor – temperature control strategy by interstage injection of cooled reaction mixture. (a) configuration of control loops; (b) optimal temperature profile on reactor length. Qualitative representation. (Adapted after [1, 8, 11])

Fig. 8.9(b)); it has a saw-teeth shape because temperature cannot be controlled within the catalyst beds. However, it is obvious that it follows the evolution of the theoretical curve.

A  $T^0_{opt} = f(Z)$  shape like that in Fig. 8.9(b) favors reversible exothermic reactions. With the advance of the plug shaped reaction mixture, the released heat constrains the process into regeneration of reactants. Hence, final conversion at reactor outlet drops. Meanwhile, if temperature is lowered with process advance, the reverse reaction will lose significance and conversion gains of up to 20%–40% can be achieved – see also chapter 2.

Lowering overall costs can be achieved by preheating the reactant flow on the expense of the products. It implies regulating the PFR inlet temperature as a function of the outlet value by means of a three-position valve (see Example 8.9 and Fig. 8.26). This is also a necessary safety measure; evacuation of gases has to be possible when pressure builds up because of overheated reactor outlet flows.

If the chemical process imposes extra constraints on temperature regulation and large heat transfer surfaces are required, a multitubular reactor [11] could be a suitable

solution. Yet, it limits the throughput because of smaller available cross sections for the fluid flow.

### 8.3.4 Temperature control in practice – batch units

Batch units operate in *cycles* that start with charging the solvent and reactants into the reactor then mixing and (usually) heating up to reaction temperature. Initiation of the chemical process can occur during heating (reaction rates speed up to a significant value due to temperature increase) or after the addition of a catalyst or an initiator (reaction mechanism switches to a less energy demanding path and occurs with measurable rate). During the course of the reaction, the desired temperature profile can be kept linear or nonlinear. After reaction completion or achievement of required conversion (process *end-point*), the unit is, as a rule, cooled off and emptied. The time amount necessary to complete one cycle ( $t_{load}$ ) is obtained by adding to the actual reaction time the durations of all other necessary operations (heat-up, cool-off, reactor emptying and cleaning, etc.).

For such a succession of operations, the process cannot be described by a steady-state, neither can the temperature controller be tuned to a single setpoint. An adequate control system should permit rapid heat-up to the reaction initializing temperature by meanwhile minimizing its overshoot. It also should adjust rapidly and accurately the load temperature to the desired (optimal) profile.

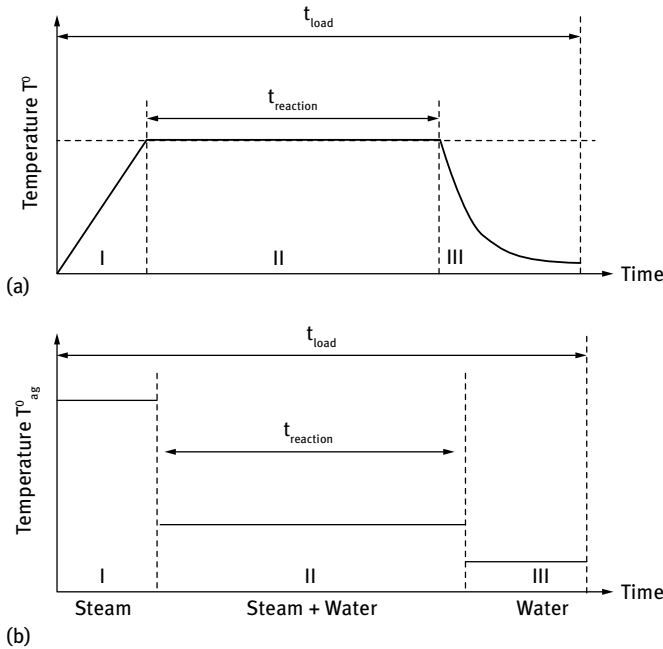
This is a complex task since the dynamics of the batch unit varies in time. Thus, process variables as well as gains and all time constants vary during reaction time. In the case of exothermic chemical reactions, the positive feedback of the process has also to be compensated by suitable self-regulating cooling controllers. In addition, batch-to-batch uniformity must also be fulfilled. The latter depends on many factors, starting with the purity of chemicals and ending with the repeatability of measurements and control equipment.

An example of synchronized temperature vs reaction time  $T^\circ = f(t_{reaction})$  profiles for both the reaction mixture (a) and the corresponding heat transfer medium (b) are illustrated in Fig. 8.10. The following operating modes can be distinguished within one batch cycle: I – heat-up, II – maintain/chemical reaction mode, and III – cool-down.

The control loop depicted in Fig. 8.4 regulates load temperature also in a BTSR by manipulating the coolant inlet valve. For batch units the jacket outlet valve could occasionally be closed during the duration of a complete cycle, whereas the other valves are operated as follows: during unit load the effluent valve is closed and during unit discharge the influent valve is usually closed. The temperature of the heat transfer agent can be adjusted by mixing for example steam and water inlet flows in various ratios (see Fig. 8.10(b)).

A three-mode conventional PID master temperature controller could be employed by tuning it to the setpoint of the reaction mode II of a batch cycle (see Fig. 8.10(a)).



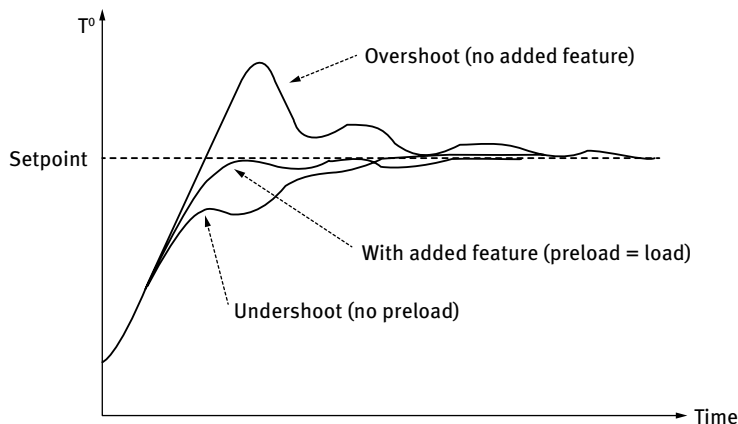


**Fig. 8.10:** Synchronized temperature vs reaction time profile for: (a) reaction mixture; (b) corresponding heat transfer medium. Operating modes within a batch cycle: I – heat-up, II – maintain/chemical reaction mode, III – cool-down. Qualitative representation. (Adapted after [8])

A regulator with a  $PB$  of 30% and a 5-min setting for both integral and derivative times would generate a significant overshoot if kept on automatic during heat-up. Therefore, regular PID controllers require some added features that minimize this inconvenient. Depending on the required proportional band the addition of a so-called *batch unit* feature (for  $PB < 50\%$ ) (see Fig. 8.11), or a *dual-mode unit* feature (for  $PB > 50\%$ ) to the controller, will ensure smarter start-up characteristics [1–2, 7–9].

The common goal during heat-up is to raise the temperature within the unit from its initial to the elevated setpoint value as quickly as possible, but by avoiding *overshoot* and oscillations. The integral term of a conventional PID regulator will cause overshoot, because it stays at the high output limit until the setpoint is reached – see top curve in Fig. 8.11. If this integral term is reduced to match the expected load, then the controller output is imposed to hold the variable at setpoint in the new steady-state. Thus, overshoot is eliminated – see middle curve in Fig. 8.11. Without preloading, the controller will *undershoot* – see bottom curve in Fig. 8.11. If derivative is not used, integration commences as soon as the output leaves its limit before crossing the setpoint. Hence, the preload setting has to be reduced somewhat below the expected load.

A dual-mode temperature controller also minimizes overshoot in a rapid heat-up. It couples an *on/off switch* controller and a preloaded PID controller. The on/off



**Fig. 8.11:** Response of a preloaded “batch unit” PID temperature controller during heat-up. Qualitative representation. (Adapted after [1, 2])

switch initiates either heating or cooling till a close value of the setpoint is reached. After this time-delay, the PID takes over and controls within a narrow temperature range. Such control is successfully applied to polystyrene synthesis reactors.

While each batch cycle benefits from well-performing heat-up control, employment of either of the above described added feature PID regulators has some drawback; settings need to be adapted from one batch to the other (even for the same unit) since heat transfer coefficients usually worsen. Yet, a preloaded PID controller having an external reset feedback from a secondary temperature is much more robust to disturbances [1, 2].

Chemical processes that are endangered to runaway behavior can be protected supplementary by setting a permissible rate of temperature rise during heat-up. In this case, the setpoint of the master loop depends on the heat-up rate. As the reaction mode temperature II is approached near the end of the heat-up mode I (see Fig. 8.10(a)), the value of the permissible rise rate is reset to lower values than at the beginning of the cycle. A model-based heat-up controller for example, would use in this case the heat input rate as the variable manipulated by the predictor block. Such an approach prevents building up of thermal inertia and keeps the process out of the potentially unsafe temperature domains.

A smart temperature control strategy ensures an *optimal* (linear or nonlinear) load temperature during reaction course. Hence, the end-point is reached in minimum of reaction time and the product distribution (described by the values of conversion and selectivity) is favorable. The optimum  $T^\circ = f(t_{\text{reaction}})$  profile is to be obtained by solving eq. (8.23), since the reaction rate  $r$  depends on both temperature  $T^\circ$  and conversion  $\xi$  and the latter on reaction time  $t_{\text{reaction}}$  (see Example 8.3).

$$\frac{dr}{dT^\circ} = 0 \quad (8.23)$$

A *nonlinear* setting is desirable for exothermic reversible (equilibrium) reactions (see also section 8.3.3 and Fig. 8.9(b)). At process beginning, the accumulation of heat accelerates both reactions but favors mostly the direct step (because of high reactant concentrations), whereas when the products accumulate, the rate of the reverse reaction gains significance. Moreover, temperature elevation will also push the mass balance towards the reactants. Therefore, a decreasing temperature profile during the course of the reaction is an adequate strategy.

Meanwhile, irreversible processes that are only slightly exothermic could benefit from an *adiabatic* operation mode. Accumulated heat and resulting temperature increase can compensate for the decay of the reaction rate due to the consumption of reactant species, but only as long as thermal safety requirements are met. As a result, reaction end-point can be reached in smaller volume units.

Jacket and heat transfer media temperature regulation can also benefit from non-linear settings for runaway reactions. Cooling must be very effective at the beginning of the reaction mode II, but as the process advances, the reaction rates decrease and hence the process generated heat rate also decays. Thus, cooling rates have to be adjusted accordingly.

Let us consider the example of a reversible reaction  $A \rightleftharpoons P$  of first-order in both directions and occurring in liquid phase. Its rate is given by eq. (8.24). Terms  $k_1$  and  $k_2$  (both expressed in  $s^{-1}$ ) stand for the first-order rate coefficients of the direct and reverse reactions, respectively, whereas  $C_A$  and  $C_P$  for the momentary molar concentration ( $\text{kmol/m}^3$ ) of reactant  $A$  and product  $P$ , respectively.

$$r = k_1 C_A - k_2 C_P = k_{01} \exp\left(-\frac{E_{a1}}{RT^{\circ}}\right) C_A - k_{02} \exp\left(-\frac{E_{a2}}{RT^{\circ}}\right) C_P \quad (8.24)$$

If the initial concentration of the product is zero ( $C_{P_0} = 0$ ), then

$$C_A = C_{A_0} (1 - \xi); \quad C_P = C_{A_0} \xi \quad (8.25)$$

and eq. (8.24) can be written as a function of  $\xi$  as follows:

$$r = k_{01} \exp\left(-\frac{E_{a1}}{RT^{\circ}}\right) C_{A_0} (1 - \xi) - k_{02} \exp\left(-\frac{E_{a2}}{RT^{\circ}}\right) C_{A_0} \xi \quad (8.26)$$

According to eq. (8.23), it further becomes

$$\frac{dr}{dT^{\circ}} = \frac{C_{A_0}}{R(T^{\circ})^2} \left[ (1 - \xi) k_{01} E_{a1} \exp\left(-\frac{E_{a1}}{RT^{\circ}}\right) - \xi k_{02} E_{a2} \exp\left(-\frac{E_{a2}}{RT^{\circ}}\right) \right] = 0 \quad (8.27)$$

and the optimum temperature value  $T^{\circ}_{opt}$  in relationship (8.29) can be calculated from it. Because the equilibrium composition is related to the rate coefficient ratio  $k_1/k_2$  by means of equilibrium constant  $K$  (see (8.28), where  $\xi_{eq}$  stands for the equilibrium value of  $\xi$ ),  $T^{\circ}_{opt}$  is controlled by both kinetic and thermodynamic data [14].

$$K = \frac{k_1}{k_2} = \frac{(C_A)_{eq}}{(C_B)_{eq}} = \frac{\xi_{eq}}{1 - \xi_{eq}} \quad (8.28)$$

$$T_{opt}^{\circ} = \frac{E_{a1} - E_{a2}}{R \ln \left[ \left( \frac{E_{a1} k_{01}}{E_{a2} k_{02}} \right) \left( \frac{1 - \xi}{\xi} \right) \right]} \quad (8.29)$$

The resulting  $T_{opt}^{\circ} = f(\xi)$  in eq. (8.29) and its correlation with  $\xi = f(t_{reaction})$  will generate the optimal  $T_{opt}^{\circ} = f(t_{reaction})$  profile (see Fig. 8.12(a) and Example 8.3).

Conversion is always defined with respect to the limiting reactant – see eq. (8.1). So that, if there are two or more reactants involved, concentration of species that affect rate should take into account the *excess ratio*  $\gamma$  of other species over the limiting one.

For the reversible reaction  $A + B \rightleftharpoons P$  of second-order on the direct reaction path and of first on the reverse,  $\gamma$  is defined as

$$\gamma = \frac{C_{Bo}}{C_{Ao}} \quad (8.30)$$

if  $A$  is the limiting and  $B$  is the *excess reactant* [5]. According to this definition,  $\gamma > 1$ . A value  $\gamma = 1$  means that reactants are mixed in stoichiometric ratios.  $T_{opt}^{\circ}$  – see result in eq. (8.33) – can be calculated by using reaction rate (8.32),  $C_A$  and  $C_P$  from eq. (8.25) together with the expression of  $C_B$  in eq. (8.31).

$$C_B = C_{Ao} (\gamma - \xi) \quad (8.31)$$

$$r = k_1 C_A C_B - k_2 C_P = k_1 (C_{Ao})^2 (1 - \xi) (\gamma - \xi) - k_2 C_{Ao} \xi \quad (8.32)$$

$$T_{opt}^{\circ} = \frac{E_{a1} - E_{a2}}{R \ln \left[ \left( \frac{E_{a1} k_{01}}{E_{a2} k_{02}} \right) \left( \frac{C_{Ao} (1 - \xi) (\gamma - \xi)}{\xi} \right) \right]} \quad (8.33)$$

Terms  $k_1$  and  $k_2$  in eqs. (8.32)–(8.33) signify rate coefficients of the direct and reverse reactions:  $k_1$  is of second ( $\text{m}^3/\text{kmol}\cdot\text{s}$ ) and  $k_2$  of first order ( $\text{s}^{-1}$ ), respectively.

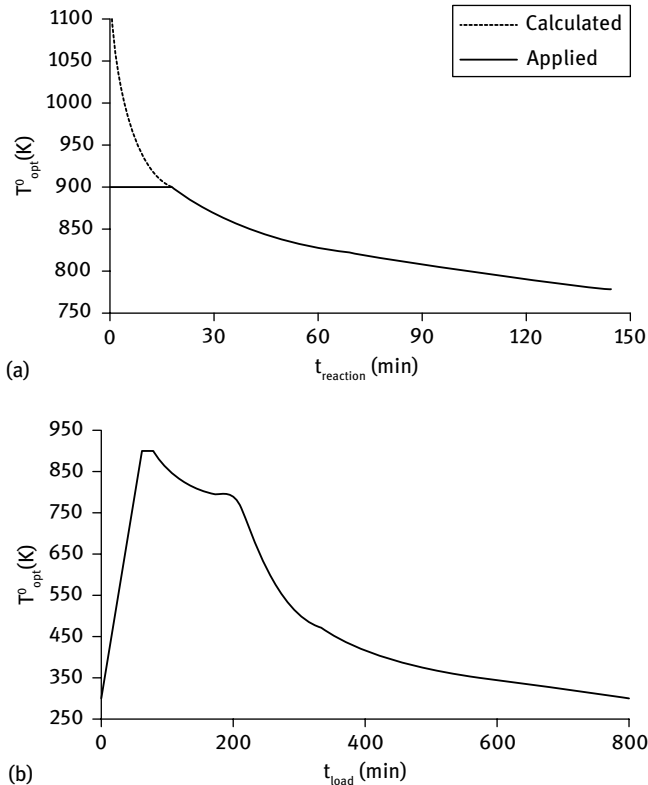
### Example 8.3

The desired conversion for the above discussed first-order reversible reaction  $A \rightleftharpoons P$  is  $\xi = 0.8$ . The maximum permitted temperature by the equipment is  $T_{max}^{\circ} = 900$  K. Kinetic data are  $k_{01} = 30 \text{ min}^{-1}$ ,  $k_{02} = 10^6 \text{ min}^{-1}$ ,  $E_{a1}/R = 6000$  K, and  $E_{a2}/R = 16000$  K. Heat-up starts from 300 K and occurs with a rate of 10 K/min.

Advancing  $\xi$  values from 0 to 0.8 yield  $T_{opt}^{\circ}$  values by relationship (8.29). These result in  $k_1$  and  $k_2$ , which further lead to  $\xi_{eq}$  and  $t_{reaction}$  [5], respectively – see eqs. (8.34)–(8.35). At each temperature,  $\xi_{eq}$  is calculated and the reversibility requirement  $\xi \leq \xi_{eq} \leq 1$  has to be fulfilled.

$$\xi_{eq} = \frac{k_1}{k_1 + k_2} \quad (8.34)$$

**i**



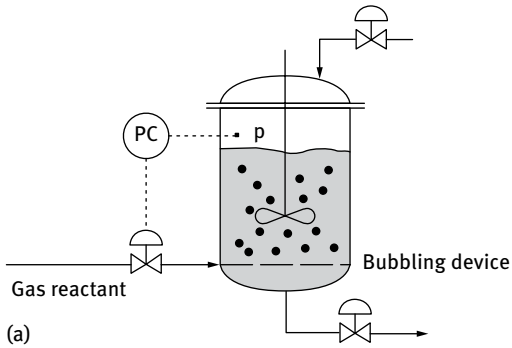
**Fig. 8.12:** Optimum temperature vs time profiles for the batch unit in Example 8.3: (a) duration of chemical process; (b) duration of a batch cycle (I – heat-up mode, II – chemical reaction mode, III – cool-down mode).

$$t_{reaction} = \frac{1}{k_1 + k_2} \ln \left( \frac{\xi_{eq}}{\xi_{eq} - \xi} \right) \quad (8.35)$$

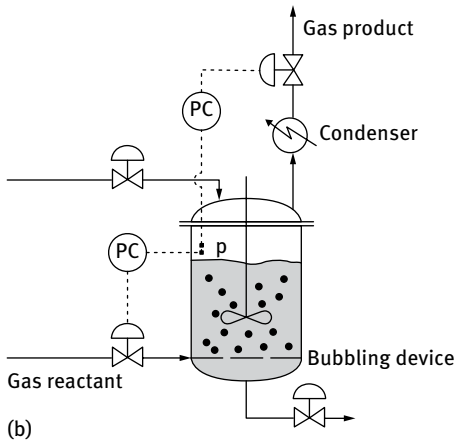
The results, correlated into a nonlinear  $T_{opt}^{\circ} = f(t_{reaction})$  profile for the duration of the chemical process (operation mode II in Fig. 8.10(a)), are presented in Fig. 8.12(a). The dotted line stays for the calculated (theoretical) optimum temperature values. Since the unit is limited to  $T_{max}^{\circ}$  of 900 K, the setpoint temperature profile corresponding to the continuous line settles at exactly this value in the region for which computed  $T_{opt}^{\circ}$  exceeds  $T_{max}^{\circ}$ . Figure 8.12(b) illustrates the entire temperature profile of a cycle for this batch unit. It may be observed that the linear setting of chemical operation mode II in Fig. 8.10(a) is replaced by an optimized nonlinear one.

## 8.4 Pressure control

Chapter 8.3.3 mentions some examples of chemical processes that are, besides temperature, also pressure sensitive. Figure 8.13 presents examples of pressure control loops



(a)



(b)

**Fig. 8.13:** Pressure control in batch units: (a) gas is reactant; (b) gas is product or both reactant and product. (Adapted after [1])

for batch units where a gas phase species is either (a) the reactant (chemical reaction is preceded by absorption of gas into liquid) or (b) the reaction product (for example carbon dioxide) or both reactant and product, respectively. In both cases, the bulk liquid phase concentration  $C$  ( $\text{kmol}/\text{m}^3$ ) of the species in question is related to its partial pressure in the vapor space  $p_{\text{gas}}$  (atm) by Henry's law [6, 14]. In relationship (8.36),  $k_{\text{Henry}}$  ( $\text{kmol}/\text{m}^3 \cdot \text{atm}$ ) is a temperature-dependent coefficient.

$$C = k_{\text{Henry}} p_{\text{gas}} \quad (8.36)$$

Hence, pressure changes are linked to reaction advance (by means of conversion  $\xi$ ) as well as to temperature changes.

Both pressure regulators depicted in Fig. 8.13 are fast and easily controlled. Depending on the measured vapor phase total pressure, they manipulate either an inlet or an outlet gas flow by means of a valve. A gas vent is usually employed for the loop in Fig. 8.13(b). Moreover, a condenser is often placed in front of the vent to minimize liquid product loss. If the chemical process occurs with both gas consumption as well as generation, the gas inlet pipe in Fig. 8.13(b) may benefit from a flow controller also manipulated by reactor pressure. This configuration is part of the necessary safety strategy (it may shut down gas inlet) for the undesired situation of pressure building up in the unit and the outlet vent being not able to properly reduce it.

In continuous units operating both on liquid and gas streams, placement of at least one of the flow controllers on ratio (see chapter 1) is useful because it ensures the desired setpoint initial mixture of reactants.

For processes requiring vacuum, its source is often a steam jet-type ejector. Such constant-capacity sources are commonly matched to the variable-capacity reactor by wasting the excess capacity of the ejector [1]. A special situation is that of vacuum stripping in polymerization units: solvent and unreacted monomers have to be removed as rapidly as possible without causing the reaction mass to foam. The load temperature depends on the balance between heating (due to steam inlet) and cooling (due to monomer vaporization). The vacuum level is tightly linked to temperature, therefore also controlled by it.

## 8.5 Liquid level control

Liquid level in an operating unit is an indicator of the reaction mass volume. Tank reactors are filled only up to 70%–80% of their total volume. This fraction depends on process characteristics; if reaction-caused foaming occurs, then it will be smaller.

Well controlled residence time for a continuous unit translates into a tight  $\frac{V}{F}$  ratio between the reaction mixture volume  $V$  and the reactor outflow  $F$ . It also ensures both desired conversion – see eqs. (8.3)–(8.4) – and product distribution – see (8.2). This is commonly accomplished in practice by a controller that approximates  $V$  through a level measurement and manipulates the outflow valve. Such a loop is depicted in Fig. 8.14(a). It operates similarly in a batch unit – see Fig. 8.14(b) – where the desired volume, thus dilution, conversion and selectivity – see eqs. (8.1)–(8.2) and (8.5) – is regulated by means of inlet flows. When the fluid level in the reactor reaches the setpoint value, inlets are shut-down.

Volume  $V$  (liquid level) in Fig. 8.14(a) is maintained constant by a proportional regulator that modifies the effluent rate  $F$  (controller output) according to:

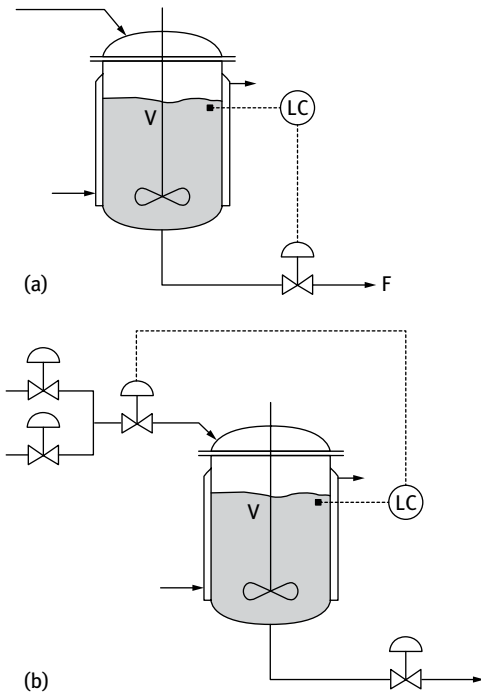
$$F = K_V (V - V_{min}) \quad (8.37)$$

The controlled variable  $V_{min}$  stands for the minimum value of  $V$  of the setpoint. For a batch unit, such as in Fig. 8.14(b),  $V_{min}$  corresponds to  $F = 0$ . The proportional gain  $K_V$  ( $s^{-1}$ ) is a coefficient depending on both regulator and valve characteristics.

## 8.6 pH control

### 8.6.1 pH and titration curves

The pH is a numeric scale used in chemistry to specify the acidity or basicity of an aqueous solution. It is defined as the negative decimal logarithm of the *hydrogen ion*



**Fig. 8.14:** Liquid level control for (a) a continuous unit (residence time control); (b) a batch unit (volume control). Simplified scheme for isotherm operation. (Adapted after [8])

*activity*. This differs from the concentration  $[H^+]$  by a factor called *activity coefficient*. The value of the latter depends on the *ionic strength* (load) of the aqueous media and lies within 0 and 1. Yet, for diluted solutions, such as wastewater, it is approximated to 1. Thus, pH can be written as

$$pH = -\log [H^+] \text{ or } [H^+] = 10^{-pH} \quad (8.38)$$

At 25°C, the neutral value of pH is 7, where  $[H^+] = [OH^-] = 10^{-7} N$ . The brackets indicate *Normal* concentrations, designated as  $N$  and expressed in g-ions/dm<sup>3</sup>. At 25°C,

$$[H^+] * [OH^-] = 10^{-14} = K_{Water} \quad (8.39)$$

$K_{Water}$  is the dissociation equilibrium constant of water. It varies between  $10^{-15}$  at 0°C and  $10^{-12}$  at 100°C [6]. When  $K_{Water} = 10^{-14}$ , the *acidic* range is covered by pH of 0–7, whereas the *basic* by 7–14 units, respectively. It is obvious that a tenfold variation of either  $[H^+]$  or  $[OH^-]$  concentrations yields only a 1 unit change in pH. This strong non-linear (logarithmic) relationship – see eq. (8.38) – makes pH so difficult to control.

In practice, pH control is required mainly for neutralization of industrial wastewaters. These contain *strong* or *weak* and *acid* or *basic* species. *Buffering* chemicals may also be present. Among the latter category are carbonates, silicates, phosphates, sulfonates, organic origin citric acid or basic amines, and metal ions such as iron,



copper, lead, chromium, cobalt, nickel, tin, zinc, cadmium, vanadium, manganese [15], and others. Buffers and weak agents, either acid or base, behave alike.

Some of the common strong acids are hydrochloric, nitric, or sulfuric acid. The strong bases include sodium or potassium hydroxide. They dissociate completely into ionic species when dissolved in water, whereas the weak agents dissociate (ionize) only partially and in more than one stage. These elementary steps are characterized by ionization constants:  $K_{acid}$  for acids,  $K_{base}$  for bases. Values of  $pK_{acid}$  and  $pK_{base}$  are commonly listed in databases; these stand for the negative decimal logarithm values of  $K_{acid}$  or  $K_{base}$ , respectively, and represent the pH value that neutralizes half of the amount in an aqueous solution.

When strong hydrochloric acid or sodium hydroxide base is added to the solution, it reaches concentrations  $C_{acid}$  and  $C_{base}$ , respectively. Negative and positive charges must balance [16], so that

$$C_{acid} + [H^+] = C_{base} + [OH^-] \quad (8.40)$$

Coupling eqs. (8.39) and (8.40) yields eq. (8.41), the titration curve equation of these two strong agents [16]. It is plotted in Fig. 8.15(a) in  $pH = f(C_{base} - C_{acid})$  coordinates.

$$C_{base} - C_{acid} = 10^{(pH - 14)} - 10^{-pH} \quad (8.41)$$

If weak acetic acid is neutralized with the same base, eq. (8.41) and its plot modify accordingly (see Fig. 8.15(b)) [16]. Depending on the number of dissociation steps and  $K_{acid}$  values, the ratio's value in eq. (8.42) will change. The same is true for bases.

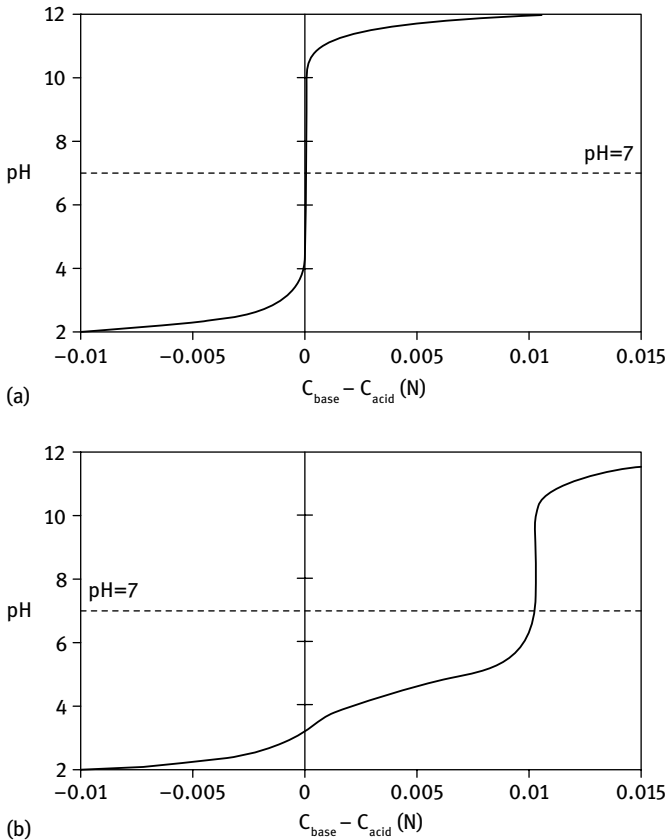
$$C_{base} - C_{acid} = 10^{(pH - 14)} - 10^{-pH} + \frac{C_{Acetic\ acid}}{1 + 10^{(pK_{Acetic\ acid} - pH)}} \quad (8.42)$$

In Fig. 8.15(a), point 0 on the abscissa means that acid and base character of the aqueous solution are in balance ( $C_{base} - C_{acid} = 0$ ) and neutralization is complete. Negative abscise values indicate an acidic, whereas positive a basic character. The reverse titration of sodium hydroxide with hydrochloric acid yields the mirror image of this curve.

The shape of curve in Fig. 8.15.a illustrates clearly the difficulties of a proper pH control. Bringing the pH in the vicinity of neutral 7 translates in crossing the huge gap, on a concentration scale of a few orders of magnitude, between the actual value of controlled variable and the setpoint. Thus, a linear regulator is not suitable.

### **i** Example 8.4

Let us consider an ion exchanger wastewater entering treatment procedure at acidic pH 3. The goal is to raise pH to 6. This requires matching the inlet concentration of  $10^{-3}$  N with an accuracy of  $\pm 10^{-6}$  N; in other words, the amount of caustic reagent added has to match the load in ratios of 1 part to 1000 [16]. Only 1 unit less in the pH of the inlet flow makes the goal more difficult; since precision has to be conserved, the amount to control is in ration of 1 part to 10,000. For the example of increasing pH 3 to 6,  $1\text{ m}^3$  of



**Fig. 8.15:** Titration curve at 25°C: sodium hydroxide neutralizes an aqueous solution of (a)  $10^{-2}$  N hydrochloric acid; (b)  $10^{-2}$  N acetic acid. (Adapted after [16])

$10^{-3}$  N hydrochloric acid is neutralized by  $1 \text{ m}^3$  of  $10^{-3}$  N sodium hydroxide (see titration curve in Fig. 8.15(a)).

Meanwhile, the shape of curve in Fig. 8.15(b) shows that weaker acids or buffers are less sensitive to added reagent amounts; thus, pH control becomes somewhat easier. Here neutralization is complete at  $C_{base} - C_{acid} = 0.01$ , the exact concentration of acetic acid.

A comparison between the curves in Fig. 8.15 leads to the apparent conclusion that any weak or buffering species present besides the strong agent in an aqueous solution to be neutralized, makes the pH control task easier. Yet, real wastewaters contain a wide variety of species in large concentration ranges. Some of these are only partially ionized but all must be neutralized. The problem lies in the fact that the pH analyzer senses all and delivers an overall value, regardless of the species generating it. This again makes pH control difficult since different amounts of reagent have to be added if the same pH is generated by a strong or a weak agent. Yet, the demand for an extremely precise material balance remains.

**i** Example 8.5

Let us consider the same wastewater (pH of 3) as in Example 8.4. If the inlet value is due to acetic acid instead of hydrochloric acid,  $1 \text{ m}^3$  of  $\approx 5.6 \times 10^{-2} \text{ N}$  sodium hydroxide is required to neutralize  $1 \text{ m}^3$  of wastewater (56-fold more concentrated than for  $10^{-3} \text{ N}$  hydrochloric acid also yielding pH 3). A similar calculation shows that a  $10^{-3} \text{ N}$  acetic acid solution has a pH of  $\approx 3.6$ , this is seemingly closer to the neutral pH 7. Yet,  $1 \text{ m}^3$  of it needs  $1 \text{ m}^3$  of base of  $\approx 9.5 \times 10^{-3} \text{ N}$  caustic agent, almost a 10-fold higher amount than expected [16].

It may be concluded that feedforward signals of pH from the inlet flow are not useful in regulating the neutralizing reagent flow because overdosing or underdosing might occur. The only exception could be the situation of known composition inlet flows.

## 8.6.2 pH regulator characteristics

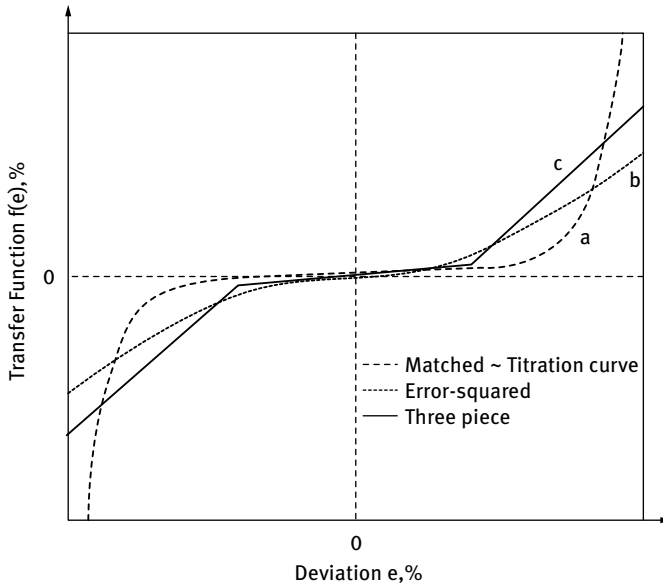
pH sensors have to fulfill constraints described generally for analyzers in section 8.7. The electrical impedance of a widely used glass sensor is in the range of  $10^2 \text{ M}\Omega$ . It contains a galvanic cell with two electrodes: a standard hydrogen electrode opposed to a reference, usually calomel or silver chloride electrode. The electromotive force of the cell depends linearly on the potential  $E_{\text{H}^+/\text{H}_2}$  (V) of the hydrogen ion-selective electrode which also depends linearly on the pH. Ideally, it obeys the Nernst equation [6, 14]:

$$E_{\text{H}^+/\text{H}_2} = E_{\text{H}^+/\text{H}_2}^{\circ} - \frac{2.303 R T^{\circ}}{F_{\text{araday}}} \text{pH} \quad (8.43)$$

In eq. (8.43), the standard electrode potential is  $E_{\text{H}^+/\text{H}_2}^{\circ} = 0$ , and the value of the slope is of 59.1 mV (also called the *Nernstian slope*) when:  $R = 3.314 \text{ J/mol}\cdot\text{K}$ ,  $T^{\circ} = 298 \text{ K}$  and  $F_{\text{araday}} = 96485 \text{ C/mol}$ . It also points out that pH is temperature dependent, so that both calibration as well as measured readings must be temperature compensated.

Because of the unique shape of titration curves, PID regulators employed in pH control need special nonlinear characterizers to linearize the loop. Therefore, there is no universal pH characterizer. However, three main types [16] are commonly in use since rigorous control has to be applied mainly around the setpoint and not over the entire curve (see Fig. 8.16).

One option is to *match* the titration curve (curve a in Fig. 8.16). This might be applicable if the influent to be neutralized is well described in terms of chemical species and their concentrations. However, these situations are rare and a great variety of titration curves overlap into a *real* one, which might even differ for the same influent. For example, municipal wastewater has different composition when household or storm water is looked at. In these cases, a very simple *three-piece* characterizer may be used (curve c in Fig. 8.16). It consists of three lines, each one applicable on a certain pH range. The *error-square* characterizer (curve b in Fig. 8.16) relies on a simple parabolic transfer function of the deviation [16],  $f(e) = e|e|$ . All



**Fig. 8.16:** Most extensively employed pH characterizers as compared to a real titration curve: (a) matched; (b) three-piece; (c) error-square. (Adapted after [16])

are qualitatively illustrated in Fig. 8.16, where the matched curve is considered to superpose the real titration curve. Other characterizers are also feasible, but the above-described types have found the most extensive use.

Among the three, the matched curve provides the best response. However, recovery time from significant load changes is still longer than the oscillation period around the setpoint. Thus, even if matching is performed by some hundreds of points, the control loop will still not behave like a linear one. The three-piece characterizer is a very crude description of the titration curve, but still performs well; it has an adjustable minimum gain, whereas the error-squared curve has a zero gain at zero deviation. Moreover, the domain for which the gain of the three-piece description is valid is also adjustable [16].

Because of the logarithmic feature of titration curves, the PID regulator's proportional band and gain have to be chosen adequately to slightly damp the loop in the vicinity of setpoint. This is to avoid cycling. When cycling occurs for a controller that operates both acid and base agent valves in sequence, the amplitude could fall out of specification range. Hence, the controller will command alternative addition of each reagent even if there is no real influent load change; only the added agents will neutralize each other (waste of chemicals) [1–2]. Limit-cycling may be acceptable, for example, in the case of wastewater where the target specification range is large (between pH 6–9) and cycling amplitude falls within it.

The integral time of the employed PID should be at least twofold higher as compared to the corresponding optimum of a linear loop, yet it is limited to the

window necessary for the pH to recover from a significant change. The fastest part of the cycle (near setpoint) yields the optimum derivative time. It is comparable to that of a corresponding linear loop.

Commercially available *self-tuning PID* controllers are usefully employed when process parameters change frequently and substantially [16]. These are tuned for the worst-case scenario and let self-tuning do the rest: observe loop behavior under upset conditions and estimate settings that improve performance. The drawback is that no distinction is made among pH readings due to load changes, dirty, de-calibrated, or malfunctioning sensors; thus, the PID may readjusts to unreal parameters.

### 8.6.3 Aspects of pH control in practice

According to Hoyle *et al.* [16], neutralization is a process that *must* be designed for controllability. There are many aspects to be considered and only a few, generally valid, recommendations are mentioned here. Preferred continuous units should be stirred tanks with influent entering at the top and effluent evacuating at the bottom. The unit diameter should equal the liquid column height. The residence time should be of at least 5 min and the dead-time maximum 1/20 of it. The mixer has to ensure a vessel turnover of maximum 1 min and a pumping rate of a minimum of 20-fold the throughput flow.

In batch units, a liquid is placed into the vessel and its pH is adjusted to the desired value by adding a certain reagent. Since no flow is continuously entering or leaving the unit, the load is zero. Thus, the controller output bias should be fixed at zero when setpoint pH is reached so that the valve dosing the neutralizing agent closes at setpoint. This involves elimination of integral action and consequently either a simple proportional or a proportional-plus-derivative regulator is recommended. The absence of a proportional controller can be replaced by building into the calculation block a (8.37)-like relationship, where  $K_{pH}$  stands for the proportional gain:

$$\text{Controller output} = K_{pH} (pH_{\text{setpoint}} - pH_{\text{controlled}}) \quad (8.44)$$

In other words, the logarithmic (nonlinear) feature of titration curves is compensated with equal-percentage valves for the neutralizing agents. The valve does not have to match the entire curve because controlled pH reaches the setpoint only once, at the end of the batch cycle when the valve is closed. The dynamic pH profile over a cycle time will show a slight undershoot because the derivative component of the regulator closes the neutralizing reagent valve just shortly before the setpoint pH is reached. Even though such course of actions might slightly increase the duration of a cycle, it is preferred over a permanent overshoot (see also section 8.7.1). Figure 8.17 illustrates such pH control loop.

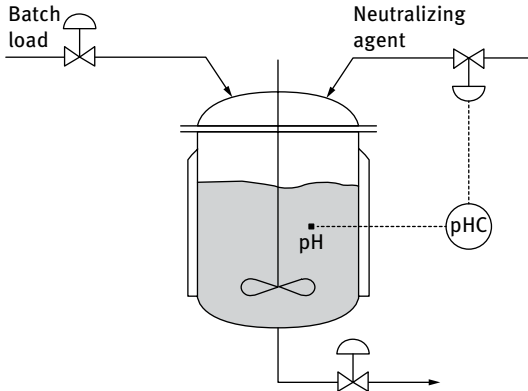


Fig. 8.17: A pH control loop for a batch unit.

Figure 8.18 illustrates a series of three consecutive CSTRs. Such a configuration alleviates the effort of adjusting pH over more than 2 units, which is over more than 100-fold concentration range. Manipulating valves and flows within a precision of  $\pm 1\%$  is thus manageable. In a single neutralizing stage, pH would most probably cycle excessively because of valve imprecision.

Units I, II, and III in Fig. 8.18 cover each approximately two pH units, starting from either strong acidic or basic media in I (pH of 1 or 13) and closing up stepwise (through II with pH of 3 or 11) to the desired pH in III (pH of 6 or 9) [16]. The successive steps units are served by smaller valves, thus each setpoint has to be very well defined. Otherwise, a failure upstream passes unmanageable disturbances (overloads) downstream.

The target specification of the final outlet flow is the setpoint of the last stage, which is responsible for the fine-tuning of pH. Each unit has its own independent pH control loop. Yet, these are still able to pass along undesired cyclings to the next loop. To avoid a cyclic disturbance upstream that matches the exact period of a controller downstream, the loops should have distinct/different oscillation periods.

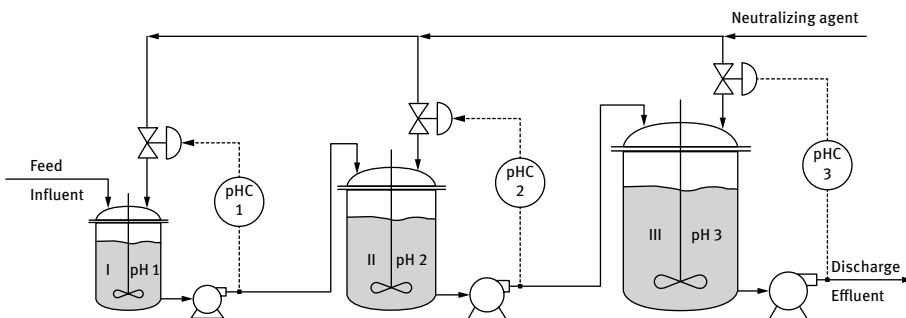


Fig. 8.18: Series of successive continuous units to adjust a wide range pH change. (Adapted after [16])

This is the reason behind sizing the units progressively larger from feed to discharge (from I to III).

Each of the three continuous units in Fig. 8.18 may also operate separately. PIDs with settings as described in section 8.6.2 are recommended; linear PI controllers can cause large overshoots in such units – see section 8.6.1.

## 8.7 End-point detection and product-quality control

Composition assessment by measuring various physical and chemical proprieties of a reaction mixture has many benefits. Analysis procedures inserted in either continuous or batch unit technological flows contribute to:

- Early detection of deviations from setpoint values (ensures desired quality products, avoids loss of chemicals);
- Proper detection of equipment malfunctions (contributes to plant safety, avoids economic loss);
- Correct identification of process end-points (reduces batch cycle time, increases productivity, avoids undesired side reactions).

Even though analytical instruments need adequate maintenance/calibration and often raise reliability questions, their use (if available and economically feasible) is recommended for end-point detection. However, other methods are also commonly in use for completion assessment in batch units. For example, the reactant flow can be terminated based on the total load weight or a decline in heat release is noticed when the limiting reactant is totally consumed. For polymerization reactions, the disappearance of the monomers is accompanied by a pressure drop.

### 8.7.1 Some analyzer types

Analyzers are used to determine some physical or chemical property of the reaction mixture. This has to be relevant for the conversion and vary enough to be detected properly if the system is disturbed. Therefore, instruments should be calibrated in a manner that senses even slight changes in the measured property. Besides pH (see section 8.6), these could be resistivity, conductivity, density, viscosity, absorbance, reflectance, transmittance, etc. [7]. Commercially available automatic on-stream detectors also involve infrared (IR) beam attenuation, gas chromatography (GC), high-performance liquid chromatography (HPLC), or mass spectrometry (MS) and even coupling of some of the mentioned techniques [7].

The sensors can be placed on-stream for continuous units, either in the reactor itself or in the outlet pipe. Reaction vessel placement is chosen for batch units. Many variations are encountered in practice: retractable or fixed, inserted sidewise or immersed from the lid of the vessel, with sample extraction or *in situ* detection, etc.

The analyzers must perform accurately and rapidly since control variables are often in their grasp (for example temperature, pressure, or pH control). Thus, automatic cleaning and rinsing is to be considered.

Some analytical procedures require grab sampling followed by bench analysis. Here sampling and sample preparation difficulties have to be overcome. These are more time consuming hence intervention in the process is delayed.

In continuous units, the measured (controlled) variable has a single setpoint value. However, the analytical property in a batch reactor is away from its setpoint for almost the entire reaction duration. Therefore, integral action must be avoided or else it leads to irrecoverable overshoot (see pH control in section 8.6, or heat-up mode in section 8.3.4). In continuous operation, overshoot is temporary; the outflow will eventually purge away its effects.

### 8.7.2 End-point detection reliability issues

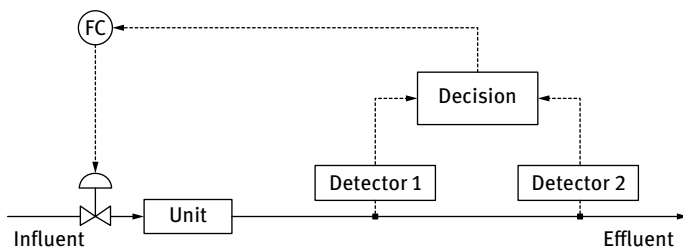
Reliable analyzers are necessary for proper operation of reaction vessels as well as for overall plant safety. Reliable *readings* are conditioned by an adequate placement of the sensor into a representative and uniformly mixed zone of the reaction mass. Malfunction or failure that might drive the unit into an unsafe state has to be avoided, therefore data provided by the analyzers should be *temperature-compensated* and the sensing devices well cleaned, maintained, and calibrated. Deposits of tars, precipitates, or biological growth on the sensor cause variable bias readings; they raise the response dead-time and alter the signal value. In the case of electrodes or pH measurements, a fluid velocity of approximately 2–3 m/s [16] keeps them clean and causes no significant abrasion.

Extra care is to be granted to pH sensors, since their ideal response to 1 unit change of pH away from 7 is of only 59.1 mV [6, 14]. In glass bulb pH sensors, electrolyte fillings of both measuring as well as reference electrodes are buffered to pH 7. Hence, neutral pH yields a zero reading. A dead-short also yields 0 mV indicating a false-neutral value.

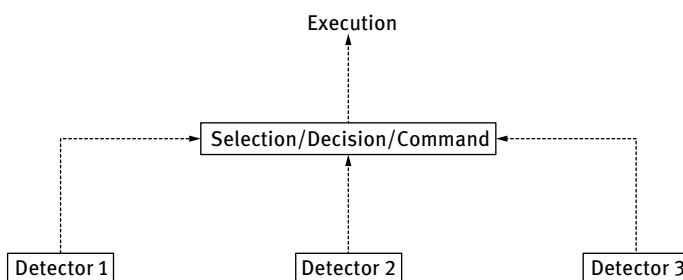
In practice such shortcomings are alleviated by use of a *multiple sensor configuration* (redundant, voting or median selector). The example in Fig. 8.19 illustrates *redundant* readings by two equal performance sensors inserted in the outflow pipe and linked to a high signal selector. If the controlled variable is for example a product concentration, the manipulated variable (inlet flow rate) is protected regardless of which analyzer fails. If detector 1 deteriorates than upstream detector 2 will take over automatically. If this one malfunctions, the selector will command automatic shut-down of the inlet. Thus, the reactor is safe until breakdown is remediated.

If no pauses in the production are permitted, a *voting* configuration (see Figure 8.20) of at least three equivalent sensors is much safer. All are placed in the same media. Reliability is gained as any reading that disagrees with the majority is disregarded by the system. Data offered by the majority is further employed in closed-loop control. Another





**Fig. 8.19:** Example of redundant analyzer configuration for adequate end-point detection and inlet flow rate regulation. (Adapted after [1])



**Fig. 8.20:** A voting/median selector configuration of three analyzers. (Adapted after [1])

option is also presented in Fig. 8.20 and relies on a *median selector* in junction with the three analyzers. This rejects both the highest and the lowest of the three readings, thus transmits the median to the control loop. Reliability is enhanced because the selector also filters noise and transients that are uncommon to two of the signals and protects the reactor against consequences of sensor failure.

In some cases, the value of controlled variable (measured property) is not evenly distributed within the reaction mass. Thus, multiple sensor readings are necessary to decide on the value to be transmitted to the control loop. For example, the temperature in a heterogeneous fixed bed reactor depends on the position of preferential flows, flow rates or catalyst age, etc. Consequently, temperature control is based on the highest reading of sensors aligned in the catalyst bed.

## 8.8 Control structure design for reaction units

### 8.8.1 Principles of control structure design

Designing the *control structure* of a chemical reactor means optimum selection of controlled and manipulated variable pairs as well as adequate choice of controller loops

and regulator configurations. Architecting control structures starts with an exhaustive characterization of the chemical process as well as of the unit that houses it, followed by a well-defined sequence of steps [17–23] – see also chapters 5 and 6. For the first task, a *generic steady-state model* is usually sufficient. It consists of an array of equations describing mass, energy and momentum balances, description of physical and chemical phenomena, characterization of various raw and construction materials, etc. *Dynamic models* are mainly needed for batch units or continuous systems that face frequent and/or significant changes [17–18, 20]. Both model types are completed with *case specific equations* when the controller parameters are defined and tuned.

### Step 1. Definition of operational and control objectives

Control structure design begins with the definition of process *operational objectives* in terms of:

- product quality specifications (desired product distribution/purity, desired molecular weight distribution for polymers, etc.);
- production specifications (desired flow rate, desired residence time, etc.);
- operational constraints (avoid overheating reaction mass, avoid overflowing the unit, avoid drying out unit, maintain certain precision and accuracy of measurements, etc.);
- environmental and economic issues (avoid pollutant overload in wastewater and emissions, save material and energy, minimize operating costs, etc.).

The *control objectives* are derived from lists as above. After their qualitative description these have to be quantified usually in terms of process output variables. For example, a continuous unit requiring thermal stability during steady-state operation translates quantitatively into core temperature that does not vary more than  $\pm 5^\circ\text{C}$  away from setpoint. A batch unit may require minimized cycle duration, that is optimized reaction time. It is demonstrated in Example 8.3 that this translates into respecting a well-defined temperature vs time profile.

A list of possible *disturbance sources* that might upset the desired objectives is helpful in the choice of adequate measured, controlled and manipulated variables. Some examples of disturbance generators are flow rate, composition, and temperature of influent stream; flow rate and temperature of coolant; pressure and purity of inlet gases; etc.

### Step 2. Identification, classification and selection of variables

The next stage of control design deals with identification, classification and selection of variables. *Input* and *output* are rather general terms to define chemical process variables; they can describe characteristics of the feed and discharge streams of either a reactor or its jacket, but can also refer to the control system. For example, a reactor exit flow rate is an output for the chemical unit, but is also the input of a liquid level controller.

Thus, when targeting control architecture, it is more straightforward to refer to *controlled* and *manipulated* variables; either may appear among process inlets or outlets.

Depending on accessibility of their values, variables may be *measurable* (on-stream, *in situ*, or by grab-sampling) or *unmeasurable*. The latter are *estimated* by means of measured values of some other parameters and controller built in mathematical models. A typical example is the liquid level measurement that yields calculated volume and residence time. Special attention is granted to various kinds of *disturbances* [7, 18–19] (step, ramp, impulse, or sinusoidal shaped); these are (often measurable) input variables that upset an operating unit.

The list of *possible* controlled (usually process outputs) as well as manipulated (usually process inputs) variables is put together by following some rules and recommendations [17–20].

The selected controlled variable:

- is not self-regulating (such as the liquid level for a unit that has a pump in the discharge pipe);
- may exceed equipment and operating constraints (such as temperature, pressure, concentration, pH, etc.);
- reflects desired quality (such as viscosity for polymerization or polycondensation reactions, pH for neutralizing processes, a certain chromatographic profile in organic synthesis, etc.);
- is directly related to and affects significantly operational objectives (such as temperature affecting selectivity and product distribution by means of side reaction rates);
- can interact seriously with other controlled variables (such as upstream pH in a series of neutralizing units – see Fig. 8.18);
- has favorable dynamic and static characteristics; it responds rapidly to adjustments imposed by the manipulated variable (such as inter-stage gas temperature for multibed heterogeneous reactors – see Figs. 8.9(a) and 8.26).

The manipulated variable:

- has significant effect on the controlled variable;
- results in rapid response of the controlled variable; in other words affects it preferably directly (dead-time associated with manipulated variable should be relatively small as compared to process time constant);
- does not recycle disturbances (such as an inlet stream that could propagate forward a disturbance, or a recycle flow that could recycle it back to the process).

The controlled variable has to be measured or estimated. However, if possible and economically feasible, it is useful to measure also manipulated variables as well as disturbances. Any measured value has to be reliable, accurate, and precise, recorded with minimum of time delay and from portions of the reaction mass, flows, etc. that significantly represent the changes of the parameter in question (see section 8.7.2). Within this context, the choice of measured variables from the list of controlled ones has to be carefully considered.

### Step 3. Freedom degree analysis

The *freedom degree*  $N_{Freedom}$  of a reactor [17–23] equals the number of variables that have to be specified to describe its operation. In other words, it is the difference between the number of independent variables  $N_{Variables}$  and independent equations  $N_{Equations}$  that interconnect these – see eq. (8.45). Any control loop raises the value of  $N_{Equations}$  (because of control laws), thus lowers  $N_{Freedom}$ . Because in the dynamic balance equations, the accumulation terms differ from zero (and are equal to zero for steady-state),  $(N_{Freedom})_{Dynamic} \geq (N_{Freedom})_{Steady}$ .

$$N_{Freedom} = N_{Variables} - N_{Equations} \quad (8.45)$$

However, for an underspecified process  $N_{Freedom} > 0$ . Some of the freedom degrees are imposed by the process environment (such as the flow rate  $F$  of an upstream unit feeding the reactor in question). Hence, the number of controlled variables  $N$  is at maximum equal to  $N_{Freedom}$ .

$$N = N_{Controlled\ Variables} = N_{Freedom} - N_{Fixed\ Variables} = N_{Manipulated\ Variables} \quad (8.46)$$

The number of independent manipulated variables has to be equal to the number of controlled ones ( $N$ ), although it is possible to design some control systems that can regulate more variables than the number derived from eq. (8.46).

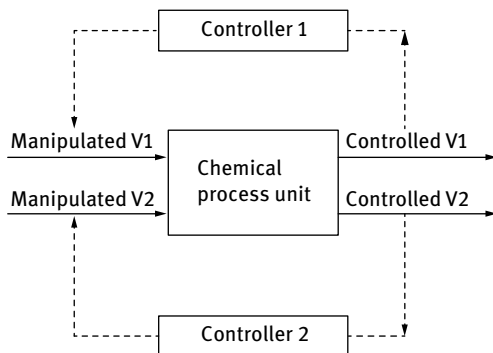
### Step 4. Possible control configurations

After identification and selection in a correct number of possible controlled and manipulated variables, they have to be *paired* in control loops. When each manipulated variable is coupled to a single controlled variable (such as conventional feedback control) the chemical unit has a *multiloop* control configuration. On the other hand, when one manipulated variable depends on two or more controlled variables, a *multivariable* control configuration has been employed (such as model predictive control) – see also chapters 2 and 5.

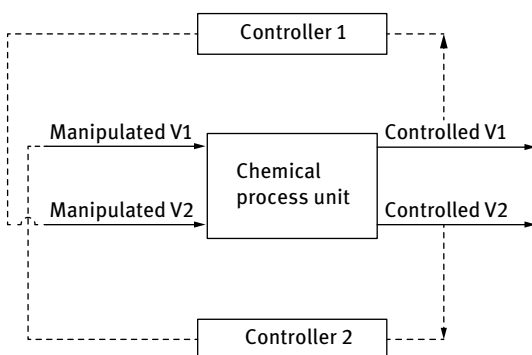
An operating unit with  $N$  controlled and  $N$  manipulated variables has  $N!$  theoretically possible different control configurations. As  $N$  increases, the task of finding the best one is more difficult, because for  $N = 2$ ,  $N! = 1 \times 2 = 2$ ; for  $N = 3$ ,  $N! = 1 \times 2 \times 3 = 6$ ; for  $N = 4$ ,  $N! = 1 \times 2 \times 3 \times 4 = 24$ ; for  $N = 5$ ,  $N! = 1 \times 2 \times 3 \times 4 \times 5 = 120$  and so on. The value of  $N$  also reflects the evolution from SISO (single-input, single-output) to MIMO (multiple-input, multiple-output) control systems.

Figure. 8.21 illustrates the two possible feedback multiloop configurations for a  $N \times N = 2 \times 2$  chemical process. *Controlled*  $V_1$  and *Controlled*  $V_2$  stay for the 2 controlled variables, whereas *Manipulated*  $V_1$  and *Manipulated*  $V_2$  for the 2 corresponding manipulated ones.

As it is probably clear for the reader by now, the most widely employed control configurations are the feedback, inferential and feedforward pairings [7–9, 17–18] (see



(a) First Configuration



(b) Second Configuration

**Fig. 8.21:** Possible feedback multiloop configurations for a  $2 \times 2$  chemical process.

also chapter 1). The *feedback loop* exploits a directly measured *controlled variable* – see Fig. 8.22(a), whereas the *feedforward loop* a measurement of the *disturbance* – see Fig. 8.22(b). The *inferential loop*, in either feedback or feedforward configuration, employs an *estimator block* instead of a direct (primary) measurement for the controlled variable; its value is computed from results of *secondary* measurements (also see Step 2). The estimator is placed on the information stream between the measurement points and the controller (see Fig. 8.22).

### Step 5. Control loop interaction analysis

Although pairing of controlled and manipulated variables follows common sense rules (see also Step 2), the “ultimate” criteria to decide upon the best control system configuration is minimization of cross-interaction among individual loops (see also chapters 4 and 5).

Bristol [24] introduced in 1966 a very simple and effective interaction assessment, the *relative gain array* (RGA). RGA calculus and pairing rules are described in detail

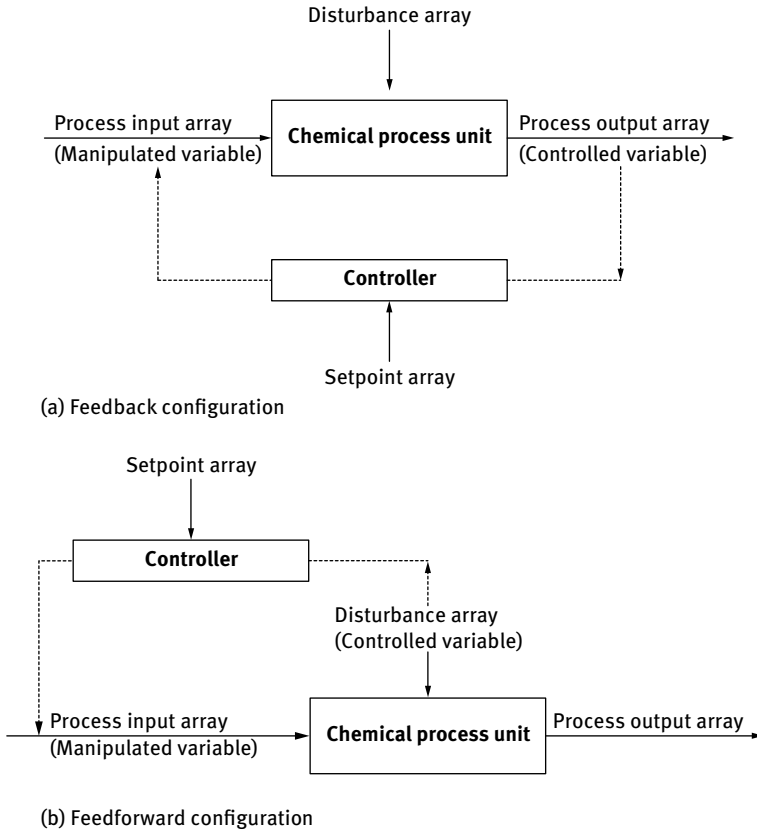


Fig. 8.22: Loop information flow for: (a) feedback control; (b) feedforward control.

in chapter 5. Because it relies on process steady-state gains, it is used in conjunction with the *Niederlinski index* NI to ensure system stability [19]. According to the RGA–NI rules [17–20], suitable controlled and manipulated variable pairs have positive and as close to one as possible RGA elements. Any pairing resulting in negative NI is unacceptable.

Even though the RGA–NI approach is commonly and widely used in the industry, there is scarce systematic information about how to treat effectively control configurations of high-dimensional processes. The RGA–NI serves well for  $2 \times 2$  processes, since negative NI values just state loop instability for  $N > 2$ , but do not prove stability for positive values; it indicates only that the system is not definitely unstable and should be tested *via* dynamic simulation before implementation. Moreover, situations with RGA elements of equal values for individual different yet feasible loops are unsolved. Therefore, improved approaches have been proposed [25], such as the dynamic RGA (DRGA) [26–28], the block relative gain (BRG) [29], the relative normalized gain array (RNGA) [30], the improved controller-dependent DRGA [31], the effective gain array (ERGA) [32] or the effective relative energy array (EREA) [33].

### Step 6. Controller design

The controller is a key element in any control loop. It is the active component that receives the information regarding the value of the controlled variable (either measured or estimated) and by means of a smart transfer function gives the command information to the manipulated variable for the process to rapidly regain its desired setpoint state.

Thus, they are chosen depending both on the identity of controlled variable (temperature, pressure, pH, conductance, color, etc.) and manipulated variable (flow rate, flow ratios, temperatures of heat transfer agents, etc). Typical regulator designs are widely described [7–9, 17–18] and the user may choose from a variety of appliances that have:

- linear or nonlinear (such as “on-off switchers”) relationships among the manipulated variable and the applied command;
- continuous or discrete actions;
- various conventional transfer functions (such as P, PI, PID), some special added features to the latter (such as a preloaded temperature PID controller used to heat-up a batch unit – see section 8.3.4 and Fig. 8.11) or combinations of these (such as the dual-mode controller also described in section 8.3.4).

### 8.8.2 Control structure design for homogeneous ideal units

To demonstrate how the above-described principles are employed, some simple chemical processes will be considered to occur in homogeneous phase in ideal chemical reaction units.

For the first-order exothermic chemical process  $A \xrightarrow{k} P$ , reaction rate  $r$  is described by relationship (8.9). Notations and significance of terms are given in section 8.3. The influent is characterized by flow rate  $F_{in}$ , temperature  $T_{in}^{\circ}$  and concentration  $C_{A0}$ , whereas the effluent by  $F$ ,  $T^{\circ}$  and  $C_A$ , respectively. The reaction mass does not suffer density and heat capacity variations during reaction course. Therefore, values of  $\rho$  and  $C_p$  respectively, are constant. Reaction heat is  $\Delta H_r$ , reaction mass volume is  $V$ . The inlet heat transfer agent has a flow rate of  $F_{ag,in}$  and a temperature of  $T_{ag,in}^{\circ}$ , is mixed perfectly (the jacket behaves like a CSTR) and leaves the jacket with  $F_{ag}$  and  $T_{ag}^{\circ}$ , respectively. The agent density and heat capacity are also constant at  $\rho_{ag}$  and  $C_{p,ag}$ , respectively.  $K_T$  and  $A_T$  refer to heat transfer characteristics of the jacket (see also section 8.3). The reactor is considered to have no thermal inertia.

#### **i** Example 8.6 The CSTR unit

The above reaction takes place in a continuously stirred tank reactor operating in steady-state as illustrated in Fig. 8.23. The generic process model consists of six

independent equations: mass balances (8.47)–(8.49) (expressed in kmol/s), the heat balances (8.50)–(8.51) (expressed in J/s), and level control law (8.37). The accumulation terms of heat and mass balances are equal to zero (steady-state model). These relationships are completed with rate law (8.9), when necessary with  $C_p$ ,  $C_{p,ag}$ , and  $\Delta H_r$ , temperature dependencies [6, 14],  $A_T - V$  connection, and other relationships that characterize the system.

$$F_{in} - F = 0 \quad (8.47)$$

$$F_{ag,in} - F_{ag} = 0 \quad (8.48)$$

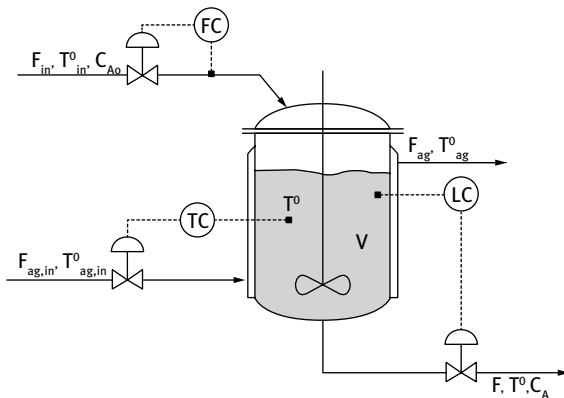
$$F_{in} C_{Ao} - F C_A - V r = 0 \quad (8.49)$$

$$F_{in} \rho C_p T_{in}^{\circ} - F \rho C_p T^{\circ} - \Delta H_r V r - K_T A_T (T^{\circ} - T_{ag}^{\circ}) = 0 \quad (8.50)$$

$$F_{ag,in} \rho_{ag} C_{p,ag} T_{ag,in}^{\circ} - F_{ag} \rho_{ag} C_{p,ag} T_{ag}^{\circ} + K_T A_T (T^{\circ} - T_{ag}^{\circ}) = 0 \quad (8.51)$$

There are 11 independent variables: process inputs  $F_{in}$ ,  $T_{in}^{\circ}$ ,  $C_{Ao}$ ,  $F_{ag,in}$ , and  $T_{ag,in}^{\circ}$  and process outputs  $F$ ,  $T^{\circ}$ ,  $C_A$ ,  $F_{ag}$ ,  $T_{ag}^{\circ}$ , and  $V$  (or residence time  $\frac{V}{F}$ ). By taking into account that  $T_{in}^{\circ}$  and  $T_{ag,in}^{\circ}$  are fixed by the environment of the reactor, three degrees of freedom will remain available for the control system, since  $N_{Variables} = 11$ ,  $N_{Equations} = 6$ , and  $N_{Fixed\ Variables} = 2$ . Thus, according to eqs. (8.45) and (8.46),  $N = 11 - 6 - 2 = 3$ .

The choice of controlled variables is made according to rules and recommendations listed in section 8.8.1. Temperature affects exponentially the rate coefficient and hence strongly modifies the conversion (operational objective related to product quality). Moreover, for exothermic reactions, extra care is required because of runaway danger (operational objective related to equipment constraints and safety).



**Fig. 8.23:** Control system configuration for a CSTR and process described in Example 8.6. (Adapted after [8])



Thus, temperature ought to be controlled. The residence time affects both conversion (product quality) and heat transfer (operational objective related to economic issues). An option of regulating it is by means of volume, in other words by liquid level control. Another parameter affecting balances (8.49) and (8.50) by means of input terms, hence both product specification as well as overall costs, is the influent flow rate. It also affects the value of  $C_{A_0}$ .

The obvious manipulated variables to keep  $T^\circ$ ,  $V$ , and  $F_{in}$  at desired values are adjusting the inlet flow rate for the jacket as well as both the reactor exit and inlet flow rates. Clearly, the measured variables should be core temperature, liquid level, and inlet flow rate.

**Tab. 8.1:** Results of RGA analysis for a  $3 \times 3$  chemical process occurring in a CSTR [8].

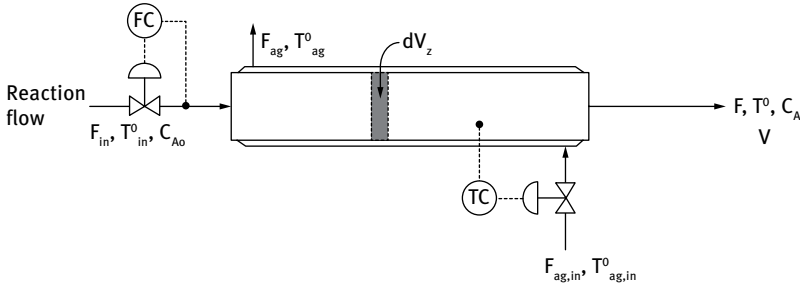
Manipulated variables	Controlled variables		
	$F_{in}$	$T^\circ$	$V$
$F_{in}$	1	0	0
$F_{ag,in}$	0	<b>1.3375</b>	-0.3375
$F$	0	-0.3375	<b>1.3375</b>

Table 8.1 presents the relative gain array for all three controlled and three manipulated variable pairings for the synthesis of ethyl acrylate [8]. Positive and close to 1 elements – see highlighted values in Tab. 8.1 – indicate suitable pairs (see also chapter 5 and section 8.8.1) [17–20]. For this particular case, the Niederlinski index is proven positive [8] and therefore suggests that the control system is not definitely unstable. Hence,

- Core temperature  $T^\circ$  is controlled by manipulating inlet flow rate  $F_{ag,in}$  of heat-transfer agent; the controller uses directly measured  $T^\circ$  values;
- Volume  $V$  (and related residence time) is controlled by manipulating effluent flow rate  $F$ ; the controller uses an estimator block to calculate  $V$  from the measured liquid level – see eq. (8.37);
- Influent flow rate  $F_{in}$  is controlled by manipulating the feed valve; the controller uses directly measured  $F_{in}$  values.

### **i** Example 8.7 The PFR unit

Let us consider that the reaction takes place in a tubular plug flow reactor and the latter has reached steady-state – see Fig. 8.24. The generic model of the previous example modifies somewhat since the reaction mass is described as a plug that advances on the longitudinal axis  $Z$  of the tube (see also section 8.2), thus parameter values are distributed along flow direction.



**Fig. 8.24:** Control system configuration for a PFR and process described in Example 8.7. (Adapted after [8])

Overall mass balances (8.47)–(8.48) are valid for the PFR as well, but eqs. (8.49)–(8.51) are rewritten for the infinitely small volume  $dV$ , which has a  $dA_{T,Z}$  heat-exchange surface. Index  $Z$  is related to position within the reactor,  $dC_{A,Z}$  stands for the variation of (limiting) reactant content within the element  $dV$  and  $dT_{ag,Z}^o$  is the heat transfer agent temperature variation that corresponds to  $dV$  and  $dA_{T,Z}$ . Since each such element behaves like a CSTR, its exit temperature  $T_Z^o$  also corresponds to its core and is the value at which all temperature dependent parameters are calculated (here  $r_Z$  and when necessary  $C_{P,Z}$ ,  $\Delta H_Z$  or others). The accumulation terms are equal to zero for a steady-state model. Integration of eqs. (8.52)–(8.54) between process imposed limits (tube length  $L$  and  $\frac{V}{F}$  residence time) yields balances for the entire reactor.

$$F dC_{A,Z} - dV_Z r_Z = 0 \quad (8.52)$$

$$F \rho C_P (T_{in,Z}^o - T_Z^o) - \Delta H_r dV_Z r_Z - K_T dA_{T,Z} (T_Z^o - T_{ag,Z}^o) = 0 \quad (8.53)$$

$$F_{ag} \rho_{ag} C_{P,ag} dT_{ag,Z}^o + K_T A_{T,Z} (T_Z^o - T_{ag,Z}^o) = 0 \quad (8.54)$$

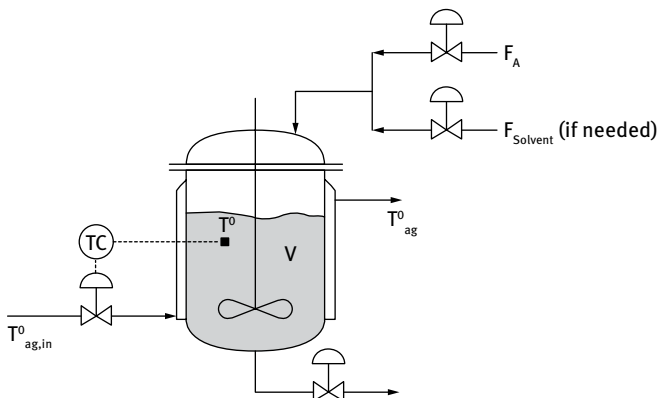
The PFR needs, like the CSTR in Example 8.6, temperature  $T^o$  and residence time  $\frac{V}{F}$  control. The first may be achieved by the usual means of manipulating the inlet jacket flow rate  $F_{ag,in}$ . Since the volume  $V$  of the reaction mixture is imposed environmentally (fixed by reactor dimensions), the only means of  $\frac{V}{F}$  control is by manipulating the inlet flow rate  $F_{in}$  – see Fig. 8.24. Measured variables are  $T^o$  and  $F_{in}$ .

### Example 8.8 The BSTR unit

The same first-order reaction is now considered to occur in a batch stirred tank reactor – see Fig. 8.25. Mass and energy balances have to be written for dynamic behavior and accumulation terms are different from zero. Since reaction mixture volume is imposed by unit dimensions,  $dV/dt = 0$  and  $F = 0$  (relationship (8.37) yields zero). Hence, only mass and energy balances for the reaction mass – relationships (8.55)–(8.56) – and the energy balance for the heat transfer agent – relationship (8.57) – have to be considered in control design. Integration is to be carried out for the time elapse corresponding to



$t_{\text{reaction}} \cdot V_{\text{ag}}$  in eq. (8.57) stands for the fixed volume of the reactor jacket. Mass balances are expressed in kmol/s, whereas energy balances in J/s.



**Fig. 8.25:** Control system configuration for a BSTR and process described in Example 8.8. (Adapted after [8])

$$\left(\frac{d}{dt}\right)(V C_A) = -V r = -V k C_A = -V k_0 \exp\left(-\frac{E}{RT^o}\right) C_A \quad (8.55)$$

$$\rho C_P \frac{d}{dt}(V T^o) = -\Delta H_r V r - K_T A_T (T^o - T_{ag}^o) \quad (8.56)$$

$$\rho_{ag} C_{P,ag} \frac{d}{dt}(V_{ag} T_{ag}^o) = K_T A_T (T^o - T_{ag}^o) \quad (8.57)$$

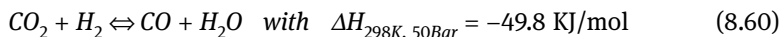
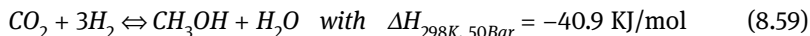
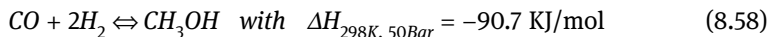
The control system in Fig. 8.25 consists of a temperature control loop, based on the direct measurement of core temperature and jacket inlet flow rate manipulation. Extra features of the regulators for batch reactors are described in section 8.3.4 (see also Fig. 8.11). During unit charging, flow rate controllers for the reactants (maybe also put on ratio adjustment – see chapter 1) are useful but not always necessary. On the other hand, the process end-time detection is both useful and necessary. Chapter 8.7 goes more into this issue's details. Thus, an end-time detector coupled with the temperature regulator may also be imaginable.

### 8.8.3 Control structure design for some heterogeneous units

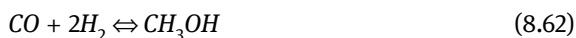


#### **Example 8.9** A multibed heterogeneous catalytic column – heterogeneous PFR

Let us consider a gas phase reaction occurring in a multibed heterogeneous catalytic column like the one in Fig. 8.26 (see also Fig. 4.2). The unit can be treated like a PFR. For example, the industrial methanol production from synthesis gas on copper based catalyst occurs according to independent reaction steps (8.59)–(8.60) [34–36].



The values of reaction enthalpies above were reported by Aasberg-Petersen *et al.* [35]. Other independent side reactions are [34]:



A variety of rate laws have been proposed depending on the employed catalyst, reaction conditions, reactor design, etc. [34–38]. All have in common the dependence of rates on both pressure and temperature; for example those employed by Machado *et al.* [34] in a comparative analysis of methanol production routes:

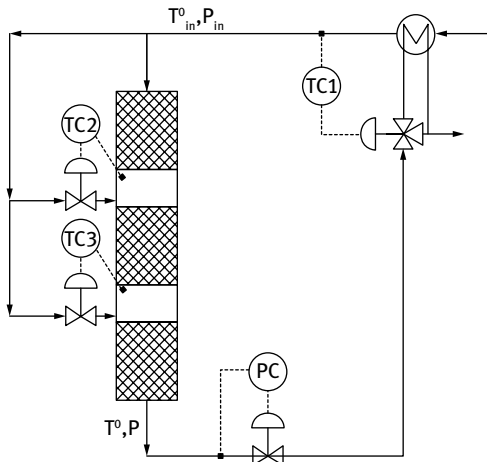
$$r_{\text{equilibrium 1}} = \frac{K_1 K_{\text{CO}} \left( f_{\text{CO}} f_{\text{H}_2}^{1.5} - \frac{f_{\text{CH}_3\text{OH}}}{f_{\text{H}_2}^{0.5} K_{P, \text{equilibrium 1}}^o} \right)}{(1 + K_{\text{CO}} f_{\text{CO}} + K_{\text{CO}_2} f_{\text{CO}_2}) \left( f_{\text{H}_2}^{0.5} + \frac{K_{\text{H}_2\text{O}}}{K_{\text{H}_2}^{0.5}} f_{\text{H}_2\text{O}} \right)} \quad (8.63)$$

$$r_{\text{equilibrium 2}} = \frac{K_2 K_{\text{CO}_2} \left( f_{\text{CO}_2} f_{\text{H}_2}^{0.5} - \frac{f_{\text{CH}_3\text{OH}} f_{\text{H}_2\text{O}}}{f_{\text{H}_2}^{0.5} K_{P, \text{equilibrium 2}}^o} \right)}{(1 + K_{\text{CO}} f_{\text{CO}} + K_{\text{CO}_2} f_{\text{CO}_2}) \left( f_{\text{H}_2}^{0.5} + \frac{K_{\text{H}_2\text{O}}}{K_{\text{H}_2}^{0.5}} f_{\text{H}_2\text{O}} \right)} \quad (8.64)$$

$$r_{\text{equilibrium 3}} = \frac{K_3 K_{\text{CO}_2} \left( f_{\text{CO}_2} f_{\text{H}_2} - \frac{f_{\text{H}_2\text{O}} f_{\text{CO}}}{K_{P, \text{equilibrium 3}}^o} \right)}{(1 + K_{\text{CO}} f_{\text{CO}} + K_{\text{CO}_2} f_{\text{CO}_2}) \left( f_{\text{H}_2}^{0.5} + \frac{K_{\text{H}_2\text{O}}}{K_{\text{H}_2}^{0.5}} f_{\text{H}_2\text{O}} \right)} \quad (8.65)$$

The significance of terms in eqs. (8.63)–(8.65) is:

- $r_{\text{equilibrium},i}$  (where  $i = 1, 2,$  and  $3$ ) stand for the rate of individual reaction steps (8.58), (8.59) and (8.60) in this sequence (kmol/m<sup>3</sup>·s);
- dimensionless  $K_{P, \text{equilibrium},i}^o$  and  $K_i$  stand for the equilibrium constants of above reaction steps for standard and employed reaction conditions, respectively;
- dimensionless  $K_{\text{CO}}, K_{\text{CO}_2}, K_{\text{CH}_3\text{OH}}, K_{\text{H}_2}$ , and  $K_{\text{H}_2\text{O}}$  stand for the adsorption equilibrium constants of gas phase species on the solid catalyst surface;
- dimensionless  $f_{\text{CO}}, f_{\text{CO}_2}, f_{\text{CH}_3\text{OH}}, f_{\text{H}_2}$ , and  $f_{\text{H}_2\text{O}}$  stand for the fugacity of each species involved in main reaction steps.



**Fig. 8.26:** Control system of a gas phase reaction in a multibed heterogeneous reaction column as described in example 8.9. (Adapted after [1, 8])

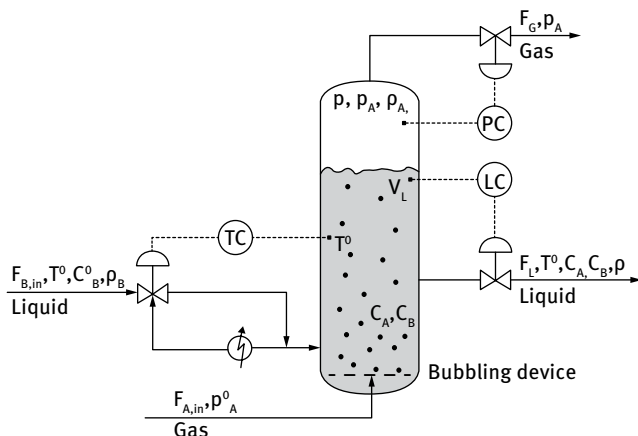
Relationships (8.63)–(8.65) prove the importance of coupled pressure and temperature control. Used together with mass and energy balances like those described in Example 8.7, they complete the steady-state model of this reaction unit – see Fig. 8.26.

*Fugacity* in gas phase is similar to the activity in condensed phases [6, 14] (see also chapter 8.6.1). Its value depends on both total gas pressure and temperature. Thus, rate laws are both pressure and temperature dependent because of both fugacity and various equilibrium constant values. As a consequence, the reactor needs temperature as well as pressure control. The control structure in Fig. 8.26 not only permits temperature control at the reactor inlet *via* loop TC1 but also ensures optimum temperature profiles (see also Fig. 8.9) by adjusting it between catalyst beds *via* loops TC2 and TC3 respectively (see also Fig. 8.9 and section 8.3.3). All loops rely on direct measurements and manipulate gas flow rates. The outlet flow can preheat the inlet reactants, thus lowers overall operating costs of the unit. This is possible by means of a three-way valve.

### **i** Example 8.10 A bubbling column – heterogeneous CSTR

Let us consider a heterogeneous gas-liquid reaction occurring in a bubble column like the one in Fig. 8.27. Such units have a simple design, are suitable for continuous and semi-continuous operation and provide high back-mixing in the liquid phase. Therefore, their model derives from that of the CSTR. One of the main advantages is that they can operate in virtually isotherm conditions because of liquid evaporation. Disadvantages are due to the limited employable temperature and pressure ranges [11].

For the case in Fig. 8.27, the gas phase reactant *A* has to dissolve first in the liquid phase *B* to react with it (oxidation of various organic liquids [11, 39], commercial



**Fig. 8.27:** Control system of a heterogeneous gas-liquid bubbling column as described in Example 8.10. (Adapted after [8])

production of acetic acid by carbonylation of methanol [40–41], carbon dioxide sequestration by caustic media [39, 42–45]). In other words, mass transfer by means of gas and liquid phase diffusion as well as absorption of gas into the liquid are a prerequisite of the actual chemical reaction [4, 39]. Hence, these steps can also be rate determining.

Let us consider the process of carbon dioxide sequestration by ammonia [42–44] or other organic amines such as mono-ethanol-amine and bi-ethanol-amine [45]. Here, diffusion and chemical reaction are much faster than absorption. Other simplifying assumptions are: the reaction occurs in the liquid bulk phase between a dissolved species  $A$  entering the unit with the gas inlet and species  $B$  contained by the liquid according to  $A_{(L)} + B_{(L)} \xrightarrow{-k} P_{(L)}$ ; the reaction rate law (8.66) is of second-order, the unit is perfectly stirred and operates under isotherm conditions at  $T^0$ . If the liquid inlet temperature corresponds to the core value and the reaction heat is negligible ( $\Delta H_r \approx 0$ ), energy balances are not necessary here.

$$r = k C_A C_B = k_0 \exp\left(-\frac{E_a}{RT^0}\right) C_A C_B \quad (8.66)$$

The rate coefficient  $k$  in eq. (8.66) is expressed in  $\text{m}^3/\text{kmol}\cdot\text{s}$  [5–6, 14]. For the significance of the other terms, see eq. (8.6). By taking into account the listed simplifying assumptions, the generic steady-state model will consist of overall mass balances (expressed in  $\text{kmol/s}$ ) for the liquid as well as gas phase in eq. (8.70) and (8.71) respectively, mass balances (also expressed in  $\text{kmol/s}$ ) for reactants  $A$  and  $B$  in eq. (8.72) and (8.73), respectively, and relationships (8.67)–(8.69) describing the mass transfer (absorption) of species  $A$  at the gas-liquid interface. The accumulation terms in eqs. (8.70)–(8.73) are equal to zero.

$$C_{A0} = \frac{p_A}{k_{\text{Henry},A}}; \quad k_{\text{Henry},A} = f(T^\circ) \quad (8.67)$$

$$p_A = \frac{n_A}{n} p; \quad \rho_A = \frac{p M_A}{R T^\circ} \quad (8.68)$$

$$N_A = k_{L,A} (C_{A0} - C_A) \quad (8.69)$$

$$F_{B,\text{in}} \rho_B + M_A N_A A_{G-L} - F_L \rho = 0 \quad (8.70)$$

$$F_{A,\text{in}} - F_G - \frac{A_{G-L} N_A M_A}{\rho_A} = 0 \quad (8.71)$$

$$A_{G-L} N_A - F_L C_A - V_L k C_A C_B = 0 \quad (8.72)$$

$$F_{B,\text{in}} C_{B0} - F_L C_B - V_L k C_A C_B = 0 \quad (8.73)$$

The significance of the terms in the above equations is (see also Fig. 8.27):

- $F_{A,\text{in}}$  and  $F_{B,\text{in}}$  are the gas and liquid inlet flow rates ( $\text{m}^3/\text{s}$ );
- $F_G$  and  $F_L$  are the gas and liquid outlet flow rates ( $\text{m}^3/\text{s}$ );
- $C_{A0}$ ,  $C_{B0}$ ,  $C_A$ , and  $C_B$  are the liquid phase concentrations ( $\text{kmol}/\text{m}^3$ ) of species  $A$  and  $B$  respectively;
- $V_L$  is the reaction mass volume ( $\text{m}^3$ );
- $\rho_A$  and  $\rho_B$  are the densities of bulk gas-phase  $A$  and liquid-phase  $B$  respectively ( $\rho$  stands for the density of reaction mixture, which is the exit liquid flow  $\text{kg}/\text{m}^3$ );
- $p$  stands for the total pressure (atm), whereas  $p_A$  stands for the partial pressure of  $A$  (atm) in the unit's vapor space;
- $k_{\text{Henry},A}$  is the temperature dependent Henry's coefficient characterizing the mass transfer of  $A$  from gas to liquid ( $\text{kmol}/\text{m}^3 \cdot \text{atm}$ );
- $A_{G-L}$  is the total mass transfer surface ( $\text{m}^2$ ) given by bubbles;
- $n$  is the total mole number, whereas  $n_A$  stands for the mole number of  $A$  in gas phase (kmol);
- $M_A$  is the molecular mass of species  $A$  ( $\text{kg}/\text{kmol}$ );
- $N_A$  is the molar flow rate of  $A$  from gas to liquid through mass transfer surface ( $\text{kmol}/\text{m}^2 \cdot \text{s}$ );
- $k_{L,A}$  is the flow rate coefficient ( $\text{m}^3/\text{m}^2 \cdot \text{s}$ );
- $T^\circ$  is the core temperature (K);
- $R$  is the universal gas constant ( $\text{m}^3 \cdot \text{atm}/\text{kmol} \cdot \text{K}$ ).

An operational objective such as constant product distribution in the liquid effluent requires the control system presented in Fig. 8.27. Since the temperature and pressure dependent mass transfer of  $A$  from gas to liquid (absorption) is assumed to be rate determining, both  $T^\circ$  and  $p$  control are required.  $T^\circ$  is measured directly within the core of the reaction mass and regulated *via* the temperature of the liquid inlet  $F_{B,\text{in}}$ . The pressure is

also measured directly, but in the unit's vapor's space. Its value is regulated by means of the gas outlet flow valve, in other words, *via* value of  $F_G$ . Residence time control ensures both the above stated operational objective and avoids flooding or drying out the column. It relies on a level measurement and a  $V_L$  estimator that accordingly manipulates the liquid effluent flow rate  $F_L$ .

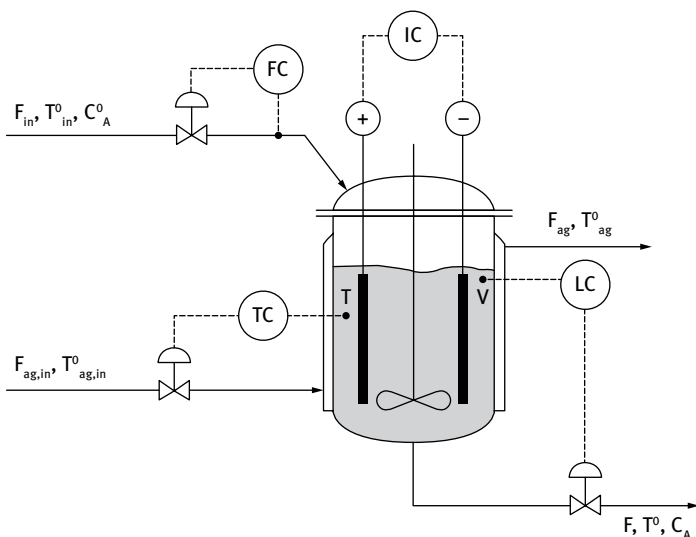
It is worth mentioning that value of  $A_{G-L}$  can be raised using various bubbling devices or stirrers or by packing the unit with inert fillers [4, 11, 39].

### Example 8.11 A stirred batch electrolysis cell – heterogeneous CSTR



Electrochemical reactors are useful when high-purity products are desired that cannot be obtained otherwise with economically feasible selectivity. These units can operate continuously either like a CSTR or a PFR, semicontinuously or in batch cycles [8, 11]. Electrolysis occurs when an externally applied current flows between two solid electrodes through a usually liquid electrolyte. As a result, it mediates electron exchange between chemical species – see electrolysis law (8.74). It has the advantage of much lower temperatures than conventional thermal or catalytic procedures. The products can be either solid deposits (metal refining), liquid (sodium hydroxide solution, adiponitril, etc.) or gaseous products (chlorine, hydrogen, ozone, etc.).

The process requires a minimum voltage to be applied; the thermodynamically predicted value is called *decomposition voltage*, yet the real voltage is higher (disadvantage of such units) because of electrode overpotential and ohmic losses



**Fig. 8.28:** Control system of a CSTR electrochemical unit as described in Example 8.11. (Adapted after [8, 11])



[6, 14]. This fact has to be considered in any steady-state model – see eqs. (8.80)–(8.83) [6, 14, 46–47]. Chemical transformations happening both on electrodes and in bulk electrolyte as well as mass transport phenomena (when occurring with significant rate [48–51]) are also parts of the model.

Let us consider the electrochemical unit illustrated by Fig. 8.28, in which the first-order reaction  $A + z_A e^- \xrightarrow{k} P$  takes place at constant temperature  $T^\circ$ . According to the first law of electrolysis [6, 14], the mass  $m$  altered at each electrode is directly proportional to the quantity of electricity transferred at it. Thus,

$$n_A = \frac{m_A}{M_A} = \frac{I \frac{V}{F}}{z_A F_{araday}} \quad (8.74)$$

The significance of terms not listed yet within this chapter is:

- $n_A$  stands for the transformed mole number of  $A$  (mol), whereas  $m_A$  corresponds to the mass (g);
- $M_A$  stands for molecular mass of  $A$  (g/mol);
- $z_A$  stands for the number of exchanged electrons by  $A$  during the electrochemical process;
- $I$  is the current intensity (A);
- terms  $\frac{V}{F}$  and  $F_{araday}$  stand for the *electrolysis* (residence) *time* (s) and Faraday constant (96485 C/mol), respectively.

Let us make some simplifying assumptions:

- The reactor is perfectly stirred and behaves like a CSTR [46];
- The chemical process does not affect significantly the liquid flow density  $\rho$  (kg/m<sup>3</sup>) and its heat capacity  $C_p$  (J/kg·K) so that these are considered constant;
- The reaction heat is negligible ( $\Delta H_r \approx 0$ ), so that reaction generated heat will not affect energy balance (8.76).

The steady-state mass balances (8.75)–(8.76) have zero accumulation terms. Mass balances (kmol/s) (8.75) and (8.47)–(8.48) as well as energy balances (J/s) (8.76) and (8.51) are completed by eqs. (8.77)–(8.83). These describe the overall voltage, electrode potentials, various overpotentials, other physical and chemical phenomena.

$$F_{in} C_{A0} - F C_A - \frac{I}{z_A F_{araday}} = 0 \quad (8.75)$$

$$F_{in} \rho C_p T_{in}^\circ - F \rho C_p T^\circ - K_T A_T (T^\circ - T_{ag}^\circ) + R_{total} \frac{I^2}{2} = 0 \quad (8.76)$$

$$C_A = C_{A0} - \frac{I}{z_A F_{araday} F} \quad (8.77)$$

$$U = \varepsilon_{anode} + \varepsilon_{cathode} + \eta_{anode} + \eta_{cathode} + \eta_{ohmic} \quad (8.78)$$

$$(\varepsilon_{ox/red})_{electrode} = (\varepsilon_{ox/red}^{\circ})_{electrode} - \frac{2.303 R T^{\circ}}{z F_{araday}} \lg \left( \frac{a_{ox}}{a_{red}} \right)_{electrode} \quad (8.79)$$

$$\eta_{electrode} = const_1 T^{\circ} + const_2 T^{\circ} \lg (i_{electrode}) \quad (8.80)$$

$$i_{electrode} = \frac{I}{A_{electrode}} \quad (8.81)$$

$$\eta_{ohmic} = I (R_{electrolyte} + R_{other}) \quad (8.82)$$

$$R_{electrolyte} = \frac{d}{\lambda_{electrolyte} A_{anode+cathode}} \quad (8.83)$$

Again, the significance of terms not listed yet within this chapter is:

- $i$  stands for the current density ( $A/m^2$ );
- $A_i$  stands for the indicated electrode surface ( $m^2$ );
- $U$  stands for the applied potential (V);
- $(\varepsilon_{ox/red}^{\circ})_{electrode}$  stands for the reduction potential (V) at the specified electrode (anode or cathode); according to the Nernst law (8.79) [6, 14] it is written as a function of the standard value  $\varepsilon_{ox/red}^{\circ}$  (V) and activities of the oxidized and reduced forms of the species that suffers electrochemical transformation ( $\alpha_{ox}$  and  $\alpha_{red}$  respectively, both expressed in  $kmol/m^3$ );
- $\eta_{electrode}$  stands for the overpotential (V) at the specified electrode (anode or cathode); according to the Butler-Volmer law (8.80) [6, 14] it is written as a function of local current density  $i_{electrode}$  and temperature  $T^{\circ}$ ;
- $\eta_{ohmic}$  stands for potential drop (V) caused by the resistance of all electrical charge transporting media (electrolyte, wires, contactors, etc.);
- $R_i$  stands for the electrical resistance ( $\Omega$ ) of the specified media:  $R_{electrolyte}$  for the electrolyte,  $R_{total}$  for electrical charge transporting media that come into contact with the electrolyte and can heat it up due to electrical current transport (Joule heating);  $R_{other}$  for the wiring, contactors, etc.;
- $\lambda_{electrolyte}$  stands for the electrolyte conductivity (S/m);
- $d$  represents the distance between anode and cathode (m).

The electrochemical cell presented in Fig. 8.28 and the CSTR in Fig. 8.23 behave alike. Hence, the electrochemical reactor requires core temperature  $T^{\circ}$ , electrolysis (residence) time  $\frac{V}{F}$  and inlet flow rate  $F_{in}$  control. This is achieved as described for Example 8.6. Because current intensity  $I$  governs mass transformation, this unit needs an additional  $I$  control loop. Its value is adjusted by manipulating the applied voltage  $U$ .

There are some comments worth making for Examples 8.6–8.11. The described control schemes serve well and under stabile conditions. Yet, protection against wide variations of influent quality ( $C_{A0}$  or  $p_A^{\circ}$  not constant, not adequate electrode

material, etc.) or undesired phenomena (worsened heat transfer because of various deposits, incorrect measurement of controlled variable, detector failure or malfunction, etc.) is not guaranteed. These would require extra control loops and more complex control structures.

Control systems illustrated in Fig. 8.23–8.28, even if simple in design, can benefit from smart built in models as well as from optimized characteristics. For example, core temperature control of the polyvinyl chloride synthesis batch reactor improves when a nonlinear model predictive controller NMPC (see chapter 2.9) is used instead of a conventional PID [52]; strong disturbances and frequent changes will be handled well within smaller time windows.

## References

- [1] Kendall, D.C., Schlegel, D.F., Hertanu, H.I., Molnár, F., Lipták, B.G., *Chapter 8.9. Chemical Reactors: Basic Control Strategies*, p. 1664–1696 in Lipták, B.G. (Ed.), *Instrument Engineers' Handbook*, Vol. II. *Process Control and Optimization*, 4<sup>th</sup> Edition, CRC Press, Boca Raton, FL, 2006.
- [2] Shinskey, F.G., *Chapter 8.10. Chemical Reactors: Control and Optimization*, p. 1697–1710 in Lipták, B.G. (Ed.), *Instrument Engineers' Handbook*, Vol. II. *Process Control and Optimization*, 4<sup>th</sup> Edition, CRC Press, Boca Raton, FL, 2006.
- [3] Simándi, B. (Ed.), Cséfalvay, E., Déak, A., Farkas, T., Hanák, L., Mika, L.T., Mizsey, P., Sawinsky, J., Simándi, B., Szánya, T., Székely, E., Vágó, E., *Vegyipari műveletek II. Anyagátadó műveletek és kémiai reaktorok*, 2<sup>nd</sup> Edition, Typotex Kiadó, Budapest, 2012.
- [4] Levenspiel, O., *Chemical Reaction Engineering*, 3<sup>rd</sup> Edition, John Wiley & Sons, New-York, 1999.
- [5] Marin, G., Yablonsky, G.S., *Kinetics of Chemical Reactions. Decoding complexity*, Wiley-VCH, Weinheim, 2011.
- [6] Atkins, P., de Paula, J., *Physical Chemistry*, 9<sup>th</sup> Edition, Oxford University Press, Oxford, 2010.
- [7] Agachi, P.S., Cristea, M.V., *Basic Process Engineering Control*, DeGruyter, Berlin/Boston, 2014.
- [8] Agachi, Ş., *Automatizarea proceselor chimice*, Casa Cărţii de Ştiinţă, Cluj-Napoca, 1994.
- [9] Shinskey, F.G., *Process Control Systems*, 4<sup>th</sup> Edition, McGraw-Hill, New-York, 1996.
- [10] Hopkins, B., *Pressure Monitored Temperature Controlled System for a Liquid-Vapor Process*, Patent US 3708658 A, Monsanto Company, 1973.
- [11] Henkel, K.D., *Reactor Types and Their Industrial Applications*, in *Ullman's Encyclopedia of Industrial Chemistry*, Wiley-VCH, Weinheim, Vol. 31, p. 293–327, 2012.
- [12] Imre-Lucaci, A., Agachi, P.Ş., *Optimizarea proceselor din industria chimică*, Editura Tehnică, Bucureşti, 2002.
- [13] Agachi, P.Ş., Vass, E., *Optimizarea reactorului de fabricare a metanolului utilizând principiul maximului lui Pontryagin* Revista de Chimie, 6, 513–521, 1995.
- [14] Motschmann, H., Hofmann, M., *Physikalische Chemie*, DeGruyter, Berlin/Boston, 2014.
- [15] Csavdári, A., *Catalytic Kinetic Methods in Analytical Chemistry. Principles and Applications*, Editura MEGA, Cluj-Napoca, 2008.
- [16] Hoyle, D.L., McMillan, G.K., Shinskey, F.G., *Chapter 8.32. pH Control*, p. 2044–2056 in Lipták, B.G. (Ed.), *Instrument Engineers' Handbook*, Vol. II. *Process Control and Optimization*, 4<sup>th</sup> Edition, CRC Press, Boca Raton, FL, 2006.
- [17] Luyben, W.L., *Chemical Reactor Design and Control*, John Wiley & Sons, Hoboken, NJ, 2007.
- [18] Seborg, D.E., Edgar, T.F., Mellichamp, D.A., *Process Dynamics and Control*, 2<sup>nd</sup> Edition, John Wiley & Sons, New York, 2004.

- [19] Suliman, M.A., Ali, E.M., Alhumaizi, K.I., Ajbar, A.H.M., *Control System Design and Structure*, Lecture Notes of the "Process Control in the Chemical Industries" Workshop, King Saud University, Riyadh, Saudi Arabia, p. 39–64, 2002.
- [20] Skogestad, S., *Control Structure Design for Complete Chemical Plants*, Computers and Chemical Engineering, 28, 219–234, 2004.
- [21] Chen, R., McAvoy, T., *Plantwide Control System Design. Methodology and Application to a Vinyl Acetate Process*, Industrial & Engineering Chemistry Research, 42, 4753–4771, 2003.
- [22] Benyahia, B., Lakerveld, R., Barton, P.I., *A Plant-Wide Dynamic Model of a Continuous Pharmaceutical Plant*, Industrial & Engineering Chemistry Research, 51, 15393–15412, 2012.
- [23] Lakerveld, R., Benyahia, B., Braatz, R., Barton, P.I., *Model-based Design of Plant-wide Control Strategy for a Continuous Pharmaceutical Plant*, AIChE Journal, 59, 3671–3685, 2013.
- [24] Bristol, E., *On a New Method of Interactions for Multivariable Process Control*, IEEE Transactions on Automatic Control, 11, 133–134, 1966.
- [25] Shen, Y., Cai, W.J., Li, S., *Multivariable Process Control: Decentralized, Decoupling or Sparse?*, Industrial & Engineering Chemistry Research, 49, 761–771, 2010.
- [26] Jain, A., Babu, B.V., *A New Measure of Process Interaction in the Domain Dynamics*, Proceedings of AIChE Annual Meeting, San Francisco, CA, USA, Paper 202d, 2013.
- [27] Tung, L.S., Edgar, T.F., *Analysis of Control-Output Interactions in Dynamic Systems*, AIChE Journal, 27, 690–693, 1981.
- [28] Witcher, M.F., McAvoy, T.J., *Interacting Control Systems: Steady-State and Dynamic Measurement of Interaction*, ISA Transactions, 16, 35–41, 1977.
- [29] Kariwala, V., Forbes, J.F., Meadows, E.S., *Block Relative Gain: Proprieties and Pairing Rules*, Industrial & Engineering Chemistry Research, 42, 4564–4574, 2003.
- [30] Chen, Q., Luan, X., Liu, F., *RNGA Loop Pairing Criterion for Multivariable Systems Subject to a Class of Reference Inputs*, Proceedings of the 19<sup>th</sup> World Congress of The International Federation of Automatic Control, Cape Town, South Africa, p. 4721–4726, 2014.
- [31] Avoy, T.M., Arkun, Y., Chen, R., Robinson, D., Schnelle, P.D., *A New Approach to Defining a Dynamic Relative Gain*, Control Engineering Practice, 11, 907–914, 2003.
- [32] Xiong, Q., Cai, W.J., He, M.J., *A Practical Loop Pairing Criterion for Multivariable Processes*, Journal of Process Control, 15, 741–747, 2005.
- [33] Naini, N.M., Fatechi, A., Khaki-Sedigh, A., *Input-output Pairing Using Effective Relative Energy Array*, Industrial & Engineering Chemistry Research, 48, 7137–7144, 2009.
- [34] Machado, C.F.R., de Medeiros, J.L., Araújo, O.F.Q., *A Comparative Analysis of Methanol Production Routes: Synthesis Gas versus CO<sub>2</sub> Hydrogenation*, Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management, p. 2981–2990, Bali, Indonesia, 2014.
- [35] Aasberg-Petersen, K., Stub Nielsen, C., Dybkaer, I., Perregaard, J., *Large Scale Methanol Production From Natural Gas*, Whitepaper, Haldor Topsoe, 2008.
- [36] Rahman, D., *Kinetic Modeling of Methanol Synthesis from Carbon Monoxide, Carbon Dioxide and Hydrogen Over a Cu/ZnO/Cr<sub>2</sub>O<sub>3</sub> Catalyst*, Master's Thesis, San José State University, USA, 2012.
- [37] Kuzsynski, M., Browne, W.I., Fontein, H.J., Westerterp, K.R., *Reaction Kinetics for the Synthesis of Methanol from CO and H<sub>2</sub> on a Copper Catalyst*, Chemical Engineering and Processing, 21, 179–191, 1987.
- [38] Panahi, P.N., Mousavi, S.M., Niaei, A., Farzi, A., Salari, D., *Simulation of Methanol Synthesis from Synthesis Gas in Fixed Bed Catalytic Reactor Using Mathematical Modeling and Neural Networks*, International Journal of Scientific & Engineering Research, 3, 1–7, 2012.
- [39] Harriot, P., *Chemical Reactor Design*, Marcel Dekker. (by Taylor & Francis Group LLC), New York, 2003.
- [40] Dake, S.B., Jaganathan, R., Chaudhari, R.V., *New Trends in the Rate Behavior of Rhodium-Catalyzed Carbonylation of Methanol*, Industrial & Engineering Chemistry Research, 28, 1107–1110, 1989.

- [41] Nowicki, L., Ledakowicz, S., Zarzycki, R., *Kinetics of Rhodium-Catalyzed Methanol Carbonylation*, Industrial & Engineering Chemistry Research, 31, 2472–2475, 1992.
- [42] Qin, F., Wang, S., Hartono, A., Svendsen, H.F., Chen, C., *Kinetics of CO<sub>2</sub> Absorption in Aqueous Ammonia Solution*, International Journal of Greenhouse Gas Control, 4, 729–738, 2010.
- [43] Liu, J., Wang, S., Qi, G., Zhao, B., Chen, C., *Kinetics and Mass Transfer of Carbon Dioxide Absorption into Aqueous Ammonia*, Energy Procedia, 4, 525–532, 2011.
- [44] Jeon, S.B., Seo, J.B., Lee, H.D., Kang, S.K., Oh, K.J., *Absorption Kinetics of Carbon Dioxide into Aqueous Ammonia Solution: Addition of Hydroxyl Groups for Suppression of Vaporization*, Korean Journal of Chemical Engineering, 30, 1790–1796, 2013.
- [45] Hsu, C.H., Chu H., Cho, C.M., *Absorption and Reaction Kinetics of Amines and Ammonia Solutions with Carbon Dioxide in Flue Gas*, Journal of Air & Waste Management Association, 53, 246–252, 2003.
- [46] Scott, K., *The Continuous Stirred Tank Electrochemical Reactor. An Overview of Dynamic and Steady-State Analysis for Design and Modelling*, Journal of Applied Electrochemistry, 21, 945–960, 1991.
- [47] Babajide, A.D., Fahidy, T.Z., *Problems in Modelling and the Study of Electrochemical Reactor Dynamics*, The Canadian Journal of Chemical Engineering, 46, 253–258, 1968.
- [48] Trinidad, P., Walsh, F., Gilroy, D., *Conversion Expressions for Electrochemical Reactors which Operate under Mass Transport Controlled Reaction Conditions, Part I: Batch Reactor, PFR and CSTR*, International Journal of Engineering Education, 14, 431–441, 1998.
- [49] Walsh, F., Trinidad, P., Gilroy, D., *Conversion Expressions for Electrochemical Reactors which Operate under Mass Transport Controlled Reaction Conditions, Part II: Batch Recycle, Cascade and Recycle Loop Reactors*, International Journal of Engineering Education, 21, 981–992, 2005.
- [50] Thilakavathi, R., Rajasekhar, D., Balasubramanian, N., Srinivasakannan, C., Al Shoaibi, A., *CFD Modeling of Continuous Stirred Tank Electrochemical Reactor*, International Journal of Electrochemical Science, 7, 1386–1401, 2012.
- [51] Rodriguez, G., Sierra-Espinosa, F.Z., Teloxa, J., Álvarez, A., Hernández, J.A., *Hydrodynamic Design of Electrochemical Reactors Based on Computational Fluid Dynamics*, Desalination and Water Treatment, DOI: 10.1080/19443994.2015.1114169, 2015.
- [52] Nagy, Z., Agachi, Ş., *Model Predictive Control of a PVC Batch Reactor*, Computers & Chemical Engineering, 21, 571–591, 1997.

## 9 Control of distillation processes

*Distillation* is a separation technique of miscible liquids. It is based on the *boiling point* (more precisely on the *volatility*) difference of components and usually appears at the end of a technological flux. Therefore, the characteristics and behavior of upstream operations (reactions, extractions, etc.) may significantly affect its operating specifications. Therefore, a well-designed and implemented automated control system (ACS) has to be robust against any upstream (or any other) disturbances.

Distillation columns exhibit slow dynamic responses as compared to other units because of the high number of mass transfer stages that may be involved. This fact combined with the numerous other factors affecting distillation raises special control issues and require more attention during design of ACS configurations [1, 2].

### 9.1 Economic constraints of distillation

One major objective of distillation is to deliver products of prescribed purity and quality. These specifications are usually imposed by marketing price. Hence, minimization of operational costs is an economic requirement. In other words, distillation columns have to operate under the most economically profitable conditions. Naturally, other operation related constraints must be obeyed as well.

Prescribed purity/quality specifications can be achieved by either using a rigorous process control (enables direct production) or by mixing two products of different purities/qualities. However, the latter option is not necessarily economically profitable since the energy spent to obtain the high-purity product is lost.

There is no direct correlation between product price and its purity. Yet, products outrunning minimal specifications do not sell better nor result in improved incomes. Thus, there is no point in surpassing these. The diagram in Fig. 9.1 demonstrates that the maximal income is obtained by the minimal required purity that still ensures product marketing. By comparing Grade 1 and Grade 2 products, it is obvious that separation costs increase with purity, whereas the selling price remains constant. Yet, if the prescribed quality is not achieved to the desired first-grade, cost penalizations are to be considered because the product's marketing price drops to that of the second-grade.

Therefore, the majority of distillation columns are operated to slightly overdo specifications, so that the actual purity never falls under the prescribed value. For this reason, distillation columns are equipped with deposition vessels that act as concentration equalizers. Their removal is economically disadvantageous.

Expressing these facts in numbers means that if, for example, the product purity is allowed  $\pm 0.2\%$  fluctuation with respect to the desired value, the concentration controller setpoint should be with  $0.2\%$  above it. The benzene-toluene separation

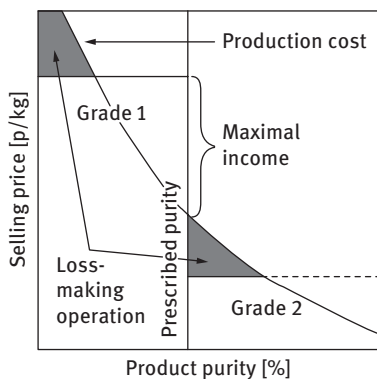


Fig. 9.1: Distillation: marketing price vs product purity. (Adapted after [1])

for instance, prescribes a toluene purity of 99.5%; thus, a  $\pm 0.2\%$  fluctuation sets the controller to 99.7%. Meanwhile, a setpoint of 99.7 instead of 99.5%, results in a 10% energy consumption increase. Moreover, a setpoint of 99.9% toluene consumes 36% more energy.

Industrial practice showed that for the benzene-toluene separation by distillation, a slightly improved setpoint increases energy consumption by at most 18% as compared to the fulfillment of minimal specifications. This is due to the nonlinear process characteristics within this concentration domain, where toluene purity is closer to 99.5 rather than to 99.9% if the controller is set to 99.7%.

Sometimes distillation yields an intermediate, not the final, product of a technological flux. Thus, the diagram in Fig. 9.1 is not valid and it makes more sense to aim for energy saving, while approximately respecting quality requirements, rather than to target rigorous specifications.

## 9.2 The recovery factor

Figure 9.2 illustrates schematically a *binary distillation* column. The valuable product is evacuated with the *distillate stream*  $D$ . The feed, distillate, liquid phase bottom (*waste*), and top vapor phase molar flow rates are symbolized by  $F$ ,  $D$ ,  $B$ , and  $V$ , respectively. The corresponding liquid phase mole fraction of the volatile compound are symbolized by  $x$ , whereas that of the vapor phase by  $y$ .

The *recovery factor*  $r$  is defined as the fraction of valuable product within the feed evacuated with the distillate stream – see eq. (9.1). Defined as such,  $r$  has values between 0% and 100%. The recovery decreases with increasing product purity (see explanations for Figs. 9.3 and 9.4)

$$r = \frac{Dx_D}{Fx_F} 100 \quad (9.1)$$

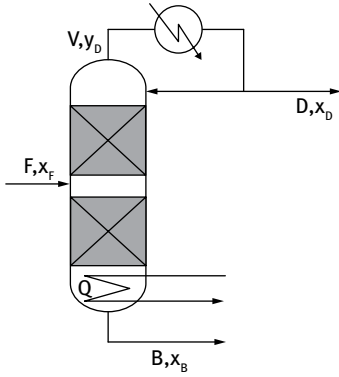


Fig. 9.2: Schematic representation of a binary distillation column.

If the feed departs the system entirely as distillate ( $F = D$ ), the concentration of the valuable species within  $D$  and  $F$  streams are equal and  $r = 100\%$ . In other words, the compound of interest has been fully recovered, but without concentrating it. If, however,  $x_F = 0.5$ ,  $x_D = 1$  and  $D = F/2$ , then a 100% recovery is coupled with a significant 50% concentration improvement.

A relationship between the desired species concentrations and the recovery factor can be established by combining information of equations (9.1) and (9.9) respectively.

$$r = \frac{x_D(x_F - x_B)}{x_F(x_D - x_B)} 100 \quad (9.2)$$

The values of  $x_B$  and  $x_D$  can be calculated as follows [3]:

$$x_D = \frac{Sx_B}{1 + x_B(S - 1)} \quad (9.3)$$

$$x_B = \frac{x_D}{S - x_D(S - 1)} \quad (9.4)$$

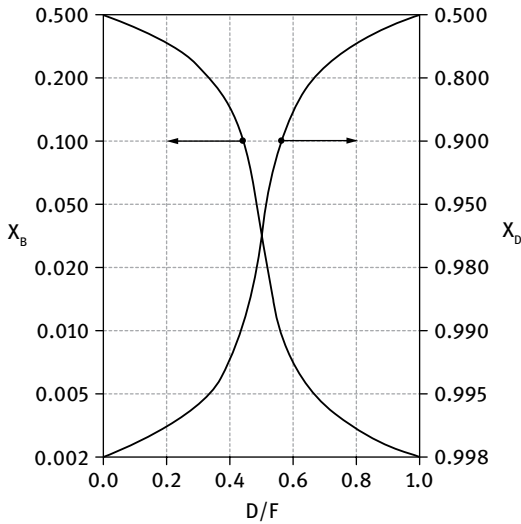
Equations (9.3) and (9.4) contain the *separation factor*  $S$ . It represents the ratio of light to heavy components between the distillate and waste [3]. For a binary distillation it can be expressed as:

$$S = \frac{\frac{x_D}{(1 - x_D)}}{\frac{x_B}{(1 - x_B)}} = \frac{x_D(1 - x_B)}{x_B(1 - x_D)} \quad (9.5)$$

The importance of separation factor in control issues will be discussed in more detail in chapter 9.3. It connects to the  $V/F$  flow rate ratio [3] – see relationship (9.6), where  $n$  stands for the number of trays,  $E$  for the tray efficiency, and  $\alpha$  for the average (mean) relative volatility.

$$\ln S = \frac{(nE)^{0.68} \alpha^{1.68} \ln \alpha}{3.5} \frac{V}{F} \quad (9.6)$$

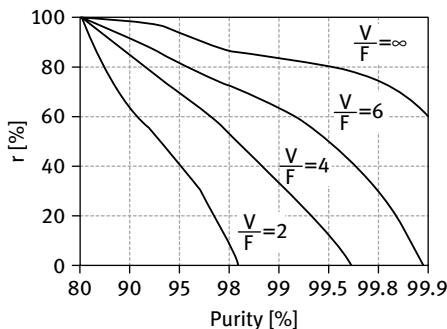




**Fig. 9.3:** Variation of the volatile species concentration in distillate and waste streams as a function of  $D/F$  ratio. The ordinates are expressed in logarithmic scales. (Adapted after [1])

The  $D/F$  flow rate ratio affects significantly the operation of a distillation column. An increase of  $D$  is followed by a decrease of its concentration  $x_D$  as well as by some enhancement of the recovery factor. Meanwhile,  $x_B$  within the bottom waste stream decreases along with  $D/F$  ratio; this causes a recovery increase. These facts are demonstrated by data presented in Fig. 9.3. It illustrates the variation of volatile species concentration in  $D$  and  $B$  streams, at various  $D/F$  values.

Another important parameter affecting distillation column performance is the  $V/F$  ratio. The vapor phase flow rate  $V$  is proportional to the heating power. For a constant  $F$ , the ratio  $V/F$  itself also reflects consumed heating power. Figure 9.4 illustrates the variation of recovery factor  $r$  against product purity at various  $V/F$  values, yet at constant  $x_F = 0.5$ . The curves were computed by using eqs. (9.2), (9.6), (9.38), and



**Fig. 9.4:** Recovery factor vs product purity at various  $V/F$  ratios and constant  $x_F = 0.5$ . The abscissa is expressed in logarithmic scale. (Adapted after [1])

(9.42), respectively. A strong decrease of  $r$  can be observed as purity increases regardless of heating power.

Data in Fig. 9.4 suggest that higher purity and recovery can be achieved if  $V/F$  ratio is kept high. However, some major issues arise when applying higher vapor phase stream flow rates  $V$ :

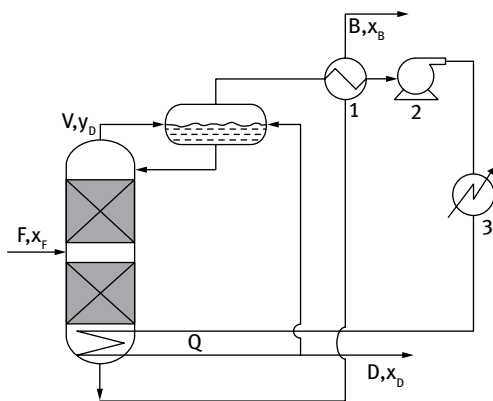
- High vapor flow rate  $V$  causes a decrease in tray efficiency;
- Extremely high  $V/F$  may lead to flooding, an unwanted phenomena;
- Higher vapor load requires higher heating loads, undesired especially when separation of temperature sensitive components is targeted;
- High vapor generation is usually coupled with high reflux rates. This leads to a significant increase of both heating and cooling demands which translates into an overall cost increase.

### 9.3 Lowering energy demand of distillation units

Operation of distillation columns requires significant energy both for cooling and heating. The over 40,000 devices operating in 2010 in the USA only demanded approximately 40%–60% of the total energy consumption of entire chemical and refining industry. It represented 19% of U.S. manufacturers as well as 6% of the total U.S. energy usage [4].

The worldwide energy consumption of distillation columns is obviously higher. Some examples of energy saving techniques employed in practice are:

- Improvement of tray efficiency by optimizing the vapor flow rate with the aid of proper control configuration [4];
- Growth of relative volatility by working at minimal pressure. Approximately 5% energy can be saved at night time and 25% in winter by only applying advanced atmospheric cooling (temperature drop) and therefore minimizing the working pressure of the column;
- Appropriate design and choice of the feeding tray; if the feeding concentration differs from that of the actual tray, the separation efficiency decreases;
- Applying internal heat integration; a heat pump system can be employed to use the condensation heat of the distillate for warming up the reboiler. Pinch analysis enables assessment of minimal heating and cooling demand as well as the extent in which it can be covered by appropriate heat integration. Figure 9.5 presents a simple internal heat pumping system, where 1 represents the heat exchanger that uses heat from the bottom stream; 2, the pump increasing the vapor stream pressure to rise both its temperature and boiling point; and 3, the supplementary external heat source. Other more complicated heat integration schemes are also encountered in practice. One example is the employment of the bottom stream heat for preheating operation of the feed [5].
- Inclusion of the distillation column in a plantwide heat integration system to minimize the energy demand of the whole process. This enables more flexibility for the heat reuse [6].

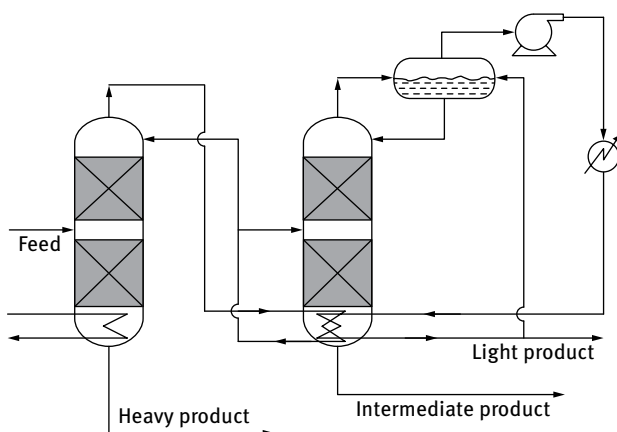


**Fig. 9.5:** Heat integration for a binary distillation column: an internal heat pumping system.

- Applying specific case heat integration for concurrent cascade distillation columns that separate ternary mixtures. Figure 9.6 illustrates such an example for which the heat used to boil up the first column is completely transferred to the second column. Here a condenser-reboiler heat pump is also employed.

The choice of a reliable energy reduction/saving technique/strategy is strongly related to cooling/heating equipment and its implementation costs. Smart solutions may demand higher initial investments, thus capital amortization rates have to be considered as well within the early design stage of distilleries.

Heat integration provides excellent opportunities to reduce overall energy demand. Meanwhile, it reduces process flexibility thus rises significant controllability concerns [7].



**Fig. 9.6:** Heat integration for concurrent cascade distillation columns.

The control of individual heat exchangers is well known, however control of a *heat exchanger network* (HEN) may be challenging. Such networks are not only of interest for thermally integrated distillation, but represent the core of any heat integrated plant. Mathisen et al. [8] presented some specific rules for bypass selection in HEN control. Glemmestad and Gundersen [9] developed a freedom degree analysis to assess whether operating HENs can be optimized. However, the most common HEN control strategies are rather simple:

- Maintaining outlet temperatures bypassing the flow of process-to-process heat exchangers as well as duties of process-to-utility heat exchangers at near setpoint values;
- Dividing flow rates using process stream splitters;
- Even simply keeping large material recycle streams that can store and carry energy, has been discussed as beneficial for cascades of heated systems such as the one illustrated by Fig. 9.6 [10].

## 9.4 General control of continuous distillation columns

### 9.4.1 Mass and energy balance imposed control issues

The design of an adequate ACS for a distillation column is based on the device's behavior analysis, that is on heat and mass balances. Figure 9.2 illustrates schematically a binary distillation column and the corresponding symbols of main operating characteristics.

In steady-state operation mode, the global mass balance is described by eq. (9.7). According to it, input mass of the feed is completely evacuated within the output distillate and bottom streams.

$$F = D + B \quad (9.7)$$

The mass balance for the valuable chemical species is expressed by eq. (9.8). Similarly to eq. (9.7), it states that the fed volatile compound is split between  $D$  and  $B$  streams.

$$F x_F = D x_D + B x_B \quad (9.8)$$

Replacing  $B$  from eq. (9.7) into eq. (9.8) leads to the  $D/F$  ratio as a function of concentrations:

$$\frac{D}{F} = \frac{x_F - x_B}{x_D - x_B} \quad (9.9)$$

Nevertheless,  $B$  or  $F$  can also be expressed from eq. (9.7) and other ratios, such as  $B/F$  or  $B/D$  may be obtained from eq. (9.8). Yet from the operator's point of view,  $D/F$  relates easier to the column efficiency.

**i** Example 9.1

A distillation column has a feed stream characterized by  $x_F = 0.5$ . The distillate is required  $x_D = 0.95$  and  $x_B$  within the bottom stream should not exceed 0.05. Let us calculate the  $D/F$  ratio that ensures the minimum required purity.

The ratio  $D/F$  can be obtained by means of eq. (9.9):

$$\frac{D}{F} = \frac{0.5 - 0.05}{0.95 - 0.05} = 0.5$$

Let us recalculate  $D/F$  if the feed concentration drops with 10%. Thus,  $x_{F_{new}} = 0.9 \cdot x_F$  and

$$\frac{D}{F} = \frac{0.5 \cdot 0.9 - 0.05}{0.95 - 0.05} = 0.44$$

The results show that a negative 10% disturbance demands a distillate flow rate decrease of 12% to maintain the prescribed distillate and bottom compositions.

Equation (9.9) yields the  $x_D = f(x_F)$  relationship (9.10) from which the distillate concentration can be computed quickly as a function of operational parameters. Derivation generates the  $x_D$  dependence on bottom concentration  $x_B$  – see (9.11).

$$x_D = x_B + \frac{F(x_F - x_B)}{D} \quad (9.10)$$

$$\frac{dx_D}{dx_B} = 1 - \frac{F}{D} \quad (9.11)$$

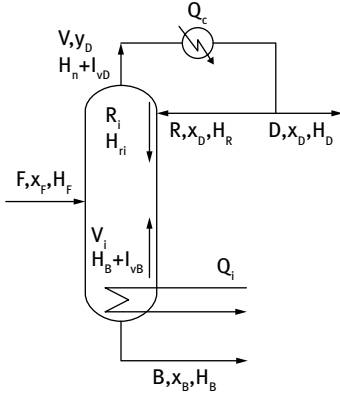
Equations (9.9)–(9.11) reveal that:

- the distillate concentration  $x_D$  increases linearly with the flow rate;
- a  $D = 0$  value results in  $x_B = x_F$ , thus operating columns at  $x_B > x_F$  is impossible.

The mechanism of evacuated flows  $D$ ,  $V$  and  $B$  influencing values of  $x_D$ ,  $x_B$  and  $y_D$  respectively, is related to both heat and mass balances.

For instance, the internal vapor flow rate  $V_i$  would ideally equal the external vapor flow rate  $V$  if the reflux would recirculate at the exact boiling point temperature of the tray to which it reenters the column (ideal case). In other words, neither vapor condensation nor liquid evaporation would occur. In reality, however, the reflux temperature is usually lower than that of the tray into which it is fed. Thus, a part of the internal vapor flow is condensed while the reflux stream is heated up to the boiling point. Consequently, the internal and external vapor fluxes have different flow rates and the condensed internal vapors will contribute to the internal reflux by increasing it.

The interconnections between mass and energy balances are schematically illustrated in Fig. 9.7, where subscript  $i$  symbolizes internal flux,  $H$  stands for the liquid enthalpy,  $l_v$  is the latent evaporation heat, index  $n$  stands for the  $n^{\text{th}}$  tray,  $H_n$  is the



**Fig. 9.7:** Interconnection between energy and mass balances. Schematic representation for a distillation column.

enthalpy at boiling point on tray  $n$ , and  $Q_c$  represents the amount of heat removed by the condenser.

The mass balance at the top of the distillation column (see Fig. 9.7) is:

$$V_i + R = V + R_i \quad (9.12)$$

The left-hand side of eq. (9.12) describes input mass, whereas the right-hand side stands for the output. The corresponding energy balance is described by eq. (9.13),

$$V_i(H_n + l_{vD}) + RH_R = V(H_n + l_{vD}) + R_i H_{Ri} \quad (9.13)$$

By combining information in eqs. (9.12) and (9.13), the internal reflux may be expressed as a function of external reflux in (9.14). It confirms the assumption that the subcooled external reflux (when  $H_n > H_R$ ) leads to an increase of internal reflux.

$$R_i = R \left( 1 + \frac{H_n - H_R}{l_{vD}} \right) \quad (9.14)$$

The *specific enthalpy* is defined in eq. (9.15) as a ratio between the flow's heat  $H$  and boiling point enthalpy  $H_n$  difference and the vaporization heat  $l_v$ :

$$q = \frac{H - H_n}{l_v} \quad (9.15)$$

Since the liquid temperature cannot exceed the boiling point,  $q \leq 0$ . By replacing (9.15) in (9.14) a simpler form is obtained.  $q_R$  stands for the specific enthalpy of the reflux.

$$R_i = R(1 - q_R) \quad (9.16)$$

Equation (9.16) reveals that the internal reflux cannot be lower than the external reflux flow rate.

The global heat balance is described by eq. (9.17), where  $Q_i$  stands for the heat entered through the reboiler and  $Q_c$  stands for the heat evacuated in the condenser.

This equation shows the way energy consumption depends on flow rates  $F$ ,  $D$ , and  $B$  as well as on the heat amount these carry.

$$FH_F + Q_i = Q_c + DH_D + BH_B \quad (9.17)$$

Cooling and heating demands can be calculated as a function of internal and external vapor flow rates:

$$Q_i = V_i l_{vB} \quad (9.18)$$

$$Q_c = V(H_n + l_{vD}) - VH_D \quad (9.19)$$

From (9.17)–(9.19), the internal vapor flow rate can be expressed as:

$$V_i = \frac{V(H_n + l_{vD}) - VH_D + DH_D + BH_B - FH_F}{l_{vB}} \quad (9.20)$$

Equation (9.20) can be reformulated by introducing the external reflux, a much more easier measured and controlled parameter. Hence,

$$V_i = \frac{F(H_n + l_{vD} - H_F) + R(H_n + l_{vD} - H_D) + B(H_n + l_{vD} + H_B)}{l_{vB}} \quad (9.21)$$

Relationship (9.21) shows that with raising reflux, the internal vapor flow rate also elevates. Thus, the process heating energy demand increases.

The global heat balance (9.17) may be expressed also as a function of specific enthalpies of individual flows:

$$Fq_F + V_i \frac{l_{vB}}{l_{vD}} = V - Rq_R + Bq_B \quad (9.22)$$

To achieve constant distillate purity, the  $D/F$  ratio has to be varied independently. Otherwise,  $D/F$  values depend on both  $R$  and  $V$ . Thus, small variations of  $R$  and/or  $V$  for columns operating at high  $R/D$  generates significant variations of  $D/F$ . Moreover,  $V$  cannot be controlled independently as, according to equation (9.23), derived from (9.22), it is a function of  $V_i$ ,  $R$ ,  $q_R$ ,  $F$  and  $q_F$ .

$$V = V_i \frac{l_{vB}}{l_{vD}} + Fq_F + Rq_R - Bq_B \quad (9.23)$$

If  $D$  is not directly controlled but results as the difference  $D = V - R$ , then eq. (9.23) becomes

$$D = V_i \frac{l_{vB}}{l_{vD}} + Fq_F - R(1 - q_R) - Bq_B \quad (9.24)$$

that yields when using the global mass balance (9.7):

$$D = \frac{V_i \frac{l_{vB}}{l_{vD}} + F(q_F - q_B) - R(1 - q_R)}{1 - q_B} \quad (9.25)$$

The extent to which variations of  $F$  and  $R$  affect the distillate flow rate  $D$  can be assessed from the differential of eq. (9.25). Relationship (9.26) is obtained under the assumption that the bottom-specific enthalpy  $q_B$  does not vary significantly since the outlet temperature is near the boiling point. Therefore,  $dq_B = 0$ .

$$dD = \frac{V_i \frac{l_{vB}}{l_{vD}} + dF(q_F - q_B) + Fdq_F - dR(1 - q_R) + Rdq_R}{1 - q_B} \quad (9.26)$$

Division by  $D$  turns eq. (9.26) into eq. (9.27), the expression of the relative variation of distillate rate.

$$\frac{dD}{D} = \frac{\frac{V_i}{D} \frac{dV_i}{V_i} \frac{l_{vB}}{l_{vD}} + \frac{dR}{R} \frac{R}{D} (1 - q_R) + \frac{R}{D} dq_R + \frac{dF}{F} \frac{F}{D} (q_F - q_B) + \frac{F}{D} dq_F}{1 - q_B} \quad (9.27)$$

A numeric analysis of the above eq. (9.27) reveals that a 1% variation of  $V_i$  for a column working at  $R/D = 10$  generates an 11% variation of  $dD/D$ . If the feed-specific enthalpy  $q_F$  changes with 1%, the variation of  $dD/D$  is 2% for  $F/D = 2$  and 5% for  $F/D = 5$ , respectively. These observations suggest that for columns working at high  $R/D$ , the  $R$  has to be controlled by a level controller of the reflux reservoir, and  $D$  should be controlled independently.

From all above it may be concluded that:

1. The  $V/F$  ratio has to be kept constant for economic (energy saving) reasons;
2. The  $D/F$  ratio has to be kept constant to ensure a desired distillate concentration  $x_D$ ;
3. The bottom flow concentration  $x_B$  requires additional equations to all above presented.

The separation factor  $S$  was already defined in equation (9.5). It also respects condition (9.28), where  $\alpha_m$  stands for *average (mean) relative volatility*,  $E$  for tray efficiency and  $n$  for the number of trays, respectively. The minimum value  $S = 1$  corresponds to the case for which there is no separation; in other words, the bottom and distillate concentrations are equal ( $x_B = x_D$ ).

$$1 \leq S \leq \alpha_m^{-nE} \quad (9.28)$$

The relative volatility  $\alpha$  of components is expressed as [1, 2]:

$$\alpha_{ij} = \frac{y_i}{x_i} \frac{P_j^0}{y_j} = \frac{P_i^0}{P} = \frac{K_i}{K_j} = \frac{P_i^0}{P_j^0} \quad (9.29)$$

where  $P_i^0$  and  $P_j^0$  are the partial pressures of  $i$  and  $j$  pure components and  $P$  is the total pressure in the column.  $K_{i,j}$  stands for the equilibrium ratio at vaporization of



$i$  and  $j$  components (the so-called  $K$  factor):  $K_i = \frac{y_i}{x_i} = \frac{P_i^0}{P}$  and  $K_j = \frac{y_j}{x_j} = \frac{P_j^0}{P}$ . Equation (9.29) demonstrates that for any  $\alpha = 1$ , separation is impossible because the components are identically volatile.

For a binary mixture, the relative volatility (9.29) will express as:

$$\alpha = \frac{y(1-x)}{x(1-y)} \quad (9.30)$$

As a rule, the vapor pressure  $P$  increases with temperature as described by the Antoine equation (9.31), in which  $A$ ,  $B$ , and  $C$  are material-specific constants.

$$\ln P = A - \frac{B}{T + C} \quad (9.31)$$

The following two equations can be derived from eq. (9.29):

$$\alpha_{ij} = \frac{P_i^0}{P_j^0} \leftrightarrow \ln \alpha_{ij} = \ln \frac{P_i^0}{P_j^0} \quad (9.32)$$

$$\ln \alpha_{ij} = \ln P_i^0 - \ln P_j^0 \quad (9.33)$$

Relationships (9.33) and (9.31) show that the relative volatility decreases when temperature rises.

Due to the pressure drop along the column,  $\alpha$  varies from tray to tray. However, for basic modeling and design purposes, it can be expressed as the arithmetic average of the distillate and bottom values – see eq. (9.34). Meanwhile, for stronger variations, the geometric mean is applied – see eq. (9.35).

$$\alpha_m = \frac{\alpha_B + \alpha_D}{2} \quad (9.34)$$

$$\alpha_m = \sqrt{\alpha_D \alpha_B} \quad (9.35)$$

The tray efficiency  $E$  is defined in eq. (9.36) as the ratio between theoretical ( $n_t$ ) and practical ( $n_p$ ) tray numbers necessary to achieve a desired separation.

$$E = \frac{n_t}{n_p} \quad (9.36)$$

A typical value of tray efficiency is 0.7. If the tray efficiency decreases due to high vapor flow caused stripping, the product purity cannot be considerably improved by increasing the reflux ratio. In this situation, the achievable concentration limit of the column has been reached. Meanwhile, if the tray efficiency is small due to low vapor flow caused weeping, the energy consumption cannot be further reduced (lower  $V$  is not applicable). In this case, the lowest energy consumption has been reached.

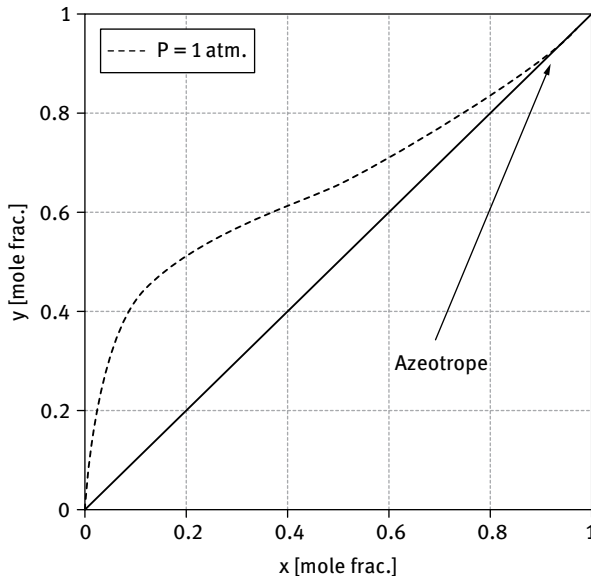


Fig. 9.8: Vapor-liquid equilibrium curve for the ethanol-water mixture.

For a binary mixture, the liquid-vapor phase equilibrium is generally described as  $y = f(x)$  relationship, such as

$$y = \frac{x\alpha_m^{n_i}}{1 + x(\alpha_m^{n_i} - 1)} \quad (9.37)$$

Figure 9.8 presents the vapor-liquid equilibrium curve for the ethanol-water mixture at atmospheric pressure. It represents the gas phase volatile ethanol concentration  $y$  versus its liquid phase concentration  $x$  in the mixture. The difference at any given  $x$  between diagonal (where  $x = y$ ) and equilibrium curve values of  $y$  stands for the maximal concentration difference that can be achieved by one distillation unit having a 100% tray efficiency. The curve and diagonal exhibit closer  $y$  values at  $x$  near to 0 or to 1. This translates into more trays needed to achieve the same concentration difference. The curve crosses the diagonal in the so-called azeotropic point ( $y = x$ ) corresponding to the azeotrope mixture. It is also known as *constant boiling mixture* because when boiling it, the vapor and liquid phases have identical compositions and therefore cannot be separated by means of simple distillation. As a result, at constant pressure distillation this  $y$  equaling  $x$  value cannot be exceeded.

Equation (9.38) describes the vapor-liquid equilibrium (9.37) in terms of the separation factor  $S$ . It yields a similar relationship to that of Fenske's equation (9.39) [11].

$$x_D = \frac{x_B S}{1 + x_B(S - 1)} \quad (9.38)$$

$$y = \frac{x\alpha}{1 + x(\alpha - 1)} \quad (9.39)$$

Combining eqs. (9.9) and (9.38) gives the mole fraction of interest  $x_D$  (the more volatile species at the column top).

$$x_D = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \quad (9.40)$$

The terms in eq. (9.40) are:

$$\begin{aligned} a &= \frac{D}{F}(S - 1) \\ b &= - \left[ \left( \frac{D}{F} + x_F \right) (S - 1) + 1 \right] \\ c &= Sx_F \end{aligned} \quad (9.41)$$

Further, the value of  $x_B$  can be obtained using  $x_D$

$$x_B = \frac{x_D}{S - x_D(S - 1)} \quad (9.42)$$

The following expression of  $S$  was proved to be valid if the feeding stream  $F$  is introduced into the column into the optimal tray ( $q_F = 0$ ) [12]:

$$S = \left( \frac{\alpha_m}{\sqrt{1 + \frac{D}{Rx_F}}} \right)^{nE} \quad (9.43)$$

Otherwise, if  $q_F \neq 0$ , above (9.43) becomes:

$$S = \left( \alpha_m \sqrt{1 - \frac{1 - q_F + \frac{R}{D}}{\left(1 + \frac{R}{D}\right)\left(1 - q_F + \frac{x_F R}{D}\right)}} \right)^{nE} \quad (9.44)$$

From eq. (9.44), the  $D/R$  ratio can be expressed as a function of separation factor  $S$  and distillate concentration  $x_D$  – see eq. (9.45). Thus,  $D/R$  values can be calculated that keep both  $S$  and  $x_D$  at desired values.

$$\frac{D}{R} = x_F \left[ \left( \frac{\alpha_m}{S^{1/nE}} \right)^2 - 1 \right] \quad (9.45)$$

Shinskey [11] showed that the separation factor can also be linked to the  $V/F$  – see equation (9.6).

### **i** Example 9.2

A binary mixture of equal 0.5 mole fractions is subjected to separation in a distillation column that has 30 theoretical trays. The average relative volatility is 1.72. If the tray

efficiency is  $E = 1$ , let us calculate the  $V/F$ ,  $D/F$  and  $R/D$  ratios that ensure  $x_D = 0.95$  and  $x_B = 0.05$ , respectively. Let us also calculate the recovery factor.

Equation (9.5) can be used to calculate the separation factor from given  $x_D = x_B = 0.05$ :

$$S = \frac{x_D(1-x_B)}{x_B(1-x_D)} = \frac{0.95 \times (1-0.05)}{0.05 \times (1-0.95)} = 361$$

Further, eq. (9.45) yields the  $D/R$  ratio, whereas  $D/F$  is computed from eq. (9.9):

$$\frac{D}{R} = x_F \left[ \left( \frac{\alpha_m}{S^{1/nE}} \right)^2 - 1 \right] = 0.5 \left[ \left( \frac{1.72}{361^{1/30}} \right)^2 - 1 \right] \approx 0.5$$

$$\frac{D}{F} = \frac{x_F - x_B}{x_D - x_B} = \frac{0.5 - 0.05}{0.95 - 0.05} = 0.5$$

Rearrangement of eq. (9.6) permits calculation of the  $V/F$  ratio:

$$\frac{V}{F} = \frac{3.5 \ln S}{(nE)^{0.68} \alpha_m^{1.68} \ln \alpha_m} \approx 1.5$$

while the recovery factor is obtained from eq. (9.2):

$$r = \frac{x_D(x_F - x_B)}{x_F(x_D - x_B)} 100 = \frac{0.95 \times (0.5 - 0.05)}{0.5 \times (0.95 - 0.05)} 100 = 95\%$$

The result shows that 95% of the volatile component fed into the column is evacuated within the distillate stream.

### Example 9.3



Let us assume that the distillation column presented in Example 9.2 is operated at the same heating power ( $V/F = \text{constant}$ ), but the  $D/F$  ratio modifies to 0.48. Let us assess the changes in both concentrations and recovery factor.

The  $V/F$  ratio can be expressed based on the mass balance at the column's top. From it, the  $D/R$  ratio can be further obtained:

$$\frac{V}{F} = \frac{R_i}{F} + \frac{D}{F} \leftrightarrow \frac{V}{F} = \frac{D}{F} \left( \frac{R}{D} + 1 \right)$$

$$\frac{D}{R} = \frac{\frac{D}{F}}{\frac{V}{F} - \frac{D}{F}} = \frac{0.48}{1.5 - 0.48} = 0.47$$

The result shows that the  $D/R$  value decreased as compared to that in Example 9.2. This suggests that the separation factor has to increase. From eq. (9.43), the value of  $S$  is accessible and its value proves to be indeed higher.

$$S = \left( \frac{\alpha_m}{\sqrt{1 + \frac{D}{R x_F}}} \right)^{nE} = \left( \frac{1.72}{\sqrt{1 + \frac{0.47}{0.5}}} \right)^{30} = 560.8$$

Equations (9.40) and (9.41) permit calculation of  $x_D$

$$\begin{aligned} a &= \frac{D}{F}(S - 1) = 0.48 \times (560.8 - 1) = 268.7 \\ b &= -\left[ \left( \frac{D}{F} + x_F \right) (S - 1) + 1 \right] = -[(0.48 + 0.5) \times (560.8 - 1) + 1] = -549.6 \\ c &= 0.5 \times 560.8 = 280.4 \\ x_D &= \frac{549.6 - \sqrt{549.6^2 - 4 \times 268.7 \times 280.4}}{2 \times 268.7} = 0.974 \end{aligned}$$

Thus, the higher separation factor results in a higher distillate concentration. The value of  $x_B$  is:

$$x_B = \frac{x_D}{S(1 - x_D) + x_D} = \frac{0.974}{560.8 \times (1 - 0.974) + 0.974} = 0.062$$

The bottom concentration also increased. This is due to the decrease of distillate flow-rate  $D$ . The recovery factor is:

$$r = \frac{x_D(x_F - x_B)}{x_F(x_D - x_B)} 100 = \frac{0.974 \times (0.5 - 0.062)}{0.5 \times (0.974 - 0.062)} 100 = 93.65\%$$

Consequently, higher distillate concentration leads to lower recovery.

### **i** Example 9.4

Let us repeat the calculations for  $D/F = 0.5$ , but with  $V/F$  (heating energy) increasing to 1.8.

The  $D/R$  ratio can be calculated as follows:

$$\frac{D}{R} = \frac{\frac{D}{F}}{\frac{V}{F} - \frac{D}{F}} = \frac{0.5}{1.8 - 0.5} = 0.385$$

The separation factor is yield by eq. (9.43) and results in the highest value of all examples so far.

$$S = \left( \frac{\alpha_m}{\sqrt{1 + \frac{D}{R x_F}}} \right)^{nE} = \left( \frac{1.72}{\sqrt{1 + \frac{0.385}{0.5}}} \right)^{30} = 2219.5$$

The distillate concentration is computed as above:

$$a = \frac{D}{F}(S - 1) = 0.5 \times (2219.5 - 1) = 1109.3$$

$$b = -\left[\left(\frac{D}{F} + x_F\right)(S - 1) + 1\right] = -[(0.5 + 0.5) \times (2219.5 - 1) + 1] = -2219.5$$

$$c = 0.5 \times 2219.5 = 1109.8$$

$$x_D = \frac{2219.5 - \sqrt{2219.5^2 - 4 \times 1109.3 \times 1109.8}}{2 \times 1109.3} = 0.981$$

It is obvious that the highest separation factor  $S$  resulted in the highest value of  $x_D$ . Thus, increasing  $V/F$  ratios favors distillate concentration.

The value  $x_B$  of the bottom concentration is:

$$x_B = \frac{x_D}{S(1 - x_D) + x_D} = \frac{0.981}{2219.5 \times (1 - 0.981) + 0.981} = 0.0000084$$

The results show that a higher heating power almost completely stripes the volatile compound from the reboiler. The recovery factor is:

$$r = \frac{x_D(x_F - x_B)}{x_F(x_D - x_B)} 100 = \frac{0.981 \times (0.5 - 0.0000084)}{0.5 \times (0.981 - 0.0000084)} 100 = 99.99\%$$

and indicates that practically the whole quantity of the feed volatile compound is evacuated with the distillate.

It can be concluded that an increasing heating power ( $V/F$  ratio) has a positive effect on both the distillate concentration  $x_D$  and recovery factor  $r$ . Nevertheless, this requires additional heating (costs) – see sections 9.1 and 9.3. A further increase of  $V/F$  may lead to evaporation of the less volatile species, which will, in turn, decrease  $x_D$ .

### 9.4.2 Control solutions

When summarizing all process characteristics and economic aspects discussed so far, the following general control strategy guidelines are drawn for a distillation column:

- An ACS is required to keep  $V/F$  ratios constant by means of heating power.
- Because at constant feed composition  $x_F$ , the variations of  $x_D$  and  $x_B$  are related to those of flow rate  $F$ , a concentration control system is required at the column's top. It should manipulate independently the distillate flow rate  $D$  to keep  $D/F$  constant.
- The reflux vessel requires a level control system. This should control the  $D/R$  ratio and thus ensure the mass balance closure ( $V = D + R$ ) within this column region.

- The column bottom requires a level control system. This settles the  $B/F$  ratio thus enabling bottom mass balance to close – see eq. (9.46).

Figure 9.9 presents a very general control scheme of a distillation column based on above statements. Various other control solutions are employed in practice. Some of these will be briefly presented and discussed.

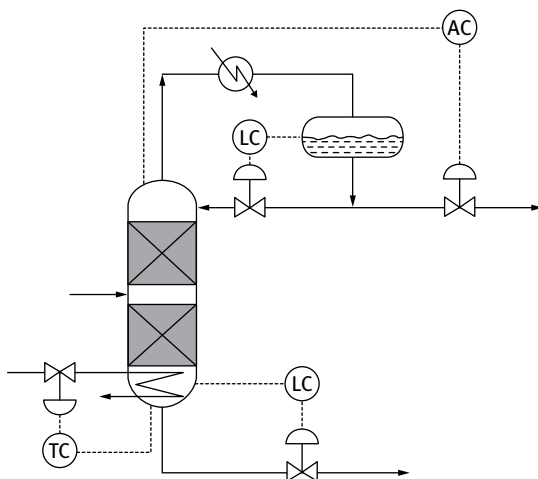


Fig. 9.9: The general control scheme of a distillation column.

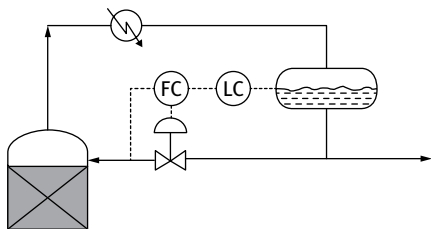
### Level control

Various methods can be applied to ensure mass balance closure in both bottom and reflux vessels. The bottom mass balance requires the sum of generated vapor flow rate  $V_i$  and bottom flow rate  $B$  to be equal to the internal reflux rate entering the reboiler  $R_{i,B}$ :

$$R_{i,B} = B + V_i \quad (9.46)$$

If  $B$  is low, the bottom liquid level control can be achieved by manipulating the heating energy and hence controlling the evaporation rate (it does not introduce significant deviation of vapor flow as compared to its prescribed value). The temperature control can be solved through the bottom flow rate  $B$  since the boiling point is a function of composition. This strategy is applied to avoid the possible inverse response of the reboiler (see also chapter 9.5 and Fig. 9.14).

Level control in the reflux vessel can be achieved by a cascade control (see Fig. 9.10) which ensures fast response to variations of  $D$ . Because the majority of columns are operated at high(er)  $R/D$  values, this solution is handy since low disturbances in  $D$  require significant variations of  $R$  to maintain the prescribed  $R/D$ . The same dynamic effect can be achieved when manipulating flow rate  $D$ .



**Fig. 9.10:** Level control in a reflux vessel by means of a cascade control system.

### Concentration control

A wide palette of instrumentation is in use for the measurement of liquid concentrations [2]. Some of the more advanced techniques involve high (or ultra-high) pressure liquid chromatography, gas chromatography, viscosity or density based methods, UV/VIS and IR spectroscopy, etc. However, the cost of these devices limits their wide on-stream implementation. In contrast, low-cost thermometers are widely used in distillation columns for concentration measuring purposes. These are based on the fact that boiling point is a function of pressure and composition.

The gas phase mole fraction of a volatile component in ideal mixtures is expressed as a function of its vapor pressure  $P_i$  at a given pressure  $P$  by eq. (9.47). The Raoult equation (9.48) links the partial pressure of the pure component  $P_i^*$  to its actual concentration  $x_i$ :

$$y_i = \frac{P_i}{P} \quad (9.47)$$

$$P_i = P_i^* x_i \quad (9.48)$$

If for a binary A – B mixture, species A is more volatile, then the combination of eqs. (9.47) and (9.48) yields for A:

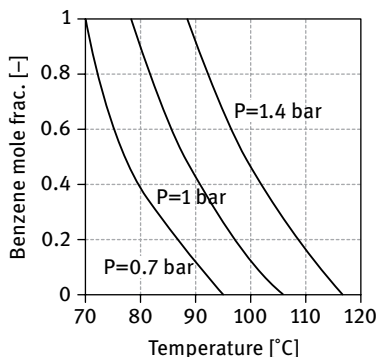
$$x_A = \frac{P - P_B}{P_A - P_B} \quad (9.49)$$

The vapor pressure is known to depend exponentially on temperature. Thus, boiling point, pressure, and composition can be correlated. Figure 9.11 presents such an example for the benzene-toluene mixture, where temperature dependencies are almost linear.

Such correlations can translate a simple temperature measurement into an estimate concentration value. However, they ought to be used with caution because in some cases, the boiling point difference of pure components is rather small and falls within the measurement error domain of the thermometer. These errors can further propagate into the control system.

For example, a 1% molar concentration variation in a *n*-butane/*iso*-butane mixture generates just 0.25°C boiling point variation. A similar deviation is yield by ~ 140 mmHg pressure increase or ~0.6% *iso*-pentane impurity. A very usual 0.5% precision class thermoresistance with 40°C maximal temperature has a 0.2°C





**Fig. 9.11:** Correlation among boiling point, pressure and composition for the benzene-toluene mixture. (Adapted after [1])

designed measuring error. It is obvious that this thermometer is not recommended for boiling point based concentration determinations in this case.

Nevertheless, thermometers are widely used in distillation columns. To shorten their response time, they should always be merged into the liquid. A sensitive measurement means that any disturbance is sensed quickly and reliably. Therefore, it is recommended to install the thermometer into the tray having the maximal temperature gradient within the column's vertical temperature profile. This places it usually at the top of the column, right under the reflux tray (see Fig. 9.17.b). However, measuring temperature and thus estimating concentration for another tray than for the one that requires control might degrade control performance. Yet, measurements directly on the controlled tray will result in more accurate concentration values but a more vulnerable to disturbances control system.

### Pressure control

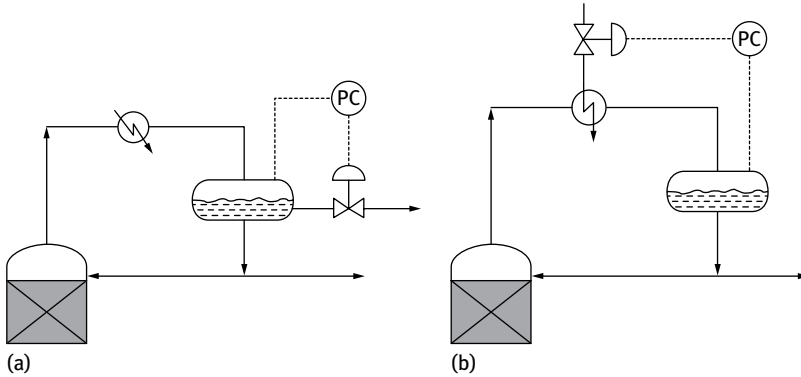
Because the boiling point depends on pressure, the latter has to be carefully controlled. Pressure control is carried out by taking into account both the nature and properties of the gas that generates overpressure. Two limiting cases generally occur:

- Overpressure caused only by noncondensing gases when a gas purge is added to the reflux vessel – see Fig. 9.12.a.
- Overpressure caused only by condensing gases when the coolant flow rate is manipulated – see Fig. 9.12.b. The heat transfer area of the condenser can also be varied by its partial flooding.

### Recommendations for control system configuration

The conclusion of this chapter is that adequate distillation control requires  $R/D$ ,  $V/F$ , and  $D/F$  ratios, reboiler level, and working pressure control to ensure product purity specifications and reduce process energy demand.

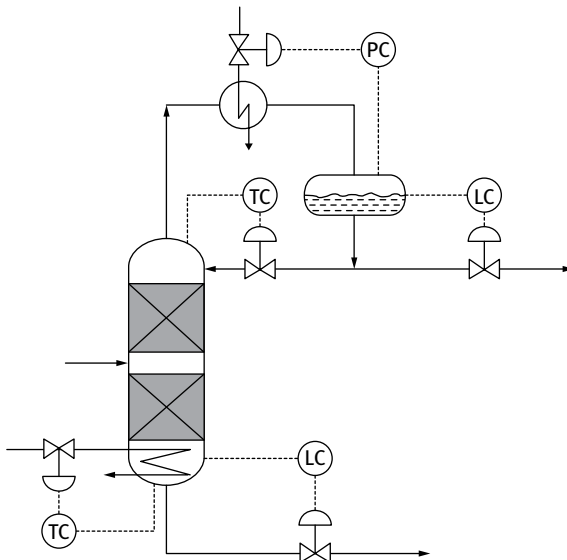
Numerous debates are described by the literature [3, 13] with regard to suitable control configuration for distillation columns. Yet all authors agree on the necessity



**Fig. 9.12:** (a) Pressure control by means of a gas purge. (b) Pressure control by coolant flow rate manipulation.

of controlling bottom and distillate concentrations by manipulating the  $R$ ,  $V$ ,  $D$  and  $B$ , respectively. Two of these flows are used to control concentrations and the other two for mass balance closure and recovery maximization. These four values can be paired in  $C_4^2 = 6$  ways. Since  $D$  and  $B$  cannot be varied simultaneously because they are linked *via* global mass balance (9.7), only 5 possible pairings remain to control concentrations.

Moreover, if the ratios ( $R/V$ ,  $R/D$ ,  $R/B$ ,  $V/D$ ,  $V/B$ , and  $D/B$ ) are also considered as manipulated variables, a total of  $C_{10}^2 = 45$  possible pairings result. Since,  $D$ ,  $B$ , and  $D/B$  cannot be paired, that leaves 42 possibilities.



**Fig. 9.13:** Control structure configuration of a distillation column based on the  $V-R$  pairing as manipulated variables.

Selecting adequate manipulated variables from these numerous possibilities, hence designing a well-performing control system, is strongly related to the process requirements and operating conditions. Usually, the Bristol [3] method is used (see also chapter 5); it involves the matrix of relative gains which highlights the dynamic interactions between individual control loops. However, employment of the  $V - R$  couple is most widely applied in practice (see Fig. 9.13.).

## 9.5 Control issues of continuous distillation column dynamics

Chapter 9.4 focused on steady-state operation of distillation columns. Their dynamic behavior, however, brings about significant delays that can lead to undesired instability of the ACSs.

Any distillation column is composed of a given number of mass transfer units (trays). Thus, overall mass transfer presents capacitive dynamic behavior and the distillation column behaves like a series of capacitive elements. Yet, distinction is to be made between hydrodynamic and composition dynamics.

The trays and the reboiler contain both vapor and liquid. However, the liquid mass is with some orders of magnitude higher than that of the vapor phase. Hence, the latter may be neglected when discussing column dynamics. Time delays may appear when the liquid phase is transferred from one tray to the one positioned underneath it. Such mass transfer occurs when increased internal flow rates  $R_i$  cause liquid accumulation and crossing of the tray's weir.

For a perpendicular-shaped weir, the liquid flow rate leaving the tray is given by the Francis equation below, where  $l$  is the weir width and  $h$  stands for the liquid level exceeding it [14, 15]:

$$f_i = 1.873lh^{\frac{3}{2}} \quad (9.50)$$

The retained liquid volume within the tray is described by eq. (9.51), where  $f$  represents the evacuating flow rate over the weir.

$$V_T = Ah = A \left( \frac{f}{1.873l} \right)^{\frac{2}{3}} \quad (9.51)$$

The tray's hydraulic time constant  $T_h$  is defined as the variation of retained volume with evacuating flow rate. After doing the math,  $T_h$  is written as:

$$dT_h = \frac{dV_T}{df} = \frac{0.442A}{l^{\frac{2}{3}}f^{\frac{1}{3}}} \quad (9.52)$$

When talking about hydraulic dynamics of a distillation column, it is important to mention the potential inverse response [2]. If the reboiler level is controlled by means of heating power, then increasing it may result in a quick concentration drop of the volatile compound. In some mixtures, this entails a similarly fast mixture density

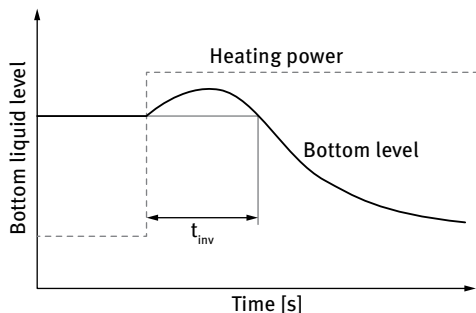


Fig. 9.14: Inverse response of bottom liquid level to increased heating power.

decrease. This inherently leads to an increased holdup volume which finally causes the *inverse response*. Naturally, due to the increased vapor generation rate, after a certain time the reboiler volume will start to decrease again (see Fig. 9.14). This time elapse is called *inversion time*  $t_{inv}$  and can be as high as some minutes or hours, in other words not negligible.

A rare consequence of high inversion times is the controller's failure in regulating the reboiler liquid level. In such cases, level control is achieved *via* bottom flow-rate  $B$ . If, however  $B$  is too low, level control must be carried out by means of reflux flow rate. This is a highly undesirable situation, since the system's response in mass transport comes after the maximal possible time delay, which is the sum of hydraulic delays of all trays.

Variation in liquid and vapor flow rates affect tray compositions. Thus, concentration variation rate depends on flow rates as well as on liquid amounts beholden by the tray. Therefore, concentration dynamics is slower than hydraulic dynamics.

The dynamics of a packed bed distillation column that separates methanol-water mixtures has been assessed [12] by using Raschig rings of 25-mm diameter. The experimental results were processed by means of following equations:

$$T_{C,ij} = \frac{H_C}{R + (R + D) \frac{P_{ij}}{P}} \quad (9.53)$$

$$T_{S,ij} = \frac{H_S}{(R + D) + (W - F - R) \frac{P_{ij}}{P}} \quad (9.54)$$

where

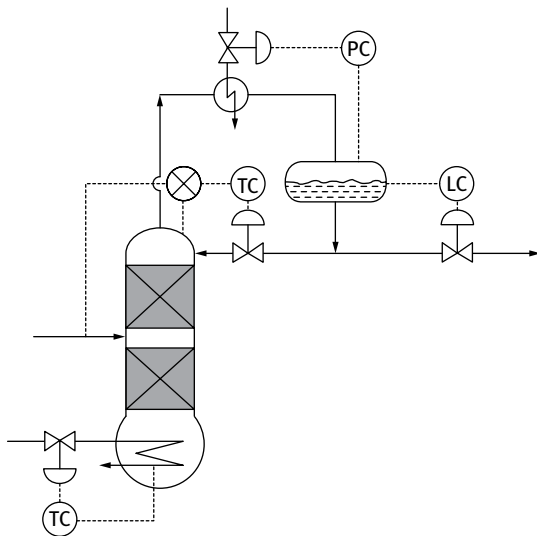
- $T_C$  and  $T_S$  are the time constants of the rectifying and stripping sections, respectively;
- $H_C$  and  $H_S$  are the liquid holdups within the rectifying and stripping sections, respectively. Their values can be calculated by means of the Schulmann, Ulrich, Weils, and Proulx relationships [16].
- $P_{ij}$  is the partial pressure of  $i$  component at  $j^{\text{th}}$  packed transfer unit temperature.

The time lapse of the response to a certain disturbance depends on the nature of disturbing parameter. A disturbance of feed flow rate  $F$  can result in up to 8 hours delay in the response of the distillate concentration. Fig. 9.15 illustrates a feedforward control loop used to attenuate such kind of disturbances by adjusting flow rate  $R$ . Meanwhile, a distillate flow rate  $D$  disturbance causes only a 15-minute delay in the value of  $x_D$ . The time constant of a tray is of about 20 minutes and does not depend on the disturbance's location ( $F$  or  $D$ ).

High-purity distillation columns [10] show three time scales for their dynamic behavior:

- Fast time scale for both the condenser and reboiler heat duties. The coolant and steam flows are manipulated variables that regulate temperatures of reboiler and condenser.
- Intermediate time scale for the compositions and holdups of individual stages. Manipulated variables are the large reflux rate  $R$  and the vapor flow rate  $V$ .
- Slow time scale for global input-output column behavior. The outputs are product purity and overall material balance. The manipulated inputs (for a model-based nonlinear controller) are the small bottom and distillate flows as well as the set-points of level controllers (a total of four variables: two flow rates and two references for the two level regulators, the bottom and the reflux vessel).

Each distillation system may require individual proper control strategies such as feedforward, cascade or model predictive control (MPC) [17–19]. The latter is also able to incorporate process nonlinearity aspects (see also chapter 2). State estimators ensure



**Fig. 9.15:** A feedforward control loop that reduces the effect of time delays occurring in the product quality when the feed flow rate is disturbed.

good quality control even under plant-model mismatch operating conditions and disturbances [20].

Chapter 2 explains that MPC performance strongly relates to the process model. The transfer functions of process dynamics, eqs. (9.53) and (9.54) are not able to describe the start-up and shut-down conditions. Because first-principle models describe better process dynamics, MPC applications often require their employment. Such a dynamic model (see also Fig. 9.17) for the pilot-scale continuous distillation column of ethanol-water mixtures illustrated by Fig. 9.18 was developed and validated by Agachi et al. [21, 22].

Model calibration [21, 23] was carried out by defining and solving a process optimization problem. The objective function is the sum squared error (SSE) described by eq. (9.55), where  $DV$  denotes the vector of decision variables defined in eq. (9.56).

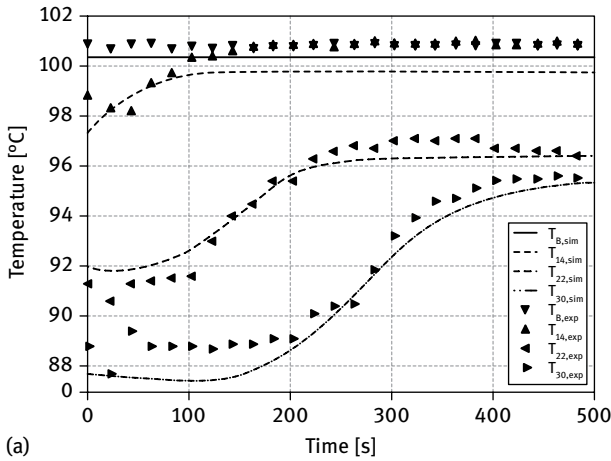
$$SSE(DV) = \sum_{i=1}^N \sum_{j=1}^P (T_{i,j, sim} - T_{i,j, exp})^2 + w \sum_{i=1}^N (x_{i, sim} - x_{i, exp})^2 \stackrel{!}{=} \min \quad (9.55)$$

$$DV = [E_r, E_s, h_r, h_s, Q_b, Q_c] \quad (9.56)$$

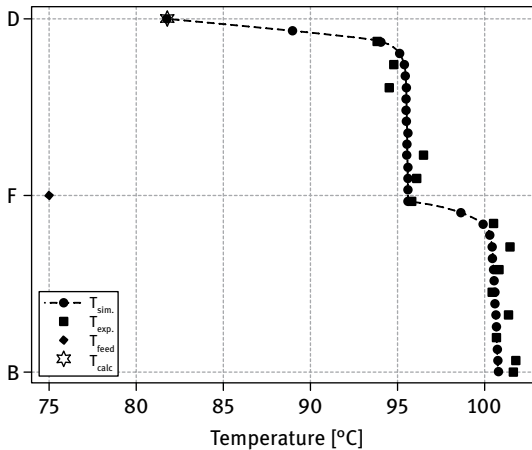
The decision variables within eq. (9.55) are the tray efficiency of the rectifying ( $E_r$ ) and stripping ( $E_s$ ) sections respectively, the liquid level of the trays within the rectifying ( $h_r$ ) and stripping ( $h_s$ ) sections as well as the heat losses of both bottom ( $Q_b$ ) and condenser ( $Q_c$ ). The latter were included because neither the bottom nor the condenser is thermally insulated (see Fig. 9.16).



**Fig. 9.16:** Pilot-scale distillation column (*Illudest*) at the Faculty of Chemistry and Chemical Engineering, Babeş-Bolyai University of Cluj-Napoca, Romania. (Adapted after [20].)



(a)



(b)

**Fig. 9.17:** First-principle model validation for a pilot-scale continuous distillation column of ethanol-water mixtures. (a) Experimental and simulated dynamic data; (b) Steady-state experimental and simulated temperature profiles of distillate and feed streams.

The first part of objective function (9.55) minimizes the calculated and measured temperature ( $T_{i,j,sim}$  and  $T_{i,j,exp}$ ) differences in each tray ( $P$ ) and each measured moment of time ( $N$ ). The second part is aimed to fit the time evolution of calculated and measured distillate concentrations ( $x_{i,sim}$  and  $x_{i,exp}$ ).  $w$  is a weighting factor. The bottom concentration was measured and used to verify mass balance (9.8). Figure 9.17(a) presents the correlation between simulated and experimental dynamic data and demonstrates their fairly good superposition. Figure 9.17(b) depicts steady-state measured and simulated temperature profiles of distillate and feed streams [23]. Again, data fit well.

## 9.6 Control issues of batch distillation columns

Discontinuous distillation columns are rarely used at industrial scale. However, fine chemical and pharmaceutical industry often operates batch-wise, thus batch distillation may still occasionally be in use. Next to the production scale limitations, batch distillation is un-favored by operational reasons as well. Among these are time consuming auxiliary operations such as loading the mixture, un-loading the bottom, reaching boiling temperatures, cleaning operations, etc. Therefore, this topic is only briefly addressed to within this chapter.

If the column is operated at constant distillate flow rate, its concentration varies with the bottom concentration, but the separation factor is constant in time. Thus, global mass and component balances are written as follows:

$$M_B = M_{B,0} + tD \quad (9.57)$$

$$M_B x_B = M_{B0} x_{B0} + D \int x_D dt \quad (9.58)$$

The bottom concentration  $x_B$  and the concentration  $x_D$  of collected distillate are hence:

$$x_B = \frac{M_{B0} x_{B0} - D \int x_D dt}{M_{B0} - tD} \quad (9.59)$$

$$\bar{x}_D = \frac{\int x_D dt}{t} \quad (9.60)$$

The process end-point implies the setpoint  $\bar{x}_D$  to be reached ( $\bar{x}_D = x_D$ ). When operating at high  $D$  and low  $R/D$  ratio, the distillation stops early. In contrast, when  $R/D$  is high and  $D$  is low the separation is good, the product quality is high but the energy demand

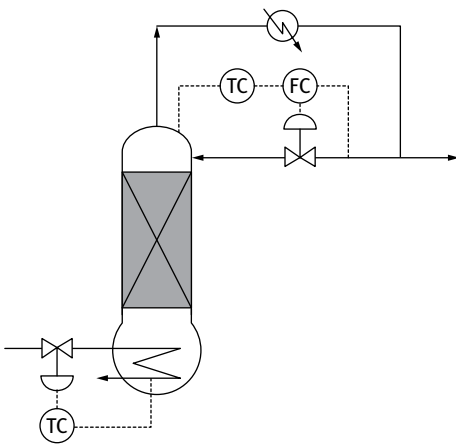


Fig. 9.18: Control scheme of a batch distillation column.



is also high. Therefore, for a better control, the value of  $D$  should be high at the beginning of the process, then gradually decline to compensate the decreasing bottom concentration  $x_B$ . Figure 9.18 presents the control scheme of a batch distillation column.

Classic operation modes of batch columns are either at constant reflux flow rate  $R$  or at constant distillate concentration  $x_D$ , respectively. When  $R$  is constant, the value of  $x_D$  will decrease in time, thus employed  $R$  has to be chosen carefully to ensure the prescribed end-point concentration. Constant  $x_D$  operation requires time-variable  $R$ ; at the beginning while  $x_B$  is relatively high,  $R$  can be low. Afterward, it must be gradually increased while  $x_B$  drops.

The optimal control strategy of batch distillation columns has been a debated topic during the last decades [24, 25]. Even so, variation of  $R$  between the two limiting cases (constant reflux or concentration) [26, 27] is the most encountered control solution. Reflux optimization may be carried out by dynamic programming. An example is presented and discussed in detail in chapter 4.2 – see also explanations for Fig. 4.5.

Model predictive control is often used to simultaneously take into account disturbances, control objectives and process constraints. However, an earlier study [28] has proved that for more complicated separations, a significant time improvement can be achieved with two stage distillation.

## References

- [1] Agachi, Ş, *Automatizarea proceselor chimice*, Casa Cărţii de Ştiinţă, Cluj-Napoca, 1994.
- [2] Luyben, W.L., *Practical Distillation Control*, Springer Science & Business Media, Berlin/Heidelberg/Dordrecht/New York City, 2012.
- [3] Shinskey, F.G., *Distillation Control: For Productivity and Energy Conservation*. McGraw-Hill, New York, 1984.
- [4] Porru, M., Baratti, R., Alvarez, J., *Energy saving through control in an industrial multicomponent distillation column*, IFAC-PapersOnLine, 48(8), 1138–1143, 2015.
- [5] Gadalla, M., Olujić, M.Z., Sun, L., De Rijke, A., Jansens, P.J., *Pinch analysis-based approach to conceptual design of internally heat-integrated distillation columns*, Chemical Engineering Research and Design, 83, 987–993, 2005.
- [6] Linnhoff, B., Dunford, H., Smith, R., *Heat integration of distillation columns into overall processes*, Chemical Engineering Science, 38(8), 1175–1188, 1983.
- [7] Iancu, M.H., *Advanced control of the heat integrated complex plants*, PhD thesis abstract, Babeş-Bolyai University of Cluj-Napoca, Romania, 2010; doctorat.ubbcluj.ro/sustinerea\_publica/rezumat/2010/inginerie-chimica/lancu\_Mihaela\_en.pdf.
- [8] Mathisen, K.M., Skogestad, S., Wolff, E.A., *Bypass selection for control of heat exchanger networks*, Computers & Chemical Engineering, 16, S263–S272, 1992.
- [9] Glemmestad, B., Gundersen, T., *A systematic procedure for optimal operation of heat exchanger networks*, AIChE Symposium Series, 1998, 320(94).
- [10] Baldea, M., Daoutidis, P., *Dynamics and Nonlinear Control of Integrated Process Systems*, Cambridge University Press, Cambridge, 2012.
- [11] Shinskey, F.G., *Energy Conservation Through Control*, Academic Press, New York, 1978.
- [12] Douglas, J.M., Jafarey, A., Seemann, R., *Short-cut techniques for distillation column design and control. 2. Column operability and control*, Industrial & Engineering Chemistry Process Design and Development, 18(2), 203–210, 1979.

- [13] Skogestad, S., Morari, M., *Control configuration selection for distillation columns*, *AIChE Journal*, 33(10), 1620–1635, 1987.
- [14] Kister, H.Z., Mathias, P.M., Steinmeyer, D.E., Penney, W.R., Crocker., B.B., Fair, J.R., *Chapter 14. Equipment for distillation, gas absorption, phase dispersion, and phase separation*, in Green, W.G., Perry, R.H. (Eds.), *Perry's Chemical Engineering Handbook, 8<sup>th</sup> Edition*, McGraw-Hill, New York, 2007.
- [15] Wittgens, B., Skogestad, S., *Evaluation of dynamic models of distillation columns with emphasis on the initial response*, *Modeling, Identification and Control*, 21(2), 83–103, 2000.
- [16] Shulman, H.L., Ullrich, C.F., Wells, N., Proulx, Z., *Performance of packed columns. III. Holdup for aqueous and nonaqueous systems*, *AIChE Journal*, 1(2), 259–264, 1955.
- [17] Gokhale, V., Hurowitz, S., Riggs, J.B., *A Comparison of Advanced Distillation Control Techniques for a Propylene/Propane Splitter*, *Chemical Engineering Research and Design*, 34(12), 4413–4419, 1995.
- [18] Kiss, A.A., *Advanced Distillation Technologies: Design, Control and Applications*, John Wiley & Sons, Chichester, 2013.
- [19] Rewagad, R.R., Kiss, A.A., *Dynamic optimization of a dividing-wall column using model predictive control*, *Chemical Engineering Science*, 68(1), 132–142, 2012.
- [20] Assandri, A.D., de Prada, C., Rueda, A., Martínez, J.L., *Nonlinear parametric predictive temperature control of a distillation column*, *Control Engineering Practice*, 21(12), 1795–1806, 2013.
- [21] Agachi, P.Ş., Cristea, V.M., *Basic Process Engineering Control*, De Gruyter, Berlin/Boston, 2014.
- [22] Agachi, P.Ş., Cristea, V.M., Nagy, Z.K., Imre-Lucaci, A., *Model Based Control: Case Studies in Process Engineering*, John Wiley & Sons, Weinheim, 2007.
- [23] Szavuly, M., Szilágyi, B., Toos, A., Imre-Lucaci, A., *Experimental validation of a distillation process mathematical model*, *CAPE Forum 2012*, p. 58, 2012.
- [24] Macchietto, S., Mujtaba, I.M., *Design of operation policies for batch distillation*, in *Batch Processing Systems Engineering SE – 9*, vol. 143, Reklaitis, G, Sunol, A., Rippin, D.T., Hortaçsu, Ö., (Eds), Springer Verlag, Berlin/Heidelberg, p. 174–215, 1996.
- [25] Zavala, J.C., Coronado, C., *Optimal control problem in batch distillation using thermodynamic efficiency*, *Industrial & Engineering Chemistry Research*, 47(8), 2788–2793, 2008.
- [26] Logsdon, J.S., Biegler, L.T., *Accurate determination of optimal reflux policies for the maximum distillate problem in batch distillation*, *Industrial & Engineering Chemistry Research*, 32(4), 692–700, 1993.
- [27] Diwekar, U.M., Malik, R.K., Madhavan, K.O., *Optimal reflux rate policy determination for multicomponent batch distillation columns*, *Computers & Chemical Engineering*, 11(6), 629–637, 1987.
- [28] Christensen, F.M., Jørgensen, S.B., *Optimal control of binary batch distillation with recycled waste cut*, *Chemical Engineering Journal*, 34(2), 57–64, 1987.

## 10 Control of absorption processes

*Absorption* is a separation process of gas mixtures based on selective dissolution of one or more gas components in a liquid solvent. The inverse process, the stripping of absorbed gas from the solvent, is called *desorption*. It is usually carried out in *absorbers*, differently structured chemical equipment designed to carry out the gas-liquid mass transfer. However, the most frequent type encountered in practice is the *absorption column* (either packed or with trays – see Fig. 10.1).

The polluted gas stream is fed at the column bottom with high concentration (mole fraction  $y_F$ ), and after dissolving the soluble component, the purified gas stream is evacuated at the top (with a mole fraction of  $y_V$ ). The liquid solvent enters at the top with low concentration of dissolvable gas (mole fraction  $x_R$ ) and flows in counter-current with the gas. While being in contact with it, the soluble species will dissolve; thus, the evacuation is more concentrated (mole fraction  $x_B$ ). The corresponding stream flow rates, expressed in mol/s, are liquid inlet  $R$  and outlet  $B$ , gas inlet  $F$  and outlet  $V$ , respectively.

The mass balance (mol/s) equations for an absorber under steady state are

$$F + R = V + B \quad (10.1)$$

and

$$Fy_F + Rx_R = Vy_V + Bx_B \quad (10.2)$$

Expressing  $B$  from eq. (10.1) and substituting it to eq. (10.2), the following results:

$$Fy_F + Rx_R = Vy_V + (F + R - V)x_B \quad (10.3)$$

The nondissolving gases leave the column within flux  $V$ :

$$F(1 - y_F) = V(1 - y_V) \quad (10.4)$$

Expressing  $V$  from eq. (10.4), substituting it to eq. (10.3), and after rearrangement, the following is obtained:

$$y_V = \frac{y_F(1 - x_B) + \frac{R}{F}(x_R - x_B)}{(1 - x_B) + \frac{R}{F}(x_R - x_B)} \quad (10.5)$$

Equation (10.5) reveals a connection between outlet gas stream purity and operational parameters such as gas and liquid flow rate as well as inlet concentrations.

Meanwhile, for dilute systems, the Kremser-Brown-Sounders [1] equation can be used to calculate the outlet gas stream concentration:

$$\frac{y_{F,i} - Y_{V,i}}{y_{F,i} - K_i X_{R,i}} = \frac{1 - A_i^{-n}}{1 - A_i^{-(n+1)}} \quad (10.6)$$

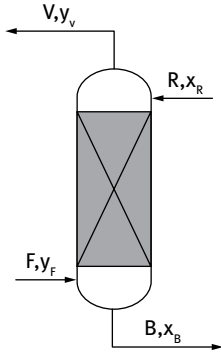


Fig. 10.1: Scheme of a continuous absorption column.

where

- $i$  denotes a certain component/chemical species;
- $Y_{V,i} = y_{V,i}V/F$  is the corrected gas phase mole fraction (dimensionless since mol/mol); similarly, the liquid is characterized by  $X_{B,i}$ ;
- $P_{i,0}$  is the vapor pressure of pure component at working temperature (atm);
- $K_i = P_{i,0}/P$  the pressure fraction of component “ $i$ ” (dimensionless);
- $A_i = R/(K_iV)$  is the absorption factor (dimensionless);
- $n$  stands for the theoretical number of trays.

The  $V - F$  difference is a result of gas absorption. Hence, if more than one species is absorbed, mass balance (10.1) becomes eq. (10.7). A similar relationship stands for the liquid stream.

$$V - F = V \sum_i (y_{F,i} - Y_{V,i}) \tag{10.7}$$

The  $i$  component mass balance (within the numerous individual mass balances) can be written instead of (10.2) as:

$$V(y_{F,i} - Y_{V,i}) = R(X_{B,i} - x_{R,i}) \tag{10.8}$$

After processing eqs. (10.7)–(10.8), absorber outlet stream composition characteristics in the specified species  $i$  are obtained – see eq. (10.9). These values are helpful in absorbance yield and performance assessment. They also represent control variables.

$$y_{V,i} = \frac{Y_{V,i}}{1 - \sum_i (y_{F,i} - Y_{V,i})} \tag{10.9}$$

$$x_{B,i} = \frac{X_{B,i}}{1 - \sum_i (X_{B,i} - x_{R,i})}$$

Since the gas is not *pure*, when assuming dilute systems, i.e. low liquid-phase concentrations, the *liquid equilibrium* concentrations (mole fractions)  $x_B^*$  and  $x_R^*$  can be expressed by Henry's law [2, 3].  $T_F$  and  $T_V$  stand for the stream temperatures of  $F$  and  $V$ , respectively.

$$x_B^* = \frac{P_{F,i}}{H(T_F)} \quad (10.10)$$

$$x_R^* = \frac{P_{V,i}}{H(T_V)}$$

The Henry's constant temperature dependence is described by the Van't Hoff relationship [2] as,

$$\frac{d \ln H}{dT} = \frac{E_{abs}}{R_{universal} T^2}, \quad (10.11)$$

where  $E_{abs}$  stands for the activation energy of absorption and  $R_{universal}$  is the universal gas constant.

Relationships (10.10) assume that at the top (contact point of  $V$  and  $R$  streams) and bottom (contact point of  $F$  and  $B$  streams), the phases are in equilibrium. Despite the fact that thermodynamic equilibrium rarely occurs in practice, this assumption is widely applied in mathematical description of absorbers.

The information presented above demonstrates that the outlet gas stream concentration (value of  $y_V$ ) in need to be controlled depends on inlet  $x_R$  and  $y_F$ , inlet flow-rates  $F$  and  $R$ , temperature, and overall pressure.

### **i** Example 10.1

An absorption column with four theoretical units works at 5-bar pressure and 38°C temperature. The feed and reflux stream compositions as well as  $K_i$  values (see eq. (10.6)) are given in Tab. 10.1. Let us calculate the variation in composition of outlet streams if the pressure increases with 5% for  $R/V = 0.5$ .

**Tab. 10.1:** Feed and reflux stream compositions of an absorption column as well as the  $K_i$  values.

	$y_F$	$x_R$	$K_i$	
			P = 5 bar	P = 5.25 bar
Propane (1)	0.070	0.0002	2.15	2.13
Isobutane (2)	0.900	0.0004	1.00	0.95
<i>n</i> -Butane (3)	0.030	0.0005	0.74	0.70

The right hand side of eq. (10.6) can be computed for all compounds from the above listed data:

$$A_1 = \frac{0.5}{2.25} = 0.222 \quad A_2 = \frac{0.5}{1} = 0.5 \quad A_3 = \frac{0.5}{0.74} = 0.675$$

Further:

$$\frac{1 - A_1^4}{1 - A_1^5} = 0.221 \quad \frac{1 - A_2^4}{1 - A_2^5} = 0.483 \quad \frac{1 - A_3^4}{1 - A_3^5} = 0.622$$

The corrected gas composition is according to eq. (10.6):

$$Y_{V,1} = 0.07 - 0.221 \times (0.07 - 0.221 \times 0.0002) = 0.054$$

$$Y_{V,2} = 0.9 - 0.483 \times (0.9 - 0.483 \times 0.0004) = 0.465$$

$$Y_{V,3} = 0.03 - 0.622 \times (0.03 - 0.622 \times 0.0005) = 0.011$$

Further, the sum on the right hand side of eq. (10.7) becomes:

$$\sum_i (y_{F,i} - Y_{V,i}) = (0.07 - 0.054) + (0.9 - 0.465) + (0.03 - 0.011) = 0.47$$

This result coupled with the use of eq. (10.9) easily yields the outlet gas stream composition:

$$y_{V,1} = \frac{0.054}{1 - 0.47} = 0.102 \quad y_{V,2} = \frac{0.465}{1 - 0.47} = 0.877 \quad y_{V,3} = \frac{0.011}{1 - 0.47} = 0.020$$

A similar algorithm leads to the liquid phase composition. From eq. (10.8), the corrected liquid values can be calculated as follows:

$$X_{B,1} = \frac{1}{0.5} (0.07 - 0.054) + 0.0002 = 0.0322$$

$$X_{B,2} = \frac{1}{0.5} (0.9 - 0.465) + 0.0004 = 0.8704$$

$$X_{B,3} = \frac{1}{0.5} (0.03 - 0.011) + 0.0005 = 0.0385$$

Further, the sum appearing in eq. (10.9) is

$$\sum_i (X_{B,i} - x_{R,i}) = (0.0322 - 0.0002) + (0.8704 - 0.0004) + (0.0385 - 0.0005) = 0.94.$$

Outlet liquid mole fractions are computed using eq. (10.9):

$$x_{B,1} = \frac{0.0322}{1 + 0.94} = 0.016 \quad x_{B,2} = \frac{0.8704}{1 + 0.94} = 0.4486 \quad x_{B,3} = \frac{0.0385}{1 + 0.94} = 0.019$$

The pressure variation involves the modification of  $K_i$  values (see Tab. 10.1). After similar calculations, the modified mole fraction values are

$$\begin{aligned} y_{V,1} &= 0.102 & y_{V,2} &= 0.875 & y_{V,3} &= 0.019 \\ x_{B,1} &= 0.0172 & x_{B,2} &= 0.459 & x_{B,3} &= 0.02 \end{aligned}$$

**Tab. 10.2:** Calculated outlet mole fraction values in liquid and gas streams at 5 and 5.25 bar and the net percentage change after the pressure variation.

Parameter	Composition					
	$y_{V,1}$	$y_{V,2}$	$y_{V,3}$	$x_{B,1}$	$x_{B,2}$	$x_{B,3}$
P = 5 bar	0.102	0.877	0.020	0.0160	0.4486	0.019
P = 5.25 bar	0.102	0.875	0.019	0.0172	0.4590	0.020
Net change (%)	0	-0.228	-5.000	-7.5000	+2.3180	+5.000

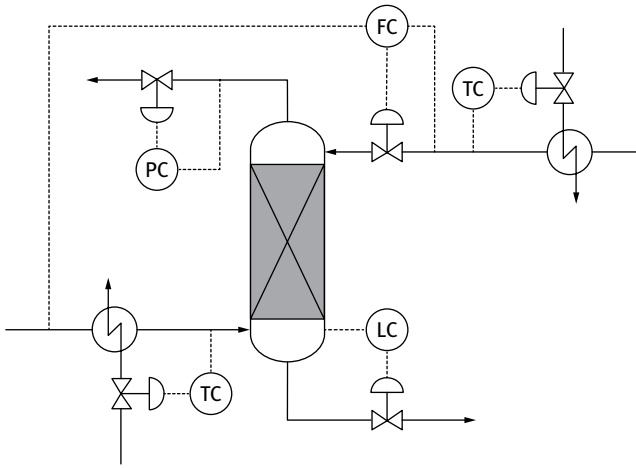
Table 10.1 summarizes calculated values in Example 10.1. The results demonstrate that increased pressure shifts the gas-liquid equilibrium of species  $i$  toward the liquid. Thus, gas solubility increases with pressure [2].

The industrial practice of absorber operation proved that the most significant part (around 80%) of total mass transfer occurs in the first and last mass transfer unit (tray). Meanwhile, the temperature profile in the column can vary due to dissolution heat release and heat transfer between the different temperature gas and liquid inlet streams. This causes a temperature gradient along the column that affects the gas-liquid mass transfer rate through both Henry's law and  $K_i$  factor values.

When inlet stream temperatures and compositions are constant, the  $R/F$  ratio controls the composition of both currents. If the feed temperature increases, solubility will decrease. Hence, the value of  $R/F$  has to increase to compensate for this negative effect. Similarly, if the amount of soluble component increases in the feed, the values of  $R/F$  should also increase (more mass has to be dissolved). These observations are quantitatively described by equations (10.6) and (10.9).

Consequently, the control structure of an absorption column (see Fig. 10.2) must contain the following elements:

- $R/F$  ratio control;
- temperature control of both inlet streams;
- level control at the bottom (to close the mass balance);
- pressure control at the top.



**Fig. 10.2:** General control configuration of a continuous absorption column. The flow rate controller FC is on ratio.

However, more advanced control strategies can be employed. Feedforward or cascade controllers are applied [4], for instance, to attenuate the effects of disturbances in temperature or composition of feed streams.

A huge variety of control configurations can be imagined. The choice of a reasonable controller configuration based purely on theoretical considerations, requires sensitivity, and parametric analysis [5]. Such model-based analysis reveals steady state and dynamic behavior patterns.

In several cases, a chemical reaction accompanies the physical dissolution (absorption). This happens during the  $\text{SO}_2$  absorption in water:  $\text{SO}_2$  is dissolved physically then enters a chemical reaction with water.



In such situations, the *chemical equilibrium* is described by *equilibrium constant*  $K_{eq}$ . For diluted solutions, the water concentration is virtually unchanged and the equilibrium constant becomes

$$K_{eq} = \frac{[\text{H}^+][\text{HSO}_3^-]}{[\text{SO}_2]} \quad (10.13)$$

$K_{eq}$  depends on temperature and pressure; these functions are described by thermodynamic and kinetic equations [2]. Equation (10.13) can be used to calculate the amount of  $\text{SO}_2$  consumed by chemical reaction.

The reaction leads to an apparent increase in  $\text{SO}_2$  dissolution; the total quantity of  $\text{SO}_2$  transferred from the gas into liquid will increase as compared to the thermodynamically expected value (Henry's law).



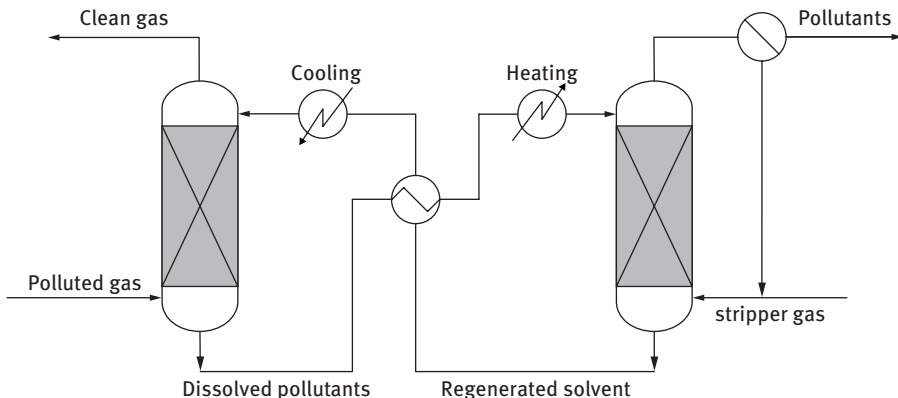


Fig. 10.3: Absorption-desorption system with heat integration.

The dissolved mass (after absorption) has to be stripped from the liquid. This occurs in a desorption column. In contrast with the absorbers, the desorbers operate under higher temperature and lower pressure (conditions that favor stripping).

Absorption and desorption columns are generally operated in couples. Such a system is presented in Fig. 10.3 and has multiple advantages. Among these, the following aspect can be mentioned:

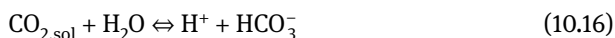
- The liquid phase leaving the desorber is recycled to the absorber. This reduces necessary fresh solvent amounts. Some specific absorption processes that use expensive and/or harmful solvents are deliberately conducted under such strategy.
- The simultaneous operation opens new doors for thermic integration; the energy demand, as compared to the isolated operation, is lower.

During the last decade, CO<sub>2</sub> absorption from post-combustion gases was granted special attention. The usual absorber works under pressure and employs an organic liquid (often aqueous mono-ethanol amine solution [6]) for capturing CO<sub>2</sub>. This strategy is applied mainly to reduce emissions of fossil power plants and to produce CO<sub>2</sub> in desired purity as valuable by-product. The process is based on following elementary physical and chemical steps:

1. Physical dissolution of CO<sub>2</sub>:



2. Chemical reaction between solute CO<sub>2</sub> and mono-ethanol amine and/or water, respectively:



After desorption, CO<sub>2</sub> is compressed and transported to the user. The absorption/desorption/compression system is thermally integrated with the power plant. This yields significant overall performance improvement compared to the isolated thermal operation of the absorption/desorption system. However, the more advanced heat integration reduces the process flexibility and therefore controllability. Meanwhile, model predictive controllers (see chapter 2) could improve control performance of such rigid systems [7].

## References

- [1] Agachi, Ș., *Automatizarea Proceselor Chimice*, Casa Cărții de Știință, Cluj-Napoca, 1994.
- [2] Tosun, I., *The Thermodynamics of Phase and Reaction Equilibria*, Elsevier, Amsterdam, 2013.
- [3] Kister, H.Z., Mathias, P.M., Steinmeyer, D.E., Penney, W.R., Crocker, B.B., Fair, J.R., *Chapter 14. Equipment for distillation, gas absorption, phase dispersion, and phase separation in* Green, W.G., Perry, R.H. (Eds.), *Perry's Chemical Engineering Handbook, Eighth Edition*, McGraw-Hill, New York, 2007.
- [4] Govindarajan, A., Jayaraman, S.K., Sethuraman, V., Raul, P.R., Rhinehart, R.R., *Cascaded process model based control: packed absorption column application*, *ISA Transactions*, 53, 391–401, 2014.
- [5] Ungureanu, S., *Stabilitatea sistemelor dinamice*, Editura Tehnică, București, 1988.
- [6] Pădurean, A., Cormoș, C.C., Cormoș, A.M., Agachi, P.Ș., *Multicriterial analysis of post-combustion carbon dioxide capture using alkanolamines*, *International Journal of Greenhouse Gas Control*, 5, 676–685, 2011.
- [7] Hossein Sahraei, M., Ricardez-Sandoval, L.A., *Controllability and optimal scheduling of a CO<sub>2</sub> capture plant using model predictive control*, *International Journal of Greenhouse Gas Control*, 30, 58–71, 2014.

## 11 Control of extraction processes

*Extraction* is a complete or partial separation process of a chemical called *solute*, from a liquid or solid mixture called *matrix*, by means of another species called *solvent*. It is useful when liquid individual components cannot be separated by means of distillation because of their close boiling points, low volatility, thermal instability, or formation of an azeotrope. In case of solid mixtures, it is a handy tool for the recovery of useful components.

If the feed matrix **F** contains the solute species **A** to be separated by employing solvent **S**, then the following requirements have to be fulfilled for the extraction to be possible [1, 2]:

- Liquid solvent **S** is immiscible (or partially miscible) with **F** if this is a liquid, or **S** does not dissolve **F** in case the latter is a solid;
- Solute **A** dissolves in solvent **S**;
- At equilibrium, the concentration of solute **A** in solvent **S** should be (considerable) higher than in initial matrix **F**.

Any extraction demands the following sequence of steps:

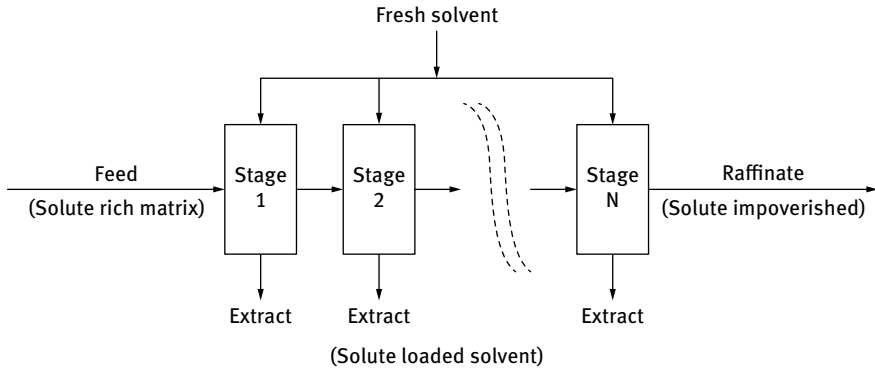
- Contact between matrix and solvent;
- Transfer of solute from the matrix into the solvent until (ideally) the equilibrium concentrations are reached due to partition of solute between matrix and solvent;
- Separation of resulting extract from the raffinate or residue, respectively.

The raffinate is further processed to recover both solute (when desired) and solvent. The transfer of solute from the loaded solvent into another phase is called *stripping* or *back-extraction* and serves the purpose of obtaining it. Meanwhile, recovery of the solvent is compulsory because of economic as well as environmental reasons.

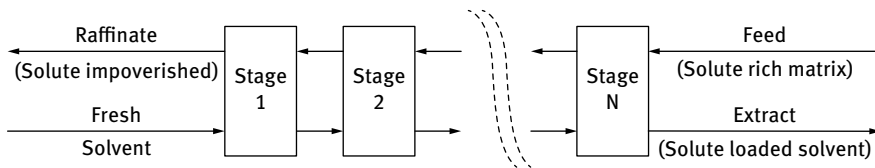
If both the matrix and the solvent are liquids, the procedure is called *liquid-liquid (L-L) extraction*, *solvent extraction* or *refining*. The solute loaded solvent is called *extract* and the impoverished matrix *raffinate*, respectively. If the matrix is a solid, the technique is called *solid-liquid (S-L) extraction* or *percolation* [1, 2]. The resulting liquid and solid phases are called *extract* and *residue*, respectively. In the case of L-L extraction, the raffinate may not contain any solvent, whereas for S-L extraction, the solid residue will always be wetted by it. The extract may also require post-extraction treatment because of solid residue particles.

Liquid-liquid extraction is suitable for continuous operation [1, 2] even at throughputs of 100,000 m<sup>3</sup>/h [3] and is therefore applied at large scale in the oil industry [3, 4]. Meanwhile, solid-liquid extraction is applied mainly batchwise and at much smaller scale in the food and medical industry. Therefore, this chapter focuses mainly on L-L extraction.

The most widely employed techniques involve either *crosscurrent* (see Fig. 11.1(a)) or *countercurrent* (see Fig. 11.1(b)) L-L extraction. Both involve a cascade of *stages* or *extractors*. Within the first option, fresh solvent is added to the raffinate of each stage; the extract is collected from each stage. The second option implies the solvent and the raffinate to flow in countercurrent through the entire sequence of stages [1].



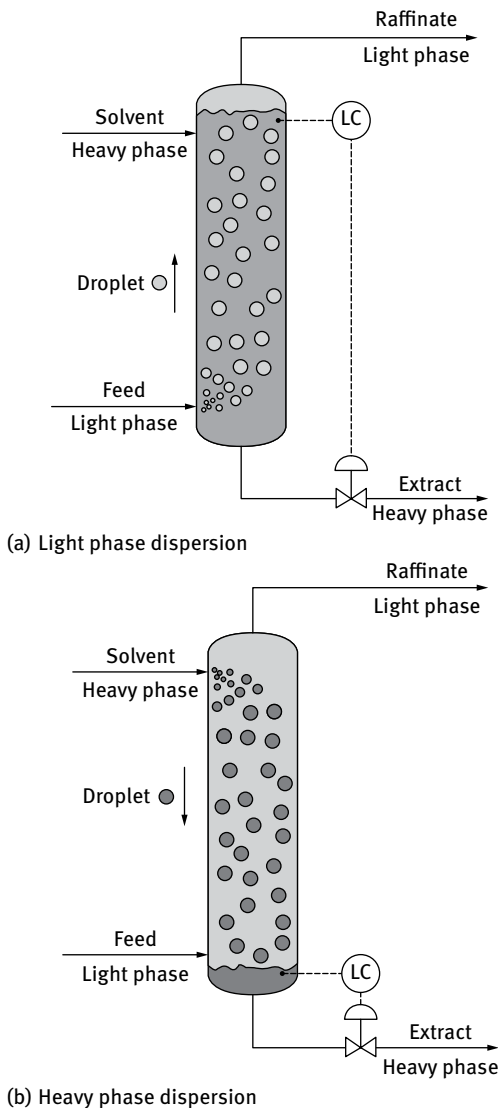
(a) Crosscurrent L-L extraction



(b) Countercurrent L-L extraction

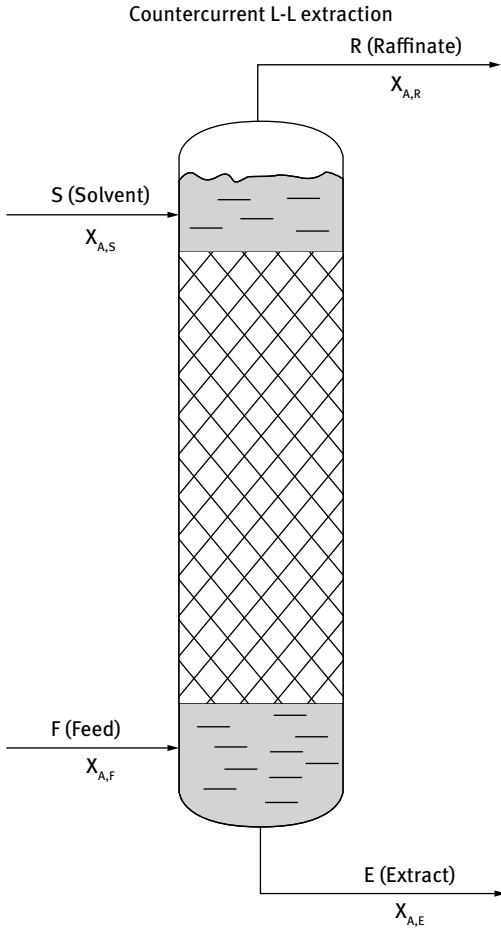
**Fig. 11.1:** (a) crosscurrent and (b) countercurrent liquid-liquid extraction. (Adapted after [1])

Mass transfer between the feed (or raffinate for a multistage extraction) and the solvent (or extract for the case of multistep countercurrent extraction) includes droplet formation/dispersing, transport, and settling of the phases. The feed can be dispersed into the solvent (Fig. 11.2(a)) or the solvent dispersed in it (Fig. 11.2(b)) [3]. Figure 11.2 reveals these differences and depicts the feed matrix as a light-phase that accumulates at the top of the heavy one. Dispersion, and settling, of phases is based on the difference between the densities of feed and solvent. Density difference is also responsible for the spontaneous countercurrent flows. Thus, light-phase droplets will move upward and heavy ones in the opposite direction along the height of an *extraction tower*. The types of dispersion remain as depicted in Fig. 11.2 for the case of a heavy-phase feed entered at the top of the tower. A widely employed industrial scale L-L extractor is the agitated extraction column. It requires, as a limiting condition, at least a  $0.05\text{-kg/m}^3$  density difference between the feed and the solvent [3, 4]. Figure 11.2 also illustrates the differences imposed for the heavy liquid (extract of interest) level control (see also explanations for Fig. 11.4).



**Fig. 11.2:** Liquid level control for extraction towers with (a) light phase and (b) heavy phase dispersion. Light gray stands for the light/low density and dark gray stands for the heavy/high density liquid. (Adapted from [3])

Figure 11.3 presents the scheme of mass balance for a countercurrent L-L packed extraction tower, with light-phase dispersion and in continuous steady-state operation mode. It describes a *binary mixture* that shares a *single solute*. The species of interest is **A**, to be extracted from feed matrix **F** into solvent **S**. The flow rates ( $\text{m}^3/\text{h}$  or  $\text{kg}/\text{h}$ ) are symbolized with  $F$ ,  $S$ ,  $E$ , and  $R$  for the feed, solvent, extract, and raffinate respectively. The dimensionless mass or molar fraction of solute **A** in each of these streams is symbolized with  $x_{AF}$ ,  $x_{AS}$ ,  $x_{AE}$ , and  $x_{AR}$ , respectively. Ideally,  $x_{AR} = 0$  (100% extraction efficiency) and  $x_{AS} = 0$  (pure solvent, 100% stripping efficiency).



**Fig. 11.3:** Mass-balance of a countercurrent L-L packed extraction tower, with light phase dispersion and in continuous steady-state operation mode.

The total mass balance and those of solute **A** and solvent **S** are expressed by eqs. (11.1) and (11.2)–(11.3), respectively.

$$F + S = E + R \quad (11.1)$$

$$FX_{AF} + Sx_{AS} = Ex_{AE} + Rx_{AR} \quad (11.2)$$

$$S(1 - x_{AS}) = E(1 - x_{AE}) \quad (11.3)$$

Operational objectives of continuous L-L extraction control include, among others, extract quality specifications. This translates into maintenance of a desired  $x_{AE}$  value for variables  $F$ ,  $x_{AF}$ , and  $x_{AS}$ . The mass balances above leads to (11.4), and further to the expression (11.5) of  $x_{AE}$ , the mass/molar fraction of solute **A** in extract stream **E**.

$$x_{AE}[S(1 - x_{AR}) + F(x_{AF} - x_{AR})] = Sx_{AS}(1 - x_{AR}) + F(x_{AF} - x_{AR}) \quad (11.4)$$

$$x_{AE} = \frac{\frac{S}{F}x_{AS}(1-x_{AR}) + (x_{AF} - x_{AR})}{\frac{S}{F}(1-x_{AR}) + (x_{AF} - x_{AR})} \quad (11.5)$$

The flow rate ratio  $S/F$  affects  $x_{AE}$ ; hence, input stream rates have to be put on ratio control (see Fig. 11.4). Thus,  $S$  will automatically adjust to disturbances of  $F$ . Flow rate control also ensures productivity specifications, such as desired hourly throughputs. In addition, it avoids drying out of the extractor or its flooding by either the light or heavy phase. Liquid level control (see Figs. 11.2 and 11.4) helps to ensure the latter operational constraint but also provides constant (steady-state) volumes of light and heavy phases, respectively. These contribute to the desired  $x_{AE}$  and  $x_{AR}$  values.

The driving force of mass transfer in L-L extraction is the solute concentration gradient between the steady-state bulk phase ( $x_{AE}$  and  $x_{AR}$ ) and the L-L interface ( $x_{AE}^i$  and  $x_{AR}^i$ ). The solute mass stream  $N$  (kg/h) between the two liquid phases is expressed by relationship (11.6), where  $k_E$  and  $k_R$  stand for global mass transfer coefficients (kg/h) in the extract and solvent, respectively. They depend on the nature and composition of streams E and R as well as temperature. The end-point for batch extraction implies  $N = 0$ ; in other words,  $x_{AE} = x_{AE}^i$  and  $x_{AR} = x_{AR}^i$  (bulk solute concentrations in E and R correspond to the equilibrium values):

$$N = k_E(x_{AE}^i - x_{AE}) = k_R(x_{AR} - x_{AR}^i) \quad (11.6)$$

Separation by L-L extraction (that is, values of  $x_{AE}^i$  and  $x_{AR}^i$ ) is governed by the thermodynamic partition *equilibrium* of the solute between the two liquid phases in contact [1–2, 5–6]. It is usually expressed by the Nernst distribution law below – see (11.7), where  $K_A$  is the dimensionless equilibrium constant called *partition coefficient*.

$$K_A = \frac{x_{AE}^i}{x_{AR}^i} \quad (11.7)$$

Equation (11.7) describes a linear correlation between the solute contents in the resulting extract and raffinate streams, respectively. It is valid only for diluted solutions. Defined as such, good performance extraction implies  $K_A > 1$ , since  $x_{AE} > x_{AR}$  ought to be true. It is determined experimentally for each solute and feed/solvent pair, at various temperatures and vapor pressures. Therefore, partition relationships are also important in the choice of correct  $S/F$  ratios [1, 2].

The temperature dependence of  $K_A$  can be expressed by the Van't Hoff equation [5, 6] (11.8), where  $\Delta H$  is the equilibrium phase transfer enthalpy of the solute (J/mol),  $T^\circ$  is the temperature (kelvin), and  $R$  is the universal gas constant.

$$\frac{d \ln(K_A)}{dT^\circ} = \frac{\Delta H}{R(T^\circ)^2} \quad \text{or} \quad \ln(K_A) = -\frac{\Delta H}{RT^\circ} \quad (11.8)$$

Depending on  $\Delta H$  value,  $K_A$  is more or less strongly affected by temperature (see Tab. 11.1). Therefore, the necessity of temperature control has to be checked before investing in expensive equipment (see Examples 11.1 and 11.2), whereas feed/solvent flow rate ratio and heavy-phase level control are always recommended (see Fig. 11.4).

Figure 11.4 presents a complete control system for a countercurrent L-L packed extraction tower operated continuously under steady-state conditions [7, 8]. The loops control the S/F flow rate ratio, temperatures of both F and S streams, and extract volume. Only the latter is based on a liquid level reading; the others measure directly the controlled variable. All manipulated variables are flow rates. Input streams benefit of temperature adjustment before entering the extractor. Improved performance control can be achieved using various model-based techniques [9–12].

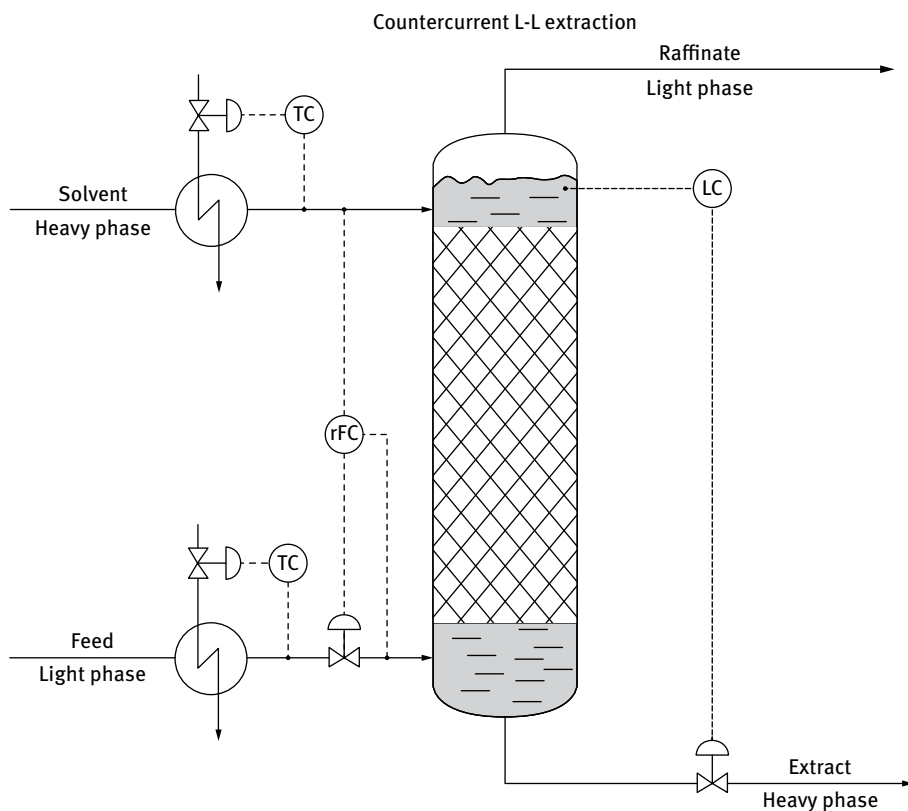


Fig. 11.4: Control scheme for the extraction tower in Fig. 11.3. (Adapted after [7])

Figures 11.3 and 11.4 as well as relationships (11.1)–(11.5) refer to binary mixtures of immiscible or partially/totally miscible liquids that share a single solute. Sometimes, such a mixture can share *two or more solutes*. If the feed matrix contains, for example,



species **A** and **B**, both soluble and extractable by the solvent, another indicator of process performance is also used in practice. It is called *selectivity*  $\beta$  or *relative separation* and refers to the capacity of the solvent to extract preferentially a certain solute from the feed [2, 5].  $\beta$  is defined as the ratio between the partition coefficients  $K_A$  and  $K_B$  of **A** and **B**, respectively – see eq. (11.7); thus, it is dimensionless.

$$\beta = \frac{K_A}{K_B} \quad (11.9)$$

If  $\beta > 1$ , the solvent has a higher affinity toward **A**, whereas if  $\beta < 1$ , it prefers **B**. If  $\beta = 1$ , the employed solvent is nonselective, and separation of **A** or **B** from matrix **F** cannot be carried out by means of extraction. According to relationship (11.8), both  $K_A$  and  $K_B$  are temperature dependent; hence, adjustment of  $\beta$  to a desired value can only benefit from temperature control (see Fig. 11.4).

*Multicomponent* liquid mixtures are often encountered in practice [1]. In these cases, mass balances have to consider all components. Partition equilibrium will only seldom obey linear relationships such as that in eq. (11.7). Therefore, complex partition graphs are in use. Some examples are the Gibbs, Janecke, or tetrahedral diagrams for ternary or quaternary systems. These are drawn at constant temperature and vapor pressure [1, 2, 5].

### **i** Example 11.1

Let us consider an extraction tower such as presented in Figs. 11.3 and 11.4. The goal is to extract at 300 K ethylamine from an aqueous solution, available at  $F = 100 \text{ m}^3/\text{h}$  flow rate, by using methylbenzene (density of  $\rho_S = 870 \text{ kg}/\text{m}^3$ ). The dispersed phase is the solvent (see Fig. 11.2(a)). The feeding streams are characterized by  $x_{AF} = 0.15$  and  $x_{AS} = 0.01$ , respectively [7]. The distribution of solute between aqueous and organic phases is described by the Nernst equation (11.7) by a  $K_A = 1.360$  at 300 K. The setpoint is at  $x_{AE} = 0.05$ .

Let us estimate the change in the necessary solvent flow rate as well as in the resulted hourly extracted ethylamine quantity if the temperature increases by 25 K.

Equation (11.5) will lead to the value of S/F ratio. By replacing in it  $x_{AR}$  according to eq. (11.7), the following will be obtained:

$$\frac{S}{F} = \frac{(x_{AF} - x_{AR})(1 - x_{AE})}{(x_{AE} - x_{AS})(1 - x_{AR})} = \frac{\left(x_{AF} - \frac{x_{AE}}{K_A}\right)(1 - x_{AE})}{(x_{AE} - x_{AS})\left(1 - \frac{x_{AE}}{K_A}\right)} \quad (11.10)$$

Hence,

$$\frac{S}{F} = \frac{\left(0.15 - \frac{0.05}{1.36}\right)(1 - 0.05)}{(0.05 - 0.01)\left(1 - \frac{0.05}{1.36}\right)} = 2.79$$

and  $S = 2.79 F = 2.79 \times 100 = 279 \text{ m}^3/\text{h}$  at 300 K.

The extract flow rate can be computed by means of mass balance (11.3):

$$E = \frac{S(1 - x_{AS})}{(1 - x_{AE})} = \frac{279(1 - 0.01)}{(1 - 0.05)} = 290.75 \text{ m}^3/\text{h} \quad (11.11)$$

The hourly extracted ethylamine quantity is calculated under the simplifying assumption that the solvent density is not affected significantly by low solute contents.

$$E x_{AE} \rho_S = 290.75 \text{ m}^3/\text{h} \cdot 0.05 \cdot 870 \text{ kg/m}^3 = 12647.63 \text{ kg/h} \quad (11.12)$$

Hence, hourly productivity is of  $\approx 12.65$  tons/h.

At 325 K (25 K temperature increase), the partition coefficient will be computed by using eq. (11.8). The necessary  $\Delta H$  value is obtained from  $K_A$  at 300 K.

$$\Delta H = -RT^\circ \ln(K_A) \quad (11.13)$$

Thus,

$$\Delta H = -8.314 \text{ J/mol K} \cdot 300\text{K} \cdot \ln(1.36) = -766.93 \text{ J/mol} \approx -767 \text{ J/mol}$$

Further,

$$\ln(K_A) = -\frac{-767 \text{ J/mol}}{8.314 \text{ J/mol K} \cdot 325\text{K}} = 0.284 \text{ and } K_A = e^{-0.284} = 1.328.$$

Values at 325 K lead to  $\frac{S}{F} = 2.77$ ,  $S = 277 \text{ m}^3/\text{h}$ , and  $E = 288.66 \text{ m}^3/\text{h}$ , and a productivity of  $\approx 12.56$  tons/h removed solute.

With increased temperature,  $K_A$  decreases slightly (with 2.35%) and less solute will be removed by the solvent. Accordingly, to keep the setpoint  $x_{AE}$ , the flow rate  $S$  has to be slightly adjusted (with 0.7%). The system responds with the same 0.7% less productivity. If these fairly small deviations fall within the range of admitted specifications, then temperature control is not necessary. However, if the solute is expensive or raises environmental/health concerns, the setpoint  $x_{AE}$  might be tightly controlled and temperature regulation is necessary. However, the example illustrates the importance of  $S/F$  flow rate ratio control.

### Example 11.2

Let us consider the extractor described in the previous example but for a ten-fold higher as well as lower value of phase transfer enthalpy:  $\Delta H_1 = -7670.0 \text{ J/mol}$  and  $\Delta H_2 = -76.7 \text{ J/mol}$ . The input streams are set on constant ratio. The data are  $S/F = 3$ ,  $x_{AF} = 0.15$ , and  $x_{AS} = 0.01$ , respectively. Depending on the season, the



temperature of the input streams can vary between 5°C and 60°C. Let us assess whether temperature control is a necessity with respect to setpoint  $x_{AE}$ .

Equation (11.8) is employed for the calculus of  $K_A$  at the extremities of the temperature disturbance interval. By replacing eq. (11.7) in eq. (11.5), the value of  $x_{AE}$  can be obtained for each set of data. The results are cumulated in Tab. 11.1.

**Tab. 11.1:** Effect of temperature on the solute mass fraction of the resulting extract.

Temperature		$\Delta H_1 = -7670 \text{ J/mol}$		$\Delta H_2 = -76.7 \text{ J/mol}$	
°C	Kelvin	$K_A$	$x_{AE}$	$K_A$	$x_{AE}$
5	278	27.619	0.0566	1.034	0.0453
60	333	15.965	0.0562	1.028	0.0452

Higher  $K_A$ , as defined in eq. (11.7), favors extraction of solute and increases  $x_{AE}$ . Temperature changes from 5°C to 60°C bring about less than 1% variation in  $x_{AE}$  values. Thus, the system is virtually temperature independent and its regulation is unnecessary.

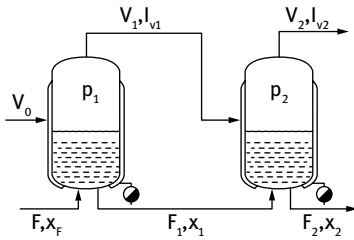
## References

- [1] Robbins, L.A., Cusak, R.W., *Section 15: Liquid-Liquid Operations and Equipment*, p. 15-1-15-47 in Perry, R.H., Green, W.G., Maloney, J.O., (Eds.), *Perry's Chemical Engineering Handbook*, Seventh Edition, McGraw-Hill, New York, 1997.
- [2] Tudose, R.Z., Ibănescu, I., Vasiliu, M., Stancu, A., Cristian, Gh., Lungu, M., *Procese, operații, utilaje în industria chimică*, Editura Didactică și Pedagogică, București, 1977.
- [3] Schultz + Partner Verharenstechnik GmbH, *Liquid-liquid extraction*, prospectus, <http://www.schulzpartner.com/en/downloads/>
- [4] Sulzer Chemtech, *Liquid-liquid extraction technology*, prospectus, <http://website-box.net/site/www.sulzer.com>
- [5] Atkins, P., de Paula, J., *Physical Chemistry*, 9<sup>th</sup> Edition, Oxford University Press, Oxford, 2010.
- [6] Motschmann, H., Hofmann, M., *Physikalische Chemie*, DeGruyter, Berlin, 2014.
- [7] Agachi, Ș., *Automatizarea proceselor chimice*, Casa Cărții de Știință, Cluj-Napoca, 1994.
- [8] Weinstein, O., Semait, R., Lewin, D.R., *Modeling, simulation and control of liquid-liquid extraction columns*, *Chemical Engineering Science*, 53, 325–329, 1998.
- [9] Mjalli, F.S., Abdel-Jabbar, N.M., Fletcher, J.P., *Modeling, simulation and control of a scheibel liquid-liquid contactor. Part 1. Dynamic analysis and system identification*, *Chemical Engineering and Processing*, 44, 543–555, 2005.
- [10] Mjalli, F.S., Abdel-Jabbar, N.M., Fletcher, J.P., *Modeling, simulation and control of scheibel a liquid-liquid contactor. Part 2. Model-based control synthesis and design*, *Chemical Engineering and Processing*, 44, 531–542, 2005.
- [11] Mjalli, F.S., *Neuronal network model-based predicted control of liquid-liquid extraction contactors*, *Chemical Engineering and Processing*, 60, 239–253, 2005.
- [12] Djurovic, J., *An inverse control of the extraction column*, *International Journal of Mathematical Models and Methods in Applied Sciences*, 5, 67–76, 2011.

## 12 Control of evaporation processes

*Evaporation* is a separation process of a solid substance dissolved in a liquid solvent. The procedure is to heat the initial solution till the solvent is evaporated to the desired degree. This process is a very old one, and it is being used to obtain salt or fresh water from sea water. Usually, it is employed to concentrate the solution where the solid product is more valuable than the solvent (see NaOH dissolved in  $H_2O$  from the electrolysis of the brine) [1].

The evaporation takes place in *evaporators* with *simple (single)* or *multiple evaporation stages (effects)* (Fig. 12.1) [2]. The evaporators usually contain a fascicule of tubes through which the solution to be concentrated is circulated through natural or forced convection. These are surrounded by a steam chest. In the multiple effect evaporator, the temperature, pressure, and heat transfer coefficients vary with each consequent effect, decreasing with each effect.



**Fig. 12.1:** Evaporator plant with two effects.  $F, F_{i(1,2)}$  are the mass flows to or from the effects;  $V_{i(0,1,2)}$  are the steam and vapor mass flows;  $p_{i(1,2)}$  are the vapor pressures inside the evaporators;  $x_{i(F,1,2)}$  are the solid mass fractions in the feed and evaporator's solution;  $l_{vi(0,1,2)}$  is the latent heat of vaporization in each evaporator, either steam or solvent (the vapor flow is not always water, the solution may contain another solvent).

To save energy, vapors that result from the evaporation of the solution in each effect are used as heating agent for the following ones.

To establish the control solution, one has to describe the process through a steady-state mathematical model that links the output variable  $x_2$ , that is the final solid product mass fraction, to the input variables, by using the following hypotheses:

- the evaporators are always at the boiling temperature of the solution by manipulation of pressures  $p_1$  and  $p_2$ ;
- the feed concentration is kept at the steady-state value of the concentration in the first effect;
- the vapors do not entrain solid particles; this is not entirely true if we are looking at the deposit of solid on the vapor outlet piping;
- the steam used for heating always condenses completely at the condensing point by transferring only the latent heat of vaporization;

- the heat losses to the environment are considered null; in reality, a good insulation has a heat loss of 5% from the heat content of the equipment;
- the solution level in the evaporators is constant.

With these simplifying assumptions, the mathematical model is formed of mass (eqs. (12.1) and (12.2)) and heat balances (eq. (12.3)) expressed for both evaporators:

$$F = V_1 + F_1 \text{ and } F_1 = V_2 + F_2 \quad (12.1)$$

$$Fx_F = F_1x_1 \text{ and } F_1x_1 = F_2x_2 \quad (12.2)$$

$$V_0l_{v0} = V_1l_{v1} \text{ and } V_1l_{v1} = V_2l_{v2} \quad (12.3)$$

These mass/heat balances are valid only in the conditions imposed by the constraints (simplifying hypotheses) above. Eqs. (12.1) can be written as such if the level is constant (hypothesis 6); eqs. (12.2) respects hypotheses 2 and 3; eqs. (12.3) respects conditions 1, 3, 4, and 5; the entire heat transferred by the heating agent ( $V_0l_{v0}$ ) is used to vaporize the solution in the first effect, and the heat transferred by the heating agent to the second effect ( $V_1l_{v1}$ ) is used to vaporize the solution in it.

By replacing in the second part of eq. (12.2) the expression of  $F_2 = F_1 - V_2$  (eq. (12.1)), one obtains:

$$x_2 = x_1 \frac{F_1}{F_2} = x_1 \frac{F_1}{F_1 - V_2}, \quad (12.4)$$

and by further replacing  $V_2 = V_1 \frac{l_{v1}}{l_{v2}}$  (eq. (12.3)), one obtains:

$$x_2 = x_1 \frac{F_1}{F_1 - V_1 \frac{l_{v1}}{l_{v2}}}, \quad (12.5)$$

Consequently, by using the same logic:

$$x_2 = x_F \frac{F}{F - V_0 l_{v0} \left( \frac{1}{l_{v1}} + \frac{1}{l_{v2}} \right)} \quad (12.6)$$

Generalizing for an  $n$ -effect evaporator leads to:

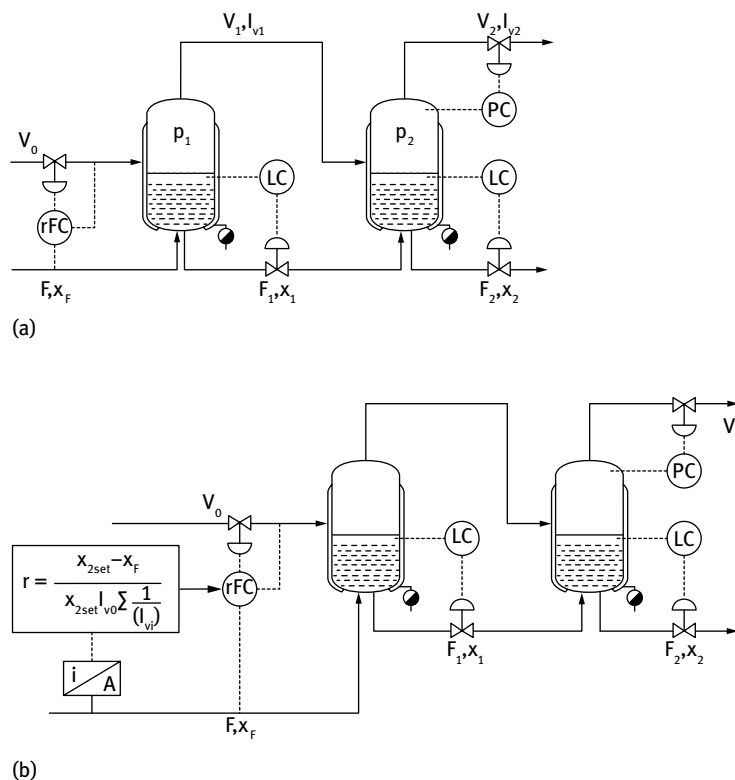
$$x_n = x_F \frac{1}{1 - \frac{V_0}{F} l_{v0} \sum_{i=1}^n \frac{1}{l_{vi}}}. \quad (12.7)$$

By analyzing the expression of the final mass fraction, one observes that it depends on  $x_F$ , which is considered to be an uncontrolled variable (disturbance), as well as on the ratio  $\frac{V_0}{F}$ . However, since the mass balance in steady state is written in the form of eq. (12.1), the level control in both evaporators has to be ensured by two level control loops.

The other issue is that of keeping the boiling temperature constant and appropriate; this is carried out by using a pressure control loop. An additional control loop for the feed temperature control, can be added to the entire automation scheme. Optionally, a ratio control can be applied to the  $\frac{V_0}{F}$  flow rate ratio (Fig. 12.2(a) and (b)) if the RGA, or any sensitivity analysis, mentions the ratio as being an important control variable, as expressed in eqs. (12.7) and (12.15). If the disturbance  $x_F$  is really important and frequent, the ratio can be calculated by an inferential scheme with:

$$\frac{V_0}{F} = \frac{x_{2\text{ set}} - x_F}{x_{2\text{ set}} l_{v0} \sum_{i=1}^n \frac{1}{l_{vi}}} \quad (12.8)$$

Relationship (12.8) is obtained from eq. (12.6), or by means of a feedforward controller. Especially the feedforward controllers are requested for small holdup evaporators that are difficult to control with independent control loops. At these evaporator plants, variations in steam and feed, as in feed concentration, cause high loss of valuable product.



**Fig. 12.2:** Possible control schemes for the evaporation process. (a) Classic evaporation control scheme. (b) Advanced control scheme.

### Sensitivity analysis relative to pressure, steam/feed flow rates, and feed concentration variations

In order to adopt a control solution for an evaporating system, a sensitivity analysis has to be done [3].

#### **i** Example 12.1

Consider a multitubular evaporator for a concentrated NaOH solution (Fig. 12.3). It has the following characteristics (technological parameters are given at nominal operation point):

Number of pipes of the boiler:  $n = 1860$

Pipelines length:  $l = 8.5$  m

Inner diameter:  $d_i = 0.033$  m

Outer diameter:  $d_o = 0.038$  m

Solution level in the evaporator:  $h = 7.1$  m

Inner tubes wall thermal conductivity:  $\lambda_w = 50$  W/m · K

Input temperature:  $T_F^\circ = 140$  °C

Input NaOH mass fraction:  $x_F = 0.1$  kg/kg

Output NaOH mass fraction:  $x_1 = 0.1568$  kg/kg

Operation pressure:  $p = 5.5$  bar

Solution feed flow rate:  $F = 37$  kg/s

Steam flow rate:  $V_0 = 10$  kg/s

Steam pressure:  $p_{st} = 12$  bar

Solution thermal conductivity:  $\lambda_s = 0.48$  W/m · K

Steam thermal conductivity:  $\lambda_{st} = 0.65$  W/m · K

Solution viscosity:  $\eta_s = 3 \cdot 10^{-3}$  Pa · s

Steam viscosity:  $\eta_{st} = 1.4 \cdot 10^{-5}$  Pa · s

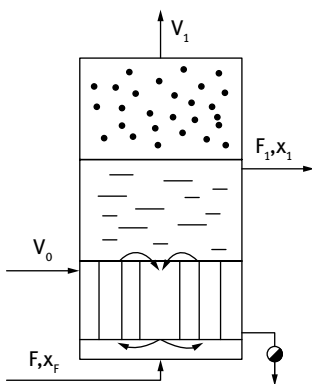


Fig. 12.3: Multitubular evaporator with single effect.

The change in output concentration with the variation of the pressure in the evaporator, with  $\pm 0.5$  bar relative to the operational pressure, has to be calculated. Same

evaluations are to be done for a steam pressure change from 12 to 8 bar and for a steam flow increase from 10 to 12 kg/s. Comments on the control system in these conditions are required.

The sodium hydroxide solution circulates inside the tubes while the steam flows outside them. The equations describing the operation of the evaporator are:

$$\begin{aligned}
 Fc_{pF,sol}T_F^\circ - F_1c_{p,sol}T_{b,sol}^\circ - V_1i_{v,sol} + K_T A_T (T_{st}^\circ - T_{b,sol}^\circ) &= 0 \\
 F - F_1 - V_1 &= 0 \\
 Fx_F - F_1x_1 &= 0 \\
 c_{p,NaOH} &= \left(\frac{10^3}{40}\right) \left(7.34 + 125e^{-T_{b,sol}^\circ} + \frac{13.88 \cdot 10^5}{T_{b,sol}^2}\right) \\
 c_{p,sol} &= c_{p,NaOH}x_{NaOH} + (1 - x_{NaOH}) \cdot 4180 \\
 i_{v,sol} &= c_{p,sol}T_{b,sol}^\circ + l_v \\
 T_{b,sol}^\circ &= T_{b,H_2O}^\circ + \Delta = T_{b,H_2O}^\circ + 0.0162 \frac{T_b^2}{l_v} \Delta^0
 \end{aligned}$$

where  $c_{p,sol}$  is the heat capacity of the NaOH solution,  $T_{b,sol}^\circ$  is the boiling temperature of the solution at a certain pressure  $p$ ,  $i_{v,sol}$  is the enthalpy of the solution,  $K_T$  and  $A_T$  are the heat transfer coefficient and heat transfer area, respectively,  $T_{st}^\circ$  is the temperature of the steam at 12 bar,  $l_v$  is the water latent heat of vaporization,  $x_{i(F,1,NaOH)}$  is the mass fraction of the components in different points,  $\Delta$  is the ebullioscopic temperature increase at pressure  $p$ ,  $\Delta^0$  is the ebullioscopic temperature increase at atmospheric pressure, and  $T_{b,H_2O}^\circ$  is the boiling temperature of water.

Parameters calculated by means of a computer program are given in Tab. 12.1.

**Tab. 12.1:** Operational parameters of the evaporator in Example 12.1 when operating under steady state conditions.

No.	$x_F$ (%)	$p$ (bar)	$F$ ( $\frac{kg}{s}$ )	$F_{st}$ ( $\frac{kg}{s}$ )	$p_{st}$ (bar)	$T_F^\circ$ (°C)	$x_1$ (%)	$V_1$ ( $\frac{kg}{s}$ )	$F_1$ ( $\frac{kg}{s}$ )
1	10	5.5	37	10	12	140	15.68	13.40	23.60
2	8	5.5	37	10	12	140	12.71	13.70	23.30
3	10	6.0	37	10	12	140	15.32	12.85	24.15
4	10	5.0	37	10	12	140	16.06	13.95	23.05
5	10	5.5	40	10	12	140	14.91	13.17	26.83
6	10	5.5	37	12	12	140	18.04	16.49	20.51
7	10	5.5	37	10	8	140	10.84	2.87	34.13
8	10	5.5	37	10	12	140	15.69	13.42	23.68



The nominal operation regime parameter values are presented in the first row (row 1) of the table.

It is clear that a change of 20% (row 2) of the input concentration is affecting with approximately 19% the output concentration. At the same time, if a change of the feed flow rate occurs (+8% – row 5), the concentration decreases with 5%. This demonstrates the importance of a ratio control system with inferential recalculation of the ratio  $F_{st}/F$  when such variations occur.

At the same time, when a change in pressure occurs ( $\pm 9\%$  – rows 3 and 4), the decrease/increase in concentration at the output is of  $\mp 2.4\%$ . To evaporate at this pressure the same quantity of solvent in order to obtain the desired nominal concentration, the additional consumption of steam (row 3) is of +0.55 kg/s. This translates into a supplementary heat consumption of 1 Gcal/h (at  $l_v = 525$  kcal/kg). At an amount of 7200 hours of operation/year and at a cost of 70 USD/Gcal, this means an additional expenditure of ~500,000 USD/year. The importance of a tight pressure control is hence demonstrated.

The steam pressure variation influences substantially the quality of the product (row 8): 25% of steam pressure decrease produce a reduction of mass concentration from 15.68% to 10.84% (cca. 30%), which means the steam station has to have a tight output pressure control.

One has to look at the important aspect of technical specifications. Let us suppose that the plant has to deliver NaOH solution at 15.68%. If an accidental increase of 20% of the steam flow rate occurs (row 7), the consequence is the unnecessary concentration of the solution to 18.04% with a supplementary heat consumption of 1837 kcal/s. If the producer is obliged to dilute the solution at the required mass fraction, the supplementary energy is lost together with its costs. The sensitivity to disturbances is an important issue. The efficiency of the control scheme resides not only in the decrease of the energy consumption but also in the stability of the quality of the final product in the conditions of variable feed rate and heat and/or composition input.

By considering the general case of  $n$  effects, and by generalizing eq. (12.2), the following is obtained:

$$Fx_F = F_n x_n, \quad (12.9)$$

The overall material balance includes all vapor flows removed from each effect,

$\sum_{i=1}^n V_i$ ; thus:

$$F = F_n + \sum_{i=1}^n V_i. \quad (12.10)$$

The total vapor flow to be extracted in order to reach the desired concentration is obtained when combining eq. (12.9) and (12.10):

$$\sum_{i=1}^n V_i = F \left( 1 - \frac{x_F}{x_n} \right), \quad (12.11)$$

so that:

$$x_n = x_F \left( 1 - \frac{\sum_{i=1}^n V_i}{F} \right) \quad (12.12)$$

The differential of above expression as a function of  $\frac{\sum_{i=1}^n V_i}{F}$ , indicates the sensitivity of the final concentration relative to the steam/feed ratio  $\left(\frac{V_0}{F}\right)$ .

$$\frac{dx_n}{d\left(\frac{\sum_{i=1}^n V_i}{F}\right)} = \frac{x_F}{\left(1 - \frac{\sum_{i=1}^n V_i}{F}\right)} = \frac{x_n^2}{x_F} \quad (12.13)$$

If one wants to see which is the sensitivity of the concentration at a given percentage of steam/feed ratio, one may calculate as follows:

$$\frac{dx_n}{d\left(\frac{\sum_{i=1}^n V_i}{F}\right)\left(\frac{\sum_{i=1}^n V_i}{F}\right)} = \frac{x_n(x_n - x_F)}{x_F} \quad (12.14)$$

At the same time, the sensitivity of the process can be determined not only by the steam/feed ratio, but also by the input/output (final) concentration ratio  $\frac{x_n}{x_F}$  as well. This is the main reason for the variation of product quality. The differential as a function of concentration is:

$$\frac{dx_n}{dx_F} = \frac{1}{1 - \frac{\sum_{i=1}^n V_i}{F}} = \frac{x_n}{x_F} \quad (12.15)$$

It shows that the feed concentration change has a greater influence than that of the steam supply change. In such cases, the inferential control scheme illustrated by Fig. 12.2.(b) can be used.

The extended approach of evaporation process control is described in [4].

Evaporators are used not only in concentrating “orange juice” or “caustic soda”, but quite recently in air conditioning (AC) as well, in automotive industries and in households. The reason is simple: in Europe, 75%–80% of the new cars are fitted with AC systems, as compared to 12% in 1990, while in the USA “surveys by the Department of Energy in USA suggest that the yearly energy expenditure for residential air-conditioning in 1997 was 0.42 Quadrillion BTU (~0.1 Quadrillion kcal)” and still continues to rise [5]. Chapter 5 of this volume presents new methods of multi-variable adaptive control. More recent work of the center for AC at the University of Illinois is described in both [5] and [6]. In [7], a fuzzy control algorithm is used in the same field of interest.

## References

- [1] Agachi, P. Ş., Constantinescu, D. M., Macedon, D., Oniciu, L., Topan, V., Neacsu, I., *Contributions at the diminishing energetic consumption of the brine electrolysis process*, Paper in the volume of the 11-th Meeting the Scientific Research Center Rm.Vilcea 10–12 October 1995, 1–12.
- [2] McCabe, L. W., Smith, J. C., Harriot, P., *Unit operations of Chemical Engineering*, McGraw-Hill Book, 1993, 465.
- [3] Shinskey F. G., *Energy conservation through control*, Academic Press, (1978), 195.
- [4] Shinskey F. G., *Energy conservation through control*, Academic Press, (1978), 183.
- [5] Shah, R., Alleyne, A. G., Bullard, C. W., Rasmussen, B. P., Hrnjak, P. S., *Dynamic Modeling And Control of Single and Multi-Evaporator Subcritical Vapor Compression Systems*, Air Conditioning & Refrigeration Center, Mechanical & Industrial Engineering Dept. University of Illinois, 2003, 1.
- [6] Rasmussen, B. P., Alleyne, A. G., *Dynamic Modeling and Advanced Control of Air Conditioning and Refrigeration Systems*, ACRC TR-244, 2006, 207.
- [7] Wua, C., Xingxib, Z., Shiminga, D., *Development of control method and dynamic model for multi-evaporator air conditioners (MEAC)*, Energy Conversion and Management, **46**(3), 451–465, 2005.

## 13 Control of drying processes

*Drying* is the separation of *moisture* from a solid substance through exposure to heated air or nonsaturated gas in (most cases). Another option is to use the heat produced in electric dryers (e.g. washing machines, drying chambers) by respecting Joule's first law. There are other nonconventional drying procedures [1], such as freezing, microwave exposure, and dielectric drying. These are not approached here. There are multiple options for drying, depending on the specific characteristics of the material to be dried [1]. Hereby, we discuss processes for which the drying agent is air or gas, but also mention some recent developments.

In terms of energy consumption, drying is a highly energy-intensive operation and represents 10% to 25% of the national industrial energy production in developed countries [1]. Therefore, it is natural to find control solutions to save money. The reports mention savings between 1.5% and 30% of total costs, depending on how close to the optimum the process is operated [7].

Notions of *psychrometry* are given in [2]. The drying is either continuous (e.g. tile-drying tunnels) or in batches (e.g. electrical insulator or fruit drying chambers). The heat necessary for drying is transferred to the drying agent before it comes into contact with the humid solid to be dried. In this situation, the drying is adiabatic.

The relative humidity depends on the pressure at which the drying process takes place [2]. Consequently, the measurements of relative humidity through measuring the dew point can be applied only at constant pressure. Once the pressure decreases (e.g. an atmospheric front), the relative humidity increases for a given dew point. The humidity content of solid materials is in equilibrium with the relative humidity of the air. Inside drying chambers, the environment is far from being in equilibrium with the solids to be dried, hence creating the drive for drying. From the thermal transfer point of view, the solvent/water evaporation rate is proportional with the temperature difference between the solid and its drying agent. The temperature of the drying agent is equal to the dry-bulb temperature ( $T_{db}^{\circ}$ ) while that of the solid is closer to the wet-bulb temperature ( $T_{wb}^{\circ}$ ), because the wet bulb is a wetted solid on which the liquid evaporates adiabatically. Thus,  $T_{wb}^{\circ}$  is significant from the point of view of the drying rate.

It is easy to measure  $T_{wb}^{\circ}$  of the ambient, but it is a real problem to measure it at continuously high temperatures as in the drying chambers/tunnels because they favor the fast evaporation of water in the wet-bulb. Therefore,  $T_{wb}^{\circ}$  can be estimated either when knowing  $T_{db}^{\circ}$  and the absolute humidity by means of an iterative process [3], or by empirically found expressions based on measurement data [4]:

$$T_{wb,c}^{\circ} = T^{\circ} \tan^{-1} [0.151977(\varphi\% + 8.313659)^{\frac{1}{2}}] + \tan^{-1}(T^{\circ} + \varphi\%) - \tan^{-1}(\varphi\% - 1.676331) + 0.00391838(\varphi\%)^{\frac{3}{2}} \tan^{-1}(0.023101\varphi\%) - 4.686035 \quad (13.1)$$

$T^{\circ} = T_{db}^{\circ}$  is the ambient temperature. The relative humidity can be found from psychrometric charts as in Fig. 7.61 from [2]. For this calculus, a first guess of  $T_{wb,0}^{\circ}$  is required to estimate the difference  $T^{\circ} - T_{wb,0}^{\circ}$  and thus the first approximation

of  $\varphi_0\%$ . If  $T_{wb,c}^\circ = T_{wb,0}^\circ$ , the calculus ends; if not, the value of  $T_{wb,0}^\circ$  is replaced with the calculated value and the wet-bulb temperature is recalculated once again until the difference between  $T_{wb,c}^\circ$  at step  $j$  and  $T_{wb,c}^\circ$  at step  $j-1$  is less than  $0.1^\circ\text{C}$  [3].

The drying rate  $R$  depends on the air speed along the solid, the dimension of its particles, its water content, the dryer characteristics, and the temperature difference  $T^\circ - T_{wb}^\circ$ . When a solid particle absorbs solvent, its surface is wet. In a dry ambient the liquid starts to evaporate. If  $T^\circ - T_{wb}^\circ$  is kept constant, the evaporation rate of the liquid is constant until dry spots on the particle appear. This is the constant rate drying period that lasts until the critical moisture content  $w_c$  is reached. It represents the humidity of a system's atmosphere above which a crystal of a water-soluble salt will always become damp (absorb moisture from the atmosphere) and below which it will always stay dry (release moisture to the atmosphere) (Fig. 13.1).

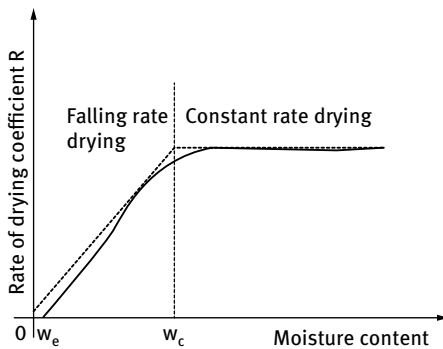


Fig. 13.1: Evolution of the drying rate  $R$  as a function of the moisture of the solid.

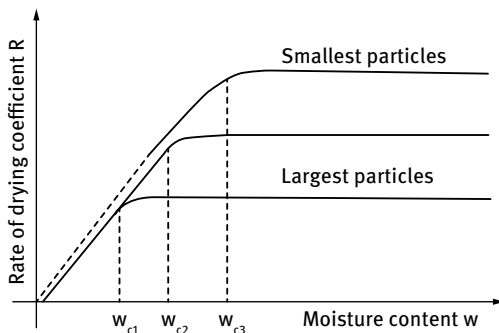


Fig. 13.2: The drying rate as function of the particle size of dried material.

Since dry spots appear on the surface, the drying rate decreases because the liquid, in order to be evaporated, has to diffuse from the interior to the solids surface. The rate decreases to 0 when the absolute humidity of the solid reaches the equilibrium state with the environment,  $w_c$ . The critical moisture content depends on the dimension of the particles [5, 6]. Thus, for particles with a diameter of  $d_1 = 90 - 160\mu$ ,  $w_c = 5\%$ ; for particles with  $d_2 = 60 - 90\mu$ ,  $w_c = 10\%$ ; and for particles with  $d_3 < 60\mu$ ,  $w_c = 21\%$ , respectively (Fig. 13.2).

## 13.1 Batch drying control

### 13.1.1 Conventional batch drying control

Many products are dried batch-wise: fruits and vegetables, construction materials (sand or bricks), electrical insulators, pharmaceutical products, etc. An example of a batch dryer is given in Fig. 13.3.

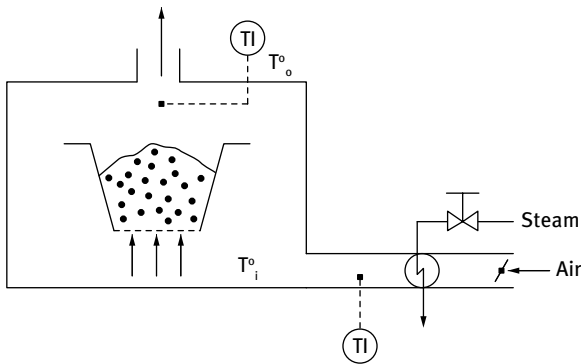


Fig. 13.3: Scheme of a batch dryer.

During a batch drying, both solid and outlet air temperatures increase in time (Fig. 13.4).

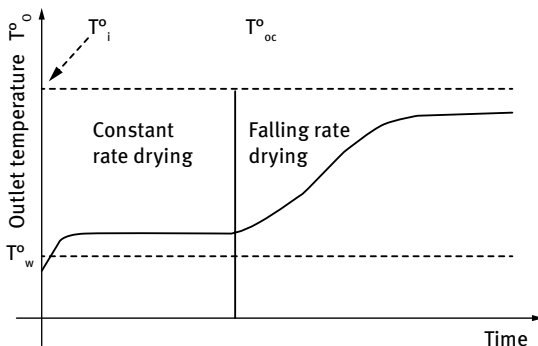


Fig. 13.4: Drying agent temperature profile during batch drying.

If the solid's temperature is smaller than  $T_{wb}^\circ$ , part of the received heat is used to raise the temperature to the value of  $T_{wb}^\circ$ . This phenomenon causes a slight increase of the outlet temperature  $T_o^\circ$ , and after that, during the constant rate drying period, the outlet temperature stays at the value imposed by the surface evaporation,  $T_{oc}^\circ$ . During the last period of variable drying rate, the temperature of the heating agent increases asymptotically to the dryer's inlet temperature  $T_i^\circ$ .

The heat balance during the constant rate drying period is:

$$F_g c_{pg} (T_i^\circ - T_o^\circ) = V l_v \quad (13.2)$$

where  $F_g$  is the drying agent mass flow (air),  $c_{pg}$  is the drying agent specific heat,  $V$  is the vapor mass flow,  $l_v$  is the latent heat of vaporization.

The vaporization rate  $dV$  from a particle may be written as:

$$dV = k_m R \cdot dA (T^\circ - T_{wb}^\circ) \quad (13.3)$$

and is proportional with the drying rate  $R$ , heat transfer area  $dA$ , mass transfer coefficient  $k_m$ , and with the temperature difference between  $T^\circ$  and  $T_{wb}^\circ$ .

As long as the air passes over the particle to be dried, the temperature decreases due to the evaporation:

$$F_g c_{pg} dT^\circ = -dV l_v \quad (13.4)$$

From both eqs. (13.3) and (13.4) it results that:

$$dA = \frac{F_g c_{pg}}{k_m R l_v} \cdot \frac{dT^\circ}{T^\circ - T_{wb}^\circ}, \quad (13.5)$$

After integration the relationship above becomes:

$$A = \frac{F_g c_{pg}}{k_m R l_v} \ln \left( \frac{T_i^\circ - T_{wb}^\circ}{T_o^\circ - T_{wb}^\circ} \right) \quad (13.6)$$

In the period of variable drying rate (Fig. 13.4),  $R = kx \leq kx_c$  and from (13.6) the following can be obtained:

$$x = \frac{F_g c_{pg}}{k_m k A l_v} \ln \left( \frac{T_i^\circ - T_{wb}^\circ}{T_o^\circ - T_{wb}^\circ} \right) \quad (13.7)$$

Equation (13.7) contains impossible to evaluate constants,  $k_m$  and  $k$ . Therefore, in the constant drying rate zone,  $x = x_c$  and:

$$\frac{k_m k A l_v}{F_g c_{pg}} x_c = K x_c = \ln \left( \frac{T_i^\circ - T_{wb}^\circ}{T_{oc}^\circ - T_{wb}^\circ} \right) \quad (13.8)$$

Thus, eq. (13.7) can be solved for any output temperature at which a certain humidity  $x^*$  can be attained:

$$x^* = \frac{1}{K} \ln \left( \frac{T_i^\circ - T_{wb}^\circ}{T_o^\circ - T_{wb}^\circ} \right) \text{ or} \quad (13.9)$$

$$T_{oc}^{\circ*} = T_{wb}^\circ + e^{-Kx^*} (T_i^\circ - T_{wb}^\circ) \quad (13.10)$$

Eliminating  $T_{wb}^\circ$  from eq. (13.8) it results that:

$$T_o^{\circ*} = T_{oc}^\circ \left( \frac{1 - e^{-Kx^*}}{1 - e^{-Kx_c}} \right) + T_i^\circ \left( 1 - \frac{1 - e^{-Kx^*}}{1 - e^{-Kx_c}} \right) \quad (13.11)$$

or,

$$T_o^{o*} = K^* T_{oc}^o + (1 - K^*) T_i^o \quad (13.12)$$

This way, the temperature control system should implement eq. (13.11) or its simplified version, eq. (13.12), to stop the drying process at the desired value  $T_o^{o*}$ . One condition to be satisfied is that the inlet temperature  $T_i^o$ , has to be kept constant (meaning an inlet temperature control).

### Example 13.1 [7]



Let us consider the batch drying of a material with  $x_c = 0.1$  to the desired moisture content of  $x^* = 0.02$ . The inlet temperature is  $T_i^o = T_{i,db}^o = 93^\circ\text{C}$  and  $T_{wb}^o = 34^\circ\text{C}$ . The outlet temperature during the constant rate drying period is  $T_{oc}^o = 40^\circ\text{C}$ . What is the value of the temperature,  $T_{o1}^o$ , at which drying should be stopped? What happens if the material-specific area is twofold smaller? What happens if, in this second case, the temperature at which the drying is stopped is  $T_{o1}^o$ ?

From eq. (13.7):

$$K = \frac{1}{0.1} \ln \left( \frac{93 - 34}{40 - 34} \right) = 22.81$$

From eq. (13.11):

$$K^* = \frac{1 - e^{-22.81 \cdot 0.02}}{1 - e^{-22.81 \cdot 0.1}} = \frac{0.367}{0.898} = 0.408$$

From eq. (13.12):

$T_{o1}^{o*} = 0.408 \cdot 40 + (1 - 0.408) \cdot 93 = 72^\circ\text{C}$ . This is the temperature at which we stop drying.

If the specific area of the solid material diminishes twofold,  $K$  (eq. (13.7)) decreases to its half. Thus (eq. (13.10)):

$$T_{oc,2}^o = 34 + e^{-\frac{22.81}{2} \cdot 0.1} (93 - 34) = 52.8^\circ\text{C}. K^* \text{ stays at the same value and then:}$$

$$T_{o2}^{o*} = 0.408 \cdot 52.8 + (1 - 0.408) \cdot 93 = 76.6^\circ\text{C}.$$

This is the temperature at which the drying has to be stopped in the second case.

The attained moisture content (eq. (13.8)) is:

$$x_2^* = \frac{2}{22.81} \cdot \ln \left( \frac{93 - 34}{76.6 - 34} \right) = 0.028$$

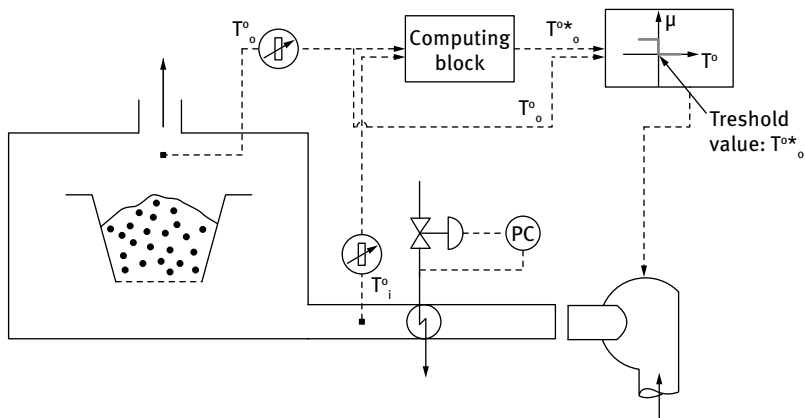
If the “stopping” temperature is kept at  $72^\circ\text{C}$  in the second case, then:

$$x^* = \frac{2}{22.81} \cdot \ln \left( \frac{93 - 34}{72 - 34} \right) = 0.038,$$

and a higher than desired moisture content is retained by the solid.

Shinsky suggests that the control system should be inferential and should calculate  $T_o^{o*}$  in order to interrupt the drying process when  $T_o^o \geq T_o^{o*}$  (Fig. 13.5). It has to be mentioned that in addition to the “stopping” scheme, the input temperature has to be kept constant by means of a temperature control system (to respect eq. (13.12)).





**Fig. 13.5:** Control scheme for stopping the drying process at a certain desired value of the moisture in the solid to be dried.

Several suggestions of control are given in [8]. Dufour identified control solutions that are currently in industrial use. Our research in advanced control drying [9, 10] proposed solutions that targeted cost reduction while preserving the standard quality of the dried material.

### 13.1.2 Advanced batch drying control

#### **i** Example 13.2 [9]

The high-voltage electric insulator production implies a two-stage batch drying process. During the first step, the moisture content of the target product is reduced from 18%–20% to 0.4% in special gas-heated chambers. The second step is carried out in high-temperature ovens, to achieve an even lower moisture content. In [9], two advanced algorithms (fuzzy and MPC) are proposed to control the drying process of the insulators. The original notations are kept within this example.

In Fig. 13.6, both the structure and the operation of the drying chamber are described.

Nomenclature:

$x$  – moisture content of the air coming from the insulator

$x_f$  – moisture content of the burned gases

$x_o$  – moisture content of the outflow gases

$x_{ext}$  – moisture content of the air absorbed in the burner

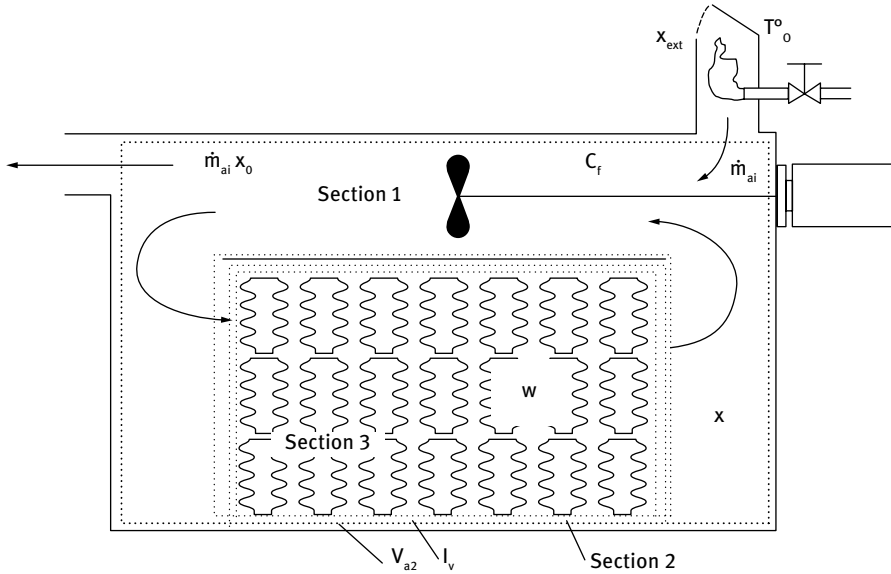
$w$  – moisture content of the solid

$w_c$  – critical moisture content of the solid

$\dot{m}_a$  – mass outflow rate of burned gas

$\dot{m}_{ai}$  – mass flow rate of fresh air

$\dot{m}_{st}$  – mass flow of evaporated steam



**Fig. 13.6:** Scheme of the drying chamber. Section 1 represents the air volume within the drying chamber; Section 2 represents the direct surroundings of the drying product; Section 3 represents the drying product itself.

$m_s$  – mass of the dried solid

$V_{ach}$  – free volume of the drying chamber

$V_{a2}$  – air volume in section 2

$A_s$  – evaporation area of the solid

$A_{ch}$  – heat transfer area of the drying chamber

$T_i^\circ$  – input temperature of the burned gases

$T_o^\circ$  – output temperature of the mixture of gases and steam from the drying chamber

$T_{ext}^\circ$  – outer temperature

$K_A$  – heat transfer coefficient through the walls of the drying chamber

$\rho$  – density of the gas mixture in the immediate vicinity of the solid subjected to drying

$\rho_a$  – density of the mixture of burned gas and vapor in the drying chamber

$l_v$  – latent heat of vaporization

$\dot{V}_F$  – flow rate of the methane gas burned in the burner

$H_F$  – heat of combustion

Mass and energy balance equations are used to describe the dynamic behavior of the system. The main studied outputs of the model are: the moisture content of the drying product  $w$ , the outlet air temperature  $T_o^\circ$ , and the air humidity  $x_o$ . The input variables are the natural gas flow rate  $\dot{V}_F$  and the mass flow rate of fresh air  $\dot{m}_{ai}$ . The mass balance of steam within section 1 is described by:

$$\dot{m}_{ai}x_f + \dot{m}_a x - (\dot{m}_a + \dot{m}_{ai})x_o = V_{ach}\rho_a \frac{dx_o}{dt}$$

with  $V_{ach}$  being the volume of air in section 1. In section 2, the steam flows around the drying product are modeled by:

$$\dot{m}_a(x_o - x) - m_s \frac{dw}{dt} = \frac{d}{dt}(V_{a2}\rho x)$$

with  $V_{a2}$  being the infinitesimal small volume of air in section 2 and  $m_s \frac{dw}{dt}$  the steam flow coming from the solid. Because of this, the last term of the equation above can be neglected, which results in the differential equation:

$$\frac{dw}{dt} = (x_o - x) \frac{\dot{m}_a}{\dot{m}_s}$$

In section 3, the behavior of the drying good itself is described with a normalized diagram by means of the following equation [10, 11]:

$$\frac{dw}{dt} = -\frac{\dot{m}_a}{\dot{m}_s} A_s$$

The drying velocity during the three periods of the entire drying process of a hygroscopic material is characterized by the diagrams in Fig. 13.7 [11].

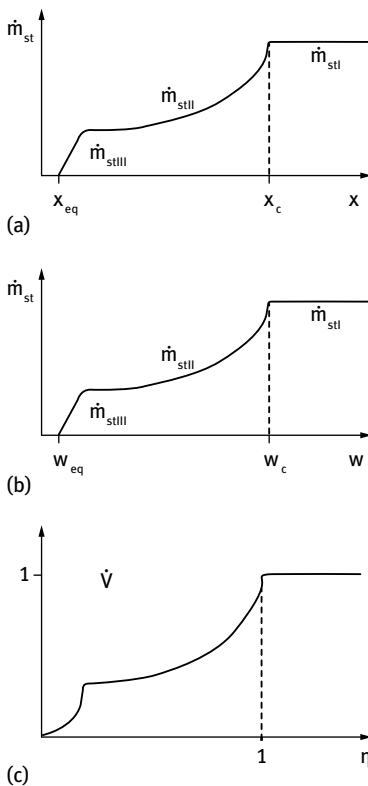


Fig. 13.7: Drying rate depending on the moisture content of the drying product: (a) absolute and, (b) normalized.

Diagram (a), only available by experiments and valid for certain conditions, can be normalized to (b) according to:

$$\dot{v}(\eta) = \frac{\dot{m}_{st}}{\dot{m}_{stl}} \text{ and } \eta = \frac{w - w_{equ}}{w_c - w_{equ}}$$

It is assumed that  $w_c$  is constant and does not depend on the drying conditions and that  $w_{equ}$  only depends on the relative air humidity, without being affected by other factors. It is also assumed that all diagrams of the drying velocity for different drying conditions are geometrically similar.

The equilibrium humidity  $w_{equ}$  dependence on the relative air humidity  $\phi$ , was described for clay by means of a correlation equation. The saturation humidity of the air,  $x_{sat}$ , is dependent on the temperature  $T_o^\circ$ . For low partial pressures of steam, a simple equation for  $\dot{m}_{stl}$  (Fig. 13.7) was considered:

$$\dot{m}_{stl} = k(x_{sat} - x)$$

where the mass transfer coefficient  $k$  is determined experimentally.

Two energy balance equations, one for the drying chamber:

$$\begin{aligned} \dot{m}_{ai}[c_{pa}(T_i^\circ + T_o^\circ) + x_f(l_v + c_{pst}T_i^\circ) - x_o(l_v + c_{pst}T_o^\circ)] + m_s \frac{dw}{dt}(l_v + c_{pst}T_o^\circ) - \\ K_A A_{ch}(T_o^\circ - T_{ext}^\circ) = V_{ach} \rho_a [(c_{pa} + x_o c_{pst}) \frac{dT_o^\circ}{dt} + c_{pst} \frac{dx_o}{dt} T_o^\circ] \end{aligned}$$

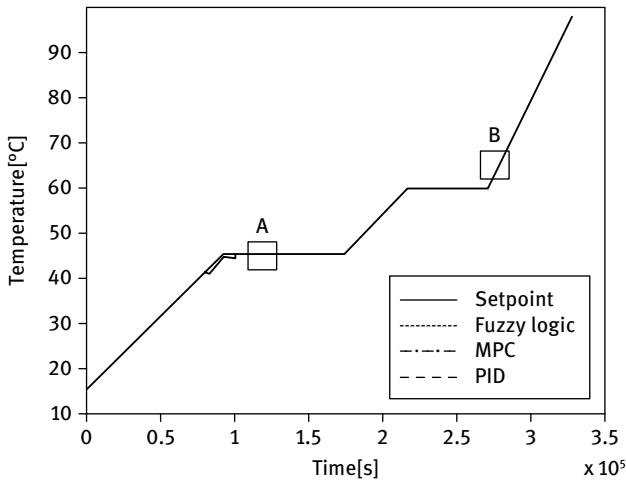
and the other for the burner:

$$\begin{aligned} \dot{m}_{ai}[c_{pa}T_{ext}^\circ + x_{ext}(l_v + c_{pst}T_{ext}^\circ)] + (c_{pF}T_{ext}^\circ + H_F) \frac{M_{FPF}}{R(T_{ext}^\circ + 273)} \dot{V}_F \\ = \dot{m}_{ai}[c_{pa}T_i^\circ + x_f(l_v + c_{pst}T_i^\circ)] \end{aligned}$$

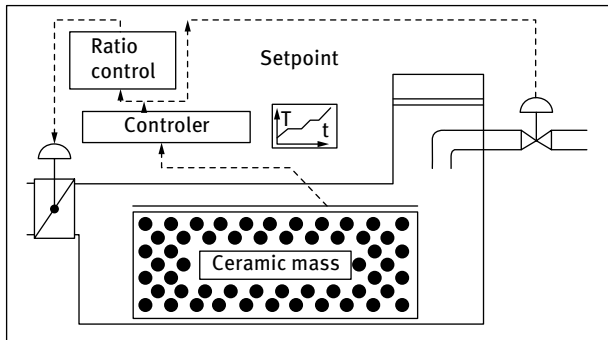
are used to describe the outlet temperature change.

The fuzzy controller proposed for the batch drying chamber receives a measured value from the system, it fuzzifies it (assigns it a membership value), applies the system's rules, computes an overall result of all the rules and then defuzzifies the result, converting it into a number which is an appropriate command for the system it controls (see also chapter 3). The mission of the model predictive controller (see also chapter 2) is accomplished by anticipating (predicting) the outputs of the system with the aid of the mathematical model. The controller output is generated based on the anticipated behavior of the system. All control methods investigated in this paper obey the current control practice, i.e. driving the evolution of the moisture content of the drying product in the desired way by means of controlling the air temperature inside the chamber. Usually, the desired decreasing profile of the drying product moisture content is obtained by imposing an increasing ramp-constant profile on the air temperature. The setup of the simulated system is shown in Fig. 13.8.

The scaled dynamic sensitivity analysis of the output variables with respect to the studied inputs pointed out the natural gas flow rate as the most important manipulated variable (about 10 times more important than the mass flow rate of fresh air). The control system was designed accordingly. The control scheme is presented in Fig. 13.9.



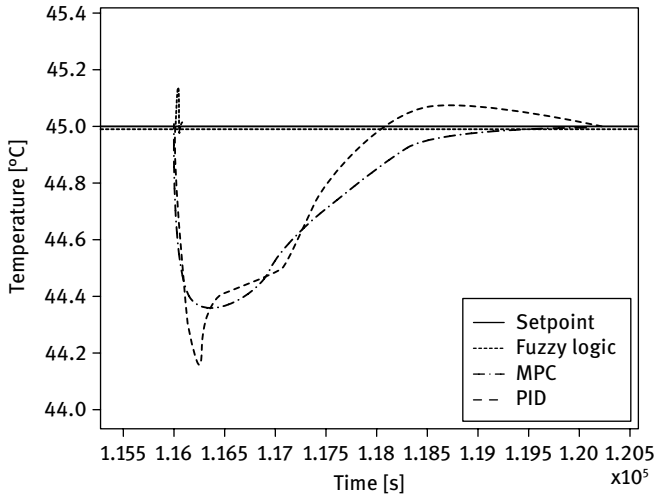
**Fig. 13.8:** Comparative behavior of FL, MPC, and PID control in the presence of the heating power disturbance. A and B are the below magnified windows corresponding to constant temperature and to ramp temperature change respectively.



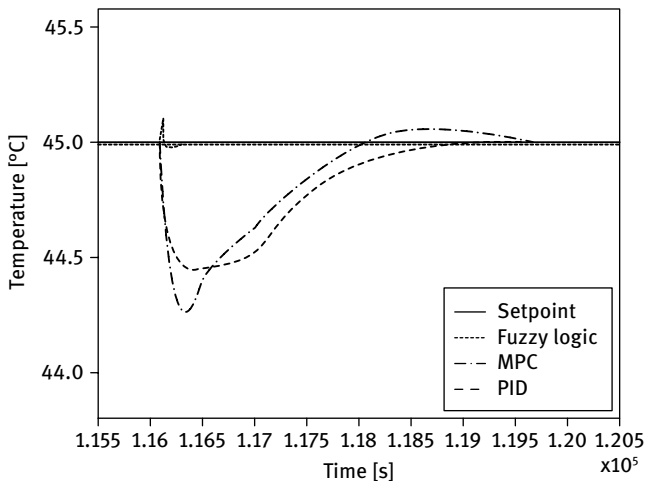
**Fig. 13.9:** Structure of the control system of the drying chamber described in Example 13.2.

The authors studied the comparative behavior of three control algorithms: PID, fuzzy logic (FL), and MPC; the results are presented in Figs. 13.10–13.12.

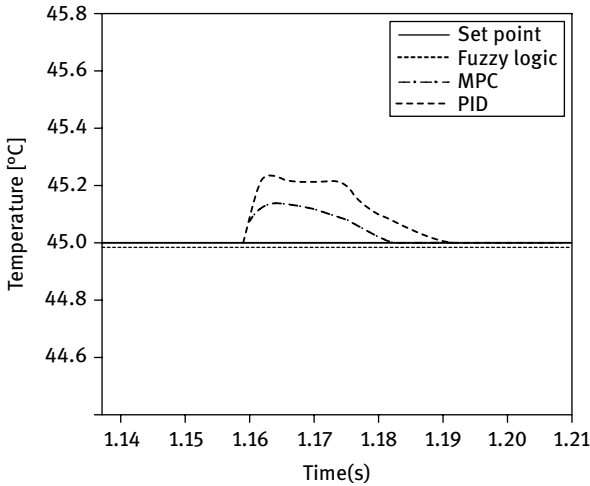
The performance tests were carried out for three significant disturbances typically occurring in industrial practice: a  $10^\circ\text{C}$  inlet air temperature  $T_{ext}^o$  drop (from  $16^\circ\text{C}$  to  $6^\circ\text{C}$ ), a 10% heating power capacity  $H_F$  drop of the natural gas, and a 10% rise in the moisture content of the inlet air. The disturbances were applied as steps at time  $t = 116,000$  s.



**Fig. 13.10:** Comparative behavior of FL, MPC, and PID control in the presence of a heating power disturbance (detailed).

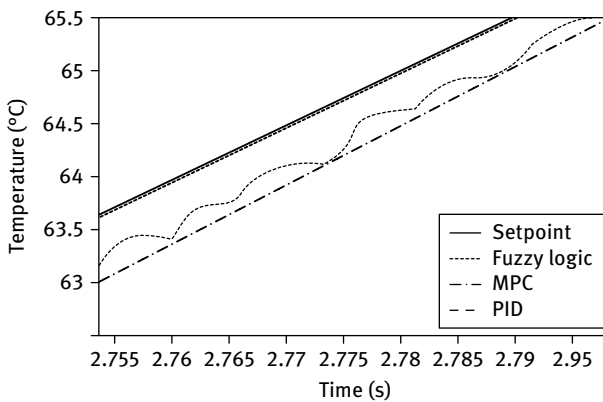


**Fig. 13.11:** Comparative behavior of FL, MPC, and PID control in the presence of an air inlet temperature drop disturbance (detailed).



**Fig. 13.12:** Comparative behavior of FL, MPC, and PID control in the presence of an air inlet humidity increase disturbance (detailed).

With respect to setpoint tracking performance, the results reveal a good behavior in case of PID and MPC, yet FL control proves superior abilities. The figure below illustrates that FLC is very accurate; it follows with precision both the constant and the ramp sections of the temperature setpoint scheduling function (Fig. 13.13).



**Fig. 13.13:** Detailed presentation of the ramp setpoint following performance of FLC, MPC, and PID control.

All control methods exhibit a low offset behavior for the constant parts of the setpoint function. For the ramp sections, as in Fig. 13.13 (detail B on Fig. 13.8), MPC and PID control proved to be less accurate than FL by showing a larger offset.

The conclusion of the authors, when referring to the FL controller performances, are that the accuracy of FL control is largely due to the asymmetrical membership function definition. It takes into account the need for an asymmetric amplitude of the manipulated variable change (i.e. a controller response of higher amplitude to a negative error compared to a lower amplitude response for a positive error) in the ramp section of the setpoint function. With respect to disturbance rejection performance, FL control showed a considerably shorter (more than 10 times) response time and smaller (more than five times) overshoot than other control strategies. It is worth mentioning that these new techniques are extremely valuable when food is processed, so that nutritive proprieties of goods are preserved [12, 13].

### 13.2 Continuous adiabatic drying

Continuous drying can be carried out in fluidized beds or in longitudinal dryers. Fig. 13.14 presents the control scheme of a fluidized bed dryer [14].

The air flow is maintained in a way that ensures particle's fluidization. At the same time, the air flow rate should be kept at a value that does not entrain solid particles to the exit. The differential pressure is a measure of fluidization and it is used to control the air flow rate. The humid product enters the dryer, it is homogenized inside it and leaves the equipment at the opposite side in a dry form. If the product to be dried has at exit a higher moisture content than the critical value, then the dryer is not self regulated; consequently, it functions in the constant drying rate regime and according to eq. (13.8), the output temperature depends on the critical moisture content. In this situation, the control system presented in Fig. 13.14 cannot control the product humidity.

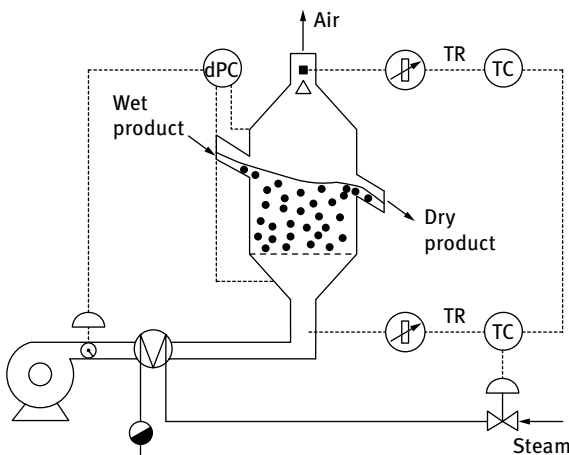
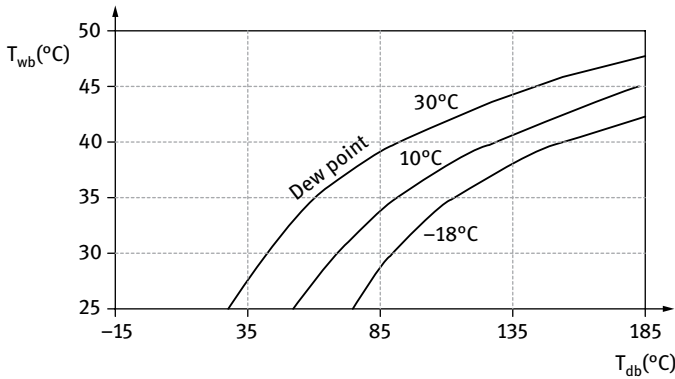


Fig. 13.14: Control scheme of a continuous fluidized bed dryer.



Normally, the product leaves the dryer with  $w < w_c$  in such a way that the output humidity corresponds to the value expressed in eq. (13.7). It is necessary to ensure the residence time imposed by diagram 13.15.



**Fig. 13.15:** Wet bulb temperature ( $T_{wb}^{\circ}$ ) as a function of the dry bulb temperature ( $T_{db}^{\circ}$ ) and the dew point.

The increase of the material inflow rate, or of its moisture content, will induce an increase of the air humidity  $x$ . The temperature controller manipulates the butterfly valve in order to increase  $T_i^{\circ}$  with the consequence of increasing  $T_o^{\circ}$  to the desired value. The deficiencies of this control system are explained through eq. (13.7): at a higher evaporation load (rate), both  $T_i^{\circ}$  and  $T_{wb}^{\circ}$  are increased. If  $T_o^{\circ}$  is kept constant,  $x$  has to increase to satisfy eq. (13.7). To keep  $x$  at  $x^*$ ,  $T_o^{\circ}$  has to be controlled at the value of  $T_o^{*\circ}$  expressed by eq. (13.12), but for which value of  $T_{wb}^{\circ}$  has to be either measured or estimated (see eq. (13.10) and approximation formula (13.1)). For water, the diagrams in Fig. 13.15 can be used; the curves cover the temperatures of the dew point in the temperate zone.

The moisture of the fluidized bed can be controlled either through the heat inflow or through the solid material flow rates; usually, the solid material is more difficult to be manipulated and thus the input temperature is controlling the moisture content.

The curves in Fig. 13.16, show that the output temperature can be approximated with a straight line:

$$T_o^{*\circ} = aT_i^{\circ} + b \quad (13.13)$$

with the value of  $b$  adjusted to the values of the dew point [15]:

$$b = b_0 + f(T_{dew}^{\circ})$$

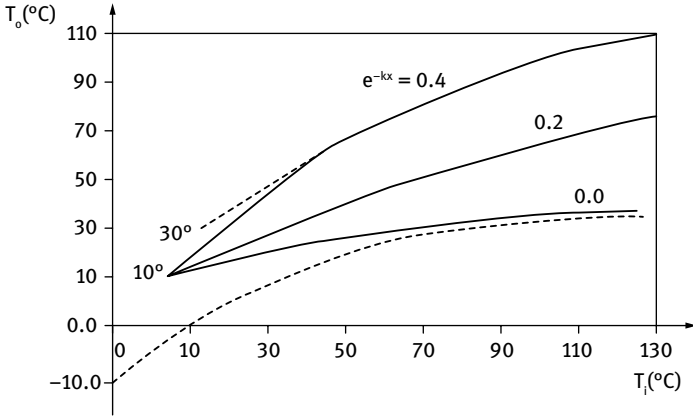


Fig. 13.16: The solutions of eq. (13.11) for  $T_{wb}^o = 10^\circ\text{C}$  and different values of  $e^{-Kx}$ .

To implement the control scheme, a calibration to determine the values of  $e^{-Kx}$  in the operating conditions of the dryer is needed:

$$e^{-Kx} = \frac{T_o^o - T_{wb}^o}{T_i^o - T_{wb}^o} \tag{13.14}$$

The automatic control system capable to keep constant the desired moisture content is illustrated in Fig. 13.17.

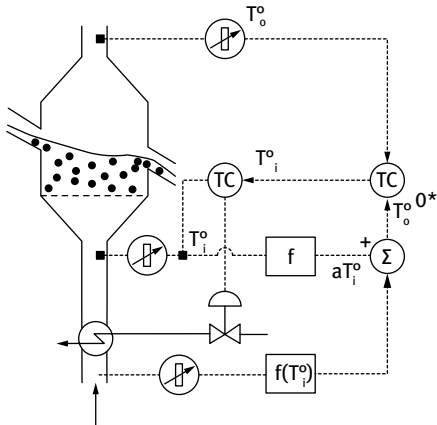


Fig. 13.17: Automatic control system of the moisture content for a solid material.

As one may see, the control of the moisture content is based on indirect measurements, temperatures, and moreover, some of them have to be estimated. Recently, researchers from UMIP [16] have developed a tool to allow the online measurement

and consequently control of moisture in a fluidized bed dryer. The system uses electrical capacitance tomography (ECT) to take images of gas-solid distribution and to provide online moisture measurements. This way, the setpoint can be fixed directly to the desired value of  $w^*$ .

More recent developments refer to freeze-drying for sensitive biological substances. Researchers from Italy, USA, and Japan [17, 18] reported the monitoring and control of freeze-drying for pharmaceuticals at very low temperatures (around  $-40^{\circ}\text{C}$ ).

## References

- [1] Mujumdar, A., Editor, *Handbook of Industrial Drying*, Fourth Edition, CRC Press, Taylor & Francis Group, 4, 2015.
- [2] Agachi, P.S., Cristea, V.M., *Basic Process Engineering Control*, Walter de Gruyter GmbH, Berlin/Boston, 217, 2014.
- [3] Martinez, A.T., *On the evaluation of the wet bulb temperature as a function of dry bulb temperature and relative humidity*, *Atmosfera*, 7, 179, 1993.
- [4] Stull, R., *Wet-bulb temperature from relative humidity and air temperature*, American Meteorological Society, 2267, 2011.
- [5] Fonyó, Z., Fábry, G., *Vegyipari műveletani alapismeretek (Basic knowledge in operations in industrial chemistry)*, Nemzeti Tankönyvkiadó, Budapest, 905, 2004.
- [6] Perry, R., Green, D., Maloney, J., Eds., *Perry's Chemical Engineer's Handbook*, McGraw-Hill, Ch. 12–34, Table 12.7, 1999.
- [7] Shinskey, F.G., *Energy Conservation through Control*, Academic Press, New York, 212, 1978.
- [8] Dufour, P., *Control engineering in drying technology: review and trends*, special issue of *Drying Technology on Progress in Drying technologies* (5), 24(7), 889–904, 2006.
- [9] Bâldea, M., Cristea, V.M., Agachi, P.S., *A fuzzy logic approach to the control of the drying process*, *Hungarian Journal of Industrial Chemistry*, 30, 167–179, 2002.
- [10] Van Meel, D.A., *Adiabatic convection batch drying with recirculation of air*, *Chemical Engineering Science*, 9, 36–44, 1958.
- [11] Krischer, O., Kast, W., *Die wissenschaftlichen Grundlagen der Trocknungstechnik*, Springer-Verlag, Berlin, 1992.
- [12] Cristea, V.M., Irimiță, A., Ostace, G., Agachi, P.Ș., *Control of forced convection drying in food slabs*, *Proceedings of the 22<sup>nd</sup> European Symposium of Computer Aided Process Engineering*, 17–20 June 2012, London, 932–936.
- [13] Yuzgeca, U., Beceriklib, Y., Turker, M., *Nonlinear predictive control of a drying process using genetic algorithms*, *ISA Transactions*, 45(4), 589–602, 2006.
- [14] Shinskey, F.G., *Energy Conservation through Control*, Academic Press, New York, 226, 1978.
- [15] Senbon, T., Hanabuchi, F., *Instrumentation Systems: Fundamentals and Applications*, Berlin, Chapter 9, 598, 1991.
- [16] Yang, W., Wang, H., *Methods and apparatus relating to fluidized beds*, University of Manchester, UMIP, US Patent 8,461,852 B2, 2013.
- [17] Barresi, A., Pisano, R., Rasetto, Valeria, Fissore, D., Marchisio, D., *Model-based monitoring and control of industrial freeze-drying processes: effect of batch nonuniformity*, *Drying Technology: An International Journal*, 28(5), 577–590, 2010.
- [18] Patel, S., Doen, T., Pikal, M., *Determination of end point of primary drying in freeze-drying process control*, *AAPS PharmSciTech*, 11(1), 73–84, 2010.

## 14 Control of crystallization processes

### 14.1 The process of crystallization

*Crystallization* is a separation, purification, and particle formation technique. It is widely used in the industry of fine chemicals and pharmaceuticals, but it is also used in many other industries. The driving force of crystallization is the thermodynamic instability caused by *supersaturation*. It means that the actual solute concentration is higher than the solubility under given thermodynamic conditions. Depending on the way supersaturation is generated, different types of crystallization are distinguished:

- *Crystallization by cooling* is based on the fact that solubility generally increases with temperature, thus decreases during cooling. If a warm saturated solution is cooled, it automatically becomes supersaturated – see the temperature drop from 90°C to 60°C in Fig. 14.1(a).
- *Anti-solvent crystallization (salting-out)* is based on the different solubility of the same species in various solvents or solvent mixtures. If another significantly lower solubility solvent is added continuously to a saturated solution prepared in a high solubility solvent, the mixture becomes supersaturated (see Fig. 14.1(b)) and the excess solute “salts out”. The low solubility solvent is called anti-solvent.
- *Reaction crystallization (precipitation)* occurs during a fast (usually ionic) chemical reaction that yields a hardly soluble product. Its solubility is generally some orders of magnitude lower than the expected product concentration; thus, at reaction endpoint, the mixture is supersaturated and the excess precipitates. An example is the synthesis of hydroxyapatite: the chemical reaction between ammonium phosphate and calcium nitrate, followed by the precipitation of calcium phosphate [1].

The supersaturation is quantitatively characterized either by the *supersaturation ratio* ( $S$ ) or the *relative supersaturation* ( $\sigma$ ) – see eq. (14.1). Defined as such, both ( $S$ ) and ( $\sigma$ ) are dimensionless  $C$  and  $C_s$  stand for the actual and the saturation solute concentrations, respectively.

$$\sigma = S - 1; S = \frac{C}{C_s} \quad (14.1)$$

Regardless of the supersaturation method, the crystallization process consists of two main mechanisms: *nucleation* and *crystal growth*. These coexist in the slurry and occur either successively and/or competitively.

Nucleation describes the emergence of new crystals. From the industrial point of view, there are two types: *primary* and *secondary* nucleation. Primary nucleation occurs when a new crystal is formed directly from the bulk solution from individually solvated molecules. Secondary nucleation is a collective name for the phenomena in

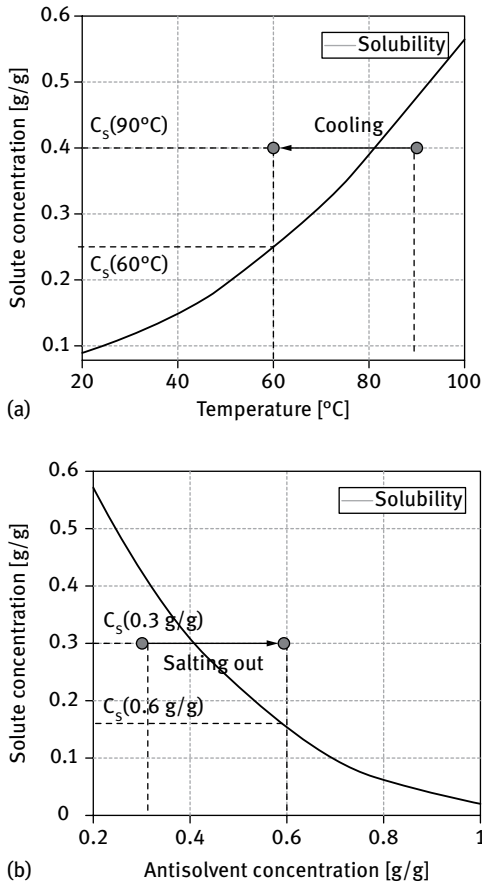


Fig. 14.1: Typical temperature (a) and solvent composition (b) dependencies of solubility.

which existing crystals generate new nucleons under the action of shear stresses or crystal-impeller/wall and crystal-crystal collisions. The primary nucleation occurs at the beginning of crystallization until the formation of the first crystals and is further gradually replaced by the secondary nucleation [2]. In continuous crystallizers of industrial importance, the secondary nucleation is governing.

Numerous models have been proposed to describe these mechanisms. Among them two popular equations for the primary and secondary nucleation rates are [3]

$$B_p = k_p \exp\left(-\frac{k_e}{ln^2(S)}\right) \exp\left(-\frac{E_{a,p}}{RT}\right) \quad (14.2)$$

$$B_s = k_s \sigma^b V_C^j \exp\left(-\frac{E_{a,s}}{RT}\right) \quad (14.3)$$

The significance of terms in eqs. (14.2)–(14.3) is:

- $B_p$ , primary nucleation rate ( $\#/m^3 \cdot s$ );
- $B_s$ , secondary nucleation rate ( $\#/m^3 \cdot s$ );
- $k_p$ , primary nucleation rate constant ( $\#/m^3 \cdot s$ );
- $k_s$ , secondary nucleation rate constant ( $\#/m^3 \cdot s$ );
- $k_e$ , a constant characterizing  $B_p$  dependence on supersaturation ( $\#/m^3 \cdot s$ );
- $b$ , a constant characterizing  $B_s$  dependence on supersaturation (dimensionless);
- $j$ , a constant characterizing  $B_s$  dependence on crystal volume (dimensionless);
- $V_c$ , fraction of total slurry volume occupied by formed crystals ( $m^3/m^3$ );
- $E_{a,p}$  and  $E_{a,s}$ , activation energy of primary and secondary nucleation, respectively ( $J/mol \cdot K$ ).

Equations above show that ( $B_p$ ) and ( $B_s$ ) depend on supersaturation ratio ( $S$ ), temperature, and some material/system-dependent constants. The latter may be sensitive to mixing conditions and sometimes to very small amounts (ppm order of magnitude) of impurities.

The nucleons and existing crystals are growing as long as the solution is supersaturated. The deposition of solvated molecules (or atoms/ions) on the crystal surface consists of three mass transfer steps:

- Convective transfer (diffusion) from the bulk solution toward the mass transfer boundary layer;
- Conductive transfer (diffusion) through the boundary layer toward crystal surface;
- Integration into the crystal lattice (deposition).

Deposition occurs according to two major types of surface integration mechanisms:

- *Monolayer* crystal growth: new layers are built only after the previous ones are completely finished;
- *Polylayer* crystal growth: more than one crystal layer is being built at the same time. This mechanism occurs at higher supersaturation ratios.

The precise mathematical description of crystal growth is complicated because it strongly depends on liquid phase mass transfer as well as on the nature of crystal growth. However, the following empirical formula can successfully be applied for the macroscopic description [3]:

$$G = k_g \sigma^g \exp\left(-\frac{E_{a,g}}{RT}\right) \quad (14.4)$$

where the significance of terms is:

- $G$ , linear crystal growth rate (m/s)
- $k_g$ , crystal growth rate constant (m/s)
- $g$ , a constant characterizing  $G$  dependence on supersaturation (dimensionless);
- $E_{a,g}$ , activation energy of crystal growth ( $J/mol \cdot K$ ).

Equation (14.4) shows that, similarly to nucleation, crystal growth depends on supersaturation, temperature, and system/material-specific constants. Generally speaking, the supersaturation dependency of crystal growth is lower than that of nucleation; hence, lower ( $S$ ) values favor crystal growth. For diffusion-controlled growth, the exponent  $g$  is around 1, but for a surface integration-limited growth, the  $g$  is near 2 [3].

The global crystallization rate is determined by both the nucleation and growth rates. If the nucleation rate is small and that of the growth is high, fewer crystals of bigger dimensions are formed. The reverse situation yields a high crystal number having smaller sizes. The goal of the crystallization process control is to adjust both nucleation and growth to achieve the desired purity and crystal size distribution of the product.

The analogue of the well-known continuous stirred tank reactor (CSTR) is the continuous mixed suspension mixed product removal (CMSMPR) crystallizer [3]. The liquid phase is assumed to be perfectly mixed at all three scales (micro-, meso- and macro-scale), the crystal size distribution (CSD) is identical in each point of the slurry volume and the suspension removal is unclassified (the CSD of the outlet stream is identical to that of the crystallizer slurry).

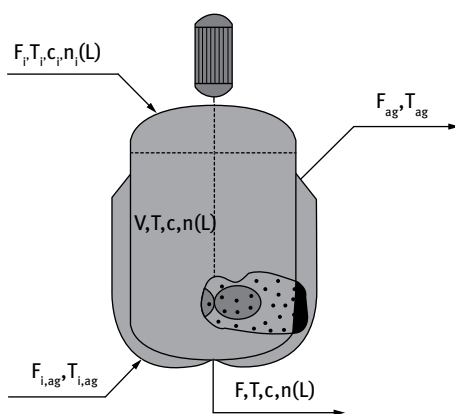


Fig. 14.2: Scheme of a CMSMPR crystallizer.

The scheme of a cooling CMSMPR crystallizer is presented in Fig. 14.2. The unit is cooled *via* a jacket and is continuously fed with an inlet stream ( $F_i$ , expressed in  $\text{m}^3/\text{s}$ ) of high temperature ( $T_i$ , expressed in either in  $^\circ\text{C}$  or  $\text{K}$ ) and solute concentration ( $C_i$ , expressed in  $\text{kg}/\text{m}^3$ ). The value of ( $C_i$ ) is close to that of the solubility, but still somewhat lower to avoid fouling (undesired deposition in pipes). CSD in the inlet stream is denoted with ( $n_i(L)$ ). The core temperature ( $T$ ) is considerably lower than that of the feeding stream, thus supersaturation is generated as illustrated by Fig. 14.1(a).

The crystal size distribution is characterized by the ( $n(L)$ ) size density (expressed in  $\#/\text{m}^3$ ) which gives the number of crystals having size  $L$  (expressed in  $\text{m}$ ). The value

refers to a volume of  $1 \text{ m}^3$ . Assuming a constant slurry volume  $V$  (expressed in  $\text{m}^3$ ), the population balance model governing the size distribution of crystals is described by [4]

$$\frac{\partial n(L)}{\partial t} + \frac{\partial Gn(L)}{\partial L} = \frac{F}{V} [n_i(L) - n(L)] + (B_p + B_s) \delta(L - L_n) \quad (14.5)$$

where the initial condition is set by

$$n_{in}(L|t=0) = n_0(L) \quad (14.6)$$

In eq. (14.5),  $L_n$  stands for the nucleon size. The subscript  $i$  stands for the inlet and  $in$  for the initial values. The first term on the left side of eq. (14.5) stands for the time evolution of CSD and the second term takes into consideration the effects of crystal growth. The right side of eq. (14.5) describes the effects of inlet and outlet streams on CSD as well as that of both nucleation rates. The Dirac-delta function  $\delta(L - L_n)$  of the last term in eq. (14.5) implies that nucleation affects only the number of crystals having sizes of  $L_n$ , but has no effect on the number of bigger crystals. In other words, the value of the function  $\delta(L - L_n) = 1$  if  $L = L_n$  and  $\delta(L - L_n) = 0$  if  $L \neq L_n$  (has the role of a boundary condition).

The solute mass balance in a CMSMPR is written as follows:

$$\frac{dC}{dt} = \frac{F}{V} [C_i - C] - 3k_v \rho_c \int_0^{\infty} L^2 Gn(L) dL \quad (14.7)$$

The initial condition is given by

$$C(0) = C_{in} \quad (14.8)$$

As mentioned above, solute concentrations are symbolized by  $C$ .  $\rho_c$  in eq. (14.7) stands for the crystal density (expressed in  $\text{kg}/\text{m}^3$ ) and  $k_v$  is a dimensionless volume shape factor used to compute the volume of crystals. The volume of a single crystal can be expressed in the function of its size  $L$  [3] as

$$V_c = k_v L^3 \quad (14.9)$$

Defined as such,  $k_v = 1$  for the cube and  $k_v = \pi/6$  for a sphere.

Process model equations (14.5) and (14.7) as well as information presented in Fig. 14.1(a) demonstrate that a cooling crystallization can be manipulated by means of temperature. It influences the process through solubility as well as through nucleation and growth rates. Because solubility decreases with temperature (see Fig. 14.1(a)) higher overall yields may be achieved when operating at lower temperatures. The feed flow rate also affects the resulting CSD: high flow rates translate into shorter average residence times; thus, crystals do not have a long enough time to grow, and as a result, the yield drops.

Consequently, the control of a CMSMPR crystallizer consists of the following automatic control systems (loops):

- a temperature control loop to adjust solubility as well as rates of crystal nucleation and growth;



- an inlet flow rate control loop to adjust the average residence time;
- a level control loop to ensure the constant slurry volume.

Such a control structure is presented in Fig. 14.3.

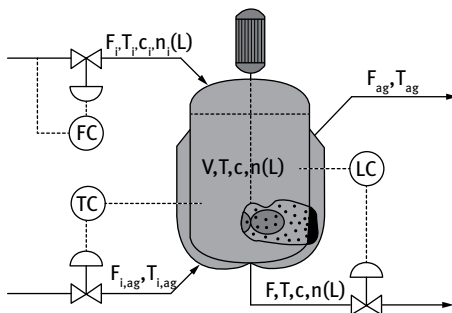


Fig. 14.3: General control structure of a CMSMPR crystallizer.

## 14.2 Crystal size distribution control

Several macroscopic properties of particulate products depend on crystal size distribution: specific surface, porosity, biological activity, dissolution rate, some mechanical properties, etc. Moreover, the CSD affects the downstream operations such as filtration, granulation/milling, drying. Therefore, achievement of target CSD already within the crystallization stage, eliminates downstream crystal size adjustment, reduces production cost, and lowers environmental impact. More importantly, it ensures higher-quality crystals since these secondary size-modifying operations generally do not result in flat and clear crystal facets. Consequently, CSD control is an important aspect of modern chemical industry. Its significance becomes obvious also in terms of overall economic efficiency if the remaining solution after filtration is concentrated and recirculated to the crystallizer.

State-of-the art CSD control strategies can be divided into two major groups:

- *Model-free* control strategies
- *Model-based* control strategies

Crystallization, as a required size particle formation technique, is mainly applied in the fine chemical and pharmaceutical industry. These fields of production are dominated by discontinuous batch crystallizers. Thus, within the following, focus is on batch crystallization control. The goals resume to maintaining of either an optimum cooling profile or an optimum anti-solvent addition rate (or both) to achieve the desired CSD while respecting various economic and technological constraints.

### 14.2.1 Model-free crystal size distribution control

The quick spread of *process analytical technologies* (PAT) enables the fast, easy and accurate online monitoring of the crystallization process. The real-time measurements cover the tracking of liquid and solid phase properties using a combination of the following PAT instruments:

1. *Focused beam reflectance measurement* (FBRM). The FBRM is a specific transducer of particulate systems. It emits a laser beam within a sensor inserted in the slurry. The beam rotates with a specified frequency during which it “intersects” (meets) the crystals. The length of this intersection, called *chord length* (CL) is then detected by the sensor. Up to a few thousands of CLs are measured per second. The number of FBRM count is proportional with the crystal number per volume unit ( $\#/m^3$ ). The CL distribution (CLD) can be constructed from the length of individual CL values and is proportional with the CSD. It is like a fingerprint of the particulate system. The measuring principle of FBRM is presented in Fig. 14.4.

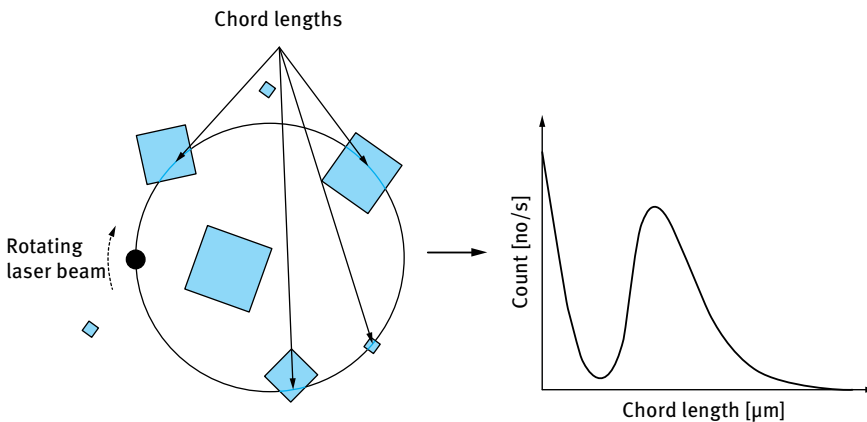
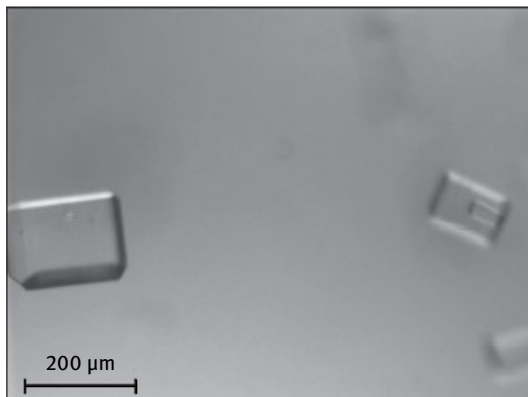


Fig. 14.4: Measuring principle of FBRM and a typical chord length distribution (CLD).

2. *Process vision microscopy* (PVM). The PVM is an *in situ* process microscope that is able to record images generally in the 1- to 1000- $\mu\text{m}$  size range. Various tools of image analysis can be applied to analyze the recorded images. Similarly to the FBRM, a relative crystal number can be extracted from these. Moreover, shape information can also be gathered after a more thorough image analysis. A typical PVM image is presented in Fig. 14.5.



**Fig. 14.5:** Typical PVM image, captured during the crystallization of an inorganic material, Loughborough University, 2014.08.12.

3. The temperature is monitored by means of conventional temperature transducer.
4. Depending on the nature of the solid and liquid phases, the solute concentration is usually assessed *via* online UV/VIS/NIR, FTIR, or Raman spectrometers.

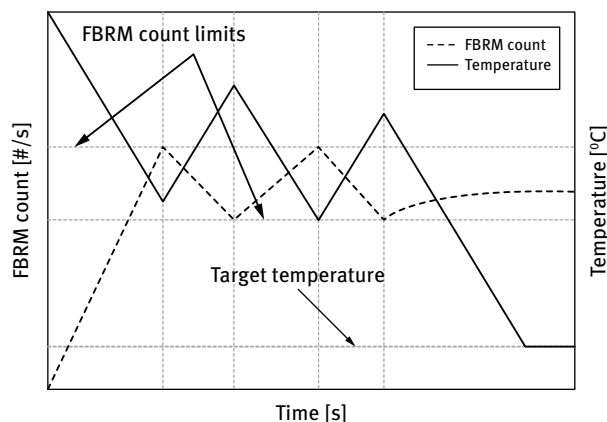
#### Direct nucleation control (DNC)

The DNC is a model-free PAT-based control strategy applied for batch cooling crystallization processes. The main idea behind the DNC is the fact that after a certain degree of cooling the crystallized mass remains constant, so that the average crystal size is a function of the crystal number. Variation of the crystal number results in the adjustment of the crystal size: if fewer crystals are produced they will inevitably be bigger and *vice-versa*. The controlled variable is the relative count provided by the FBRM detector.

The working principle of DNC is simple [5]. Crystallization is started with a high temperature saturated solution; it is cooled with a specified rate. When supersaturation is reached (see Fig. 14.1(a)), nucleation and growth of crystals begin. Meanwhile, the CSD is monitored *via* FBRM. When the upper relative FBRM count limit is exceeded, a heating stage begins. Because of the heating, the solution becomes undersaturated and the dissolution of crystals commences. This reduces the crystal number, and thus the FBRM count. According to the Oswald ripening phenomenon [2], the smaller crystals dissolve faster, and thus the FBRM count decreases faster as compared to the global (mass) dissolution. When FBRM count lower limit is reached, the second cooling loop is started. It triggers again supersaturation.

Within this second cooling stage, in contrast to the first, some crystals already exist in the slurry and start to grow immediately. As it has already been mentioned, the lower supersaturation favors the growth; thus, as long as the growth of these

crystals consumes supersaturation, the nucleation is not favored. However, a too slow cooling prolongs the batch time; hence, higher cooling rates are recommended. Yet, this may lead to enhanced nucleation, which in return leads again to an FBRM count increase. Consequently, heating and cooling are alternatively repeated until the final target temperature and a relatively stable FBRM count (translated into CSD) are reached. Fig. 14.6 presents the typical time diagram of a DNC control.



**Fig. 14.6:** Presentation of the DNC concept: the controller maintains the FBRM count between the predefined limits by heating and cooling cycles.

The main advantage of DNC is that if the appropriate settings are applied (cooling and heating rates), a stable control can be achieved in most cases without any prior system information (solubility, nucleation, growth kinetics, etc.). It is robust to disturbances generally caused by impurities or by deposit detachments from the crystallizer wall or stirrer. However, in some situations, high batch times may result and CSD is indirectly affected *via* crystal number. In other words, the average crystal size and number can be satisfactory, but the dispersion may be high and the crystal shape can hardly be controlled.

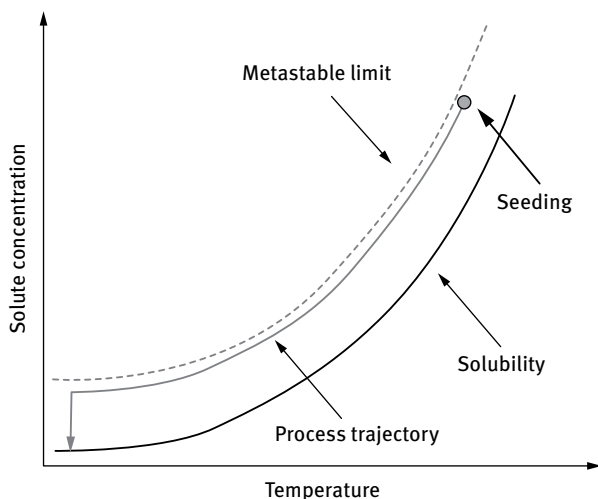
### The supersaturation control (SSC)

Figure 14.7 illustrates in other terms the practical observation that low supersaturation favors crystal growth: a second characteristic line can be defined and tracked in the phase diagram, at higher than solubility values. It is called the *metastable limit* [6]. Under it, nucleation is negligible, but crystal growth occurs. The zone between the metastable limit and the solubility line is called the *metastable zone*.

The main idea behind the supersaturation control (SSC) is the addition of *seed crystals*, in a well-specified quantity and size distribution, to a solution characterized by the metastable zone. The supersaturation is further controlled by manipulating

the temperature or the anti-solvent amount to keep the process in this zone. Improved growth rate is ensured by keeping supersaturation near to the metastable limit. Thus, similarly to the DNC, crystal size distribution is controlled indirectly.

Numerous studies [7] showed that a well-designed SSC can perform almost as well as optimal control. SSC conducts the process in the vicinity of the maximum allowed supersaturation, and thus eliminates the formation of small crystals and shortens batch time [5] as compared to the DNC. Meanwhile, it requires preliminary knowledge of solubility and metastable limit. The accurate and proper determination of these may be experimentally time consuming.



**Fig. 14.7:** The SSC concept: the process is conducted near the metastable limit, which favors crystal growth without generating significant nucleation.

### Combined DNC and SSC

Hybrid techniques combine the advantages of DNC and SSC [6]. For example, the process starts with control by applying DNC for *in situ* seed crystal generation. Further, these are grown faster *via* employment of SSC. This approach benefits from both the robust crystal number control of DNC and the efficiency of SSC. Moreover, it avoids the necessity of seeding.

The market offers a variety of software designed especially for process control of crystallization. For example, the commercially available CryMOCO was put forward for PAT-based crystallization and includes both DNC and SSC strategies [5].

The presented model-free CSD control strategies have their strengths, yet will not always be able to satisfy constraints imposed by crystallization non-linearity. Therefore, better performing model based control that explicitly includes the target CSD is needed [6].

### 14.2.2 Model-based crystal size distribution control

The model-based crystal size distribution control strategies use the mathematical model of crystallization to find the optimal temperature profile or anti-solvent addition rate for the achievement of target crystal quality (yield, purity, CSD, shape, etc.) with respect to given constraints. The process model generally consists of the population balance governing the CSD and the mass balance for the solute – see eq. (14.5) and (14.7), respectively. Kinetic laws (14.2)–(14.4) are either available in the literature or have to be assessed from experimental data; they can rarely be predicted by molecular modeling. Therefore, adequate model architecting and validation often involve significant engineering effort and constitutes a major part of control structure design.

However, once the model is available, crystallization control can be divided into open-loop and closed-loop model-based control. Various goals can be formulated for example for the batch crystallizer (also see chapter 2 on MPC): batch time minimization, crystal size distribution optimization, cost reduction, etc. Multiobjective goals can also be formulated. However, due to the nonlinear feature of crystallization, a nonlinear MPC (N-MPC) is required to provide reasonable controller performance [6]. Chapter 2 is dedicated to model predictive control, thus only a general description of these concepts' application to batch crystallization control will be subsequently presented.

#### Open-loop model-based control

In the case of any open-loop control, the process output is not taken into account in the calculation algorithm. The control signal (temperature profile or the anti-solvent addition rate) is determined by solving a process optimization problem that employs the model. By optimizing the control signal, the constraints are explicitly taken into account and the CSD can be directly manipulated. The advantage of open-loop model based (optimal) control over the PAT-based technique is the possibility of constraint satisfaction, direct CSD manipulation, and cost reduction (no need for expensive instrumentation). The major disadvantage of this approach is that the process may be extremely sensitive to disturbances; without feedback, the process can easily depart from the optimal path after relative small disturbances (the control signal is often pre-defined and optimization is carried out before the process starts).

#### The closed-loop online MPC

In the case of closed-loop MPC, the process model is used in *real time* to optimize the control signal with respect to global performance requirements. This permits incorporating harmful effects of disturbances as well as model uncertainties into the control strategy. This generally results in better control performance. Yet, the process simulation has to be fast enough to enable a real time optimization.

During the duration of a batch cycle, data are available from the initiation of the process until the certain moment when the measurement is carried out. These can be used in the MPC algorithm to improve its performance. The state model incorporates

CSD values, yet these cannot be measured online by any PAT tool presented in section 14.2.1. Thus, state estimators are needed to assess the unmeasurable system states [8]. For a batch crystallizer that operates with a fixed batch time, the prediction horizon is decreasing, but the quantity of experimental data is increasing during the duration of a batch cycle. Therefore, the MPC is referred to as *shrinking horizon MPC* and the state estimator as to *receding horizon estimator* (RHE). Figure 14.8 presents this principle of a closed-loop MPC for a cooling batch crystallization with state estimation.

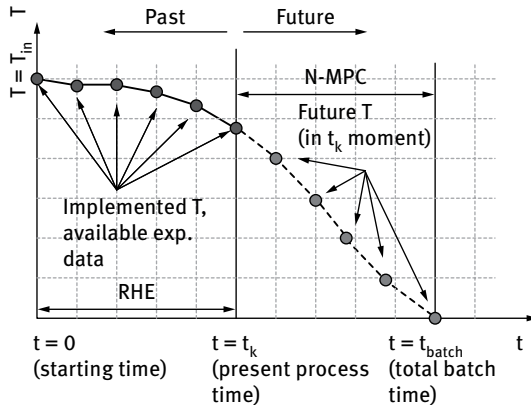


Fig. 14.8: The scheme of batch crystallization MPC: a shrinking horizon MPC with RHE.

The RHE readjusts the model in each MPC-involved iteration to fit the available experimental data, then gives the starting point for the MPC simulation (control signal optimization). As a result, the MPC uses a model that describes most appropriately the process from the beginning to the actual process time. Thus, the configuration of shrinking horizon MPC coupled with a receding horizon estimator is a powerful tool and is expected to give good control performance [9]. However, this method is computationally demanding, so that computational power can often be limiting in its application. The block diagram of such a controller is presented in Fig. 14.9.

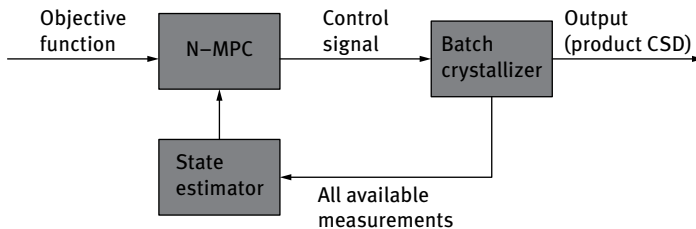


Fig. 14.9: The block diagram of a shrinking horizon MPC with receding horizon RHE applied to the control of batch crystallizers.

It is worth mentioning that the most sophisticated model-free and model-based techniques are both able to control not only crystal size distribution, but also polymorph transformations and co-crystallization. These strategies are based on the difference between the solubility of various of polymorphic forms of the same species or various solutes, respectively [10].

### Example 14.1



A chemical species having temperature dependent solubility is crystallized in a batch unit. The solubility is described with the following equation. It is valid within 20°C and 80°C.

$$C_s [\text{kg/kg}] = 0.223 - 9.6 \times 10^{-4}T + 1.17 \times 10^{-4}T^2 \quad (14.11)$$

The unit is charged with a 100-kg saturated solution at 60°C. Let us calculate the expected mass of crystals if the filtration and drying losses are of 5% each and the final batch temperature is 20°C.

According to eq. (14.11), the solubility concentrations at the initial and final operating temperatures are

$$\text{for } (T = 20); C_s = 0.223 - 9.6 \times 10^{-4} \times 20 + 1.17 \times 10^{-4} \times 20^2 = 0.2506 \text{ kg/kg}$$

$$\text{for } (T = 60); C_s = 0.223 - 9.6 \times 10^{-4} \times 60 + 1.17 \times 10^{-4} \times 60^2 = 0.5866 \text{ kg/kg}$$

The difference between the two values is

$$\Delta C_s = 0.336 \text{ kg/kg}$$

It represents the quantity of crystallized matter, when 100% crystallization yield is assumed, for 1 kg solution that is cooled from 60°C to 20°C. Then the quantity of crystallized matter for the entire batch unit is:

$$m_c = 100 \times 0.336 = 33.6 \text{ kg}$$

After taking into account the losses of filtration and drying (5% each), the final product quantity is obtained.

$$m_{\text{prod}} = 33.6 \times 0.95^2 = 30.324 \text{ kg}$$

### Example 14.2



A CMSMPR (see Fig. 14.2) crystallizes a species for which the solubility is described by eq. (14.11). After a total loss of 15% in the downstream operations, the plant yields a 250-kg/h product. The feed flow rate is of 1000 kg/h, has a temperature of 80°C, and a solute concentration with 5% lower than the corresponding solubility. The unit operates at 30°C. Let us calculate the yield of crystallization. What operational changes can be proposed to improve its value?



The solubility concentrations at the input and operating temperatures are

$$\text{for } (T = 30); C_s = 0.223 - 9.6 \times 10^{-4} \times 30 + 1.17 \times 10^{-4} \times 30^2 = 0.2995 \text{ kg/kg}$$

$$\text{for } (T = 80); C_s = 0.223 - 9.6 \times 10^{-4} \times 80 + 1.17 \times 10^{-4} \times 80^2 = 0.895 \text{ kg/kg}$$

The inlet solute concentration at 80°C can be calculated based on the inlet solubility.

$$C_{in} = 0.95 C_s = 0.95 \times 0.895 = 0.801 \text{ kg/kg}$$

The inlet mass is

$$m_{in} = C_{in} F = 0.801 \left[ \frac{\text{kg}}{\text{kg}} \right] \times 1000 \left[ \frac{\text{kg}}{\text{h}} \right] = 801 \left[ \frac{\text{kg}}{\text{h}} \right]$$

The hourly crystal mass production can be calculated as follows:

$$m_{out} = \frac{250}{0.85} = 294.12 \left[ \frac{\text{kg}}{\text{h}} \right]$$

If a 100% crystallization yield is assumed, the productivity at 30°C would be

$$m_{out,100} = (C_{in} - C_s) F = (0.801 - 0.2995) \times 1000 = 501.5 \left[ \frac{\text{kg}}{\text{h}} \right]$$

By considering the real production of 294.12 kg/h, the value of crystallization yield can be computed.

$$\eta = \frac{m_{out}}{m_{out,100}} 100 = \frac{294.12}{501.5} = 58.65\%$$

Yield improvement can be achieved by

- reducing the operating temperature to increase supersaturation;
- increasing the average residence time, either by reducing the feed thus yield or by increasing the crystallizer volume.

## References

- [1] Szilágyi, B., Muntean, N., Barabás, R., Ponta, O., Lakatos, B.G., *Reaction precipitation of amorphous calcium phosphate: population balance modelling and kinetics*, Chemical Engineering Research and Design, 93, 278–286, 2015.
- [2] Mersmann, A., *Crystallization technology handbook*, Drying Technology, 13, 1037–1038, 2001.
- [3] Randolph, A., Larson, M., *Theory of Particulate Processes, Analysis and Techniques of Continuous Crystallization*, Academic Press, Salt Lake City, UT, 1988.
- [4] Ramkrishna, D., *Population Balances*, Academic Press, London, 2000.
- [5] Nagy, Z.K., Fevotte, G., Kramer, H., Simon, L.L., *Recent advances in the monitoring, modelling and control of crystallization systems*, Chemical Engineering Research and Design, 91, 1903–1922, 2013.
- [6] Nagy, Z.K., Braatz, R.D., *Advances and new directions in crystallization control*, Annual Review of Chemical and Biomolecular Engineering, 3, 55–75, 2012.

- [7] Qinglin, S., Hermanto, M.W., Braatz, R.D., Chiu, M.S., *A new extended prediction self-adaptive control (EPSAC) strategy for batch control*, AIChE Annual Meeting, Conference Proceedings, 2012.
- [8] Mesbah, A., Huesman, A.E.M., Kramer, H.J.M., Van den Hof, P.M.J., *A comparison of nonlinear observers for output feedback model-based control of seeded batch crystallization processes*, Journal of Process Control, 21, 652–666, 2011.
- [9] Forgione, M., Birpoutsoukis, G., Bombois, X., Mesbah, A., Daudey, P.J., Van den Hof, P.M.J., *Batch-to-batch model improvement for cooling crystallization*, Control Engineering Practice, 41, 72–82, 2015.
- [10] Hermanto, M.W., Chiu, M.S., Braatz, R.D., *Nonlinear model predictive control for the polymorphic transformation of L-glutamic acid crystals*, AIChE Journal, 55, 2631–2645, 2009.

## 15 Problems and exercises

### 15.1 Advanced process control



1. A first-order reaction takes place in a cascade of three equal-dimension CSTRs. Concentration transducers are placed after each reactor, and they measure the concentration with a dead time  $\tau = 1$  min. An important load disturbance is the feed reactant concentration. The final concentration is modified *via* input flow to the reactor. Compare the performances (stability, overshoot) of a feedback control system with those of a cascade control with (a1) inner measurement done after the second reactor; (a2) inner measurement done after the first reactor.

Technical data are:  $V_1 = V_2 = V_3 = 1 \text{ m}^3$ ,  $F_1 = F_2 = F_3 = 3 \text{ m}^3/\text{h}$ , and  $k_1 = k_2 = k_3 = 0.1 \text{ h}^{-1}$ ; conversion measurement is done by each transducer with a dead time of 1 min; the transducer gains are of 2 kmol/mA each; the control valve gain and time constant are  $12 \text{ m}^3/\text{h}\cdot\text{mA}$  and 0 s, respectively.



2. The automatic control system (ACS) from Fig. 15.1 is tuned for a quarter amplitude decay ratio. Compare in terms of stability and steady-state error, the performances of a simple feedback control system – see controller  $K_1$  in Fig. 15.1, with that of the cascade of two controllers presented in Fig 15.1.

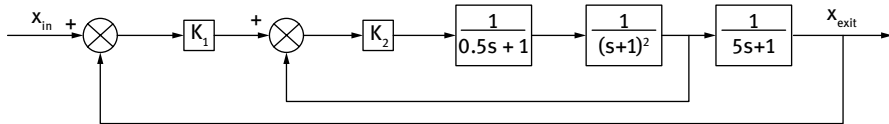


Fig. 15.1: A cascade ACS.



3. The ACS for concentration presented in Fig. 15.2 has the following characteristics: chemical reaction  $A \rightarrow B$  is of first order with the rate constant  $k = 0.5 \text{ min}^{-1}$ ; the residence time in each reactor is of  $V/F = 2$  min; the molar concentrations in steady-state are  $C_{A1} = 0.4 \text{ kmol/m}^3$ ;  $C_{A2} = 0.2 \text{ kmol/m}^3$ ;  $C_{A3} = 0.1 \text{ kmol/m}^3$ ;  $C_{AO} = 1.8 \text{ kmol/m}^3$ . The main load disturbance is the variation of input concentration entering through the feed tank placed ahead of the reactors and having a time constant of 5 min. The composition in the tank is measured with a composition disturbance transducer with the gain 1 and dead time 2 min. The actuating device, which controls the composition through the concentrate flow, has the total transfer function  $H_{AD}(s) = 1$ .

Synthesize the disturbance controller transfer function by knowing that the manipulating variable is the concentrate flow.

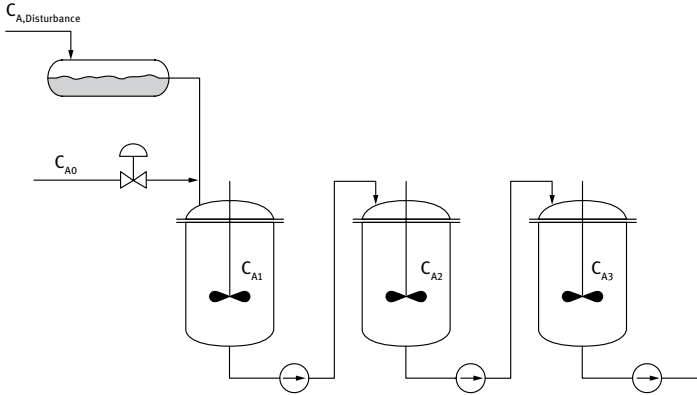


Fig. 15.2: Cascade of 3 CSTR units.

4. The transfer functions of a binary distillation column are the following:

$$\frac{X_D(s)}{F(s)} = \frac{K_F e^{-\tau_r s}}{(T_F s + 1)^2}; \quad \frac{X_D(s)}{X_F(s)} = \frac{K_X e^{-\tau_x s}}{(T_X s + 1)^2}; \quad \frac{X_D(s)}{R(s)} = \frac{K_R e^{-\tau_r s}}{(T_R s + 1)^2}$$

where  $x_D$  is the output that has to be kept constant,  $F$  and  $x_F$  are the disturbances, and  $R$  is the manipulated variable. Carry out a numerical simulation with the following values:  $K_F = 0.51 \frac{\text{kmol/m}^3}{\text{kmol/min}}$ ,  $K_X = 0.12 \frac{\text{kmol/m}^3}{\text{kmol/m}^3}$ ,  $K_R = 0.22 \frac{\text{kmol/m}^3}{\text{kmol/min}}$ ,  $T_F = 5 \text{ min}$ ,  $T_X = 6 \text{ min}$ ;  $T_F = 2 \text{ min}$ ,  $\tau_F = \tau_X = 0.5 \text{ min}$ , and  $\tau_R = 0.1 \text{ min}$ . Steady-state values are  $F = 1 \frac{\text{kmol}}{\text{min}}$ ,  $R = 0.4 \frac{\text{kmol}}{\text{min}}$ ,  $x_F = 0.5$ , and the disturbances are 10% of the nominal values. The transfer functions of the transducers and control valves are included in the given transfer functions.

5. Determine the transfer functions that link  $Y$  to  $Y_1$  within the interacting control systems presented in Fig. 15.3. Calculate the damping factor  $\zeta$  if  $T_1 = 8 \text{ min}$ ,  $T_2 = 8 \text{ min}$ ,  $K_1 = K_2 = 10$ ,  $K_3 = 3$ , and  $K_4 = 5$ . Calculate the transfer functions of the decoupling elements.

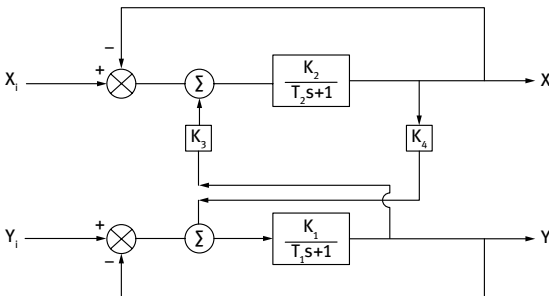


Fig. 15.3: Two interacting control systems.



6. Determine the transfer function of the disturbance controller for the heat-transfer device in Fig. 15.4, when the main disturbance is  $T_{in}^{\circ}$ . The temperature transducer has the following transfer function:  $H(s) = H_{DT}(s) = \frac{0.12 \text{ mA}/^{\circ}\text{C}}{0.024s + 1}$ , whereas of the steam flow-regulating valve is  $H_{AD}(s) = \frac{5.261 \text{ l/s}/\text{mA}}{0.083s + 1}$ . For the transfer process on the path  $F_{ag, in} \rightarrow T_{out}^{\circ}$ , it is  $H_{pr m}(s) = \frac{2^{\circ}\text{C}/\text{l/s}}{(0.017s + 1)(0.432s + 1)}$  whereas for the  $T_{in}^{\circ} \rightarrow T_{out}^{\circ}$  path, it is  $H_{pr p}(s) = \frac{0.2e^{-0.1s}}{(0.8s + 1)(0.5s + 1)}$ , respectively. Assess whether the transfer function of the controller is physically feasible and place the control scheme on the figure.

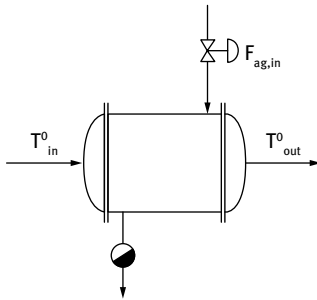


Fig. 15.4: Tubular heat exchanger.



7. Explain how the feedforward control may be included in the MPC algorithm. Are there any additional tuning considerations to be considered for the feedback-feedforward MPC controller?



8. Can the MPC controller cope with processes showing inverse response? What simplifications may be brought to the MPC performance index for controlling processes with large pure time delay?



9. What are the tuning guidelines for achieving zero steady-state offset for the unconstrained MPC design?



10. How would the multivariable MPC tuning take into account the control of output variables having values of different orders of magnitude?



11. Is it possible to apply inferential control to the MPC control algorithm, i.e. to control nonmeasurable (primary) output variables using other measured (secondary) outputs? Please elaborate.
























12. The MPC controller can use a higher number of control variables than the number of controlled variables. What practical circumstances may motivate such control approach?





13. In a mass transfer reactor, gas  $A$  is dispersed in aqueous saline solution  $B$ . The reactor has a jacket with cooling water. The temperature is measured with a

thermoresistor with sheath, with the temperature controller modifying the water flow in the jacket. The level transducer is a differential pressure gauge and the level controller manipulates the outflow from the reactor. Gas A flow is used to control the conversion in the reactor which has to be at 50%. The pressure is kept at 2 bar by manipulating the output unreacted gas flow and of the vapors at the operating temperature of 110°C. Design the overall control block scheme of the system and describe its interactions.





14. Explain the reason for considering the fuzzy controller to have good robust characteristics. 
15. How is it possible to include the integral mode in the fuzzy controller design? 
16. The fuzzy controller may be designed so as to include adaptive features. Elaborate on the way the form of the membership functions, fuzzy sets selection, and fuzzy rules formulation may be used for accomplishing this adaptive capability. 
17. Explain the way a fuzzy controller may be designed to operate with different control performance in the cases of the rising and descending change of the controlled variable to control processes showing integral behavior. 
18. What measures can be applied for the fuzzy controller tuning to reduce oscillations of the controlled variable? 
19. Explain why the RGA matrix is scale independent and the sum of its elements on each row or column is equal to unity. 
20. How many decouplers are necessary for a  $5 \times 5$  multi-loop control system? If some of the decouplers result as physically unrealizable, what solution still exists to implement the decoupling? 
21. How does the nonlinearity affect the loop pairing based on steady-state RGA? 
22. Elaborate a methodology for determining the RGA matrix based on an experimental approach. 
23. What are the incentives of the centralized control approach, compared to the decentralized one? 
24. Which of the layers of the hierarchical structure for controlling the whole plant should consider multiobjective requirements? 
25. Explain how you would integrate the following plantwide control requirements: minimizing the utility, energy, raw material, and environment protection costs and maximizing the throughput. 
26. Explain the importance of the active constraint variables and the self-optimizing variables for the plantwide control structure design methodology. 

-  27. Please motivate why the selection of the primary and secondary controlled variables features a mutual implication relationship.
-  28. How and at what part of the plantwide control structure design methodology should be constraints taken into account?
-  29. What is the effect of sampling on the design of the discrete controllers? What are the effects of increasing the sampling time on the control performance?
-  30. Explain why the PID velocity algorithm does not need initialization. How is initialization performed for the PID position algorithm.
-  31. How can continuous tuning rules, such as Ziegler-Nichols and Cohen-Coon, be used for tuning discrete controllers? What is the effect of the sampling time value on tuning discrete controllers?
-  32. What are the design solutions for limiting the ringing effect?
-  33. What are the design solutions to avoid getting physically unrealizable controllers, in the case of the controller design imposing the desired closed loop Z transform response  $Y(z)$ ?
-  34. How can a discrete PID controller be designed to have adaptive features in order to respond to changes (nonlinearity) in the steady-state and dynamic behavior of the controlled process?





## 15.2 Applied process engineering control

-  1. Write the steady-state model and simulate a CSTR in function for the assumption of three control loops such as those illustrated in Fig. 8.23. Consider the data specified for Example 8.1. Other technical data are: jacket volume  $V_{jacket} = 220$  l, heat capacity of heat-transfer agent  $C_{p,ag} = 1$  kJ/kg·grd, density  $\rho_{ag} = 1000$  kg/m<sup>3</sup>, and inlet temperature  $T_{ag,in}^o$  at a value that ensures thermal stability. Feed and discharge occurs on pipes with 50-mm inner diameter. The pressure drops are 0.5 and 0.2 bar on the inlet and outlet pipes, respectively. The transducer time constants are  $T_{TT} = 2$  min for temperature,  $T_{TF} = 5$  s for flow rate, and  $T_{TL} = 2$  s for level control, respectively. The corresponding amplification factors are  $K_{TT} = 0.08$  mA/°C,  $K_{TF} = 0.8$  mA/m<sup>3</sup>·h, and  $K_{TL} = 4$  mA/m, respectively.
-  2. Check the thermal stability of a CSTR for an exothermic first-order reaction such as considered in Example 8.1 and calculate the oscillation period. The reactor has a heat-transfer surface of  $A_r = 4$  m<sup>2</sup> characterized by  $K_r = 2.10^3$  kJ/m<sup>2</sup>·h·°C. The reaction heat is temperature independent and has a value of  $\Delta H_r = -12000$  kJ/kg. The inlet flow has a rate of  $F_{in} = 2000$  kg/h and contains  $C_{Ao} = 0.2$  kmol/m<sup>3</sup> at  $\rho = 10^3$  kg/m<sup>3</sup>. Its inlet temperature is  $T_{in}^o = 95^\circ\text{C}$ . The reaction mass does not suffer significant density changes and keeps its heat capacity constant at




$C_p = 935 \text{ J/kg}\cdot^\circ\text{C}$ . The setpoint is adjusted at  $100^\circ\text{C}$  and 2 hours residence time (see Fig. 8.1(b)). The heat-transfer agent inlet temperature into the jacket is of  $T_{ag,in}^\circ = 85^\circ\text{C}$ . Comment on the results. How does operation alter if (a)  $F_{in}$  and (b)  $C_{A0}$  drop to half?

3. Recalculate the parameters  $D/R$ ,  $V/F$ ,  $D/F$  of the distillation column presented in Example 9.2 that allows an  $r = 0.95$  recovery factor if the feeding concentration decreases to  $x_F = 0.4$ . 
4. Calculate the optimal  $D/F$  ratio that minimizes the losses generated by contamination of distillate and bottom product. The distillate product can be sold at 5.3 USD/kg, but the bottom product at 2.3 USD/kg. The feeding stream composition is  $x_F = 0.5$ , the separation factor is  $S = 361$ , whereas the molar masses of bottom and distillate species are of  $M_B = 92$  and  $M_D = 78 \text{ kg/kmol}$ , respectively. 
5. Calculate the batch time for a discontinuous distillation required to achieve  $x_B = 0.1$  when the loaded mixture is of 100 kmol with an  $x_F = 0.3$  (initial) concentration. The mean distillate concentration at the end of the batch should be of 0.7. The distillate flow rate is of 10 kmol/h. 
6. An absorption column has to ensure absorption of species A. Design a feedforward control system for the following situations: 
  - (a) it attenuates 0.2% variation of  $y_F$
  - (b) it attenuates 0.8-bar variation of operating pressure.

The operating pressure is kept constant.

7. Consider a countercurrent L-L packed extraction tower, with light phase dispersion and in continuous steady-state operation mode at 300 K (see Fig. 11.3 and 11.4). It is characterized by the following data:  $F = 100 \text{ m}^3/\text{h}$ ,  $\rho_F = 1000 \text{ kg/m}^3$ ,  $\rho_S = 870 \text{ kg/m}^3$ ,  $x_{AF} = 0.15$ ,  $x_{AS} = 0.01$ ,  $K_A = 1.36$  and setpoint  $x_{AE} = 0.05$ . Design a suitable control system that countervails  $\pm 30\%$  disturbances in  $x_{AF}$  and  $x_{AS}$ . 
8. Consider a countercurrent L-L packed extraction tower, in continuous steady-state operation mode at 300 K (see Fig. 11.3 and 11.4). Technical data are  $F = S = 5 \text{ kmol/h}$  and  $K_A = 4.76$ . Assess whether temperature control is necessary. 
9. An evaporator concentrates a solution from 20% initial mass concentration to a final concentration of 80%. Calculate the effect on the final concentration of (a) 1% change in steam flow or feed flow rate and (b) 1% change in feed concentration. 
10. A single effect evaporator concentrates an aqueous solution of NaOH (sodium hydroxide) from a 0.2-kg/kg solution to a 0.7-kg/kg solution. The feed flow rate of diluted solution is  $F = 70.4 \text{ kg/h}$  and the latent heat of vaporization of water is  $l_v = 525 \text{ kcal/kg}$ . Calculate the setpoint of a steam flow controller that ensures the vaporization of the unnecessary water. 




-  11. Calculate for the evaporator in Example 12.1 the gains on the transfer paths  $V_0 \rightarrow x_0$  and  $x_F \rightarrow x_0$ .
-  12. Consider an aspirin powder batch dryer, operating to bring the moisture content from  $w_c = 0.1$  to desired  $w^* = 0.05$ , having the input air temperature  $T_i^o = 93^\circ\text{C}$  and the wet bulb temperature of  $T_{wb}^o = 34^\circ\text{C}$ . Calculate what additional energy consumption and what additional costs are involved if the drying of 500 kg powder stops at  $80^\circ\text{C}$ . The cost is of 70 USD/Gcal.
-  13. A chemical engineer determines the following solubility concentrations for a certain chemical species:

**Tab. 15.1:** Solubility concentrations at different temperatures for a chemical species.

Temperature [ $^\circ\text{C}$ ]	10	30	60
Solubility [g/g]	0.013	0.046	0.182

If the batch crystallizer is loaded at  $40^\circ\text{C}$  with saturated solution and cooled to  $15^\circ\text{C}$ , what is the expected crystal mass per kilogram suspension?

-  14. A CMSMPR crystallizer is operated at  $T = 20^\circ\text{C}$ , and fed with a  $T = 40^\circ\text{C}$  solution. After some structural modification, the crystallizer is fed with a second-inlet, saturated stream. Let us calculate the required temperature of the second inlet flux that doubles the specific solid content (g/g) of the outlet stream. The yield is of  $\eta = 90\%$ , whereas the solubility is characterized by data in Tab. 15.1.

# Index

- Absorber 262, 263, 264, 266, 268
- absorption 262
  - of carbon dioxide 268
  - column 262, 263, 266
  - control 266
- acid 203
  - strong 203
  - weak 203
- activation energy 97, 181, 264
- additivity 46
- advanced control 2, 292, 318
- analytical model 75
- analyzer 210
- analyzer configuration 211
  - median selector 212
  - redundant 211
  - voting 211
- anti-solvent 303
- applied control 177, 178, 322
- automatic control system (ACS) 3, 301, 318
- automatic mode 123, 128
- average (mean) relative volatility 243
- azeotrope mixture 245, 270
  
- back-extraction 270
- balance
  - energy 191, 221, 228, 241, 295
  - heat 219, 290, 293
  - mass 184, 219, 221, 228, 239, 262, 273, 293, 307
  - population 307, 313
- base 203
  - strong 203
  - weak 203
- batch 180, 195
  - crystallization 308
  - distillation 103, 259
  - process 102
  - unit 180, 195, 201, 203
- batch stirred tank reactor (BSTR) 179
- binary mixture 272
- Black-Nichols diagram 7
- block
  - corrector 191
  - estimator 216
  - predictor 191
- boiling point 233, 251
- boiling temperature 283
- Boolean inference 85
- buffer 203
  
- CARIMA model 49, 65
- cascade control 3, 8, 10, 11, 190, 318
- centralized control 120, 133
  - internal model control 134
  - inverse of the process transfer matrix 133
- chord length 309
- chord length distribution 309
- cold stream injection 193
- column
  - absorption 262, 263, 266
  - distillation 233, 235, 239, 259
- concentration control 251
- condition number 128, 129
- constant
  - equilibrium 203
- constraint 62, 185, 186, 187
  - softening 62
- continuous
  - unit 188
- continuous mixed suspension mixed product removal (CMSMPR) 306, 324
- continuous signal 145
- continuous stirred tank reactor (CSTR) 3, 179, 184, 209, 318, 322
- control algorithm 2
- control function 117
- controller 5
  - gain 5, 8, 14, 188
  - parameters 5
- control solution 249
- control strategy 178
- control structure 15, 212, 215, 321
- coolant 188, 189
- crisp set 75, 76, 80
- cross-over frequency 8
- crystal 303, 308, 313
  - density 307
  - growth 303
  - monolayer growth 305
  - polylayer growth 305
  - seed 311
  - size distribution 306, 308, 313

- crystallization 303, 324
  - batch 324
  - control 307
  - cooling 303
- damping time 20
- dead time 14, 15, 19, 188
- decentralized control 120
- decoupling control 319, 321
- defuzzification 89, 295
- desorber 268
- desorption 262, 268, 269
- dew point 300
- differential pressure 299
- digital control system 145
  - analogue-to-digital conversion 145
  - digital-to-analogue conversion 145, 146
- direct nucleation control (DNC) 310
- discrete controller 173, 322
  - Dahlin controller 174
  - deadbeat controller 174
- discrete convolution 167, 168
- discrete PID controller 171, 322
  - anti-reset windup 173
  - integral (or reset) windup 173
  - position form 171
  - ringing effect 172
  - tuning 171, 172
  - velocity form 171, 172, 173
- discrete signal 145
- discrete system 145
  - difference equation model 156, 159, 160
  - discrete model of the continuous first order system 159
  - discrete time model of the second order system 160
  - linear and discrete time invariant model 156
- dispersion 271
- distillate 234
- distillation 233, 323
  - batch 259, 323
  - binary 234
  - column 18, 20, 28, 114, 233, 235, 239, 259, 323
  - control 249, 252
- disturbance 3, 5, 15, 17, 21, 213, 214, 216, 256
- dry bulb 300
- dry-bulb temperature 287
- dryer 287, 289, 299
- drying 287, 324
  - adiabatic 299
  - batch 289, 292, 324
  - chamber 287, 292
  - continuous 299
  - control 287, 289, 292
- dynamic 254
  - optimal control 102
  - programming 103
  - state 254
- ebullioscopic temperature 283
- electrode 206
  - potential 206
- end-point 210, 211
- end-time 222
- enthalpy 223, 274, 283
- equilibrium 96
  - constant 96
- equivalent gain 19
- evaporation 279
  - control 279, 281
  - heat 240
- evaporator 279, 282, 323
  - multiple effect 279
  - simple / single effect 279, 282, 323
- excess factor 26
- exponential smoothing filter 157
- extract 270
- extraction 270, 323
  - column 323
  - column / tower 271
  - continuous 275
  - control 270
  - countercurrent 271, 323
  - crosscurrent 271
  - liquid-liquid (L-L) 270, 323
  - multistage 271
  - residue 270
  - selectivity 276
  - solid-liquid (S-L) 270
  - solvent 270
  - stage 271
- extractor 271
- factor
  - amplification 186, 188
  - recovery 234

- separation 235, 245
- volume shape 307
- feasible solution 33, 62
- feedback 34
  - control 6, 8, 10, 11, 14, 34, 318
- feedforward control 15, 24, 320, 323
- fluidized bed 299
- focused beam reflectance measurement 309
- freedom degree 215
- fugacity 96, 224
- fuzzification 84, 295
- fuzzy 75, 292, 295, 321
  - aggregation 86, 87
  - algebraic product operation 82
  - algebraic sum operation 82
  - antecedent 84, 85, 86
  - bell-shaped membership function 79
  - complement operation 81, 82
  - consequent 84, 89
  - controller design 75, 87, 90
  - fuzzy system 86, 87
  - if-then rule 84, 86
  - implication 84, 85
  - inference 84, 85, 87, 89
  - intersection operation 81, 82
  - linguistic attributes 83, 85, 86
  - linguistic variable 81, 83, 84, 88
  - logic 83, 84, 87
  - rules 84, 86, 87, 89
  - saturation membership function 78, 80
  - set 75, 76, 80, 82, 83
  - singleton membership function 79
  - trapezoidal membership function 78
  - triangular membership function 77
  - trueness 83
  - union operation 81, 82
- generalized model predictive control 41
- Gibbs diagram 276
- Hamiltonian 94, 100
- heat
  - capacity 283
  - integration 237, 238, 268, 269
  - interstage exchange 193
  - transfer agent 188
  - transfer area 283
  - transfer coefficient 185, 283
  - transfer surface 185
- heat exchanger 239, 320
  - network 239
- Henry
  - coefficient 201
  - law 201
- hierarchical control structure 136
  - advanced control layer 137
  - diagnosis and fault detection layer 137
  - instrumentation control layer 136
  - optimization layer 137
  - plant planning and scheduling layer 137
  - regulatory control layer 136, 140, 142
  - supervision control layer 137, 140
  - supervisory control layer design 143
- homogeneity 46
- inference engine 87
- inferential control 27, 216, 320
- integral error
  - IAE, Integral Absolute Error 129
  - ISE, Integral Square Error 129
  - ITAE, Integral of the Time weighted Absolute Error 129
- inverse response 254, 320
- inverse Z transform 163
  - definition 163
  - determination by direct division 163, 164
  - determination by partial fraction expansion 164
- inversion time 255
- Janecke diagram 276
- latent heat
  - of vaporization 279, 283, 290
- level control 202, 203, 250, 272, 321
- look ahead 39
- loop
  - feedback 6, 8, 10, 11, 14, 216, 217
  - feedforward 15, 216, 217
  - inner 3
  - master 12, 13, 190
  - multi 193, 216
  - outer 3, 15
  - primary 11
  - secondary 11
  - slave 12, 13, 190

- Mamdani inference 85, 89
- manual mode 123, 128
- mathematical model 19, 97, 213, 279, 322
  - dynamic 213
  - steady state 213, 219, 220, 228
- matrix 114, 270
  - function 114
- maximum principle 92
- metastable 311
  - limit 311
  - zone 311
- model predictive control (MPC) 32, 292, 295, 313, 320
  - analytical solution 60, 62
  - constraints 32, 33, 42, 60, 61, 67, 69
  - contraction constraint 67
  - controller gain matrix 64
  - feedforward 34, 39, 57
  - free run output 51, 52, 59, 64
  - impulse response model 43, 44, 47
  - input (control) horizon 39, 40, 41, 65
  - input weighting matrix 41, 42, 65
  - model horizon 65
  - model prediction 50, 65
  - optimization 32, 34, 35, 39, 40, 41, 60, 61, 62, 64, 67, 69
  - output weighting matrix 41, 65
  - performance index 32, 35, 39, 40, 41, 42, 60, 67, 70, 71
  - prediction horizon 35, 38, 40, 65, 67
  - predictor matrix 51, 52, 57, 59
  - receding horizon control 35, 38
  - shrinking horizon 314
  - stability 65, 66
  - state space model 49, 56
  - state weighting matrix 42
  - step response model 43, 44, 47, 48, 51, 57
  - terminal penalty term 67
  - terminal point equality 67
  - terminal region 67
  - time series model 49, 59
  - tuning 64
- moisture 287, 292, 324
  - critical content 288
- multiloop control 120, 321
  - complete decoupling 131, 132
  - Hadamard product 126
  - interaction 128
  - interaction among control loops 120
    - interaction assessment 124, 125, 128
    - pairing alternatives 123
    - pairing the control loops 120, 122, 126
    - partial decoupling 131
    - PID 120
    - relative gain array (RGA) 126, 127, 128, 129, 178, 216, 220, 321
    - single loop controller 120
    - stability 125, 127, 128
    - tuning 128, 129, 132
- multiple-input multiple-output (MIMO)
  - process 119, 215
- multivariable control 119, 215, 320
  - control objectives 119
  - interaction between process variables 119, 120, 121, 126
- Nernst
  - distribution law 274
  - equation 206
- Nernstian slope 206
- Niederlinski index (NI) 128, 129, 217
- nonlinear model predictive control (NMPC) 68, 313
  - discrete time control signal 69
  - discrete-time model 69
  - linearization 70
  - nominal steady state point 68
  - sequential algorithm 70
  - simultaneous algorithm 71
  - state-space model 69
- norm 42
  - 1-norm 42
  - infinity 42
  - square 42
- nucleation 303
  - primary 303
  - rate 304
  - secondary 303
- objective 213
  - control 213
  - function 257
  - operational 213, 273
- objective function 93, 100, 105
- open loop optimal control 33
- operation mode
  - adiabatic 191, 198
  - dynamic 254, 293
  - isotherm 189, 191

- non-isotherm 4
- steady state 239
- optimal 313
  - allocation 105, 109
  - decision 108, 109
  - policy 110
  - solution 116
  - temperature profile 100, 102, 193, 194, 200, 313
- optimal control 37, 92, 102, 260
  - dynamic 92, 102, 111
  - open loop 35
  - steady-state 92
- oscillation
  - period 191, 209
- overshoot 10, 196
  
- partition
  - coefficient 274
  - equilibrium 274
- percolation 270
- performance index 320
- pH 202, 203
  - control 202, 205, 208
  - regulator 206, 207
  - sensor 206
- phase shift 186, 187, 188
- plantwide control 137, 142, 321
  - active constraints variables 141
  - bottom-up methodology 140, 142
  - control design requirements 138
  - controller design 218
  - economic optimization index 140
  - mass, energy, and momentum balance 138
  - mixed integer optimization problem 143
  - off-design operating conditions 143
  - operating training system 139
  - optimization layer design 143
  - primary control variables 137, 140, 141
  - real time optimization 139
  - regulatory control 136, 142
  - secondary controlled variables 142, 143
  - selection of the control variables 140, 143
  - selection of the primary controlled variables 141
  - self optimizing variables 141
  - self-regulating (-optimizing) control structure 140
  - servo control 143
  - top-down methodology 140
  - traditional control practice 138
- plug flow 180
- plug flow reactor (PFR) 184
- Pontryagin's principle 92
- precipitation 303
- pressure 96, 252
  - control 200, 201, 252
  - partial 96
  - vapor 201, 244
- process
  - batch 180, 195, 201, 203
  - continuous 111, 188, 203
  - function 13
  - in parallel 10, 12, 14
  - in series 14
  - self-regulating 182
  - underspecified 215
- process analytical technologies (PAT) 309
- process vision microscopy 309
- proportional band (PB) 186, 188, 191
- proportional gain 202, 208
- proportional integral derivative (PID) 207, 208
  - batch-unit 196
  - controller 195, 207
  - dual-mode unit 196
- psychrometry 287
  
- quadratic performance index 33, 60, 111
  
- raffinate 270
- rate
  - coefficient 180, 184, 198
  - drying 288
  - evacuated heat 184
  - evaporation 287, 288
  - feed rate 189
  - flow rate 188, 274
  - generated heat 184
  - law 184, 223, 225
- rate determining step 97
- ratio 24
  - control 24, 274
  - controller 25
  - excess 199
- reactant
  - excess 180, 199
  - limiting 180
- reaction 97, 267
  - control 179

- conversion 180, 181, 183, 192
- endothermic 182
- end-point 195
- equilibrium 193
- equilibrium (reversible) 198, 199
- exothermic 21, 182, 193, 218, 322
- first-order 180, 184, 198, 199, 218
- rate 97, 184
- reversible 25
- runaway 182, 197
- second-order 225
- selectivity 180
- temperature 182
- time 180
- unit 179
- reactor 179
  - batch stirred tank reactor (BSTR) 179, 180, 221
  - continuous stirred tank reactor (CSTR) 21, 25, 179, 180, 209, 218, 224, 227
  - dynamic operation 179, 183, 187
  - electrochemical 227
  - fed-batch 180
  - ideal 179
  - jacket 218
  - multibed 194, 222, 224
  - operation cycle 195
  - plug flow reactor (PFR) 29, 179, 180, 191, 220, 222
  - pure-batch 180
  - steady state operation 179, 183, 187
- reboiler 237, 250
- reconstruction of continuous signals 148, 153
- First-Order Hold (FOH) element 154
- Zero-Order Hold (ZOH) element 153, 154, 155, 168
- refining 270
- reflux 240, 242, 260
- relative disturbance gain 129
- relative humidity 287
- residence time 179, 202, 307
  
- sampling 145, 161, 172
  - continuous signals 147
  - ideal sampling element 151, 152
- sampling time 19, 137, 145
  - first order system 148
  - first order system with dead time 149
  - underdamped second order system 150
- saturation 303
- selective control 28
- selector switch 28
  
- sensitivity 178, 282, 296
  - analysis 282, 296
- setpoint 3
- settling 271
- Shannon theorem 65, 151
- single-input single-output (SISO) process 215
- singular value decomposition 128
- slack variable 62
- solute 270
- solvent 270, 279
- statistical model 75
- steady state 239, 279, 262
  - offset 129
  - (static) gain matrix 122, 123
- stripping 270
  - efficiency 272
- supersaturation 303
- supersaturation control (SSC) 311
  
- Tagaki-Sugeno-Kang inference 85
- temperature control 182, 188, 195, 275, 321, 323
- tetrahedral diagram 276
- thermal
  - equilibrium 185
  - instability 182, 184
  - stability 182, 183, 188, 322
- thermometer 251
- time constant 14, 19, 186, 188
- titration curve 205, 207
- transfer
  - function 8, 11, 218, 318, 320
  - matrix 122, 126, 133
- tray 11, 12, 19, 254, 266
  - efficiency 244
- tubular plug flow reactor (PFR) 179
- tuning 5, 13, 14
  
- undershoot 196
- universe of discourse 76, 80
  
- Van't Hoff equation 274
- vapor 279
  - flow rate 240, 242
  - pressure 201, 244
- variable 213
  - controlled 17, 213, 214, 220
  - input 213, 214
  - manipulated 10, 213, 214, 220
  - measurable 214

- output 11, 213
- unmeasurable 214
- volatility 233
  
- waste 234
- wet bulb 287
- wet bulb temperature 287, 324
  
- Zadeh inference 85
- zeta transform 168
  - inverse 168
  
- Ziegler-Nichols method 322
- Z transfer function 166, 167
  - of interconnected systems 169
  - of sampled systems 168
  - poles 167, 172
  - zeros 167
- Z transform 160, 161, 322
  - definition 161
  - inverse 163
  - of simple functions 162



